

19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 688 021**

51 Int. Cl.:

G10L 19/012 (2013.01)

G10L 19/00 (2013.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

86 Fecha de presentación y número de la solicitud internacional: **19.12.2013 PCT/EP2013/077527**

87 Fecha y número de publicación internacional: **26.06.2014 WO14096280**

96 Fecha de presentación y número de la solicitud europea: **19.12.2013 E 13814127 (0)**

97 Fecha y número de publicación de la concesión europea: **18.07.2018 EP 2936486**

54 Título: **Adición de ruido de confort para modelar ruido de fondo a bajas tasas de bits**

30 Prioridad:

21.12.2012 US 201261740883 P

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

30.10.2018

73 Titular/es:

**FRAUNHOFER-GESELLSCHAFT ZUR
FÖRDERUNG DER ANGEWANDTEN
FORSCHUNG E.V. (100.0%)
Hansastraße 27c
80686 München, DE**

72 Inventor/es:

**FUCHS, GUILLAUME;
LOMBARD, ANTHONY;
RAVELLI, EMMANUEL;
DÖHLA, STEFAN;
LECOMTE, JÉRÉMIE y
DIETZ, MARTIN**

74 Agente/Representante:

ARIZTI ACHA, Monica

ES 2 688 021 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).

Adición de ruido de confort para modelar ruido de fondo a bajas tasas de bits

DESCRIPCIÓN

5 La presente invención se refiere a procesamiento de señales de audio y, en particular, a la codificación de voz ruidosa y la adición de ruido control a las señales de audio.

Habitualmente se utilizan generadores de ruido de confort en la transmisión discontinua (DTX) de señales de audio, en particular, de señales de audio con contenido de voz. En ese modo, la señal de audio es clasificada, en primer lugar, en tramas activas e inactivas por un detector de actividad de voz (VAD). Se puede encontrar un ejemplo de VAD en [1]. Basándose en el resultado del VAD, solo se codifican y transmiten las tramas de voz activa a la tasa de bits nominal. Durante pausas prolongadas, en que solo está presente el ruido de fondo, la tasa de bits se reduce o se ajusta a cero y el ruido de fondo es codificado de manera episódica y paramétrica. Es así como se reduce significativamente la tasa de bits promedio. El ruido se genera durante las tramas inactivas del lado del decodificador por medio de un generador de ruido de confort (CNG). Por ejemplo, los codificadores de voz AMR-WB [2] e ITU G.718 [1] tienen la posibilidad de funcionar en ambos casos en el modo DTX. Otro decodificador de voz de ese tipo se conoce a partir del documento [3].

20 A partir del documento [4] se conoce un decodificador, que comprende una herramienta de relleno para insertar líneas espectrales en posiciones de una trama decodificada, que se cuantifican a cero en el sitio del codificador.

La codificación de voz y especialmente la voz ruidosa a bajas tasas de bits es propensa a alteraciones. Los codificadores de voz se basan por lo general en un modelo de producción de voz que ya no se sostiene en presencia de ruido de fondo. En ese caso, la codificación pierde eficiencia y se reduce la calidad de la señal de audio decodificada. Más aun, ciertas características de la codificación de voz pueden ser especialmente molestas al manejar la voz ruidosa. En efecto, a bajas tasas de bits, la cuantificación gruesa de los parámetros de codificación produce cierta fluctuación en el tiempo, y las fluctuaciones son perceptualmente fastidiosas con la codificación de voz sobre ruido de fondo fijo.

30 La reducción de ruido es una técnica muy conocida para intensificar la inteligibilidad de la voz y mejorar la comunicación en presencia de ruido de fondo. También se la ha adoptado en la codificación de voz. Por ejemplo, el codificador G.718 utiliza la reducción de ruido para deducir ciertos parámetros de codificación como el tono de la voz. Tiene también la posibilidad de codificar la señal intensificada en lugar de la señal original. Luego la voz es más predominante en comparación con el nivel de ruido en la señal decodificada. Sin embargo, habitualmente suena más degradada o menos natural, ya que la reducción de ruido podría distorsionar los componentes de voz y causar alteraciones de ruido musical audible además de las alteraciones de codificación. El objetivo de la presente invención es proporcionar conceptos mejorados para el procesamiento de señales de audio. El objetivo de la presente invención se obtiene por medio de un decodificador de acuerdo con la reivindicación 1, un codificador de acuerdo con la reivindicación 21, un sistema de acuerdo con la reivindicación 22, un método de acuerdo con la reivindicación 23 y 24, un flujo de bits de acuerdo con la reivindicación 25 y un programa de ordenador de acuerdo con la reivindicación 26. En un aspecto, la invención proporciona un decodificador que está configurado para procesar un flujo de bits de audio codificado, en el que el decodificador comprende:

45 un decodificador de flujos de bits configurado para derivar una señal de audio decodificada del flujo de bits, en el que la señal de audio decodificada comprende por lo menos una trama decodificada;
 un dispositivo de estimación de ruido configurado para producir una señal de estimación de ruido que contiene una estimación del nivel y/o la forma espectral o de un ruido en la señal de audio decodificada;
 un dispositivo generador de ruido de confort configurado para derivar una señal de ruido de confort de la señal de estimación de ruido; y
 50 un combinador configurado para combinar la trama decodificada de la señal de audio decodificada y la señal de ruido de confort para obtener una señal de audio de salida.

El decodificador de flujos de bits puede ser un dispositivo o un programa de ordenador con capacidad para decodificar un flujo de bits de audio, que es un flujo de datos digitales que contiene información de audio. El proceso de decodificación da origen a una señal digital de audio decodificada, que puede ser alimentada a un conversor A/D para producir una señal de audio analógica, que luego puede ser alimentada a un altavoz para producir una señal audible.

60 La señal de audio decodificada se divide en las denominadas tramas, en el que cada una de estas tramas contiene información de audio referente a un determinado intervalo de tiempo. Dichas tramas se pueden clasificar en tramas activas y tramas inactivas, donde una trama activa es una trama que contiene componentes deseados de la información de audio, como por ejemplo voz o música, en tanto que una trama inactiva es una trama que no contiene ningún componente deseado de la información de audio. Las tramas inactivas aparecen por lo general durante las pausas, donde no hay presencia de ningún componente deseado, tal como música o voz. Por lo tanto, las tramas inactivas habitualmente contienen únicamente ruido de fondo.

En la transmisión discontinua (DTX) de señales de audio solo se obtienen las tramas activas de la señal de audio decodificada decodificando el flujo de bits, ya que, durante las tramas inactivas, el codificador no transmite la señal de audio dentro del flujo de bits.

5 En la transmisión no discontinua (no-DTX) de señales de audio se obtienen las tramas activas y también las tramas inactivas decodificando el flujo de bits.

10 Las tramas que se obtienen mediante la decodificación del flujo de bits por el decodificador de flujos de bits se denominan tramas decodificadas.

15 El dispositivo de estimación de ruido está configurado para producir una señal de estimación de ruido que contiene una estimación del nivel y/o la forma espectral de un ruido incluido en la señal de audio decodificada. Más aun, el dispositivo generador de ruido de confort está configurado para derivar una señal de ruido de confort de la señal de estimación de ruido. La señal de estimación de ruido puede ser una señal que contiene información con respecto a las características del ruido contenido en la señal de audio decodificada en forma paramétrica. La señal de ruido de confort es una señal de audio artificial, que corresponde al ruido contenido en la señal de audio decodificada. Estas características permiten que el ruido de confort suene como el ruido de fondo real sin requerir información complementaria alguna con respecto al ruido de fondo contenido en el flujo de bits.

20 El combinador está configurado para combinar la trama decodificada de la señal de audio decodificada y la señal de ruido de confort para obtener una señal de audio de salida. Como resultado de eso la señal de audio de salida comprende tramas decodificadas, que comprenden ruido artificial. El ruido artificial contenido en las tramas decodificadas permite enmascarar las alteraciones de la señal de audio de salida especialmente cuando se transmite el flujo de bits a bajas tasas de bits. Alisa las fluctuaciones observadas habitualmente y al mismo tiempo enmascara las alteraciones de codificación predominantes.

25 A diferencia de la técnica anterior, la presente invención aplica el principio de adición del ruido de confort artificial a las tramas decodificadas. El concepto de la invención se puede aplicar tanto a los modos DTX como no DTX.

30 La invención proporciona un método para intensificar la calidad de la voz ruidosa codificada y transmitida a bajas tasas de bits. A bajas tasas de bits, la codificación de la voz ruidosa, es decir voz grabada con ruido de fondo, habitualmente no es tan eficiente como la codificación de la voz limpia. La síntesis decodificada generalmente es propensa a alteraciones. Los dos tipos de orígenes, el ruido y la voz, no pueden ser codificados de manera eficiente por un esquema de codificación que se basa en un modelo de origen único. La presente invención proporciona un concepto para modelar y sintetizar el ruido de fondo del lado del decodificador y requiere muy poca o ninguna información complementaria. Esto se obtiene estimando el nivel y la forma espectral del ruido de fondo del lado del decodificador y generando artificialmente un ruido de confort. El ruido generado se combina con la señal de audio decodificada y permite enmascarar las alteraciones de codificación.

35 Además, el concepto se puede combinar con un esquema de reducción de ruido aplicado del lado del codificador. La reducción de ruido intensifica el nivel de la relación de señal a ruido (SNR) y mejora la eficiencia de la codificación de audio subsiguiente. Luego se compensa la cantidad de ruido faltante en la señal de audio decodificada mediante el ruido de confort del lado del decodificador. Sin embargo, habitualmente suena más degradado o menos natural, ya que la reducción de ruido podría distorsionar los componentes de audio y causar distorsiones de ruido musical además de las alteraciones por codificación. Un aspecto de la presente invención consiste en enmascarar dichas distorsiones desagradables mediante la adición de un ruido de confort del lado del decodificador. Cuando se usa un esquema de reducción de ruido, la adición de ruido de confort no deteriora la SNR. Más aun, el ruido de confort oculta una gran parte del molesto ruido musical típico de las técnicas de reducción de ruido.

40 En una realización preferida de la invención, la trama decodificada es una trama activa. Esta característica extiende el principio de adición de ruido de confort a las tramas activas decodificadas.

