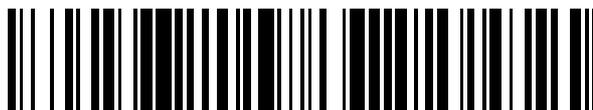


19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 586 635**

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 13818850 (3)**

97 Fecha y número de publicación de la concesión europea: **15.06.2016 EP 2959480**

54 Título: **Métodos y aparatos para Hangover de 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:

17.10.2016

73 Titular/es:

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

72 Inventor/es:

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

74 Agente/Representante:

DE ELZABURU MÁRQUEZ, Alberto

ES 2 586 635 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 Hangover de DTX en codificación de audio

Campo técnico

5 La solución descrita en la presente memoria se refiere en general a codificación de audio, y en particular a tramas de hangover asociadas a transmisión discontinua (DTX) en codificación de audio.

Antecedentes

10 Los estándares actuales de codificación de audio o de habla como el 3GPP AMR (3GPP TS 26.071) y el AMR-WB (3GPP TS 26.171), así como varios estándares ITU-T de codificación de habla (por ejemplo, Recomendación G.729 de ITU-T, Recomendación G.718 de ITU-T), incluyen un esquema de transmisión discontinua (DTX) que suspende la transmisión de habla durante la inactividad de habla, y en cambio transite tramas de Descriptor de Inserción de Silencio (SID) a tasas de bits y tasas de transmisión de trama significativamente reducidas en comparación con las usadas para habla activa codificada. El propósito de DTX es incrementar la eficacia de la transmisión, lo que a su vez reduce el coste de la comunicación de habla Y/o incrementa el número de conexiones telefónicas simultáneamente posibles en un sistema de comunicación dado.

15 Los sistemas de comunicación del estado de la técnica actual con DTX transmiten tramas regulares de codificación de habla durante segmentos de habla activa. Durante segmentos inactivos, por ejemplo las pausas de habla, estos sistemas transmiten en cambio tramas de SID a partir de las cuales el receptor genera lo que se conoce como ruido de confort como señal de sustitución respecto a la señal de inactividad. Con el fin de conseguir la mejor eficacia de DTX posible, resulta deseable que se transmitan solamente tramas de codificación de habla durante el habla activa y no en segmentos de habla inactiva, por ejemplo durante las pausas de habla.

20 Con el fin de hacer esta distinción entre habla e inactividad, se usa un detector de actividad de voz (VAD) en el lado de codificación, o de envío. Durante tramas correspondientes a segmentos de habla activa, se eleva una banderola de VAD. Este concepto adolece en la práctica, y especialmente en situaciones de habla en ruido de fondo, de errores de clasificación de VAD. Es decir, los períodos de inactividad se clasifican como períodos de habla activa, y/o viceversa. Uno de los problemas principales de los VADs es la detección de los puntos extremos del habla, es decir, el momento preciso en que la señal cambia desde habla activa a inactividad. La principal razón de este problema es que muchas desviaciones de habla decaen lentamente antes de que el habla se interrumpa realmente, de tal modo que la finalización de las ráfagas de conversación pueden ser cubiertas muy bien por el ruido de fondo. La consecuencia de este problema puede ser que tales desviaciones de habla sean clasificadas como inactividad, lo que puede dar como resultado que las tramas de señal correspondientes no sean codificadas, transmitidas y reconstruidas como habla activa sino por el contrario como señal de silencio, para la que se generan tramas de ruido de confort. Esto significa que las desviaciones de habla (fin de los períodos de habla) pueden ser percibidas y cortadas, lo que conduce a una calidad, e incluso una inteligibilidad significativamente reducidas del habla reconstruida. En otras palabras, esto puede conducir a una mala experiencia del usuario.

35 Los códecs del estado de la técnica actual como el AMR y el AMR-WB resuelven este problema simplemente retardando el inicio de la operación de DTX con síntesis de ruido de confort un número de tramas tras la desviación detectada de VAD. Esto se hace con una lógica de control de DTX en el codificador, que extiende o añade un período de tiempo durante el que se codifica una señal de entrada como habla activa incluso aunque la banderola de VAD indique inactividad. Este período se denomina período de hangover y en caso de AMR y de AMR-WB, el período de hangover tiene una longitud de 7 tramas.

45 El período de hangover no se usa solamente como un medio para evitar un recorte de la parte final (o desviación) del habla sino también para análisis de parámetro de trama de SID. En caso de AMR y de AMR-WB, los primeros parámetros de trama de SID tras un tramo de conversación (suficientemente largo) no se transmiten, sino que en su caso son calculados por el descodificador a partir de los parámetros de trama de habla recibidos y almacenados durante el período de hangover (3GPP TS 26.092; 3GPP TS 26.192). El propósito de realizar el cálculo de parámetro de trama del SID en base a los parámetros de trama de habla recibida durante el período de hangover es el de ahorrar recursos de transmisión que en otro caso habrían sido consumidos en transmisión de trama de SID, y el de minimizar el efecto de errores potenciales de transmisión sobre los primeros parámetros de trama de SID.

50 El problema principal del período de hangover en las soluciones descritas del estado de la técnica es que compromete la eficacia del esquema del DTX. Las tramas de hangover se codifican como habla activa a pesar de que son probablemente tramas de inactividad. Si el habla comprende ráfagas frecuentes de conversación separadas entre períodos de inactividad, entonces se codifica un número significativo de tramas con una alta tasa de bits, así como tramas de habla, en vez de como tramas de ruido de confort.

55 Un problema relativo se presenta si el período de hangover se acorta con el fin de mejorar la eficacia del esquema de DTX. Cuanto más corto sea el período de hangover, mayor probabilidad hay de que no represente apropiadamente la señal de ruido de inactividad. Esto puede conducir entonces a degradaciones audibles de la síntesis de ruido de confort inmediatamente después de las ráfagas de conversación.

En AMR y AMR-WB, el codificador y el descodificador mantienen el rastreo de las tramas de hangover de DTX usando una máquina de estado que necesita ser síncrona en el codificador y en el descodificador.

Sumario

5 Sería deseable, en el lado de un descodificador de audio, generar ruido de confort que sea representativo del ruido de fondo en el lado de un codificador de audio. Además, es deseable hacer esto de una forma eficiente, usando solamente un mínimo de recursos. Así, un objetivo de la solución sugerida en la presente memoria consiste en permitir la generación de ruido de confort que sea representativo de ruido de fondo en el lado de un codificador, y hacerlo usando una cantidad limitada de recursos.

10 La solución sugerida en la presente memoria aumenta la eficacia de las transmisiones de habla con DTX sin comprometer la calidad de la síntesis de ruido de confort al final de ráfagas de conversación.

15 Según un primer aspecto, se proporciona un método llevado a cabo por un nodo de transmisión o nodo de codificación. El nodo de transmisión es operable para codificar audio, tal como habla, y para comunicar con otros nodos o entidades, por ejemplo en una red de comunicación. El nodo de transmisión es además operable para aplicar un esquema de DTX que comprende la transmisión de tramas de SID durante inactividad de habla. El método comprende determinar, a partir de un número N de tramas de hangover, un conjunto Y de tramas que son representativas de ruido de fondo. El método comprende además transmitir las N tramas de hangover, que comprenden dicho conjunto Y de tramas, a un nodo de recepción. El método comprende además transmitir una primera trama de SID al nodo de recepción asociada a la transmisión de las N tramas de hangover, donde la trama de SID comprende información que indica el conjunto Y determinado de tramas de hangover al nodo de recepción. El método anterior permite que el nodo de recepción genere ruido de confort en base al conjunto Y de tramas de hangover.

25 Según un segundo aspecto, se proporciona un método llevado a cabo por un nodo de recepción o nodo de descodificación. El nodo de descodificación es operable para descodificar audio, tal como habla, y para comunicar con otros nodos o entidades, por ejemplo en una red de comunicación. El nodo de descodificación es además operable para aplicar un esquema de DTX que comprende recepción de tramas de SID y generación de ruido de confort durante inactividad de habla. El método comprende recibir N tramas de hangover desde un nodo de transmisión. Además, se recibe una primera trama de SID asociada a las N tramas de hangover. Se determina un conjunto Y de tramas de hangover, entre el número N de tramas de hangover recibidas, en base a la información de la trama de SID recibida. Además, se genera ruido de confort en base al conjunto Y de tramas de hangover.

30 Según un tercer aspecto, se proporciona un nodo de transmisión o de codificación. El nodo de transmisión es operable para codificar audio, tal como habla, y es operable para comunicar con otros nodos o entidades, por ejemplo en una red de comunicación. El nodo de transmisión es además operable para aplicar un esquema de DTX que comprende la transmisión de tramas de SID durante inactividad de habla. El nodo de transmisión comprende medios de procesamiento, por ejemplo en forma de un procesador y una memoria, en donde dicha memoria contiene instrucciones ejecutables por dicho procesador. Los medios de procesamiento son operativos para determinar, entre un número N de tramas de hangover, un conjunto Y de tramas que son representativas de ruido de fondo. Los medios de procesamiento son además operativos para transmitir las N tramas de hangover, comprendiendo dicho conjunto Y de tramas, hasta un nodo de recepción; y, además, para transmitir una primera trama de SID al nodo de recepción asociada a la transmisión de las N tramas de hangover, donde la trama de SID comprende información que indica el conjunto determinado Y de tramas de hangover al nodo de recepción.

45 Según un cuarto aspecto, se proporciona un nodo de recepción o nodo de descodificación. El nodo de recepción es operable para descodificar audio, tal como habla, y es operable para comunicar con otros nodos o entidades. El nodo de transmisión es además operable para aplicar un esquema de DTX que comprende recibir tramas de SID durante inactividad de habla. El nodo de recepción comprende medios de procesamiento, por ejemplo en forma de un procesador y una memoria, y en donde dicha memoria contiene instrucciones ejecutables por dicho procesador. Los medios de procesamiento son operativos para recibir N tramas de hangover desde un nodo de transmisión; y, además, para recibir una primera trama de SID asociada a las N tramas de hangover. Los medios de procesamiento son además operativos para determinar, en base a la información de la trama de SID recibida, un conjunto Y de tramas de hangover, entre el número N de tramas de hangover; y, para generar ruido de confort en base al conjunto Y de tramas de hangover.

Según un quinto aspecto, se proporciona un programa de ordenador, que comprende un código de programa de ordenador, el cual, cuando se ejecuta en un nodo de transmisión, provoca que el nodo de transmisión lleve a cabo el método conforme al primer aspecto.

55 Según un sexto aspecto, se proporciona un programa de ordenador, que comprende un código de programa de ordenador, el cual, cuando se ejecuta en un nodo de recepción, provoca que el nodo de recepción lleve a cabo el método conforme al segundo aspecto.

Según un séptimo aspecto, se proporciona un producto de programa de ordenador que comprende el programa de ordenador conforme al quinto aspecto.

Según un octavo aspecto, se proporciona un producto de programa de ordenador que comprende el programa de ordenador conforme al sexto aspecto.

Breve descripción de los dibujos

5 Los que anteceden y otros objetos, características y ventajas de la solución divulgada en la presente memoria, resultarán evidentes a partir de la descripción más particular que sigue de realizaciones según se ha ilustrado en los dibujos que se acompañan. Los dibujos no están necesariamente a escala, poniéndose en cambio énfasis en la ilustración de los principios de la solución divulgada en la presente memoria.

La Figura 1 es un diagrama de bloques de un codificador. El codificador comprende un VAD y un codificador de hangover.

