

19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 748 144**

51 Int. Cl.:

G10L 19/012 (2013.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

96 Fecha de presentación y número de la solicitud europea: **12.12.2013 E 16173655 (8)**

97 Fecha y número de publicación de la concesión europea: **12.06.2019 EP 3086319**

54 Título: **Métodos y aparatos para retención DTX en codificación de audio**

30 Prioridad:

22.02.2013 US 201361768028 P

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

13.03.2020

73 Titular/es:

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

72 Inventor/es:

**BRUHN, STEFAN;
SEHLSTEDT, MARTIN y
JANSSON TOFTGÅRD, TOMAS**

74 Agente/Representante:

LINAGE GONZÁLEZ, Rafael

ES 2 748 144 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

Métodos y aparatos para retención DTX en codificación de audio

5 **Campo técnico**

La solución descrita en el presente documento se refiere generalmente a la codificación de audio, y en particular a las tramas de retención (del inglés "hangover") asociadas con la transmisión discontinua (DTX) en la codificación de audio.

10

Antecedentes

Los estándares actuales de codificación de audio o conversación como 3GPP AMR (3GPP TS 26.071) y AMR-WB (3GPP TS 26.171), así como varios estándares de codificación de conversación ITU-T (por ejemplo, la recomendación ITU-T G.729, la recomendación ITU-T G.718) incluyen un esquema de transmisión discontinua (DTX) que suspende la transmisión de conversación durante la inactividad de conversación, y en su lugar transmite tramas de descriptor de inserción de silencio (SID) a una velocidad de bits y velocidad de transmisión de trama significativamente reducidas en comparación con las usadas para la conversación activa codificada. El propósito de DTX es aumentar la eficiencia de transmisión, lo que a su vez reduce el costo de la comunicación de conversación y/o aumenta el número de conexiones telefónicas posibles simultáneamente en un sistema de comunicación dado.

20

Los actuales sistemas de comunicación del estado de la técnica con DTX transmiten tramas de codificación de conversación regulares durante los segmentos de conversación activos. Durante segmentos inactivos, por ejemplo, pausas de conversación, estos sistemas transmiten más bien tramas SID a partir de las cuales el receptor genera el llamado ruido de confort como una señal de sustitución de la señal de inactividad. Para lograr la mejor eficiencia DTX posible, es deseable que las tramas de codificación de conversación solo se transmitan durante la conversación activa y no en segmentos inactivos, por ejemplo, durante pausas de conversación.

25

Para hacer esta distinción entre conversación e inactividad, se usa un detector de actividad de voz (VAD) en el lado de codificación o envío. Durante las tramas correspondientes a segmentos de conversación activos, se levanta una marca VAD. Este concepto sufre en la práctica, y especialmente en situaciones de conversación en ruido de fondo, de errores de clasificación VAD. Es decir, los períodos de inactividad se clasifican como períodos de conversación activa y/o viceversa. Uno de los principales problemas de los VAD es la detección de los puntos finales de conversación, es decir, el punto preciso en el tiempo en el que la señal cambia de conversación activa a la inactividad. La razón principal de este problema es que muchos desplazamientos de conversación están decayendo lentamente antes de que la conversación realmente se detenga, de modo que el final de los arranques de habla puede muy bien estar cubierto por el ruido de fondo. La consecuencia de este problema puede ser que tales desplazamientos de conversación se clasifiquen como inactividad, lo que puede dar como resultado que las tramas de señal correspondientes no se codifiquen, transmitan y reconstruyan como conversación activa, sino más bien como una señal de silencio para la que se generan tramas de ruido de confort. Esto significa que los desplazamientos de conversación (fin de períodos de conversación) pueden percibirse como cortadas, lo que lleva a una calidad significativamente reducida e incluso a la inteligibilidad de la conversación reconstruida. En otras palabras, esto puede conducir a una mala experiencia del usuario.

30

35

40

Los códecs de estado de la técnica actuales como AMR y AMR-WB resuelven este problema simplemente retrasando el inicio de la operación DTX con síntesis de ruido de confort varias tramas después del desplazamiento detectado por VAD. Esto se hace con una lógica de control DTX en el codificador, que extiende o agrega un período de tiempo durante el cual una señal de entrada se codifica como conversación activa, aunque la marca VAD indique inactividad. Este período se llama período de retención y en el caso de AMR y AMR-WB el período de retención es de 7 tramas de largo.

50

El período de retención no solo se usa como un medio para evitar el corte final (o desplazamiento) de conversación, sino también para el análisis de parámetros de trama SID. En el caso de AMR y AMR-WB, los primeros parámetros de trama SID después de un arranque de habla (suficientemente largo) no se transmiten, sino que el decodificador los calcula a partir de los parámetros de trama de conversación recibidos y almacenados durante el período de retención (3GPP TS 26.092; 3GPP TS 26.192). El propósito de realizar el cálculo del parámetro de trama SID basándose en los parámetros de trama de conversación recibidos durante el período de retención es ahorrar recursos de transmisión que de otro modo deberían haberse gastado en la transmisión de trama SID y minimizar el efecto de posibles errores de transmisión en los primeros parámetros de trama SID.

55

60

El principal problema con el período de retención en las soluciones de estado de la técnica descritas es que compromete la eficiencia del esquema DTX. Las tramas de retención están codificadas como conversación activa a pesar de que probablemente sean tramas de inactividad. Si la conversación comprende frecuentes arranques de habla separados entre períodos de inactividad, entonces se codifica un número significativo de tramas con alta velocidad de bits, por lo tanto, como tramas de conversación, en lugar de tramas de ruido de confort.

65

Un problema relacionado surge si el período de retención se acorta para mejorar la eficiencia del esquema DTX. Cuanto más corto sea el período de retención, más probable es que no represente adecuadamente la señal de ruido de inactividad. Esto puede conducir a degradaciones audibles de la síntesis de ruido de confort inmediatamente al final de los arranques de habla.

5 En AMR y AMR WB, el codificador y el decodificador realizan un seguimiento de las tramas de retención DTX usando una máquina de estado que debe estar sincronizada en el codificador y el decodificador.

10 El documento US 2010/106490 divulga el ajuste de la duración del período de retención basándose en los valores de energía de tramas de conversación en el período de retención DTX.

Sumario

15 Sería deseable, en un lado del decodificador de audio, generar ruido de confort, que es representativo del ruido de fondo en un lado del codificador de audio. Además, es deseable hacer esto de manera eficiente, usando solo un mínimo de recursos. Por lo tanto, un objetivo de la solución sugerida en el presente documento es permitir la generación de ruido de confort que sea representativo del ruido de fondo en un lado del codificador, y hacerlo usando una cantidad limitada de recursos.

20 La solución sugerida en el presente documento aumenta la eficiencia de las transmisiones de conversación con DTX sin comprometer la calidad de la síntesis de ruido de confort al final de los arranques de habla.

25 De acuerdo con un primer aspecto, se proporciona un método realizado por un codificador. El codificador funciona para codificar audio, como conversación, y para comunicarse con otros nodos o entidades, por ejemplo, en una red de comunicación. El codificador también funciona para aplicar un esquema DTX que comprende la transmisión de tramas SID durante la inactividad de conversación. El método comprende determinar un número N de tramas de retención, en el que el número N de tramas de retención es variable. El método comprende además transmitir las N tramas de retención a un decodificador. El método comprende además transmitir una primera trama SID al decodificador después de un período de retención, donde la primera trama SID comprende información que indica el número determinado N de tramas de retención. El método anterior permite que el decodificador genere ruido de confort basándose en las N tramas de retención.

35 De acuerdo con un segundo aspecto, se proporciona un método realizado por un decodificador. El decodificador funciona para decodificar audio, como conversación, y para comunicarse con otros nodos o entidades, por ejemplo, en una red de comunicación. El decodificador también funciona para aplicar un esquema DTX que comprende la recepción de tramas SID y la generación de ruido de confort durante la inactividad de conversación. El método comprende recibir N tramas de retención de un codificador. Además, se recibe una primera trama SID después de recibir las N tramas de retención. Un número N de tramas de retención se determina basándose en la información en la primera trama SID recibida. Además, el ruido de confort se genera basándose en las N tramas de retención.

40 De acuerdo con un tercer aspecto, se proporciona un codificador. El codificador funciona para codificar audio, como conversación, y funciona para comunicarse con otros nodos o entidades, por ejemplo, en una red de comunicación. El codificador también funciona para aplicar un esquema DTX que comprende la transmisión de tramas SID durante la inactividad de conversación. El codificador comprende medios de procesamiento, por ejemplo en forma de un procesador y una memoria, en el que dicha memoria contiene instrucciones ejecutables por dicho procesador. Los medios de procesamiento son operativos para determinar un número N de tramas de retención, en los que el número N de tramas de retención es variable. Los medios de procesamiento son además operativos para transmitir las N tramas de retención a un decodificador; y además para transmitir una primera trama SID al decodificador después de un período de retención, donde la primera trama SID comprende información que indica el número determinado N de tramas de retención.

55 De acuerdo con un cuarto aspecto, se proporciona un decodificador. El decodificador funciona para decodificar audio, como conversación, y funciona para comunicarse con otros nodos o entidades. El decodificador funciona además para aplicar un esquema DTX que comprende la recepción de tramas SID durante la inactividad de conversación. El decodificador comprende medios de procesamiento, por ejemplo en forma de un procesador y una memoria, y en el que dicha memoria contiene instrucciones ejecutables por dicho procesador. Los medios de procesamiento son operativos para recibir N tramas de retención de un codificador; y además recibir una primera trama SID después de recibir las N tramas de retención. Los medios de procesamiento son además operativos para determinar, basándose en la información en la primera trama SID recibida, un número N de tramas de retención; y para generar ruido de confort basándose en las N tramas de retención.

60 De acuerdo con un quinto aspecto, se proporciona un programa informático que comprende un código de programa informático que, cuando se ejecuta en un codificador, hace que el codificador realice el método de acuerdo con el primer aspecto.

65

De acuerdo con un sexto aspecto, se proporciona un programa informático que comprende un código de programa informático que, cuando se ejecuta en un decodificador, hace que el decodificador realice el método de acuerdo con el segundo aspecto.

- 5 De acuerdo con un séptimo aspecto, se proporciona un producto de programa informático, que comprende el programa informático de acuerdo con el quinto aspecto.

De acuerdo con un octavo aspecto, se proporciona un producto de programa informático, que comprende el programa informático de acuerdo con el sexto aspecto.