45 En una realización preferida de la invención, la trama decodificada es una trama activa. Esta característica extiende el principio de adición de ruido de confort a las tramas inactivas decodificadas.

50 En una realización preferida de la invención, el dispositivo de estimación de ruido comprende un dispositivo de análisis espectral configurado para crear una señal de análisis que contiene el nivel y la forma espectral del ruido presente en la señal de audio decodificada y un dispositivo para producir estimaciones de ruido configurado para producir la señal de estimación de ruido sobre la base de la señal de análisis.

60 En una realización preferida de la invención, el dispositivo generador de ruido de confort comprende un generador de ruido configurado para crear una señal de ruido de confort en el dominio de la frecuencia sobre la base de la señal de estimación de ruido y un sintetizador espectral configurado para crear la señal de ruido de confort sobre la

base de la señal de ruido de confort en el dominio de la frecuencia.

5 En una realización preferida de la invención, el decodificador comprende un dispositivo conmutador configurado para conmutar el decodificador de manera alternada a un primer modo de operación o a un segundo modo de operación, en el que, en el primer modo de operación, la señal de ruido de confort es alimentada al combinador, en tanto que la señal de ruido de confort no es alimentada al combinador en el segundo modo de operación. Estas características permiten abandonar el uso del ruido de confort artificial en situaciones en que no es necesario.

10 En una realización preferida de la invención, el decodificador comprende un dispositivo de control configurado para controlar el dispositivo conmutador en forma automática, donde el dispositivo de control comprende un detector de ruido configurado para controlar el dispositivo conmutador dependiendo de una relación de señal a ruido de la señal de audio decodificada, en el que en condiciones de baja relación de señal a ruido el decodificador se conmuta al primer modo de operación y en condiciones de alta relación de señal a ruido al segundo modo de operación. En virtud de estas características se puede activar el ruido de confort solo en situaciones de voz ruidosa, es decir, no en situaciones de voz limpia ni música limpia. Para discriminar entre condiciones de baja relación de señal a ruido y condiciones de alta relación de señal a ruido se puede definir y utilizar un umbral para la relación de señal a ruido.

20 En una realización preferida de la invención, el dispositivo de control comprende un receptor de información complementaria configurado para recibir información complementaria contenida en el flujo de bits, que corresponde a la relación de señal a ruido de la señal de audio decodificada, y configurado para crear una señal de detección de ruido, donde el detector de ruido controla el dispositivo conmutador dependiendo de la señal de detección de ruido. Estas características permiten controlar el dispositivo conmutador basándose en un análisis de la señal realizado por un dispositivo externo que produce y/o procesa el flujo de bits recibido. El dispositivo externo puede ser especialmente un codificador producir el flujo de bits.

25 En una realización preferida de la invención, la información complementaria que corresponde a la relación de señal a ruido de la señal de audio decodificada consiste en por lo menos un bit dedicado del flujo de bits. En general, un bit dedicado es un bit que contiene, solo o junto con otros bits dedicados, información definida. En este contexto, el bit dedicado puede indicar si la relación de señal a ruido es superior o inferior a un umbral predefinido.

30 En una realización preferida de la invención, el dispositivo de control comprende un estimador de energía de la señal deseada configurado para determinar la energía de una señal deseada de la señal de audio decodificada, un estimador de energía del ruido configurado para determinar la energía de un ruido de la señal de audio decodificada y un estimador de relación de señal a ruido configurado para determinar la relación de señal a ruido de la señal de audio decodificada sobre la base de la energía de la señal deseada y sobre la base de la energía del ruido, donde el dispositivo conmutador se conmuta dependiendo de la relación de señal a ruido determinada por el dispositivo de control. En este caso no es necesaria ninguna información complementaria en el flujo de bits. Como la energía de la señal deseada por lo general es superior a la energía del ruido de la señal decodificada, la energía total de la señal de audio decodificada, que incluye la energía de la señal deseada como así también la energía del ruido brinda una estimación aproximada de la energía de señal deseada de la señal de audio decodificada. Por esta razón, se puede calcular la relación de señal a ruido por aproximación dividiendo la energía total de la señal de audio decodificada por la energía del ruido de la señal decodificada.

45 En una realización preferida de la invención, el flujo de bits contiene tramas activas y tramas inactivas, donde el dispositivo de control está configurado para determinar la energía de señal deseada de la señal de audio decodificada durante las tramas activas y para determinar la energía del ruido de la señal de audio decodificada durante las tramas inactivas. De esta manera, se puede obtener una gran precisión al estimar la relación de señal a ruido de manera sencilla.

50 En una realización preferida de la invención, el flujo de bits contiene tramas activas y tramas inactivas, en el que el decodificador comprende un receptor de información complementaria configurado para discriminar entre las tramas activas y las tramas inactivas sobre la base de información complementaria contenida en el flujo de bits, que indica si la trama actual es activa o inactiva. Gracias a esta característica se pueden identificar las tramas activas o las tramas inactivas respectivamente, sin esfuerzo de cálculo.

55 En una realización preferida de la invención, la información complementaria que indica si la trama actual es activa o inactiva consiste en por lo menos un bit dedicado en el flujo de bits.

60 En una realización preferida de la invención, el dispositivo de control está configurado para determinar la energía de señal deseada de la señal de audio decodificada sobre la base de la señal de análisis. En este caso la señal de análisis, que habitualmente debe ser computada para realizar la estimación de ruido, puede ser reutilizada, por lo cual se puede reducir la complejidad.

En una realización preferida de la invención el dispositivo de control está configurado para determinar la energía del

ruido de la señal de audio decodificada sobre la base de la señal de estimación de ruido. En esa realización, la señal de estimación de ruido, que por lo general tiene que ser computada para generar ruido de confort, puede ser reutilizada, por lo cual se puede reducir aún más la complejidad.

5 En una realización preferida de la invención, el dispositivo generador de ruido de confort está configurado para crear la señal de ruido de confort sobre la base de una señal de nivel objetivo de ruido de confort. El nivel de ruido de confort agregado debe ser limitado para preservar la inteligibilidad y la calidad. Esto se puede lograr escalando el ruido de confort empleando una señal de ruido objetivo que indica un nivel de ruido objetivo predeterminado.

10 En una realización preferida de la invención la señal de nivel objetivo de ruido de confort es ajustada dependiendo de una tasa de bits del flujo de bits. Por lo general, la señal de audio decodificada exhibe una relación de señal a ruido más elevada que la señal de entrada original, especialmente a bajas tasas de bits donde las alteraciones por codificación son más severas. Esta atenuación del nivel de ruido en la codificación de voz proviene de un paradigma de modelado de origen que espera tener la voz como entrada. De lo contrario, la codificación por modelado de origen no es totalmente apropiada y no puede reproducir toda la energía de los componentes que no son de voz. Por ende, se puede ajustar la señal de nivel objetivo de ruido de confort dependiendo de la tasa de bits para compensar de manera aproximada la atenuación de ruido introducida de manera inherente por el proceso de codificación.

20 En una realización preferida de la invención, se ajusta la señal de nivel objetivo de ruido de confort dependiendo de un nivel de atenuación de ruido causada por un método de reducción de ruido aplicado al flujo de bits. Gracias a estas características, se puede compensar la atenuación del ruido causada por una reducción del módulo de ruido en un codificador.

25 En una realización preferida de la invención, la energía de la señal de ruido de confort en el dominio de la frecuencia del ruido aleatorio $w(k)$ se ajusta dependiendo de la señal de nivel objetivo de ruido de confort, que indica un nivel objetivo de ruido de confort g_{tar} , por cada frecuencia k de la siguiente manera: $E_w(k) = \max\{g_{tar} - \hat{E}_n(k); 0\}$, en el que $\hat{E}_n(k)$ se refiere a una estimación de la energía del ruido de la señal de audio decodificada a la frecuencia k , proporcionada por el dispositivo para producir estimaciones de ruido. Mediante estas características se puede intensificar la inteligibilidad y calidad de la señal de salida.

30 En una realización preferida de la invención, el decodificador comprende un decodificador de flujos de bits adicional, donde el decodificador de flujos de bits y el decodificador de flujos de bits adicional son de diferentes tipos, en el que el decodificador comprende un conmutador configurado para alimentar la señal decodificada procedente del decodificador de flujos de bits o la señal decodificada procedente del decodificador de flujos de bits adicional al dispositivo de estimación de ruido y al combinador. Como la adición de ruido de confort se efectúa cuando se utiliza el decodificador de flujos de bits, como así también cuando se utiliza el decodificador de flujos de bits adicional, se pueden reducir al mínimo las alteraciones de transición al conmutar entre el decodificador de flujos de bits y el decodificador de flujos de bits adicional. Por ejemplo, el decodificador de flujos de bits puede ser un decodificador de flujos de bits por predicción lineal excitada por códigos algebraicos (ACELP), en tanto que el decodificador de flujos de bits adicional puede ser un decodificador de flujos de bits de núcleo basado en transformadas (TCX).

La invención proporciona además un codificador para el procesamiento de señales de audio que está configurado para producir un flujo de bits de audio, en el que el codificador comprende:

45 un codificador de flujos de bits configurado para producir una señal de audio codificada que corresponde a una señal de audio de entrada y para derivar el flujo de bits de la señal de audio codificada;

50 un analizador de señales que consta de un estimador de relación de señal a ruido configurado para determinar la relación de señal a ruido de la señal de audio de entrada sobre la base de la energía de señal deseada de la señal de audio determinada por un estimador de energía de la señal deseada y sobre la base de la energía de un ruido de la señal de audio de entrada determinada por un estimador de energía de ruido;

un dispositivo de reducción de ruido configurado para producir una señal de audio con reducción de ruido; y

55 un dispositivo conmutador configurado para alimentar, dependiendo de la relación de señal a ruido determinada de la señal de audio de entrada, ya sea la señal de audio de entrada o la señal de audio con reducción de ruido al codificador de flujos de bits con el propósito de codificar la señal respectiva, en el que el codificador de flujos de bits está configurado para transmitir una información complementaria, que indica si se ha codificado la señal de audio de entrada o la señal de audio con reducción de ruido dentro del flujo de bits.