10 La Figura 2 es un diagrama de bloques de un descodificador que opera en DTX.

La Figura 3 es un diagrama de bloques de la lógica de determinación de VAD y de hangover.

La Figura 4 es un diagrama de bloques de un codificador de hangover.

La Figura 5 es un diagrama de flujo para el codificador de hangover.

Las Figuras 6a y 6b son diagramas de flujo para el descodificador de hangover.

15 Las Figuras 7a y 7b son diagramas de flujo que ilustran ejemplos de realizaciones de un método llevado a cabo por un nodo de transmisión o codificación, acordes con la solución sugerida en la presente memoria.

La Figura 8 es un diagrama de flujo que ilustra un ejemplo de realización de un método llevado a cabo por un nodo de recepción o descodificación, acorde con la solución sugerida en la presente memoria.

20 Las Figuras 9-10 son diagramas de bloques que ilustran ejemplos de realizaciones de un nodo de transmisión, acordes con la solución sugerida en la presente memoria.

Las Figuras 11-12 son diagramas de bloques que ilustran ejemplos de realizaciones de un nodo de recepción, acordes con la solución sugerida en la presente memoria.

Descripción detallada

25 Según se ha mencionado con anterioridad: En sistemas de comunicación que utilizan transmisión discontinua (DTX), la eficacia de transmisión se reduce cuando se usan técnicas de hangover para evitar la degradación de calidad debido a decisiones incorrectas del detector de actividad de voz (VAD).

30 En lo que se conoce como segmentos de señal inactiva, por ejemplo pausas de voz, se genera ruido de confort, en el lado de un descodificador, usando información transmitida en tramas de descriptor de inserción de silencio (SID). Si se usa también el período de hangover para análisis de parámetro de SID, la longitud de éste no es, con preferencia, tan larga como se requiere para cubrir decisiones de VAD incorrectas, sino ligeramente más largas para capturar características de señales de fondo. En general, la probabilidad de generación de ruido de confort apropiado se incrementará con períodos de hangover más largos. Por otra parte, los períodos de hangover largos rebajan la eficacia del sistema de comunicación que utiliza DTX puesto que las tramas de señal inactiva serán transmitidas como tramas de señal de habla a una tasa de bits y una tasa de transmisión de trama más altas. En sistemas de comunicación que usan estas técnicas, existe en consecuencia un compromiso entre la eficacia de transmisión y la probabilidad de ruido de confort representativo.

35 Un período de hangover tras una desviación de habla puede ser adaptativo. Respecto al codificador, esto significa que tras una conmutación de decisión de un VAD desde 1 (= habla activa) a 0 (= inactividad), se añade un período de hangover adaptativo. La información que especifica las tramas pertenecientes al período de hangover puede ser transmitida con la primera trama de SID tras el período de hangover. En la Figura 1 se ha representado un diagrama esquemático de bloques de un codificador de ese tipo.

40 El descodificador puede recibir, por ejemplo con la primera trama de SID, la indicación de cuáles de las tramas de habla activa recibidas previamente pertenecen al período de hangover. La información de habla codificada de las tramas pertenecientes al período de hangover puede ser usada posteriormente para el cálculo del parámetro de SID en el lado del descodificador. En la Figura 2, se muestra un diagrama esquemático de bloques del descodificador.

45 En la descripción que sigue, a efectos de explicación y no de limitación, se exponen detalles específicos tales como arquitecturas particulares, interfaces, técnicas, etc., con el fin de proporcionar una comprensión completa del concepto descrito en la presente memoria. Sin embargo, resultará evidente para los expertos en la materia que el concepto descrito puede ser puesto en práctica en otras realizaciones que se aparten de esos detalles específicos. Es decir, los expertos en la materia estarán capacitados para idear diversas disposiciones que, aunque no se describan o muestren explícitamente en la presente memoria, materializan los principios del concepto descrito y

están incluidas dentro de su espíritu y de su alcance. En algunos casos, se omiten descripciones detalladas de dispositivos, circuitos y métodos bien conocidos con el fin de no oscurecer la descripción conforme al presente concepto con detalles innecesarios. Todas las afirmaciones de la presente memoria que se refieren a principios, aspectos y realizaciones del concepto descrito, así como los ejemplos específicos de los mismos, están destinados a abarcar equivalentes tanto estructurales como funcionales de los mismos. Adicionalmente, se pretende que tales equivalentes incluyan tanto equivalentes actualmente conocidos como también equivalentes desarrollados en el futuro, por ejemplo cualesquiera elementos que se desarrollen que lleven a cabo la misma función, con independencia de su estructura.

Así, por ejemplo, los expertos en la materia podrán apreciar que los diagramas de bloques de la presente memoria pueden representar vistas conceptuales de circuitería ilustrativa o de otras unidades funcionales que materialicen los principios de la solución. De forma similar, se apreciará que cualesquiera diagramas de flujo, diagramas de transición de estado, seudo-código, y similares, representan diversos procesos que pueden ser representados sustancialmente en un medio legible con ordenador y por tanto ejecutados mediante un ordenador o procesador, se haya representado o no explícitamente tal ordenador o procesador.

Las funciones de los diversos elementos que incluyen bloques funcionales, incluyendo aunque sin limitación los etiquetados o descritos como, por ejemplo, "ordenador", "procesador", o "controlador", pueden ser proporcionados con el uso de hardware tal como hardware de circuitos y/o hardware capaz de ejecutar software en forma de instrucciones codificadas, almacenadas en un medio legible con ordenador. De ese modo, tales funciones y bloques funcionales ilustrados han de ser entendidos ya sea como implementados en hardware y/o implementados en ordenador, y por tanto implementados con máquina.

En términos de implementación de hardware, los bloques funcionales pueden incluir o abarcar, sin limitación, hardware de procesador de señal digital (DSP), equipo de conjunto reducido de instrucciones, circuitería de hardware (por ejemplo, digital o analógica) que incluya, aunque sin limitación, Circuito(s) Integrado(s) Específico(s) de la Aplicación (ASICs), y (cuando sea adecuado) máquinas de estado capaces de realizar tales funciones.

En ejemplos de realización de la solución sugerida en la presente memoria, la longitud de un período de hangover, es decir, el número de tramas de hangover, puede ser variable y adaptativo. Se puede generar un período de hangover adaptativo, por ejemplo, en respuesta a la decisión de VAD y un indicador adicional. En la Figura 3, se muestra un diagrama esquemático de bloques del VAD. La decisión inmediata de VAD puede ser una banderola correspondiente al habla inmediata/clasificación de inactividad del VAD. Siempre que el VAD clasifica una trama de señal como habla activa, puede emerger esta banderola, y en otro caso puede ser bajada. Se puede introducir una banderola de hangover para controlar la longitud del período de hangover añadido después de que se haya bajado la banderola de VAD inmediata. Esto se hace preferiblemente de tal modo que se asegure que la señal de las tramas de hangover comprende principalmente una porción representativa del ruido de fondo y que las porciones de habla potencialmente restantes son despreciables. Esto se hace con el propósito de permitir una estimación fiable del parámetro de SID en el lado de descodificación, cuya estimación es representativa en cuanto a la señal de ruido de inactividad, y la cual no está afectada por las porciones de habla potencialmente restantes. Una medición útil para basar la banderola de hangover, es la relación estimada de señal respecto a ruido (SNR), la cual compara el nivel estimado del habla restante con el nivel de ruido de inactividad estimado. Por ejemplo, cuando esta estimación de SNR está por encima de un cierto umbral, la banderola de hangover puede ser levantada, y cuando cae por debajo de dicho umbral el período de hangover puede terminar. Debe apreciarse que la lógica de determinación de hangover puede generar una banderola de VAD final que podría ser diferente de la banderola de VAD inmediata presente en su entrada.

Por ejemplo, la longitud del período de hangover puede ser adaptada en respuesta a la SNR estimada. Esto supone que la SNR disminuye al final de una ráfaga de conversación. La adaptación tiene en cuenta que el grado de reducción de SNR puede ser variable de una ráfaga de conversación a otra. El resultado es que la longitud del período de hangover en tramas es un parámetro variable. Según un ejemplo de realización, esta longitud de hangover, es decir, el indicador de hangover, se codifica y se transmite al descodificador. Un diagrama esquemático de bloques de un codificador de hangover ha sido representado en la Figura 4. Adicionalmente a las banderolas de VAD y de hangover, el codificador del ejemplo de hangover usa una primera banderola de SID. La primera banderola de SID puede indicar si la trama actual es el primer SID a continuación de de la codificación de señal activa. Debe apreciarse que las banderolas no tienen necesariamente que ser señaladas explícitamente como variables específicas, sino que podrían ser implícitas, por ejemplo derivables de otras variables de estado de codificador. La longitud codificada del período de hangover puede ser transmitida como parte de la información comprendida en la primera trama de SID transmitida después del final de la transmisión de tramas de habla activa. La Figura 5 muestra un diagrama de flujo genérico para el codificador de indicador de hangover.

Según un ejemplo de realización de la solución sugerida en la presente memoria, la longitud del período de hangover tras la caída de la banderola de VAD inmediata está adaptada de tal modo que el conjunto de tramas a considerar para estimación de parámetro de SID es una variable. Es decir, el número de tramas de hangover puede ser fijo o variable, pero el conjunto de tramas a considerar para la determinación de parámetros de SID para generación de ruido de confort no es necesariamente igual al número de tramas de hangover. En esta alternativa, se supone que existe una medición que indica la idoneidad de cada trama del período de hangover a continuación de la

caída de la banderola de VAD inmediata para la estimación de parámetro de SID. Por ejemplo, las tramas para las que la medición está por encima de un cierto umbral, pueden ser consideradas como representativas del ruido de fondo y por lo tanto adecuadas para la estimación de parámetro de SID. La medición, al igual que en lo que antecede, puede estar basada en estimaciones de SNR. Entonces, conforme a esta realización, la primera trama de SID tras el final de la transmisión de tramas de habla activa puede contener información acerca del conjunto específico de tramas que van a ser usadas para estimación de parámetro de SID.

Como ejemplo, el conjunto puede comprender las n tramas que preceden a la primera trama de SID. La codificación de las tramas que se van a usar para la estimación de parámetro de SID, puede realizarse entonces con una palabra de código de N -max-bits, donde cada bit representa una trama respectiva que precede a la primera trama de SID. Si se establece un bit en la palabra de código ($= 1$), la trama representada por el bit podrá ser usada para estimación de parámetro de SID, en otro caso no.

La medición de SNR que se usa en las realizaciones anteriores, es solamente un ejemplo. Además, son posibles mediciones más avanzadas. En general, una medición 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 medición más avanzada de ese tipo puede compararse, por ejemplo, la potencia o las propiedades espectrales de la trama actual con las propiedades correspondientes de tramas recientes o de otras tramas recientes en las que se haya identificado que contienen ruido.

Podría parecer como posibilidad en la corriente normal de bits de tramas codificadas incluir un bit para señalar si la trama codificada es una trama de hangover o no. Sin embargo, esto se considera menos ventajoso puesto que podría significar que tendría que reservarse un bit de cada trama de habla para información que se usa solamente después de la terminación de una ráfaga de habla.

Mientras que los párrafos anteriores discuten el hangover específico de DTX, también es habitual que ya el VAD añada algún hangover para evitar el recorte en la desviación de habla. Debería ser entonces posible permitir que el hangover específico de VAD y el hangover de DTX se solapen. Por ejemplo, el análisis de señal puede contribuir a la terminación anticipada del hangover si existe un número de tramas suficiente para generar un ruido de confort estable, con independencia de si las últimas tramas proceden del hangover de VAD o del hangover de DTX.