- 10 **Breve descripción de los dibujos**

Los anteriores y otros objetos, características y ventajas de la solución divulgada en el presente documento serán evidentes a partir de la siguiente descripción más particular de realizaciones como se ilustra en los dibujos adjuntos. Los dibujos no están necesariamente a escala, sino que se hace hincapié en ilustrar los principios de la solución divulgada en el presente documento.

- 15 La figura 1 es un diagrama de bloques del codificador. El codificador comprende un VAD y un codificador de retención.

- 20 La figura 2 es un diagrama de bloques del decodificador que funciona en DTX.

La figura 3 es un diagrama de bloques de VAD y lógica de determinación de retención.

- 25 La figura 4 es un diagrama de bloques del codificador de retención.

La figura 5 es un diagrama de flujo para el codificador de retención.

- 30 Las figuras 6a y 6b son diagramas de flujo para el decodificador de retención.

Las figuras 7a y 7b son diagramas de flujo que ilustran realizaciones de ejemplo de un método realizado por un nodo de transmisión o codificación, de acuerdo con la solución sugerida en el presente documento.

- 35 La figura 8 es un diagrama de flujo que ilustra una realización de ejemplo de un método realizado por un nodo de recepción o de decodificación, de acuerdo con la solución sugerida en el presente documento.

Las figuras 9-10 son diagramas de bloques que ilustran realizaciones de ejemplo de un nodo de transmisión, de acuerdo con la solución sugerida en el presente documento.

- 40 Las figuras 11-12 son diagramas de bloques que ilustran realizaciones de ejemplo de un nodo de recepción, de acuerdo con la solución sugerida en el presente documento.

Descripción detallada

- 45 La invención se define en las reivindicaciones adjuntas. Todos los casos de la palabra "realización o realizaciones", excepto los correspondientes con las reivindicaciones, se refieren a ejemplos útiles para comprender la invención que fueron presentados originalmente pero que no representan realizaciones de la invención reivindicada en el presente. Esos ejemplos se muestran solo con fines ilustrativos.

- 50 Como se mencionó anteriormente: en los sistemas de comunicación que utilizan transmisión discontinua (DTX), la eficiencia de transmisión se reduce cuando se usan técnicas de retención para evitar la degradación de la calidad debido a decisiones incorrectas del detector de actividad de voz (VAD).

- 55 En los llamados segmentos de señal inactivos, por ejemplo, pausas de conversación, se genera ruido de confort, en el lado del decodificador, usando información transmitida en tramas de descriptor de inserción de silencio (SID). Si el período de retención también se usa para el análisis de parámetros SID, la duración del mismo preferiblemente no es tan larga como se requiere para cubrir decisiones del VAD incorrectas, sino un poco más larga para capturar las características de la señal de fondo. En general, la probabilidad de generar un ruido de confort adecuado aumentará con períodos de retención más largos. Por otro lado, los períodos de retención largos disminuyen la eficiencia del sistema de comunicación que utiliza DTX, ya que las tramas de señal inactivas se transmitirán como tramas de señal de conversación a una velocidad de bits y velocidad de transmisión de tramas más alta. En los sistemas de comunicación que usan estas técnicas, existe un compromiso entre la eficiencia de transmisión y la probabilidad de un ruido de confort representativo.

- 65 Un período de retención después de un desplazamiento de conversación puede ser adaptativo. Para el codificador, esto significa que, tras una decisión del VAD que cambia de 1 (= conversación activa) a 0 (= inactividad), se agrega

un período de retención adaptativo. La información que especifica las tramas que pertenecen al período de retención puede transmitirse con la primera trama SID después del período de retención. En la figura 1, se muestra un diagrama de bloques esquemático de tal codificador.

5 El decodificador puede recibir, por ejemplo, con la primera trama SID, la indicación de cuál de las tramas de conversación activas recibidas anteriormente pertenecen al período de retención. La información de conversación codificada de las tramas que pertenecen al período de retención puede usarse posteriormente para el cálculo del parámetro SID del lado del decodificador. En la figura 2, se muestra un diagrama de bloques esquemático del decodificador.

10 En la siguiente descripción, para fines de explicación y no de limitación, se exponen detalles específicos tales como arquitecturas particulares, interfaces, técnicas, etc. para proporcionar una comprensión profunda del concepto descrito en el presente documento. Sin embargo, será evidente para los expertos en la técnica que el concepto descrito se puede practicar en otras realizaciones que se apartan de estos detalles específicos. Es decir, los expertos en la técnica podrán diseñar varias disposiciones que, aunque no se describen o muestran explícitamente en el presente documento, incorporan los principios del concepto descrito y se incluyen dentro de su alcance. En algunos casos, se omiten descripciones detalladas de dispositivos, circuitos y métodos conocidos para no oscurecer la descripción de acuerdo con el presente concepto con detalles innecesarios.

15 Así, por ejemplo, los expertos en la técnica apreciarán que los diagramas de bloques en el presente documento pueden representar vistas conceptuales de circuitería ilustrativa u otras unidades funcionales que incorporan los principios de la solución. De manera similar, se apreciará que cualquier diagrama de flujo, diagrama de transición de estado, pseudocódigo y similares representan diversos procesos que pueden estar representados sustancialmente en un medio legible por computadora y ejecutados por una computadora o procesador, tanto si tal computadora o procesador se muestra explícitamente como si no.

20 Las funciones de los diversos elementos, incluidos los bloques funcionales, incluidos, entre otros, los etiquetados o descritos como por ejemplo, "computadora", "procesador" o "controlador", se pueden proporcionar mediante el uso de hardware como hardware de circuito y/o hardware capaz de ejecutar software en forma de instrucciones codificadas almacenadas en un medio legible por computadora. Por lo tanto, tales funciones y bloques funcionales ilustrados deben entenderse como implementados por hardware y/o implementados por computadora, y por lo tanto implementados por máquina.

25 En términos de implementación de hardware, los bloques funcionales pueden incluir o abarcar, sin limitación, hardware de procesador de señal digital (DSP), procesador de conjunto de instrucciones reducido, circuitería de hardware (por ejemplo, digital o analógico) que incluyen, entre otros, circuito o circuitos integrados de aplicación específica (ASIC) y (cuando corresponda) máquinas de estado capaces de realizar tales funciones.

30 En realizaciones de ejemplo de la solución sugerida en el presente documento, la duración de un período de retención, es decir, el número de tramas de retención, puede ser variable y adaptativa. Se puede generar un período de retención adaptativo, por ejemplo, en respuesta a la decisión del VAD y un indicador adicional. En la figura 3 se muestra un diagrama de bloques esquemático del VAD. La decisión inmediata del VAD puede ser una marca correspondiente a la clasificación inmediata de conversación/inactividad del VAD. Siempre que el VAD clasifique una trama de señal como conversación activa, esta marca se puede levantar y, de lo contrario, puede reducirse. Se puede introducir una marca de retención para controlar la duración del período de retención agregado después de que se haya bajado la marca VAD inmediata. Esto se hace preferiblemente de modo que se garantice que la señal de las tramas de retención comprende principalmente una porción representativa del ruido de fondo y que las porciones de conversación potencialmente restantes son insignificantes. Esto se hace con el propósito de permitir una estimación confiable del parámetro SID en un lado de decodificación, cuya estimación es representativa de la señal de ruido de inactividad y que no se ve afectada por las porciones de conversación potencialmente restantes. Una medida útil para basar la marca de retención es la relación señal/ruido estimada (SNR), que compara el nivel estimado de conversación restante con el nivel de ruido de inactividad estimado. Por ejemplo, cuando esta estimación de SNR está por encima de cierto umbral, la marca de retención se puede levantar y cuando cae por debajo de dicho umbral, el período de retención puede terminar. Cabe señalar que la lógica de determinación de retención puede generar una marca VAD final que podría ser diferente de la marca VAD inmediato en su entrada.

35 Por ejemplo, la duración del período de retención puede adaptarse en respuesta a la SNR estimada. Esto supone que la SNR disminuye al final de un arranque de habla. La adaptación tiene en cuenta que el grado de disminución de SNR puede variar de un arranque de habla a otro. El resultado es que la duración del período de retención en tramas es un parámetro variable. De acuerdo con una realización de ejemplo, esta duración de retención, es decir, el indicador de retención, se codifica y se transmite al decodificador. En la figura 4 se presenta un diagrama de bloques esquemático de un codificador de retención. Además de las marcas VAD y de retención, el codificador de retención de ejemplo usa una primera marca SID. La primera marca SID puede indicar si la trama actual es el primer SID que sigue a la codificación de señal activa. Cabe señalar que las marcas no necesariamente tienen que estar explícitamente señaladas como variables específicas, sino que pueden ser implícitas, por ejemplo, derivables de otras variables de estado del codificador. La duración codificada del período de retención puede transmitirse como

parte de la información comprendida en la primera trama SID transmitida después del final de la transmisión de tramas de conversación activas. La figura 5 muestra un diagrama de flujo genérico para el codificador del indicador de retención.

5 De acuerdo con una realización de ejemplo de la solución sugerida en el presente documento, la duración del período de retención después de la marca VAD inmediato descendente está adaptada de tal manera que el conjunto de tramas a considerar para la estimación del parámetro SID es una variable. Es decir, el número de tramas de retención puede ser fijo o variable, pero el conjunto de tramas a considerar para determinar los parámetros SID para la generación de ruido de confort no es necesariamente igual al número de tramas de retención. En este enfoque, se
10 supone que hay una medida que indica la idoneidad de cada trama del período de retención después de la marca VAD inmediato descendente para la estimación del parámetro SID. Por ejemplo, las tramas para las cuales esta medida está por encima de cierto umbral pueden considerarse representativas del ruido de fondo y, por lo tanto, adecuadas para la estimación del parámetro SID. La medida puede, como anteriormente, basarse en estimaciones de SNR. Entonces, de acuerdo con esta realización, la primera trama SID después del final de la transmisión de
15 tramas de conversación activas puede contener información sobre el conjunto específico de tramas que se usarán para la estimación de parámetros SID.

Como ejemplo, el conjunto puede comprender las n tramas que preceden a la primera trama SID. La codificación de qué tramas usar para la estimación del parámetro SID se puede hacer con una palabra de código de N bits
20 máximos, donde cada bit representa una trama respectiva que precede a la primera trama SID. Si se establece un bit en la palabra de código (= 1), la trama representada por el bit se usará para la estimación del parámetro SID, de lo contrario, no.