60 El codificador de flujos de bits puede ser un dispositivo o un programa de ordenador con capacidad para codificar una señal de audio, que es una señal de datos digitales que contiene información de audio. El proceso de codificación resulta en un flujo de bits digital, que se puede transmitir por un vínculo de datos digitales a un decodificador situado en un punto remoto.

La señal de audio de entrada es directamente codificada por el codificador de flujos de bits. El codificador de flujos de bits puede ser un codificador de voz o una conmutación de esquemas de bajo retardo entre un codificador de voz ACELP y un codificador de audio basado en transformadas TCX. El codificador de flujos de bits es responsable de la codificación de la señal de audio de entrada y la generación del flujo de bits necesario para decodificar la señal de audio. En paralelo, la señal de entrada es analizada por un módulo denominado analizador de señales. En una realización preferida el análisis de señales es igual al utilizado en G.718. Consiste en un dispositivo de análisis espectral seguido por el dispositivo para producir estimaciones de ruido. Los espectros de la señal original y del ruido estimado son enviados como entrada al módulo de reducción de ruido. La reducción de ruido atenúa el nivel de ruido de fondo en el dominio de la frecuencia. La cantidad de reducción está dada por el nivel pretendido de atenuación. La señal intensificada en el dominio del tiempo (señal de audio con reducción de ruido) se genera después de la síntesis espectral. La señal se utiliza para deducir algunas características, como por ejemplo la estabilidad del tono, que luego es aprovechada por el VAD para discriminar entre tramas activas e inactivas. El resultado de la clasificación puede ser utilizado además por el módulo codificador. En la realización preferida, se utiliza un modo de codificación específico para manejar tramas inactivas. De esta manera, el decodificador puede decir la bandera VAD del flujo de bits sin necesidad de un bit dedicado.

Para evitar distorsiones innecesarias en situaciones sin ruido (voz limpia o música limpia), la reducción de ruido se aplica sólo en caso de la voz ruidosa y de lo contrario se la omite. La discriminación entre señales con y sin ruido se obtiene estimando la energía a largo plazo tanto del ruido como de la señal deseada (voz o música). La energía a largo plazo se computa mediante un filtrado autorregresivo de primer orden de la energía de entrada de trama (durante las tramas activas) o usando la salida del módulo de estimación de ruido (durante las tramas inactivas). De esta manera se puede calcular una estimación de la relación de señal a ruido, que se define como la relación de la energía a largo plazo de la voz o música durante la energía a largo plazo del ruido. Si la relación de señal a ruido es inferior a un umbral predeterminado, la trama se considera ruidosa, de lo contrario se la clasifica como voz limpia. Como el codificador de flujos de bits está configurado para transmitir información complementaria dentro del flujo de bits, que indica si se está codificando la señal de audio de entrada o la señal de audio con reducción de ruido, el decodificador puede ajustar la señal de nivel objetivo de ruido de confort automáticamente al modo de operación del codificador.

En la realización preferida de la invención, durante las tramas activas, solo se actualiza la estimación de energía a largo plazo de voz/música. Durante las tramas inactivas, solo se actualiza la estimación de energía del ruido.

La invención proporciona además un sistema que comprende un decodificador para procesamiento de señales de audio y un codificador para el procesamiento de señales de audio, donde el decodificador está diseñado de acuerdo con la invención reivindicada y/o el codificador está diseñado de acuerdo con la invención reivindicada.

En otro aspecto la invención presenta a un método para decodificar un flujo de bits de audio, en el que el método comprende:

40 derivar una señal de audio decodificada del flujo de bits, en el que la señal de audio decodificada comprende por lo menos una trama decodificada;

45 producir una señal de estimación de ruido que contiene una estimación del nivel y/o la forma espectral de un ruido en la señal de audio decodificada;

derivar una señal de ruido de confort de la señal de estimación de ruido y

50 combinar la trama decodificada de la señal de audio decodificada y la señal de ruido de confort para obtener una señal de audio de salida.

La invención presenta además un método para codificar señales de audio para producir un flujo de bits de audio, donde el método comprende:

55 determinar la relación de señal a ruido de una señal de audio de entrada sobre la base de una energía determinada de señal deseada de la señal de audio de entrada y una energía determinada de un ruido de la señal de audio de entrada;

60 producir una señal de audio con reducción de ruido;

producir una señal de audio codificada que corresponde a la señal de audio de entrada, donde, dependiendo de la relación de señal a ruido determinada de la señal de audio de entrada, se codifica la señal de audio de entrada o la señal de audio con reducción de ruido;

derivar el flujo de bits de la señal de audio codificada; y

transmitir una información complementaria, que indica si se está codificando la señal de audio de entrada o la señal de audio con reducción de ruido dentro del flujo de bits.

5 La invención proporciona además un flujo de bits producido de acuerdo con el método antes mostrado. El flujo de bits reivindicado contiene información complementaria, que indica si se está codificando la señal de audio de entrada o la señal de audio con reducción de ruido.

10 En un aspecto adicional, la invención proporciona un programa de ordenador para poner en práctica los métodos de la invención al ejecutarse en un ordenador o un procesador.

A continuación, se describen las realizaciones preferidas de la invención con respecto a los dibujos adjuntos, en los cuales:

15 La Fig. 1 ilustra un codificador de acuerdo con la técnica anterior;

La Fig. 2 ilustra una primera y una segunda realización de un codificador de acuerdo con la invención;

20 La Fig. 3 ilustra una primera y una segunda realización de un decodificador de acuerdo con la invención.

La Fig. 3 ilustra una primera realización de un decodificador 1 de acuerdo con la invención. El decodificador 1 está configurado para procesar un flujo de bits de audio codificado BS, en el que el decodificador 1 comprende:

25 un decodificador de flujos de bits 2 configurado para derivar una señal de audio decodificada DS del flujo de bits BS, en el que la señal de audio decodificada DS comprende por lo menos una trama decodificada;

30 un dispositivo de estimación de ruido 3 configurado para producir una señal de estimación de ruido NE que contiene una estimación del nivel y/o la forma espectral de un ruido N presente en la señal de audio decodificada DS;

un dispositivo generador de ruido de confort 4 configurado para derivar una señal de audio de ruido de confort CN de la señal de estimación de ruido NE; y

35 un combinador 5 configurado para combinar la trama decodificada de la señal de audio decodificada DS y la señal de ruido de confort CN para obtener una señal de audio de salida OS.

40 El decodificador de flujos de bits 2 puede ser un dispositivo o un programa de ordenador con capacidad para decodificar un flujo de bits de audio BS, que es un flujo de datos digitales que contiene información de audio. El proceso de decodificación da origen a una señal digital de audio decodificada DS, que puede ser alimentada a un conversor A/D para producir una señal de audio analógica, que luego puede ser alimentada a un altavoz para producir una señal audible.

45 La señal de audio decodificada DS comprende las denominadas tramas, en el que cada una de estas tramas contiene información de audio referente a un determinado momento. Dichas tramas se pueden clasificar en tramas activas y tramas inactivas, donde una trama activa es una trama que contiene componentes deseados WS de la información de audio, a la que también se hace referencia como señal WS deseada, como por ejemplo voz o música, en tanto que una trama inactiva es una trama que no contiene ningún componente deseado de la información de audio. Las tramas inactivas aparecen por lo general durante las pausas, donde no hay presencia de ningún componente deseado, tal como música o voz. Por lo tanto, las tramas inactivas habitualmente contienen únicamente ruido de fondo N.

50 El dispositivo de estimación de ruido 3 está configurado para producir una señal de estimación de ruido NE que contiene una estimación del nivel y/o la forma espectral de un ruido incluido en la señal de audio decodificada DS. Además, el dispositivo generador de ruido de confort 4 está configurado para derivar una señal de audio de ruido de confort CN de la señal de estimación de ruido NE. La señal de estimación de ruido NE puede ser una señal que contiene información con respecto a las características del ruido N contenido en la señal de audio decodificada DS en forma paramétrica. La señal de ruido de confort CN es una señal de audio artificial, que corresponde al ruido N contenido en la señal de audio decodificada DS. Estas características permiten que el ruido de confort CN suene como el ruido de fondo real N sin requerir información complementaria alguna con respecto al ruido de fondo N contenido en el flujo de bits BS.

60 El combinador 5 está configurado para combinar la trama decodificada de la señal de audio decodificada DS y la señal de ruido de confort CN para obtener una señal de audio de salida OS. Como resultado de eso la señal de audio de salida OS comprende tramas decodificadas, que comprenden ruido artificial CN. El ruido artificial CN

contenido en las tramas decodificadas permite enmascarar las alteraciones de la señal de audio de salida OS especialmente al transmitir el flujo de bits BS a bajas tasas de bits.

5 A diferencia de la técnica anterior, la presente invención aplica el principio de adición de ruido artificial de confort CN a las tramas activas o no activas decodificadas. El concepto de la invención se puede aplicar tanto a los modos DTX como no DTX.

10 La invención proporciona un método para intensificar la calidad de la voz ruidosa codificada y transmitida a bajas tasas de bits. A bajas tasas de bits, la codificación de la voz ruidosa, es decir la voz grabada con ruido de fondo N, por lo general no es tan eficiente como la codificación de voz limpia WS. La síntesis decodificada habitualmente es propensa a alteraciones. Los dos tipos de fuentes, el ruido N y la voz WS, no pueden ser codificados con eficiencia mediante un esquema de codificación que se basa en un único modelo de origen. La presente invención ofrece un concepto para modelar y sintetizar el ruido de fondo N del lado del decodificador y requiere muy poca o ninguna información complementaria. Esto se obtiene estimando el nivel y la forma espectral del ruido de fondo N del lado del decodificador y generando artificialmente un ruido de confort CN. El ruido generado CN se combina con la señal de audio decodificada DS y permite enmascarar las alteraciones por codificación durante las tramas decodificadas.

