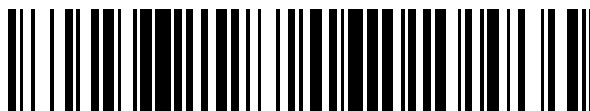


19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 768 179**

51 Int. Cl.:

G10L 19/20 (2013.01)

G10L 21/038 (2013.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

86 Fecha de presentación y número de la solicitud internacional: **28.01.2014 PCT/EP2014/051641**

87 Fecha y número de publicación internacional: **07.08.2014 WO14118185**

96 Fecha de presentación y número de la solicitud europea: **28.01.2014 E 14701755 (2)**

97 Fecha y número de publicación de la concesión europea: **13.11.2019 EP 2951822**

54 Título: **Codificador de audio, decodificador de audio, método para proporcionar una información de audio codificada, método para proporcionar una información de audio decodificada, programa informático y representación codificada utilizando una ampliación de ancho de banda adaptada a la señal**

30 Prioridad:
29.01.2013 US 201361758205 P

45 Fecha de publicación y mención en BOPI de la traducción de la patente:
22.06.2020

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

72 Inventor/es:
**DISCH, SASCHA;
HELMRICH, CHRISTIAN;
HILPERT, JOHANNES;
ROBILLIARD, JULIEN;
SCHMIDT, KONSTANTIN y
WILDE, STEPHAN**

74 Agente/Representante:
ARIZTI ACHA, Monica

ES 2 768 179 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

5 Codificador de audio, decodificador de audio, método para proporcionar una información de audio codificada, método para proporcionar una información de audio decodificada, programa informático y representación codificada utilizando una ampliación de ancho de banda adaptada a la señal

Campo técnico

10 Todas las siguientes apariciones de la palabra "realización(es)", si se refiere a combinaciones de características diferentes de las definidas por las reivindicaciones independientes, se refieren a ejemplos que se presentaron originalmente pero que no representan realizaciones de la invención actualmente reivindicada; estos ejemplos todavía se muestran con fines ilustrativos únicamente.

15 Las realizaciones de acuerdo con la invención están relacionadas con un codificador de audio para proporcionar una información de audio codificada basándose en una información de audio de entrada.

Otras realizaciones de acuerdo con la invención están relacionadas con un decodificador de audio para proporcionar una información de audio decodificada basándose en una información de audio codificada.

20 Otras realizaciones de acuerdo con la invención están relacionadas con un método para proporcionar una información de audio codificada basándose en una información de audio de entrada.

Otras realizaciones de acuerdo con la invención están relacionadas con un método para proporcionar una información de audio decodificada basándose en una información de audio codificada.

25 Otras realizaciones de acuerdo con la invención están relacionadas con un programa informático para llevar a cabo uno de dichos métodos.

30 Otras realizaciones de acuerdo con la invención están relacionadas con una representación de audio codificada que representa una información de audio.

35 Algunas realizaciones de acuerdo con la invención están relacionadas con una ampliación de ancho de banda de audio genérico con coeficiente de información secundaria que se adapta a la señal para la codificación de audio de muy baja tasa de bits.

Antecedentes de la invención

40 En los últimos años, se ha desarrollado una creciente demanda de una codificación y decodificación de contenidos de audio. Si bien las tasas de bits y capacidades de almacenamiento disponibles para la transmisión y almacenamiento de contenidos de audio codificados han aumentado considerablemente, todavía existe una demanda de una codificación, transmisión, almacenamiento y decodificación, eficientes en tasas de bits, para contenidos de audio a una calidad razonable, especialmente de señales de habla en escenarios de comunicación.

45 Los sistemas de codificación del habla modernos son capaces de codificar contenido de audio digital de banda ancha (WB), es decir, señales con frecuencias de hasta 7-8 kHz, con tasas de bits bajas, como de 6 kbps. Los ejemplos más ampliamente expuestos son las recomendaciones de ITU-T G.722.2 (consúltese, por ejemplo, la referencia [1]), como también el G.718, más recientemente desarrollado (consúltese, por ejemplo, las referencias [4] y [10]) y el códec de habla y audio unificado de MPEG xHE-AAC (consúltese, por ejemplo, la referencia [8]). Tanto el G.722.2, también conocido como AMR-WB, como el G.718 emplean técnicas de ampliación de ancho de banda (BWE) de entre 6,4 y 7 kHz para permitir que el codificador de núcleo ACELP subyacente se "centre" en las frecuencias más bajas perceptivamente más relevantes (en particular aquellas a las que el sistema auditivo humano es sensible a la fase), y de este modo logran una calidad suficiente, especialmente con tasas de bits muy bajas. En el xHE-AAC, se utiliza la replicación de banda espectral mejorada (eSBR) para la ampliación de ancho de banda (BWE). En términos generales, el proceso de la ampliación del ancho de banda puede dividirse en dos enfoques conceptuales:

- BWE "a ciegas" o "artificial", en el que los componentes de alta frecuencia (HF) son componentes se reconstruyen a partir de la señal de codificador de núcleo de baja frecuencia (LF) solamente, es decir, sin necesidad de requerir información secundaria transmitida desde el codificador. Este esquema se utiliza por AMR-WB y G.718 a 16 kbps y menos, así como también por algunos sistemas de post-procesamiento de ampliación de ancho de banda compatibles hacia atrás que operan en el habla telefónica de banda estrecha tradicional (consúltese, por ejemplo, las referencias [5] y [9]).
- BWE "guiado", que difiere de la ampliación del ancho de banda a ciegas por el hecho de que algunos de los parámetros utilizados para la reconstrucción del contenido de alta frecuencia (HF) se transmiten al decodificador

como información secundaria en lugar de estimarse a partir de la señal de núcleo decodificada. Este enfoque se utiliza por AMR-WB, G.718, xHE-AAC, así como también por algunos otros códecs (consúltese, por ejemplo, las referencias [2], [7] y [11]) que utilizan este enfoque, pero no con tasas de bit muy bajas.

5 El artículo "A scalable bandwidth extension algorithm" de V. Berisha y A. Spanias describe que las técnicas de ampliación de ancho de banda más modernas predicen la banda de alta frecuencia basada en características extraídas de la banda inferior. Se describe que, aunque esto funciona para algunas señales, surgen problemas cuando la correlación entre la banda baja y alta es insuficiente. Se describe que, en estas situaciones, la información de banda alta debe enviarse a un decodificador. El artículo propone un método de codificación de habla escalable
10 basado en los principios de la ampliación de ancho de banda. La selección de tasa está basada en criterios psicoacústicos explícitos, mientras que se realiza la ampliación de ancho de banda usando una técnica de estimación de MMSE restringida.

15 El artículo "Bandwidth Extension of Audio Based on Partial Loudness Criteria" de V. Berisha y A. Spanias describe un algoritmo de ampliación de ancho de banda de fuente-filtro que hace uso de conceptos psicoacústicos para determinar los beneficios perceptuales que una trama de audio particular gana de una representación más exacta de la banda alta.

20 Sin embargo, se ha encontrado que es difícil proporcionar una ampliación de ancho de banda adecuada con bajas tasas de bits para proporcionar una calidad suficientemente buena en la reconstrucción del contenido de audio.

Por lo tanto, hay una necesidad de un concepto de ampliación de ancho de banda que aporte un compromiso mejorado entre tasa de bits y calidad de audio.

25 **Sumario de la invención**

Una realización de acuerdo con la invención crea un codificador de audio para proporcionar una información de audio codificada basándose en una información de audio de entrada. El codificador de audio comprende un
30 codificador de baja frecuencia configurado para codificar una porción de baja frecuencia de la información de audio de entrada para obtener una representación codificada de la porción de baja frecuencia. El codificador de audio comprende también un proveedor de información de ampliación de ancho de banda configurado para proporcionar información de ampliación de ancho de banda basándose en la información de audio de entrada. El codificador de audio está configurado para incluir selectivamente información de ampliación de ancho de banda en la información de audio codificada de una manera que se adapta a la señal.

35 Esta realización de acuerdo con la invención se basa en el hallazgo de que, para algunos tipos de contenido de audio, e incluso para algunas porciones de una pieza contigua de contenido de audio, puede lograrse una ampliación de ancho de banda de buena calidad basándose en la representación codificada de la porción de baja frecuencia sin ninguna información secundaria sobre la ampliación de ancho de banda, o con solamente una
40 pequeña cantidad de información secundaria sobre el ancho de banda (por ejemplo, un pequeño número de parámetros de ampliación de ancho de banda, que están incluidos en la información de audio codificada). Sin embargo, el concepto también se basa en el hallazgo de que, para otros tipos de contenido de audio, e incluso para otras porciones de una pieza contigua de contenido de audio, puede ser necesario (o al menos muy deseable) incluir una información secundaria de la ampliación de ancho de banda (por ejemplo, parámetros de ampliación de ancho de banda especializados), o una mayor cantidad de información secundaria de la ampliación de ancho de banda (por
45 ejemplo, cuando se compara con el caso mencionado anteriormente) dentro de la información de audio codificada, porque de lo contrario una ampliación del ancho de banda del lado del decodificador no proporciona una calidad de audio satisfactoria.

50 Mediante la inclusión selectiva de información de ampliación de ancho de banda en la información de audio codificada (por ejemplo, variando selectivamente una cantidad de información de ampliación de ancho de banda o de parámetros de ampliación de ancho de banda incluidos en la información de audio codificada, o mediante la conmutación selectiva entre una inclusión de información de ampliación de ancho de banda en la información de audio codificada y una omisión de dicha inclusión de información de ampliación de ancho de banda en la información
55 de audio codificada), se puede evitar que una información "innecesaria" de la ampliación de ancho de banda consuma valiosas tasas de bits en el caso en que una ampliación de ancho de banda en el lado del decodificador realmente no requiera información de la ampliación de ancho de banda, y no obstante, se pueda asegurar que la información de la ampliación del ancho de banda (o un aumento de la cantidad de información de ampliación del ancho de banda) se incluya en la información de audio codificada si la información de la ampliación de ancho de banda se requiere realmente para una ampliación del ancho de banda en el lado del decodificador, es decir, para
60 una reconstrucción, en el lado del decodificador, del contenido de audio.

Por lo tanto, mediante la inclusión selectiva de información de la ampliación de ancho de banda en la información de audio codificada de una manera que se adapta a la señal, es decir, cuando la información de ampliación de ancho

de banda es realmente necesaria para alcanzar una calidad suficientemente buena de una representación de la señal de audio decodificada, es posible reducir la tasa de bits promedio mientras aún se mantiene la posibilidad de obtener una buena calidad de audio.

- 5 En otras palabras, el codificador de audio puede, por ejemplo, conmutar entre una provisión de una información de ampliación de ancho de banda, lo que permite una ampliación de ancho de banda guiada por parámetros en el lado de un decodificador de audio, y una omisión de la provisión de la información de ampliación de ancho de banda, lo que requiere el uso de una ampliación de ancho de banda a ciegas en el lado de un decodificador de audio.
- 10 Por lo tanto, mediante el concepto anteriormente descrito es posible lograr un compromiso particularmente bueno entre tasa de bits y calidad de audio.

15 En una realización preferida, el codificador de audio comprende un detector configurado para identificar porciones de la información de audio de entrada que no pueden decodificarse con una calidad suficiente o deseada (por ejemplo, en términos de una medida de calidad predeterminado) basándose en la representación codificada de la porción de baja frecuencia, y mediante el uso de una ampliación de ancho de banda a ciegas. En este caso, el codificador de audio está configurado para incluir selectivamente información de ampliación de ancho de banda en la información de audio codificada para porciones de la información de audio de entrada identificadas por el detector. Mediante la determinación, o la estimación (por ejemplo, basándose en rasgos de la información de audio de entrada, o basándose en una reconstrucción parcial o completa de la información de audio en el lado del codificador de audio), qué porciones de la información de audio de entrada no pueden decodificarse con una calidad suficiente (o deseada) basándose en la representación codificada de la porción de baja frecuencia, y mediante el uso de una ampliación de ancho de banda a ciegas, se obtiene un criterio significativo para decidir si se debe incluir o no información de ampliación de ancho de banda en la información de audio codificada para porciones (por ejemplo, tramas) de la información de audio de entrada (o de forma equivalente, para tramas o porciones de la información de audio codificada). En otras palabras, el criterio anteriormente mencionado, que se evalúa por el detector, permite un buen compromiso entre la impresión de audición, que puede lograrse mediante la decodificación de la información de audio codificada, y la tasa de bits de la información de audio codificada

30 En una realización preferida, el codificador de audio comprende un detector configurado para identificar porciones de la información de audio de entrada para las que los parámetros de ampliación de ancho de banda no se pueden estimar basándose en la porción de baja frecuencia con una exactitud suficiente o deseada. En este caso, el codificador de audio está configurado para incluir selectivamente información de ampliación de ancho de banda en la información de audio codificada para porciones de la información de audio de entrada identificadas por el detector.

35 Esta realización de acuerdo con la invención se basa en el hallazgo de que una determinación en cuanto a si los parámetros de ampliación de ancho de banda se pueden estimar basándose en una parte de baja frecuencia con la exactitud suficiente o deseada o no constituye un criterio que puede evaluarse con un esfuerzo computacional moderado, y que sin embargo constituye un buen criterio para decidir si se debe incluir o no información de la ampliación de ancho de banda en la información de audio codificada.

40 En una realización preferida, el codificador de audio comprende un detector configurado para identificar porciones de la información de audio de entrada en función de si las porciones son porciones temporalmente estacionarias y en función de si las porciones tienen un carácter de paso bajo. Por otra parte, el codificador de audio está configurado para omitir selectivamente una inclusión de información de ampliación de ancho de banda en la información de audio codificada para porciones de la información de audio de entrada identificadas por el detector como porciones temporalmente estacionarias que tiene un carácter de paso bajo.

45 Esta realización de acuerdo con la invención se basa en el hallazgo de que típicamente no es necesario incluir información de ampliación de ancho de banda en la información de audio codificada para porciones de la información de audio de entrada que son temporalmente estacionarias y que comprenden un carácter de paso bajo, ya que una ampliación de ancho de banda a ciegas (que no se basa en información de ampliación de ancho de banda ni en parámetros de la corriente de bits) típicamente permite una reconstrucción suficientemente buena de tales porciones de la señal. En consecuencia, hay un criterio que puede evaluarse, de una manera computacionalmente eficiente, y que sin embargo permite obtener buenos resultados (en términos de un compromiso entre la tasa de bits y la calidad de audio).

50 En una realización preferida, el detector está configurado para identificar porciones de la información de audio de entrada en función de si las porciones comprenden habla vocalizada, y/o en función de si las porciones comprenden ruido ambiental (por ejemplo, de un automóvil), y/o en función de si las porciones comprenden música sin instrumentos de percusión. Se ha encontrado que tales porciones, que comprenden habla vocalizada, o que comprenden ruido ambiental, o que comprenden música sin instrumentación de percusión, por lo general pueden reconstruirse utilizando una ampliación de ancho de banda a ciegas con suficiente calidad de audio, de tal manera que es recomendable omitir la inclusión de información de ampliación de ancho de banda en la información de audio codificada para tales porciones.

En una realización preferida, el codificador de audio comprende un detector configurado para identificar porciones de la información de audio de entrada en función de si una diferencia entre una envolvente espectral de una porción de baja frecuencia y una envolvente espectral de una porción de alta frecuencia es mayor que o igual a una medida de diferencia predeterminada. En este caso, el codificador de audio está configurado para incluir selectivamente información de ampliación de ancho de banda en la información de audio codificada para porciones de la información de audio de entrada identificadas por el detector.

Se ha encontrado que porciones de la información de audio de entrada, que comprenden una gran diferencia entre una envolvente espectral de una porción de baja frecuencia y una envolvente espectral de una porción de alta frecuencia, típicamente no pueden reconstruirse bien usando una ampliación de ancho de banda a ciegas, dado que una ampliación de ancho de banda a ciegas a menudo proporciona envolventes espectrales similares en la porción de alta frecuencia (es decir, en la señal de ampliación de ancho de banda) en comparación con la respectiva porción de baja frecuencia. Por lo tanto, se ha encontrado que una evaluación de la diferencia entre la envolvente espectral de la porción de baja frecuencia y la envolvente espectral de la porción de alta frecuencia constituye un buen criterio para decidir si se debe incluir o no información de ampliación de ancho de banda en la información de audio codificada.

En una realización preferida, el detector está configurado para identificar porciones de la información de audio de entrada en función de si las porciones comprenden habla no vocalizada, y/o en función de si las porciones comprenden sonidos de percusión. Se ha encontrado que las porciones que comprenden habla no vocalizada y porciones que comprenden sonidos de percusión típicamente comprenden espectros en los que la envolvente espectral de la porción de baja frecuencia difiere sustancialmente de la envolvente espectral de la porción de alta frecuencia. Por lo tanto, la detección de habla no vocalizada y/o de sonidos de percusión ha demostrado ser un buen criterio para decidir si se debe incluir o no información de ampliación de ancho de banda en la información de audio codificada.

En una realización preferida, el codificador de audio comprende un detector configurado para determinar una inclinación espectral de porciones de la información de audio de entrada, y para identificar porciones de la información de audio de entrada en función de si la inclinación espectral determinada es mayor que o igual a un valor umbral de inclinación variable. En este caso, el codificador de audio está configurado para incluir selectivamente información de ampliación de ancho de banda en la información de audio codificada para porciones de la información de audio de entrada identificadas por el detector. Se ha encontrado que es posible derivar una inclinación espectral con un esfuerzo computacional moderado y seguir proporcionando un buen criterio para la decisión de si se debe incluir o no información de ampliación de ancho de banda en la información de audio codificada. Por ejemplo, si la inclinación espectral alcanza o supera un valor de umbral de inclinación, se puede concluir que el espectro tiene un carácter de paso alto y que no puede reconstruirse bien mediante la ampliación de ancho de banda a ciegas. En particular, la ampliación de ancho de banda a ciegas típicamente no puede reconstruir espectros que comprenden una inclinación positiva (en la que se hace énfasis en una porción de alta frecuencia con respecto a una porción de baja frecuencia) con una buena exactitud. Además, dado que una porción de alta frecuencia es de particular relevancia perceptiva en el caso de una inclinación espectral positiva, es recomendable en tales casos incluir la información de ampliación de ancho de banda en la representación de audio codificada.

En una realización preferida, el detector está configurado además para determinar un coeficiente de cruce en cero de porciones de la información de audio de entrada, y para identificar porciones de la información de audio de entrada también en función de si el coeficiente de cruce en cero determinado es mayor que o igual a un valor fijo o valor umbral variable del coeficiente de cruce en cero. Se ha encontrado que el coeficiente del cruce en cero es también un buen criterio para detectar porciones de la información de audio de entrada que no pueden reconstruirse bien usando una ampliación de ancho de banda a ciegas, de tal manera que tiene sentido (en términos de lograr un buen compromiso entre la tasa de bits y la calidad del audio) incluir la información de ampliación de ancho de banda en la información de audio codificada.