La medida SNR que se usa en las realizaciones anteriores es solo un ejemplo. Además, son posibles medidas más
25 avanzadas. En general, una medida adecuada debe ser un buen indicador de si la trama correspondiente contiene ruido que sea bien representativo de la señal de ruido de inactividad. Una de estas medidas más avanzadas puede, por ejemplo, comparar la potencia o las propiedades espectrales de la trama actual con las propiedades correspondientes de tramas recientes o de otras tramas recientes que se han identificado que contienen ruido.

30 Puede aparecer como una posibilidad en el flujo de bits normal de tramas codificadas incluir un bit para señalar si la trama codificada es una trama de retención o no. Sin embargo, esto se considera menos ventajoso ya que significaría que un bit en cada trama de conversación debería reservarse para la información que solo se usa después del final de una ráfaga de conversación.

35 Si bien los párrafos anteriores explican el retención DTX específico, también es común que el VAD ya agregue algo de retención para evitar el corte en el desplazamiento de conversación. Entonces sería posible permitir que el retención VAD específico y el retención DTX se superpongan. Por ejemplo, el análisis de la señal puede contribuir a la terminación temprana de retención si hay un número suficiente de tramas para generar un ruido de confort estable, independientemente de si las últimas tramas son del retención VAD o el retención DTX.

40 En la figura 6a, un diagrama de flujo esquemático muestra un decodificador de indicador de retención del lado del decodificador de ejemplo. En el ejemplo de 6a, se puede indicar en cada trama si es una trama de retención o no, y las tramas de retención se almacenan después. A partir del indicador de retención decodificado, se puede determinar cuál de las tramas de retención almacenadas deben usarse como base para el ruido de confort.
45 Alternativamente, la decisión en 601a, de si una trama es una trama de retención o no, no se toma hasta que el indicador de retención se decodifique en 602a. Para tomar la decisión después de la decodificación 602a, un conjunto de las tramas recibidas más recientemente debe almacenarse en un búfer, por ejemplo, de la duración N_max (número máximo de tramas de retención). En el último caso, las tramas de retención pueden identificarse en el conjunto de tramas actualmente almacenadas en el búfer, basándose en el indicador de retención decodificado, y
50 por lo tanto pueden almacenarse parámetros de al menos parte de las tramas de retención. Esto es quizás más claro en la figura 6b, que muestra el almacenamiento de 601b de las últimas tramas N_max. Cuando el indicador de retención se decodifica en 602b, las tramas de retención están presentes entre las tramas almacenadas, y los parámetros de ruido de confort pueden determinarse 603b basándose en las tramas de retención indicadas por el indicador de retención. Entonces el ruido de confort puede generarse 604b basándose en los parámetros. Como en
55 el codificador, la primera marca SID puede indicar si la trama actual es el primer SID después de la codificación de señal activa. La primera marca SID no necesariamente debe almacenarse en una variable, sino que puede derivarse de otras variables de estado del decodificador.

60 Los parámetros SID típicos son parámetros de ganancia y parámetros espectrales predictivos lineales como los parámetros de frecuencia espectral de línea (LSF). En una realización de ejemplo, el decodificador puede tomar estos parámetros de las 5 tramas anteriores y calcular sus promedios. Estos parámetros promediados pueden usarse posteriormente en la síntesis de ruido de confort del sistema DTX. Alternativamente, los parámetros SID usados para la síntesis de ruido de confort pueden determinarse a partir de un conjunto específico de las tramas de retención indicadas. El conjunto específico puede derivarse en el lado del decodificador usando, por ejemplo, el
65 parámetro de duración de retención recibido y los parámetros de las tramas recibidas previamente que se han almacenado en una memoria.

Aunque los parámetros derivados de un conjunto de tramas de retención se denominan principalmente parámetros SID en el presente documento, también sería posible usar otros parámetros, denominados de manera diferente, pero que tienen el mismo propósito, a saber, ser una base para la generación de ruido de confort.

5 El decodificador puede obtener, por ejemplo, del indicador de retención en la primera trama SID después de una secuencia de tramas de conversación activas, información sobre el conjunto específico de tramas anteriores que se usarán para el cálculo de parámetros SID. Después, los parámetros SID pueden calcularse usando, por ejemplo, la ganancia y los parámetros espectrales de las tramas identificadas por el código recibido. Suponiendo que se usa
10 una palabra de código de $n = 8$ bits como indicador de retención y esta palabra de código contiene la secuencia de bits "0 1 0 1 1 1 1", entonces se usan los 5 tramas directamente anteriores y la 7ª trama anterior. La ganancia y los parámetros espectrales de estas tramas se pueden promediar y usar posteriormente en la síntesis de ruido de confort del sistema DTX.

15 En los siguientes párrafos, se describirán con más detalle diferentes aspectos de la solución divulgada en el presente documento con referencias a ciertas realizaciones y a los dibujos adjuntos. Para fines de explicación y no de limitación, se establecen detalles específicos, tales como escenarios y técnicas particulares, con el fin de proporcionar una comprensión profunda de las diferentes realizaciones. Sin embargo, otras realizaciones pueden apartarse de estos detalles específicos.

20 Método de ejemplo realizado por un nodo de transmisión/codificación, figura 7

A continuación se describirá un método de ejemplo realizado por un nodo de transmisión o un nodo de codificación con referencia a la figura 7a. El nodo de transmisión funciona para codificar audio, como conversación, y para comunicarse con otros nodos o entidades, por ejemplo, en una red de comunicación. El nodo de transmisión funciona además para aplicar un esquema DTX que comprende la transmisión de tramas SID durante la inactividad de conversación. El nodo de transmisión puede ser, por ejemplo, un teléfono celular, una tableta, una computadora o cualquier otro dispositivo capaz de comunicación por cable y/o inalámbrica y de codificación de audio.

30 La figura 7a ilustra el método que comprende determinar 703a, entre un número N de tramas de retención, un conjunto Y de tramas que es representativo del ruido de fondo. El método comprende además transmitir 704a las N tramas de retención, que comprenden dicho conjunto Y de tramas, a un nodo de recepción. El método comprende además transmitir 705a una primera trama SID al nodo de recepción en asociación con la transmisión de las N tramas de retención, donde la trama SID comprende información que indica el conjunto Y determinado de tramas de
35 retención al nodo de recepción. El método anterior permite que el nodo de recepción genere ruido de confort basándose en el conjunto Y de tramas de retención.

El orden de las acciones en las figuras 7a y b solo es un ejemplo. Por ejemplo, el conjunto Y podría determinarse después de que se hayan transmitido las N tramas de retención.

40 Las tramas comprendidas en el conjunto Y de tramas de retención deben ser representativas del ruido de fondo. Por lo tanto, de entre número N de tramas de retención, las más adecuadas para determinar o calcular parámetros para la generación de ruido de confort, por ejemplo, los llamados parámetros SID deben ser identificadas. Las tramas del conjunto Y podrían determinarse o identificarse, por ejemplo, basándose en un nivel SNR de la señal comprendida
45 en cada trama, y cuando este nivel SNR cumple un cierto criterio, se determina que la trama es adecuada para su uso como base para el cálculo, por ejemplo, de los parámetros SID. Algunas de las N tramas de retención pueden ser menos representativas del ruido de fondo. Por ejemplo, algunas de las tramas de retención pueden comprender, al menos en parte, conversación o ruido transitorio, lo que las hace inadecuadas como base para derivar parámetros relacionados con la generación de ruido de confort. Por ejemplo, las tramas de conversación generalmente tienen
50 estructuras formantes, que no se ven en el ruido de fondo; y las tramas de ruido transitorio pueden tener mayor energía que el ruido de fondo promedio. Tales tramas de retención, que no representan el ruido de fondo, no deben incluirse en el conjunto Y.

El conjunto Y de tramas puede indicarse de diferentes maneras en la primera trama SID, que se describirá más adelante. Por "primera trama SID" se entiende la primera trama SID en un período DTX, que típicamente indica el inicio del período DTX. Por período DTX se entiende aquí un período de inactividad de conversación, durante el cual las tramas codificadas se envían desde el nodo de transmisión al nodo de recepción a una velocidad de bits y/o velocidad de trama más baja que durante los períodos no DTX. Por período DTX se entiende aquí el período entre ráfagas de conversación activas, período que se reemplaza por ruido de confort. Estos períodos comienzan con el primer SID para marcar la transición al ruido de confort. Esto generalmente es seguido por períodos de varios tramas "NO_DATA", que como su nombre lo indica no contienen ningún dato, y tramas SID (o SID_UPDATE). Las tramas SID se transmiten con mayor frecuencia a intervalos regulares, denominados "intervalo SID", hasta que el siguiente enunciado desencadena una transición de regreso a la codificación de conversación activa. Es decir, con un intervalo SID de 8, el período DTX se codificaría como: primer SID seguido de 7 tramas NO_DATA antes del
60 SID_UPDATE. Esta secuencia con 7 tramas NO_DATA seguidas de una actualización SID se repite hasta que se produce la transición a la conversación activa.

Una ventaja del método descrito anteriormente es, como se describió anteriormente, que permite a un nodo de recepción derivar parámetros para el ruido de confort a partir de tramas que se determina que son adecuadas para este propósito. Esto mejora la calidad del ruido de confort generado y, por lo tanto, mejora la experiencia del usuario.

5 El conjunto Y se indica además al nodo de recepción de una manera muy eficiente en el uso de recursos, utilizando la primera trama SID para este propósito. Es una ventaja determinar las tramas de retención adecuadas en el nodo de transmisión, ya que en este nodo, los datos reales de la señal de audio son accesibles, mientras que en el nodo de recepción, solo está disponible una versión cuantificada de los datos.

10 La información que indica el conjunto Y puede comprender un número, lo que implica una serie de tramas de retención en secuencia; una palabra de código o mapa de bits que indica las posiciones de las tramas que pertenecen al conjunto Y, entre las N tramas de retención; una palabra de código o mapa de bits que indica algunas de las N tramas de retención que se incluyen en el conjunto Y, y/o una palabra de código o mapa de bits que indica cuáles de las N tramas de retención que no se incluyen en el conjunto Y.

15 Por ejemplo, la trama SID podría comprender un número, por ejemplo, 5, que debe ser interpretado por el nodo de recepción, por ejemplo, ya que las últimas cinco tramas de retención deben usarse para determinar los parámetros para la generación de ruido de confort. Alternativamente, el número podría interpretarse como algún otro grupo de cinco tramas entre las N tramas de retención, como las últimas cinco excepto una. El número N de tramas de retención podría ser, por ejemplo, 6, 7, 8 o 9. En un caso especial, el número N de tramas de retención podría ser igual al número indicado en la trama SID, es decir, los parámetros deben determinarse basándose en todas las tramas de retención.