20 Además, el concepto se puede combinar con un esquema de reducción de ruido aplicado del lado del codificador. La reducción de ruido intensifica el nivel de la relación de señal a ruido (SNR) y mejora la eficiencia de la codificación de audio subsiguiente. Luego se compensa la cantidad de ruido faltante N en la señal de audio decodificada DS mediante el ruido de confort CN del lado del decodificador. Sin embargo, habitualmente suena más degradada o menos natural, ya que la reducción de ruido podría distorsionar los componentes de voz y causar alteraciones audibles de ruido musical además de las alteraciones de codificación. Un aspecto de la presente invención consiste en enmascarar dichas distorsiones desagradables mediante la adición de un ruido de confort CN del lado del decodificador. Cuando se usa un esquema de reducción de ruido, la adición de ruido de confort no deteriora la SNR. Más aun, el ruido de confort oculta una gran parte del molesto ruido musical típico de las técnicas de reducción de ruido.

30 En una realización preferida de la invención, la trama decodificada es una trama activa. Esta característica extiende el principio de adición de ruido de confort a las tramas activas decodificadas.

En una realización preferida de la invención la trama decodificada es una trama activa. Esta característica extiende el principio de adición de ruido de confort a las tramas inactivas decodificadas.

35 En una realización preferida de la invención, el dispositivo de estimación de ruido 3 comprende un dispositivo de análisis espectral 6 configurado para crear una señal de análisis AS que contiene el nivel y la forma espectral del ruido presente en la señal de audio decodificada DS y un dispositivo para producir estimaciones de ruido 7 configurado para producir la señal de estimación de ruido NE sobre la base de la señal de análisis AS.

40 En una realización preferida de la invención, el dispositivo generador de ruido de confort 4 comprende un generador de ruido 8 configurado para crear una señal de ruido de confort en el dominio de la frecuencia FD sobre la base de la señal de estimación de ruido NE y un sintetizador espectral 9 configurado para crear la señal de ruido de confort CN sobre la base de la señal de ruido de confort en el dominio de la frecuencia FD.

45 En una realización preferida de la invención, el decodificador 1 comprende un dispositivo conmutador 10 configurado para conmutar el decodificador 1 de manera alternada a un primer modo de operación o a un segundo modo de operación donde, en el primer modo de operación, la señal de ruido de confort CN es alimentada al combinador, en tanto que la señal de ruido de confort CN no es alimentada al combinador 5 en el segundo modo de operación. Estas características permiten abandonar el uso del ruido de confort CN artificial en situaciones en que no es necesario.

50 En una realización preferida de la invención, el decodificador 1 comprende un dispositivo de control 11 configurado para controlar el dispositivo conmutador 10 en forma automática, donde el dispositivo de control 10 comprende un detector de ruido 12 configurado para controlar el dispositivo conmutador 10 dependiendo de una relación de señal a ruido de la señal de audio decodificada DS en el que, en condiciones de baja relación de señal a ruido, el decodificador se conmuta al primer modo de operación y en condiciones de alta relación de señal a ruido al segundo modo de operación. En virtud de estas características se puede activar el uso de ruido de confort CN sólo en situaciones de voz ruidosa, es decir, no en situaciones de voz limpia o música limpia. Para discriminar entre condiciones de baja relación de señal a ruido y condiciones de alta relación de señal a ruido se puede definir y utilizar un umbral para la relación de señal a ruido.

60 En una realización preferida de la invención, el dispositivo de control 11 comprende un receptor de información complementaria 13 configurado para recibir información complementaria contenida en el flujo de bits BS, que corresponde a la relación de señal a ruido de la señal de audio decodificada DS, y configurado para crear una señal

de detección de ruido ND, donde el detector de ruido 12 conmuta el dispositivo conmutador 11 dependiendo de la señal de detección de ruido ND. Estas características permiten controlar el dispositivo conmutador 10 sobre la base de un análisis de la señal ejecutado por un dispositivo externo que produce y/o procesa el flujo de bits recibido BS. El dispositivo externo puede ser, especialmente, un codificador que produce el flujo de bits BS.

5 En una realización preferida de la invención, la información complementaria que corresponde a la relación de señal a ruido de la señal de audio decodificada DS consiste en por lo menos un bit dedicado en el flujo de bits BS. En general, un bit dedicado es un bit que contiene, solo o junto con otros bits dedicados, información definida. En este contexto, el bit dedicado puede indicar si la relación de señal a ruido es superior o inferior a un umbral predefinido.

10 En una realización preferida de la invención, el dispositivo generador de ruido de confort 4 está configurado para crear la señal de ruido de confort CN sobre la base de una señal de nivel objetivo de ruido de confort TNL. El nivel de ruido de confort CN agregado debe ser limitado para preservar la inteligibilidad y calidad. Esto se puede obtener escalando el ruido de confort CN usando una señal de ruido objetivo TNL que indica un nivel de ruido objetivo predeterminado.

15 En una realización preferida de la invención, la señal de nivel objetivo de ruido de confort TNL es ajustada dependiendo de una tasa de bits del flujo de bits BS. Por lo general, la señal de audio decodificada DS exhibe una relación de señal a ruido más elevada que la señal de entrada original, especialmente a bajas tasas de bits en que las alteraciones por codificación son más severas. Esta atenuación del nivel de ruido en la codificación de voz proviene del paradigma de modelado de origen que estima tener la voz como entrada. Por lo demás, la codificación por modelado de origen no es totalmente apropiada y no puede reproducir toda la energía de los componentes que no son de voz. Por ende, se puede ajustar la señal de nivel objetivo de ruido de confort TNL dependiendo de la tasa de bits para compensar de manera aproximada la atenuación de ruido introducida inherentemente por el proceso de codificación.

20 En una realización preferida de la invención la señal de nivel objetivo de ruido de confort TNL es ajustada dependiendo de un nivel de atenuación de ruido causado por un método de reducción de ruido aplicado al flujo de bits BS. Mediante estas características, se puede compensar la atenuación de ruido causada por un módulo de reducción de ruido en un codificador.

25 En una realización preferida de la invención, una energía de la señal de ruido de confort en el dominio de la frecuencia FD del ruido aleatorio $w(k)$ es ajustada dependiendo de la señal de nivel objetivo de ruido de confort TNL, que indica un nivel objetivo de ruido de confort g_{tar} , por cada frecuencia k de la siguiente manera: $E_w(k) = \max\{g_{tar} - \hat{E}_n(k); 0\}$, en el que $\hat{E}_n(k)$ se refiere a una estimación de la energía del ruido N de la señal de audio decodificada DS a la frecuencia k , emitida por el dispositivo para producir estimaciones de ruido 7. Mediante estas características, se puede intensificar la inteligibilidad y calidad de la señal de salida OS.

30 La Fig. 3 ilustra una segunda realización de un decodificador 1 de acuerdo con la invención. La segunda realización del decodificador 1 se basa en el decodificador 1 de la primera realización. A continuación, solo se exponen y explican las diferencias con respecto a la primera realización.

35 En una realización preferida de la invención, el dispositivo de control comprende un estimador de energía de la señal deseada 14 configurado para determinar la energía de la señal deseada WS de la señal de audio decodificada DS, un estimador de energía del ruido 15 configurado para determinar la energía de un ruido N de la señal de audio decodificada DS y un estimador de relación de señal a ruido 16 configurado para determinar la relación de señal a ruido de la señal de audio decodificada DS sobre la base de la energía de la señal deseada WS y sobre la base de la energía del ruido N, donde el dispositivo conmutador 10 se conmuta dependiendo de la relación de señal a ruido determinada por el dispositivo de control 11. En este caso, no es necesaria información complementaria alguna en el flujo de bits con respecto a la relación de señal a ruido. Por lo tanto, tampoco es necesario el receptor de información complementaria 13 de acuerdo con la primera realización.

40 En una realización preferida de la invención, el flujo de bits BS contiene tramas activas y tramas inactivas, donde el dispositivo de control 11 está configurado para determinar la energía de señal deseada WS de la señal de audio decodificada DS durante las tramas activas y para determinar la energía del ruido N de la señal de audio decodificada DS durante las tramas inactivas. De esta manera, se puede obtener de manera sencilla una gran precisión al estimar la relación de señal a ruido.

45 En una realización preferida de la invención, el flujo de bits BS contiene tramas activas y tramas inactivas, en el que el decodificador 1 comprende un receptor de información complementaria 17 configurado para discriminar entre las tramas activas y las tramas inactivas sobre la base de información complementaria contenida en el flujo de bits que indica si la trama actual es activa o inactiva. Mediante esta característica, se pueden identificar las tramas activas o las tramas inactivas, respectivamente, sin esfuerzo de cálculo.

En la realización preferida de la invención, el receptor de información complementaria 17 puede estar configurado para controlar y un conmutador 17a, que alimenta de manera alternada una señal de salida OW del estimador de energía de la señal deseada 14 o una señal de salida ON del estimador de energía del ruido 15 al estimador de relación de señal a ruido 16, en el que la señal de salida OW de un estimador de energía de la señal deseada 14 es alimentada al estimador de relación de señal a ruido 16 durante las tramas activas y en el que la señal de salida ON del cálculo estimativo de energía de ruido de 15 es alimentada al estimador de relación de señal a ruido 16 durante las tramas inactivas. Mediante estas características, se puede calcular la relación de señal a ruido de manera fácil y precisa.

10 En una realización preferida de la invención, el dispositivo de control 11 está configurado para determinar la energía de señal deseada de la señal de audio decodificada sobre la base de la señal de análisis AS. En este caso, la señal de análisis AS, que habitualmente debe ser computada para la estimación del ruido, puede ser reutilizada, por lo que se puede reducir la complejidad.

15 En una realización preferida de la invención, el dispositivo de control 11 está configurado para determinar la energía del ruido N de la señal de audio decodificada DS sobre la base de la señal de estimación de ruido NE. En esa realización, se puede reutilizar la señal de estimación de ruido NE, que por lo general se debe computar para la generación del ruido de confort, por lo que se puede reducir aún más la complejidad.