En la Figura 6a, un diagrama esquemático de flujo muestra un ejemplo de descodificador de indicador de hangover del lado del descodificador. En el ejemplo de 6a, puede estar indicado en cada trama si ésta es una trama de hangover o no, y las tramas de hangover se almacenan después. A partir del indicador de hangover descodificado, se puede determinar cuál de las tramas de hangover almacenadas podría haber sido usada como base para ruido de confort. Alternativamente, la decisión en 601a de si una trama es una trama de hangover o no, no se adopta hasta que el indicador de hangover ha sido descodificado en 602a. Para que la decisión se tome después de la descodificación 602a, un conjunto de las tramas recibidas más recientemente necesita ser almacenada en una memoria intermedia, por ejemplo, de longitud N_{max} (número máximo de tramas de hangover). En este último caso, las tramas de hangover pueden ser identificadas en el conjunto de tramas que esté almacenado en ese momento en la memoria intermedia, en base al indicador de hangover descodificado, y de ese modo, los parámetros de al menos parte de las tramas de hangover pueden ser almacenados. Esto se aprecia quizás más claramente en la Figura 6b, la cual muestra el almacenaje 601b de las últimas N_{max} tramas. Cuando el indicador de hangover se descodifica en 602b, las tramas de hangover están presentes entre las tramas almacenadas, y los parámetros de ruido de confort pueden ser determinados 603b en base a las tramas de hangover indicadas por el indicador de hangover. A continuación se puede generar 604b ruido de confort en base a los parámetros. Al igual que en el codificador, la primera banderola de SID puede indicar si la trama actual es el primer SID que sigue a la codificación de señal activa. La primera banderola de SID no tiene necesariamente que estar almacenada en una variable, sino que puede ser derivada de otras variables de estado del descodificador.

Los parámetros de SID típicos son parámetros de ganancia y parámetros espectrales predictivos lineales como los parámetros de frecuencia espectral de línea (LSF). En un ejemplo de realización, el descodificador puede obtener esos parámetros a partir de las 5 tramas precedentes y calcular los promedios de las mismas. Los parámetros promediados pueden ser usados a continuación en la síntesis de ruido de confort del sistema de DTX. Alternativamente, los parámetros de SID usados para síntesis de ruido de confort pueden ser determinados a partir de un conjunto específico de las tramas de hangover indicadas. El conjunto específico puede ser derivado en el lado del descodificador usando, por ejemplo, el parámetro de longitud de hangover recibido y parámetros de las tramas previamente recibidas que han sido almacenadas en una memoria.

Incluso aunque los parámetros derivados de un conjunto de tramas de hangover sean mencionados en la mayor parte de los casos como parámetros de SID en el presente documento, también debería ser posible usar otros parámetros, indicados de forma diferente, pero que sirvan para los mismos propósitos, en particular el de ser una base para la generación de ruido de confort.

El descodificador puede obtener, por ejemplo a partir del indicador de hangover en la primera trama de SID tras una secuencia de tramas de habla activa, información acerca del conjunto específico de tramas precedentes que van a ser usadas para el cálculo del parámetro de SID. A continuación, los parámetros de SID pueden ser calculados

usando, por ejemplo, los parámetros de ganancia y espectrales de las tramas que sean identificadas mediante el código recibido. Suponiendo que se use una palabra de código de $n=8$ bits como indicador de hangover y que esta palabra de código contiene la secuencia de bits "0 1011111", entonces se usan las 5 tramas directamente precedentes y la 7ª trama precedente. Los parámetros de ganancia y espectrales de estas tramas pueden ser promediados y usados posteriormente en la síntesis de ruido de confort del sistema de DTX.

En los párrafos que siguen, se van a describir con mayor detalle diferentes aspectos de la solución divulgada en la presente memoria, con referencia a determinadas realizaciones y a los dibujos que se acompañan. Con fines explicativos y no limitativos, se exponen detalles específicos, tal como escenarios y técnicas particulares, con el fin de proporcionar una comprensión completa de las diferentes realizaciones. Sin embargo, otras realizaciones pueden apartarse de esos detalles específicos.

Ejemplo de método llevado a cabo por un nodo de transmisión/codificación, Figura 7

Un ejemplo de método llevado a cabo por un nodo de transmisión o nodo de codificación, va a ser explicado en lo que sigue con referencia a la Figura 7a. El nodo de transmisión es operable para codificar audio, tal como habla, y para comunicar con otros nodos o entidades, por ejemplo en una red de comunicación. El nodo de transmisión es además operable para aplicar un esquema de DTX que comprende transmisión de tramas de SID durante inactividad de habla. El nodo de transmisión puede ser, por ejemplo, un teléfono celular, una tableta, un ordenador o cualquier otro dispositivo capacitado para comunicación alámbrica y/o inalámbrica, y para codificación de audio.

La Figura 7a ilustra el método que comprende determinar 703a, a partir de un número N de tramas de hangover, un conjunto Y de tramas que sean representativas de ruido de fondo. El método comprende además transmitir 704a las N tramas de hangover, comprendiendo dicho conjunto Y de tramas, hasta un nodo de recepción. El método comprende además transmitir 705a una primera trama de SID hasta el nodo de recepción asociada a la transmisión de las N tramas de hangover, donde la trama de SID comprende información indicativa del conjunto Y determinado de tramas de hangover, hasta el nodo de recepción. El método anterior permite que el nodo de recepción genere ruido de confort en base al conjunto Y de tramas de hangover.

El orden de las acciones de las Figuras 7a y b es solamente un ejemplo. Por ejemplo, el conjunto Y podría ser determinado después de que las N tramas de hangover hayan sido transmitidas.

Las tramas comprendidas en el conjunto Y de tramas de hangover deberán ser representativas de ruido de fondo. De ese modo, fuera del número N de tramas de hangover, se podrán identificar las que sean más adecuadas para determinar o calcular los parámetros para la generación de ruido de confort, por ejemplo los llamados parámetros de SID. Las tramas del conjunto Y pueden ser determinadas o identificadas, por ejemplo en base a un nivel de SNR de la señal comprendida en cada trama, y cuando este nivel de SNR cumpla un determinado criterio, se determina que la trama es adecuada para su uso como base para el cálculo de, por ejemplo, parámetros de SID. Algunas de las N tramas de hangover pueden ser menos representativas de ruido de fondo. Por ejemplo, algunas de las tramas de hangover pueden comprender, al menos parcialmente, habla o ruido transitorio, lo que hace que las mismas sean inadecuadas como base para la extracción de parámetros relacionados con la generación de ruido de confort. Por ejemplo, las tramas de habla tienen en general estructuras formantes, las cuales no se aprecian en el ruido de fondo; y las tramas de ruido transitorio pueden tener energía más alta que el ruido de fondo promedio. Tales tramas de hangover, no representativas de ruido de fondo, no deben ser incluidas en el conjunto Y .

El conjunto Y de tramas puede ser indicado de diferentes maneras en la primera trama de SID, lo cual va a ser mejor descrito más adelante. Mediante "primera trama de SID" se entiende la primera trama de SID en un período de DTX, típicamente indicadora del inicio del período de DTX. Mediante período de DTX se entiende aquí un período de inactividad de habla, durante el que se envían tramas codificadas desde el nodo de transmisión hasta el nodo de recepción a una tasa de bits y/o una tasa de trama más bajas que durante los períodos de no DTX. Mediante período de DTX se entiende aquí el período entre ráfagas de habla activa, cuyo período se reemplaza por ruido de confort. Estos períodos empiezan con el primer SID para marcar la transición a ruido de confort. Esto va seguido normalmente por períodos de un número de tramas de "NO_DATA", las cuales, según indica su nombre, no contienen ningún dato, ni las tramas de SID (SID_UPDATE). Las tramas de SID son transmitidas, con mayor frecuencia, a intervalos regulares, indicados como "intervalo de SID", hasta que la siguiente conversación dispare una transición de nuevo a codificación de habla activa. Es decir, con un intervalo de SID de 8, el período de DTX podría ser codificado como: primer SID seguido de 7 tramas de NO_DATA antes del SID_UPDATE. Esta secuencia con 7 tramas de NO_DATA seguida de una actualización de SID se repite después hasta que ocurre la transición a habla activa.

Una ventaja del método descrito anteriormente consiste, según se ha descrito con anterioridad, en que permite que un nodo de recepción deduzca parámetros para ruido de confort a partir de tramas que se ha determinado que son adecuadas para este propósito. Esto mejora la calidad del ruido de confort generado, y por lo tanto mejora la experiencia del usuario. El conjunto Y se indica además al nodo de recepción de una manera muy eficiente en cuanto a recursos, utilizando la primera trama de SID para este propósito. Es una ventaja determinar las tramas de hangover adecuadas en el nodo de transmisión, puesto que en este nodo, los datos reales de la señal de audio son accesibles, mientras que en el nodo de recepción solamente está disponible una versión cuantificada de los datos.

La información indicativa del conjunto Y puede comprender un número, el cual implica un número de tramas de hangover en secuencia; una palabra de código o mapa de bits, que indica las posiciones de las tramas pertenecientes al conjunto Y, entre las N tramas de hangover; una palabra de código o mapa de bits, que indica algunas de las N tramas de hangover que están comprendidas en el conjunto Y, y/o una palabra de código o mapa de bits, que indica cuales de las N tramas de hangover no están comprendidas en el conjunto Y.

Por ejemplo, la trama de SID podría comprender un número, por ejemplo 5, que deberá ser interpretado, por el nodo de recepción, por ejemplo, como que las últimas cinco tramas de hangover deberían ser usadas para determinar parámetros para la generación de ruido de confort. Alternativamente, el número podría ser interpretado como algún otro grupo de cinco tramas entre las N tramas de hangover, tal como las cinco últimas menos una. El número N de tramas de hangover podría ser, por ejemplo, 6, 7, 8 ó 9. En un caso especial, el número N de tramas de hangover podría ser igual al número indicado en la trama de SID, es decir, los parámetros podrían ser determinados entonces en base a todas las tramas de hangover.

Alternativamente, o adicionalmente, la trama de SID podría comprender una palabra de código o mapa de bits/máscara de bits que indique las posiciones de las tramas pertenecientes al conjunto Y. Dicha palabra de código podría estar configurada de diferentes formas. Se podría usar un sistema de código, en el que tanto el nodo transmisor como el nodo receptor tengan conocimiento del significado de los códigos, por ejemplo ambos lados tengan acceso a un libro de códigos que especifique, por ejemplo, que la palabra de código "01" mapea en tramas de hangover, en la trama k, k-1, k-2, k-4 y k-6 entre las N tramas de hangover. Alternativamente, se podría usar un mapa de bits/máscara de bits. Este mapa de bits podría cubrir la totalidad de las N posiciones de las N tramas de hangover, o un subconjunto de las N posiciones. El nodo de recepción podría, en algún momento, haber sido informado previamente del carácter del mapa de bits/máscara de bits. Por ejemplo, si N=8, un ejemplo de mapa de bits/máscara de bits tal como "11011000" podría estar comprendido en la trama de SID, indicando que la 4ª, 5ª, 7ª y 8ª tramas previas podrían ser usadas para determinar parámetros para ruido de confort. Alternativamente, el mapa de bits/máscara de bits "11011" podría estar comprendido en la primera trama de SID, con el mismo significado que en el ejemplo previo. Alternativamente, las posiciones de las tramas de hangover que no estén comprendidas en el conjunto Y podrían ser indicadas. En analogía con el ejemplo anterior, un mapa de bits/máscara de bits correspondiente podría ser entonces "00100111" o "00100" o "100111".