25 Alternativamente o además, la trama SID podría comprender una palabra de código o mapa de bits/máscara de bits que indica las posiciones de las tramas que pertenecen al conjunto Y. Tal palabra de código podría configurarse de diferentes maneras. Se podría usar un sistema de códigos, donde tanto el nodo de transmisión como el nodo de recepción tienen conocimiento del significado de los códigos, por ejemplo, ambos lados tienen acceso a un libro de códigos que especifica, por ejemplo, que la palabra de código "01" se mapea a tramas de retención, en la trama k; k-1, k-2, k-4 y k-6 entre las N tramas de retención. Alternativamente, podría usarse un mapa de bits/máscara de bits. Tal mapa de bits podría cubrir todas las N posiciones de las N tramas de retención o un subconjunto de las N posiciones. El nodo de recepción debería, en algún momento, haber sido informado previamente del carácter del mapa de bits/máscara de bits. Por ejemplo, si N = 8, un mapa de bits/máscara de bits de ejemplo como "11011000" podría estar comprendido en la trama SID, lo que indica que las 4^a, 5^a, 7^a y 8^a tramas anteriores deben usarse para determinar los parámetros del ruido de confort. Alternativamente, el mapa de bits/máscara de bits "11011" podría estar comprendido en la primera trama SID, teniendo el mismo significado que el ejemplo anterior. Alternativamente, podrían indicarse las posiciones de las tramas de retención que no están comprendidos en el conjunto Y. En analogía con el ejemplo anterior, un mapa de bits/máscara de bits correspondiente podría ser "00100111" o "00100" o "100111".

40 Estas son todas las diferentes realizaciones de información que podrían estar comprendidas en la primera trama SID para indicar cuál de las tramas de retención debe usarse. Generalmente, cuantos menos bits sean necesarios para indicar el conjunto Y, mejor.

45 El concepto explicado anteriormente de transmitir, en la primera trama SID, una identificación del conjunto de tramas de retención para basar la generación de ruido de confort, se puede combinar con la transmisión de parámetros SID como parte de la primera trama SID. Es decir, la primera trama SID puede comprender además parámetros SID. Estos parámetros SID darán una indicación de cómo se ve la señal en la trama actual. Esta información podría, por ejemplo, ponderarse más que la información de las tramas de retención anteriores. Por supuesto, las tramas de retención ya se pueden ponderar de manera diferente sin considerar los parámetros de señal de la trama SID, pero de todos modos la decisión de no ir a DTX en la trama anterior debe indicar que no estamos lo suficientemente seguros de que esta trama represente inactividad/solo fondo ruido.

50 El número N de tramas de retención puede ser dinámicamente variable, como se describió anteriormente. El número N podría determinarse basándose en las propiedades de una señal de audio de entrada. Por ejemplo, el número N podría depender del sonido de conversación que renuncia al período DTX y/o el carácter del ruido de fondo. Al usar un número dinámico de tramas de retención, el número de tramas de retención que deben transmitirse a un nodo de recepción podría mantenerse al mínimo y, por lo tanto, podrían ahorrarse recursos, en comparación con tener un número estático de tramas de retención.

60 Algunas acciones, que pueden preceder al método ilustrado en la figura 7a, se ilustran en la figura 7b. En la figura 7b, se determina en una acción 701b si una trama de un flujo de audio, por ejemplo, un segmento de una señal de audio, cuya señal comprende al menos en parte conversación, comprende conversación activa o no. Esto a menudo se conoce como detección de actividad de voz, VAD. Cuando se determina que una o más tramas no comprenden conversación activa, se transmitirán varias tramas de retención, por ejemplo, para reducir la probabilidad de cortar un sonido de conversación, como se describió anteriormente. Cuando se aplica un número dinámico de tramas de retención, se puede analizar la señal comprendida en las primeras tramas que se determina que no comprenden

conversación activa, y se puede determinar un número adecuado de tramas de retención en una acción 702b. Posiblemente, también se pueden tener en consideración las propiedades de las últimas tramas determinadas para comprender la conversación activa al determinar un número apropiado N de tramas de retención, por ejemplo, para determinar una SNR o una disminución de energía entre tramas adyacentes.

5 Es decir, se puede determinar un número, N, de tramas de retención basándose en una propiedad de la señal comprendida en las tramas antes y/o después de una decisión de inactividad de conversación. Además, o alternativamente, las propiedades de tramas de señal anteriores determinadas para comprender solo ruido de fondo podrían tenerse en consideración al determinar N.

10 Como se mencionó anteriormente, la determinación de una serie de tramas de retención podría basarse en una característica de una disminución de SNR o energía dentro y/o entre tramas de señal. El número N de tramas de retención puede ser estático, semiestático o dinámico, y podría ser diferente para diferentes desplazamientos de conversación.

15 Las tramas de retención transmitidas al nodo de recepción, por ejemplo, en la acción 704b, pueden codificarse de acuerdo con la codificación de tramas que comprenden conversación activa, como se describió previamente. Cuando el número N de tramas de retención es dinámico, el número N también podría indicarse al nodo de recepción, por ejemplo, en la primera trama SID.

20 Método de ejemplo realizado por un nodo de decodificación, figura 8

A continuación se describirá un método de ejemplo realizado por un nodo de recepción o un nodo de decodificación con referencia a la figura 8. El nodo de decodificación funciona para decodificar audio, como conversación, y para comunicarse con otros nodos o entidades, por ejemplo, en una red de comunicación. El nodo de decodificación funciona además para aplicar un esquema DTX que comprende la recepción de tramas SID y la generación de ruido de confort durante la inactividad de conversación. El nodo de decodificación puede ser por ejemplo, un teléfono celular, una tableta, una computadora o cualquier otro dispositivo capaz de comunicación por cable y/o inalámbrica y de decodificación de audio.

30 El método de ejemplo ilustrado en la figura 8 comprende recibir 801 N tramas de retención desde un nodo de transmisión. Además, se recibe 802 una primera trama SID en asociación con las N tramas de retención. Un conjunto Y de tramas de retención, entre el número N de tramas de retención, se determina 803, basándose en la información en la trama SID recibida. Además, el ruido de confort se genera 805, al menos parcialmente basándose en el conjunto Y de las tramas de retención.

40 La trama SID podría recibirse después de que se haya recibido la última de las N tramas de retención, lo que indica el inicio de un período DTX. Sin embargo, la trama SID también podría recibirse antes de las tramas de retención, o entre dos tramas de retención, si esto fuera permitido y regulado en el protocolo de transmisión para el esquema DTX.

45 El número N de tramas de retención podría indicarse en la primera trama SID, sin embargo, esto es opcional. El número N podría establecerse alternativamente en un valor predeterminado, por ejemplo, 7, lo que implica que las 7 últimas tramas recibidas, sin contar la trama SID, antes de un período DTX serían tramas de retención. Además, cuando se aplica un número dinámico de tramas de retención, hay otras formas de señalar el número N de tramas de retención. Por ejemplo, el número podría señalarse implícitamente a través de las propiedades de la señal de audio, por ejemplo, una medida SNR a largo plazo. Tal medida podría generarse basándose en la señal de audio decodificada y, por lo tanto, podría estar disponible en el decodificador.

50 La trama SID comprende, como se describió anteriormente, información que indica un conjunto Y de tramas, de entre las N tramas de retención, seleccionadas por el nodo de transmisión como representativas del ruido de fondo. Por lo tanto, es posible que el nodo de recepción determine el conjunto Y de tramas basándose en la primera trama SID. Es decir, basándose en la información comprendida en la primera trama SID que indica el conjunto Y. La información podría ser explícita o implícita, y se ejemplificó anteriormente al describir el método realizado por un

55 nodo de transmisión.

60 El nodo de recepción debe generar ruido de confort durante los períodos DTX silenciosos, es decir, durante los períodos en que no se reciben tramas de conversación desde un nodo de transmisión. El ruido de confort debe imitar preferiblemente el ruido de fondo en el nodo de transmisión. Para generar un ruido de confort tan auténtico como sea posible, el nodo de recepción debe estimar el ruido de fondo basándose en las tramas de retención que son más representativas del ruido de fondo. Alternativamente o además, el nodo de recepción podría recibir una estimación del ruido de fondo del nodo de transmisión, por ejemplo, en forma de parámetros SID. Las tramas SID están codificadas a una velocidad de bits significativamente menor que las tramas de señal activas. Por lo tanto, las características del ruido de fondo se capturan mejor, en el lado del codificador, durante el retención (desde las

65 tramas de retención) que en el SID. Sin embargo, la inclusión de parámetros SID en la primera trama SID puede ser ventajosa para tener una transición suave de tramas de retención a la generación de ruido de confort.

5 El nodo de recepción estima o deriva parámetros para la generación de ruido de confort, basándose en el conjunto Y de tramas. Los parámetros están asociados con el ruido de fondo en el lado del nodo de transmisión. Al hacerlo, el ruido de confort generado basado en dichos parámetros reflejará el ruido de fondo en el lado del nodo de transmisión de buena manera, y así logrará una experiencia de usuario buena/deseada. La selección del conjunto Y en el lado del transmisor es ventajosa, ya que en ese lado, la información de audio completa es accesible, en lugar de la versión reducida y cuantificada que está disponible en el lado del nodo de recepción.

10 Como se describió anteriormente, la información que indica el conjunto Y puede comprender uno o más de: un número, lo que implica una serie de tramas de retención en secuencia; una palabra de código o mapa de bits que indica las posiciones de las tramas que pertenecen al conjunto Y, entre las N tramas de retención; una palabra de código o mapa de bits que indica cuáles de las N tramas de retención están al menos comprendidas en el conjunto Y; y una palabra de código o mapa de bits que indica cuáles de las N tramas de retención no están comprendidas en el conjunto Y.

15 Además, la primera trama SID puede comprender además parámetros SID. El número N de tramas de retención puede variar dinámicamente basándose en las propiedades de una señal de audio de entrada, como se describió anteriormente.