20 En una realización preferida de la invención, el decodificador 1 comprende un decodificador adicional de flujos de bits (que no se ilustra en las figuras), en el que el decodificador de flujos de bits 2 y el decodificador adicional de flujos de bits son de diferentes tipos, en el que el decodificador 1 comprende un conmutador (que no se ilustra en las figuras) configurado para alimentar ya sea la señal decodificada DS procedente del decodificador de flujos de bits 2 o la señal decodificada procedente del decodificador adicional de flujos de bits al dispositivo de estimación de ruido 3 y al combinador 5. Como la adición de ruido de confort se realiza cuando se utiliza el decodificador de flujos de bits 2, como así también cuando se utiliza el decodificador adicional de flujos de bits, se pueden reducir al mínimo las alteraciones por transición al conmutar entre el decodificador de flujos de bits 2 y el decodificador adicional de flujos de bits. Por ejemplo, el decodificador de flujos de bits 2 puede ser un decodificador de flujos de bits por predicción lineal excitada por códigos algebraicos (ACELP), en tanto que el decodificador de flujos de bits adicional puede ser un decodificador de flujos de bits de núcleo basado en transformadas (TCX).

En la figura 3 se describe el decodificador 1 de la invención, donde la adición del ruido de confort se realiza a ciegas en el dominio de la frecuencia. Para contar con un ruido de confort CN que se asemeja al ruido de fondo real N, se utiliza un dispositivo de estimación de ruido 3 en el decodificador 1 para determinar el nivel y la forma espectral del ruido de fondo N, sin que se necesaria información complementaria alguna.

El dispositivo generador de ruido de confort 4 el ruido de confort solo se puede activar en situaciones de voz ruidosa, es decir, no en situaciones de voz limpia ni música limpia. La discriminación se puede basar en la detección realizada en el codificador. En este caso, la decisión se debe transmitir empleando un bit dedicado. En una realización preferida, por el contrario, se aplica un dispositivo para producir estimaciones de ruido 7 que es similar al dispositivo de estimación de ruido usado en el codificador. Consiste en la estimación de una relación de señal a ruido a largo plazo adaptando por separados las estimaciones a largo plazo de la energía del ruido N o la energía de la señal deseada WS, como por ejemplo voz y/o música, dependiendo de la decisión del VAD. Esta última se puede deducir directamente del índice de los modos ACELP y TCX. En efecto, TCX y ACELP pueden ser ejecutadas en un modo específico denominado TCX-NA y ACELP-NA, respectivamente, cuando la señal es de tramas no activas de voz/música, es decir tramas con ruido de fondo solamente. Todos los modos de ACELP y TCX se refieren a tramas activas. Por lo tanto, se puede evitar la presencia de un bit de VAD dedicado en el flujo de bits.

El nivel de ruido de confort agregado debe ser limitado para preservar la inteligibilidad y calidad. Por lo tanto, se escala el ruido de confort para alcanzar un nivel de ruido objetivo predeterminado. Si g_{tar} denota el nivel objetivo de amplificación de ruido, se ajusta la energía E_w del ruido aleatorio $w(k)$ correspondiente a cada frecuencia k de la siguiente manera:

$$E_w(k) = \max\{g_{tar} - \hat{E}_n(k); 0\},$$

donde $\hat{E}_n(k)$ se refiere a una estimación de la energía de ruido presente en la salida de audio decodificada a la frecuencia k , emitida por el módulo de estimación de ruido.

Por lo general, la señal de audio decodificada DS exhibe una relación de señal a ruido más elevada que la señal de entrada original, especialmente a bajas tasas de bits en que las alteraciones por codificación son más severas. Esta atenuación del nivel de ruido en la codificación de voz proviene de un paradigma de modelado de origen que estima tener la voz como entrada. Por lo demás, la codificación por modelado de origen no es totalmente apropiada y no puede reproducir toda la energía de los componentes que no son de voz. Por lo tanto, con referencia al primer aspecto de la invención, utilizando el codificador ilustrado en la figura 1, ajusta el nivel objetivo de ruido de confort

g_{tar} dependiendo de la tasa de bits para compensar aproximadamente la atenuación de ruido introducida en forma inherente por el proceso de codificación.

5 En cuanto al segundo aspecto de la invención que utiliza el codificador ilustrado en la figura 2, el nivel objetivo de ruido de confort g_{tar} debe tener en cuenta, además, la atenuación de ruido causada por el módulo de reducción de ruido incluido en el codificador.

10 Más aun, la adición de ruido de confort descrita en la presente permite alisar las alteraciones de transición entre un tipo de codificación (por ejemplo) y el otro (por ejemplo, TCX) mediante la adición uniforme de un ruido de confort en todas las tramas.

La figura 1 ilustra un codificador de acuerdo con la técnica anterior que se puede utilizar en combinación con los decodificadores ilustrados en la figura 3.

15 La señal de entrada IS es codificada directamente por el codificador de flujos de bits 20. El codificador de flujos de bits 20 puede ser un codificador de voz o una conmutación de esquemas de bajo retardo entre un codificador de voz ACELP y un codificador de audio basado en transformadas TCX. El codificador de flujos de bits 20 comprende un codificador de señales 21 para codificar la señal IS y un productor de flujos de bits 22 para generar el flujo de bits BS necesario para producir la señal decodificada DS en el decodificador 1. En paralelo, la señal de entrada IS es
20 analizada por un módulo denominado analizador de señales 23, que comprende un dispositivo de estimación de ruido 24. En la realización preferida, el dispositivo de estimación de ruido 24 es igual al utilizado en G.718. Consiste en un dispositivo de análisis espectral 25 seguido por un dispositivo para producir estimaciones de ruido 26. El espectro SI de la señal original 15 y el producir estimaciones de ruido 26. El espectro SI de la señal original 15 y el espectro NI del ruido estimado son ingresados al módulo de reducción de ruido 27. El módulo de reducción de ruido
25 27 atenúa el nivel de ruido de fondo en la señal intensificada en el dominio de la frecuencia FS. La cantidad de reducción está dada por la señal de nivel de atenuación TAS. La señal intensificada en el dominio del tiempo (señal de audio con reducción de ruido) TS se genera una vez realizada la síntesis espectral por el dispositivo de síntesis espectral 28. La señal TS se utiliza para deducir algunas características, como por ejemplo la estabilidad del tono, que luego es aprovechada por el detector de actividad de señal 29 para TS se utiliza para deducir algunas
30 características, como por ejemplo la estabilidad del tono, que luego es aprovechada por el detector de actividad de señal 29 para el tono, que luego es aprovechada por el detector de actividad de señal 29 para discriminar entre tramas activas e inactivas. El resultado de la clasificación puede ser utilizado a su vez por el módulo codificador 18. En una realización preferida, se utiliza un modo de codificación específico para manejar las tramas inactivas. De esta manera, el decodificador 1 puede deducir la bandera de actividad de señal (bandera VAD) del flujo de bits sin
35 necesidad de un bit dedicado.

La figura 2 ilustra una primera realización de un codificador 18 de acuerdo con la invención. El codificador 18 ilustrado en la figura 2 se basa en el codificador 18 mostrado en la figura 1.

40 El codificador 18 mostrado en la figura 2 está configurado para producir un flujo de bits de audio BS, en el que el codificador 18 comprende:

45 un codificador de flujos de bits 20 configurado para producir una señal de audio codificada ES que corresponde a una señal de audio de entrada IS y para derivar el flujo de bits BS de la señal de audio codificada ES;

50 un analizador de señales 19 que tiene un estimador de relación de señal a ruido 33 configurado para determinar la relación de señal a ruido de la señal de audio de entrada IS sobre la base de la energía de señal deseada WS de la señal de audio de entrada IS determinada por un estimador de energía de la señal deseada 31 y sobre la base de la energía de un ruido N de la señal de audio de entrada IS determinada por el estimador de energía de ruido 32;

un dispositivo de reducción de ruido 27, 28 configurado para producir una señal de audio con reducción de ruido TS y

55 un dispositivo conmutador 35 configurado para alimentar, dependiendo de la relación de señal a ruido determinada de la señal de audio de entrada IS, la señal de audio de entrada IS o la señal de audio con reducción de ruido TS al codificador de flujos de bits 20 para codificar la señal respectiva IS, TS, en el que el codificador de flujos de bits 20 está configurado para transmitir una información complementaria dentro del flujo de bits, que indica si se está codificando la señal de audio de entrada IS o la señal de audio con reducción de
60 ruido TS.

El codificador de flujos de bits 20 puede ser un dispositivo o un programa de ordenador con capacidad para codificar una señal de audio, que es una señal de datos digitales que contiene información de audio. El proceso de codificación da origen a un flujo de bits digital, que se puede transmitir a través de un vínculo de datos digitales a un

decodificador existente en un sitio remoto.

La parte del codificador de acuerdo con una realización de la invención se da en la figura 4. La principal diferencia en comparación con la figura 3 reside en el hecho de que codifica la salida de la reducción de ruido, es decir, la señal TS intensificada. Para evitar las distorsiones innecesarias en situaciones sin ruido (voz limpia o música limpia), se aplica la reducción de ruido sólo en el caso ruido (voz limpia o música limpia), se aplica la reducción de ruido sólo en el caso de voz ruidosa y de lo contrario se la omite. La discriminación entre señales con y sin ruido se obtiene mediante la estimación de la energía a largo plazo de la señal deseada WS (voz o música) realizada por el estimador de energía de la señal deseada 31 y mediante la estimación de la energía a largo plazo del ruido N realizada por el estimador de energía del ruido 32. Para este fin, el estimador de energía de la señal deseada 31 recibe la señal del espectro SI correspondiente a la señal de entrada IS provista por el dispositivo de análisis espectral 25. A su vez, el estimador de energía del ruido recibe la señal de estimación de ruido NI correspondiente a la señal de entrada IS provista por el dispositivo para producir estimaciones de ruido 26. Durante las tramas activas, a largo plazo de voz/música WE. Durante las tramas inactivas, solo se actualiza la estimación de energía de ruido NE. La energía a largo plazo se computa mediante un filtrado autorregresivo de primer orden de la energía de entrada de trama (durante las tramas activas) o usando la salida del módulo de estimación de ruido (durante las tramas inactivas). De esta manera, el estimador de relación de señal a ruido 33 puede calcular una estimación de la relación de señal a ruido SR, que contiene la relación de la energía a largo plazo de la voz o música WS durante la energía a largo plazo del ruido N. La relación de señal a ruido RS es alimentada a un detector de ruido 34 que determina si la trama actual contiene una señal de audio ruidosa o una señal de audio limpia. Si la relación de señal a ruido RS es inferior a un umbral predeterminado, se considera que la trama es de voz ruidosa, de lo contrario, se la clasifica como voz limpia.

