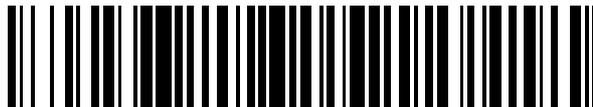


19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 690 256**

51 Int. Cl.:

G10L 19/20 (2013.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

86 Fecha de presentación y número de la solicitud internacional: **23.07.2015 PCT/EP2015/066953**

87 Fecha y número de publicación internacional: **04.02.2016 WO16016105**

96 Fecha de presentación y número de la solicitud europea: **23.07.2015 E 15741215 (6)**

97 Fecha y número de publicación de la concesión europea: **25.07.2018 EP 3175453**

54 Título: **Decodificador de audio, método y programa de computación que utiliza una respuesta de entrada cero para obtener una transición suave**

30 Prioridad:

28.07.2014 EP 14178830

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

20.11.2018

73 Titular/es:

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

72 Inventor/es:

**RAVELLI, EMMANUEL;
FUCHS, GUILLAUME;
DISCH, SASCHA;
MULTRUS, MARKUS;
PIETRZYK, GRZEGORZ y
SCHUBERT, BENJAMIN**

74 Agente/Representante:

ARIZTI ACHA, Monica

ES 2 690 256 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).

Decodificador de audio, método y programa de computación que utiliza una respuesta de entrada cero para obtener una transición suave

DESCRIPCIÓN

5 **1. Campo Técnico**

Una forma de realización de acuerdo con la invención se relaciona con un decodificador de audio para emitir una información de audio decodificada sobre la base de una información de audio codificada.

10 Otra forma de realización de acuerdo con la invención se relaciona con un método para emitir una información de audio decodificada sobre la base de una información de audio codificada.

15 Otra forma de realización de acuerdo con la invención se relaciona con un programa de computación para ejecutar dicho método.

En general, las formas de realización de acuerdo con la invención se relacionan con el tratamiento de una transición de un códec CELP a un códec basado en MDCT en la codificación de audio conmutada.

20 **2. Antecedentes de la invención**

En los últimos años ha existido una creciente demanda de transmisión y almacenamiento de información de audio codificada. También existe una creciente demanda de codificación de audio y decodificación de audio de señales de audio que comprenden tanto voz como audio general (como, por ejemplo, música, ruido de fondo y demás).

25 Para mejorar la calidad de la codificación y también para mejorar la eficiencia de la tasa de bits, se han introducido códecs de audio conmutado (o con conmutación) que conmutan entre diferentes esquemas de codificación de tal manera que, por ejemplo, se codifique una primera trama utilizando un primer concepto de codificación (por ejemplo, un concepto de codificación basado en CELP, por sus siglas en inglés Code Excited Linear Prediction, Predicción Lineal Excitada por Código), y de tal manera que se codifique una segunda trama de audio subsiguiente utilizando un segundo concepto de codificación (por ejemplo, un concepto de codificación basado en MDCT (por sus siglas en inglés, Modified Discreet Cosine Transform, Transformada Discreta de Coseno Modificada)). Dicho de otro modo, puede haber una conmutación entre una codificación en el dominio de codificación de predicción lineal (por ejemplo, empleando un concepto de codificación basada en CELP) y una codificación en el dominio de la frecuencia (por ejemplo, una codificación que se basa en una transformada del dominio del tiempo al dominio de la frecuencia o una transformada del dominio de la frecuencia al dominio del tiempo, como por ejemplo una transformada FFT (por sus siglas en inglés, FAST FOURIER TRANSFORM, transformada rápida de Fourier), una transformada FFT inversa, una transformada MDCT (por sus siglas en inglés, Modified Discreet Cosine Transform, Transformada Discreta de Coseno Modificada) o una transformada MDCT inversa). Por ejemplo, el primer concepto de codificación puede ser un concepto de codificación basado en CELP, un concepto de codificación basado en ACELP (por sus siglas en inglés Algebraic Code Excited Linear Prediction, Predicción Lineal Basada en Código Algebraico), un concepto de codificación basado en transformada del dominio de predicción lineal excitada por código o similar. El segundo concepto de codificación puede ser, por ejemplo, un concepto de codificación basado en FFT, un concepto de codificación basado en MDCT, un concepto de codificación basado en AAC (por sus siglas en inglés, Advanced Audio Coding, Codificación Avanzada de Audio) o un concepto de codificación que se pueda considerar como concepto sucesor del concepto de codificación basado en AAC.

50 A continuación se describen algunos ejemplos de codificadores (codificadores y decodificadores) de audio convencionales.

Los códecs de audio conmutados, como por ejemplo, MPEG USAC (por sus siglas en inglés MPEG USAC, Moving Picture Experts Group Unified Speech and Audio Coding, Codificación Unificada de Voz y Audio por el Grupo de Expertos en Imágenes en Movimiento), se basan en dos esquemas principales de codificación de audio. Un esquema de codificación es, por ejemplo, un códec CELP, destinado a las señales de voz. El otro esquema de codificación es, por ejemplo, un códec basado en MDCT (simplemente denominado MDCT en lo sucesivo), dirigido a todas las demás señales de audio (por ejemplo, música, ruido de fondo). En cuando a las señales de contenido mixto (por ejemplo, voz sobre música), el codificador (y en consecuencia, también el decodificador) con frecuencia conmuta entre los dos esquemas de codificación. Es entonces necesario evitar todo defecto (por ejemplo, un click debido a una discontinuidad) al conmutar de un modo (o esquema de codificación) a otro.

60 Los códecs de audio conmutados pueden comprender, por ejemplo, problemas causados por las transiciones de CELP a MDCT.

Las transiciones de CELP a MDCT por lo general introducen dos problemas. Se puede introducir un "aliasing"

(solapamiento) debido a la trama de MDCT anterior faltante. Se puede introducir una discontinuidad en el borde entre la trama de CELP y la trama de MDCT, debido a la naturaleza de codificación de forma de onda no perfecta de los dos esquemas de codificación que operan a tasas de bits bajas/medias.

- 5 Ya existen varias estrategias para solucionar los problemas introducidos por las transiciones de CELP a MDCT, y se las describe a continuación.

Una estrategia posible es la descrita en el artículo “Efficient cross-fade las ventanas for transitions between LPC-based and non-LPC based audio coding” de Jeremie Lecomte, Philippe Gournay, Ralf Geiger, Bruno Bessette y Max Neuendorf (presentado en la 126a Convención de AES, mayo de 2009, documento 771). Este artículo describe una estrategia, en la sección 4.4.2 “modo ACELP a no LPD”. También se hace referencia, por ejemplo, a la Fig. 8 de dicho artículo. El problema de aliasing se soluciona en primer lugar aumentando la longitud de la MDCT (en este caso de 1024 a 1152) de tal manera que el punto de pliegue izquierdo de la MDCT se corra a la izquierda del borde entre las tramas de CELP y de MDCT, luego cambiando la parte izquierda de la ventana de MDCT de tal manera que se reduzca el solapamiento y, por último, introduciendo artificialmente el aliasing faltante utilizando la señal de CELP y una operación de solapamiento y suma. El problema de la discontinuidad se soluciona al mismo tiempo mediante la operación de solapamiento y suma.

Esta estrategia da buen resultado, pero tiene la desventaja de introducir un retardo en el decodificador CELP, retardo que es igual a la longitud del solapamiento (en este caso: 128 muestras).

Se describe otra estrategia en el documento US 8.725.503 B2, con fecha 13 de mayo de 2014 y titulado “Forward time domain aliasing cancellation with application in weighted or original signal domain” de Bruno Bessette.

25 En esta estrategia, no se modifica la longitud de la MDCT (ni la forma de ventana de MDCT). El problema de solapamiento se soluciona en este caso codificando la señal de corrección de aliasing con un codificador separado basado en transformadas. Se envían bits de información lateral adicionales al flujo de bits. El decodificador reconstruye la señal de corrección de aliasing y la suma a la trama de MDCT decodificada. Además, se utiliza la respuesta de entrada cero (ZIR, por sus siglas en inglés) del filtro de síntesis de CELP para reducir la amplitud de la señal de corrección de aliasing y para mejorar la eficiencia de codificación. La ZIR también contribuye a reducir significativamente el problema de las discontinuidades.

Esta estrategia también da buen resultado, aunque la desventaja es que requiere una cantidad significativa de información lateral adicional y el número de bits necesario es generalmente variable, lo que no es apropiado para un códec de tasa de bits constante.

Otra estrategia es la descrita en la solicitud de patente de Estados Unidos US 2013/0289981 A1 del 31 de octubre de 2013 titulada “Low-delay sound-encoding alternating between predictive encoding and transform encoding” de Stephane Ragot, Balazs Kovesi y Pierre Berthet. De acuerdo con dicha estrategia, la MDCT no se modifica, pero sí se cambia la parte izquierda de la ventana de MDCT para reducir la longitud del solapamiento. Para solucionar el problema de “aliasing” (alteración o defecto por solapamiento), se codifica el comienzo de la trama de MDCT utilizando un códec CELP, y luego se utiliza la señal del CELP para cancelar el aliasing, ya sea reemplazando por completo la señal de MDCT o introduciendo artificialmente el componente de aliasing faltante (de manera similar al artículo antes mencionado de Jeremie Lecomte et al.). El problema de la discontinuidad se soluciona mediante la operación de solapamiento y suma si se usa una estrategia similar a la del artículo de Jeremie Lecomte et al.; de lo contrario se soluciona mediante una sencilla operación de atenuación cruzada entre la señal de CELP y la señal de MDCT.

Como ocurre con el documento US 8.725.503 B2, esta estrategia generalmente da buen resultado, aunque tiene la desventaja de que requiere una cantidad significativa de información lateral, introducida por la CELP adicional.

En vista de las soluciones convencionales antes descritas, existe el deseo de contar con un concepto que comprenda características mejoradas (por ejemplo, una compensación mejorada entre el consumo de tasa de bits, retardo y complejidad) conmutando entre los diferentes modos de codificación.

55

3. Sumario de la invención

Una forma de realización de acuerdo con la invención crea un decodificador de audio para emitir una información de audio decodificada sobre la base de una información de audio codificada. El decodificador de audio comprende un decodificador en el dominio de la predicción lineal configurado para emitir una primera información de audio decodificada sobre la base de una trama de audio codificada en el dominio de la predicción lineal y un decodificador en el dominio de la frecuencia configurado para emitir una segunda información de audio decodificada sobre la base de una trama de audio codificada en el dominio de la frecuencia. El decodificador de audio comprende asimismo un procesador de transición. El procesador de transición está configurado para obtener una respuesta de entrada cero

60

de un filtrado de predicción lineal, donde el estado inicial del filtrado de predicción lineal se define de conformidad con la primera información de audio decodificada y la segunda información de audio decodificada. El procesador de transición también está configurado para modificar la segunda información de audio decodificada, que se emite sobre la base de una trama de audio codificada en el dominio de la frecuencia posterior a una trama de audio codificada en el dominio de la predicción lineal, dependiendo de la respuesta de entrada cero, para obtener una transición suave entre la primera información de audio decodificada y la segunda información de audio decodificada modificada.

Este decodificador de audio se basa en el hallazgo de que se puede obtener una transición suave entre una trama de audio codificada en el dominio de la predicción lineal y una trama de audio subsiguiente codificada en el dominio de la frecuencia utilizando una respuesta de entrada cero de un filtro de predicción lineal para modificar la segunda información de audio decodificada, con la condición de que el estado inicial del filtrado de predicción lineal considere tanto la primera información de audio decodificada como la segunda información de audio decodificada. En consecuencia, la segunda información de audio decodificada puede ser adaptada (modificada) de tal manera que el comienzo de la segunda información de audio decodificada modificada sea similar al final de la primera información de audio decodificada, lo que contribuye a reducir, o incluso a evitar, discontinuidades sustanciales entre la primera trama de audio y la segunda trama de audio. En comparación con el decodificador de audio antes descrito, el concepto es aplicable, en general, incluso si la segunda información de audio decodificada no comprende aliasing en absoluto. Más aún, se debe tener en cuenta que el término "filtrado de predicción lineal" puede designar tanto una aplicación única de un filtro de predicción lineal como múltiples aplicaciones de filtros de predicción lineal, donde cabe señalar que por lo general una única aplicación de filtrado de predicción lineal es equivalente a múltiples aplicaciones de filtros de predicción lineal idénticos, puesto que los filtros de predicción lineal son típicamente lineales.

Para concluir, el decodificador de audio antes citado permite obtener una transición suave entre una primera trama de audio codificada en un dominio de la predicción lineal y una segunda trama de audio subsiguiente codificada en el dominio de la frecuencia (o dominio de la transformada), donde no se introduce retardo alguno y donde el esfuerzo informático es comparativamente bajo.

Otra forma de realización de acuerdo con la invención genera un decodificador de audio para emitir una información de audio decodificada sobre la base de una información de audio codificada. El decodificador de audio comprende un decodificador en el dominio de la predicción lineal configurado para emitir una primera información de audio decodificada sobre la base de una trama de audio codificada en un dominio de la predicción lineal (o, de modo equivalente, en una representación en el dominio de la predicción lineal). El decodificador de audio comprende asimismo un decodificador en el dominio de la frecuencia configurado para emitir una segunda información de audio decodificada sobre la base de una trama de audio codificada en un dominio de la frecuencia (o, de modo equivalente, en una representación en el dominio de la frecuencia). El decodificador de audio comprende asimismo un procesador de transición. El procesador de transición está configurado para obtener una primera respuesta de entrada cero de un filtro de predicción lineal en respuesta a un primer estado inicial del filtro de predicción lineal definido por la primera información de audio decodificada, y para obtener una segunda respuesta de entrada cero del filtro de predicción lineal en respuesta a un segundo estado inicial del filtro de predicción lineal definido por una versión modificada de la primera información de audio decodificada, que se emite con un aliasing artificial, y que comprende la contribución de una porción de la segunda información de audio decodificada. Por otro lado, el procesador de transición está configurado para obtener una respuesta combinada de entrada cero del filtro de predicción lineal en respuesta a un estado inicial del filtro de predicción lineal definido por una combinación de la primera información de audio decodificada y de una versión modificada de la primera información de audio decodificada que se emite con un aliasing artificial, y que comprende la contribución de una porción de la segunda información de audio decodificada. El procesador de transición también está configurado para modificar la segunda información de audio decodificada, que se emite sobre la base de una trama de audio codificada en el dominio de la frecuencia que sigue a una trama de audio codificada en el dominio de la predicción lineal, dependiendo de la primera respuesta de entrada cero y la segunda respuesta de entrada cero o dependiendo de la respuesta combinada de entrada cero, para obtener una transición suave entre la primera información de audio decodificada y la segunda información de audio decodificada modificada.