20 Ejemplo de nodo de transmisión, figura 9

25 Las realizaciones descritas en el presente documento también se refieren a un nodo de transmisión o un nodo de codificación. El nodo de transmisión está asociado con las mismas características técnicas, objetos y ventajas que el método descrito anteriormente e ilustrado, por ejemplo, en las figuras 7a y 7b. El nodo de transmisión se describirá brevemente para evitar repeticiones innecesarias. El nodo de transmisión podría ser, por ejemplo, un dispositivo o UE, como un teléfono inteligente, una tableta, una computadora o cualquier otro dispositivo capaz de comunicación por cable o inalámbrica y de codificación de conversación.

30 A continuación, se describirá un nodo 900 de transmisión de ejemplo, adaptado para permitir el rendimiento de un método descrito anteriormente adaptado para realizar al menos una realización del método en un nodo de transmisión descrito anteriormente, con referencia a la figura 9.

35 El nodo de transmisión funciona para codificar audio, como conversación, y funciona para comunicarse con otros nodos o entidades, por ejemplo, en una red de comunicación. El nodo de transmisión funciona además para aplicar un esquema DTX que comprende la transmisión de tramas SID durante la inactividad de conversación. El nodo de transmisión puede funcionar para comunicarse, por ejemplo, en un sistema de comunicación inalámbrico, como GSM, UMTS, E-UTRAN o CDMA 2000, y/o en un sistema de comunicación por cable.

40 La parte del nodo de transmisión que está relacionada principalmente con la solución sugerida en el presente documento se ilustra como una disposición 901 rodeada por una línea intermitente/discontinua. La disposición y posiblemente otras partes del nodo de transmisión están adaptadas para permitir el rendimiento de uno o más de los métodos o procedimientos descritos anteriormente e ilustrados, por ejemplo, en las figuras 7a y 7b.

45 El nodo de transmisión ilustrado en la figura 9 comprende medios de procesamiento, en este ejemplo en forma de un procesador 903 y una memoria 904, en el que dicha memoria contiene instrucciones 905 ejecutables por dicho procesador. Los medios de procesamiento son operativos para determinar, entre un número N de tramas de retención, un conjunto Y de tramas representativas del ruido de fondo. Los medios de procesamiento son además operativos para transmitir las N tramas de retención, que comprenden al menos dicho conjunto Y de tramas, a un nodo de recepción; y para transmitir una primera trama SID al nodo de recepción en asociación con la transmisión de las N tramas de retención, donde la trama SID comprende información que indica el conjunto determinado Y de tramas de retención al nodo de recepción.

50 El nodo de transmisión permite que un nodo de recepción genere ruido de confort basándose en el conjunto Y de tramas de retención, lo que permite generar ruido de confort de alta calidad.

55 La información que indica el conjunto Y podría configurarse de diferentes maneras, y la primera trama SID podría comprender además parámetros SID; y el número N de tramas de retención podría ser variable o fijo, como se describió anteriormente.

60 El nodo 900 de transmisión se ilustra para comunicarse con otras entidades a través de una unidad 902 de comunicación, que puede considerarse que comprende medios convencionales para comunicación inalámbrica y/o por cable de acuerdo con un estándar de comunicación dentro del cual el nodo de transmisión funciona. La disposición y/o el nodo de transmisión pueden comprender además otras unidades funcionales 909, para proporcionar, por ejemplo, funciones de nodo de transmisión regulares, como por ejemplo, procesamiento de señales en asociación con la codificación de conversación.

65

La disposición 901 puede implementarse alternativamente y/o describirse esquemáticamente como se ilustra en la figura 10. La disposición 1001 comprende una unidad 1004 de determinación, para determinar, un conjunto Y de tramas, de entre un número N de tramas de retención, que es representativo del ruido de fondo. La disposición 1001 comprende además una unidad de transmisión para transmitir las N tramas de retención, que comprende, al menos, dicho conjunto Y de tramas, a un nodo de recepción; y además para transmitir una primera trama SID al nodo de recepción en asociación con la transmisión de las N tramas de retención, donde la trama SID comprende información que indica el conjunto Y determinado de tramas de retención al nodo de recepción.

La disposición 1001 puede comprender una unidad VAD, para determinar si una trama de señal comprende conversación activa o no. Alternativamente, dicha unidad VAD puede ser parte de las otras unidades funcionales 1008.

La disposición 1001 y otras partes del nodo de transmisión podrían implementarse, por ejemplo, por uno o más de: un procesador o un microprocesador y un software y almacenamiento adecuados, por lo tanto, un dispositivo lógico programable (PLD) u otro componente o componentes electrónicos/circuito o circuitos de procesamiento configurado para realizar las acciones mencionadas anteriormente.

Nodo de recepción/decodificación de ejemplo, figura 11

Las realizaciones descritas en el presente documento también se refieren a un nodo de recepción o nodo de decodificación. El nodo de recepción está asociado con las mismas características técnicas, objetos y ventajas que el método descrito anteriormente e ilustrado, por ejemplo, en la figura 8. El nodo de recepción se describirá brevemente para evitar repeticiones innecesarias. El nodo de recepción podría ser, por ejemplo, un dispositivo o UE, como un teléfono inteligente, una tableta, una computadora o cualquier otro dispositivo capaz de comunicación por cable o inalámbrica y de codificación de audio.

A continuación, se describirá con referencia a la figura 11 un nodo de recepción 1100 de ejemplo, adaptado para permitir el rendimiento de un método descrito anteriormente adaptado para realizar al menos una realización del método en un nodo de recepción descrito anteriormente.

El nodo de recepción funciona para decodificar audio, como conversación, y funciona para comunicarse con otros nodos o entidades, por ejemplo, en una red de comunicación. El nodo de transmisión funciona además para aplicar un esquema DTX que comprende la recepción de tramas SID durante la inactividad de conversación. El nodo de recepción puede funcionar para comunicarse en un sistema de comunicación inalámbrico, como GSM, UMTS, E-UTRAN o CDMA 2000, y/o en un sistema de comunicación por cable.

La parte del nodo de recepción que se relaciona principalmente con la solución sugerida en el presente documento se ilustra como una disposición 1101 rodeada por una línea intermitente/discontinua. La disposición y posiblemente otras partes del nodo de recepción están adaptadas para permitir el rendimiento de uno o más de los métodos o procedimientos descritos anteriormente e ilustrados, por ejemplo, en la figura 8.

El nodo de recepción ilustrado en la figura 11 comprende medios de procesamiento, en este ejemplo en forma de un procesador 1103 y una memoria 1104 y en el que dicha memoria contiene instrucciones 1105 ejecutables por dicho procesador. Los medios de procesamiento son operativos para recibir N tramas de retención de un nodo de transmisión; y además recibir una primera trama SID en asociación con las N tramas de retención. Los medios de procesamiento son además operativos para determinar, basándose en la información en la trama SID recibida, un conjunto Y de tramas de retención, de entre el número N de tramas de retención; y para generar ruido de confort al menos parcialmente basándose en el conjunto Y de las tramas de retención.

De este modo, el nodo de recepción está habilitado para generar ruido de confort basándose en el conjunto Y de tramas de retención, y por lo tanto habilitado para generar ruido de confort de alta calidad.

La información que indica el conjunto Y podría configurarse de diferentes maneras, y la primera trama SID podría comprender además parámetros SID; y el número N de tramas de retención podría ser variable o fijo, como se describió anteriormente.

El nodo 1100 de recepción se ilustra para comunicarse con otras entidades a través de una unidad 1102 de comunicación, que puede considerarse que comprende medios convencionales para comunicación inalámbrica y/o por cable de acuerdo con un estándar de comunicación dentro del cual el nodo de recepción funciona. La disposición y/o el nodo de recepción pueden comprender además una o más unidades 1106 de almacenamiento. La disposición y/o el nodo de recepción pueden comprender además otras unidades funcionales 1107, para proporcionar, por ejemplo, funciones de nodo de recepción regulares, como por ejemplo, procesamiento de señales en asociación con la decodificación de conversación.

La disposición 1101 y otras partes del nodo de recepción o de decodificación podrían implementarse, por ejemplo, por uno o más de: un procesador o un microprocesador y un software y almacenamiento adecuados, por lo tanto, un

dispositivo lógico programable (PLD) u otro componente o componentes electrónicos/circuito o circuitos de procesamiento configurados para realizar las acciones mencionadas anteriormente.

5 La disposición 1101 puede implementarse alternativamente y/o describirse esquemáticamente como se ilustra en la figura 12. La disposición 1201 comprende una unidad 1203 de recepción para recibir N tramas de retención desde un nodo de transmisión; y además para recibir una primera trama SID en asociación con las N tramas de retención. La disposición comprende además una unidad 1204 de determinación para determinar, basándose en la información en la primera trama SID recibida, un conjunto Y de tramas de retención, entre el número N de tramas de retención; y además un generador 1205 de ruido para generar ruido de confort basándose en el conjunto Y de tramas de retención.

15 La disposición 1201 puede comprender además una unidad de estimación para estimar los parámetros para la generación de ruido de confort, como por ejemplo, parámetros SID. El generador de ruido puede generar ruido de confort basándose en los parámetros estimados de generación de ruido.

Se supone que la disposición 1201 y/o alguna otra parte del nodo 1200 de decodificación comprenden unidades funcionales o circuitos adaptados para realizar la decodificación de audio.

20 La disposición 1201 y otras partes del nodo de recepción o de decodificación podrían implementarse, por ejemplo, por uno o más de: un procesador o un microprocesador y un software y almacenamiento adecuados, por lo tanto, un dispositivo lógico programable (PLD) u otro componente o componentes electrónicos/circuito o circuitos de procesamiento configurados para realizar las acciones mencionadas anteriormente.

25 Debe entenderse que la elección de unidades o módulos que interactúan, así como la denominación de las unidades son solo para fines de ejemplo, y los nodos de cliente y servidor adecuados para ejecutar cualquiera de los métodos descritos anteriormente pueden configurarse en una pluralidad de formas alternativas para poder ejecutar las acciones de proceso sugeridas.

30 También debe tenerse en cuenta que las unidades o módulos descritos en esta divulgación deben considerarse como entidades lógicas y no necesariamente como entidades físicas separadas.

35 Mediante el uso de la solución sugerida en el presente documento, la eficiencia de las transmisiones de conversación con DTX puede aumentarse sin comprometer la calidad de la síntesis de ruido de confort al final de los arranques de habla.