Todas estas son realizaciones diferentes de información que podrían estar comprendidas en la primera trama de SID con el fin de indicar cuáles de las tramas de hangover deberán ser usadas. En general, cuantos menos bits se necesiten para indicar en el conjunto Y, mejor será.

El concepto de transmisión discutido en lo que antecede, en la primera trama de SID, una identificación del conjunto de tramas de hangover en la que basar la generación de ruido de confort, puede ser combinada con parámetros SID de transmisión como parte de la primera trama de SID. Es decir, la primera trama de SID puede comprender además parámetros de SID. Estos parámetros de SID proporcionarán una indicación sobre cómo se observa la señal en la trama actual. Esta información podría, por ejemplo, ser más ponderada que la información de las tramas de hangover anteriores. Por supuesto, las tramas de hangover podrían ya ser ponderadas de forma diferente sin considerar los parámetros de señal de la trama de SID, pero en cualquier caso la decisión de no ir a DTX en la trama previa podría indicar que no estamos suficientemente seguros de que esta trama represente inactividad/solamente ruido de fondo.

El número N de tramas de hangover puede ser dinámicamente variable, según se ha descrito con anterioridad. El número N podría ser determinado en base a propiedades de una señal de audio entrante. Por ejemplo, el número N podría depender del sonido de habla que precede al período de DTX y/o del carácter del ruido de fondo. Usando un número dinámico de tramas de hangover, el número de tramas de hangover que necesita ser transmitido a un nodo de recepción podría ser mantenido en un mínimo, y de ese modo se podrían ahorrar recursos, en comparación con tener un número estático de tramas de hangover.

Algunas acciones, que pueden preceder al método ilustrado en la Figura 7a, han sido ilustradas en la Figura 7b. En la Figura 7b se determina, en una acción 701b, si una trama de una corriente de audio, por ejemplo un segmento de una señal de audio, cuya señal comprende al menos parcialmente habla, comprende habla activa o no. Esto se denomina con frecuencia Detección de Actividad de Voz, VAD. Cuando se determina que una o más tramas no comprenden habla activa, debe transmitirse un número de tramas de hangover, por ejemplo con el fin de reducir la probabilidad de cortar un sonido del habla, según se ha descrito con anterioridad. Cuando se aplica un número dinámico de tramas de hangover, la señal comprendida en las primeras tramas que se ha determinado que no comprenden habla activa, puede ser analizada, y se puede determinar un número adecuado de tramas de hangover en la acción 702b. Posiblemente, también se pueden obtener propiedades de las últimas tramas en las que se ha determinado que comprenden habla activa cuando se determina un número N apropiado de tramas de hangover, por ejemplo a efectos de determinar una SNR o una disminución de energía de trama entre tramas adyacentes.

Es decir, se puede determinar un número, N, de tramas de hangover en base a una propiedad de la señal comprendida en las tramas antes y/o después de una decisión de inactividad de habla. Además, o de forma alternativa, las propiedades de las tramas de señal previas que se ha determinado que comprenden solamente ruido de fondo podrían ser tomadas en consideración cuando se determina N.

Según se ha mencionado con anterioridad, la determinación de un número de tramas de hangover podría estar basada en una característica de disminución de SNR o de energía dentro de y/o entre tramas de señal. El número N de tramas de hangover puede ser estático, semiestático o dinámico, y podría ser diferente para diferentes desviaciones de habla.

- 5 Las tramas de hangover transmitidas al nodo de recepción, por ejemplo en la acción 704b, pueden ser codificadas en conformidad con la codificación de tramas que comprenden habla activa, según se ha descrito con anterioridad. Cuando el número N de tramas de hangover es dinámico, el número N podría ser indicado también al nodo de recepción, por ejemplo en la primera trama de SID.

Ejemplo de método llevado a cabo por un nodo de descodificación, Figura 8

- 10 Un ejemplo de método llevado a cabo por un nodo de recepción o nodo de descodificación, va a ser descrito a continuación con referencia a la Figura 8. El nodo de descodificación es operable para descodificar audio, tal como habla, y para comunicar con otros nodos o entidades, por ejemplo en una red de comunicación. El nodo de descodificación es operable además para aplicar un esquema de DTX que comprende recepción de tramas de SID y generación de ruido de confort durante inactividad de habla. El nodo de descodificación puede ser, por ejemplo, un
15 teléfono celular, una tableta, un ordenador o cualquier otro dispositivo capacitado para comunicación alámbrica y/o inalámbrica y para descodificación de audio.

- El ejemplo de método ilustrado en la Figura 8 comprende recibir, 801, N tramas de hangover desde un nodo de transmisión. Además, se recibe 802 una primera trama de SID asociada a las N tramas de hangover. Se determina
20 803, un conjunto Y de tramas de hangover, entre el número N de tramas de hangover, en base a la información contenida en la trama de SID recibida. Además, se genera 805 ruido de confort, en base al menos parcialmente al conjunto Y de tramas de hangover.

- La trama de SID podría ser recibida después de que se haya recibido la última de las N tramas de hangover, indicando el inicio de un período de DTX. Sin embargo, la trama de SID podría también ser recibida antes de las
25 tramas de hangover, o entre dos tramas de hangover, si esto está permitido y regulado en el protocolo de transmisión para el esquema de DTX.

- El número N de tramas de hangover podría ser indicado en la primera trama de SID, aunque esto es opcional. El número N podría ser establecido alternativamente en un valor por defecto, por ejemplo 7, lo que implica que las
30 últimas 7 tramas recibidas, sin contar la trama de SID, con anterioridad al período de DTX, serían las tramas de hangover. Además, cuando se aplica un número dinámico de tramas de hangover, existen otras formas de señalar el número N de tramas de hangover. Por ejemplo, el número podría ser señalado implícitamente a través de propiedades de la señal de audio, por ejemplo una medición de SNR a largo plazo. Tal medición podría ser generada en base a la señal de audio descodificada y podría estar por tanto disponible en el descodificador.

- La trama de SID comprende, según se ha descrito con anterioridad, información que indica un conjunto Y de tramas, entre las N tramas de hangover, seleccionadas por el nodo de transmisión como representativas de ruido de fondo.
35 Por lo tanto, es posible que el nodo de recepción determine el conjunto Y de tramas en base a la primera trama de SID. Es decir, en base a la información comprendida en la primera trama de SID que indica el conjunto Y. La información podría ser explícita o implícita, y ha sido ejemplificada más arriba cuando se ha descrito el método llevado a cabo por un nodo de transmisión.

- El nodo de recepción sirve para generar ruido de confort durante períodos silentes de DTX, es decir, durante
40 períodos en que no se reciben tramas de habla desde un nodo de transmisión. El ruido de confort debe mimetizar preferiblemente el ruido de fondo en el nodo de transmisión. Con el fin de generar un ruido de confort tan auténtico como sea posible, el nodo de recepción deberá estimar el ruido de fondo en base a las tramas de hangover que sean más representativas del ruido de fondo. Alternativamente o adicionalmente, el nodo de recepción podría recibir una estimación del ruido de fondo desde el nodo de transmisión, por ejemplo en forma de parámetros de SID. Las
45 tramas de SID están codificadas a una tasa de bits significativamente más baja que las tramas de señal activa. Las características del ruido de fondo son por lo tanto mejor capturadas, en el lado del codificador, durante el hangover (a partir de las tramas de hangover) que en el SID. Sin embargo, la inclusión de parámetros de SID en la primera trama de SID puede ser ventajosa a efectos de tener una transición suave desde tramas de hangover a generación de ruido de confort.

- El nodo de recepción estima o deduce parámetros para la generación de ruido de confort, en base al conjunto Y de
50 tramas. Los parámetros son asociados al ruido de fondo en el lado del nodo de transmisión. Haciéndolo así, el ruido de confort generado en base a dichos parámetros reflejará el ruido de fondo en el lado del nodo transmisor de buena forma, y de ese modo se obtendrá una experiencia de usuario buena/deseada. Seleccionar el conjunto Y en el lado del transmisor resulta ventajoso, puesto que en ese lado, la información de audio completa es accesible, en vez de
55 la versión reducida, cuantificada, que está disponible en el lado del nodo receptor.

Según se ha descrito con anterioridad, la información que indica el conjunto Y puede comprender una o más de un número que implica un número de tramas de hangover en secuencia; una palabra de código o mapa de bits que indica las posiciones de las tramas pertenecientes al conjunto Y, entre las N tramas de hangover; una palabra de

código o mapa de bits indicativo de cuáles de las N tramas de hangover están al menos comprendidas en el conjunto Y; y, una palabra de código o mapa de bits indicativa de cuáles de las N tramas de hangover no están comprendidas en el conjunto Y.

- 5 Además, la primera trama de SID puede comprender además parámetros de SID. El número N de tramas de hangover puede ser dinámicamente variable en base a propiedades de una señal de audio de entrada, según se ha descrito con anterioridad.

Ejemplo de nodo de transmisión, Figura 9

10 Realizaciones descritas en la presente memoria se refieren también a un nodo de transmisión, o nodo de codificación. El nodo de transmisión está asociado a las mismas características técnicas, objetos y ventajas que el método descrito con anterioridad e ilustrado, por ejemplo, en las Figuras 7a y 7b. El nodo de transmisión va a ser descrito de forma resumida, con el fin de evitar repeticiones innecesarias. El nodo de transmisión podría ser, por ejemplo, un dispositivo o UE, tal como un teléfono inteligente, una tableta, un ordenador o cualquier otro dispositivo capacitado para comunicación alámbrica y/o inalámbrica, y para codificación de habla.

15 En lo que sigue, se va a describir un ejemplo de nodo de transmisión 900, adaptado para habilitar la actuación de un método descrito en lo que antecede, adaptado para llevar a cabo una realización del método en un nodo de transmisión descrito en lo que antecede, con referencia a la Figura 9.

20 El nodo de transmisión es operable para codificar audio, tal como habla, y es operable para comunicar con otros nodos o entidades, por ejemplo en una red de comunicación. El nodo de transmisión es además operable para aplicar un esquema de DTX que comprende la transmisión de tramas de SID durante inactividad de habla. El nodo de transmisión puede ser operable para comunicar, por ejemplo, en un sistema de comunicación inalámbrica, tal como GSM, UMTS, E.UTRAN o CDMA 2000, y/o en un sistema de comunicación alámbrica.

25 La parte del nodo de transmisión que está principalmente relacionada con la solución sugerida en la presente memoria, ha sido ilustrada como disposición 901 rodeada por una línea discontinua/de puntos. La disposición y posiblemente otras partes del nodo de transmisión, están adaptadas para habilitar la actuación de uno o más de los métodos o procedimientos descritos anteriormente e ilustrados por ejemplo en las Figuras 7a y 7b.

30 El nodo de transmisión ilustrado en la Figura 9 comprende medios de procesamiento, que en este ejemplo tienen forma de un procesador 903 y una memoria 904, en donde dicha memoria contiene instrucciones 905 ejecutables por dicho procesador. Los medios de procesamiento están operativos para determinar, entre un número N de tramas de hangover, un conjunto Y de tramas que sean representativas de ruido de fondo. Los medios de procesamiento están además operativos para transmitir las N tramas de hangover, comprendiendo al menos dicho conjunto Y de tramas, hasta un nodo de recepción; y para transmitir una primera trama de SID hasta el nodo de recepción junto con la transmisión de las N tramas de hangover donde la trama de SID comprende información indicativa del conjunto Y determinado de tramas de hangover para el nodo de recepción.