Esta forma de realización de acuerdo con la invención se basa en el hallazgo de que se puede obtener una transición suave entre una trama de audio codificada en el dominio de la predicción lineal y una trama de audio subsiguiente codificada en el dominio de la frecuencia (o, en general, en el dominio de la transformada) modificando la segunda información de audio decodificada sobre la base de una señal que es una respuesta de entrada cero de un filtro de predicción lineal, un estado inicial del cual es definido tanto por la primera información de audio decodificada como por la segunda información de audio decodificada. Se puede utilizar una señal de salida de ese tipo de filtro de predicción lineal para adaptar la segunda información de audio decodificada (por ejemplo, una porción inicial de la segunda información de audio decodificada, que sigue inmediatamente a la transición entre la primera trama de audio y la segunda trama de audio), de tal manera que haya una transición suave entre la primera información de audio decodificada (asociada a una trama de audio codificada en el dominio de la predicción lineal) y

la segunda información de audio decodificada modificada (asociada a una trama de audio codificada en el dominio de la frecuencia o en el dominio de la transformada) sin necesidad de enmendar la primera información de audio decodificada.

5 Se ha descubierto que la respuesta de entrada cero del filtro de predicción lineal es muy adecuada para emitir una transición suave, puesto que el estado inicial del filtro de predicción lineal se basa en la primera información de audio decodificada y la segunda información de audio decodificada, donde se compensa un aliasing incluido en la segunda información de audio decodificada con el aliasing artificial, que se introduce en la versión modificada de la primera información de audio decodificada.

10 Además se ha encontrado que no es necesario retardo alguno de la decodificación mediante la modificación de la segunda información de audio decodificada sobre la base de la primera respuesta de entrada cero y la segunda respuesta de entrada cero, o dependiendo de la respuesta de entrada cero combinada, pero dejando sin cambios la primera información de audio decodificada, ya que la primera respuesta de entrada cero y la segunda respuesta de entrada cero, o la respuesta de entrada cero combinada, son muy adecuadas para alisar la transición entre la trama de audio codificada en el dominio de la predicción lineal y la trama de audio subsiguiente codificada en el dominio de la frecuencia (o en el dominio de la transformada) sin cambiar la primera información de audio decodificada, dado que la primera respuesta de entrada cero y la segunda respuesta de entrada cero, o la respuesta de entrada cero combinada, modifican la segunda información de audio decodificada de tal manera que la segunda información de audio decodificada sea sustancialmente similar a la primera información de audio decodificada por lo menos en la transición entre la trama de audio codificada en el dominio de la predicción lineal y la trama de audio subsiguiente codificada en el dominio de la frecuencia.

25 Para concluir, la forma de realización antes descrita de acuerdo con la presente invención permite producir una transición suave entre una trama de audio codificada en el dominio de codificación de predicción lineal y una trama de audio subsiguiente codificada en el dominio de la frecuencia (o en el dominio de la transformada), donde se evita la introducción de un retardo adicional puesto que solo se modifica la segunda información de audio decodificada (asociada a la trama de audio subsiguiente codificada en el dominio de la frecuencia), y donde se puede obtener una buena calidad de la transición (sin defectos sustanciales) mediante el uso de la primera respuesta de entrada cero y la segunda respuesta de entrada cero, o la respuesta de entrada cero combinada, lo que da lugar a la consideración tanto de la primera información de audio decodificada como de la segunda información de audio.

30 En una forma de realización preferida, el decodificador en el dominio de la frecuencia está configurado para ejecutar una transformada solapada inversa, de tal manera que la segunda información de audio decodificada comprenda un aliasing. Se ha descubierto que los conceptos de la invención antes citados actúan con especial efectividad incluso en caso de que el decodificador en el dominio de la frecuencia (o el decodificador en el dominio de la transformada) introduzca el aliasing. Se ha encontrado que se puede cancelar dicho aliasing con moderado esfuerzo y buenos resultados mediante la inclusión de un aliasing artificial en la versión modificada de la primera información de audio decodificada.

40 En una forma de realización preferida, el decodificador en el dominio de la frecuencia está configurado para ejecutar una transformada solapada inversa, de tal manera que la segunda información de audio decodificada comprenda un aliasing en una porción de tiempo que está temporalmente solapada con una porción de tiempo respecto de la cual el decodificador en el dominio de la predicción lineal emite la primera información de audio decodificada, y de tal manera que la segunda información de audio decodificada esté libre de aliasing en una porción de tiempo posterior a la porción de tiempo respecto de la cual el decodificador en el dominio de la predicción lineal emite la primera información de audio decodificada. Esta forma de realización de acuerdo con la invención se basa en la idea de que es ventajoso emplear una transformada solapada (o una transformada solapada inversa) y una generación de ventanas que mantenga sin aliasing la porción de tiempo, respecto de la cual no se emite una primera información de audio decodificada. Se ha encontrado que se puede producir la primera respuesta de entrada cero y la segunda respuesta de entrada cero, o la respuesta de entrada cero combinada, con poco esfuerzo informático si no es necesario emitir una información de cancelación de aliasing con respecto a un período de tiempo respecto del cual no se emite una primera información de audio decodificada. Dicho de otro modo, es preferible producir la primera respuesta de entrada cero y la segunda respuesta de entrada cero, o la respuesta de entrada cero combinada, sobre la base de un estado inicial, estado inicial en el cual se cancela sustancialmente el aliasing (por ejemplo, utilizando el aliasing artificial). En consecuencia, la primera respuesta de entrada cero y la segunda respuesta de entrada cero, o la respuesta de entrada cero combinada, están sustancialmente exentas de aliasing, por lo que es ventajoso no tener aliasing dentro de la segunda información de audio decodificada correspondiente al período de tiempo posterior al período de tiempo respecto del cual el decodificador en el dominio de la predicción lineal emite la primera información de audio decodificada. Con respecto a este punto, se debe tener presente que por lo general se emite la primera respuesta de entrada cero y la segunda respuesta de entrada cero, o la respuesta de entrada cero combinada, con respecto a dicho período de tiempo posterior al período de tiempo respecto del cual el decodificador en el dominio de la predicción lineal emite la primera información de audio decodificada (dado que la primera respuesta de entrada cero y la segunda respuesta de entrada cero, o la respuesta de entrada cero combinada,

constituyen sustancialmente una continuación en decrecimiento gradual de la primera información de audio decodificada, tomando en cuenta la segunda información de audio decodificada y, por lo general, el aliasing artificial que compensa el aliasing incluido en la segunda información de audio decodificada correspondiente al período de tiempo "solapado").

5 En una forma de realización preferida, la porción de la segunda información de audio decodificada, que se utiliza para obtener la versión modificada de la primera información de audio decodificada, comprende un aliasing. Al dejar que exista cierto grado de aliasing dentro de la segunda información de audio decodificada, se puede mantener sencillo el enventanado y evitar un excesivo aumento de la información necesaria para codificar la trama de audio
10 codificada en el dominio de la frecuencia. El aliasing que se incluye en la porción de la segunda información de audio decodificada que se utiliza para obtener la versión modificada de la primera información de audio decodificada se puede compensar mediante el aliasing artificial antes citado, de manera que no se incurre en una degradación severa de la calidad del audio.

15 En una forma de realización preferida, el aliasing artificial que se utiliza para obtener la versión modificada de la primera información de audio decodificada compensa, por lo menos parcialmente, el aliasing que se incluye en la porción de la segunda información de audio decodificada, que se utiliza para obtener la versión modificada de la primera información de audio decodificada. En consecuencia, se puede obtener una buena calidad de audio.

20 En una forma de realización preferida, el procesador de transición está configurado para aplicar un primer enventanado a la primera información de audio decodificada, para obtener una versión incluida en ventana de la primera información de audio decodificada, y para aplicar un segundo enventanado a una versión espejada en tiempo de la primera información de audio decodificada, para obtener una versión incluida en ventana de la versión espejada en tiempo de la primera información de audio decodificada. En este caso, el procesador de transición
25 puede estar configurado para combinar la versión incluida en ventana de la primera información de audio decodificada y la versión incluida en ventana de la versión espejada en tiempo de la primera información de audio decodificada, a fin de obtener la versión modificada de la primera información de audio decodificada. Esta forma de realización de acuerdo con la invención se basa en la idea de que se debe aplicar cierto grado de generación de ventanas para obtener una correcta cancelación del aliasing en la versión modificada de la primera información de audio decodificada, que se utiliza como entrada para la emisión de la respuesta de entrada cero. En consecuencia,
30 se puede lograr que la respuesta de entrada cero (por ejemplo, la segunda respuesta de entrada cero o la respuesta de entrada cero combinada) sea muy adecuada para alisar la transición entre la información de audio codificada en el dominio de codificación de la predicción lineal y la trama de audio subsiguiente codificada en el dominio de la frecuencia.

35 En una forma de realización preferida, el procesador de transición está configurado para combinar, en forma lineal, la segunda información de audio decodificada con la primera respuesta de entrada cero y la segunda respuesta de entrada cero, o con la respuesta combinada de entrada cero, correspondiente a una porción de tiempo respecto de la cual no se emite una primera información de audio decodificada por el decodificador en el dominio de la predicción lineal, a fin de obtener la segunda información de audio decodificada modificada. Se ha encontrado que una sencilla combinación lineal (por ejemplo, una sencilla suma y/o resta, o una combinación lineal ponderada, o una combinación lineal con atenuación cruzada) es muy adecuada para la producción de una transición sin altibajos.

45 En una forma de realización preferida, el procesador de transición está configurado para dejar la primera información de audio decodificada sin cambios mediante la segunda información de audio decodificada cuando se emite una información de audio decodificada correspondiente a una trama de audio codificada en un dominio de la predicción lineal, de tal manera que la información de audio decodificada provista respecto de una trama de audio codificada en el dominio de la predicción lineal se emita independientemente de la información de audio decodificada correspondiente a una trama de audio subsiguiente codificada en el dominio de la frecuencia. Se ha encontrado que el concepto de acuerdo con la presente invención ni exige cambiar la primera información de audio decodificada sobre la base de la segunda información de audio decodificada para obtener una transición suficientemente suave. Por consiguiente, dejando la primera información de audio decodificada sin cambios por la segunda información de audio decodificada, se puede evitar el retardo, puesto que como consecuencia de eso se puede emitir la primera información de audio decodificada para la renderización (por ejemplo, para un oyente) incluso antes de completarse
50 la decodificación de la segunda información de audio decodificada (asociada a la trama de audio subsiguiente codificada en el dominio de la frecuencia). Por el contrario, la respuesta de entrada cero (primera y segunda respuestas de entrada cero o respuesta de entrada cero combinada) se puede computar en cuanto esté disponible la segunda información de audio decodificada. De esa manera se puede evitar el retardo.

60 En una forma de realización preferida, el decodificador de audio está configurado para emitir una información de audio completamente decodificada correspondiente a una trama de audio codificada en el dominio de la predicción lineal, a la que sigue una trama de audio codificada en el dominio de la frecuencia, antes de decodificar (o antes de completar la decodificación) de la trama de audio codificada en el dominio de la frecuencia. Este concepto es posible debido a que la primera información de audio decodificada no se modifica sobre la base de la segunda información

de audio decodificada y esto ayuda a evitar toda demora.

En una forma de realización preferida, el procesador de transición está configurado para enventanar la primera respuesta de entrada cero y la segunda respuesta de entrada cero, o la respuesta combinada de entrada cero, antes de modificar la segunda información de audio decodificada dependiendo de la primera respuesta de entrada cero en ventana y la segunda respuesta de entrada cero en ventana, o dependiendo de la respuesta de entrada cero combinada en ventana. En consecuencia, la transición se puede llevar a cabo de manera especialmente suave. Además, se puede evitar cualquier problema que pudiera surgir como resultado de una respuesta de entrada cero muy larga.

En una forma de realización preferida, el procesador de transición está configurado para enventanar la primera respuesta de entrada cero y la segunda respuesta de entrada cero, o la respuesta de entrada cero combinada, empleando una ventana lineal. Se ha encontrado que el uso de una ventana lineal es un concepto sencillo que, de todos modos, trae aparejada una impresión auditiva favorable.

Una forma de realización de acuerdo con la invención crea un método para emitir una información de audio decodificada sobre la base de una información de audio codificada. El método comprende la ejecución de una decodificación en el dominio de la predicción lineal para emitir una primera información de audio decodificada sobre la base de una trama de audio codificada en un dominio de la predicción lineal. El método comprende asimismo ejecutar una decodificación en el dominio de la frecuencia para emitir una segunda información de audio decodificada sobre la base de una trama de audio codificada en un dominio de la frecuencia. El método comprende asimismo obtener una primera respuesta de entrada cero de un filtrado de predicción lineal en respuesta a un primer estado inicial del filtrado de predicción lineal definido por la primera información de audio decodificada y obtener una segunda respuesta de entrada cero del filtrado de predicción lineal en respuesta a un segundo estado inicial del filtrado de predicción lineal definido por una versión modificada de la primera información de audio decodificada, que se emite con un aliasing artificial, y que comprende la contribución de una porción de la segunda información de audio decodificada. Por otro lado, el método comprende obtener una respuesta de entrada cero combinada del filtrado de predicción lineal en respuesta a un estado inicial del filtrado de predicción lineal definido por una combinación de la primera información de audio decodificada y de una versión modificada de la primera información de audio decodificada, que se emite con un aliasing artificial, y que comprende la contribución de una porción de la segunda información de audio decodificada. El método comprende además modificar la segunda información de audio decodificada, que se emite sobre la base de una trama de audio codificada en el dominio de la frecuencia posterior a una trama de audio codificada en el dominio de la predicción lineal, dependiendo de la primera respuesta de entrada cero y la segunda respuesta de entrada cero, o dependiendo de la respuesta de entrada cero combinada, para obtener una transición suave entre la primera información de audio decodificada y la segunda información de audio decodificada modificada. Este método se basa en consideraciones similares a las referentes al decodificador de audio antes descrito y trae aparejadas las mismas ventajas.