ABREVIATURAS

AMR: multitasa adaptativo

40 DTX: transmisión discontinua

ITU-T: sector de estandarización de telecomunicaciones de la unión internacional de telecomunicaciones

45 LSF: frecuencia espectral lineal

VAD: detector de actividad de voz

3GPP: proyecto asociación de tercera generación

50 SID: descriptor de inserción de silencio

SNR: relación señal/ruido

55 WB: banda ancha

REIVINDICACIONES

- 1.- Un método realizado por un codificador, el codificador funcionando para codificar la conversación y aplicar un esquema de transmisión discontinua, DTX, que comprende la transmisión de las tramas de descriptor de inserción de silencio, SID, durante la inactividad de conversación, comprendiendo el método:
- 5
- determinar (703a) un número N de tramas de retención, en el que el número N de tramas de retención es variable;
 - transmitir (704a) las N tramas de retención a un decodificador;
- 10
- transmitir (705a) una primera trama SID al decodificador después de un período de retención, donde la primera trama SID comprende información que indica el número determinado N de tramas de retención,
- 15
- permitiendo así que el decodificador genere ruido de confort basándose en las N tramas de retención.
- 2.- El método de acuerdo con la reivindicación 1, en el que el número N de tramas de retención es dinámicamente variable basándose en las propiedades de una señal de audio de entrada.
- 3.- El método de acuerdo con la reivindicación 1 o 2, en el que la primera trama SID comprende además parámetros SID.
- 20
- 4.- Un método realizado por un decodificador que funciona para decodificar la conversación y aplicar un esquema de transmisión discontinua, DTX, que comprende la recepción de las tramas de descriptor de inserción de silencio, SID, y la generación de ruido de confort durante la inactividad de conversación, comprendiendo el método:
- 25
- recibir (801) N tramas de retención de un codificador;
 - recibir (802) una primera trama SID después de recibir las N tramas de retención;
- 30
- determinar (803), basándose en la información en la primera trama SID recibida, un número N de tramas de retención; y
 - generar (804) ruido de confort basándose en las N tramas de retención.
- 35
- 5.- El método de acuerdo con la reivindicación 4, en el que el número N de tramas de retención es variable dinámicamente basándose en las propiedades de una señal de audio de entrada.
- 6.- El método de acuerdo con la reivindicación 4 o 5, en el que la primera trama SID recibida comprende además parámetros SID.
- 40
- 7.- Un codificador (900, 1000), que funciona para codificar conversación y aplicar un esquema de transmisión discontinua, DTX, que comprende la transmisión de las tramas de descriptor de inserción de silencio, SID, durante la inactividad de conversación, el codificador comprendiendo medios de procesamiento operativos para:
- 45
- determinar un número N de tramas de retención, en el que el número N de tramas de retención es variable;
 - transmitir las N tramas de retención a un decodificador; y para
 - transmitir una primera trama SID al decodificador después de un período de retención,
- 50
- donde la primera trama SID comprende información que indica el número determinado N de tramas de retención.
- 8.- El codificador de acuerdo con la reivindicación 7, en el que los medios de procesamiento comprenden un procesador (903) y una memoria (904) y en el que dicha memoria contiene instrucciones (905) ejecutables por dicho procesador.
- 55
- 9.- El codificador de acuerdo con la reivindicación 7 u 8, en el que el número N de tramas de retención varía dinámicamente basándose en las propiedades de una señal de audio de entrada.
- 60
- 10.- Un decodificador (1100, 1200) que funciona para decodificar la conversación y aplicar un esquema de transmisión discontinua, DTX, que comprende la recepción de tramas de descriptor de inserción de silencio, SID, y la generación de ruido de confort durante la inactividad de conversación, el decodificador comprendiendo medios de procesamiento operativos para:
- 65
- recibir N tramas de retención de un codificador;

ES 2 748 144 T3

- recibir una primera trama SID después de recibir las N tramas de retención;
- determinar, basándose en la información en la primera trama SID recibida, un número N de tramas de retención; y

5 - generar ruido de confort basándose en las N tramas de retención.

11.- El decodificador de acuerdo con la reivindicación 10, en el que los medios de procesamiento comprenden un procesador (1103) y una memoria (1104) y en el que dicha memoria contiene instrucciones (1105) ejecutables por dicho procesador.

10 12.- El decodificador de acuerdo con la reivindicación 10 u 11, en el que el número N de tramas de retención varía dinámicamente basándose en las propiedades de una señal de audio de entrada.

15 13.- Un programa informático (905, 1105), que comprende el código del programa informático, que cuando se ejecuta en un codificador hace que el codificador realice el método de acuerdo con cualquiera de las reivindicaciones 1 a 3.

20 14.- Un programa informático (905, 1105), que comprende un código de programa informático, que cuando se ejecuta en un decodificador hace que el decodificador realice el método de acuerdo con cualquiera de las reivindicaciones 4 a 6.

15.- Un producto de programa informático que comprende un programa informático (905, 1105) de acuerdo con la reivindicación 13 o 14.