35 El nodo de transmisión permite que un nodo de recepción genere ruido de confort en base al conjunto Y de tramas de hangover, permitiendo con ello la generación de ruido de confort de alta calidad.

La información indicativa del conjunto Y podría ser configurada de diferentes maneras, y la primera trama de SID podría comprender además parámetros de SID; y, el número N de tramas de hangover podría ser variable o fijo, según se ha descrito con anterioridad.

40 El nodo de transmisión 900 ha sido ilustrado de modo que comunica con otras entidades a través de una unidad de comunicación 902, la cual puede considerarse que comprende medios convencionales para comunicación inalámbrica y/o alámbrica conforme a un estándar de comunicación dentro del cual el nodo de transmisión es operable. 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, tal como por ejemplo procesamiento de señal junto con codificación de habla.

45 La disposición 901 puede, alternativamente, ser implementada y/o descrita esquemáticamente según se ha ilustrado en la Figura 10. La disposición 1001 comprende una unidad de determinación 1004, para determinar un conjunto Y de tramas, fuera de un número N de tramas de hangover, que es representativo de ruido de fondo. La disposición 1001 comprende además una unidad de transmisión para transmitir las N tramas de hangover que comprenden, al menos dicho conjunto Y de tramas, hasta un nodo de recepción; y además para transmitir una primera trama de SID hasta el nodo de recepción junto con la transmisión de las N tramas de hangover, donde la trama de SID comprende información indicativa del conjunto Y determinado de tramas de hangover para el nodo de recepción.

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

55 La disposición 1001, y otras partes del nodo de transmisión, podrían ser implementadas, por ejemplo, mediante uno o más de entre: un procesador o un microprocesador y software y almacenaje adecuados, un Dispositivo Lógico

Programable (PLD) u otro(s) componente(s) electrónico(s)/circuito(s) de procesamiento configurados para realizar las acciones mencionadas con anterioridad.

Ejemplo de nodo de recepción/descodificación, Figura 11

5 Realizaciones descritas en la presente memoria están relacionadas con un nodo de recepción, o nodo de descodificación. El nodo de recepción se asocia a las mismas características técnicas, objetos y ventajas que el método descrito con anterioridad e ilustrado, por ejemplo, en la Figura 8. El nodo de recepción va a ser descrito de forma resumida para evitar repeticiones innecesarias. El nodo de recepción podría ser, por ejemplo, un dispositivo o UE, tal como un teléfono inteligente, una tableta, un ordenador o cualquier otro dispositivo capacitado para comunicación alámbrica y/o inalámbrica y de codificación de audio.

10 En lo que sigue, se va a describir un ejemplo de nodo de recepción 1100, adaptado para permitir la actuación de un método descrito con anterioridad, adaptado para llevar a cabo al menos una realización del método en un nodo de recepción descrito con anterioridad, con referencia a la Figura 11.

15 El nodo de recepción es operable para descodificar audio, tal como habla, y es operable para comunicar con otros nodos o entidades, por ejemplo en una red de comunicación. El nodo de transmisión es además operable para aplicar un esquema de DTX que comprende la recepción de tramas de SID durante inactividad de habla. El nodo de recepción puede ser operable para comunicar en un sistema de comunicación inalámbrica, tal como GSM, UMTS, E-UTRAN o CDMA 2000, y/o en un sistema de comunicación alámbrica.

20 La parte del nodo de recepción que está relacionada principalmente con la solución sugerida en la presente memoria, ha sido ilustrada como disposición 1101 circundada por una línea discontinua/de puntos. La disposición y posiblemente otras partes del nodo de recepción, están adaptadas para permitir la actuación de uno o más de los métodos o procedimientos descritos con anterioridad e ilustrados, por ejemplo, en la Figura 8.

25 El nodo de recepción ilustrado en la Figura 11 comprende medios de procesamiento, en este ejemplo en forma de un procesador 1103 y de una memoria 1104, y en donde dicha memoria contiene instrucciones 1105 ejecutables por dicho procesador. Los medios de procesamiento son operativos para recibir N tramas de hangover desde un nodo de transmisión; y además, para recibir una primera trama de SID junto con las N tramas de hangover. Los medios de procesamiento están además operativos para determinar, en base a la información presente en la trama de SID recibida, un conjunto Y de tramas de hangover, entre el número N de tramas de hangover; y para generar ruido de confort al menos parcialmente en base al conjunto Y de tramas de hangover.

30 El nodo de recepción está por tanto habilitado para generar ruido de confort en base al conjunto Y de tramas de hangover, y con ello habilitado para generar ruido de confort de alta calidad.

La información que indica que el conjunto Y podría ser configurado de formas diferentes, y la primera trama de SID, podrían comprender además parámetros de SID; y el número N de tramas de hangover podría ser variable o fijo, según se ha descrito con anterioridad.

35 El nodo de recepción 1100 ha sido ilustrado como dispuesto para comunicar con otras entidades por medio de una unidad de comunicación 1102, la cual puede considerarse que comprende medios convencionales para comunicación inalámbrica y/o alámbrica conforme a un estándar de comunicación dentro del cual es operable el nodo de recepción. La disposición y/o el nodo de recepción pueden comprender además una o más unidades de almacenaje 1106. 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, tal como por ejemplo procesamiento de señal junto con descodificación de habla.

40 La disposición 1101 y otras partes del nodo de recepción o de descodificación podrían ser implementadas, por ejemplo, mediante uno o más de entre: un procesador o un microprocesador y software y almacenaje adecuados, un Dispositivo Lógico Programable (PLD) u otro(s) componente(s) electrónico(s)/circuito(s) de procesamiento configurados para realizar las acciones mencionadas en lo que antecede.

45 La disposición 1101 puede ser alternativamente implementada y/o descrita esquemáticamente según se ha ilustrado en la Figura 12. La disposición 1201 comprende una unidad de recepción 1203 para recibir N tramas de hangover desde un nodo de transmisión; y además, para recibir una primera trama de SID junto con las N tramas de hangover. La disposición comprende además una unidad de determinación 1204 para determinar, en base a la información contenida en la primera trama de SID recibida, un conjunto Y de tramas de hangover, entre el número N de tramas de hangover; y además, un generador de ruido 1205 para generar ruido de confort en base al conjunto Y de tramas de hangover.

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

55 La disposición 1201 y/o alguna otra parte del nodo de descodificación 1200, se supone que comprenden unidades

funcionales o circuitos adaptados para realizar descodificación de audio.

5 La disposición 1201 y otras partes del nodo de recepción o descodificación podrían ser implementadas, por ejemplo, mediante uno o más de entre: un procesador o un microprocesador y software y almacenaje adecuados, un Dispositivo Lógico Programable (PLD) u otro(s) componente(s) electrónico(s)/circuito(s) de procesamiento configurados para realizar las acciones mencionadas con anterioridad.

10 Debe entenderse que la elección de unidades o módulos de interacción, así como la designación de las unidades, son solamente a título de ejemplo, y que se pueden configurar nodos de cliente y servidor adecuados para ejecutar cualquiera de los métodos descritos en lo que antecede según una pluralidad de formas alternativas con el fin de estar en condiciones de ejecutar las acciones sugeridas del proceso. También debe apreciarse que las unidades o módulos que se han descrito en la presente descripción deben ser entendidas como entidades lógicas y no necesariamente como entidades físicas separadas.

Mediante el uso de la solución sugerida en la presente memoria, se puede incrementar la eficacia de las transmisiones de habla con DTX sin comprometer la calidad de la síntesis de ruido de confort al final de las ráfagas de conversación.

15 Aunque la descripción que antecede contiene una multiplicidad de características específicas, éstas no deben ser interpretadas como limitativas del alcance del concepto descrito en la presente memoria sino simplemente como suministradoras de ilustraciones de algunos ejemplos de realizaciones del concepto descrito. Se apreciará que el alcance del concepto descrito en la presente abarca completamente otras realizaciones que puedan resultar obvias para los expertos en la materia, y que el alcance del concepto descrito en la presente no está en consecuencia limitado. La referencia a un elemento en singular no está destinada a que se entienda “uno y solo uno” a menos que se indique así explícitamente, sino por el contrario “uno o más”. Todos los equivalentes estructurales y funcionales de los elementos de las realizaciones descritas en lo que antecede que sean conocidos por los expertos en la materia están expresamente incorporados en la presente memoria por referencia y se entiende que están abarcados por la misma. Además, no es necesario que un dispositivo o método dirija todos y cada uno de los problemas que deben ser resueltos por el concepto descrito en la presente, para que puedan ser abarcados por la misma.

Abreviaturas

- AMR Multi-tasa Adaptativa
- DTX Transmisión Discontinua
- ITU-T Unión de Telecomunicaciones Internacionales – Sector de estandarización de telecomunicaciones
- 30 LSF Frecuencia Espectral Lineal
- VAD Detector de Actividad de Voz
- 3GPP Proyecto Partnership de Tercera Generación
- SID Descriptor de Inserción de Silencio
- SNR Relación Señal Ruido
- 35 WB Banda Ancha

REIVINDICACIONES

- 1.- Método llevado a cabo por un nodo de transmisión (900, 1000), siendo el nodo operable para codificar habla y para aplicar un esquema de transmisión discontinua, DTX, que comprende transmisión de tramas de Descriptor de Inserción de Silencio, SID, durante inactividad de habla, comprendiendo el método:
- 5 - determinar (703a), de entre un número N de tramas de hangover de DTX, un conjunto Y de tramas que comprenden ruido que es representativo de ruido de fondo;
- transmitir (704a) las N tramas de hangover, comprendiendo al menos dicho conjunto Y de tramas, hasta un nodo de recepción;
- 10 - transmitir (705a) una primera trama de SID al nodo de recepción junto con la transmisión de las N tramas de hangover, donde la trama de SID comprende información indicativa del conjunto Y determinado de tramas de hangover para el nodo de recepción,
- permitiendo de ese modo que el nodo de recepción genere ruido de confort en base al conjunto de tramas de hangover, **caracterizado porque** el número de tramas en el conjunto Y es variable en el intervalo de 1 a N.
- 2.- Método según la reivindicación 1, en donde la información indicativa del conjunto Y comprende al menos uno de:
- 15 - un número, que implica un número de tramas de hangover en secuencia;
- una palabra de código o mapa de bits, que indica las posiciones de las tramas pertenecientes al conjunto Y, entre las N tramas de hangover;
- una palabra de código o mapa de bits, que indica cuáles de las N tramas de hangover están comprendidas en el conjunto Y;
- 20 - una palabra de código o mapa de bits, que indica cuales de las N tramas de hangover no están comprendidas en el conjunto Y.
- 3.- Método llevado a cabo por un nodo de recepción (1100, 1200) operable para descodificar habla y para 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 inactividad de habla, comprendiendo el método:
- 25 - recibir (801) N tramas de hangover desde un nodo de transmisión;
- recibir (802) una primera trama de SID junto con las N tramas de hangover;
- determinar (803), en base a la información contenida en la trama de SID recibida, un conjunto Y de tramas de hangover que comprende ruido que es representativo de ruido de fondo, de entre las N tramas de hangover;
- 30 - generar (804) ruido de confort en base al conjunto Y de tramas de hangover, **caracterizado porque** el número de tramas en el conjunto Y es variable en el intervalo de 1 a N.
- 4.- Método según la reivindicación 3, en donde la información indicativa del conjunto Y comprende al menos uno de:
- un número, que implica un número de tramas de hangover en secuencia;
- una palabra de código o mapa de bits, que indica las posiciones de las tramas pertenecientes al conjunto Y, entre las N tramas de hangover;
- 35 - una palabra de código o mapa de bits, que indica cuáles de las N tramas de hangover están al menos comprendidas en el conjunto Y;
- un palabra de código o mapa de bits, que indica cuáles de las N tramas de hangover no están comprendidas en el conjunto Y.
- 5.- Un nodo de transmisión (900, 1000), operable para codificar habla y para aplicar un esquema de transmisión discontinua, DTX, que comprende transmisión de tramas de Descriptor de Inserción de Silencio, SID, durante inactividad de habla, comprendiendo el nodo de transmisión medios de procesamiento operativos para:
- 40 - determinar, a partir de entre un número N de tramas de hangover, un conjunto Y de tramas que comprenden ruido que es representativo de ruido de fondo;
- 45 - transmitir las N tramas de hangover, comprendiendo al menos dicho conjunto Y de tramas, a un nodo de recepción, y
- transmitir una primera trama de SID al nodo de recepción junto con la transmisión de las N tramas de hangover,

donde la trama de SID comprende información que indica el conjunto Y determinado de tramas de hangover al nodo de recepción, **caracterizado porque** el número de tramas en el conjunto Y es variable en el intervalo de 1 a N.