Otra forma de realización de acuerdo con la invención crea un programa de computación para ejecutar dicho método al correr el programa de computación en una computadora.

Otra forma de realización de acuerdo con la invención crea un método para emitir una información de audio decodificada sobre la base de una información de audio codificada. El método comprende emitir una primera información de audio decodificada sobre la base de una trama de audio codificada en un dominio de la predicción lineal. El método comprende asimismo emitir una segunda información de audio decodificada sobre la base de una trama de audio codificada en un dominio de la frecuencia. El método comprende asimismo obtener una respuesta de entrada cero de un filtrado de predicción lineal, donde el estado inicial del filtrado de predicción lineal se define de conformidad con la primera información de audio decodificada y la segunda información de audio decodificada. El método comprende asimismo modificar la segunda información de audio decodificada, que se emite sobre la base de una trama de audio codificada en el dominio de la frecuencia posterior a una trama de audio codificada en el dominio de la predicción lineal, dependiendo de la respuesta de entrada cero, para obtener una transición suave entre la primera información de audio decodificada y la segunda información de audio decodificada modificada.

Este método se basa en las mismas consideraciones que el decodificador de audio antes descrito.

Otra forma de realización de acuerdo con la invención comprende un programa de computación para ejecutar dicho método.

4. Breve Descripción de las Figuras

A continuación se describen las formas de realización de acuerdo con la presente invención tomando como referencia las figuras adjuntas, en las cuales:

Fig. 1 ilustra un diagrama esquemático de bloques de un decodificador de audio de acuerdo con

- una forma de realización de la presente invención;
- Fig. 2 ilustra un diagrama esquemático de bloques de un decodificador de audio de acuerdo con otra forma de realización de la presente invención;
- 5 Fig. 3 ilustra un diagrama esquemático de bloques de un codificador de audio de acuerdo con otra forma de realización de la presente invención;
- Fig. 4a ilustra una representación esquemática de las ventanas en una transición de una trama de audio codificada por MDCT a otra trama de audio codificada por MDCT;
- 10 Fig. 4b ilustra una representación esquemática de una ventana empleada para una transición de una trama de audio codificada por CELP a una trama de audio codificada por MDCT;
- 15 Figs. 5a, 5b y 5c ilustran una representación gráfica de señales de audio en un decodificador de audio convencional;
- Figs. 6a, 6b, 6c y 6d ilustran una representación gráfica de señales de audio en un decodificador de audio convencional;
- 20 Fig. 7a ilustra una representación gráfica de una señal de audio obtenida sobre la base de una trama de CELP anterior y de una primera respuesta de entrada cero;
- Fig. 7b ilustra una representación gráfica de una señal de audio, que es una segunda versión de la trama de CELP anterior, y de una segunda respuesta de entrada cero;
- 25 Fig. 7c ilustra una representación gráfica de una señal de audio que se obtiene si se subtrae la segunda respuesta de entrada cero de la señal de audio de la trama de MDCT actual;
- 30 Fig. 8a ilustra una representación gráfica de una señal de audio obtenida sobre la base de una trama de CELP anterior;
- Fig. 8b ilustra una representación gráfica de una señal de audio, que se obtiene como segunda versión de la trama de MDCT actual; y
- 35 Fig. 8c ilustra una representación gráfica de una señal de audio, que es una combinación de la señal de audio obtenida sobre la base de la trama de CELP anterior y de la señal de audio que es una segunda versión de la trama de MDCT;
- 40 Fig. 9 ilustra un gráfico de flujo de un método para emitir una información de audio decodificada, de acuerdo con una forma de realización de la presente invención; y
- Fig. 10 ilustra un gráfico de flujo de un método para emitir una información de audio decodificada, de acuerdo con otra forma de realización de la presente invención.
- 45

5. Descripción detallada de las formas de realización

5.1. Decodificador de Audio de acuerdo con la Fig. 1

50 La Fig. 1 ilustra un diagrama esquemático de bloques de un decodificador de audio 100, de acuerdo con una forma de realización de la presente invención. El codificador de audio 100 está configurado para recibir una información de audio codificada 110, que puede comprender, por ejemplo, una primera trama codificada en un dominio de la predicción lineal y una segunda trama subsiguiente codificada en el dominio de la frecuencia. El decodificador de audio 100 también está configurado para emitir una información de audio decodificada 112 sobre la base de la información de audio codificada 110.

55

El decodificador de audio 100 comprende un decodificador en el dominio de la predicción lineal 120, que está configurado para emitir una primera información de audio decodificada 122 sobre la base de una trama de audio codificada en el dominio de la predicción lineal. El decodificador de audio 100 comprende asimismo un decodificador en el dominio de la frecuencia (o un decodificador en el dominio de la transformada 130), que está configurado para emitir una segunda información de audio decodificada 132 sobre la base de una trama de audio codificada en el dominio de la frecuencia (o en el dominio de la transformada). Por ejemplo, el decodificador en el dominio de la predicción lineal 120 puede ser un decodificador CELP, un decodificador ACELP, o un decodificador similar que efectúa un filtrado de predicción lineal sobre la base de una señal de excitación y sobre la base de la representación

60

codificada de las características del filtro de predicción lineal (o los coeficientes de filtro).

5 El decodificador en el dominio de la frecuencia 130 puede ser, por ejemplo, un decodificador del tipo AAC (por sus siglas en inglés, Advanced Audio Coding, Codificación Avanzada de audio) o cualquier decodificador que se base en la decodificación tipo AAC. Por ejemplo, el decodificador en el dominio de la frecuencia (o el decodificador en el dominio de la transformada) puede recibir una representación codificada de parámetros en el dominio de la frecuencia (o parámetros en el dominio de la transformada) y emitir, sobre la base de aquélla, la segunda información de audio decodificada. Por ejemplo, el decodificador en el dominio de la frecuencia 130 puede decodificar los coeficientes en el dominio de la frecuencia (o los coeficientes en el dominio de la transformada),
10 escalar los coeficientes en el dominio de la frecuencia (o los coeficientes en el dominio de la transformada) dependiendo de factores de escala (donde los factores de escala pueden ser provistos para diferentes bandas de frecuencia y pueden ser representados en diferentes formas) y ejecutar una conversión del dominio de la frecuencia al dominio del tiempo (o una conversión del dominio de la transformada al dominio del tiempo) como, por ejemplo, una Transformada Rápida de Fourier inversa o una transformada de coseno discreta modificada (MDCT inversa).

15 El decodificador de audio 100 comprende asimismo un procesador de transición 140. El procesador de transición 140 está configurado para obtener una respuesta de entrada cero de un filtrado de predicción lineal, donde el estado inicial del filtrado de predicción lineal se define de conformidad con la primera información de audio decodificada y la segunda información de audio decodificada. Más aún, el procesador de transición 140 está configurado para
20 modificar la segunda información de audio decodificada 132, que se emite sobre la base de una trama de audio codificada en el dominio de la frecuencia que sigue a una trama de audio codificada en el dominio de la predicción lineal, dependiendo de la respuesta de entrada cero, para obtener una transición suave entre la primera información de audio decodificada y la segunda información de audio decodificada modificada.

25 Por ejemplo, el procesador de transición 140 puede comprender una determinación del estado inicial 144, que recibe la primera información de audio decodificada 122 y la segunda información de audio decodificada 132 y que emite, basándose en la misma, una información del estado inicial 146. El procesador de transición 140 comprende asimismo un filtrado de predicción lineal 148, que recibe la información del estado inicial 146 y que emite, basándose
30 en la misma, una respuesta de entrada cero 150. Por ejemplo, el filtrado de predicción lineal puede ser ejecutado por un filtro de predicción lineal, que se inicializa sobre la base de la información del estado inicial 146 y está provisto de una entrada cero. En consecuencia, el filtrado de predicción lineal produce la respuesta de entrada cero 150. El procesador de transición 140 comprende asimismo una modificación 152, que modifica la segunda información de audio decodificada 132 dependiendo de la respuesta de entrada cero 150, para obtener de esa manera una segunda información de audio decodificada modificada 142, que constituye una información de salida del procesador de
35 transición 140. La segunda información de audio decodificada modificada 142 está típicamente concatenada con la primera información de audio decodificada 122, para obtener la información de audio decodificada 112.

40 En cuanto a la funcionalidad del decodificador de audio 100, se debe considerar un caso en que a una trama de audio codificada en el dominio de la predicción lineal (primera trama de audio) le sigue una trama de audio codificada en el dominio de la frecuencia (segunda trama de audio). La primera trama de audio, codificada en el dominio de la predicción lineal, ha de ser decodificada por el decodificador en el dominio de la predicción lineal 120. En consecuencia, se obtiene la primera información de audio decodificada 122, que está asociada a la primera trama de audio. Sin embargo, la información de audio decodificada 122 asociada a la primera trama de audio queda por lo general sin efecto por ninguna información de audio decodificada sobre la base de la segunda trama de audio,
45 que es codificada en el dominio de la frecuencia. Sin embargo, la segunda información de audio decodificada 132 es emitida por el decodificador en el dominio de la frecuencia 130 sobre la base de la segunda trama de audio que es codificada en el dominio de la frecuencia.

50 Lamentablemente, la segunda información de audio decodificada 132, que está asociada a la segunda trama de audio, por lo general no comprende una transición suave con la primera información de audio decodificada 122 que está asociada a la primera información de audio decodificada.

Sin embargo, se debe tener en cuenta que la segunda información de audio decodificada es emitida con respecto a un período de tiempo que también se superpone con el período de tiempo asociado a la primera trama de audio. La porción de la segunda información de audio decodificada, que se emite durante el período de tiempo de la primera trama de audio (es decir una porción inicial de la segunda información de audio decodificada 132) es evaluada por la determinación del estado inicial 144. Más aún, la determinación del estado inicial 144 también evalúa por lo menos una porción de la primera información de audio decodificada. En consecuencia, la determinación del estado inicial 144 obtiene la información del estado inicial 146 sobre la base de una porción de la primera información de audio decodificada (porción que está asociada al tiempo de la primera trama de audio) y sobre la base de una porción de la segunda información de audio decodificada (porción de la segunda información de audio decodificada 130 que también está asociada al tiempo de la primera trama de audio). En consecuencia, la información del estado inicial 146 se emite dependiendo de la primera información decodificada 132 y también dependiendo de la segunda información de audio decodificada.

Se debe tener en cuenta que se puede emitir la información del estado inicial 146 en cuanto esté disponible la segunda información de audio decodificada 132 (o por lo menos una porción inicial de la misma necesaria para la determinación del estado inicial 144). El filtrado de predicción lineal 148 también se puede ejecutar en cuanto esté disponible la información del estado inicial 146, ya que el filtrado de predicción lineal utiliza coeficientes de filtro que ya son conocidos en virtud de la decodificación de la primera trama de audio. En consecuencia, la respuesta de entrada cero 150 puede ser emitida en cuanto esté disponible la segunda información de audio decodificada 132 (o por lo menos la porción inicial de la misma, necesaria para la determinación del estado inicial 144). Más aún, se puede emplear la respuesta de entrada cero 150 para modificar la parte de la segunda información de audio decodificada 132 que está asociada al tiempo de la segunda trama de audio (en lugar de estarlo al tiempo de la primera trama de audio). En consecuencia, una porción de la segunda información de audio decodificada, que por lo general se encuentra en el comienzo del tiempo asociado a la segunda trama de audio, se modifica. Por lo tanto, se obtiene una transición suave entre la primera información de audio decodificada 122 (que por lo general termina al final del tiempo asociado a la primera trama de audio) y la segunda información de audio decodificada modificada 142 (donde preferentemente se descarta la porción de tiempo de la segunda información de audio decodificada 132 que incluye tiempos asociados a la primera trama de audio, y por consiguiente es preferible utilizarla solo para la emisión de la información del estado inicial para el filtrado de predicción lineal). En consecuencia, se puede emitir la información de audio decodificada total 112 sin retardo, puesto que no se retarda la emisión de la primera información de audio decodificada 122 (porque la primera información de audio decodificada 122 es independiente de la segunda información de audio decodificada 132), y porque la segunda información de audio decodificada modificada 142 se puede emitir en cuanto se dispone de la segunda información de audio decodificada 132. En consecuencia, se pueden lograr transiciones suaves entre las diferentes tramas de audio dentro de la información de audio decodificada 112, aunque exista una conmutación de una trama de audio codificada en el dominio de la predicción lineal (primera trama de audio) hacia una trama de audio codificada en el dominio de la frecuencia (segunda trama de audio).

Sin embargo, se debe tener en cuenta que el decodificador de audio 100 puede ser complementado por cualquiera de las características y funcionalidades descritas en la presente.

30 5.2. Decodificador de Audio de acuerdo con la Fig. 2

La Fig. 2 ilustra un diagrama esquemático de bloques de un decodificador de audio, de acuerdo con otra forma de realización de la presente invención. El decodificador de audio 200 está configurado para recibir una información de audio codificada 210, que puede comprender, por ejemplo, una o más tramas codificadas en el dominio de la predicción lineal (o de manera equivalente, en una representación en el dominio de la predicción lineal), y una o más tramas de audio codificadas en el dominio de la frecuencia (o, de modo equivalente, en el dominio de la transformada, o de modo equivalente en una representación en el dominio de la transformada). El decodificador de audio 200 está configurado para emitir una información de audio decodificada 212 sobre la base de la información de audio codificada 210, donde la información de audio decodificada 212 puede estar, por ejemplo, en una representación en el dominio del tiempo.