En una realización preferida, el detector está configurado para aplicar una histéresis para la identificación de porciones de la información de audio de entrada, para reducir un número de transiciones entre porciones de señal identificadas (para las que la información de ampliación de ancho de banda está incluida en la representación de audio codificada) y porciones de señal no identificadas (para las que la información de ampliación de ancho de banda no está incluida en la representación de audio codificada). Se ha encontrado que es ventajoso para evitar una conmutación excesiva entre una inclusión de información de ampliación de ancho de banda en la información de audio codificada y una omisión de la inclusión de la información de ampliación de ancho de banda en la representación de audio codificada, ya que tales transiciones pueden presentarse junto con algunos artefactos, en particular si el número de transiciones es muy alto. En consecuencia, el uso de una histéresis, que puede aplicarse, por ejemplo, al valor umbral de inclinación (que es entonces un valor de umbral de inclinación variable) o al valor umbral del coeficiente de cruce en cero (que es entonces un valor umbral variable del cruce en cero), permite lograr este objetivo.

- 5 En una realización preferida, el codificador de audio está configurado para incluir selectivamente parámetros que representan una envolvente espectral de una porción de alta frecuencia de la información de audio de entrada en la información de audio codificada de una manera que se adapta a la señal. Esta realización se basa en la idea de que los parámetros que representan la envolvente espectral de la porción de alta frecuencia son particularmente importantes en una ampliación de ancho de banda guiada por parámetros, de tal manera que la inclusión de dichos parámetros que representan la envolvente espectral de la porción de alta frecuencia de la información de audio de entrada permite lograr una ampliación de ancho de banda de buena calidad sin causar una alta tasa de bits.
- 10 En una realización preferida, el codificador de baja frecuencia está configurado para codificar una porción de baja frecuencia de la información de audio de entrada que comprende frecuencias hasta una frecuencia máxima que se encuentra en un intervalo de entre 6 kHz y 7 kHz. Por otra parte, el codificador de audio está configurado para incluir selectivamente en la representación de audio codificada entre tres y cinco parámetros que describen las intensidades de porciones o subporciones de señal de alta frecuencia (por ejemplo, porciones de señal que tienen frecuencias por encima de aproximadamente 6 a 7 kHz) que tienen anchos de banda entre 300 Hz y 500 Hz. Se ha encontrado que un concepto de este tipo tiene como resultado una buena calidad de audio sin comprometer sustancialmente un esfuerzo de tasa de bits.
- 15 En una realización preferida, el codificador de audio está configurado para incluir selectivamente en la representación de audio codificada 3 - 5 parámetros cuantificados escalarmen- te que describen intensidades de cuatro porciones (o subporciones) de la señal de alta frecuencia, en el que las porciones (o subporciones) de la señal de alta frecuencia cubren rangos de frecuencia por encima de la parte de baja frecuencia. Se ha encontrado que el uso de 3 - 5 parámetros cuantificados escalarmen- te que describen intensidades de cuatro porciones de la señal de alta frecuencia es típicamente suficiente para lograr una ampliación del ancho de banda guiada por parámetro que supera una calidad de audio relativamente baja obtenible mediante una ampliación de ancho de banda a ciegas en la misma porción de señal. Por lo tanto, no hay grandes diferencias de calidad entre las porciones de señal de audio reconstruidas, independientemente de si las porciones de la señal de audio reconstruidas se reconstruyen utilizando una ampliación de ancho de banda a ciegas o una ampliación de ancho de banda guiada. Por lo tanto, el concepto anteriormente mencionado está bien adaptado al concepto que permite una conmutación entre una ampliación de ancho de banda a ciegas y una ampliación de ancho de banda guiada por parámetros.
- 20 En una realización preferida, el codificador de audio está configurado para incluir selectivamente en la representación de audio codificada una pluralidad de parámetros que describen una relación entre las energías de porciones de frecuencia espectralmente adyacentes, en el que uno de los parámetros describe una relación entre una energía de una primera porción de alta frecuencia de ampliación de ancho de banda y una porción de baja frecuencia, y en el que otro de los parámetros describe relaciones entre energías de (pares de) otras porciones de alta frecuencia de ampliación de ancho de banda. Se ha encontrado que un concepto de este tipo que describe relaciones (o diferencias) entre las energías (o, lo que es equivalente, las intensidades) de diferentes porciones de frecuencia (preferentemente adyacentes) permite una codificación eficiente de la información de ampliación de ancho de banda. También se ha encontrado que los parámetros de este tipo que describen una relación entre las energías de porciones de frecuencia espectralmente adyacentes típicamente pueden cuantificarse con solamente una pequeña cantidad de bits sin comprometer sustancialmente una calidad de audio alcanzable por una ampliación de ancho de banda.
- 25 Otra realización de acuerdo con la invención crea un decodificador de audio para proporcionar una información de audio decodificada basándose en una información de audio codificada. El decodificador de audio comprende un decodificador de baja frecuencia configurado para decodificar una representación codificada de una porción de baja frecuencia (de un contenido de audio), para obtener una representación decodificada de la porción de baja frecuencia. El decodificador de audio también comprende una ampliación de ancho de banda configurada para obtener una señal de la ampliación de ancho de banda mediante una ampliación de ancho de banda a ciegas para porciones de un contenido de audio para las que no se incluyen parámetros de ampliación de ancho de banda en la información de audio codificada, y para obtener la señal de la ampliación de ancho de banda usando una ampliación de ancho de banda guiada por parámetros para porciones del contenido de audio para las que se incluyen parámetros de ampliación de ancho de banda en la información de audio codificada.
- 30 Este codificador de audio se basa en la idea de que es posible lograr un buen compromiso entre la calidad de audio y la tasa de bits si es posible conmutar entre una ampliación de ancho de banda a ciegas y una ampliación de ancho de banda guiada por parámetros incluso dentro de una pieza contigua de contenido de audio, ya que se ha encontrado que muchas piezas típicas de contenido de audio comprenden ambas secciones para las que se puede obtener una buena calidad de audio utilizando una ampliación de ancho de banda a ciegas y secciones para las que se requiere una ampliación de ancho de banda guiada por parámetros para alcanzar la suficiente calidad de audio. Por otra parte, debería ser evidente que las mismas consideraciones explicadas anteriormente con respecto al codificador de audio también se aplican al decodificador de audio.
- 35
40
45
50
55
60

En una realización preferida, el decodificador de audio está configurado para decidir si obtener la señal de ampliación de ancho de banda utilizando una ampliación de ancho de banda a ciegas o utilizando una ampliación de ancho de banda guiada por parámetros en una base trama a trama. Se ha encontrado que dicha conmutación de buena precisión (trama a trama) entre una ampliación de ancho de banda a ciegas y una ampliación de ancho de banda guiada por parámetros ayuda a mantener la tasa de bits razonablemente baja, incluso si regularmente hay algunas tramas en las que se requiere una ampliación de ancho de banda guiada por parámetros para evitar una degradación excesiva del contenido de audio.

En una realización preferida, el decodificador de audio está configurado para conmutar entre un uso de una ampliación de ancho de banda a ciegas y una ampliación de ancho de banda guiada por parámetros dentro de una pieza contigua de contenido de audio. Esta realización se basa en el hallazgo de que incluso una sola pieza (contigua) de contenido de audio a menudo comprende pasajes (o porciones, o tramas) de diferentes tipos, algunos de los cuales deberían codificarse (y, en consecuencia, decodificarse) utilizando una ampliación de ancho de banda guiada por parámetros, mientras que otros pasajes o tramas pueden decodificarse usando una ampliación de ancho de banda a ciegas sin una degradación sustancial de la calidad de audio.

En una realización preferida, el decodificador de audio está configurado para evaluar banderas incluidas en la información de audio codificada para diferentes porciones (por ejemplo, tramas) del contenido de audio, para decidir si se debe utilizar una ampliación de ancho de banda a ciegas o una ampliación de ancho de banda guiada por parámetros (por ejemplo, para la trama a la que está asociada la bandera). En consecuencia, la decisión de si debería utilizarse una ampliación de ancho de banda a ciegas o una ampliación de ancho de banda guiada por parámetros, se mantiene sencilla, y el decodificador de audio no necesita tener una inteligencia sustancial para decidir si se debe utilizar una ampliación de ancho de banda a ciegas o una ampliación de ancho de banda guiada por parámetros.

Sin embargo, en otra realización preferida, el decodificador de audio está configurado para decidir si utilizar una ampliación de ancho de banda a ciegas o una ampliación de ancho de banda guiada por parámetros basándose en la representación codificada de la porción de baja frecuencia sin evaluar una bandera de señalización del modo de ampliación de ancho de banda. Por lo tanto, por el hecho de proporcionar inteligencia en el decodificador de audio, es posible omitir una bandera de señalización del modo de ampliación de ancho de banda, lo que reduce la tasa de bits.

En una realización preferida, el decodificador de audio está configurado para decidir si debe utilizar una ampliación de ancho de banda a ciegas o una ampliación de ancho de banda guiada por parámetros basándose en uno o más rasgos de la representación decodificada de la porción de baja frecuencia (del contenido de audio). Se ha encontrado que los rasgos de la representación decodificada de la porción de baja frecuencia constituyen cantidades que pueden utilizarse, con buena exactitud, para decidir si se debe utilizar una ampliación de ancho de banda a ciegas o una ampliación de ancho de banda guiada por parámetros. Esto es particularmente cierto si se utilizan los mismos rasgos en el lado de un codificador de audio. Por consiguiente, ya no es necesario evaluar una bandera de señalización de modo de ampliación de ancho de banda, lo que a su vez permite una reducción de la tasa de bits, ya que no es necesario incluir una bandera de señalización de modo de ampliación de ancho de banda en la representación de audio codificada en el lado de un codificador de audio.

En una realización preferida, el decodificador de audio está configurado para decidir si se debe utilizar una ampliación de ancho de banda a ciegas o una ampliación de ancho de banda guiada por parámetros, basándose en coeficientes de predicción lineales cuantificados y/o estadísticas del dominio del tiempo de la representación decodificada de la porción de baja frecuencia (del contenido de audio). Se ha encontrado que los coeficientes de predicción lineal cuantificados son fáciles de obtener en el lado de un decodificador de audio, y por el hecho de derivar una inclinación espectral, pueden por lo tanto servir como una buena indicación de si se debe utilizar una ampliación de ancho de banda a ciegas o una ampliación de ancho de banda guiada por parámetros. Por otra parte, los coeficientes de predicción lineal cuantificados también son fácilmente accesibles en el lado de un codificador de audio, de tal manera que es fácilmente posible coordinar una conmutación entre una ampliación de ancho de banda a ciegas y una ampliación de ancho de banda guiada por parámetros en el lado de un codificador de audio y en el lado de un decodificador de audio. De manera similar, las estadísticas de dominio de tiempo de la representación decodificada de la porción de baja frecuencia, tales como un coeficiente de cruce en cero, han demostrado ser una cantidad fiable para decidir si usar una ampliación de ancho de banda a ciegas o una ampliación de ancho de banda guiada por parámetros en el lado de un decodificador de audio.

En una realización preferida, la ampliación de ancho de banda está configurada para obtener la señal de ampliación de ancho de banda usando uno o más rasgos de la representación decodificada de la porción de baja frecuencia y/o usando uno o más parámetros del decodificador de baja frecuencia para porciones temporales de la información (o contenido) de audio de entrada para la que no hay parámetros de ampliación de ancho de banda incluidos en la información de audio codificada. Se ha encontrado que una ampliación de ancho de banda a ciegas de este tipo tiene como resultado una buena calidad de audio.

5 En una realización preferida, la ampliación de ancho de banda está configurada para obtener la señal de ampliación de ancho de banda utilizando una información de centroide espectral y/o usando una información de energía y/o usando una información de inclinación (espectral) y/o usando coeficientes de filtro codificados por porciones temporales de la información (o contenido) de audio de entrada para la que no hay parámetros de ampliación de ancho de banda incluidos en la información de audio codificada. Se ha encontrado que el uso de estas cantidades permite obtener una forma eficaz de obtener una ampliación de ancho de banda de buena calidad.

10 En una realización preferida, la ampliación de ancho de banda está configurada para obtener la señal de ampliación de ancho de banda utilizando parámetros de la corriente de bits que describen una envolvente espectral de una porción de alta frecuencia para porciones temporales del contenido de audio para las que hay parámetros de ampliación de ancho de banda incluidos en la información de audio codificada. Se ha encontrado que el uso de parámetros de corriente de bits que describen una envolvente espectral de la porción de alta frecuencia permite obtener una ampliación de ancho de banda guiada por parámetros con una tasa de bits eficiente con una buena calidad, en el que los parámetros de la corriente de bits que describen la envolvente espectral típicamente no requieren una elevada tasa de bits, sino que pueden codificarse con solamente un número comparativamente pequeño de bits por trama de audio. En consecuencia, incluso la conmutación hacia la ampliación de ancho de banda guiada por parámetros no da lugar a un aumento sustancial de la tasa de bits.

20 En una realización preferida, la ampliación de ancho de banda está configurada para evaluar entre tres y cinco parámetros de corriente de bits que describen intensidades de porciones de señal de alta frecuencia que tienen anchos de banda de entre 300 Hz y 500 Hz para obtener la señal de ampliación de ancho de banda. Se ha encontrado que un número comparativamente pequeño de parámetros de corriente de bits es suficiente para obtener una ampliación de ancho de banda en un intervalo perceptivamente importante, de tal manera que es posible obtener una buena calidad de audio con un pequeño aumento en la tasa de bits.

30 En una realización preferida, entre tres y cinco parámetros de corriente de bits que describen intensidades de porciones de señal de alta frecuencia que tienen anchos de banda de entre 300 Hz y 500 Hz se cuantifican escalaramente con una resolución de 2 o 3 bits de tal manera que hay entre 6 y 15 bits de parámetros de conformación espectral de ampliación de ancho de banda por cada trama de audio. Se ha encontrado que una opción de este tipo permite una eficiencia muy alta de la tasa de bits de la ampliación de ancho de banda guiada por parámetros, mientras que una calidad de la ampliación de ancho de banda es típicamente comparable con la calidad de ampliación de ancho de banda obtenible usando la ampliación de ancho de banda a ciegas para porciones "no críticas" del contenido de audio, en que la ampliación de ancho de banda a ciegas ofrece buenos resultados. En consecuencia, hay una calidad equilibrada tanto en el caso de que se aplique la ampliación de ancho de banda a ciegas como en el caso en que se aplica la ampliación de ancho de banda guiada por parámetros.

40 En una realización preferida, la ampliación de ancho de banda está configurada para llevar a cabo una suavización de las energías de la señal de ampliación de ancho de banda cuando se conmuta desde la ampliación de ancho de banda a ciegas a la ampliación de ancho de banda guiada por parámetros y/o cuando se conmuta desde la ampliación de ancho de banda guiada por parámetros a la ampliación de ancho de banda a ciegas. En consecuencia, es posible evitar los clics o "artefactos de bloqueo" que podrían causarse por las diferentes propiedades de la ampliación de ancho de banda a ciegas y la ampliación de ancho de banda guiada por parámetros.

45 En una realización preferida, la ampliación de ancho de banda está configurada para amortiguar una porción de alta frecuencia de la señal de ampliación de ancho de banda para una porción del contenido de audio al que se aplica una ampliación de ancho de banda guiada por parámetros después de una porción del contenido de audio al que se aplica una ampliación de ancho de banda a ciegas. Por otra parte, la ampliación de ancho de banda está configurada para reducir una amortiguación para una porción de alta frecuencia de la señal de ampliación de ancho de banda para una porción del contenido de audio al que se aplica una ampliación de ancho de banda a ciegas después de una porción del contenido de audio al que se aplica una ampliación de ancho de banda guiada por parámetros. Por lo tanto, el efecto de que la ampliación de ancho de banda a ciegas muestra típicamente una característica de paso bajo, si bien éste no es necesariamente el caso para la ampliación de ancho de banda guiada por parámetros, puede compensarse hasta cierto punto. En consecuencia, se reducen los artefactos en las transiciones entre porciones del contenido de audio decodificado utilizando una ampliación de ancho de banda a ciegas y el uso de una ampliación de ancho de banda guiada por parámetros.

60 Otra realización de acuerdo con la invención crea un método para proporcionar una información de audio codificada basándose en una información de audio de entrada. El método comprende codificar una porción de baja frecuencia de la información de audio de entrada para obtener una representación codificada de la porción de baja frecuencia. El método también comprende proporcionar información de ampliación de ancho de banda basándose en la información de audio de entrada. La información de ampliación de ancho de banda se incluye selectivamente en la información de audio codificada de una manera que se adapta a la señal. Este método se basa en las mismas

consideraciones que el codificador de audio descrito anteriormente.

5 Otra realización de acuerdo con la invención crea un método para proporcionar una información de audio decodificada basándose en una información de audio codificada. El método comprende la decodificación de una representación codificada de una porción de baja frecuencia para obtener una representación descodificada de la porción de baja frecuencia. El método comprende además obtener una señal de ampliación de ancho de banda utilizando una ampliación de ancho de banda a ciegas para porciones de un contenido de audio para las que no se incluyen parámetros de ampliación de ancho de banda en la información de audio codificada. El método comprende además la obtención de la señal de ampliación de ancho de banda utilizando una ampliación de ancho de banda guiada por parámetros para porciones de contenido de audio para las que se incluyen parámetros de ampliación de ancho de banda en la información de audio codificada. Este método se basa en las mismas consideraciones que el decodificador de audio descrito anteriormente.

15 Otra realización de acuerdo con la invención crea un programa informático para llevar a cabo uno de los métodos mencionados anteriormente cuando el programa informático se ejecuta en una computadora.

20 Otra realización de acuerdo con la invención crea una representación de audio codificada que representa una información de audio. La representación de audio codificada comprende una representación codificada de una porción de baja frecuencia de una información de audio y una información de ampliación de ancho de banda. La información de ampliación de ancho de banda está incluida en la representación de audio codificada de una manera que se adapta a la señal, en algunas porciones de la información de audio, pero no en todas ellas. Esta información de audio codificada se proporciona por el codificador de audio descrito anteriormente, y puede evaluarse por el decodificador de audio descrito anteriormente.