El resultado de la clasificación es emitido como salida en forma de señal de bandera de ruido NF, que se utiliza para controlar el conmutador 35. Además, la señal de bandera de ruido NF es alimentada al codificador de flujos de bits 20. El codificador de flujos de bits 20 está configurado para producir y transmitir una información complementaria sobre la base de la señal de bandera de ruido NF dentro del flujo de bits, que indica si se está codificando la señal de audio de entrada IS o la señal de audio con reducción de ruido TS. Decodificando esta bandera, un decodificador puede ajustar el nivel objetivo de ruido en forma automática sin necesidad de clasificar la señal decodificada DS como ruidosa o limpia.

La figura 2 ilustra una segunda realización de un codificador 18 de acuerdo con la invención. A continuación, se explican las características adicionales. En la figura 2, el analizador de señales 30 comprende un detector de actividad de señales 36 que recibe la señal de espectro SI correspondiente a la señal de entrada IS y la señal de estimación de ruido NI. El detector de actividad de señales 36 está configurado para discriminar entre tramas activas y tramas inactivas sobre la base de estas dos señales. El detector de actividad de señales produce una señal de actividad de señal SA que se transmite, por un lado, al codificador de flujos de bits 20 para adaptar el flujo de bits BS a la actividad de señales y, por otro lado, se la utiliza para conmutar un conmutador 37 que está configurado para alimentar en forma alternada la señal de energía de señal deseada WE o la señal de energía de ruido EN al estimador de relación de señal a ruido 33.

En una realización de un formato de trama FF del flujo de bits BS de acuerdo con la invención. La trama de acuerdo con el formato de trama FF comprende un vector de señales SV que tiene una pluralidad de bits que están situados en las posiciones de 0 a n. En la posición n+1 se encuentra un bit que es una bandera de actividad AF que indica si la trama está en una trama activa o inactiva. Asimismo, se prevé el bit de la posición n+2 que es una bandera de ruido NF que indica si la trama contiene una señal ruidosa o una señal limpia. En la posición n+3 está dispuesto un bit que es un bit de relleno PB.

En una realización preferida de la invención, la información complementaria que indica si la trama actual es activa o inactiva consiste en por lo menos un bit dedicado en el flujo de bits.

En resumen, se puede decir que, en un aspecto de la invención, se codifica la señal original en el decodificador 1. Se la codifica antes de agregarla a un ruido de confort CN generado artificialmente. El dispositivo generador de ruido de confort 4 no requiere o solo requiere una pequeña cantidad de información complementaria. En una primera realización, el dispositivo generador de ruido de confort 4 no requiere información complementaria y todo el procesamiento se realiza a ciegas. En la realización preferida, el dispositivo generador de ruido de confort 4 necesita recuperar la información VAD (resultado de la clasificación de tramas activas e inactivas) del flujo de bits BS, que ya puede estar presente en el flujo de bits, y utilizarla para otros fines. En una tercera realización, el dispositivo generador de ruido de confort 4 requiere del codificador 18 una bandera de voz ruidosa para discriminar entre voz limpia y voz ruidosa. Se puede imaginar cualquier tipo de información codificada en forma paramétrica que pueda ayudar a dirigir el dispositivo generador de ruido de confort 4.

En otro aspecto de la invención, en primer lugar, se aplica la reducción de ruido a la señal original IS y se envía una señal intensificada TS al codificador de flujos de bits 20, se la codifica y transmite. Al final de la decodificación, se

agrega un ruido de confort CN generado en forma artificial a la señal decodificada (intensificada) DS. El nivel objetivo de atenuación empleado para la reducción de ruido en el codificador es un valor fijo compartido con el módulo CNG en el decodificador. Por consiguiente, no es necesario transmitir explícitamente el nivel objetivo de atenuación.

5 Aunque se han descrito algunos aspectos en el contexto de un aparato, es obvio que estos aspectos también representan una descripción del método correspondiente, en el cual un bloque o dispositivo corresponde a una etapa del método o a una característica de un paso del método. De manera análoga, los aspectos descritos en el contexto de un paso del método también representan una descripción de un bloque o ítem correspondiente o de una característica de un aparato correspondiente. Algunas o todas las etapas del método pueden ser ejecutadas por medio de (o utilizando) un aparato de hardware, como, por ejemplo, un microprocesador, un ordenador programable o un circuito electrónico. En algunas realizaciones, una cualquiera o más de las etapas más importantes del método pueden ser ejecutados por ese tipo de aparato.

15 Dependiendo de ciertos requisitos de implementación, las realizaciones de la invención pueden ser implementadas en hardware o en software. La implementación se puede realizar empleando un medio de almacenamiento digital, por ejemplo, un disco blando, un DVD, un Blue-Ray, un CD, una ROM, una PROM, una EPROM, una EEPROM o una memoria FLASH, que tiene almacenadas en la misma señales control legibles electrónicamente, que cooperan (o tienen capacidad para cooperar) con un sistema de ordenador programable, de tal manera que se ejecute el método respectivo. Por lo tanto, el medio de almacenamiento digital puede ser legible por un ordenador.

20 Algunas realizaciones de acuerdo con la invención comprenden un transportador de datos que comprende señales de control legibles electrónicamente, con capacidad para cooperar con un sistema de computación programable de tal manera que se ejecute uno de los métodos descritos en el presente documento.

25 En general, las realizaciones de la presente invención pueden ser implementadas en forma de producto programa de ordenador con un código de programa, donde el código de programa cumple la función de ejecutar uno de los métodos al ejecutarse el programa de ordenador en un ordenador. El código de programa puede ser almacenado, por ejemplo, en un portador legible por una máquina.

30 Otras realizaciones comprenden el programa de ordenador para ejecutar uno de los métodos aquí descritos, almacenado en un portador legible por una máquina.

35 En otras palabras, una realización del método de la invención es, por lo tanto, un programa de ordenador que tiene un código de programa para realizar uno de los métodos aquí descritos al ejecutarse el programa de ordenador en un ordenador.

40 Una realización adicional del método de la invención es, por lo tanto, un portador de datos (o medio de almacenamiento digital, o medio legible por ordenador) que comprende, grabado en el mismo, el programa de ordenador para ejecutar uno de los métodos aquí descritos. El portador de datos, el medio de almacenamiento digital o el medio grabado son por lo general tangibles y/o no transitorios.

45 Otra realización del método de la invención es, por lo tanto, un flujo de datos o una secuencia de señales que representa el programa de ordenador para ejecutar uno de los métodos aquí descritos. El flujo de datos o la secuencia de señales pueden estar configurados, por ejemplo, para ser transferido a través de una conexión de comunicación de datos, por ejemplo, a través de Internet.

Otra realización comprende unos medios de procesamiento, por ejemplo, un ordenador, un dispositivo lógico programable, configurado o adaptado para ejecutar uno de los métodos aquí descritos.

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

55 Otra realización de acuerdo con la invención comprende un aparato o sistema configurado para transferir (por ejemplo, por vía electrónica u óptica) un programa de ordenador para poner en práctica uno de los métodos aquí descritos en un receptor. El receptor puede ser, por ejemplo, un ordenador, un dispositivo móvil, un dispositivo de memoria y similar. El aparato o sistema puede comprender, por ejemplo, un servidor de archivos para transferir el programa de ordenador al receptor.

60 En algunas realizaciones, se puede utilizar un dispositivo lógico programable (por ejemplo, una matriz de puertas programables en el campo) para ejecutar algunas o todas las funcionalidades de los métodos aquí descritos. En algunas realizaciones, una matriz de puertas programables en el campo puede cooperar con un microprocesador para ejecutar uno de los métodos aquí descritos. Por lo general, los métodos son ejecutados preferentemente por cualquier aparato de hardware.

Signos de referencia:

	1	decodificador
	2	decodificador de flujos de bits
5	3	dispositivo de estimación de ruido
	4	dispositivo generador de ruido de confort
	5	combinador
	6	dispositivo de análisis espectral
	7	dispositivo para producir estimaciones de ruido
10	8	generador de ruido
	9	sintetizador espectral
	10	dispositivo conmutador
	11	dispositivo de control
	12	detector de ruido
15	13	receptor de información complementaria
	14	estimador de energía de señal deseada
	15	estimador de energía de ruido
	16	estimador de relación de señal a ruido
	17	receptor de información complementaria
20	17a	conmutador
	18	codificador
	19	analizador de señales
	20	codificador de flujos de bits
	21	codificador de señales
25	22	productor de flujos de bits
	23	analizador de señales
	24	dispositivo de estimación de ruido
	25	dispositivo de análisis espectral
	26	dispositivo para producir estimaciones de ruido
30	27	módulo de reducción de ruido
	28	dispositivo de análisis espectral
	29	detector de actividad de señales
	30	analizador de señales
	31	estimador de energía de señal deseada
35	32	estimador de energía de ruido
	33	estimador de relación de señal a ruido
	34	detector de ruido
	35	conmutador
	36	detector de actividad de señales
40	37	conmutador
	BS	flujo de bits de audio codificado
	DS	señal de audio decodificada
	NE	señal de estimación de ruido
45	N	ruido
	CN	señal de ruido de confort
	OS	señal de salida de audio
	AS	señal de análisis
	FD	señal de ruido de confort en el dominio de la frecuencia
50	ND	señal de detección de ruido
	TNL	nivel objetivo de ruido de confort
	IS	señal de entrada
	ES	señal codificada
	OW	señal de salida del estimador de energía de señal deseada
55	ON	señal de salida del estimador de energía del ruido
	SI	señal de espectro correspondiente a la señal de entrada
	NI	señal de estimación de ruido correspondiente a la señal de entrada
	TAS	señal de atenuación objetivo
	FS	señal intensificada en el dominio de la frecuencia
60	TS	señal de audio con reducción de ruido
	AD	señal de detector de actividad
	WE	señal de energía de señal deseada
	EN	señal de energía de ruido
	RS	señal de relación de señal a ruido