El decodificador de audio 200 comprende un decodificador en el dominio de la predicción lineal 220, que es sustancialmente idéntico al decodificador en el dominio de la predicción lineal 120, por lo cual se aplican las explicaciones antes presentadas. Por consiguiente, el decodificador en el dominio de la predicción lineal 210 recibe tramas de audio codificadas en una representación en el dominio de la predicción lineal, que están incluidas en la información de audio codificada 210, y emite, sobre la base de una trama de audio codificada en una representación en el dominio de la predicción lineal, una primera información de audio decodificada 222, que por lo general tiene la forma de una representación de audio en el dominio del tiempo (y que típicamente corresponde a la primera información de audio decodificada 122). El decodificador de audio 200 comprende asimismo un decodificador en el dominio de la frecuencia 230, que es sustancialmente idéntico al decodificador en la frecuencia 130, por lo cual se aplican las explicaciones antes expuestas. En consecuencia, el decodificador en el dominio de la frecuencia 230 recibe una trama de audio codificada en una representación en el dominio de la frecuencia (o en una representación en el dominio de la transformada) y emite, basándose en la misma, una segunda información de audio decodificada 232, que por lo general se presenta en forma de una representación en el dominio del tiempo.

El decodificador de audio 200 comprende asimismo un procesador de transición 240, que está configurado para modificar la segunda información de audio decodificada 232, para derivar de esa manera una segunda información de audio decodificada modificada 242.

El procesador de transición 240 está configurado para obtener una primera respuesta de entrada cero de un filtro de predicción lineal en respuesta a un estado inicial del filtro de predicción lineal definido por la primera información de audio decodificada 222. El procesador de transición también está configurado para obtener una segunda respuesta de entrada cero del filtro de predicción lineal en respuesta a un segundo estado inicial del filtro de predicción lineal definido por una versión modificada de la primera información de audio decodificada, que se emite con un aliasing

artificial y que comprende la contribución de una porción de la segunda información de audio decodificada 232. Por ejemplo, el procesador de transición 240 comprende una determinación del estado inicial 242, que recibe la primera información de audio decodificada 222 y que emite una primera información de estado inicial 244 basándose en la misma. Por ejemplo, la primera información de estado inicial 244 puede reflejar simplemente una porción de la primera información de audio decodificada 222, por ejemplo una porción adyacente al final de la porción de tiempo asociada a la primera trama de audio. El procesador de transición 240 también puede comprender un (primer) filtrado de predicción lineal 246, que está configurado para recibir la primera información de estado inicial 244 como estado inicial del filtro de predicción lineal y para emitir, sobre la base de la primera información de estado inicial 244, una primera respuesta de entrada cero 248. El procesador de transición 240 comprende asimismo una combinación de modificación/adición de aliasing 250, que está configurado para recibir la primera información de audio decodificada 222, o por lo menos una porción de la misma (por ejemplo, una porción que es adyacente a un extremo de una porción de tiempo asociada a la primera trama de audio), como así también la segunda información decodificada 232, o por lo menos una porción de la misma (por ejemplo, una porción de tiempo de la segunda información de audio decodificada 232 que está temporalmente dispuesta en un extremo de una porción de tiempo asociada a la primera trama de audio, donde la segunda información de audio decodificada se emite, por ejemplo, principalmente en relación con una porción de tiempo asociada a la segunda trama de audio, aunque también en cierto grado, respecto de un extremo de la porción de tiempo asociada a la primera trama de audio que es codificada en la representación en el dominio de la predicción lineal). La combinación de modificación/adición de aliasing puede modificar, por ejemplo, la porción de tiempo de la primera información de audio decodificada, agregar un aliasing artificial sobre la base de la porción de tiempo de la primera información de audio decodificada, y también sumar la porción de tiempo de la segunda información de audio decodificada, para obtener de esa manera una segunda información de estado inicial 252. Dicho de otro modo, la combinación de modificación/adición de aliasing puede ser parte de una segunda determinación del estado inicial. La segunda información de estado inicial determina un estado inicial de un segundo filtrado de predicción lineal 254, que está configurado para emitir una segunda respuesta de entrada cero 256 sobre la base de la segunda información de estado inicial.

Por ejemplo, el primer filtrado de predicción lineal y el segundo filtrado de predicción lineal pueden utilizar una configuración de filtros (por ejemplo, coeficientes de filtro), que son producidos por el decodificador en el dominio de la predicción lineal 220 con respecto a la primera trama de audio (que es codificada en la representación en el dominio de la predicción lineal). Dicho de otro modo, el primer y segundo filtrados de predicción lineal 246, 254 pueden ejecutar el mismo filtrado de predicción lineal que también es realizado por el decodificador en el dominio de la predicción lineal 220 para obtener la primera información de audio decodificada 222 asociada a la primera trama de audio. Sin embargo, los estados iniciales del primer y segundo filtrados de predicción lineal 246, 254 pueden ser ajustados a los valores determinados por la primera determinación del estado inicial 244 y por la segunda determinación del estado inicial 250 (que comprende la combinación de modificación/adición de aliasing). Sin embargo, se puede ajustar a cero una señal de entrada de los filtros de predicción lineal 246, 254. En consecuencia, la primera respuesta de entrada cero 248 y la segunda respuesta de entrada cero 256 se obtienen de tal manera que la primera respuesta de entrada cero y la segunda respuesta de entrada cero se basen en la primera información de audio decodificada y la segunda información de audio decodificada, y se modelen usando el mismo filtro de predicción lineal utilizado por el decodificador en el dominio de la predicción lineal 220.

El procesador de transición 240 comprende asimismo una modificación 258, que recibe la segunda información de audio codificada 232 y modifica la segunda información de audio decodificada 232 dependiendo de la primera respuesta de entrada cero 248 y dependiendo de la segunda respuesta de entrada cero 256, para obtener de esa manera la segunda información de audio decodificada modificada 242. Por ejemplo, la modificación 258 puede sumar y/o restar la primera respuesta de entrada cero 248 a o de la segunda información de audio decodificada 232, y puede sumar o restar la segunda respuesta de entrada cero 256 a o de la segunda información de audio decodificada, para obtener la segunda información de audio decodificada modificada 242.

Por ejemplo, se puede emitir la primera respuesta de entrada cero y la segunda respuesta de entrada cero respecto de un período de tiempo que está asociado a la segunda trama de audio, de tal manera que solo se modifique la porción de la segunda información de audio decodificada que está asociada al período de tiempo de la segunda trama de audio. Más aún, se pueden descartar los valores de la segunda información de audio decodificada 232 asociada a una porción de tiempo que está asociada a una primera trama de audio en la producción final de la segunda información de audio decodificada modificada (sobre la base de las respuestas de entrada cero).

Más aún, el decodificador de audio 200 está configurado preferentemente para concatenar la primera información de audio decodificada 222 y la segunda información de audio decodificada modificada 242, para obtener de esa manera la información de audio decodificada total 212.

En lo que respecta a la funcionalidad del decodificador de audio 200, se hace referencia a las explicaciones anteriores del decodificador de audio 100. Más aún, a continuación se describen detalles adicionales, tomando como referencia las otras figuras.

5.3. Decodificador de Audio de acuerdo con la Fig. 3

La Fig. 3 ilustra un diagrama esquemático de bloques de un decodificador de audio 300, de acuerdo con una forma de realización de la presente invención. El decodificador de audio 300 es similar al decodificador de audio 200, por lo cual solo se describen en detalle las diferencias. Por lo demás, se hace referencia a las explicaciones antes expuestas con respecto al decodificador de audio 200.

El decodificador de audio 300 está configurado para recibir una información de audio codificada 310, que puede corresponder a la información de audio codificada 210. Más aún, el decodificador de audio 300 está configurado para emitir una información de audio decodificada 312, que puede corresponder a la información de audio decodificada 212.

El decodificador de audio 300 comprende un decodificador en el dominio de la predicción lineal 320, que puede corresponder al decodificador en el dominio de la predicción lineal 220, y un decodificador en el dominio de la frecuencia 330, que corresponde al decodificador en el dominio de la frecuencia 230. El decodificador en el dominio de la predicción lineal 320 emite la primera información de audio decodificada 322, por ejemplo sobre la base de una primera trama de audio que es codificada en el dominio de la predicción lineal. Más aún, el decodificador de audio en el dominio de la frecuencia 330 emite una segunda información de audio decodificada 332, por ejemplo sobre la base de una segunda trama de audio (que sigue a la primera trama de audio) codificada en el dominio de la frecuencia (o en el dominio de la transformada). La primera información de audio decodificada 322 puede corresponder a la primera información de audio decodificada 222, y la segunda información de audio decodificada 332 puede corresponder a la segunda información de audio decodificada 232.

El decodificador de audio 300 comprende asimismo un procesador de transición 340, que puede corresponder, en términos de su funcionalidad general, al procesador de transición 340, y que podría emitir una segunda información de audio decodificada modificada 342 sobre la base de la segunda información de audio decodificada 332.

El procesador de transición 340 está configurado para obtener una respuesta de entrada cero combinada del filtro de predicción lineal en respuesta a un estado inicial (combinado) del filtro de predicción lineal definido por una combinación de la primera información de audio decodificada y de una versión modificada de la primera información de audio decodificada, que se emite con un aliasing artificial, y que comprende la contribución de una porción de la segunda información de audio decodificada. Más aún, el procesador de transición está configurado para modificar la segunda información de audio decodificada, que se emite sobre la base de una trama de audio codificada en el dominio de la frecuencia posterior a una trama de audio codificada en el dominio de la predicción lineal, dependiendo de la respuesta de entrada cero combinada, para obtener una transición suave entre la primera información de audio decodificada y la segunda información de audio decodificada modificada.

Por ejemplo, el procesador de transición 340 comprende una combinación de modificación/adición de aliasing 342 que recibe la primera información de audio decodificada 322 y la segunda información de audio decodificada 332 y emite, basándose en la misma, una información de estado inicial combinada 344. Por ejemplo, se puede considerar que la combinación de modificación/adición de aliasing es una determinación del estado inicial. También se debe tener en cuenta que la combinación de modificación/adición de aliasing 342 puede ejecutar la funcionalidad de la determinación del estado inicial 242 y de la determinación del estado inicial 250. La información de estado inicial combinada 344 puede ser igual (o por lo menos corresponder), por ejemplo, a la suma de la primera información de estado inicial 244 y la segunda información de estado inicial 252. En consecuencia, la combinación de modificación/adición de aliasing 342 puede combinar, por ejemplo, una porción de la primera información de audio decodificada 322 con un aliasing artificial y también con una porción de la segunda información de audio decodificada 332. Más aún, la combinación de modificación/adición de aliasing 342 puede modificar asimismo la porción de la primera información de audio decodificada y/o sumar una copia enventanada de la primera información de audio decodificada 322, como se describe más adelante en forma más detallada. De esa manera se obtiene la información de estado inicial combinada 344.

El procesador de transición 340 comprende asimismo un filtrado de predicción lineal 346, que recibe la información de estado inicial combinada 344 y emite, basándose en la misma, una respuesta de entrada cero combinada 348 a una modificación 350. El filtrado de predicción lineal 346 puede ejecutar, por ejemplo, un filtrado de predicción lineal sustancialmente idéntico a un filtrado de predicción lineal que es ejecutado por el decodificador de predicción lineal 320 para obtener la primera información de audio decodificada 322. Sin embargo, un estado inicial del filtrado de predicción lineal 346 puede ser determinado por la información de estado inicial combinada 344. Además, se puede ajustar a cero una señal de entrada para emitir la respuesta de entrada cero combinada 348, de tal manera que el filtrado de predicción lineal 344 emita una respuesta de entrada cero sobre la base de la información de estado inicial combinada 344 (donde los parámetros de filtrado o coeficientes de filtro son idénticos, por ejemplo, a los parámetros de filtrado o coeficientes de filtro usados por el decodificador en el dominio de la predicción lineal 320 para emitir la primera información de audio decodificada 322 asociada a la primera trama de audio. Más aún, se utiliza la respuesta de entrada cero combinada 348 para modificar la segunda información de audio decodificada

332, para derivar así la segunda información de audio decodificada modificada 342. Por ejemplo, la modificación 350 puede sumar la respuesta de entrada cero combinada 348 a la segunda información de audio decodificada 332, o puede restar la respuesta de entrada cero combinada de la segunda información de audio decodificada.

- 5 Sin embargo, por más detalles, se hace referencia a las explicaciones de los decodificadores de audio 100, 200, y también a las siguientes explicaciones detalladas.

5.4. Descripción del Concepto de Transición

- 10 En lo sucesivo se describen algunos detalles con respecto a la transición de una trama de CELP a una trama de MDCT, que se puede aplicar a los decodificadores de audio 100, 200, 300.

Además se describen las diferencias con respecto a los conceptos convencionales.

15 MDCT y enventanado – Visión de conjunto

- En las formas de realización de acuerdo con la invención, se resuelve el problema de aliasing incrementando la longitud de la MDCT (por ejemplo, con respecto a una trama de audio codificada en el dominio de la MDCT posterior a una trama de audio codificada en el dominio de la predicción lineal) de tal manera que el punto de pliegue izquierdo (por ejemplo, de una señal de audio en el dominio del tiempo reconstruida sobre la base de una serie de coeficientes de MDCT utilizando una transformada de MDCT inversa) se corra a la izquierda del límite entre las tramas de CELP y MDCT. También se cambia la parte izquierda de la ventana de MDCT (por ejemplo, de una ventana que se aplica a una señal de audio en el dominio del tiempo reconstruida sobre la base de una serie de coeficientes de MDCT usando una transformada de MDCT inversa) (por ejemplo, en comparación con una ventana de MDCT “normal”), de tal manera que se reduzca el solapamiento.