25 **Breve descripción de las figuras**

A continuación se describen realizaciones de acuerdo con la invención haciendo referencia a las figuras adjuntas, en las que:

30 La Figura 1 muestra un diagrama esquemático de bloques de un codificador de audio, de acuerdo con una realización de la presente invención;

La Figura 2 muestra un diagrama esquemático de bloques de un codificador de audio, de acuerdo con otra realización de la presente invención;

35 La Figura 3 muestra una representación gráfica de porciones de frecuencia y la información de audio codificada asociada con las mismas;

La Figura 4 muestra un diagrama esquemático de bloques de un decodificador de audio, de acuerdo con una realización de la presente invención;

La Figura 5 muestra un diagrama esquemático de bloques de un decodificador de audio, de acuerdo con otra realización de la presente invención;

45 La Figura 6 muestra un diagrama de flujo de un método para proporcionar una representación de audio codificada, de acuerdo con una realización de la presente invención;

La Figura 7 muestra un diagrama de flujo de un método para proporcionar una representación de audio decodificada, de acuerdo con una realización de la presente invención;

50 La Figura 8 muestra una ilustración esquemática de una representación de audio codificada, de acuerdo con una realización de la presente invención.

55 **Descripción detallada de las realizaciones**

1. Codificador de audio de acuerdo con la Figura 1

60 La Figura 1 muestra un diagrama esquemático de bloques de un codificador de audio, de acuerdo con una realización de la presente invención.

El codificador de audio 100 de acuerdo con la Figura 1 recibe una información de audio de entrada 110 y proporciona, basándose en ésta, una información de audio codificada 112. El codificador de audio 100 comprende un codificador de baja frecuencia 120, que está configurado para codificar una porción de una frecuencia de la información de audio de entrada 110, para obtener una representación codificada 122 de la porción de baja

frecuencia. El codificador de audio 100 comprende también un proveedor de información de ampliación de ancho de banda 130 configurado para proporcionar información de ampliación de ancho de banda 132 basándose en la información de audio de entrada 110. El codificador de audio 100 está configurado para incluir selectivamente información de ampliación de ancho de banda 132 en la información de audio codificada 112, de una manera que se adapta a la señal.

En cuanto a la funcionalidad del codificador de audio 100, se puede decir que el codificador de audio 100 proporciona una codificación eficiente de la tasa de bits de la información de audio de entrada 110. Una porción de baja frecuencia, por ejemplo en un intervalo de frecuencias de aproximadamente 6 o 7 kHz, se codifica mediante el codificador de baja frecuencia 120, en el que es posible utilizar cualquiera de los conceptos de codificación de audio conocidos. Por ejemplo, el codificador de baja frecuencia 120 puede ser un codificador "de audio en general" (como, por ejemplo, un codificador de audio AAC) o un codificador de audio de tipo de voz (como, por ejemplo, un codificador de audio basado en predicción lineal, un codificador de audio CELP, un codificador de audio ACELP, o similares). Por consiguiente, la porción de baja frecuencia de la información de audio de entrada se codifica utilizando cualquiera de los conceptos convencionales. Sin embargo, la tasa de bits de la representación codificada 122 de la porción de baja frecuencia se mantiene razonablemente pequeña, ya que sólo se codifican los componentes de frecuencia de hasta aproximadamente 6 a 7 kHz. Por otra parte, el codificador de audio 100 es capaz de proporcionar una información de ampliación de ancho de banda, por ejemplo, en forma de parámetros de ampliación de ancho de banda que describen una porción de alta frecuencia de la información de audio de entrada 110, como, por ejemplo, una región de frecuencia que comprende frecuencias más altas que la región de frecuencia codificada por el codificador de baja frecuencia 120. Por lo tanto, el proveedor de la información de ampliación de ancho de banda 130 es capaz de proporcionar una información secundaria de la información de audio codificada 112, que puede controlar una ampliación de ancho de banda llevada a cabo en el lado de un decodificador de audio que no se muestra en la Figura. 1. La información de ampliación del ancho de banda (o la información secundaria de ampliación de ancho de banda) pueden, por ejemplo, representar una forma espectral (o envolvente espectral) de la porción de alta frecuencia de la información de audio de entrada, es decir, un intervalo de frecuencias de la información de audio de entrada que no se cubre por el codificador de baja frecuencia 120.

Sin embargo, el codificador de audio 100 está configurado para decidir, de una manera que se adapta a la señal, si la información de ampliación de ancho de banda se debería incluir en la información de audio codificada 112. Por lo tanto, el codificador de audio 100 es capaz de solamente incluir la información de ampliación de ancho de banda en la información de audio codificada 112 si se requiere la información de ampliación de ancho de banda (o al menos es deseable) para una reconstrucción de la información de audio en el lado de un decodificador de audio. En este contexto, el codificador de audio también puede controlar si la información de ampliación de ancho de banda 132 se proporciona por el proveedor de información de ampliación de ancho de banda 130 para una porción de la información de audio de entrada (o, lo que es equivalente, para una porción de la información de audio codificada), ya que naturalmente no es necesaria para proporcionar información de ampliación de ancho de banda para una porción de la información de audio de entrada (o de la información de audio codificada) si la información de ampliación de ancho de banda no se incluirá en la información de audio codificada. En consecuencia, el codificador de audio 100 es capaz de mantener la tasa de bits de la información de audio codificada 112 tan pequeña como sea posible, por el hecho de evitar la inclusión de la información de ampliación de ancho de banda 132 en la información de audio codificada 112, si se encuentra, basándose en algún proceso de análisis y/o proceso de decisión llevado a cabo por el codificador de audio 100, que la información de ampliación de ancho de banda no es necesaria para obtener cierta calidad de audio cuando se reconstruya una porción correspondiente del contenido de audio en el lado de un decodificador de audio.

Por lo tanto, el codificador de audio 100 sólo incluye la información de ampliación de ancho de banda en la información de audio codificada si es necesario (para obtener una cierta calidad de audio) en el lado de un decodificador de audio, que, por un lado, ayuda a reducir la tasa de bits de la información de audio codificada 112 y que, por otro lado, asegura que se incluye una adecuada información de ampliación de ancho de banda 132 en la información de audio codificada 112, si esto es necesario para evitar una mala calidad de audio cuando se decodifica la información de audio codificada en el lado de un decodificador de audio. Por lo tanto, se logra un compromiso mejorado entre tasa de bits y la calidad de audio por el codificador de audio 100 en comparación con las soluciones convencionales.

Por ejemplo, el decodificador de audio puede decidir, por cada trama de audio, si la información de ampliación de ancho de banda se debería incluir en la información de audio codificada 112 (o incluso si debería determinarse la información de ampliación de ancho de banda). Sin embargo, como alternativa, el decodificador de audio puede decidir, por "entrada" (por ejemplo, por cada archivo de audio o por cada corriente de audio), si debería incluirse información de ampliación de ancho de banda en la información de audio codificada 112. Para este propósito, la entrada se puede analizar (por ejemplo antes de la codificación), de tal manera que la decisión se toma de una manera que se adapta a la señal.

2. Codificador de audio de acuerdo con la Figura 2

La Figura 2 muestra un diagrama esquemático de bloques de un codificador de audio, de acuerdo con una realización de la presente invención. El codificador de audio 200 recibe una información de audio de entrada 210 y proporciona, basándose en la misma, una información de audio codificada 212. El codificador de audio 200 comprende un codificador de baja frecuencia 220, que puede ser sustancialmente idéntico al codificador de baja frecuencia 120 descrito anteriormente. El codificador de baja frecuencia 220 proporciona una representación codificada 222 de una porción de baja frecuencia de la información de audio de entrada (o, lo que es equivalente, del contenido de audio representado por la información de audio de entrada 210). El codificador de audio 200 comprende también un proveedor de información de ampliación de ancho de banda 230, que puede ser sustancialmente idéntico al proveedor de información de ampliación de ancho de banda 130 descrito anteriormente. El proveedor de información de ampliación de ancho de banda 230 normalmente recibe la información de audio de entrada 210. Sin embargo, el proveedor de información de ampliación de ancho de banda 230 también puede recibir una información de control (o información intermedia) desde el codificador de baja frecuencia 220, en el que dicha información de control (o información intermedia) puede, por ejemplo, comprender información acerca de un espectro (o de una forma espectral o de una envolvente espectral) de la porción de baja frecuencia de la información de audio de entrada 210. Sin embargo, la información de control (o la información intermedia) también puede comprender parámetros de codificación (por ejemplo, coeficientes de filtro LPC, o valores de dominio de transformada, como coeficientes de MDCT, o coeficientes de QMF) o similares. Por otra parte, el proveedor de información de ampliación de ancho de banda 230 puede, opcionalmente, recibir la representación codificada 222 de la porción de baja frecuencia, o al menos una parte de ella. Por otra parte, el codificador de audio 200 comprende un detector 240, que está configurado para decidir si la información de ampliación de ancho de banda está incluida en la información de audio codificada 212 para una porción dada de la información de audio de entrada 210 (o de una porción dada de la información de audio codificada 212). Opcionalmente, el detector 240 también puede determinar si dicha información de ampliación de ancho de banda se determina por el proveedor de la información de ampliación de ancho de banda 230 para dicha porción dada de la información de audio de entrada 210 (o de la información de audio codificada 212). El detector 240 puede, por tanto, recibir la información de audio de entrada 210, y/o información de control y/o información intermedia 224 desde el codificador de baja frecuencia 220 (por ejemplo, como se ha descrito anteriormente) y/o la representación codificada 222 de la porción de baja frecuencia. Por otra parte, el detector 240 está configurado para proporcionar una señal de control 242 que controla una provisión selectiva de la información de ampliación de ancho de banda y/o una inclusión selectiva de la información de ampliación de ancho de banda en la información de audio codificada 212.

En cuanto a la funcionalidad del codificador de audio 200, se hace referencia a las explicaciones anteriores hechas con respecto al codificador de audio 100.

Por otra parte, hay que señalar que el detector 240 comprende un papel central, ya que el detector 240 decide si la información de ampliación de ancho de banda está incluida o no en la información de audio codificada 212, y por lo tanto decide si un decodificador de audio, que recibe la información de audio codificada 212, reconstruye el contenido de audio, que se describe por la información de audio de entrada 210, utilizando una ampliación del ancho de banda a ciegas o utilizando una ampliación de ancho de banda guiada por parámetros (en la que la información de ampliación de ancho de banda representa los parámetros que guían la ampliación de ancho de banda guiada por parámetros).

En términos generales, el detector identifica porciones de la información de audio de entrada que no pueden decodificarse con la calidad suficiente o deseada basándose en la representación codificada 222 de la porción de baja frecuencia utilizando una ampliación de ancho de banda a ciegas. En otras palabras, el detector 240 debería reconocer cuándo la representación codificada de la porción de baja frecuencia 222 por sí sola no permite una ampliación de ancho de banda a ciegas con la calidad suficiente. Dicho con otras palabras, el detector 240 identifica preferentemente porciones de la información de audio de entrada para las que los parámetros ampliación de ancho de banda no se pueden estimar basándose en la porción de baja frecuencia con una exactitud suficiente (o deseada), para obtener una calidad de audio aceptable (o deseada). Por lo tanto, el detector 240 puede determinar, mediante la señal de control 242, que la información de ampliación de ancho de banda debería incluirse en la información de audio codificada para porciones de la información de audio de entrada que no pueden decodificarse con una calidad suficiente o deseada basándose en la representación codificada 222 de la parte de baja frecuencia utilizando una ampliación de ancho de banda a ciegas (es decir, sin recibir ninguna información de ampliación del ancho de banda desde el codificador). De manera equivalente, el detector puede determinar, mediante la señal de control 242, que la información de ampliación de ancho de banda debería incluirse en la información de audio codificada para porciones de la información de audio de entrada para las que los parámetros de ampliación de ancho de banda no pueden estimarse basándose en la parte de baja frecuencia (o, lo que es equivalente, la representación codificada 222 de la porción de baja frecuencia) con una exactitud suficiente o deseada.

Para identificar tales porciones, para las que la información de ampliación de ancho de banda debería estar incluida en la información codificada de audio (o, lo que es equivalente, para identificar porciones de la información de audio de entrada para las que no es necesario incluir la información de ampliación de ancho de banda en la información de

audio codificada 212), el detector 240 puede utilizar diferentes estrategias. Como se mencionó anteriormente, el detector 240 puede recibir diferentes tipos de información de entrada. En algunos casos, la decisión del detector de si la información de ampliación de ancho de banda se debería incluir o no en la información de audio codificada 212 puede basarse o no únicamente en la información de audio de entrada 210. En otras palabras, el detector 240 puede, por ejemplo, estar configurado para analizar la información de audio de entrada 210, para descubrir para cuáles porciones de la información de audio de entrada (que corresponden a porciones de la información de audio codificada 212) es necesario incluir la información de ampliación de ancho de banda de banda 232 en la información de audio codificada 212 para obtener una calidad de audio aceptable (o deseable). Sin embargo, como alternativa, la decisión del detector 240, puede estar basada en alguna información de control o información intermedia 224, proporcionada por el codificador de baja frecuencia 200. Como alternativa, o además, la decisión del detector 240 puede estar basada en la representación codificada 222 de la porción de baja frecuencia de la información de audio de entrada 210. Por lo tanto, el detector puede evaluar diferentes cantidades para determinar (o para estimar) si una ampliación de ancho de banda a ciegas en el lado de un decodificador de audio dará como resultado un audio de una calidad suficiente (o es probable que dé como resultado un audio de una calidad suficiente, o se prevé que dé como resultado en un audio de una calidad suficiente).

Por ejemplo, el detector puede determinar si porciones de la información de audio de entrada 210 son porciones temporalmente estacionarias y si las porciones de la información de audio de entrada 210 tienen un carácter de paso bajo. Por ejemplo, el detector 240 puede llegar a la conclusión de que no es necesario incluir información de ampliación de ancho de banda en la información de audio codificada 212 para aquellas porciones que han demostrado ser porciones temporalmente estacionarias y que tienen un carácter de paso bajo, ya que se ha reconocido que tales porciones de la información de audio de entrada 210 típicamente se pueden reproducir con una calidad suficientemente buena de audio en el lado de un decodificador de audio, incluso utilizando una ampliación de ancho de banda a ciegas. Esto se debe al hecho de que una ampliación de ancho de banda a ciegas típicamente funciona bien para porciones de la información (o contenido) de audio de entrada que no comprenden fuertes cambios del contenido de audio (o que no comprenden ningún transitorio ni otras variaciones fuertes del contenido de audio) y que por lo tanto se pueden considerar como temporalmente estacionarias. Por otra parte, se ha encontrado que la ampliación de ancho de banda a ciegas funciona bien para porciones del contenido de audio que comprenden un carácter de paso bajo, es decir, para una porción del contenido de audio para la que una intensidad de una porción de baja frecuencia es mayor que una intensidad de una porción de alta frecuencia, ya que se trata de un supuesto fundamental de la mayoría de los conceptos de ampliación de ancho de banda a ciegas. Por lo tanto, el detector 240 puede señalar, utilizando la señal de control 242, omitir selectivamente una inclusión de la información de ampliación de ancho de banda en la información de audio codificada 212 para tales porciones temporalmente estacionarias que tienen un carácter de paso bajo.

Por ejemplo el detector 240 puede estar configurado para identificar porciones de la información de audio de entrada que comprenden un habla vocalizada y/o porciones de la información de audio de entrada que comprenden ruido ambiental y/o porciones de la información de audio de entrada que comprenden música sin instrumentación de percusión. Tales porciones de la información de audio de entrada son típicamente temporalmente estacionarias y comprenden un carácter de paso bajo de tal manera que el detector 240 típicamente señala omitir la inclusión de información de ampliación de ancho de banda en la información de audio codificada para tales porciones.

Como alternativa, o además, el detector 240 puede analizar si una forma espectral en la porción de alta frecuencia de la información de audio de entrada puede predecirse con una exactitud razonable (por ejemplo, utilizando los conceptos aplicados por ampliación de ancho de banda a ciegas), basándose en la envolvente espectral de la porción de baja frecuencia. Por consiguiente, el detector puede, por ejemplo, estar configurado para determinar si una diferencia entre una envolvente espectral de una porción de baja frecuencia (que puede describirse, por ejemplo, por la información intermedia 224, o por la representación codificada 222 de la porción de baja frecuencia) y una envolvente espectral de una porción de alta frecuencia (que puede, por ejemplo, determinarse por el detector 240 basándose en la información de audio de entrada 210) es mayor que o igual a una medida predeterminada de la diferencia. Por ejemplo, el detector 240 puede determinar la diferencia en términos de una diferencia de intensidad, o en términos de una diferencia de forma, o en términos de una variación en la frecuencia, o en términos de otros elementos característicos de las envolventes espectrales. De acuerdo con ello, el detector 240 puede decidir (y señalar) que se incluya información de ampliación de ancho de banda 232 en la información de audio de entrada en respuesta a la comprobación de que la diferencia entre la envolvente espectral de la porción de baja frecuencia y la envolvente espectral de la porción de alta frecuencia es mayor que o igual a la medida predeterminada de la diferencia. En otras palabras, el detector 240 puede determinar en qué grado es posible predecir la envolvente espectral de la porción de alta frecuencia basándose en la envolvente espectral de la porción de baja frecuencia, y si la predicción no es posible con buenos resultados (que es el caso, por ejemplo, si la envolvente espectral prevista de la porción de alta frecuencia difiere demasiado de la envolvente espectral real de la porción de alta frecuencia) se puede concluir que la información de ampliación de ancho de banda 232 será necesaria en el lado del decodificador de audio. Sin embargo, en lugar de comparar la envolvente espectral prevista de la porción de alta frecuencia con la envolvente espectral real de la porción de alta frecuencia, el detector 240 puede, como alternativa, comparar la envolvente espectral de la porción de baja frecuencia con la envolvente espectral de la porción de alta frecuencia.

Esto tiene sentido si se supone que la envolvente espectral de la porción de alta frecuencia es típicamente similar a la envolvente espectral de la porción de baja frecuencia cuando se aplica una estimación a ciegas del ancho de banda.

5 Como alternativa o además, el detector 240 puede identificar porciones que comprenden porciones de habla y/o porciones que comprenden sonidos de percusión. Ya que en tales casos la envolvente espectral de la porción de alta frecuencia típicamente difiere considerablemente de la envolvente espectral de la porción de baja frecuencia, el detector puede señalar para que se incluya la información de ampliación de ancho de banda en la representación de audio codificada para tales porciones de la información de audio de entrada (o de la información de audio codificada) que comprende habla no vocalizada o que comprende sonidos de percusión.