	NF	bandera de ruido
	SA	señal de actividad de señal
	FF	formato de trama
	SV	vector de señal
5	AF	bandera de actividad
	NF	señal de bandera de ruido
	PB	bit de relleno

Referencias:

- 10
- [1] Recomendación ITU-T G.718: "Frame error robust narrow-band and wideband embedded variable bit-rate coding of speech and audio from 8-32 kbit/s"
- 15
- [2] 3GPP TS 26.190 "Adaptive Multi-Rate wideband speech transcoding," 3GPP Technical Specification.
- [3] WO 02/101724 A1
- [4] WO 2010/003618 A2
- 20

REIVINDICACIONES

1. Un decodificador que está configurado para procesar un flujo de bits de audio codificado (BS), en el que el decodificador (1) comprende:
- 5 un decodificador de flujos de bits (2) configurado para derivar una señal de audio decodificada (DS) del flujo de bits (BS), en el que la señal de audio decodificada (DS) comprende por lo menos una trama decodificada;
- un dispositivo de estimación de ruido (3) configurado para producir una señal de estimación de ruido (NE) que contiene una estimación del nivel y/o la forma espectral de un ruido (N) en la señal de audio decodificada (DS);
- 10 un dispositivo generador de ruido de confort (4) configurado para derivar una señal de ruido de confort (CN) de la señal de estimación de ruido (NE); y
- un combinador (5) configurado para combinar la trama decodificada de la señal de audio decodificada (DS) y la señal de ruido de confort (CN) para obtener una señal de audio de salida (OS), de modo que la trama decodificada en la señal de audio de salida (OS) comprende ruido artificial correspondiente al ruido (N) contenido en la señal de audio decodificada (DS).
- 15
2. Un decodificador de acuerdo con la reivindicación anterior, en el que la trama decodificada es una trama activa.
3. Un decodificador de acuerdo con una de las reivindicaciones anteriores, en el que la trama decodificada es una trama inactiva.
- 20
4. Un decodificador de acuerdo con una de las reivindicaciones anteriores, en el que el dispositivo de estimación de ruido (3) comprende un dispositivo de análisis espectral (6) configurado para crear una señal de análisis (AS) que contiene el nivel y la forma espectral del ruido (N) en la señal de audio decodificada (DS) y un dispositivo para producir estimaciones de ruido (7) configurado para producir la señal de estimación de ruido (NE) sobre la base de la señal de análisis (AS).
- 25
5. Un decodificador de acuerdo con una de las reivindicaciones anteriores, en el que el dispositivo generador de ruido de confort (4) comprende un generador de ruido (8) configurado para crear una señal de ruido de confort en el dominio de frecuencia (FD) sobre la base de la señal de estimación de ruido (NE) y un sintetizador espectral (9) configurado para crear la señal de ruido de confort (CN) sobre la base de la señal de ruido de confort en el dominio de frecuencia (FD).
- 30
6. Un decodificador de acuerdo con una de las reivindicaciones anteriores, en el que el decodificador (1) comprende un dispositivo conmutador (10) configurado para conmutar el decodificador en forma alternada a un primer modo de operación o a un segundo modo de operación, en el que en el primer modo de operación la señal de ruido de confort (CN) es alimentada al combinador (5), mientras que la señal de ruido de confort (CN) no es alimentada al combinador (5) en el segundo modo de operación.
- 35
7. Un decodificador de acuerdo con la reivindicación anterior, en el que el decodificador (1) comprende un dispositivo de control (11) configurado para controlar el dispositivo conmutador (10) en forma automática, en el que el dispositivo de control (11) comprende un detector de ruido (12) y está configurado para controlar el dispositivo conmutador (11) dependiendo de una relación de señal a ruido de la señal de audio decodificada (DS), en el que, en condiciones de baja relación de señal a ruido, el decodificador (1) se conmuta al primer modo de operación y en condiciones de alta relación de señal a ruido al segundo modo de operación.
- 40
8. Un decodificador de acuerdo con la reivindicación anterior, en el que el dispositivo de control (11) comprende un receptor de información complementaria (13) configurado para recibir información complementaria contenida en el flujo de bits (BS), que corresponde a la relación de señal a ruido de la señal de audio decodificada (DS), y configurado para crear una señal de detección de ruido (ND), en el que el detector de ruido (12) conmuta el dispositivo conmutador (11) dependiendo de la señal de detección de ruido (ND).
- 45
9. Un decodificador de acuerdo con la reivindicación anterior, en el que la información complementaria que corresponde a la relación de señal a ruido de la señal de audio decodificada (DS) consiste en por lo menos un bit dedicado en el flujo de bits (BS).
- 50
10. Un decodificador de acuerdo con una de las reivindicaciones 7 a 9, en el que el dispositivo de control (11) comprende un estimador de energía de señal deseada (14) configurado para determinar la energía de señal deseada (WS) de la señal de audio decodificada (DS), un estimador de energía del ruido (15) configurado para determinar la energía de un ruido (N) de la señal de audio decodificada (DS) y un estimador de relación de señal a ruido (16) configurado para determinar la relación de señal a ruido de la señal de audio decodificada (DS) sobre la base de la energía de señal deseada (WS) y sobre la base de la energía del ruido (N), en el que el dispositivo conmutador (11) se conmuta dependiendo de la relación de señal a ruido determinada por el dispositivo de control (11).
- 55
- 60

11. Un decodificador de acuerdo con una de las reivindicaciones 7 a 10, en el que el flujo de bits comprende tramas activas y tramas inactivas, en el que el dispositivo de control (11) está configurado para determinar la energía de señal deseada (WS) de la señal de audio decodificada (DS) durante las tramas activas y para determinar la energía del ruido (N) de la señal de audio decodificada (DS) durante las tramas inactivas.
- 5 12. Un decodificador de acuerdo con una de las reivindicaciones anteriores, en el que el flujo de bits comprende tramas activas y tramas inactivas, en el que el decodificador (1) comprende un receptor de información complementaria (17) configurado para discriminar entre las tramas activas y las tramas inactivas sobre la base de información complementaria contenida en el flujo de bits (BS) que indica si la trama actual es activa o inactiva.
- 10 13. Un decodificador de acuerdo con la reivindicación anterior, en el que la información complementaria que indica si la trama actual es activa o inactiva consiste en por lo menos un bit dedicado en el flujo de bits (BS).
- 15 14. Un decodificador de acuerdo con la reivindicación 4 y de acuerdo con una de las reivindicaciones 7 a 13, en el que el dispositivo de control (11) está configurado para determinar la energía de señal deseada (WS) de la señal de audio (DS) decodificada sobre la base de la señal de análisis (AS).
- 20 15. Un decodificador de acuerdo con una de las reivindicaciones 7 a 14, en el que el dispositivo de control (11) está configurado para determinar la energía del ruido (N) de la señal de audio decodificada (DS) sobre la base de la señal de estimación de ruido (NE).
- 25 16. Un decodificador de acuerdo con una de las reivindicaciones anteriores, en el que el dispositivo generador de ruido de confort (4) está configurado para crear la señal de ruido de confort (CN) sobre la base de una señal de nivel objetivo de ruido de confort (TNL).
- 30 17. Un decodificador de acuerdo con la reivindicación anterior, en el que la señal de nivel objetivo de ruido de confort (TNL) está ajustada dependiendo de una tasa de bits del flujo de bits (BS).
- 35 18. Un decodificador de acuerdo con la reivindicación 15 o 17, en el que la señal de nivel objetivo de ruido de confort (TNL) está ajustada dependiendo de un nivel de atenuación de ruido causado por un método de reducción de ruido aplicado al flujo de bits (BS).
- 40 19. Un decodificador de acuerdo con una de las reivindicaciones 16 a 18, en el que se ajusta una energía $E_w(k)$ de una banda de frecuencia k de la señal de ruido de confort en el dominio de la frecuencia (FD) dependiendo de la señal de nivel objetivo de ruido de confort (TNL), que indica un nivel objetivo de ruido de confort g_{tar} , correspondiente a cada banda de frecuencia k como $E_w(k) = \max\{g_{tar} - \hat{E}_n(k); 0\}$, en el que $\hat{E}_n(k)$ se refiere a una estimación de la energía del ruido (N) de la señal de audio decodificada (DS) en la banda de frecuencia k , emitida por el dispositivo para producir estimaciones de ruido (7).
- 45 20. Un decodificador de acuerdo con una de las reivindicaciones anteriores, en el que el decodificador (1) comprende un decodificador adicional de flujos de bits, en el que el decodificador de flujos de bits (2) y el decodificador adicional de flujos de bits son de diferentes tipos, en el que el decodificador (1) comprende un conmutador configurado para alimentar la señal decodificada (DS) procedente del decodificador de flujos de bits (2) o la señal decodificada procedente del decodificador adicional de flujos de bits al dispositivo de estimación de ruido (3) y al combinador (5).
- 50 21. Un codificador que está configurado para producir un flujo de bits de audio (BS), en el que el codificador (18) comprende:
 un codificador de flujos de bits (20) configurado para producir una señal de audio codificada (ES) que corresponde a una señal de audio de entrada (IS) y para derivar el flujo de bits (BS) de la señal de audio codificada (ES);
 un analizador de señales (30) que tiene un estimador de relación de señal a ruido (33) configurado para determinar la relación de señal a ruido de la señal de audio de entrada (IS) sobre la base de la energía de una señal deseada (WS) de la señal de audio de entrada (IS) determinada por un estimador de energía de la señal deseada (31) y en base a una energía de un ruido (N) de la señal de audio de entrada (IS) determinada por un estimador de energía de ruido (32);
 un dispositivo de reducción de ruido (27, 28) configurado para producir una señal de audio con reducción de ruido (TS); y
 un dispositivo conmutador (35) configurado para alimentar, dependiendo de la relación de señal a ruido determinada de la señal de audio de entrada (IS), ya sea la señal de audio de entrada (IS) o la señal de audio con reducción de ruido (TS), al codificador de flujos de bits (20) para codificar la señal respectiva (IS, TS), en el que el codificador de flujos de bits (20) está configurado para transmitir una información complementaria (NF), que indica si se está codificando la señal de audio de entrada (IS) o la señal de audio con reducción de ruido (TS) dentro del flujo de bits (BS).
- 60

22. Un sistema que comprende un decodificador (1) y un codificador (18), en el que el decodificador (1) está diseñado de acuerdo con una de las reivindicaciones 1 a 19 y/o el codificador (18) está diseñado de acuerdo con la reivindicación 21.