- A título de ejemplo, las Figs. 4a y 4b ilustran una representación gráfica de diferentes ventanas, donde la Fig. 4a ilustra ventanas correspondientes a una transición de una primera trama de MDCT (es decir una primera trama de audio codificada en el dominio de la frecuencia) a otra trama de MDCT (es decir una segunda trama de audio codificada en el dominio de la frecuencia). Por el contrario, la Fig. 4b ilustra una ventana que se utiliza para una transición de una trama de CELP (es decir una primera trama de audio codificada en el dominio de la predicción lineal) a una trama de MDCT (es decir una segunda trama de audio siguiente codificada en el dominio de la frecuencia).

- 35 Dicho de otro modo, la Fig. 4a ilustra una secuencia de tramas de audio que se pueden considerar como ejemplo comparativo. Por el contrario, la Fig. 4b ilustra una secuencia en que una primera trama de audio es codificada en el dominio de la predicción lineal y seguida por una segunda trama de audio codificada en el dominio de la frecuencia, donde se trata el caso de acuerdo con la Fig. 4b de manera particularmente ventajosa en las formas de realización de la presente invención.

- 40 Haciendo referencia ahora a la Fig. 4a, se debe tener en cuenta que una abscisa 410 describe el tiempo en milisegundos, y que la ordenada 412 describe la amplitud de la ventana (por ej., una amplitud normalizada de la ventana) en unidades arbitrarias. Como se puede ver, la longitud de una trama es igual a 20 ms, de tal manera que el período de tiempo asociado a la primera trama de audio se extiende entre $t = -20$ ms y $t = 0$. Un período de tiempo asociado a la segunda trama de audio se extiende desde el tiempo $t = 0$ hasta $t = 20$ ms. Sin embargo, se puede notar que una primera ventana para incluir en ventanas las muestras de audio en el dominio del tiempo provistas por una transformada de coseno discreta modificada sobre la base de los coeficientes de MDCT decodificados se extiende entre los tiempos $t = -20$ ms y $t = 8,75$ ms. Por consiguiente, la longitud de la primera ventana 420 es mayor que la longitud de la trama (20 ms). En consecuencia, aunque el tiempo entre $t = -20$ ms y $t = 0$ está asociado a la primera trama de audio, las muestras de audio en el dominio del tiempo se producen sobre la base de la decodificación de la primera trama de audio, correspondiente a los tiempos entre $t = -20$ ms y $t = 8,75$ ms. Por consiguiente, existe un solapamiento de aproximadamente 8,75 ms entre muestras de audio en el dominio del tiempo provistas sobre la base de la primera trama de audio codificada y muestras de audio en el dominio del tiempo provistas sobre la base de la segunda trama de audio decodificada. Se debe tener en cuenta que la segunda ventana lleva la designación 422 y se extiende entre el tiempo $t = 0$ y $t = 28,75$ ms.

- Más aún, se debe tener en cuenta que las señales de audio en el dominio del tiempo enventanadas provistas para la primera trama de audio y provistas para la segunda trama de audio no están exentas de aliasing. Por el contrario, la (segunda) información de audio decodificada enventanada provista respecto de la primera trama de audio comprende aliasing entre los tiempos $t = -20$ ms y $t = -11,25$ ms, y también entre los tiempos $t = 0$ y $t = 8,75$ ms. De manera similar, la información de audio decodificada enventanada provista respecto de la segunda trama de audio comprende aliasing entre los tiempos $t = 0$ y $t = 8,75$ ms, y también entre los tiempos $t = 20$ ms y $t = 28,75$ ms. Sin embargo, por ejemplo, el aliasing incluido en la información de audio decodificada provista respecto de la primera trama de audio se cancela con el aliasing incluido en la información de audio decodificada provista respecto de la

segunda trama de audio subsiguiente en la porción de tiempo entre los tiempos $t = 0$ y $t = 8,75$ ms.

Más aún, se debe tener en cuenta que en el caso de las ventanas 420 y 422, la duración temporal entre los puntos de pliegue MDCT es igual a 20 ms, que es igual a la longitud de la trama.

5 Haciendo referencia ahora a la Fig. 4b, se pasa a describir un caso diferente, es decir una ventana para una transición de una trama de CELP a una trama de MDCT que se puede utilizar en los decodificadores de audio 100, 200, 300 para emitir la segunda información de audio decodificada. En la Fig. 4b, una abscisa 430 describe el tiempo en milisegundos y una ordenada 432 describe una amplitud de la ventana en unidades arbitrarias.

10 Como se puede apreciar en la Fig. 4b, una primera trama se extiende entre el tiempo $t_1 = -20$ ms y el tiempo $t_2 = 0$ ms. Por consiguiente, la longitud de trama de la primera trama de audio, que es una trama de audio de CELP, es 20 ms. Más aún, una segunda trama de audio subsiguiente se extiende entre el tiempo t_2 y $t_3 = 20$ ms. Por consiguiente, la longitud de la segunda trama de audio, que es una trama de audio de MDCT, también es de 20 ms.

15 A continuación se describen algunos detalles con respecto a la ventana 440.

Una ventana 440 comprende una primera pendiente de ventana 442, que se extiende entre los tiempos $t_4 = -1,25$ ms y el tiempo $t_2 = 0$ ms. Una segunda pendiente de ventana 444 se extiende entre los tiempos $t_3 = 20$ ms y el tiempo $t_5 = 28,75$ ms. Se debe tener en cuenta que la transformada de coseno discreta modificada, que emite la (segunda) información de audio decodificada correspondiente a (o asociada a) la segunda trama de audio produce muestras en el dominio del tiempo correspondientes a un período de tiempo entre los tiempos t_4 y t_5 . Sin embargo, la transformada de coseno discreta modificada (o, más precisamente, la transformada de coseno discreta modificada inversa) (que se puede emplear en los decodificadores en el dominio de la frecuencia 130, 230, 330 si sigue una trama de audio codificada en el dominio de la frecuencia, por ejemplo en el dominio de MDCT a una trama de audio codificada en el dominio de la predicción lineal) produce muestras en el dominio del tiempo que comprenden un aliasing correspondiente a los tiempos entre t_4 y t_2 y a los tiempos entre el tiempo t_3 y el tiempo t_5 sobre la base de una representación en el dominio de la frecuencia de la segunda trama de audio. Por el contrario, la transformada de coseno discreta modificada inversa produce muestras sin aliasing en el dominio del tiempo correspondientes a un período de tiempo entre los tiempos t_2 y t_3 sobre la base de la representación en el dominio de la frecuencia de la segunda trama de audio. Por consiguiente, la primera pendiente de ventana 442 está asociada a las muestras de audio en el dominio del tiempo que comprenden cierta cantidad de aliasing, y la segunda pendiente de ventana 444 también está asociada a muestras de audio en el dominio del tiempo que comprenden cierta cantidad de aliasing.

35 Además, se debe tener en cuenta que el tiempo entre los puntos de pliegue de MDCT es igual a 25 ms en el caso de la segunda trama de audio, lo que implica que un número de coeficientes de MDCT codificados debe ser mayor en el caso de la situación expuesta en la Fig. 4b que en la situación expuesta en la Fig. 4a.

40 Para finalizar, los decodificadores de audio 100, 200, 300 pueden aplicar las ventanas 420, 422 (por ejemplo, para el enventanado de una salida de una transformada de coseno discreta modificada inversa en el decodificador en el dominio de la frecuencia) en caso de que tanto una primera trama de audio como una segunda trama de audio posterior a la primera trama de audio sean codificadas en el dominio de la frecuencia (por ejemplo, en el dominio de la MDCT). Por el contrario, los decodificadores de audio 100, 200, 300 pueden conmutar la operación del decodificador en el dominio de la frecuencia en caso de que una segunda trama de audio, que sigue a una primera trama de audio codificada en el dominio de la predicción lineal, sea codificada en el dominio de la frecuencia (por ejemplo, en el dominio de la MDCT). Por ejemplo, si la segunda trama de audio es codificada en el dominio de la MDCT y es posterior a una primera trama de audio anterior que es codificada en el dominio de CELP, se puede emplear una transformada de coseno discreta modificada inversa que utiliza un número incrementado de coeficientes de MDCT (lo que implica que se incluye un número incrementado de coeficientes de MDCT, en forma codificada, en la representación en el dominio de la frecuencia de una trama de audio posterior a una trama de audio anterior codificada en el dominio de la predicción lineal, en comparación con la representación en el dominio de la frecuencia de una trama de audio codificada posterior a una trama de audio anterior codificada también en el dominio de la frecuencia). Más aún, se aplica una ventana diferente, es decir la ventana 440, para enventanar la salida de la transformada de coseno discreta modificada inversa (es decir una representación de audio en el dominio del tiempo emitida por la transformada de coseno discreta modificada inversa) para obtener la segunda información de audio decodificada 132 en caso de que la segunda trama de audio (actual) codificada en el dominio de la frecuencia siga a una trama de audio codificada en el dominio de la predicción lineal (en comparación con el caso en que la segunda trama de audio (actual) siga a una trama de audio anterior también codificada en el dominio de la frecuencia).

60 Para concluir además, se puede aplicar una transformada de coseno discreta modificada inversa de longitud incrementada (en comparación con un caso normal) en el decodificador en el dominio de la frecuencia 130 en caso de que una trama de audio codificada en el dominio de la frecuencia siga a una trama de audio codificada en el dominio de la predicción lineal. Más aún, en este caso se puede usar la ventana 440 (en tanto que las ventanas 420,

422 se pueden utilizar en el caso “normal” en el que una trama de audio codificada en el dominio de la frecuencia siga a una trama de audio anterior codificada en el dominio de la frecuencia).

Con respecto al concepto de la invención, se debe tener en cuenta que la señal de CELP no se modifica para no introducir un retardo adicional, como se expone más adelante en forma más detallada. Por el contrario, las formas de realización de acuerdo con la invención generan un mecanismo para eliminar cualquier discontinuidad que se pudiera introducir en el límite entre las tramas de CELP y las tramas de MDCT. Este mecanismo alisa la discontinuidad empleando la respuesta de entrada cero del filtro de síntesis de CELP (que es utilizado, por ejemplo, por el decodificador en el dominio de la predicción lineal). A continuación se presentan los detalles.

Descripción paso por paso – Visión de conjunto

A continuación se presenta una breve descripción paso por paso. Seguidamente se presentan más detalles de esto.

Lado del Codificador

1. Cuando la trama anterior (en ocasiones también denominada “primera trama”) es de CELP (o, en general, codificada en el dominio de la predicción lineal), la trama de MDCT actual (en ocasiones también denominada “segunda trama”) (que se puede considerar un ejemplo de trama codificada en el dominio de la frecuencia o en el dominio de la transformada) es codificada con una longitud de MDCT diferente y una ventana de MDCT diferente. Por ejemplo, en este caso se puede emplear la ventana 440 (en lugar de la ventana “normal” 422).

2. La longitud de MDCT se incrementa (por ej. de 20 ms a 25 ms, véanse las Figs. 4a y 4b) de tal manera que el punto de pliegue izquierdo se corra hacia la izquierda del límite entre las tramas de CELP y de MDCT. Por ejemplo, se puede optar por una longitud de la MDCT (que puede estar definida por el número de coeficientes de MDCT) tal que la longitud de (o entre) los puntos de pliegue de MDCT sea igual a 25 ms (como se ilustra en la Fig. 4b) en comparación con la longitud “normal” entre los puntos de pliegue de MDCT de 20 ms (como se ilustra en la Fig. 4a). También se puede apreciar que el punto de pliegue “izquierdo” de la transformada de MDCT yace entre los tiempos t_4 y t_2 (en lugar de estar en un punto medio entre los tiempos $t=0$ y $t=8,75$ ms), que se pueden ver en la Fig. 4b. Sin embargo, la posición del punto de pliegue derecho de MDCT puede quedar sin modificaciones (por ejemplo, en el punto medio entre los tiempos t_3 y t_5), como se puede observar haciendo una comparación de las Figs. 4a y 4b (o, más precisamente, de las ventanas 422 y 440).

3. La parte izquierda de la ventana de MDCT se modifica de tal manera que se reduce la longitud del solapamiento (por ej. de 8,75 ms a 1,25 ms). Por ejemplo, la porción que comprende aliasing yace entre los tiempos $t_4=-1,25$ ms y $t_2=0$ (es decir antes del período de tiempo asociado a la segunda trama de audio, que se inicia en $t=0$ y finaliza en $t=20$ ms) en caso de que la trama de audio anterior sea codificada en el dominio de la predicción lineal. Por el contrario, la porción de la señal que comprende aliasing se halla entre los tiempos $t=0$ y $t=8,75$ ms en caso de que la trama de audio precedente sea codificada en el dominio de la frecuencia (por ejemplo, en el dominio de la MDCT).

Lado del Decodificador

1. Cuando la trama anterior (también denominada primera trama de audio) es de CELP (o, en términos generales, codificada en el dominio de la predicción lineal) la trama de MDCT actual (también denominada segunda trama de audio) (que es un ejemplo de trama codificada en el dominio de la frecuencia o en el dominio de la transformada) es decodificada con las mismas longitudes de MDCT y la misma ventana de MDCT utilizada del lado del codificador. Expresado de otro modo, se aplica el enventanado expuesto en la Fig. 4b a la emisión de la segunda información de audio decodificada, y también pueden ser aplicables las características antes citadas con respecto a la transformada de coseno discreta modificada inversa (que corresponden a las características de la transformada de coseno discreta modificada usada del lado del codificador).

2. Para eliminar toda discontinuidad que se pudiera producir en el límite entre las tramas de CELP y de MDCT (por ejemplo, en el límite entre la primera trama de audio y la segunda trama de audio antes citadas), se utiliza el siguiente mecanismo:

a) Se construye una primera porción de señal introduciendo artificialmente el aliasing faltante de la parte de solapamiento de la señal de MDCT (por ejemplo, de la porción de señal entre los tiempos t_4 y t_2 de la señal en el dominio del tiempo provista por la transformada de coseno discreta modificada inversa) utilizando la señal de CELP (por ejemplo, usando la primera información de audio decodificada) y una operación de solapamiento y suma. La longitud de la primera porción de señal es, por ejemplo, igual a la longitud del solapamiento (por ejemplo, 1,25 ms).

b) Se construye una segunda porción de señal restando la primera porción de señal de la señal de CELP

correspondiente (la porción situada justo antes del límite de las tramas, por ejemplo, entre la primera trama de audio y la segunda trama de audio).