15 Sin embargo, como alternativa o además, el detector 240 puede analizar una inclinación espectral de porciones de la información de audio de entrada 210. Además, el detector 240 puede utilizar una información sobre la inclinación espectral de porciones de la información de audio de entrada para decidir si la información de ampliación de ancho de banda 232 debe incluirse en la información de audio codificada 212. Tal concepto se basa en la idea de que la ampliación de ancho de banda a ciegas funciona bien para porciones de un contenido de audio para las que hay más energía (o, en términos generales, intensidad) en el intervalo de baja frecuencia en comparación con el intervalo de alta frecuencia. En contraste, si la porción de alta frecuencia (también designada como intervalo de alta frecuencia) es "dominante", es decir, comprende una cantidad sustancial de energía, típicamente la ampliación de ancho de banda a ciegas no puede reproducir bien el contenido de audio, de tal manera que la información de ampliación de ancho de banda debería estar incluida en la información de audio codificada. Por lo tanto, en algunas realizaciones el detector determina si la inclinación espectral (que describe una distribución de las energías, o, en general intensidades, sobre la frecuencia) es mayor que o igual a un valor umbral de inclinación fijo o variable. Si la inclinación espectral es mayor que o igual al valor umbral de inclinación fija o variable (lo que significa que hay una energía, o intensidad, comparativamente grande, en la porción de alta frecuencia del contenido de audio, al menos cuando se compara con un caso "normal" en el que la energía o la intensidad disminuye al aumentar la frecuencia), el detector puede decidir incluir la información sobre la ampliación de ancho de banda en la información de audio codificada.

20 Además de algunas características anteriormente mencionadas, o de todos ellos, el detector también puede evaluar un coeficiente de cruce en cero de porciones de la información de audio de entrada. Por otra parte, la decisión del detector acerca de si incluir la información de ampliación de ancho de banda también puede estar basada en si el coeficiente de cruce en cero es mayor que o igual a un valor umbral del coeficiente de cruce en cero, fija o variable. Este concepto se basa en la consideración de que un elevado coeficiente de cruce en cero indica normalmente que las frecuencias altas desempeñan un papel importante en la información de audio de entrada, lo que a su vez indica que se debería utilizar una ampliación de ancho de banda guiada por parámetros en el lado de un decodificador de audio.

30 Por otra parte, debe tenerse en cuenta que el detector 240 puede utilizar preferiblemente alguna histéresis para evitar una conmutación excesiva entre la inclusión de la información de ampliación de ancho de banda 232 en la información de audio codificada y una omisión de dicha inclusión. Por ejemplo, la histéresis se puede aplicar al valor umbral de inclinación variable, al valor umbral variable del coeficiente de cruce en cero variable o a cualquier otro valor umbral que se utilice para decidir acerca de una transición desde una inclusión de la información de ampliación de ancho de banda a una evitación de dicha inclusión, o viceversa. Por lo tanto, la histéresis puede variar un valor umbral para reducir una probabilidad para la conmutación a una omisión de la inclusión de la información de ampliación de ancho de banda cuando se incluye la información de ampliación de ancho de banda para una porción actual de la información de audio de entrada. Análogamente, el valor umbral se puede variar para reducir una probabilidad de la conmutación a la inclusión de la información de ampliación de ancho de banda cuando se evita la inclusión de la información de ampliación de ancho de banda para la porción actual de la información de audio de entrada. Por lo tanto es posible reducir los artefactos, que pueden causarse por las transiciones entre los diferentes modos.

35 A continuación se expondrán algunos detalles sobre el proveedor de la información de ampliación de ancho de banda 230. En particular, se explicará cuál información se incluye en la información de audio codificada 212 en respuesta al detector que señala que la información de ampliación de ancho de banda 232 debería estar incluida en la información de audio codificada. Para los fines de la explicación, también se hará referencia a la Figura 3, que muestra una representación esquemática de porciones de frecuencia de la información de audio de entrada y de los parámetros incluidos en la representación de audio codificada. Un eje de abscisas 310 describe una frecuencia y un eje de ordenadas 312 describe una intensidad (por ejemplo, una intensidad, tal como una amplitud o una energía) de diferentes elementos espectrales (como, por ejemplo, coeficientes de MDCT, coeficientes de QMF, coeficientes de FFT, o similares). Como se puede ver, una porción de baja frecuencia de la información de audio de entrada puede, por ejemplo, cubrir un intervalo de frecuencias desde un límite de frecuencias más bajo (por ejemplo, 0, o 50 Hz, o 300 Hz, o cualquier otro límite inferior de frecuencia razonable) hasta una frecuencia de aproximadamente 6,4 kHz. Como puede verse, la representación codificada 222 puede proporcionarse para esta porción de baja frecuencia (por

ejemplo, de 300 Hz a 6,4 kHz, o similares). Por otra parte, hay una porción de alta frecuencia que, por ejemplo, abarca de 6,4 kHz a 8 kHz. Sin embargo, una porción de alta frecuencia puede cubrir evidentemente un intervalo de frecuencias diferente, que normalmente está limitado por el intervalo de frecuencias perceptibles por un oyente humano. Sin embargo, se puede observar en la Figura 3 que, como un ejemplo, una envolvente espectral mostrada con el número de referencia 320 comprende una forma irregular en la porción de alta frecuencia. Por otra parte, puede verse que la envolvente espectral 320 comprende una energía comparativamente grande en la porción de alta frecuencia, e incluso una energía comparativamente alta entre 7,2 kHz y 7,6 kHz. A modo de comparación, en la Figura 3 se muestra también una segunda envolvente espectral 330, en el que la segunda envolvente espectral 330 muestra un decaimiento de la intensidad o de la energía (por ejemplo, por unidad de frecuencia) en la porción de alta frecuencia. Por lo tanto, la envolvente espectral 320 típicamente hará que el detector decida la inclusión de la información de ampliación de ancho de banda en la representación de audio codificada para la porción que comprende la envolvente espectral 320 mientras que la envolvente espectral 330, típicamente hará que el detector se decida por una omisión de la inclusión de la información de ampliación de ancho de banda para la porción del contenido de audio que comprende la envolvente espectral 330.

Como también se puede observar, para una porción del contenido de audio que comprende la envolvente espectral 320 se incluirán cuatro parámetros escalares en la representación de audio codificada como información de ampliación de ancho de banda. Un primer parámetro escalar puede, por ejemplo, describir la envolvente espectral (o un promedio de la envolvente espectral) para la región de frecuencia entre 6,4 kHz y 6,8 kHz, un segundo parámetro escalar puede describir la envolvente espectral 320 (o el promedio de la misma) para la región de frecuencia entre 6,8 kHz y 7,2 kHz, un tercer parámetro escalar puede describir la envolvente espectral 320 (o un promedio de la misma) para la región de frecuencia entre 7,2 kHz y 7,6 kHz, y un cuarto parámetro escalar puede describir la envolvente espectral (o un promedio de la misma) para la región de frecuencias entre 7,6 kHz y 8 kHz. Los parámetros escalares pueden describir la envolvente espectral de una manera absoluta o relativa, por ejemplo, con referencia a un intervalo (o región) de frecuencias espectralmente precedente. Por ejemplo, el primer parámetro escalar puede describir una relación de intensidad (que puede, por ejemplo, normalizarse en algún grado) entre la envolvente espectral en la región de frecuencias entre 6,4 kHz y 6,8 kHz y la envolvente espectral en una región de frecuencia inferior (por ejemplo, por debajo de 6,4 kHz). Los parámetros escalares segundo, tercero y cuarto pueden, por ejemplo, describir una diferencia (o relación) entre (las intensidades de) la envolvente espectral en intervalos de frecuencia adyacentes de manera tal que, por ejemplo, el segundo parámetro escalar pueda describir una relación entre (un valor promedio de) la envolvente espectral en el intervalo de frecuencias entre 6,8 kHz y 7,2 kHz y la envolvente espectral en el intervalo de frecuencias entre 6,4 kHz y 6,8 kHz.

Por otra parte, hay que señalar que una representación codificada de la porción de baja frecuencia, es decir, la porción de frecuencia por debajo de 6,4 kHz puede incluirse en cualquier caso. La porción de frecuencia por debajo de 6,4 kHz (porción de baja frecuencia) puede codificarse usando cualquiera de los conceptos de codificación bien conocidos, por ejemplo, utilizando una codificación de "audio en general" tal como AAC (o un derivado del mismo) o una codificación de voz (como, por ejemplo CELP, ACELP o un derivado de las mismas). Por lo tanto, para una porción del contenido de audio que comprende la envolvente espectral 320, tanto una representación codificada de la porción de baja frecuencia como cuatro parámetros de ampliación de ancho de banda escalares (que pueden cuantificarse utilizando un número relativamente pequeño de bits) se incluirán en la representación codificada de audio. En cambio, para una porción del contenido de audio que comprende la envolvente espectral 330, sólo la representación codificada de la porción de baja frecuencia estará incluida en la representación de audio codificada, pero no habrá parámetros de ampliación de ancho de banda (escalar) incluidos en la representación de audio codificada (lo que, sin embargo no causa problemas graves ya que la envolvente espectral 330 presenta una característica regular y de decaimiento (paso bajo), que puede reproducirse muy bien utilizando una ampliación de ancho de banda a ciegas)

Para concluir, el codificador de audio 200 está configurado para incluir selectivamente parámetros que representan una envolvente espectral de una porción de alta frecuencia de la información de audio de entrada en la información de audio codificada de una manera que se adapta a la señal como una información de ampliación de ancho de banda. Por ejemplo los parámetros de ampliación de ancho de banda escalares mencionados con referencia a la Figura 3 pueden incluirse en la información de audio codificada de una manera que se adapta a la señal. Hablando en términos generales, el codificador de frecuencia más baja 220 puede configurarse para codificar una porción de baja frecuencia de la información de audio de entrada 210 que comprende frecuencias de hasta una frecuencia máxima que se encuentra en un intervalo entre 6 y 7 kHz (en el que se ha utilizado un límite de 6,4 kHz en el ejemplo de la Figura 3). Por otra parte el codificador de audio puede estar configurado para incluir selectivamente en la representación de audio codificada entre tres y cinco parámetros que describen las intensidades de las porciones de señal de alta frecuencia que tienen anchos de banda de entre 300 Hz y 500 Hz. En el ejemplo de la Figura 3 se han mostrado cuatro parámetros escalares que describen las intensidades de las porciones de señal de alta frecuencia que tienen anchos de banda de aproximadamente 400 Hz. En otras palabras, el codificador de audio puede estar configurado para incluir en la representación de audio codificada cuatro parámetros cuantificados escalares que describen intensidades de cuatro porciones de la señal de alta frecuencia, cubriendo las porciones de señal de alta frecuencia intervalos de frecuencia (por ejemplo como se muestra en la Figura 3) por encima de la

porción de baja frecuencia (por ejemplo, como se explica con referencia a la Figura 3). Por ejemplo, el codificador de audio puede estar configurado para incluir selectivamente en la representación de audio codificada una pluralidad de parámetros que describen una relación entre las energías o intensidades de porciones de frecuencia espectralmente adyacentes, en las que uno de los parámetros describe una relación entre una energía o intensidad de una primera porción de ampliación de ancho de banda de alta frecuencia y una energía o intensidad de una porción de baja frecuencia, y en el que otro de los parámetros describía relaciones entre energías o intensidades de otras porciones de alta frecuencia de ampliación de ancho de banda (en el que las porciones de alta frecuencia de ampliación de ancho de banda pueden ser las porciones de frecuencia entre 6,4 y 6,8 kHz, entre 6,8 y 7,2 kHz, entre 7,2 kHz y 7,6 kHz, y entre 7,6 kHz y 8 kHz. Como alternativa, los entre tres y cinco parámetros de forma de la envolvente (que describen intensidades de porciones de señal de alta frecuencia) pueden cuantificarse en vectores. La cuantificación vectorial es típicamente un poco más eficiente que la cuantificación escalar. Por otro lado, la cuantificación vectorial es más compleja que la cuantificación escalar. En otras palabras, como alternativa, la cuantificación de los cuatro valores de energía de ampliación de ancho de banda, puede llevarse a cabo mediante una cuantificación vectorial (en lugar de usar una cuantificación escalar).

Para concluir, el codificador de audio puede estar configurado para incluir una información de ampliación de ancho de banda, comparativamente sencilla, en la representación de audio codificada, de tal manera que una tasa de bits de la representación de audio codificada se incrementa sólo ligeramente para porciones de la información de audio de entrada (o de la representación de audio codificada) para las que se encuentra, por el detector, que una ampliación de ancho de banda guiada por parámetros sería deseable.

3. Decodificador de audio de acuerdo con la Figura 4

La Figura 4 muestra un diagrama esquemático de bloques de un decodificador de audio de acuerdo con una realización de la presente invención. El decodificador de audio 400 de acuerdo con la Figura 4 recibe una información de audio codificada 410 (que puede, por ejemplo, proporcionarse por el codificador de audio 100 o por el codificador de audio 200), y proporciona, basándose en la misma, información de audio decodificada 412.

El decodificador de audio 400 comprende un decodificador de baja frecuencia 420, que recibe la información de audio codificada 410 (o al menos la representación codificada de la porción de baja frecuencia incluida en la misma), decodifica la representación codificada de la porción de baja frecuencia, y obtiene una representación decodificada 422 de la porción de baja frecuencia.

El decodificador de audio 400 también comprende una ampliación de ancho de banda 430 que está configurada para obtener una señal de ampliación de ancho de banda 432 utilizando una ampliación de ancho de banda a ciegas para porciones del contenido (representado por la información de audio codificada 410) de audio (codificado) para las que no hay parámetros de ampliación de ancho de banda incluidos en la información de audio codificada 410, y obtiene la señal de ampliación de ancho de banda 432 utilizando una ampliación de ancho de banda guiada por parámetros (haciendo uso de la información de ampliación de ancho de banda o de parámetros de ampliación de ancho de banda incluidos en la información de audio codificada 410) para porciones del contenido de audio para las que los parámetros de ampliación de ancho de banda están incluidos en la información de audio codificada (o representación de audio codificada) 410.

En consecuencia, el decodificador de audio 400 es capaz de llevar a cabo una ampliación de ancho de banda, independientemente de si hay o no parámetros de ampliación de ancho de banda incluidos en la información de audio codificada 410. Por lo tanto, el decodificador de audio se puede adaptar a la información de audio codificada 410 y permite un concepto en el que hay una conmutación entre una ampliación de ancho de banda a ciegas y una ampliación de ancho de banda guiada por parámetros. En consecuencia, el decodificador de audio 400 es capaz de manejar una información de audio codificada 410 en la que los parámetros de ampliación de ancho de banda sólo se incluyen para porciones (por ejemplo tramas) del contenido de audio que no puede ser reconstruidas con suficiente calidad utilizando una ampliación de ancho de banda a ciegas. Por lo tanto, es posible proporcionar la información de audio descodificada 412, que comprende tanto la representación decodificada de la porción de baja frecuencia como la señal de ampliación de ancho de banda (en el que este último puede, por ejemplo, añadirse en la representación decodificada 422 de la porción de baja frecuencia para obtener de esta manera la información de audio decodificada 412).

Por lo tanto, el decodificador de audio 400 ayuda a obtener un buen compromiso entre la calidad de audio y la tasa de bits.

Una mejora adicional opcional del decodificador de audio 400 se describirá a continuación, por ejemplo, haciéndose referencia a la Figura 5.

4. Decodificador de audio de acuerdo con la Figura 5

La Figura 5 muestra un diagrama esquemático de bloques de un decodificador de audio 500, de acuerdo con otra realización de la presente invención. El decodificador de audio 500 recibe una información de audio codificada (también designada como representación de audio codificada) 510 y proporciona, basándose en la misma, una información de audio decodificada (también designada como representación de audio decodificada) 512. El decodificador de audio 500 comprende un decodificador de baja frecuencia 520, que puede ser igual al decodificador de baja frecuencia 420 y puede cumplir una funcionalidad comparable. Por lo tanto, el decodificador de baja frecuencia 500 proporciona una representación decodificada 522 de una porción de baja frecuencia de un contenido de audio representado por la información de audio codificada 510. El decodificador de audio 500 también comprende una ampliación de ancho de banda 530, que puede cumplir la misma funcionalidad que la ampliación de ancho de banda 430.

La ampliación de ancho de banda 530 puede por lo tanto proporcionar una señal de ampliación de ancho de banda 532, que típicamente se combina con (por ejemplo, se añade a) la representación decodificada 522 de la porción de baja frecuencia, para obtener de este modo la información de audio decodificada 512. La ampliación de ancho de banda 530 puede, por ejemplo, recibir la representación decodificada 522 de la porción de baja frecuencia 522. Sin embargo, como alternativa, la ampliación de ancho de banda 532 puede recibir una información de control (que también se considerará como una información auxiliar o como una información intermedia) 524, que se proporciona por el decodificador de baja frecuencia 520. La información auxiliar o información de control o información intermedia 524 puede, por ejemplo, representar una forma espectral de la porción de baja frecuencia del contenido de audio, un coeficiente de cruce en cero de la representación decodificada de la porción de baja frecuencia, o cualquier otra cantidad intermedia utilizada por el decodificador de baja frecuencia 520 que es útil en el proceso de la ampliación del ancho de banda. Por otra parte, el decodificador de audio comprende un control 540, que está configurado para proporcionar una información de control 542 que indica si la ampliación de ancho de banda 530 debería llevar a cabo una ampliación del ancho de banda a ciegas o una ampliación del ancho de banda guiada por parámetros. El control 540 puede utilizar diferentes tipos de información para proporcionar la información de control 542. Por ejemplo, el control 540 puede recibir una bandera de corriente de bits del modo de ampliación de ancho de banda 510, que puede incluirse en la información de audio codificada 510. Por ejemplo, puede haber una bandera de corriente de bits en el modo de ampliación del ancho de banda para cada porción (por ejemplo, trama) de la información de audio codificada, que se puede extraer de la información de audio codificada mediante el control 540, y que se puede usar para derivar la información de control 542 (o que pueden constituir de inmediato la información de control 542). Sin embargo, como alternativa, el control 540 puede recibir una información que representa la porción de baja frecuencia, y/o que describe cómo decodificar la porción de baja frecuencia (y que por lo tanto lleva la designación de "información que decodifica la porción de baja frecuencia"). Como alternativa, o además, el control 540 puede recibir la información de control o información auxiliar o información intermedia 524 procedente desde el decodificador de baja frecuencia que puede por ejemplo llevar información acerca de una envolvente espectral de la porción de baja frecuencia, y/o una información acerca del coeficiente de cruce en cero de la representación decodificada de la porción de baja frecuencia. Sin embargo, la información de control o información auxiliar o información intermedia 524 también puede llevar una información acerca de estadísticas de la representación decodificada 522 de la porción de baja frecuencia, o puede representar cualquier otra información intermedia que se derive mediante el decodificador de baja frecuencia 520, de la representación codificada de la porción de baja frecuencia (también designada como información de decodificación de la porción de baja frecuencia).

Como alternativa, o además, el control 540 puede recibir la representación decodificada 522 de la porción de baja frecuencia y puede por sí mismo derivar valores de características (por ejemplo, una información de coeficiente de cruce en cero, una información de envolvente espectral, una información de inclinación espectral, o similares), a partir de la representación decodificada 522 de la porción de baja frecuencia.