6.- Nodo de transmisión según la reivindicación 5, en donde la información indicativa del conjunto Y comprende al menos uno de:

- 5 - un número, que implica un número de tramas de hangover en secuencia;
- una palabra de código o mapa de bits, que indica las posiciones de las tramas pertenecientes al conjunto Y, entre las N tramas de hangover;
- una palabra de código o mapa de bits, que indica cuáles de las N tramas de hangover están (al menos) comprendidas en el conjunto Y;
- 10 - una palabra de código o mapa de bits, que indica cuáles de las N tramas de hangover no están comprendidas en el conjunto Y.

7.- Nodo de transmisión según cualquiera de las reivindicaciones 5-6, en donde la primera trama de SID comprende además parámetros de SID.

- 15 8.- Nodo transmisión según cualquiera de las reivindicaciones 5-7, en donde el número N de tramas de hangover es dinámicamente variable en base a propiedades de una señal de audio de entrada.

9.- Un nodo de recepción (1100, 1200) operable para decodificar habla y para aplicar un esquema de transmisión discontinua, DTX, que comprende tramas de Descriptor de Inserción de Silencio, SID, y generación de ruido de confort durante inactividad de habla, comprendiendo el nodo de recepción medios de procesamiento operativos para:

- 20 - recibir N tramas de hangover desde un nodo de transmisión;
- recibir una primera trama de SID junto con las N tramas de hangover;
- determinar, en base a la información contenida en la trama de SID recibida, un conjunto Y de tramas de hangover que comprenden ruido que es representativo de ruido de fondo, de entre el número N de tramas de hangover, y
- 25 - generar ruido de confort en base al conjunto Y de tramas de hangover, **caracterizado porque** el número de tramas en el conjunto Y es variable en el intervalo de 1 a N.

10.- Nodo de recepción según la reivindicación 9, en donde los medios de procesamiento comprenden un procesador (1103) y una memoria (1104), y en donde dicha memoria contiene instrucciones (1105) ejecutables por dicho procesador.

- 30 11.- Nodo de recepción según la reivindicación 9 ó 10, en donde la información que indica el conjunto Y comprende al menos uno de:

- un número, que implica un número de tramas de hangover en secuencia;
- una palabra de código o mapa de bits, que indica las posiciones de las tramas pertenecientes al conjunto Y, entre las N tramas de hangover;
- 35 - una palabra de código o mapa de bits, que indica cuáles de las N tramas de hangover están (al menos) comprendidas en el conjunto Y;
- una palabra de código o mapa de bits, que indica cuáles de las N tramas de hangover no están comprendidas en el conjunto Y.

12.- Nodo de recepción según cualquiera de las reivindicaciones 9-11, en donde la primera trama de SID comprende además parámetros de SID.

- 40 13.- Nodo de recepción según cualquiera de las reivindicaciones 9-12, en donde el número N de tramas de hangover es dinámicamente variable en base a propiedades de una señal de audio de entrada.

14.- Programa de ordenador (905), que comprende un código de programa de ordenador, el cual, cuando se ejecuta en un nodo de transmisión, provoca que el nodo de transmisión lleve a cabo el método según cualquiera de las reivindicaciones 1-2.

- 45 15.- Producto de programa de ordenador que comprende un programa de ordenador (905) según la reivindicación 14.

16.- Programa de ordenador (1105), que comprende un código de programa de ordenador, el cual, cuando se

ejecuta en un nodo de recepción, provoca que el nodo de recepción lleve a cabo el método según cualquiera de las reivindicaciones 3-4.

17.- Producto de programa de ordenador que comprende un programa de ordenador (1105) según la reivindicación 16.