- 5 c) Se genera una respuesta de entrada cero del filtro de síntesis de CELP filtrando una trama de ceros y utilizando la segunda porción de señal como estados de memoria (o como estado inicial).
- d) La respuesta de entrada cero es enventanada, por ejemplo, de tal manera que se reduzca a ceros tras un número de muestras (por ej. 64).
- 10 e) La respuesta de entrada cero en ventana se suma a la porción inicial de la señal de MDCT (por ejemplo, la porción de audio que se inicia en el momento $t_2 = 0$).

Descripción Paso por Paso – Descripción Detallada de la Funcionalidad del Decodificador

15 A continuación se describe la funcionalidad del decodificador en forma más detallada.

Se aplican las siguientes notaciones: la longitud de la trama se designa N , la señal de CELP decodificada se indica con $S_C(n)$, la señal de MDCT decodificada (incluyendo la señal de solapamiento enventanada) está indicada con $S_M(n)$, la ventana usada para incluir en ventanas la parte izquierda de la señal de MDCT es $w(n)$ donde L es la longitud de la ventana, y el filtro de síntesis de CELP está indicado con $\frac{1}{A(z)}$ donde $A(z) = \sum_{m=0}^M a_m z^{-m}$ y M es el orden de filtro.

20

Descripción Detallada del Paso 1

25 Después del paso 1 del lado del decodificador (decodificación de la trama de MDCT actual con la misma longitud de MDCT y la misma ventana de MDCT que se utiliza del lado del codificador) obtenemos la trama de MDCT actual decodificada (por ejemplo, una representación en el dominio del tiempo de la “segunda trama de audio” que constituye la segunda información de audio decodificada antes mencionada. Esta trama (por ejemplo, la segunda trama) no contiene aliasing alguno puesto que el punto de pliegue izquierdo se corrió al límite izquierdo entre las tramas de CELP y MDCT (por ejemplo, usando el concepto descrito en detalle con referencia a la Fig. 4b). Esto significa que podemos obtener una reconstrucción perfecta en la trama actual (por ejemplo entre los tiempos $t_2 = 0$ y $t_3 = 20$ ms) con una tasa de bits suficientemente elevada. Sin embargo, a bajas tasas de bits, la señal no necesariamente coincide con la señal de entrada y, de esa manera, se puede introducir una discontinuidad en el límite entre la trama de CELP y MDCT (por ejemplo, en el tiempo $t=0$, como se ilustra en la Fig. 4b).

30
35 Para facilitar su comprensión, se ilustra este problema con referencia a la Fig. 5. Un trazado superior (Fig. 5a) ilustra la señal de CELP decodificada $S_C(n)$, el trazado central (Fig. 5b) ilustra la señal de MDCT decodificada (que incluye la señal de solapamiento enventanada) $S_M(n)$ y un trazado inferior (Fig. 5c) ilustra una señal de salida que se obtiene descartando la señal de solapamiento enventanada y concatenando la trama de CELP y la trama de MDCT. Existe claramente una discontinuidad en la señal de salida (expuesta en la Fig. 5c) en el límite entre dos tramas (por ejemplo, en el tiempo $t=0$ ms).

Ejemplo Comparativo del Procesamiento Adicional

45 Una solución posible a este problema es la estrategia propuesta en la referencia 1 antes citada 1 (“Efficient cross-fade windows for transitions between LPC-based and non-LPC based audio coding” de J. Lecomte et al.), que describe un concepto utilizado en MPEG USAC. En lo sucesivo se presenta una breve descripción de dicha estrategia de referencia.

50 Una segunda versión de la señal de CELP decodificada $\bar{S}_C(n)$ se inicializa en primer lugar de igual manera que la señal de CELP decodificada

$$\bar{S}_C(n) = S_C(n), n = -N, \dots, -1$$

55 luego se introduce de manera artificial el aliasing faltante introducido en la región de solapamiento

$$\bar{S}_C(n) = S_C(n)w(-n-1)w(-n-1) + S_C(-n-L-1)w(n+L)w(-n-1),$$

$$n = -L, \dots, -1$$

60

por último, se obtiene la segunda versión de la señal de CELP decodificada utilizando una operación de solapamiento y suma

$$\bar{S}_C(n) = \bar{S}_C(n) + S_M(n), n = -L, \dots, -1$$

5 Como se puede apreciar en las Figs. 6a a 6d, esta estrategia de comparación elimina la discontinuidad (véase, en particular, la Fig. 6d). El problema con esta estrategia es que introduce un retardo adicional (igual a la longitud del solapamiento), puesto que la trama anterior se modifica una vez decodificada la trama actual. En algunas aplicaciones, como la codificación de audio con bajo retardo, es conveniente (o hasta indispensable) tener un
10 retardo lo más bajo posible.

Descripción Detallada de los Pasos de Procesamiento

15 A diferencia de la estrategia convencional antes citada, la estrategia propuesta aquí para eliminar la discontinuidad no tiene retardo adicional alguno. No modifica la última trama de CELP (también denominada primera trama de audio) sino que, por el contrario, modifica la trama de MDCT actual (también denominada segunda trama de audio codificada en el dominio de la frecuencia posterior a la primera trama de audio codificada en el dominio de la predicción lineal).

20 **Paso a)**

En un primer paso, se computa una “segunda versión” de la última trama de ACELP $\bar{S}_C(n)$ como se describió anteriormente. Por ejemplo, se puede efectuar el siguiente cómputo:

25 Se inicializa en primer lugar una segunda versión de la señal de CELP decodificada $\bar{S}_C(n)$ de igual manera que la señal de CELP decodificada

$$\bar{S}_C(n) = S_C(n), n = -N, \dots, -1$$

30 luego se introduce en forma artificial el aliasing faltante en la región de solapamiento

$$\bar{S}_C(n) = S_C(n)w(-n-1)w(-n-1) + S_C(-n-L-1)w(n+L)w(-n-1),$$

$$n = -L, \dots, -1$$

35 por último, se obtiene la segunda versión de la señal de CELP decodificada empleando una operación de solapamiento y suma

$$\bar{S}_C(n) = \bar{S}_C(n) + S_M(n), n = -L, \dots, -1$$

40 Sin embargo, a diferencia de la referencia 1 (“Efficient cross-fade windows for transitions between LPC-based and non-LPC-based audio coding” de J. Lecomte et al.), no se reemplaza la señal de ACELP decodificada anterior por esta versión de la trama de ACELP anterior para no introducir un retardo adicional. Solo se utiliza como señal intermedia para modificar la trama de MDCT actual como se describe en los siguientes pasos.

45 Expresado en forma diferente, la determinación del estado inicial 144, la combinación de modificación/adición de aliasing 250 o la combinación de modificación/adición de aliasing 342 puede producir, por ejemplo, la señal $\bar{S}_C(n)$ como contribución a la información del estado inicial 146 o a la información de estado inicial combinada 344, o como segunda información de estado inicial 252. Por consiguiente, la determinación del estado inicial 144, la combinación de modificación/adición de aliasing 250 o la combinación de modificación/adición de aliasing 342 puede aplicar, por
50 ejemplo, un enventanado a la señal de CELP decodificada S_C (multiplicación por valores de ventana $w(-n-1)w(-n-1)$), suma una versión espejada en tiempo de la señal de CELP decodificada ($S_C(-n-L-1)$) escalada con un enventanado ($w(n+L)w(-n-1)$) y suma la señal de MDCT decodificada $S_M(n)$, para obtener de esa
55 manera una contribución a la información del estado inicial 146, 344, o incluso para obtener la segunda información de estado inicial 252.

Paso b)

60 El concepto comprende asimismo generar dos señales computando la respuesta de entrada cero (ZIR) del filtro de síntesis de CELP (que se puede considerar, en términos generales, como filtro de predicción lineal) usando dos

memorias diferentes (también denominadas estados iniciales) para los filtros de síntesis de CELP.

La primera ZIR $s_Z^1(n)$ se genera usando la señal de CELP decodificada anterior $S_C(n)$ como memorias para el filtro de síntesis de CELP.

5

$$s_Z^1(n) = S_C(n), n = -L, \dots, -1$$

$$s_Z^1(n) = - \sum_{m=1}^M a_m s_Z^1(n-m), n = 0, \dots, N-1$$

10 *donde $M \leq L$*

Se genera la segunda ZIR $s_Z^2(n)$ utilizando la segunda versión de la señal de CELP decodificada anterior $\bar{S}_C(n)$ como memorias para el filtro de síntesis de CELP.

15

$$s_Z^2(n) = \bar{S}_C(n), n = -L, \dots, -1$$

$$s_Z^2(n) = - \sum_{m=1}^M a_m s_Z^2(n-m), n = 0, \dots, N-1$$

donde $M \leq L$

20

Se debe tener en cuenta que la primera respuesta de entrada cero y la segunda respuesta de entrada cero pueden ser computadas por separado, donde se puede obtener la primera respuesta de entrada cero sobre la base de la primera información de audio decodificada (por ejemplo, usando la determinación de estado inicial 242 y filtrado de predicción lineal 246) y donde se puede computar la segunda respuesta de entrada cero, por ejemplo, usando la combinación de modificación/adición de aliasing 250, que puede producir la "segunda versión de la trama de CELP anterior $\bar{S}_C(n)$ " dependiendo de la primera información de audio decodificada 222 y la segunda información de audio decodificada 232, y también empleando el segundo filtrado de predicción lineal 254. Por otro lado, sin embargo, se puede aplicar un único filtrado de síntesis de CELP. Por ejemplo, se puede aplicar un filtrado de predicción lineal 148, 346, donde se utiliza una suma de $S_C(n)$ y $\bar{S}_C(n)$ como entrada para dicho filtrado de predicción lineal (combinado).

25

30

Esto se debe a que el filtrado de predicción lineal es una operación lineal, por lo que la combinación se puede llevar a cabo antes del filtrado o después del filtrado sin cambiar el resultado. Sin embargo, dependiendo de los signos, también se puede utilizar una diferencia entre $S_C(n)$ y $\bar{S}_C(n)$ como estado inicial (en el caso de $n = -L, \dots, -1$) del filtrado de predicción lineal (combinado).

35

Para finalizar, se puede obtener la primera información de estado inicial $s_Z^1(n)$, $n = -L, \dots, -1$ y la segunda información de estado inicial $s_Z^2(n)$, $n = -L, \dots, -1$ en forma individual o en forma combinada. Además, la primera y segunda respuestas de entrada cero pueden ser obtenidas merced a un filtrado de predicción lineal individual de la información de estado inicial individual, o bien empleando un filtrado de predicción lineal (combinado) sobre la base de una información de estado inicial combinada.

40

Como se demuestra en los trazados de la Fig. 7, que se pasa a explicar en detalle a continuación, $S_C(n)$ y $s_Z^1(n)$ son continuos, $\bar{S}_C(n)$ y $s_Z^2(n)$ son continuos. Más aún, como $\bar{S}_C(n)$ y $S_M(n)$ también son continuos, $S_M(n) - s_Z^2(n)$ es una señal que se inicia con un valor muy próximo a 0.

45

Haciendo referencia ahora a la Fig. 7, se explican algunos detalles.

La Fig. 7a ilustra una representación gráfica de una trama de CELP anterior y de una primera respuesta de entrada cero. Una abscisa 710 describe un tiempo en milisegundos y la ordenada 712 describe una amplitud en unidades arbitrarias.

50

Por ejemplo, se ilustra una señal de audio emitida respecto de la trama de CELP anterior (también denominada

primera trama de audio) entre los tiempos t_{71} y t_{72} . Por ejemplo, la señal $s_c(n)$ correspondiente a $n < 0$ puede aparecer entre los tiempos t_{71} y t_{72} . Más aún, se puede mostrar la primera respuesta de entrada cero entre los tiempos t_{72} y t_{73} . Por ejemplo, se puede presentar la primera respuesta de entrada cero $s_z^1(n)$ entre los tiempos t_{72} y t_{73} .

5 La Fig. 7b ilustra una representación gráfica de la segunda versión de la trama de CELP anterior y la segunda respuesta de entrada cero. Una abscisa está indicada por el número 720, e ilustra el tiempo en milisegundos. Una ordenada recibe la designación 722 e ilustra una amplitud en unidades arbitrarias. Se ilustra una segunda versión de la trama de CELP anterior entre los tiempos t_{71} (-20 ms) y t_{72} (0 ms), y se presenta la segunda respuesta de entrada
 10 cero entre los tiempos t_{72} y t_{73} (+20 ms). Por ejemplo, se ilustra la señal $\bar{s}_c(n)$, $n < 0$, entre los tiempos t_{71} y t_{72} . Más aún, la señal $s_z^2(n)$ correspondiente a $n \geq 0$ aparece entre los tiempos t_{72} y t_{73} .

Más aún, la diferencia entre $s_M(n)$ y $s_z^2(n)$ está expuesta en la Fig. 7c, donde una abscisa 730 designa un tiempo en milisegundos y donde una ordenada 732 designa una amplitud en unidades arbitrarias.

15 Más aún, se debe tener en cuenta que la primera respuesta de entrada cero $s_z^1(n)$ correspondiente a $n \geq 0$ es una continuación (sustancialmente) estable de la señal $s_c(n)$ correspondiente a $n < 0$. De manera similar, la segunda respuesta de entrada cero $s_z^2(n)$ correspondiente a $n \geq 0$ es una continuación (sustancialmente) estable de la señal $\bar{s}_c(n)$ correspondiente a $n < 0$.