5
23. Un método para decodificar un flujo de bits de audio (BS), en el que el método comprende:
derivar una señal de audio decodificada (DS) del flujo de bits (BS), en el que la señal de audio decodificada (DS) comprende por lo menos una trama decodificada;
10 producir una señal de estimación de ruido (NE) que contiene una estimación del nivel y/o la forma espectral de un ruido (N) en la señal de audio decodificada (DS);
derivar una señal de ruido de confort (CN) de la señal de estimación de ruido (NE); y
combinar la trama decodificada de la señal de audio decodificada (DS) y la señal de ruido de confort (CN) para
15 obtener una señal de audio de salida (OS), de tal manera que la trama decodificada en la señal de audio de salida (OS) comprenda ruido artificial correspondiente al ruido (N) contenido en la señal de audio decodificada (DS).

24. Un método de codificación de señales de audio para producir un flujo de bits de audio (BS), en el que el método comprende:

20 determinar la relación de señal a ruido de una señal de audio de entrada (IS) sobre la base de una energía determinada de señal deseada (WS) de la señal de audio de entrada (IS) y una energía determinada de un ruido (N) de la señal de audio de entrada (IS);
producir una señal de audio con reducción de ruido (TS);
25 producir una señal de audio codificada (ES) que corresponde a la señal de audio de entrada (IS), en el que, dependiendo de la relación de señal a ruido determinada de la señal de audio de entrada (IS), se codifica la señal de audio de entrada (IS) o la señal de audio con reducción de ruido (TS);
derivar el flujo de bits (BS) de la señal de audio codificada (ES); y
30 transmitir una información complementaria (NF), que indica si se está codificando la señal de audio de entrada (IS) o la señal de audio con reducción de ruido (TS) dentro del flujo de bits (BS).

25. Un flujo de bits producido de acuerdo con el método de la reivindicación 24.

35 26. Un programa de ordenador para realizar, al ejecutarse en una computadora o un procesador, el método de acuerdo con la reivindicación 23 o 24.

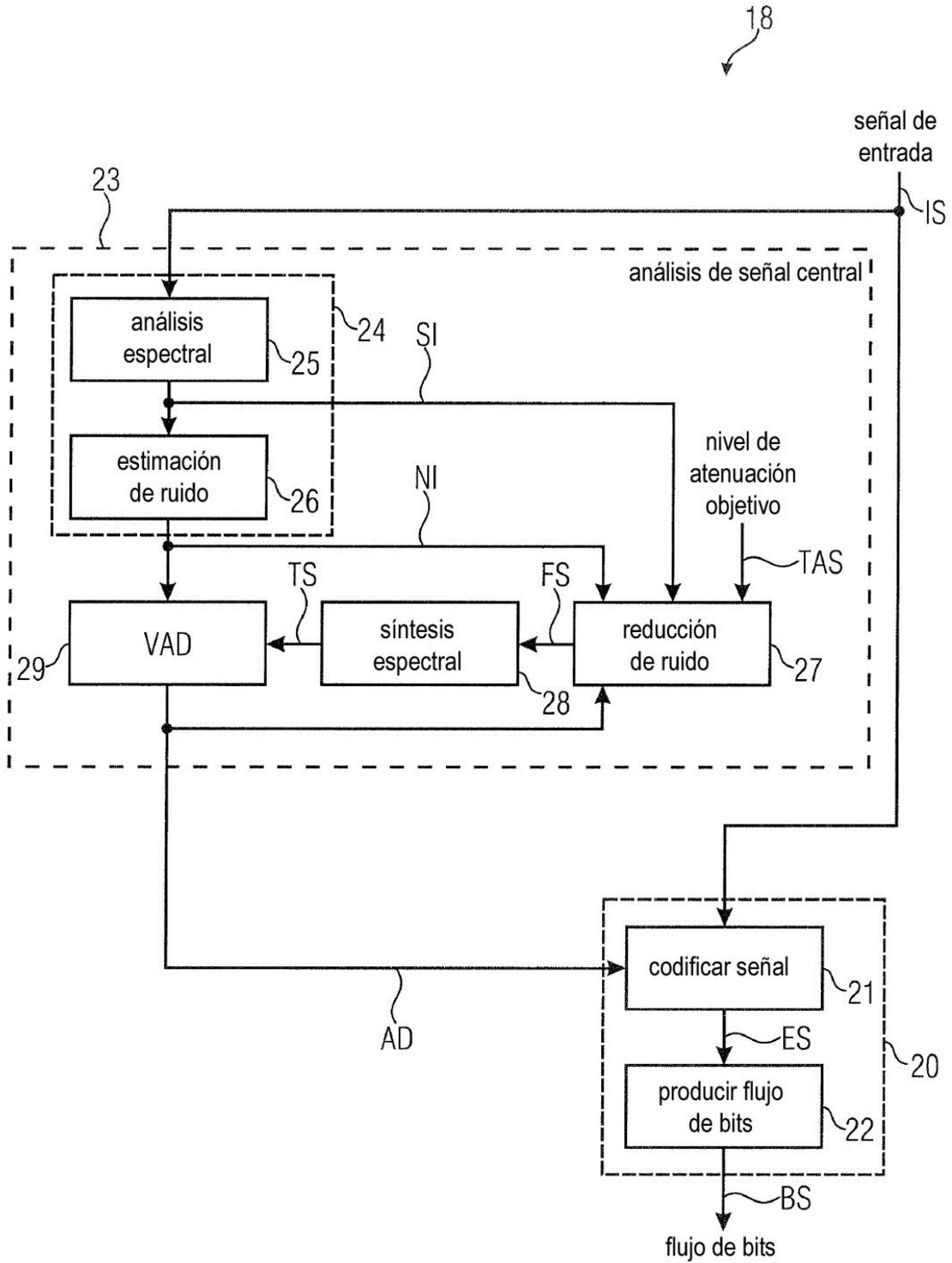


FIGURA 1
(TÉCNICA ANTERIOR)

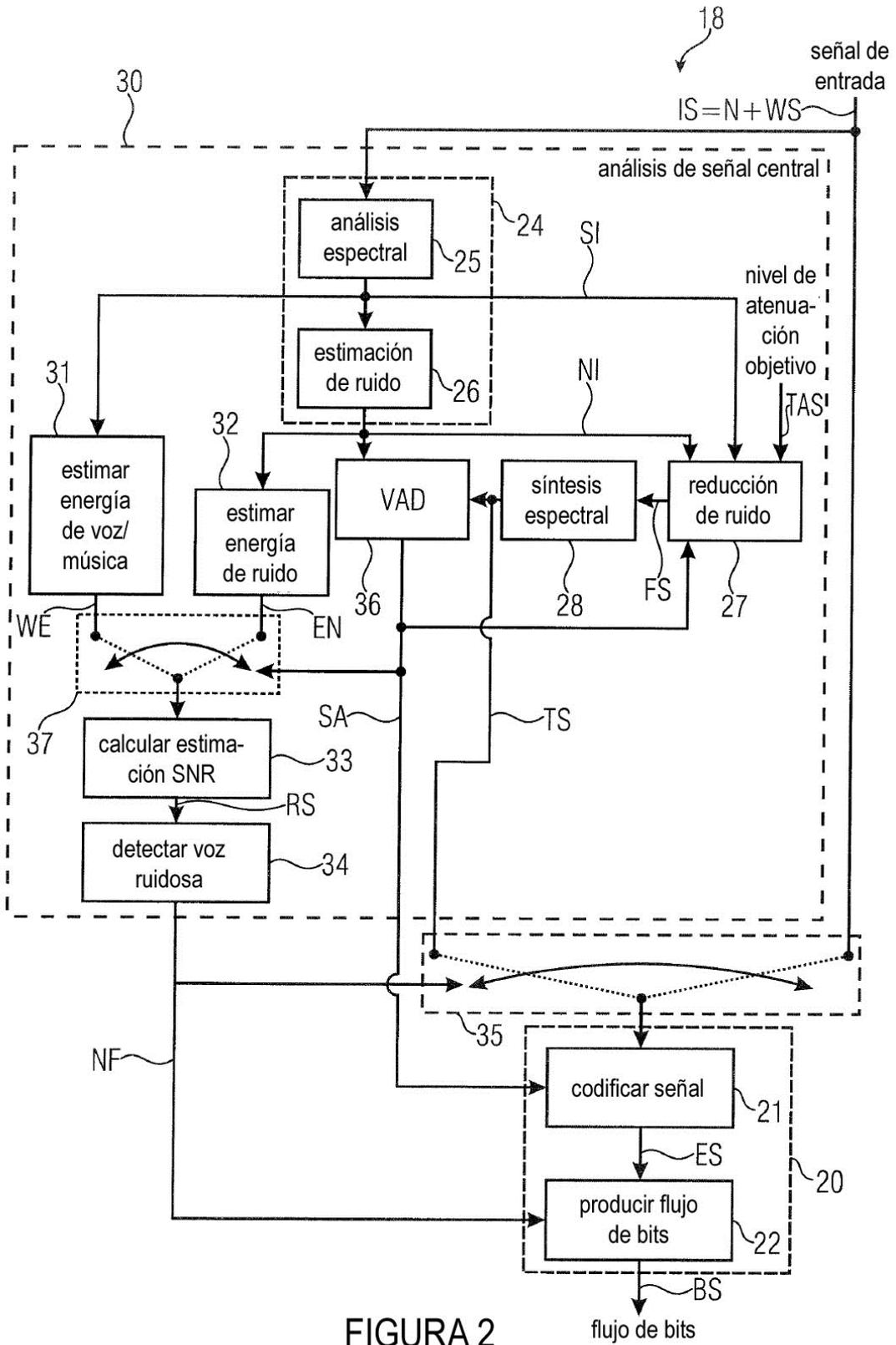


FIGURA 2