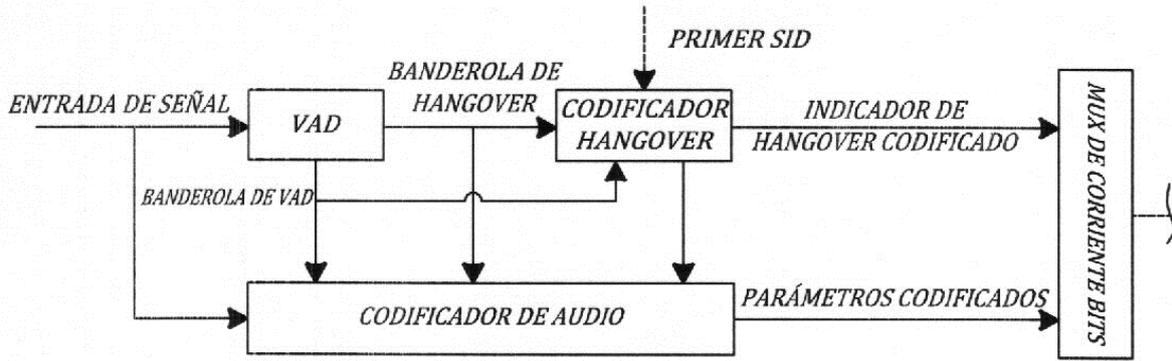


Figura 1

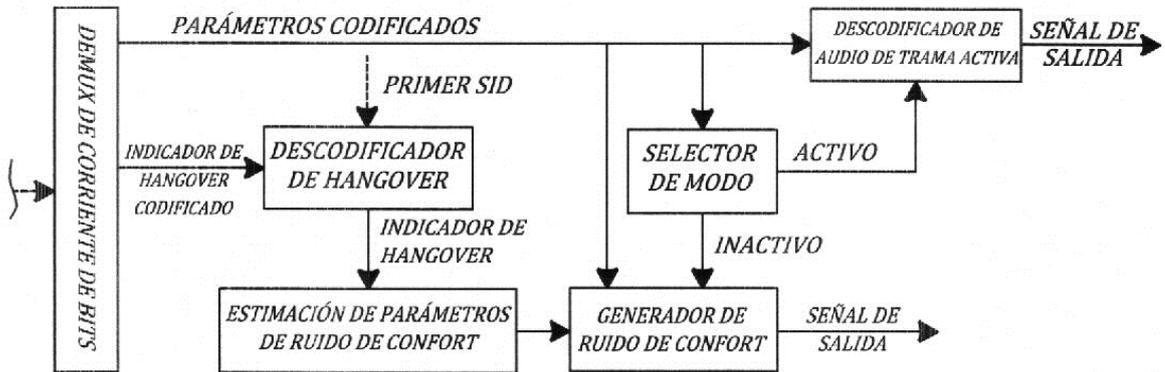


Figura 2

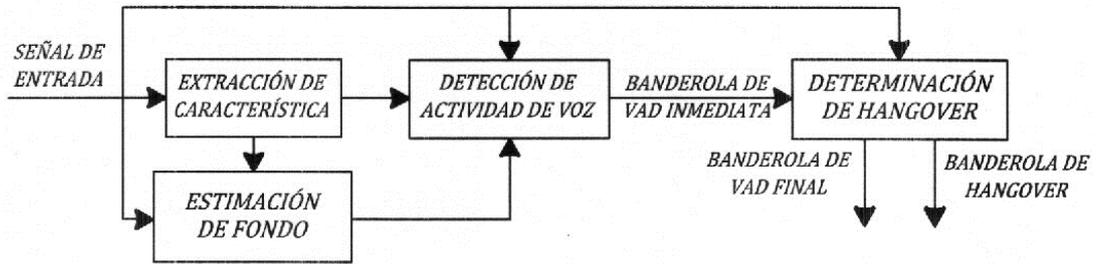


Figura 3

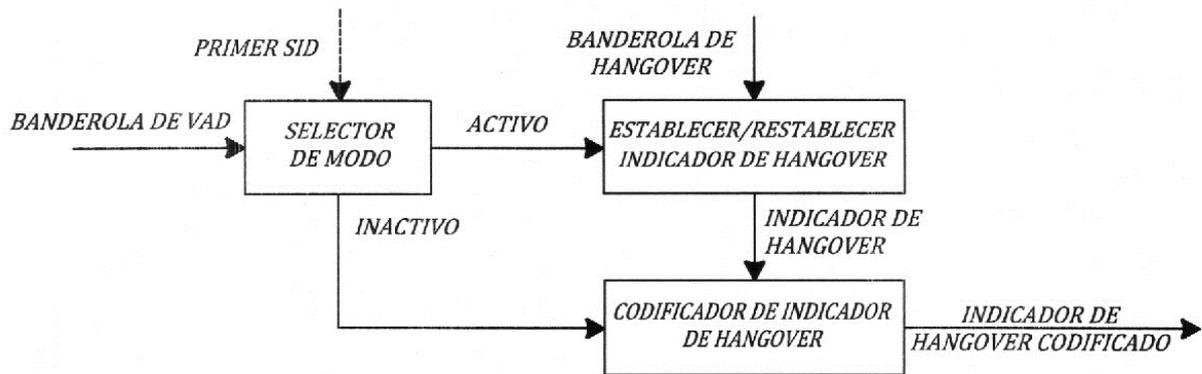


Figura 4

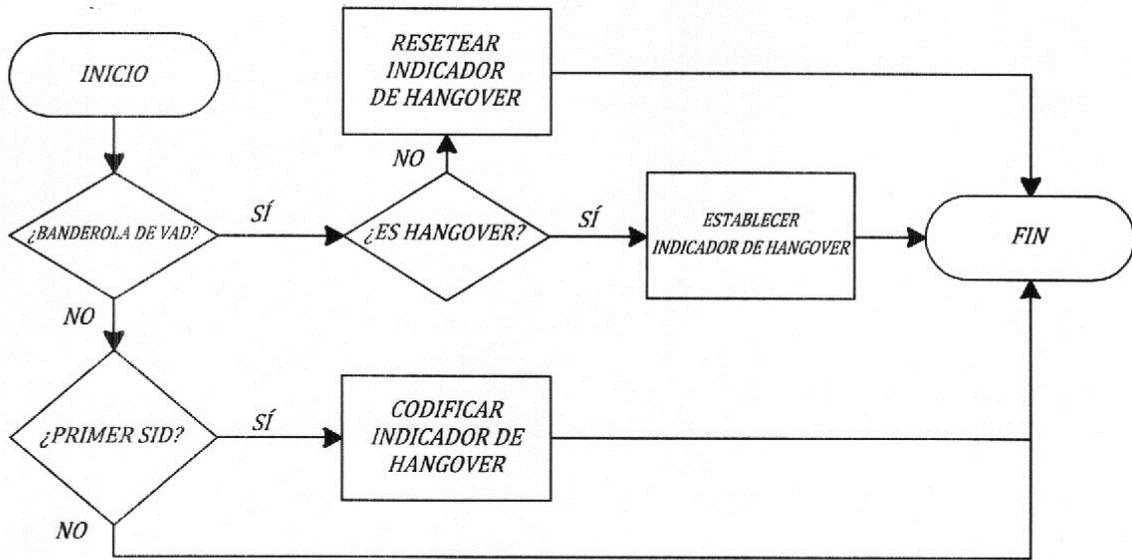


Figura 5

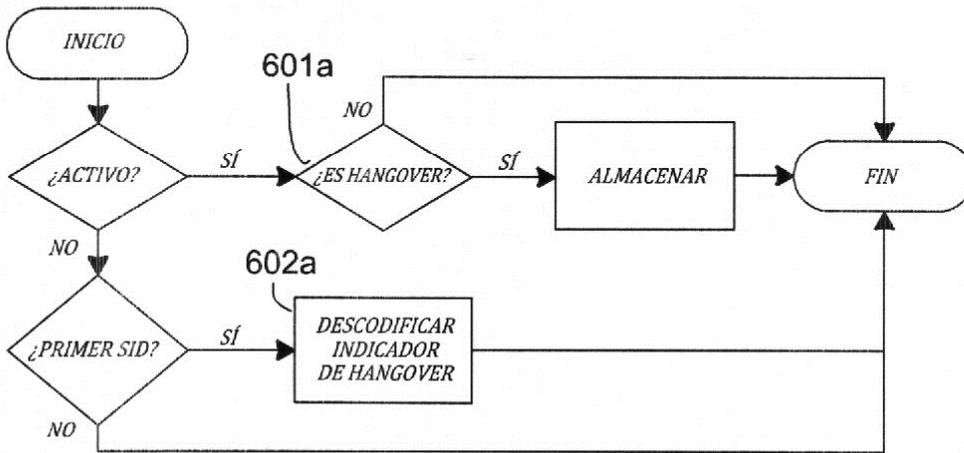


Figura 6a

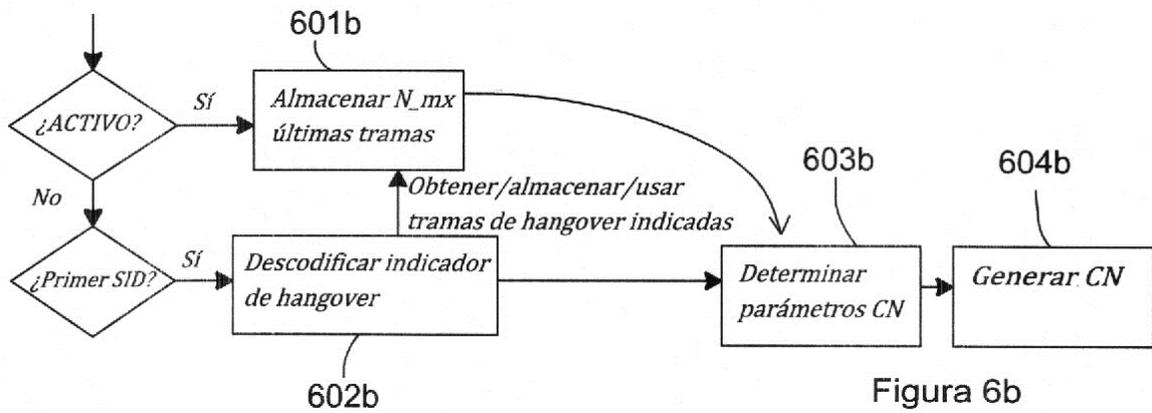


Figura 6b