Por lo tanto, el control 540 puede evaluar una bandera de corriente de bits para proporcionar la información de control a ciegas / guiada por parámetros 542, si una bandera de corriente de bits de este tipo (que señala si debería utilizarse una ampliación de ancho de banda a ciegas o una ampliación de ancho de banda guiada por parámetros) se halla incluida en la información de audio codificada 510. Sin embargo, si no hay tal bandera de flujo de bits incluida en la información de audio codificada 510 (por ejemplo, para ahorrar tasa de bits) el control 540 típicamente determina si se utiliza una ampliación de ancho de banda a ciegas o una ampliación de ancho de banda guiada por parámetros basándose en otra información. Para este fin, la información de decodificación de la porción de baja frecuencia (que puede ser igual a la representación codificada de la porción de baja frecuencia, o a un subconjunto de la misma) puede evaluarse por el control 540. Como alternativa, o adicionalmente, el control puede considerar la representación decodificada 522 de la porción de baja frecuencia para tomar una decisión acerca de si se debe utilizar una ampliación del ancho de banda a ciegas o una ampliación de ancho de banda guiada por parámetros, es decir, para proporcionar la información de control 542. Por otra parte, el control 540 puede, opcionalmente, utilizar la información de control o la información auxiliar o la información intermedia 524 proporcionada por el decodificador de baja frecuencia 520, siempre y cuando el decodificador de baja frecuencia 520 proporcione cualquier cantidad que sea utilizable por el control 540.

En consecuencia, el control 540 puede conmutar la ampliación de ancho de banda entre la ampliación de ancho de

banda a ciegas y la ampliación de ancho de banda guiada por parámetros.

5 En el caso de una ampliación de ancho de banda a ciegas, la ampliación de ancho de banda 530 puede proporcionar la señal de ampliación de ancho de banda 532 basándose en la representación decodificada 522 de la porción de baja frecuencia sin evaluar ningún parámetro de corriente de bits adicional. En cambio, en el caso de una ampliación del ancho de banda guiada por parámetros, la ampliación de ancho de banda 530 puede proporcionar la señal de la ampliación de ancho de banda 532 teniendo en consideración parámetros de corrientes de bits de ancho de banda adicionales (especializados), que ayudan a determinar las características de la porción de alta frecuencia del contenido de audio (es decir, las características de la señal de la ampliación de ancho de banda). Sin embargo, 10 la ampliación de ancho de banda 530 también puede utilizar la representación decodificada 522 de la porción de baja frecuencia, y/o la información de control o la información auxiliar o la información intermedia 524 proporcionada por el decodificador de baja frecuencia 520, para proporcionar la señal 532.

15 Por lo tanto, la decisión entre el uso de una ampliación de un ancho de banda a ciegas y una ampliación de ancho de banda guiada por parámetros determina de manera efectiva si los parámetros de ampliación de ancho de banda especializado (que típicamente no se utilizan por el decodificador de baja frecuencia 520 para proporcionar la representación decodificada de la porción de baja frecuencia) se aplican para obtener la señal de ampliación del ancho de banda (que típicamente describe la porción de alta frecuencia del contenido de audio representado por la información de audio codificada).

20 Para resumir lo anterior, el decodificador de audio 500 puede estar configurado para decidir si obtener la señal de ampliación de ancho de banda 532 utilizando una ampliación de ancho de banda a ciegas o utilizando una ampliación de ancho de banda guiada por parámetros en una base trama a trama (en el que una "trama" es un ejemplo de una porción del contenido de audio, y en el que una trama puede, por ejemplo, comprender una duración de entre 10 ms y 40 ms, y puede tener preferiblemente una duración de aproximadamente $20 \text{ ms} \pm 2 \text{ ms}$). Por lo tanto, el decodificador de audio puede estar configurado para conmutar entre una ampliación de ancho de banda a ciegas y una ampliación de ancho de banda guiada por parámetros con una granularidad temporal muy precisa.

30 Además, cabe señalar que el decodificador de audio 500 es típicamente capaz de conmutar entre un uso de una ampliación de ancho de banda a ciegas y una ampliación de ancho de banda guiada por parámetros dentro de una pieza contigua de contenido de audio. Por lo tanto, la conmutación entre la ampliación de ancho de banda a ciegas y la ampliación de ancho de banda guiada por parámetros puede realizarse sustancialmente en cualquier momento (considerando naturalmente la alineación de tramas) dentro de una pieza contigua de contenido de audio, para adaptar la ampliación del ancho de banda a las características (cambiantes) de las diferentes porciones de una sola 35 pieza de contenido de audio.

40 Tal como se mencionó anteriormente, el decodificador de audio (preferiblemente el control 540) puede configurarse para evaluar banderas (por ejemplo, una bandera de un solo bit por trama) incluidas en la información de audio codificada 510 para diferentes porciones (por ejemplo tramas) del contenido de audio, para decidir acerca de si se debe utilizar una ampliación de ancho de banda a ciegas o una ampliación del ancho de banda guiada por parámetros. En este caso, el control 540 puede mantenerse muy sencillo, a expensas de que es necesario incluir una bandera de señalización en la información de audio codificada para cada porción del contenido de audio. Sin embargo, como alternativa, el control 540 puede estar configurado para decidir acerca de si se debe utilizar una ampliación de ancho de banda a ciegas o una ampliación de ancho de banda guiada por parámetros basándose en 45 la representación codificada de la porción de baja frecuencia (que puede incluir el uso de la información de control o de la información auxiliar o de la información intermedia 524 derivada por el decodificador de baja frecuencia 520 a partir de dicha representación codificada de la porción de baja frecuencia, y que también puede incluir el uso de la representación decodificada 522, que se deriva de la representación codificada de la porción de baja frecuencia por el decodificador de baja frecuencia 520) sin evaluar una bandera de señalización (especializada) del modo de 50 ampliación del ancho de banda. Por lo tanto, es posible llevar a cabo una conmutación entre la ampliación del ancho de banda a ciegas y la ampliación del ancho de banda guiado por parámetros aun sin una sobrecarga de señalización en la corriente de bits.

55 El decodificador de audio (o el control 540) puede estar configurado para decidir acerca de si se debe utilizar una ampliación de ancho de banda a ciegas o una ampliación de ancho de banda guiada por parámetros basándose en una o más características de la representación decodificada de la porción de baja frecuencia. Tales características, como, por ejemplo, una información de inclinación espectral, una información del coeficiente de cruce en cero, o similares, pueden extraerse de la representación decodificada 522 de la porción de baja frecuencia, o pueden señalizarse mediante la información de control/información auxiliar/información intermedia 524. Por ejemplo, el 60 decodificador de audio (o el control 540) pueden estar configurados para decidir acerca de si se debe utilizar una ampliación de ancho de banda a ciegas o una ampliación de ancho de banda guiada por parámetros basándose en coeficientes de predicción lineales cuantificados (que pueden, por ejemplo, estar incluidos en la información de control/información auxiliar/información intermedia 524) y/o en función de estadísticas en el dominio del tiempo de la representación decodificada 522 de la porción de baja frecuencia.

A continuación, se describirán algunos conceptos de cómo lograr la ampliación del ancho de banda. Por ejemplo, la ampliación del ancho de banda puede estar configurada para obtener la señal de ampliación de ancho de banda 532 utilizando uno o más rasgos de la representación decodificada 522 de la porción de baja frecuencia y/o uno o más parámetros del decodificador de baja frecuencia 520 (que puede señalizarse mediante la información de control/información auxiliar/información intermedia 524) para porciones temporales del contenido de audio (de entrada) para las que no hay parámetros de ampliación incluidos en la información de audio codificada. Por lo tanto, la ampliación del ancho de banda. 530 puede llevar a cabo una ampliación de ancho de banda a ciegas, que se basa en la idea de concluir a partir de la representación decodificada de la porción de baja frecuencia hacia la porción de alta frecuencia del contenido de audio representado por la información de audio codificada. Por ejemplo, la ampliación de ancho de banda 530 puede estar configurada para obtener la señal de ampliación de ancho de banda 532 utilizando información de centroide espectral, y/o usando una información de energía, y/o usando coeficientes de filtro (por ejemplo, codificados) para porciones temporales del contenido de audio de entrada para las que no hay parámetros de ampliación incluidos en la información de audio codificada 510. En consecuencia, es posible lograr una buena ampliación de ancho de banda a ciegas.

Sin embargo, por supuesto es también posible aplicar diferentes conceptos para la ampliación de ancho de banda a ciegas.

Sin embargo, la ampliación del ancho de banda puede estar configurada para obtener la señal de ampliación de ancho de banda 532 utilizando los parámetros de corrientes de bits que describen una envolvente espectral de una porción de alta frecuencia para porciones temporales del contenido de audio para las que hay parámetros de corrientes de bits de ancho de banda incluidos en la información de audio codificada. En otras palabras, es posible lograr la ampliación del ancho de banda guiada mediante parámetros que describen la envolvente espectral de la porción de alta frecuencia. Los parámetros de corrientes de bits que describen la envolvente espectral de la porción de alta frecuencia pueden soportar la ampliación del ancho de banda guiada mediante parámetros (que pueden, sin embargo, adicionalmente basarse en algunas o en la totalidad de las cantidades utilizadas por la ampliación de ancho de banda a ciegas).

Por ejemplo, se ha descubierto la ampliación de ancho de banda debería estar configurada preferiblemente para evaluar entre tres y cinco parámetros de corrientes de bits que describen intensidades de porciones de señales de alta frecuencia que tienen anchos de banda entre 300 Hz y 500 Hz, para obtener la señal de ampliación de ancho de banda. El uso de un número comparativamente pequeño de este tipo de parámetros de corrientes de bits no aumenta sustancialmente la tasa de bits, sino que todavía da origen a una mejora suficiente de la ampliación del ancho de banda guiada en el caso de porciones de señales "difíciles", de tal manera que la calidad alcanzable mediante la ampliación de ancho de banda así guiada para porciones de señales "difíciles" es comparable a la calidad obtenible para porciones de señales "fáciles" que utilizan la ampliación del ancho de banda guiada a ciegas (en el que las porciones de señales "difíciles" son porciones de señales para las que la ampliación de ancho de banda a ciegas no daría lugar a un audio de una calidad buena o aceptable, mientras que las porciones de señales "fáciles" son porciones de señales para las que la ampliación de ancho de banda a ciegas conlleva resultados suficientes).

En consecuencia, se prefiere que se cuantifiquen escalaramente entre tres y cinco parámetros de corrientes de bits que describen intensidades de porciones de señales de alta frecuencia que tienen anchos de banda de entre 300 Hz y 500 Hz con una resolución de dos o tres bits, de tal manera que haya entre 6 y 15 bits de parámetros de conformación espectral de ampliación de ancho de banda por trama. Se ha descubierto que una tasa de bits baja de este tipo ya es suficiente para obtener un ancho de banda razonablemente bueno en el caso de porciones "difíciles" de contenido de audio.

Opcionalmente, la ampliación de ancho de banda 530 puede estar configurada para llevar a cabo un suavizado de las energías de la señal de ampliación de ancho de banda cuando se cambia de ampliación de ancho de banda a ciegas a la ampliación de ancho de banda guiada por parámetros y/o cuando se cambia de ampliación de ancho de banda guiada por parámetros a ampliación de ancho de banda a ciegas. En consecuencia, se reducen las discontinuidades en la forma espectral cuando se cambia entre la ampliación de ancho de banda a ciegas y la ampliación de ancho de banda guiada por parámetros. Por ejemplo, la ampliación de ancho de banda a ciegas puede estar configurada para amortiguar una porción de alta frecuencia de la señal del ancho de banda para una porción del contenido de audio al que se aplica una ampliación de ancho de banda guiada por parámetros después de una porción del contenido de audio al que se aplica un ancho de banda a ciegas. Además, la ampliación de ancho de banda puede estar configurada para reducir una amortiguación para una porción de alta frecuencia de la señal del ancho de banda (es decir, para enfatizar un tanto una porción de alta frecuencia de la señal de ampliación del ancho de banda) para una porción del contenido de audio al que se aplica una ampliación del ancho de banda a ciegas después de una porción del contenido de audio al que se aplica una ampliación de ancho de banda guiada por parámetros. Sin embargo, también es posible llevar a cabo una suavización mediante cualquier otra operación que reduzca las discontinuidades de la forma espectral de la porción de alta frecuencia cuando se cambia entre los

modos de ampliación del ancho de banda. Por lo tanto, se mejora una calidad de audio mediante la reducción de artefactos.

Para concluir, el decodificador de audio 500 permite una buena calidad de decodificación de un contenido de audio tanto en el caso en que se proporciona una información de ampliación de ancho de banda en la información de audio codificada como en el caso en que no se proporciona ninguna información de ampliación del ancho de banda en la información de audio codificada. El decodificador de audio puede cambiar entre una ampliación del ancho de banda a ciegas y una ampliación de ancho de banda guiada por parámetros con una granularidad temporal precisa (por ejemplo, en una base trama a trama) en la que los artefactos se mantienen pequeños.

5. Método para proporcionar una información de audio codificada basándose en una información de audio de entrada, de acuerdo con la Figura 6

La Figura 6 muestra un diagrama de flujo de un método 600 para proporcionar una información de audio codificada basándose en una información de audio de entrada. El método 600 comprende codificar 610 una porción de baja frecuencia de la información de audio de entrada para obtener una representación codificada de la porción de baja frecuencia. El método 600 también comprende proporcionar información de ampliación de ancho de banda 620 basándose en la información de audio de entrada, en el que la información de ampliación de ancho de banda se incluye selectivamente en la información de audio codificada de una manera que se adapta a la señal.

Cabe señalar que el método 600 de acuerdo con la Figura 6 se puede complementar con cualquiera de los rasgos y funcionalidades descritos en el presente documento con respecto al codificador de audio (y también con respecto al decodificador de audio).

6. Método para proporcionar una información de audio decodificada de acuerdo con la Figura 7

La Figura 7 muestra un diagrama de flujo de un método para proporcionar una información de audio decodificada, de acuerdo con una realización de la invención. El método 700 comprende la decodificación 710 de una representación codificada de una porción de baja frecuencia para obtener una representación decodificada de la porción de baja frecuencia. El método 700 también comprende la obtención 720 de una señal de ampliación de ancho de banda utilizando una ampliación de ancho de banda a ciegas para porciones de un contenido de audio para las que no hay parámetros de ampliación de ancho de banda incluidos en la información de audio codificada. Además, el método 700 comprende obtener 730 la señal de ampliación de ancho de banda utilizando una ampliación de ancho de banda guiada por parámetros para porciones del contenido de audio para las cuales hay parámetros de ampliación de ancho de banda incluidos en la información de audio codificada.

Cabe observar que el método 700 de acuerdo con la Figura 7 puede complementarse mediante cualquiera de los rasgos y funcionalidades descritos en el presente documento con respecto al decodificador de audio (y también con respecto al codificador de audio).

7. Representación de audio codificado de acuerdo con la Figura 8

La Figura 8 muestra una ilustración esquemática de un audio codificado 800 que representa una información de audio.

La representación de audio codificada (también designada como información de audio codificada) comprende una representación codificada de una porción de baja frecuencia de la información de audio. Por ejemplo, se proporciona una representación codificada 810 de una porción de baja frecuencia de una información de audio para una primera porción de la información de audio, por ejemplo, para una primera trama de la información de audio. Por otra parte, también se proporciona una representación codificada de una porción de baja frecuencia de la información de audio para una segunda porción (por ejemplo, una segunda trama) de la información de audio. Sin embargo, la representación de audio codificada 800 también comprende una información de ampliación de ancho de banda, en la que la información de ampliación de ancho de banda está incluida en la representación de audio codificada de una manera que se adapta a la señal para algunas de, pero no para todas, las porciones de la información de audio. Por ejemplo, una información de ampliación de ancho de banda 812 está incluida para la primera porción de la información de audio. En cambio, no se proporciona ninguna información de ampliación de ancho de banda para la segunda porción de la información de audio.

Para concluir, la representación de audio codificada 800 típicamente se proporciona por los codificadores de audio descritos en el presente documento, y se evalúa por los decodificadores de audio descritos en el presente documento. Naturalmente, la representación de audio codificada puede almacenarse en un medio legible por ordenador, no transitorio, o similares. Por otra parte, hay que señalar que la representación de audio codificada 800 puede complementarse por cualquiera de los rasgos, elementos de información, etc., descritos con respecto al codificador de audio y con respecto al decodificador de audio.

8. Conclusiones y otros aspectos

5 Las realizaciones de acuerdo con la presente invención abordan los problemas de la ampliación de ancho de banda convencional en la codificación de audio con una tasa de bits muy baja y los inconvenientes de las técnicas de ampliación de ancho de banda convencionales existentes mediante la propuesta de una ampliación de ancho de banda "mínimamente guiada" como una combinación, que se adapta a la señal, de una ampliación de ancho de banda guiada por parámetros, que

- 10 • utiliza una ampliación de ancho de banda guiada, es decir, transmite unos pocos bits de información secundaria por cada 20 ms (por ejemplo, por cada trama de audio), solamente si el contenido de alta frecuencia (por ejemplo, la porción de alta frecuencia) del audio de entrada no puede reconstruirse suficientemente bien a partir del audio de baja frecuencia (por ejemplo, la porción de baja frecuencia del contenido de audio),
- 15 • utiliza una ampliación de ancho de banda a ciegas, es decir, una reconstrucción clásica de componentes de alta frecuencia (por ejemplo, de una porción de alta frecuencia) a partir de rasgos de núcleo de baja frecuencia (por ejemplo, rasgos de una porción de baja frecuencia reconstruida) tales como centroide espectral, energía, inclinación, coeficientes de filtro codificados, de lo contrario;
- 20 • presenta una complejidad computacional muy baja gracias al hecho de utilizar la cuantificación escalar en lugar de la cuantificación vectorial de la información secundaria y por el hecho de evitar operaciones que involucran grandes cantidades de puntos de datos, tales como las transformadas de Fourier y la autocorrelación y/o cálculos de filtro
- es robusta con respecto a las características de las señales de entrada, es decir, no está optimizada para determinadas señales de entrada, tales como el habla de los adultos en ambientes silenciosos, para poder funcionar bien en todos los tipos de habla, así como también la música.

25 La cuestión de cuál(es) parámetro(s) deben transmitirse como información secundaria en la parte de la ampliación de ancho de banda guiada de las realizaciones de acuerdo con la presente invención, y de cuándo transmitir los parámetros, sigue pendiente de respuesta.