20 **Paso c)**

La señal de MDCT actual (por ejemplo, la segunda información de audio decodificada 132, 232, 332) es reemplazada por una segunda versión 142, 242, 342 de la señal de MDCT actual (es decir de la señal de MDCT asociada a la segunda trama de audio actual).

$$\bar{s}_M(n) = s_M(n) - s_z^2(n) + s_z^1(n)$$

Por lo tanto es sencillo demostrar que $s_c(n)$ y $\bar{s}_M(n)$ son continuas: $s_c(n)$ y $s_z^1(n)$ son continuas, $s_M(n) - s_z^2(n)$ se inicia con un valor muy cercano a 0.

Por ejemplo, $\bar{s}_M(n)$ puede ser determinado por la modificación 152, 258, 350 dependiendo de la segunda información de audio decodificada 132, 232, 323 y dependiendo de la primera respuesta de entrada cero $s_z^1(n)$ y la segunda respuesta de entrada cero $s_z^2(n)$ (por ejemplo como se ilustra en la Fig. 2), o dependiendo de una
 35 respuesta de entrada cero combinada (por ejemplo, la respuesta de entrada cero combinada $s_z^1(n) - s_z^2(n)$, 150, 348). Como se puede apreciar en los tratados de la Fig. 8, la estrategia propuesta elimina la discontinuidad.

Por ejemplo, la Fig. 8a ilustra una representación gráfica de las señales correspondientes a la trama de CELP anterior (por ejemplo, de la primera información de audio decodificada), donde una abscisa 810 describe un tiempo en milisegundos, y donde una ordenada 812 describe una amplitud en unidades arbitrarias. Como se puede apreciar, la primera información de audio decodificada se emite (por ejemplo, en virtud de la decodificación en el dominio de la predicción lineal) entre los tiempos t_{81} (-20 ms) y t_{82} (0 ms).

Más aún, como se puede apreciar en la Fig. 8b, la segunda versión de la trama de MDCT actual (por ejemplo, la segunda información de audio decodificada modificada 142, 242, 342) es emitida solo a partir del tiempo t_{82} (0 ms), aunque la segunda información de audio decodificada 132, 232, 332 se emite por lo general a partir de t_4 (como se ilustra en la Fig. 4b). Se debe tener en cuenta que la segunda información de audio decodificada 132, 232, 332 emitida entre los tiempos t_4 y t_2 (como se ilustra en la Fig. 4b) no se utiliza directamente para la emisión de la segunda versión de la trama de MDCT actual (señal $\bar{s}_M(n)$) aunque se utiliza simplemente para la emisión de
 50 componentes de la señal $s_z^2(n)$. Por motivos de claridad, se debe tener en cuenta que una abscisa 820 designa el tiempo en milisegundos, y que una ordenada 822 designa una amplitud en términos de unidades arbitrarias.

La Fig. 8c ilustra una concatenación de la trama de CELP anterior (como se ilustra en la Fig. 8a) y de la segunda versión de la trama de MDCT actual (como se ilustra en la Fig. 8b). Una abscisa 830 describe un tiempo en milisegundos, y una ordenada 832 describe una amplitud en términos de unidades arbitrarias. Como se puede apreciar, hay una transición sustancialmente continua entre la trama de CELP anterior (entre los tiempos t_{81} y t_{82} y la

segunda versión de la trama de MDCT actual (comenzando en el tiempo t_{82} y terminando, por ejemplo, en el tiempo t_5 , como se ilustra en la Fig. 4b). De esa manera, se evitan las distorsiones audibles en una transición de la primera trama (que es codificada en el dominio de la predicción lineal) a la segunda trama (que es codificada en el dominio de la frecuencia).

5 También resulta sencillo demostrar que se obtiene una reconstrucción perfecta a altas tasas: a tasas elevadas, $s_c(n)$ y $\bar{s}_c(n)$ son muy similares y ambas son similares a la señal de entrada, luego las dos ZIR son muy similares; en consecuencia, la diferencia de las dos ZIR se acerca mucho a 0 y, por último, $\bar{s}_M(n)$ es muy similar a $s_M(n)$ y ambas son muy similares a la señal de entrada.

10 **Paso d)**

Opcionalmente, se puede aplicar una ventana a las dos ZIR, para no afectar la totalidad de la trama de MDCT actual. Esto es útil, por ej., para reducir la complejidad, o si la ZIR no es aproximada a 0 al final de la trama de MDCT.

Un ejemplo de ventana es una sencilla ventana lineal $v(n)$ de una longitud P

$$v(n) = \frac{P-n}{P}, n = 0, \dots, P-1$$

donde por ej. $P = 64$.

20 Por ejemplo, la ventana puede procesar la respuesta de entrada cero 150, las respuestas de entrada cero 248, 256 o la respuesta de entrada cero combinada 348.

5.8. Método de Acuerdo con la Fig. 9

25 La Fig. 9 ilustra un gráfico de flujo de un método para emitir una información de audio decodificada sobre la base de una información de audio codificada. El método 900 comprende emitir 910 una primera información de audio decodificada sobre la base de una trama de audio codificada en un dominio de la predicción lineal. El método 900 comprende asimismo emitir 920 una segunda información de audio decodificada sobre la base de una trama de audio codificada en el dominio de la frecuencia. El método 900 comprende asimismo obtener 930 una respuesta de entrada cero de un filtrado de predicción lineal, donde el estado inicial del filtrado de predicción lineal se define de conformidad con la primera información de audio decodificada y la segunda información de audio decodificada.

35 El método 900 comprende además modificar 940 la segunda información de audio decodificada, que se emite sobre la base de una trama de audio codificada en el dominio de la frecuencia posterior a una trama de audio codificada en el dominio de la predicción lineal, dependiendo de la respuesta de entrada cero, para obtener una transición suave entre la primera información de audio decodificada y la segunda información de audio decodificada modificada.

40 El método 900 puede ser complementado por cualquiera de las características y funcionalidades aquí descritas, también en lo que respecta a los decodificadores de audio.

5.10. Método de Acuerdo con la Fig. 10

45 La Fig. 10 ilustra un gráfico de flujo de un método 1000 para emitir una información de audio decodificada sobre la base de una información de audio codificada T.

El método 1000 comprende ejecutar 1010 una decodificación en el dominio de la predicción lineal para emitir una primera información de audio decodificada sobre la base de una trama de audio codificada en un dominio de la predicción lineal.

50 El método 1000 comprende asimismo ejecutar 1020 una decodificación en el dominio de la frecuencia para emitir una segunda información de audio decodificada sobre la base de una trama de audio codificada en un dominio de la frecuencia.

55 El método 1000 comprende asimismo obtener 1030 una primera respuesta de entrada cero de un filtrado de predicción lineal en respuesta a un primer estado inicial del filtrado de predicción lineal definido por la primera información de audio decodificada y obtener 1040 una segunda respuesta de entrada cero del filtrado de predicción lineal en respuesta a un segundo estado inicial del filtrado de predicción lineal definido por una versión modificada de la primera información de audio decodificada, que se emite con un aliasing artificial, y que comprende la contribución de una porción de la segunda información de audio decodificada.

60 Por otro lado, el método 1000 comprende obtener 1050 una respuesta de entrada cero combinada del filtrado de

predicción lineal en respuesta a un estado inicial del filtrado de predicción lineal definido por una combinación de la primera información de audio decodificada y de una versión modificada de la primera información de audio decodificada, que se emite con un aliasing artificial, y que comprende la contribución de una porción de una segunda información de audio decodificada.

5 El método 1000 comprende asimismo modificar 1060 la segunda información de audio decodificada, que se emite sobre la base de una trama de audio codificada en el dominio de la frecuencia que sigue a una trama de audio codificada en el dominio de la predicción lineal, dependiendo de la primera respuesta de entrada cero y la segunda respuesta de entrada cero, o dependiendo de la respuesta de entrada cero combinada, para obtener una transición suave entre la primera información de audio decodificada y la segunda información de audio decodificada modificada.

Se debe tener en cuenta que el método 1000 puede ser complementado por cualquiera de las características y funcionalidades aquí descritas, también en lo que respecta a los decodificadores de audio.

6. Conclusiones

Para concluir, las formas de realización de acuerdo con la invención se relacionan con las transiciones de CELP a MDCT. Estas transiciones introducen en general dos problemas:

1. Aliasing debido a la trama de MDCT anterior faltante; y
2. Discontinuidad en el límite entre la trama de CELP y la trama de MDCT, debido a la naturaleza no perfecta de codificación de la forma de onda de los dos esquemas de codificación operando a tasas de bits bajas/medianas.

En las formas de realización de acuerdo con la invención, el problema de aliasing se soluciona incrementando la longitud de MDCT de tal manera que el punto de pliegue izquierdo se corre a la izquierda del límite entre las tramas de CELP y de MDCT. La parte izquierda de la ventana de MDCT también se modifica de tal manera que se reduzca el solapamiento. A diferencia de las soluciones convencionales, la señal de CELP no se modifica para no introducir ningún retardo adicional. Por el contrario, se crea un mecanismo para eliminar cualquier discontinuidad que se pudiera introducir en el límite entre las tramas de CELP y de MDCT. Este mecanismo alisa la discontinuidad utilizando la respuesta de entrada cero de los filtros de síntesis de CELP. Aquí se describen los detalles adicionales.

7. Alternativas de Implementación

Si bien se han descrito algunos aspectos en el contexto de un aparato, es obvio que estos aspectos también representan una descripción del método correspondiente, en el cual un bloque o dispositivo corresponde a un paso del método o a una característica de un paso del método. De manera análoga, los aspectos descritos en el contexto de un paso del método también representan una descripción de un bloque o artículo correspondiente o de una característica de un aparato correspondiente. Algunos o todos los pasos del método pueden ser ejecutados por (o utilizando) un aparato de hardware, como por ejemplo un microprocesador, una computadora programable o un circuito electrónico. En algunas formas de realización, uno o más de los pasos más importantes del método pueden ser ejecutados por ese tipo de aparato.

La señal de la invención transmitida o codificada puede ser almacenada en un medio de almacenamiento digital o puede ser transmitida por un medio de transmisión tal como un medio de transmisión inalámbrico o un medio de transmisión por cable tal como Internet.

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

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

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

Otras formas de realización comprenden el programa de computación para ejecutar uno de los métodos aquí descritos, almacenado en un portador legible por una máquina.

Dicho de otro modo, una forma de realización del método de la invención consiste, por lo tanto, en un programa de computación que consta de un código de programa para realizar uno de los métodos aquí descritos al ejecutarse el programa de computación en una computadora.

5 Otra forma de realización de los métodos de la invención consiste, por lo tanto, en un portador de datos (o medio de almacenamiento digital, o medio legible por computadora) que comprende, grabado en el mismo, el programa de computación para ejecutar uno de los métodos aquí descritos. El portador de datos, el medio de almacenamiento digital o el medio grabado son por lo general tangibles y/o no transitorios.

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

15 Otra forma de realización comprende un medio de procesamiento, por ejemplo una computadora, o un dispositivo lógico programable, configurado o adaptado para ejecutar uno de los métodos aquí descritos.

20 Otra forma de realización comprende una computadora en la que se ha instalado el programa de computación para ejecutar uno de los métodos aquí descritos.

25 Otra forma de realización de acuerdo con la invención comprende un aparato o un sistema configurado para transferir (por ejemplo por vía electrónica u óptica) un programa de computación para transferir uno de los métodos aquí descritos a un receptor. El receptor puede ser, por ejemplo, una computadora, un dispositivo móvil, un dispositivo de memoria o similar. El aparato o sistema puede comprender, por ejemplo, un servidor de archivos para transferir el programa de computación a un receptor.

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

35 El aparato aquí descrito puede ser implementado utilizando un aparato de hardware, o utilizando una computadora, o utilizando una combinación de un aparato de hardware y una computadora.

Los métodos descritos en la presente se pueden poner en práctica utilizando un aparato de hardware, o utilizando una computadora, o utilizando una combinación de un aparato de hardware y una computadora.

40 Las formas de realización precedentemente descritas son meramente ilustrativas de los principios de la presente invención. Se entiende que las modificaciones y variaciones de las disposiciones y los detalles aquí descritos han de ser evidentes para otras personas con capacitación en la técnica. Por lo tanto, solo se pretende limitarse al alcance de las siguientes reivindicaciones de patente y no a los detalles específicos presentados a manera de ilustración y explicación de las formas de realización de la presente.