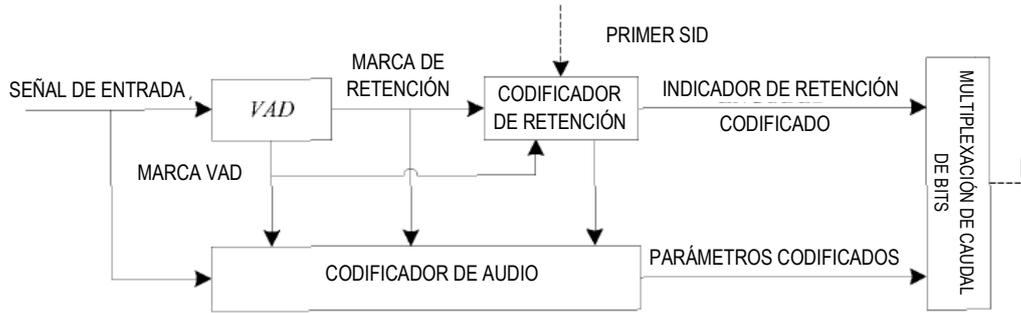


Figura 1

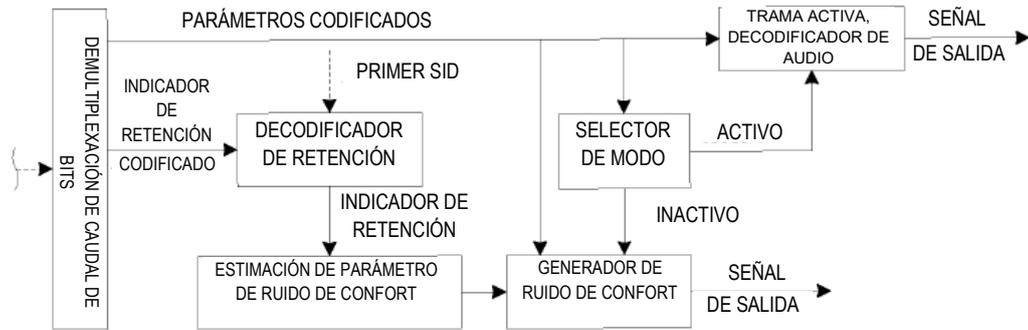


Figura 2

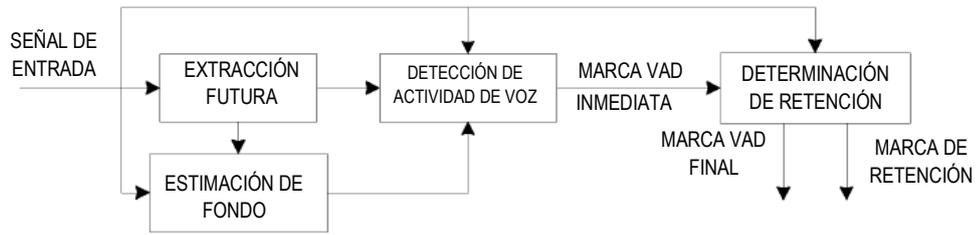


Figura 3

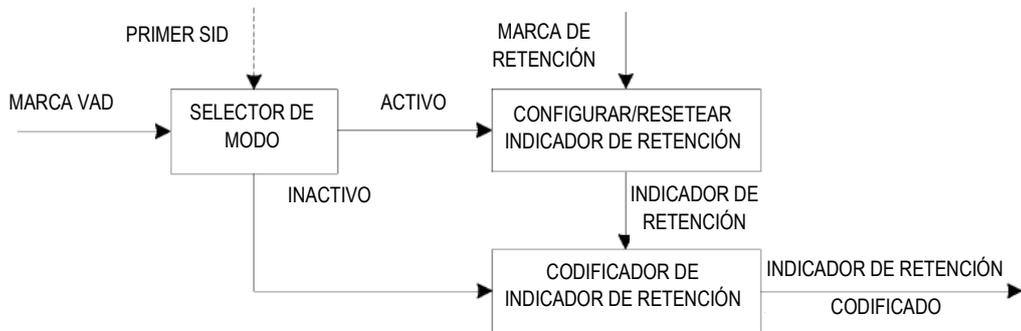


Figura 4

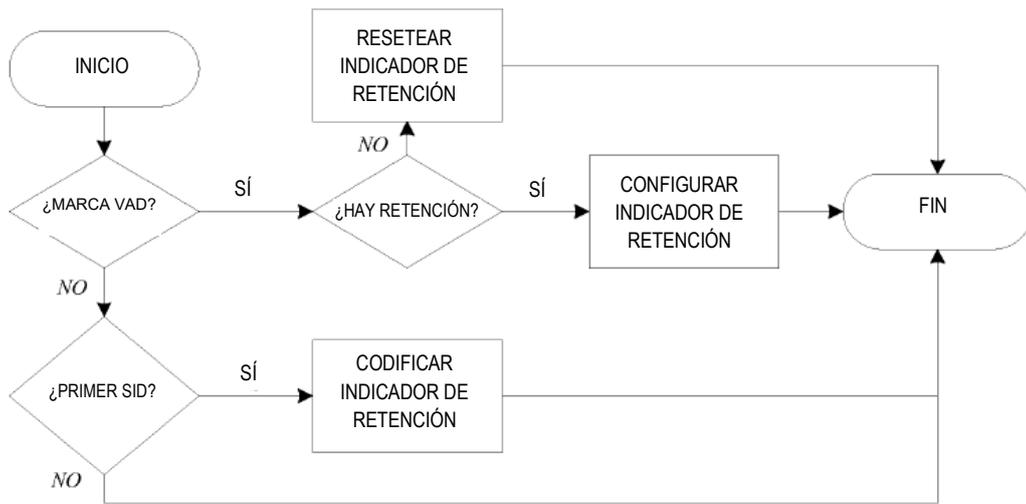


Figura 5

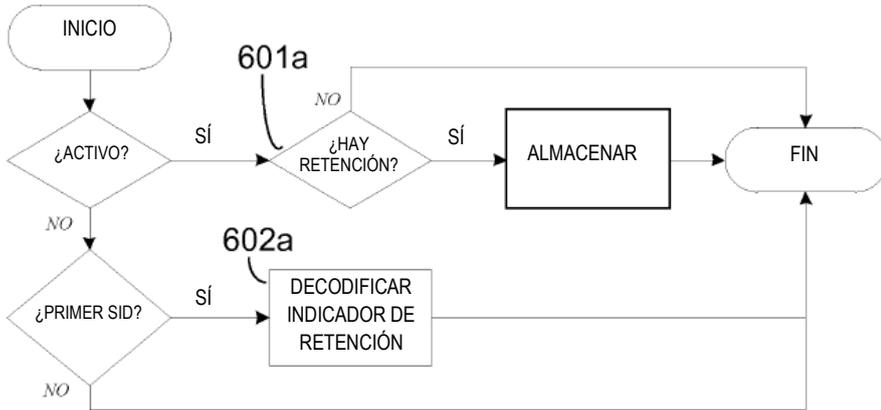


Figura 6a

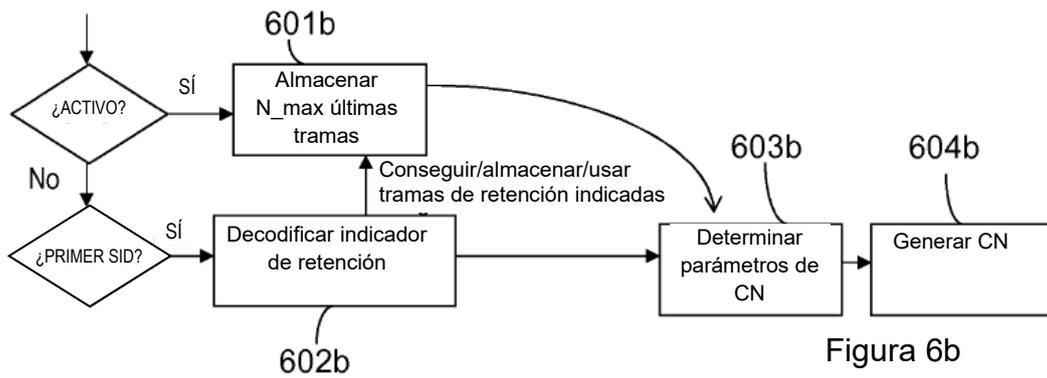


Figura 6b