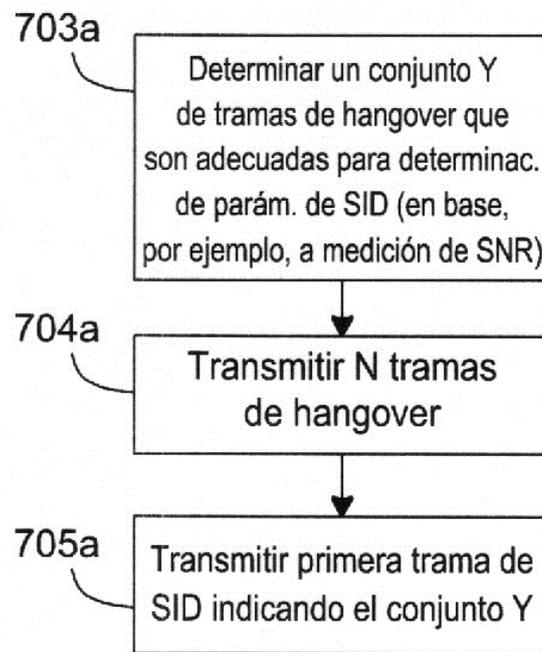


Figura 7a

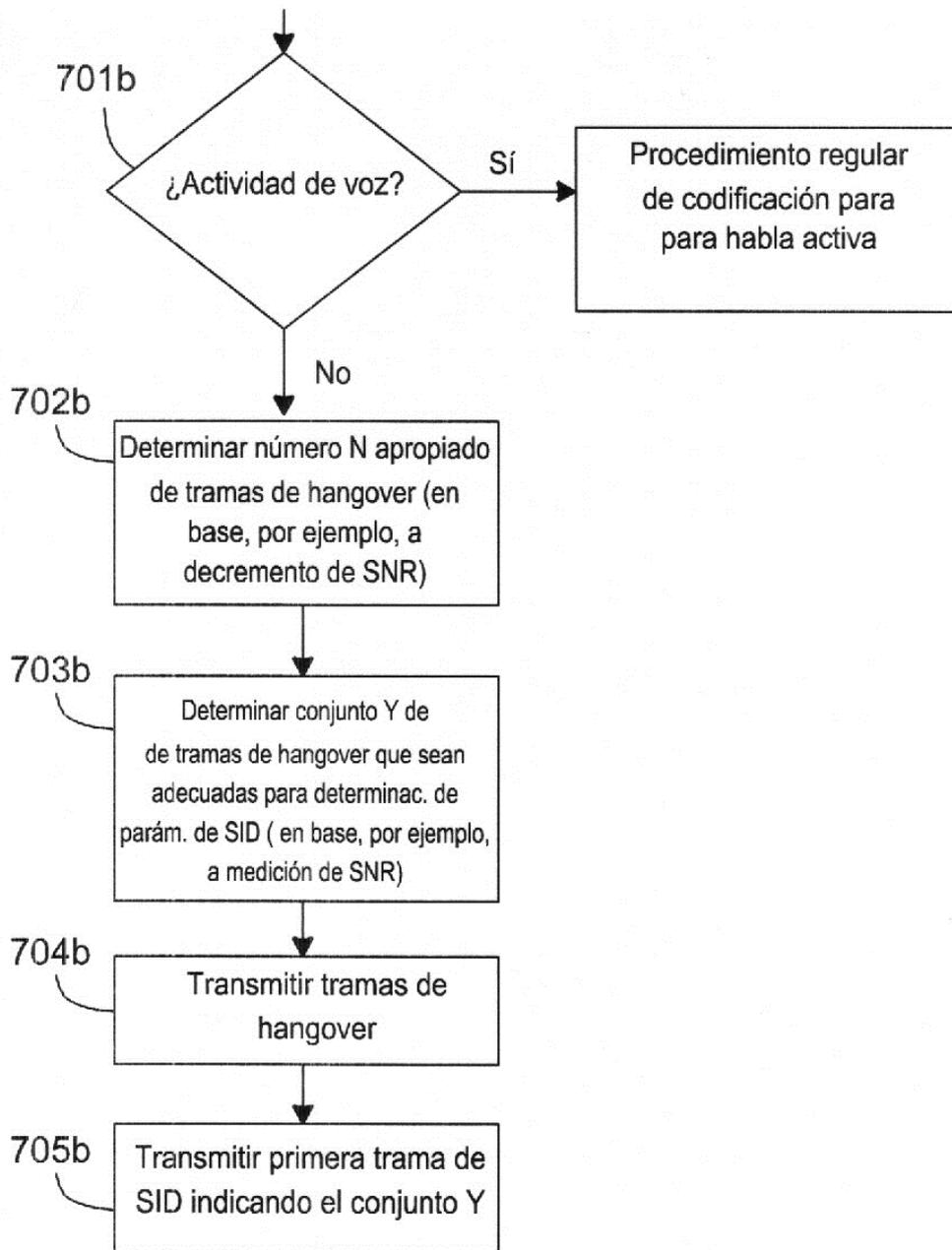


Figura 7b

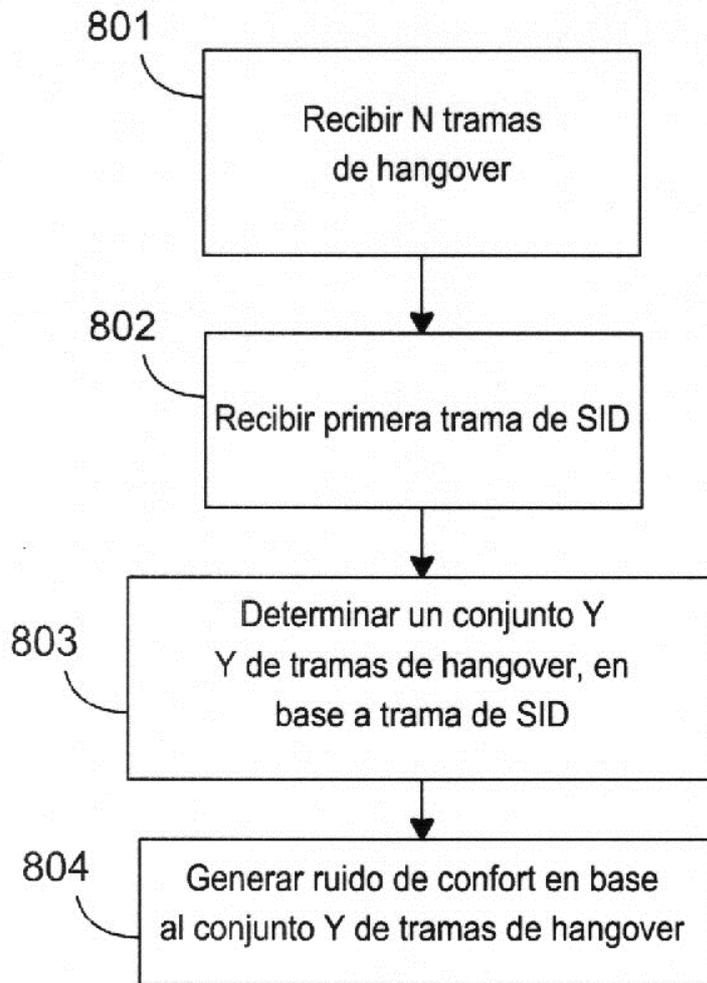


Figura 8

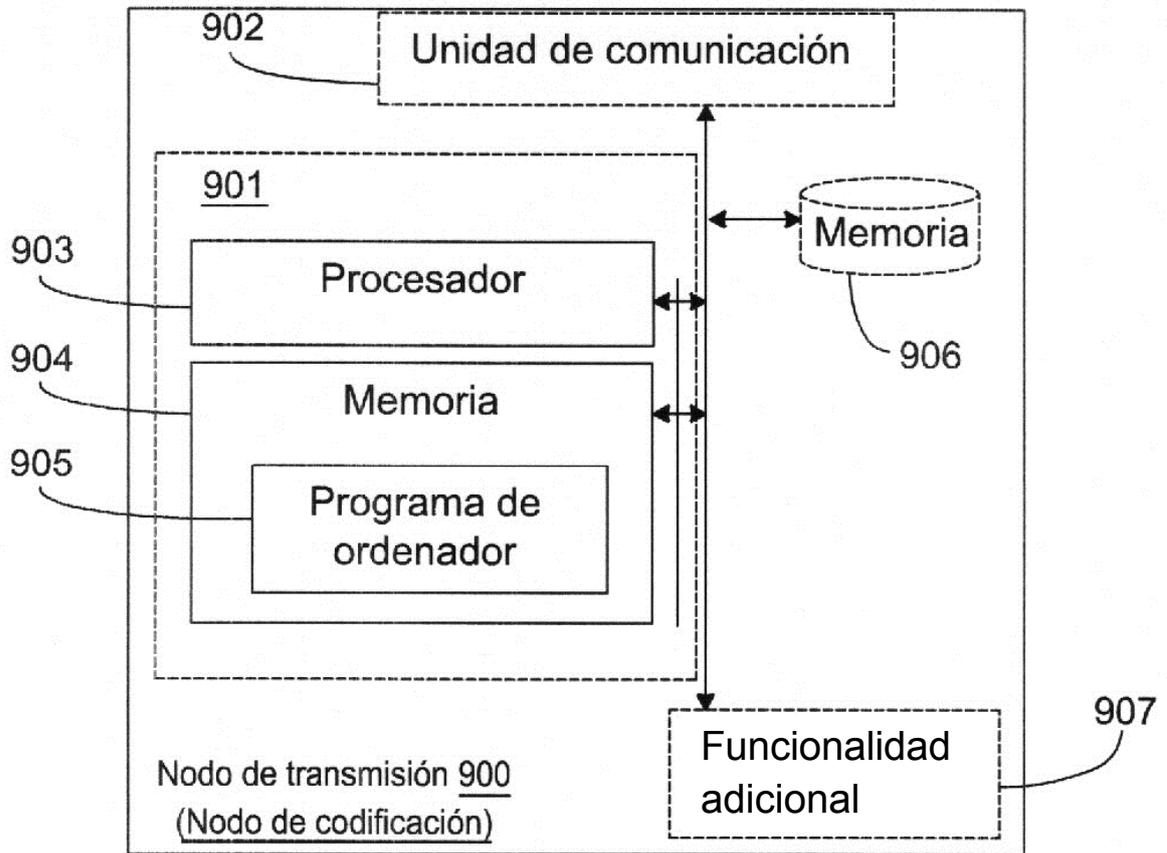


Figura 9

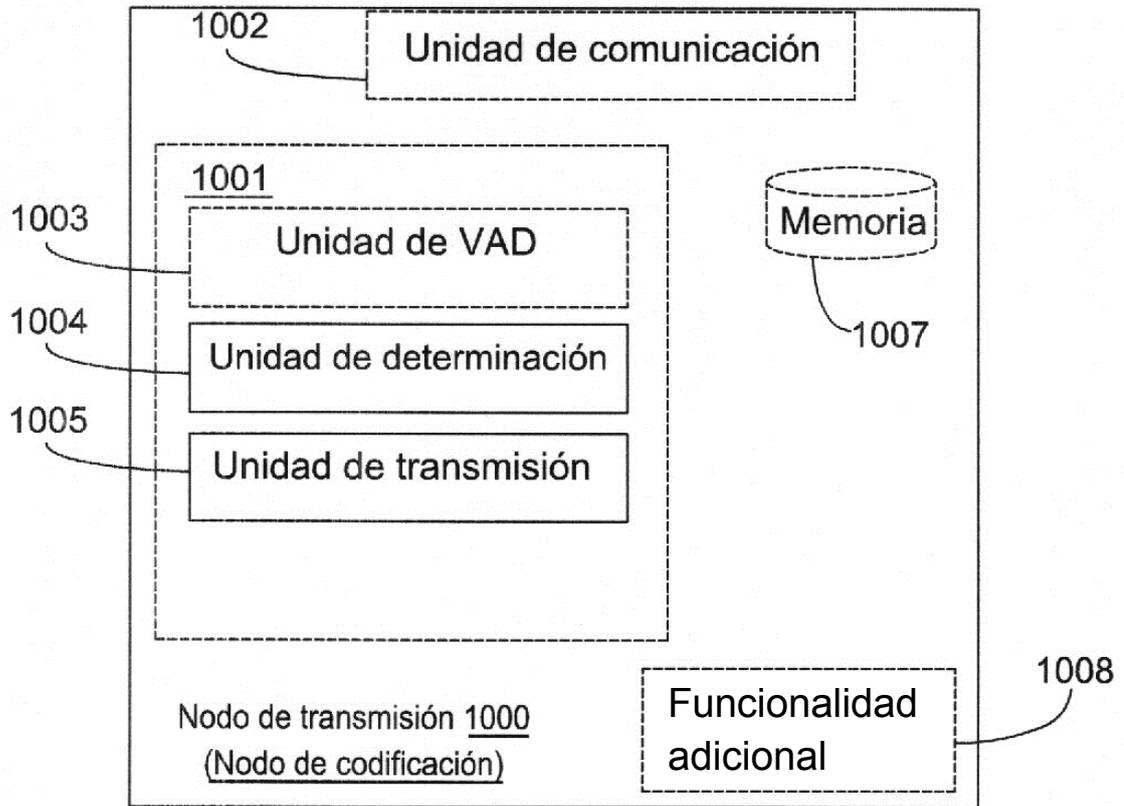


Figura 10

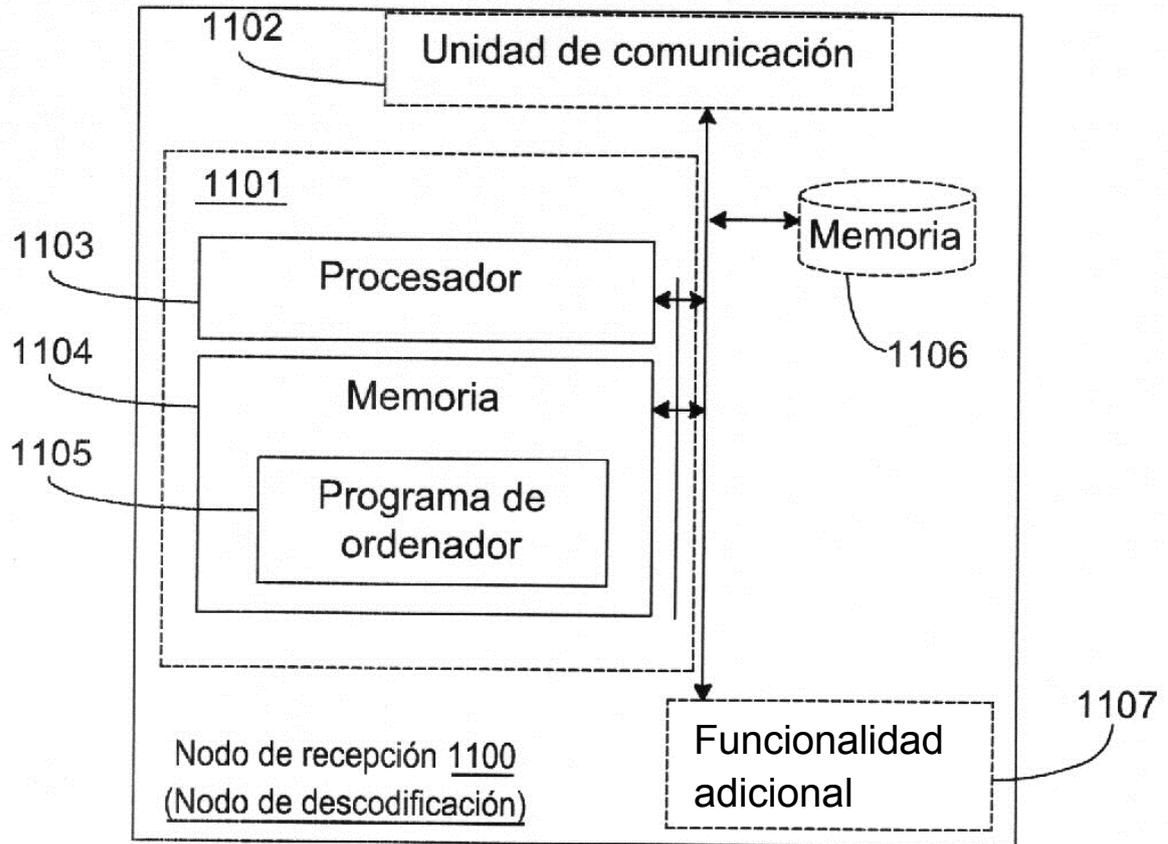


Figura 11

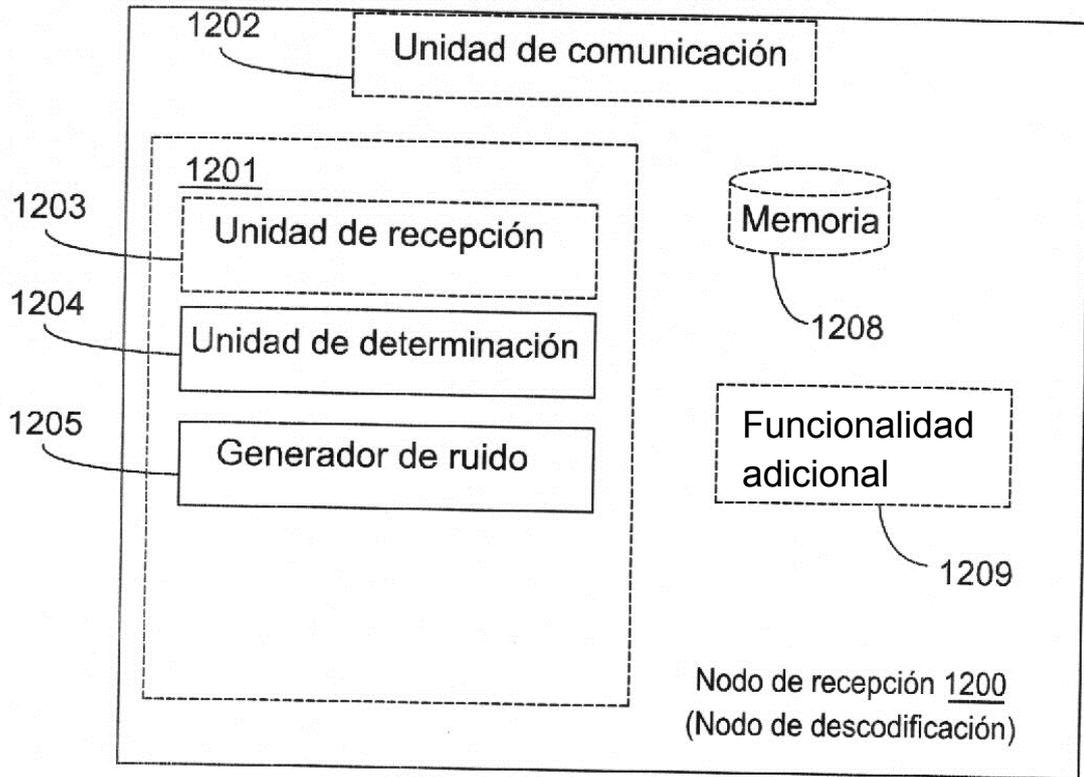


Figura 12