30 Se comprobó que en los códecs de banda ancha tales como AMR-WB, la envolvente espectral de la región de alta frecuencia por encima de la región del codificador de núcleo representa los datos más críticos necesarios (o deseables) para llevar a cabo la ampliación de ancho de banda con una calidad adecuada. Todos los otros parámetros, tales como la estructura espectral precisa y la envolvente temporal se pueden derivar de la señal de núcleo decodificada con bastante precisión, o son de poca importancia perceptiva. Por ello, la parte guiada de ampliación del ancho de banda mínimamente guiada, descrita aquí, sólo transmite la envolvente espectral de alta frecuencia como información secundaria (por ejemplo, como información de ampliación de ancho de banda). Esto ayuda a mantener bajo la tasa de la información secundaria de la ampliación del ancho de banda. Además, se descubrió experimentalmente que las extensiones de ancho de banda a ciegas, proporcionan una calidad suficiente, es decir, al menos aceptable, sobre los pasajes de señales temporalmente estacionarias con un carácter de paso bajo más o menos pronunciado. El habla vocalizada, el ruido ambiental y las secciones de música sin instrumentación de percusión, son ejemplos comunes. De hecho, la mayor entrada en un sistema de voz de banda ancha y de codificación de audio típicamente recae en esta categoría.

45 Sin embargo, los segmentos de señal, cuyos espectros instantáneos presentan una envolvente muy diferente en la región de alta frecuencia (por ejemplo, en la porción de alta frecuencia) que en la región (o porción de baja frecuencia) de baja frecuencia (núcleo-codificador) se han de codificar, preferentemente, por medio de una ampliación de ancho de banda guiada que transmite una representación cuantificada de la envolvente espectral de alta frecuencia como información secundaria (por ejemplo, como información de ampliación de ancho de banda). La razón es que en tales constituciones espectrales, las extensiones de ancho de banda a ciegas, son generalmente incapaces de predecir la progresión de la envolvente espectral de alta frecuencia a partir de la envolvente de la señal de núcleo, como se indica mediante los coeficientes de filtro codificados o mediante la señal residual en forma espectral (también conocida como excitación en los codificadores de voz). Ejemplos destacados son el habla no vocalizada, especialmente los fricativos y los africados fuertes tales como la "s" o la "z" alemana, así como también determinados sonidos de percusión que se encuentran principalmente en la música moderna. En realizaciones de acuerdo con la presente invención, la ampliación del ancho de banda guiada, por lo tanto se activa solamente para tales espectros de alta frecuencia "impredecibles".

60 Una ampliación de ancho de banda mínimamente guiada de acuerdo con la presente invención se llevó a cabo en el contexto del LD-USAC, una versión de bajo retardo de xHE-AAC, para ampliar el ancho de banda de señal de banda ancha codificada (codificada en WB), a 13,2 kbits/s de 6,4 a 8,0 kHz. En el lado del codificador, la decisión a ciegas/guiada se calcula por cada trama de códec de 20 ms a partir de la inclinación espectral de la señal de entrada en una escala de frecuencia perceptiva (un rasgo existente que también se utiliza en la ruta de acceso de codificación por ACELP), así como también rasgos en el dominio del tiempo tales como el cambio en el coeficiente de cruce en cero de la señal de entrada proporcionada por un detector de transitorios existente (que también se

utiliza para otras decisiones en el modo de codificación). Más específicamente, si la inclinación espectral es positiva, lo que significa que la energía espectral tiende a aumentar al aumentar la frecuencia, y por encima de un umbral especificado, y al mismo tiempo el coeficiente de cruce en cero se ha incrementado en una cierta relación o se halla por encima de un determinado umbral, lo que significa que la trama actual representa el inicio o se encuentra dentro de un paso de forma de onda de ruido, entonces se elige y señala la ampliación del ancho de banda guiada. En caso contrario, se selecciona la ampliación del ancho de banda a ciegas. En cuanto a los umbrales mencionados anteriormente, se aplica además una sencilla histéresis para reducir la probabilidad de la conmutación de ida y vuelta entre la ampliación de ancho de banda guiada y a ciegas. Una vez adoptado el modo de la ampliación de ancho de banda para una trama, los umbrales de decisión a utilizarse en tramas sucesivas se reducen un poco para que sea más probable que el códec se mantenga en el modo guiado. Una vez que se ha decidido cambiar de nuevo al modo a ciegas, los umbrales originales se restauran, lo que hace que sea menos probable que la decisión de la ampliación del ancho de banda oscile de inmediato de nuevo al modo guiado.

El resto del procedimiento para la ampliación de ancho de banda a tramas se resume como sigue:

1. Si la ampliación de ancho de banda se halla en modo a ciegas, transmitir un "0" por medio de un bit en la corriente de bits para señalar este modo al decodificador. Opcionalmente, no transmitir ningún bit y dejar que el descodificador identifique la trama utilizando el modo de ampliación de ancho de banda a ciegas mediante un análisis del lado del decodificador de la señal núcleo.

2. Si la ampliación de ancho de banda está en el modo guiado, transmitir un "1" utilizando un bit en la corriente de bits. Seguidamente el codificador calcula cuatro índices de ganancia, cada uno de los cuales cubre 400 Hz de la señal de entrada, para permitir una conformación espectral exacta de la región de ampliación de ancho de banda de 6,4 a 8 kHz. En una realización de USAC de bajo retardo, cada uno de los cuatro índices es el resultado de una cuantificación escalar de una de las cuatro energías de QMF de ampliación de ancho de banda con respecto a la energía de QMF precedente (o con respecto a la energía del espectro de QMF de 4,8 - 6,4 kHz, en el caso de primera ganancia de la ampliación de ancho de banda). Dado que se utiliza un cuantificador de elevación media de 2 bits con un tamaño de escalón de 2 dB, las ganancias abarcan un intervalo de valores de -3 ...3 dB y consumen 8 bits por trama. Esto permite obtener una información secundaria total de 9 bits por trama de ancho de banda guiado u, opcionalmente, de 8 bit si se excluye la señalización como en la etapa 1.

3. En el decodificador correspondiente, se lee el, primer bit de la ampliación de ancho de banda. Si es "0", se utiliza la ampliación de ancho de banda a ciegas, en caso contrario se llena con 8 bits más y se utiliza la ampliación de ancho de banda guiada. Opcionalmente, la lectura del primer bit de la ampliación de ancho de banda se omite (ya que este bit no se halla presente en la corriente de bits), y la decisión a ciegas/guiada se lleva a cabo localmente mediante análisis de señal de núcleo, como se mencionó en la etapa 1.

4. Si en el decodificador se determinó la ampliación de ancho de banda a ciegas, se lleva a cabo una ampliación de ancho de banda utilizando solamente rasgos de la señal de núcleo decodificada. Esta ampliación de ancho de banda sigue esencialmente el concepto de ampliación de ancho de banda descrito en uno de los siguientes: [2], [3], [6] y [9] pero en el dominio de QMF en lugar de en el de DFT y con solamente rasgos de baja complejidad derivados del espectro QMF de núcleo, por ejemplo, centroide espectral/inclinación.

5. Si se ha seleccionado el modo de ampliación de ancho de banda guiada en el decodificador, los cuatro índices de ganancia de 2 bits se cuantifican a la inversa en ganancias de energía de QMF y se aplican para la conformación espectral de las bandas de región de ampliación de ancho de banda de QMF que se reconstruyen como en la etapa 4. En otras palabras, se emplea una ampliación de ancho de banda a ciegas aquí también, con la salvedad de que la conformación espectral se realiza a través de factores de escala transmitidos en la corriente de bits, en lugar del escalamiento extrapolado a partir de la señal de núcleo (que, como consecuencia, constituye una ampliación de ancho de banda de audio guiada por parámetros).

6. Cuando se cambia de ampliación de ancho de banda guiada y a ciegas, desde una trama a la siguiente, se lleva a cabo un sencillo suavizado de las energías de alta frecuencia para minimizar los artefactos de conmutación (discontinuidades de energía de alta frecuencia) causados por el comportamiento similar a paso bajo de la ampliación de ancho de banda a ciegas. El suavizado esencialmente funciona como un fundido cruzado entre las extensiones de ancho de banda a ciegas y guiada: una primera trama de ampliación de ancho de banda guiada que sigue a unas varias tramas de ampliación de ancho de banda a ciegas se amortigua un poco en su región de alta frecuencia, mientras que la amortiguación de alta frecuencia de una primera trama de ampliación de ancho de banda a ciegas después de alguna ampliación o ampliaciones de ancho de banda guiadas se reduce un poco.

En contenido de voz telefónico típico y en música popular, los experimentos han demostrado que alrededor del 13 % de todas las tramas de 20 ms utilizan la ampliación de ancho de banda guiada en LD-USAC. Por consiguiente, la tasa promedio de información secundaria de la ampliación de ancho de banda asciende a más o menos 2 bits por

trama o 0,1 kbit/s. Esto es mucho menos que las tasas de (e) SBR (consúltese, por ejemplo, la referencia [8]), o que cualquiera de las extensiones de ancho de banda del codificador de habla guiada a las que se hace referencia en el presente documento.

5 Por otro lado se observará que, como se sugiere como método opcional en la descripción etapa por etapa anteriormente en esta sección, la señalización de 1 bit de la decisión de modo de ampliación de ancho de banda para el decodificador puede evitarse si tanto el codificador como el decodificador pueden derivar dicha decisión a partir de la señal codificada en núcleo en un modo exacto de bits. Esto se puede lograr si el codificador selecciona el modo de ampliación de ancho de banda basándose en algunos rasgos derivados de la señal de núcleo decodificada localmente, ya que ésta es la única señal disponible en el decodificador. Suponiendo que no se produjo ningún error de transmisión en una determinada trama y que tanto el codificador como el decodificador determinan el modo de ampliación de ancho de banda a partir de exactamente los mismos rasgos de la señal de núcleo (tales como coeficientes de LPC cuantificados o estadísticas en el dominio del tiempo a partir de la señal residual decodificada como el coeficiente de cruce en cero, como se señaló anteriormente), la decisión de modo es idéntica en el codificador y el decodificador.

Las realizaciones de acuerdo con la invención superan un cierto dilema de calidad en cuanto a los códecs de banda ancha que se pueden observar a tasas de bits de 9-13 kbit/s. Se ha encontrado que, por una parte, dichas tasas ya son demasiado bajas para justificar la transmisión de incluso cantidades moderadas de datos de ampliación de ancho de banda, descartando los sistemas de ampliación de ancho de banda guiadas típicos con 1 kbit/s o más de información secundaria. Por otro lado, se ha encontrado que una ampliación de ancho de banda a ciegas factible tendrá un sonido significativamente peor en al menos algunos tipos de habla o de material de música debido a la incapacidad de una predicción adecuada de parámetros a partir de la señal de núcleo. Se ha encontrado que es por lo tanto deseable reducir la tasa de información secundaria de un esquema de ampliación de ancho de banda guiada a un nivel muy por debajo de 1 kbit/s, lo que permite su adopción incluso en la codificación de tasa de bits muy baja. El enfoque, que se utiliza en realizaciones de acuerdo con la invención, consiste en identificar los segmentos de las señales de entrada típicos que son mal reconstruidos o reconstruidos de manera subóptima por la ampliación de ancho de banda a ciegas, y para transmitir solamente para estos segmentos la información secundaria necesaria para mejorar la calidad de la reconstrucción de alta frecuencia a un nivel aceptable (o al menos a un nivel que se halle en el intervalo de la calidad del ancho de banda en dicha señal). En otras palabras: las partes de la señal de entrada de alta frecuencia que se recrean razonablemente bien por una ampliación de ancho de banda a ciegas deberían codificarse con muy poca o ninguna información secundaria de ampliación de ancho de banda, y solamente los pasajes sobre que un ancho de banda a ciegas degradaría la impresión general de la calidad códec deberían tener sus componentes de alta frecuencia reproducidos por una ampliación de ancho de banda guiada. Un diseño de este tipo para la ampliación de ancho de banda, que ajusta la tasa de información secundaria de una forma que se adapta a la señal, es el objeto de la presente invención y se denomina "ampliación de ancho de banda mínimamente guiada".

Las realizaciones de acuerdo con la invención superan múltiples enfoques de ampliación de ancho de banda que se han documentado en los últimos años (véase, por ejemplo, las referencias [1], [2], [3], [4], [5], [6], [7], [8], [9] y [10]). En general, todos estos o bien son completamente a ciegas o están completamente guiados en un punto de funcionamiento dado, independientemente de las características instantáneas de la señal de entrada. Además, todas las implementaciones de extensiones de ancho de banda a ciegas (consúltese, por ejemplo las referencias [1], [3], [4], [5], [9] y [10]) están optimizadas exclusivamente para señales de voz y como tales es poco probable que proporcionen una calidad satisfactoria en otras entradas tales como música (lo que inclusive fue objeto de comentarios en algunas publicaciones). Por último, la mayoría de las realizaciones de ampliación de ancho de banda convencionales son relativamente complejas, ya que emplean transformadas de Fourier, cálculos de filtros de LPC, o cuantificación vectorial de la información secundaria. Esto puede causar una desventaja en la adopción de la nueva tecnología de codificación en los mercados de telecomunicaciones móviles, dado que la mayoría de los dispositivos móviles proporcionan una potencia de cálculo muy limitada.

Como conclusión adicional, las realizaciones de acuerdo con la invención crean un codificador de audio o un método para la codificación de audio o un programa informático relacionado, como se ha descrito anteriormente.

55 En otras realizaciones de acuerdo con la invención se crea un decodificador de audio o un método de decodificación de audio o un programa informático relacionado, como se ha descrito anteriormente.

En realizaciones adicionales de la invención se crea una señal de audio codificada o un medio de almacenamiento en el que se ha almacenado la señal de audio codificada anteriormente mencionada.

60

9. Alternativas de implementación

Si bien se han descrito algunos aspectos en el contexto de un aparato, es evidente que estos aspectos también representan una descripción del método correspondiente, donde un bloque o dispositivo corresponde a una etapa de

procedimiento o un rasgo de una etapa de procedimiento. Análogamente, los aspectos descritos en el contexto de una etapa del método también representan una descripción de un bloque o elemento o aspecto correspondientes de un aparato. Algunas, o todas las etapas del método pueden ser ejecutadas por (o usarse) un aparato de hardware, tal como por ejemplo un microprocesador, una computadora programable o un circuito electrónico. En algunas realizaciones, una o más de las etapas más importantes del método pueden ejecutarse por un aparato de este tipo.

La señal de audio codificada de la invención puede almacenarse en un medio de almacenamiento digital o puede transmitirse en un medio de transmisión tal como un medio de transmisión inalámbrico o un medio de transmisión por cable, tal como Internet.

En función de determinados requisitos para la implementación, las realizaciones de la invención pueden implementarse en hardware o en software. La implementación puede realizarse utilizando un medio de almacenamiento digital, por ejemplo un disquete flexible, un DVD, un Blu-Ray, un CD, una ROM, una PROM, una EPROM, una EEPROM o una memoria FLASH, que tenga señales de control legibles electrónicamente almacenadas en el mismo, que cooperan (o que son capaces de cooperar) con un sistema informático programable de tal manera que se lleva a cabo el método respectivo. Por lo tanto, el medio de almacenamiento digital puede ser legible por ordenador.

Algunas realizaciones de acuerdo con la invención comprenden un portador de datos que tiene señales de control legibles electrónicamente, que son capaces de cooperar con un sistema informático programable, de tal manera que se lleva a cabo uno de los métodos descritos en el presente documento.

En general, las realizaciones de la presente invención pueden implementarse como un producto de programa informático con un código de programa, siendo el código de programa operativo para llevar a cabo uno de los métodos, cuando el producto de programa informático se ejecuta en una computadora. El código de programa puede almacenarse por ejemplo en un soporte legible por máquina.

Otras realizaciones comprenden el programa informático para llevar a cabo uno de los métodos anteriormente descritos, almacenado en un portador legible por máquina.

En otras palabras, una realización del método de acuerdo con la invención consiste por lo tanto en un programa informático que tiene un código de programa para llevar a cabo uno de los métodos descritos en el presente documento, cuando el programa informático se ejecuta en una computadora.

Por lo tanto, otra realización de los métodos de acuerdo con la invención consiste en un portador de datos (o un medio de almacenamiento digital, o un medio legible por ordenador) que comprende, grabado en el mismo, el programa informático para llevar a cabo uno de los métodos descritos en el presente documento. El portador de datos, el medio de almacenamiento digital o el medio grabado, son típicamente tangibles y/o no transitorios.

Otra realización del método de acuerdo con la invención es por lo tanto una corriente de datos o una secuencia de señales que representa el programa informático para llevar a cabo uno de los métodos descritos en el presente documento. La corriente de datos o la secuencia de señales pueden, por ejemplo, estar configurados para transferirse mediante una conexión de comunicación de datos, por ejemplo por intermedio de Internet.

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

Otra realización comprende una computadora que tiene instalada en la misma el programa informático para llevar a cabo uno de los métodos descritos en el presente documento.

Otra realización de acuerdo con la invención comprende un aparato o un sistema configurado para transferir (por ejemplo, electrónicamente u ópticamente) un programa informático para llevar a cabo uno de los métodos descritos en el presente documento en un receptor. El receptor puede ser, por ejemplo, una computadora, un dispositivo móvil, un dispositivo de memoria o similares. El aparato o sistema pueden, por ejemplo, comprender un servidor de archivos para transferir el programa informático al receptor.

En algunas realizaciones, es posible utilizar un dispositivo lógico programable (por ejemplo, un campo de matrices de puertas programables) para llevar a cabo algunas o todas las funcionalidades de los métodos descritos en el presente documento. En algunas realizaciones, un campo de matrices de puertas programables puede cooperar con un microprocesador para llevar a cabo uno de los métodos descritos en el presente documento. En términos generales, es preferible llevar a cabo los métodos mediante cualquier aparato de hardware.

El aparato descrito en el presente documento puede implementarse mediante un aparato de hardware, o mediante una computadora, o utilizando una combinación de un aparato de hardware y una computadora.

Los métodos descritos en el presente documento pueden implementarse mediante un aparato de hardware, o utilizando una computadora, o utilizando una combinación de un aparato de hardware y una computadora.

- 5 Las realizaciones anteriormente descritas son meramente ilustrativas de los principios de la presente invención. Se da por entendido que las modificaciones y variaciones de las disposiciones y de los detalles descritos en el presente documento serán evidentes para otros expertos en la materia. Por lo tanto, la finalidad es que se limite solamente por los alcances de las reivindicaciones adjuntas y no por los detalles específicos presentados a modo de descripción y explicación de las realizaciones en el presente documento.