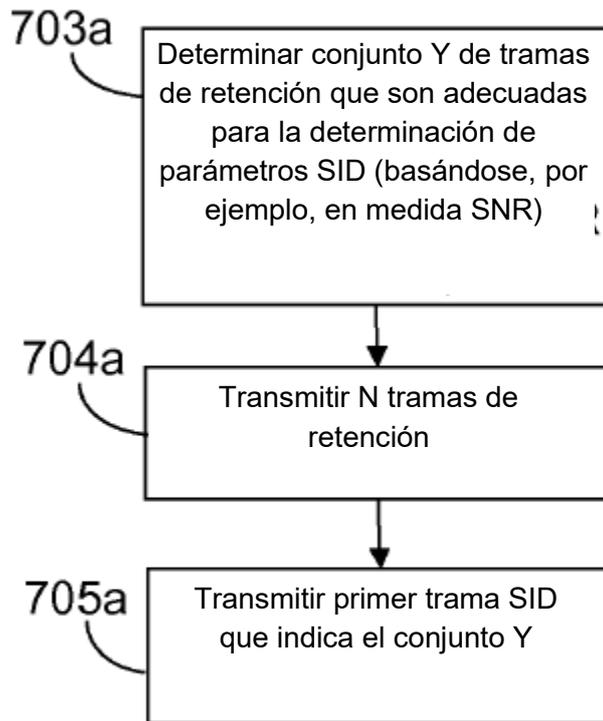


Figura 7a

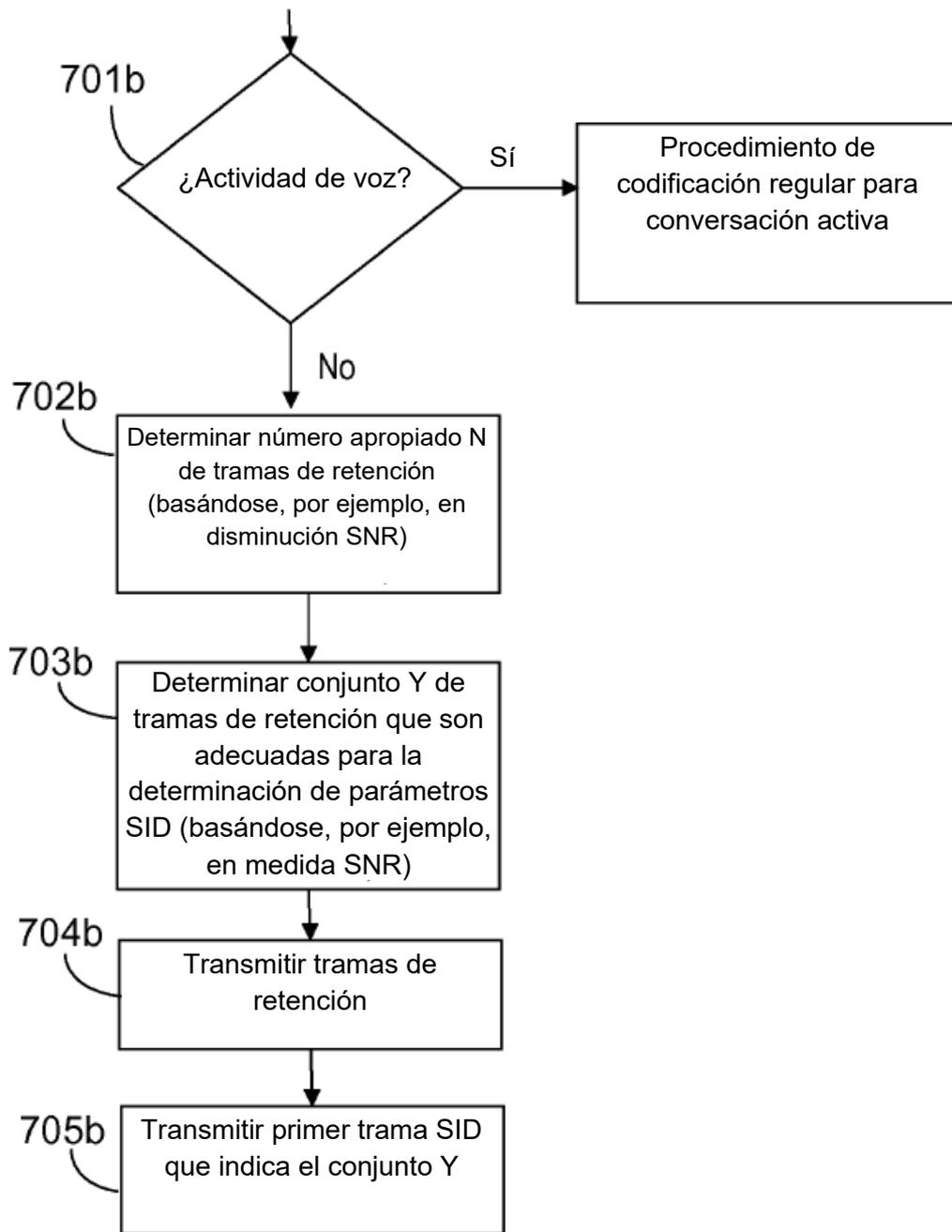


Figura 7b

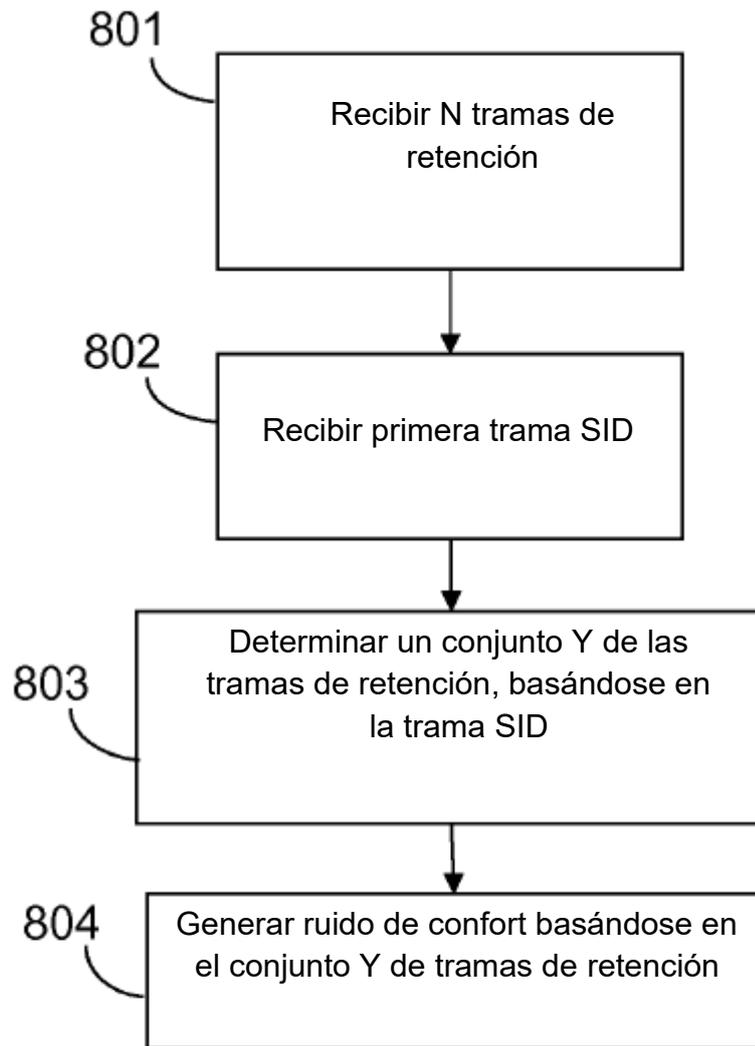


Figura 8

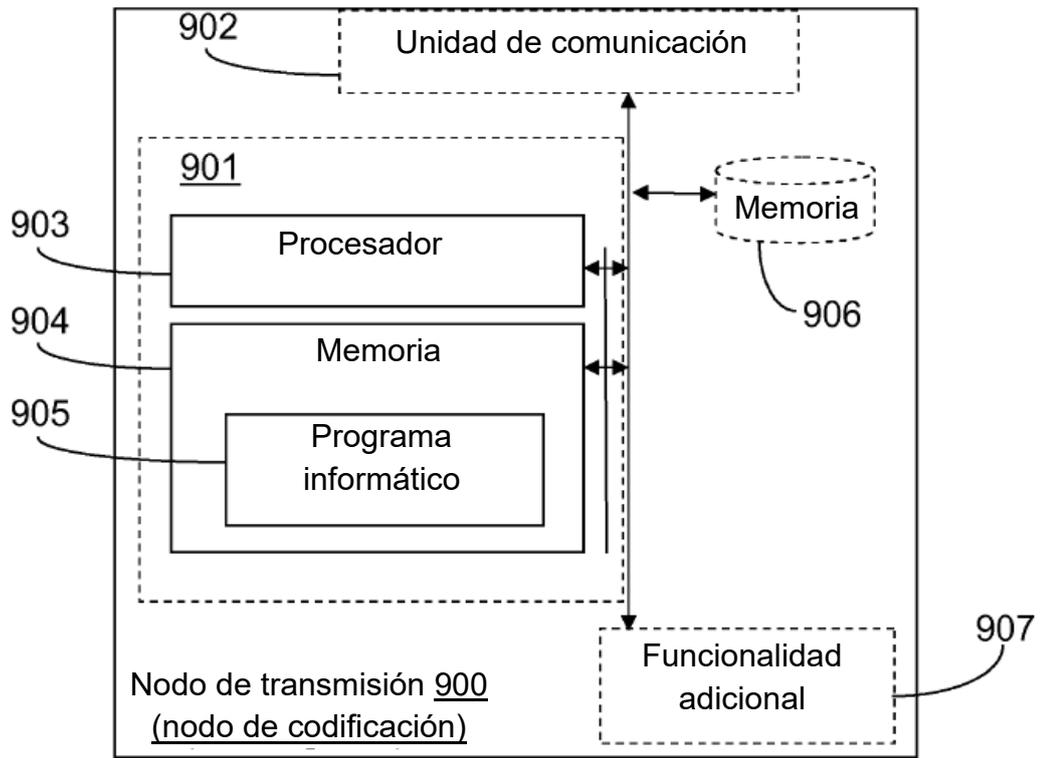


Figura 9

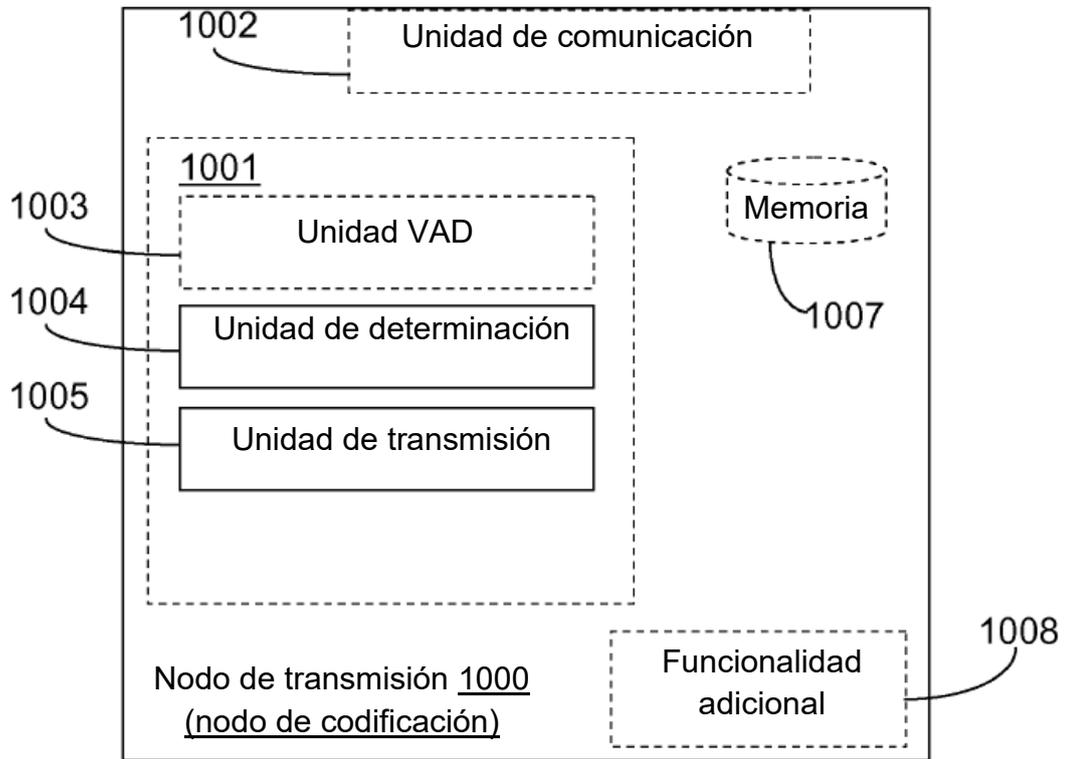


Figura 10

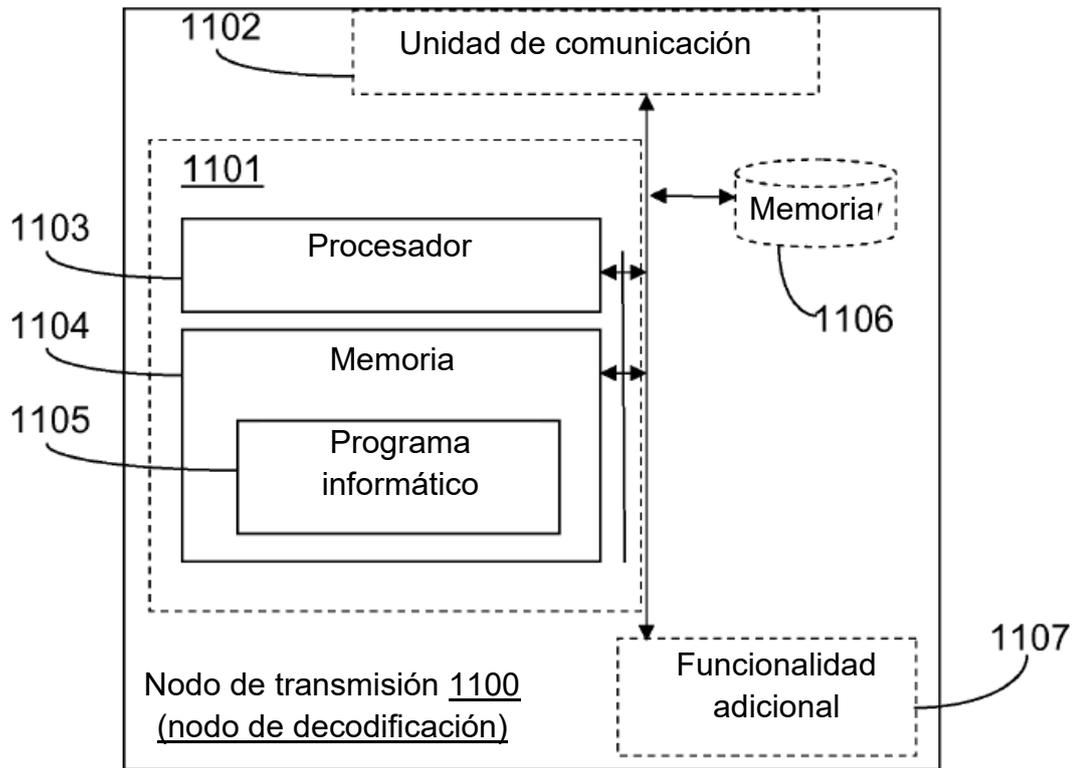


Figura 11

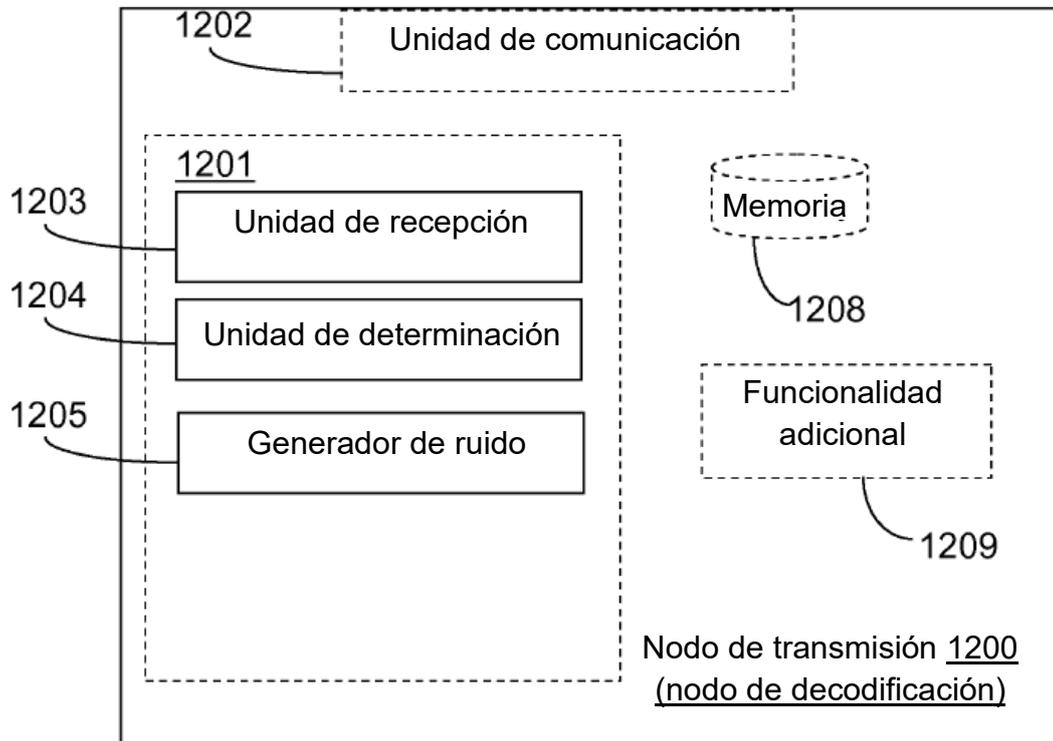


Figura 12