10

Referencias

- [1] B. Bessette et al., "The Adaptive Multi-rate Wideband Speech Codec (AMR-WB)", IEEE Trans. on Speech and Audio Processing, Vol. 10, N.º 8, noviembre de 2002.
- 15 [2] B. Geiser et al., "Bandwidth Extension for Hierarchical Speech and Audio Coding in ITU-T Rec. G.729.1", IEEE Trans. on Audio, Speech, and Language Processing, Vol. 15, N.º 8, noviembre de 2007.
- [3] B. Iser, W. Minker y G. Schmidt, Bandwidth Extension of Speech Signals, Springer Lecture Notes in Electrical Engineering, Vol. 13, Nueva York, 2008.
- 20 [4] M. Jelinek y R. Salami, "Wideband Speech Coding Advances in VBR-WB Standard", IEEE Trans. on Audio, Speech, and Language Processing, Vol. 15, N.º 4 de mayo de 2007.
- [5] I. Katsir, I. Cohen y D. Malah, "Speech Bandwidth Extension Based on Speech Phonetic Content and Speaker Vocal Tract Shape Estimation", in Proc. EUSIPCO 2011, Barcelona, España, septiembre de 2011.
- [6] E. Larsen y R. M. Aarts, Audio Bandwidth Extension: Application of Psycho-acoustics, Signal Processing and Loudspeaker Design, Wiley, Nueva York, 2004.
- 25 [7] J. Mäkinen et al., "AMR-WB+: A New Audio Coding Standard for 3rd Generation Mobile Audio Services", in Proc. ICASSP 2005, Filadelfia, Estados Unidos, marzo de 2005.
- [8] M. Neuendorf et al., "MPEG Unified Speech and Audio Coding - The ISO/MPEG Standard for High-Efficiency Audio Coding of All Content Types", in Proc. 132nd AES Convention, Budapest, Hungría, abril de 2012. También aparece en el diario de la AES, 2013.
- 30 [9] H. Pulakka y P. Alku, "Bandwidth Extension of Telephone Speech Using a Neural Network and a Filter Bank Implementation for Highband Mel Spectrum", IEEE Trans. on Audio, Speech, and Language Processing, Vol. 19, No. 7, septiembre de 2011.
- [10] T. Vaillancourt et al., "ITU-T EV-VBR: A Robust 8-32 kbit/s Scalable Coder for Error Prone Telecommunications Channels", in Proc. EUSIPCO 2008, Lausana, Suiza, agosto de 2008.
- 35 [11] L. Miao et al., "G.711.1 Annex D and G.722 Annex B: New ITU-T Superwideband codecs", in Proc. ICASSP 2011, Praga, República Checa, mayo de 2011.

REIVINDICACIONES

1. Un decodificador de audio (400; 500) para proporcionar una información de audio decodificada (412; 512) basándose en una información de audio codificada (410; 510), comprendiendo el decodificador de audio:
 5 un decodificador de baja frecuencia (420; 520) configurado para decodificar una representación codificada de una porción de baja frecuencia para obtener una representación decodificada (422; 522) de la porción de baja frecuencia; y
 una ampliación de ancho de banda (430; 530) configurada para obtener una señal de ampliación de ancho de banda (432; 532) utilizando una ampliación de ancho de banda a ciegas para porciones de un contenido de audio para las
 10 que no hay parámetros de ampliación de ancho de banda incluidos en la información de audio codificada, y para obtener la señal de ampliación de ancho de banda utilizando una ampliación de ancho de banda guiada por parámetros para porciones del contenido de audio para las que hay parámetros de ampliación de ancho de banda incluidos en la información de audio codificada,
 en donde el decodificador de audio se configura para decidir si usar una ampliación de ancho de banda a ciegas o
 15 una ampliación de ancho de banda guiada por parámetros basándose en la representación codificada de la porción de baja frecuencia sin evaluar una bandera de señalización de modo de ampliación de ancho de banda.
2. El decodificador de audio (400; 500) de acuerdo con la reivindicación 1, en donde el decodificador de audio se configura para decidir si obtener la señal de ampliación de ancho de banda usando una ampliación de ancho de banda a ciegas o usando una ampliación de ancho de banda guiada por parámetros basándose en tramas.
3. El decodificador de audio (400; 500) de acuerdo con la reivindicación 1 o 2, en donde el decodificador de audio se configura para cambiar entre un uso de una ampliación de ancho de banda a ciegas y una ampliación de ancho de banda guiada por parámetros dentro de una pieza contigua de contenido de audio.
- 25 4. El decodificador de audio (400; 500) de acuerdo con una de las reivindicaciones 1 a 3, en donde el decodificador de audio se configura para evaluar banderas incluidas en la información de audio codificada para diferentes porciones del contenido de audio, para decidir si usar una ampliación de ancho de banda a ciegas o una ampliación de ancho de banda guiada por parámetros.
- 30 5. El decodificador de audio (400; 500) de acuerdo con una de las reivindicaciones 1 a 4, en donde el decodificador de audio se configura para decidir si usar una ampliación de ancho de banda a ciegas o una ampliación de ancho de banda guiada por parámetros basándose en una o más características de la representación decodificada de la porción de baja frecuencia.
- 35 6. El decodificador de audio (400; 500) de acuerdo con una de las reivindicaciones 1 a 5, en donde el decodificador de audio se configura para decidir si usar una ampliación de ancho de banda a ciegas o una ampliación de ancho de banda guiada por parámetros basándose en coeficientes de predicción lineal y/o basándose en estadísticas de dominio de tiempo de la representación decodificada de la porción de baja frecuencia.
- 40 7. El decodificador de audio (400; 500) de acuerdo con una de las reivindicaciones 1 a 6, en donde la ampliación de ancho de banda se configura para obtener la señal de ampliación de ancho de banda usando una o más características de la representación decodificada de la porción de baja frecuencia y/o usando uno o más parámetros del decodificador de baja frecuencia para porciones temporales del contenido de audio de entrada para el que no se incluye ningún parámetro de ampliación de ancho de banda en la información de audio codificado.
- 45 8. El decodificador de audio (400; 500) de acuerdo con una de las reivindicaciones 1 a 7, en donde la ampliación de ancho de banda se configura para obtener la señal de ampliación de ancho de banda usando una información de centroide espectral y/o usando una información de energía, y/o usando una información de inclinación, y/o usando coeficientes de filtro para porciones temporales del contenido de audio de entrada para el que no se incluyen parámetros de ampliación de ancho de banda en la información de audio codificada.
- 50 9. El decodificador de audio (400; 500) de acuerdo con una de las reivindicaciones 1 a 8, en donde la ampliación de ancho de banda se configura para obtener la señal de ampliación de ancho de banda usando parámetros de corriente de bits que describen una envolvente espectral de una porción de alta frecuencia para porciones temporales del contenido de audio para las que hay parámetros de ampliación de ancho de banda incluidos en la información de audio codificada.
- 55 10. El decodificador de audio (400; 500) de acuerdo con la reivindicación 9, en donde la ampliación de ancho de banda se configura para evaluar entre tres y cinco parámetros de corriente de bits que describen intensidades de las porciones de señal de alta frecuencia con anchos de banda entre 300 Hz y 500 Hz, para obtener la señal de ampliación de ancho de banda.
- 60 11. El decodificador de audio (400; 500) de acuerdo con la reivindicación 10, en donde los entre tres y cinco

parámetros de corriente de bits que describen intensidades de porciones de señal de alta frecuencia se cuantifican de forma escalar con 2 o 3 bits de resolución, tal que existen entre 6 y 15 bits de parámetros de moldeo espectral de ampliación de ancho de banda por trama de audio.

- 5 12. El decodificador de audio (400; 500) de acuerdo con una de las reivindicaciones 1 a 11, en donde la ampliación de ancho de banda se configura para realizar un alisamiento de energías de la señal de ampliación de ancho de banda cuando cambia de la ampliación de ancho de banda a ciegas a la ampliación de ancho de banda guiada por parámetros y/o cuando cambia de la ampliación de ancho de banda guiada por parámetros a la ampliación de ancho de banda a ciegas.
- 10 13. El decodificador de audio (400; 500) de acuerdo con la reivindicación 12, en donde la ampliación de ancho de banda está configurada para amortiguar una porción de alta frecuencia de la señal de ampliación de ancho de banda para una porción del contenido de audio al que se aplica una ampliación de ancho de banda guiada por parámetros después de una porción del contenido de audio a la que se aplica una ampliación de ancho de banda a ciegas; y
- 15 en donde la ampliación de ancho de banda está configurada para reducir una amortiguación o para aumentar un nivel para una porción de alta frecuencia de la señal de ampliación de ancho de banda para una porción del contenido de audio al que se aplica una ampliación de ancho de banda a ciegas después de una porción del contenido de audio al que aplica una ampliación de ancho de banda guiada por parámetros.
- 20 14. Un método (700) para proporcionar una información de audio decodificada basándose en una información de audio codificada, comprendiendo el método:
- decodificar (710) una representación codificada de una porción de baja frecuencia para obtener una representación decodificada de la porción de baja frecuencia; y
- 25 obtener (720) una señal de ampliación de ancho de banda utilizando una ampliación de ancho de banda a ciegas para porciones de un contenido de audio para las que no hay parámetros de ampliación de ancho de banda incluidos en la información de audio codificada; y
- 30 obtener (730) la señal de ampliación de ancho de banda utilizando una ampliación de ancho de banda guiada por parámetros para porciones del contenido de audio para las que hay parámetros de ampliación de ancho de banda incluidos en la información de audio codificada;
- en donde el método comprende decidir si usar una ampliación de ancho de banda a ciegas o una ampliación de ancho de banda guiada por parámetros basándose en la representación codificada de la porción de baja frecuencia sin evaluar una bandera de señalización de modo de ampliación de ancho de banda.
- 35 15. Un programa informático para llevar a cabo el método de acuerdo con la reivindicación 14, cuando el programa informático se ejecuta en una computadora.

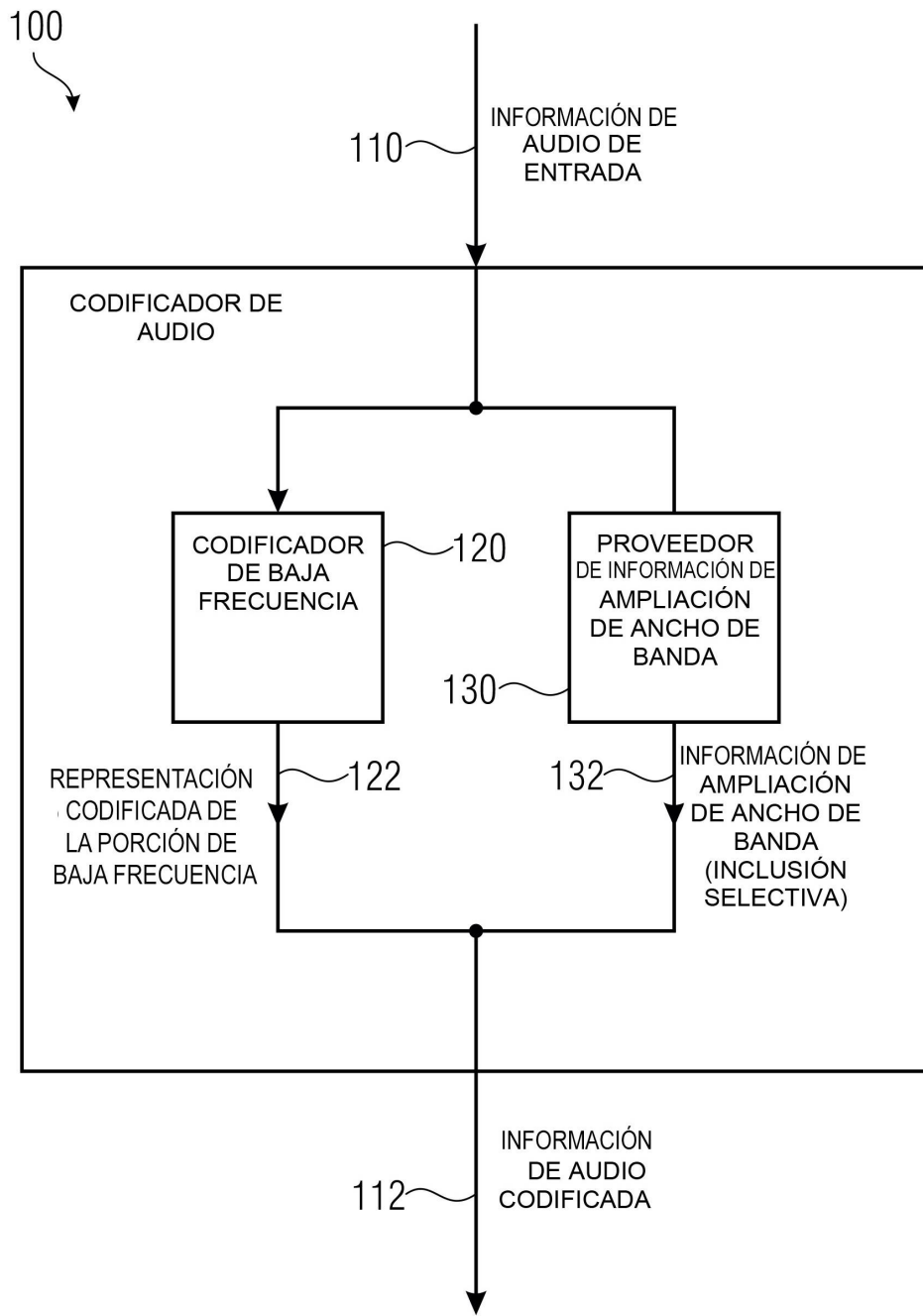


FIGURA 1

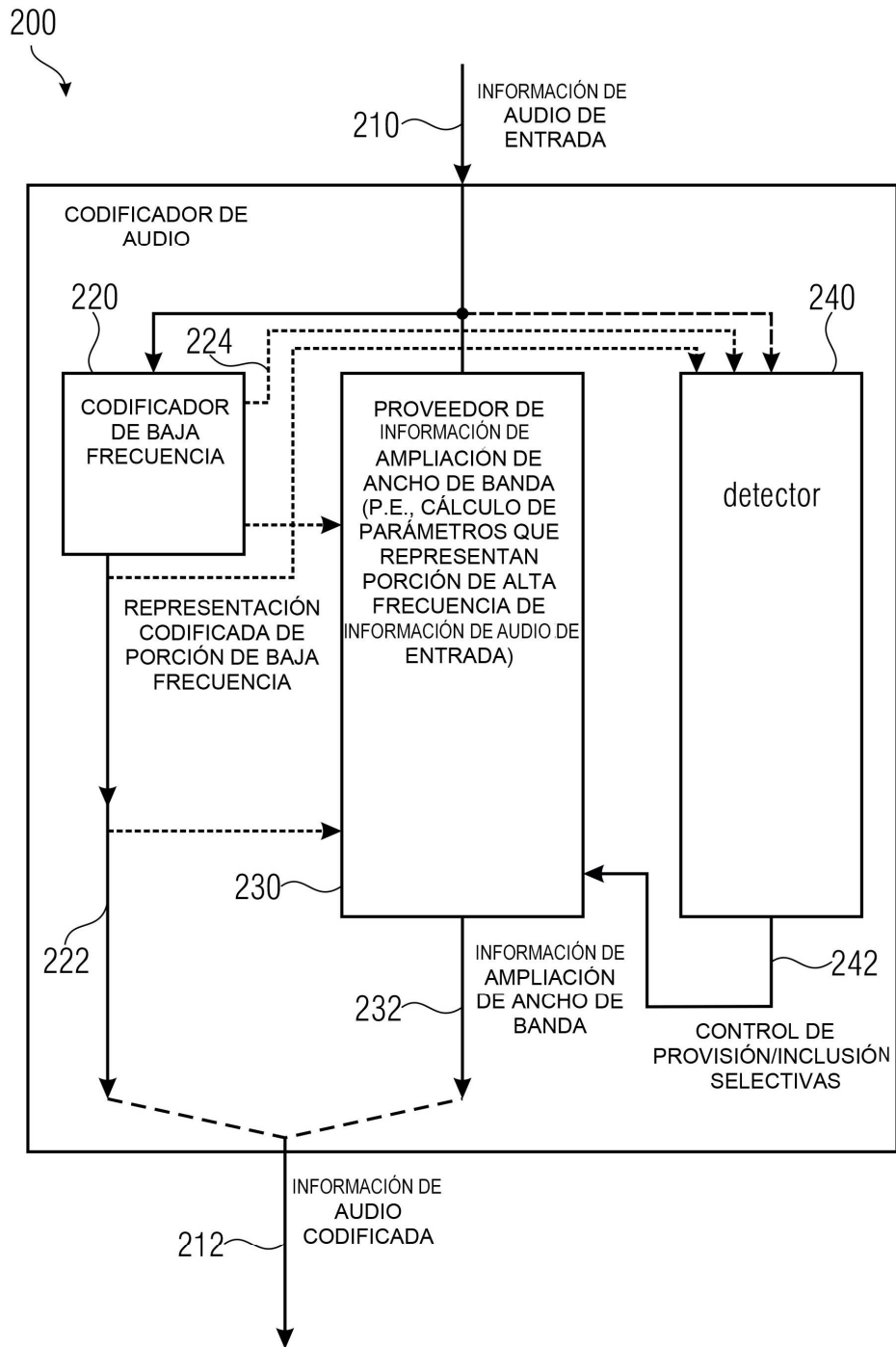


FIGURA 2

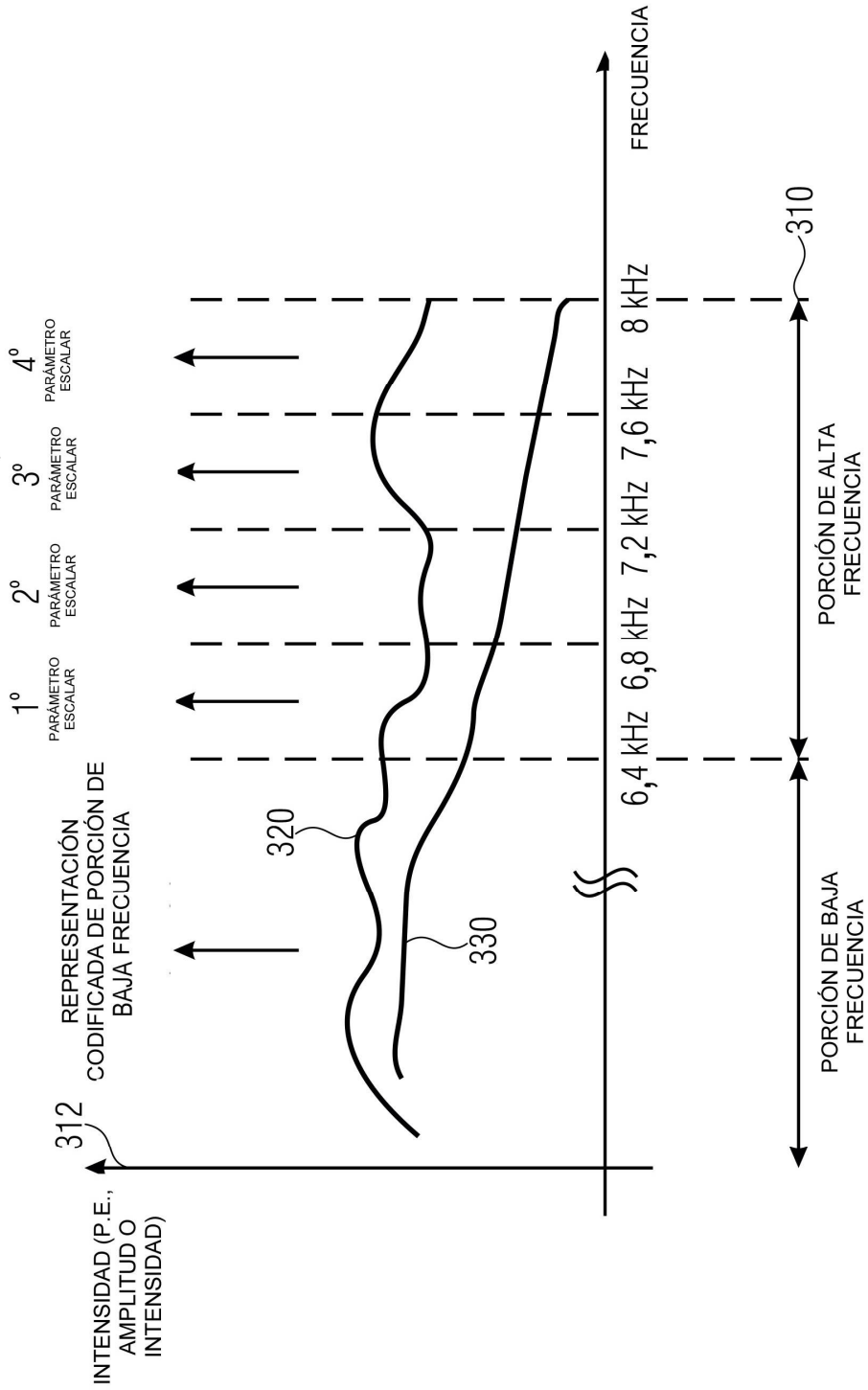


FIGURA 3

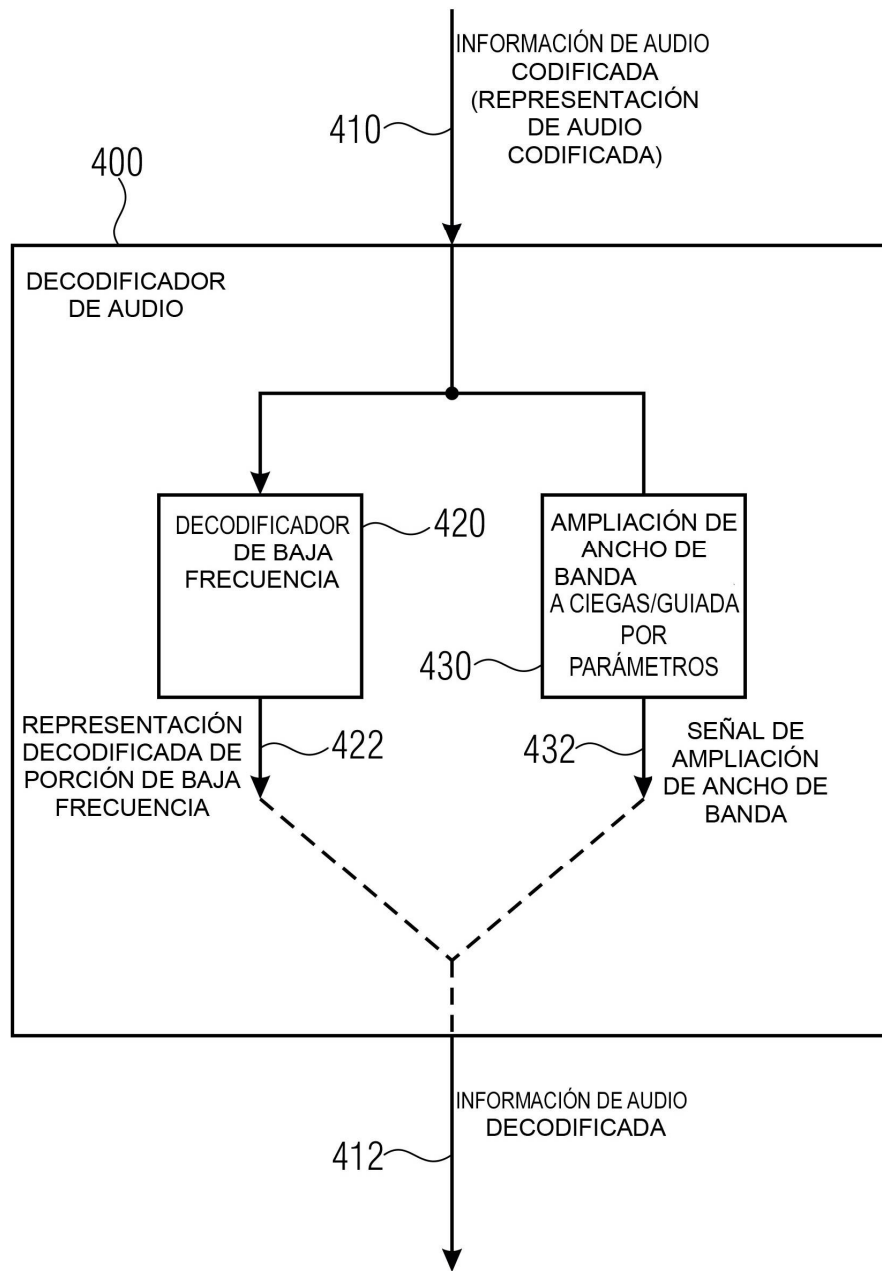


FIGURA 4

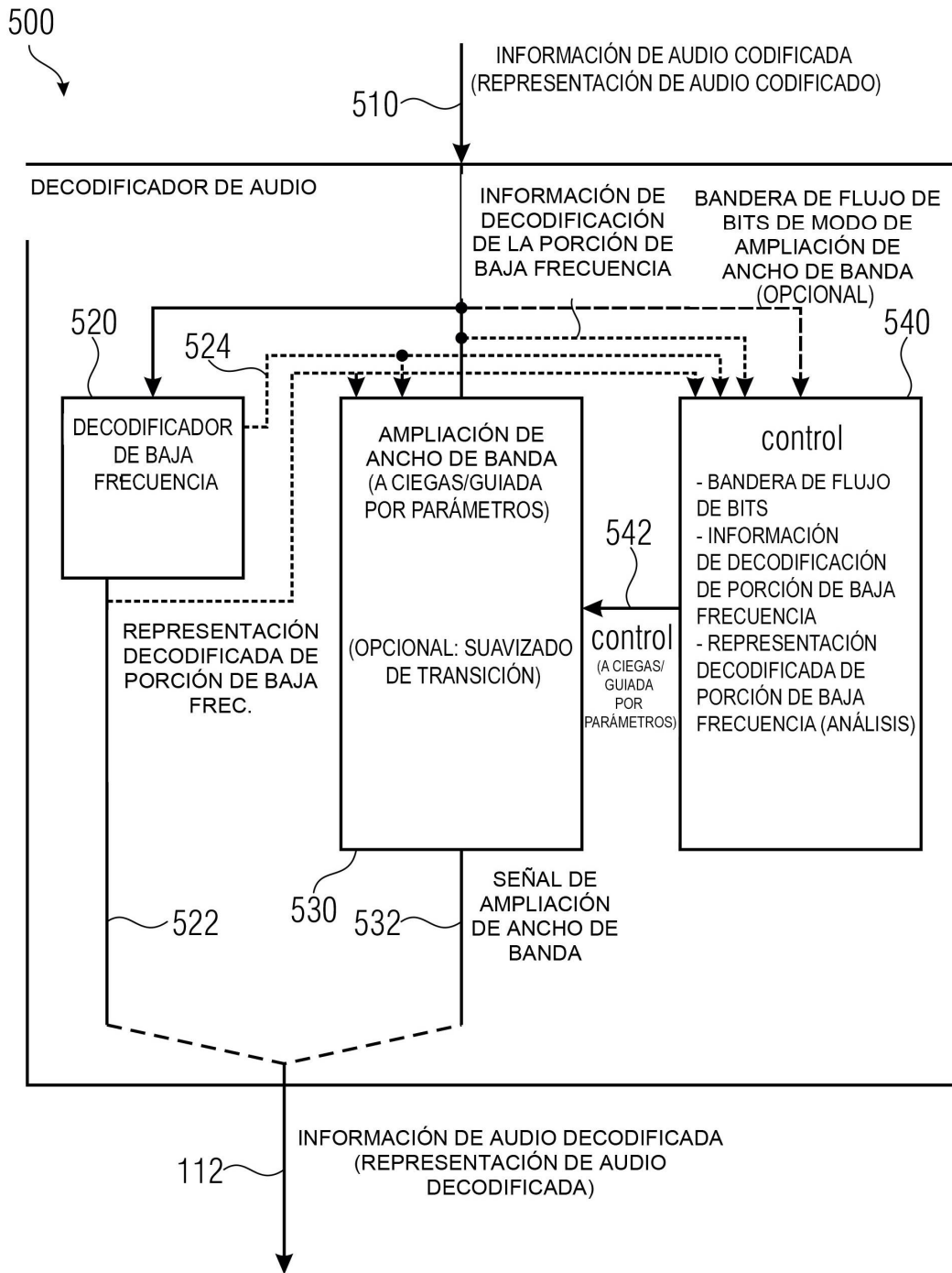


FIGURA 5

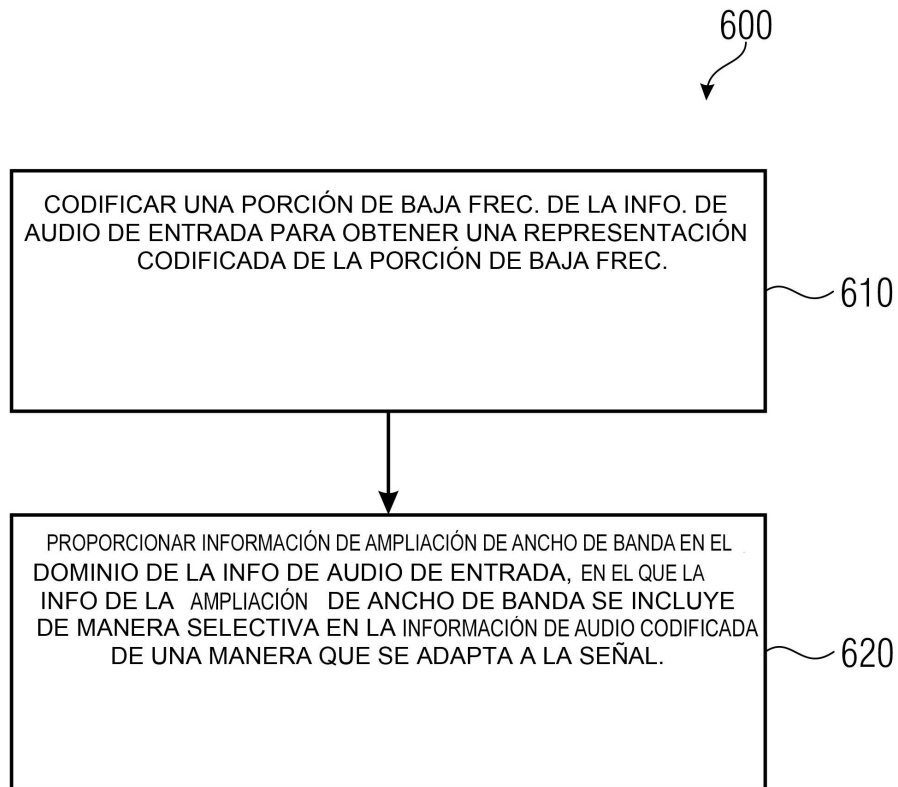


FIGURA 6

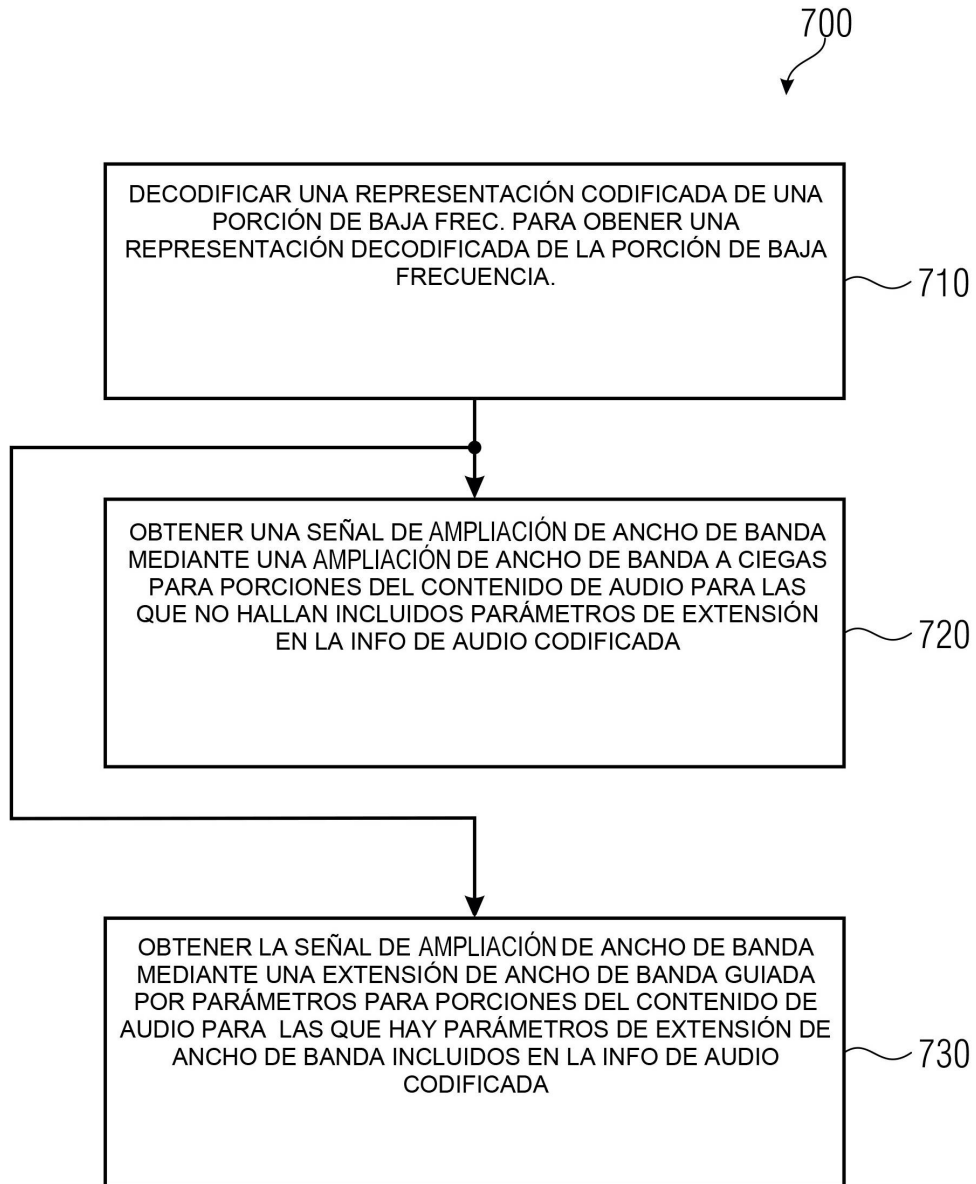


FIGURA 7

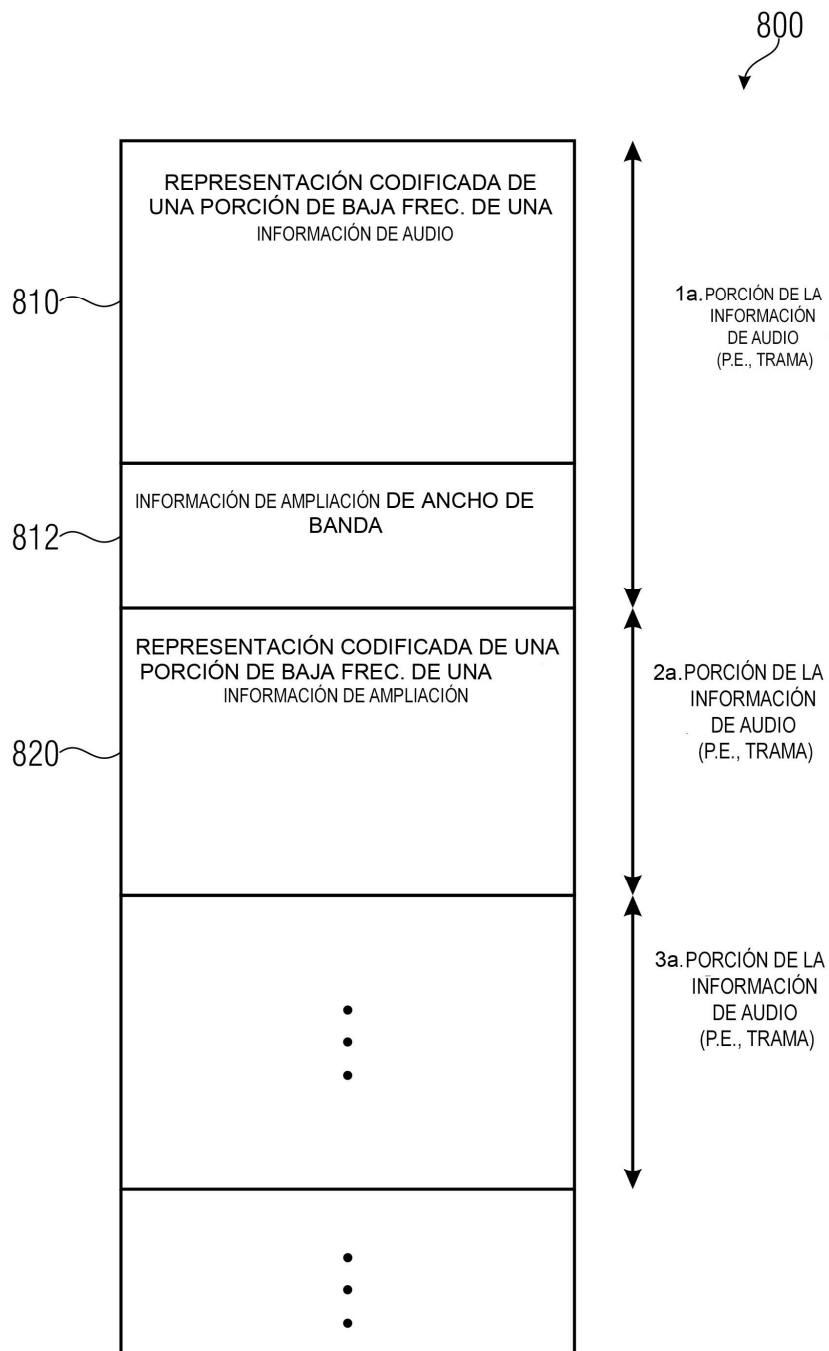


FIGURA 8