45

REIVINDICACIONES

1. Un decodificador de audio (100;200;300) para emitir una información de audio decodificada (112;212;312) sobre la base de una información de audio codificada (110;210;310), donde el decodificador de audio comprende:
- 5 un decodificador en el dominio de la predicción lineal (120;220;320) configurado para emitir una primera información de audio decodificada (122;222;322; $S_C(n)$) sobre la base de una trama de audio codificada en un dominio de la predicción lineal;
- 10 un decodificador en el dominio de la frecuencia (130;230;330) configurado para emitir una segunda información de audio decodificada (132;232;332; $S_M(n)$) sobre la base de una trama de audio codificada en un dominio de la frecuencia y
- 15 un procesador de transición (140; 240;340), donde el procesador de transición está configurado para obtener una respuesta de entrada cero (150; 256;348) de un filtrado de predicción lineal (148; 254; 346), donde se define un estado inicial (146;252;344) del filtrado de predicción lineal de conformidad con la primera información de audio decodificada y la segunda información de audio decodificada y
- 20 donde el procesador de transición está configurado para modificar la segunda información de audio decodificada (132; 232; 332; $S_M(n)$), que se emite sobre la base de una trama de audio codificada en el dominio de la frecuencia que sigue a una trama de audio codificada en el dominio de la predicción lineal, dependiendo de la respuesta de entrada cero, para obtener una transición suave entre la primera información de audio decodificada ($S_C(n)$) y la segunda información de audio decodificada modificada ($\bar{S}_M(n)$).
2. El decodificador de audio de acuerdo con la reivindicación 1, en el cual el procesador de transición está configurado para obtener una primera respuesta de entrada cero (248; $s_z^1(n)$) de un filtro de predicción lineal (246) en respuesta a un primer estado inicial (244; $S_C(n)$) del filtro de predicción lineal definido por la primera información de audio decodificada (222; $S_C(n)$) y
- 25 donde el procesador de transición está configurado para obtener una segunda respuesta de entrada cero (256; $s_z^2(n)$) del filtro de predicción lineal (254) en respuesta a un segundo estado inicial (252) del filtro de predicción lineal definida por una versión modificada ($\bar{S}_C(n)$) de la primera información de audio decodificada (222, $S_C(n)$), que se emite con un aliasing artificial y que comprende la contribución de una porción de la segunda información de audio decodificada (232, $S_M(n)$) o
- 30 donde el procesador de transición está configurado para obtener una respuesta combinada de entrada cero (150; 348) del filtro de predicción lineal (148;346) en respuesta a un estado inicial (146;344) del filtro de predicción lineal definida por una combinación de la primera información de audio decodificada (122;322; $S_C(n)$) y de una versión modificada ($\bar{S}_C(n)$) de la primera información de audio decodificada (122;322; $S_C(n)$), que se emite con un aliasing artificial, y que comprende la contribución de una porción de la segunda información de audio decodificada (132;332; $S_M(n)$);
- 35 donde el procesador de transición está configurado para modificar la segunda información de audio decodificada (132;232;332; $S_M(n)$), que se emite sobre la base de una trama de audio codificada en el dominio de la frecuencia que sigue a una trama de audio codificada en el dominio de la predicción lineal, dependiendo de la primera respuesta de entrada cero (248; $s_z^1(n)$) y la segunda respuesta de entrada cero (256; $s_z^2(n)$), o dependiendo de la respuesta combinada de entrada cero (150; $s_z^1(n) - s_z^2(n)$), para obtener una transición suave entre la primera información de audio decodificada (122; 222;322; $S_C(n)$) y la segunda información de audio decodificada modificada
- 45 (142;242;342; $\bar{S}_M(n)$).
3. El decodificador de audio (100; 200; 300) de acuerdo con la reivindicación 1 o 2, donde el decodificador en el dominio de la frecuencia (130; 230;330) está configurado para ejecutar una transformada solapada inversa, de tal manera que la segunda información de audio decodificada (132;232;332) comprenda un aliasing.
- 50
4. El decodificador de audio (100;200;300) de acuerdo con la reivindicación 1 o la reivindicación 2 o la reivindicación 3, donde el decodificador en el dominio de la frecuencia (130;230;330) está configurado para ejecutar una transformada solapada inversa, de tal manera que la segunda información de audio decodificada (132;232;332) comprenda un aliasing en una porción de tiempo que está temporalmente solapada con una porción de tiempo respecto de la cual el decodificador en el dominio de la predicción lineal (120;220;320) emite una primera información de audio decodificada (122;222;322), y de tal manera que la segunda información de audio decodificada esté exenta de aliasing en una porción de tiempo siguiente a la porción de tiempo respecto del cual el decodificador en el dominio de la predicción lineal emite una primera información de audio decodificada.
- 55
- 60 5. El decodificador de audio (100; 200;300) de acuerdo con una de las reivindicaciones 2 a 4, en el cual la porción de

la segunda información de audio decodificada (132;232;332), que se utiliza para obtener la versión modificada ($\bar{S}_c(n)$) de la primera información de audio decodificada, comprende un aliasing.

6. El decodificador de audio (100; 200;300) de acuerdo con la reivindicación 5, en el cual el aliasing artificial, que se utiliza para obtener la versión modificada ($\bar{S}_c(n)$) de la primera información de audio decodificada, compensa por lo menos parcialmente el aliasing que se incluye en la porción de la segunda información de audio decodificada (132;232;332), que se utiliza para obtener la versión modificada de la primera información de audio decodificada.

7. El decodificador de audio (100; 200;300) de acuerdo con una de las reivindicaciones 2 a 6, en el cual el procesador de transición (140;240;340) está configurado para obtener la primera respuesta de entrada cero $s_z^1(n)$, o un primer componente $s_z^1(n)$ de la respuesta de entrada cero combinada, de acuerdo con

$$s_z^1(n) = - \sum_{m=1}^M a_m s_z^1(n-m), n = 0, \dots, N-1$$

o de acuerdo con

$$s_z^1(n) = + \sum_{m=1}^M a_m s_z^1(n-m), n = 0, \dots, N-1$$

con

$$s_z^1(n) = S_c(n), n = -L, \dots, -1$$

$$M \leq L$$

donde n designa un índice de tiempo,

donde $s_z^1(n)$ en el caso de $n=0, \dots, N-1$ designa la primera respuesta de entrada cero (248) correspondiente al índice de tiempo n, o un primer componente de la respuesta combinada de entrada cero (150;348) correspondiente al índice de tiempo n;

donde $s_z^1(n)$ en el caso $n=-L, \dots, -1$ designa el primer estado inicial (244) correspondiente al índice de tiempo n, o un primer componente del estado inicial (146;344) correspondiente al índice de tiempo n;

donde m designa una variable de actividad,

donde M designa una longitud de filtro del filtro de predicción lineal;

donde a_m designa coeficientes de filtro del filtro de predicción lineal;

donde $S_c(n)$ designa un valor anteriormente decodificado de la primera información de audio decodificada (122;222;322) correspondiente al índice de tiempo n;

donde N designa una duración del procesamiento.

8. El decodificador de audio (100; 200;300) de acuerdo con una de las reivindicaciones 2 a 7, en el cual el procesador de transición (140;240;340) está configurado para aplicar un primer enventanado ($(w(-n-1)w(-n-1))$) a la primera información de audio decodificada (122;222;322; $S_c(n)$), para obtener una versión incluida en ventana de la primera información de audio decodificada, y para aplicar un segundo enventanado ($(w(n+L)w(-n-1))$) a una versión espejada en tiempo ($S_c(-n-L-1)$) de la primera información de audio decodificada (122;222;322; $S_c(n)$), para obtener una versión incluida en ventana de la versión espejada en tiempo de la primera información de audio decodificada y donde el procesador de transición está configurado para combinar la versión incluida en ventana de la primera información de audio decodificada y la versión incluida en ventana de la versión espejada en tiempo de la primera información de audio decodificada, a fin de obtener la versión modificada ($\bar{S}_c(n)$) de la primera información de audio decodificada.

9. El decodificador de audio de acuerdo con una de las reivindicaciones 2 a 8, donde el procesador de transición (140; 240;340) está configurado para obtener la versión modificada $\bar{S}_c(n)$ de la primera información de audio decodificada $S_c(n)$ de acuerdo con

$$\bar{S}_c(n) = S_c(n)w(-n-1)w(-n-1) + S_c(-n-L-1)w(n+L)w(-n-1) + S_M(n),$$

$$n = -L, \dots, -1,$$

donde n designa un índice de tiempo,

donde $w(-n-1)$ designa un valor de una función de ventana correspondiente al índice de tiempo $(-n-1)$;

donde $w(n+L)$ designa un valor de una función de ventana correspondiente al índice de tiempo $(n+L)$;

5 donde $S_c(n)$ designa un valor anteriormente decodificado de la primera información de audio decodificada (122;222;322) correspondiente al índice de tiempo (n);

donde $S_c(-n-L-1)$ designa un valor anteriormente decodificado de la primera información de audio decodificada correspondiente al índice de tiempo $(-n-L-1)$;

10 donde $S_M(n)$ designa un valor decodificado de la segunda información de audio decodificada (132;232;332) correspondiente al índice de tiempo n; y

donde L describe la longitud de una ventana.

10 El decodificador de audio (100;200;300) de acuerdo con una de las reivindicaciones 2 a 9, en el cual el procesador de transición (140;240;340) está configurado para obtener la segunda respuesta de entrada cero

15 $(256; \bar{s}_z^2(n))$, o un segundo componente $s_z^2(n)$ de la respuesta combinada de entrada cero (150;348) de acuerdo con

$$s_z^2(n) = - \sum_{m=1}^M a_m s_z^2(n-m), n = 0, \dots, N-1$$

20 o de acuerdo con

$$s_z^2(n) = + \sum_{m=1}^M a_m s_z^2(n-m), n = 0, \dots, N-1$$

con

25

$$s_z^2(n) = \bar{s}_c^2(n), n = -L, \dots, -1$$

$$M \leq L$$

30 donde n designa un índice de tiempo,

donde $s_z^2(n)$ en el caso $n=0, \dots, N-1$ designa la segunda respuesta de entrada cero correspondiente al índice de tiempo n, o un segundo componente de la respuesta combinada de entrada cero correspondiente al índice de tiempo n;

35 donde $s_z^2(n)$ en el caso $n=-L, \dots, -1$ designa el segundo estado inicial (252) correspondiente al índice de tiempo n, o un segundo componente del estado inicial (146; 344) correspondiente al índice de tiempo n;

donde m designa una variable de actividad,

donde M designa una longitud de filtro del filtro de predicción lineal (148; 254; 346);

donde a_m designa coeficientes de filtro del filtro de predicción lineal;

40 donde $\bar{s}_c^2(n)$ designa los valores de la versión modificada de la primera información de audio decodificada correspondiente al índice de tiempo n;

donde N designa una duración del procesamiento.

11. El decodificador de audio (100;200;300) de acuerdo con una de las reivindicaciones 2 a 10, en el cual el procesador de transición (140;240;340) está configurado para combinar en forma lineal la segunda información de audio decodificada (132;232;332) con la primera respuesta de entrada cero (248) y la segunda respuesta de entrada cero (256), o con la respuesta combinada de entrada cero (150;348), correspondiente a una porción de tiempo respecto de la cual el decodificador en el dominio de la predicción lineal (120;220;320) no emite una primera información de audio decodificada (122;222;322), a fin de obtener la segunda información de audio decodificada modificada.

50

12. El decodificador de audio (100; 200; 300) de acuerdo con una de las reivindicaciones 2 a 11, donde el procesador de transición (140; 240; 340) está configurado para obtener la segunda información de audio decodificada modificada $\bar{s}_M^2(n)$ de acuerdo con

55
$$\bar{s}_M^2(n) = S_M(n) - s_z^2(n) + s_z^1(n), \text{ en el caso de } n=0, \dots, N-1,$$

o de acuerdo con

$$\bar{S}_M(n) = S_M(n) - v(n)s_Z^2(n) + v(n)s_Z^1(n), \text{ en el caso de } n=0, \dots, N-1,$$

5 donde

donde n designa un índice de tiempo;

donde $S_M(n)$ designa los valores de la segunda información de audio decodificada correspondiente al índice de tiempo n;

10 donde $s_Z^1(n)$ en el caso de $n=0, \dots, N-1$ designa la primera respuesta de entrada cero correspondiente al índice de tiempo n, o un primer componente de la respuesta combinada de entrada cero correspondiente al índice de tiempo n y

donde $s_Z^2(n)$ en el caso de $n=0, \dots, N-1$ designa la segunda respuesta de entrada cero correspondiente al índice de tiempo n, o un segundo componente de la respuesta combinada de entrada cero correspondiente al índice de tiempo n;

15 donde $v(n)$ designa los valores de una función de ventana;
donde N designa una duración del procesamiento.

13. El decodificador de audio (100;200;300) de acuerdo con una de las reivindicaciones 1 a 12, en el cual el procesador de transición (140;240;340) está configurado para dejar la primera información de audio decodificada (122;222;322) sin cambios efectuados por la segunda información de audio decodificada (132;232;332) al emitir una información de audio decodificada correspondiente a una trama de audio codificada en un dominio de la predicción lineal, de tal manera que la información de audio decodificada emitida respecto de una trama de audio codificada en el dominio de la predicción lineal sea emitida en forma independiente de la información de audio decodificada provista respecto de una trama de audio subsiguiente codificada en el dominio de la frecuencia.

14. El decodificador de audio (100;200;300) de acuerdo con una de las reivindicaciones 1 a 13, donde el decodificador de audio está configurado para emitir una información de audio completamente decodificada (122;222;322) correspondiente a una trama de audio codificada en el dominio de la predicción lineal seguida por una trama de audio codificada en el dominio de la frecuencia, antes de decodificar la trama de audio codificada en el dominio de la frecuencia.

15. El decodificador de audio (100;200;300) de acuerdo con una de las reivindicaciones 2 a 14, en el cual el procesador de transición (140;240;340) está configurado para enventanar la primera respuesta de entrada cero (248) y la segunda respuesta de entrada cero (256), o la respuesta combinada de entrada cero (150;348), antes de modificar la segunda información de audio decodificada (132;232;332) dependiendo de la primera respuesta de entrada cero enventanada y la segunda respuesta de entrada cero enventanada, o dependiendo de la respuesta de entrada cero combinada en ventana.

16. El decodificador de audio (100;200;300) de acuerdo con la reivindicación 15, en el cual el procesador de transición está configurado para enventanar la primera respuesta de entrada cero y la segunda respuesta de entrada cero, o la respuesta combinada de entrada cero, utilizando una ventana lineal.

17. Un método (900) para emitir una información de audio decodificada sobre la base de una información de audio codificada, método que comprende:

45 emitir (910) una primera información de audio decodificada ($S_C(n)$) sobre la base de una trama de audio codificada en un dominio de la predicción lineal;

emitir (920) una segunda información de audio decodificada ($S_M(n)$) sobre la base de una trama de audio codificada en un dominio de la frecuencia y

50 obtener (930) una respuesta de entrada cero de un filtrado de predicción lineal, donde el estado inicial del filtrado de predicción lineal se define de conformidad con la primera información de audio decodificada y la segunda información de audio decodificada y

55 modificar (940) la segunda información de audio decodificada ($S_M(n)$), que se emite sobre la base de una trama de audio codificada en el dominio de la frecuencia que sigue a una trama de audio codificada en el dominio de la predicción lineal, dependiendo de la respuesta de entrada cero, para obtener una transición suave entre la primera información de audio decodificada ($S_C(n)$) y la segunda información de audio decodificada modificada ($\bar{S}_M(n)$).

18. Un programa de computación configurado para poner en práctica el método de acuerdo con la reivindicación 17 al ejecutarse el programa de computación en una computadora.

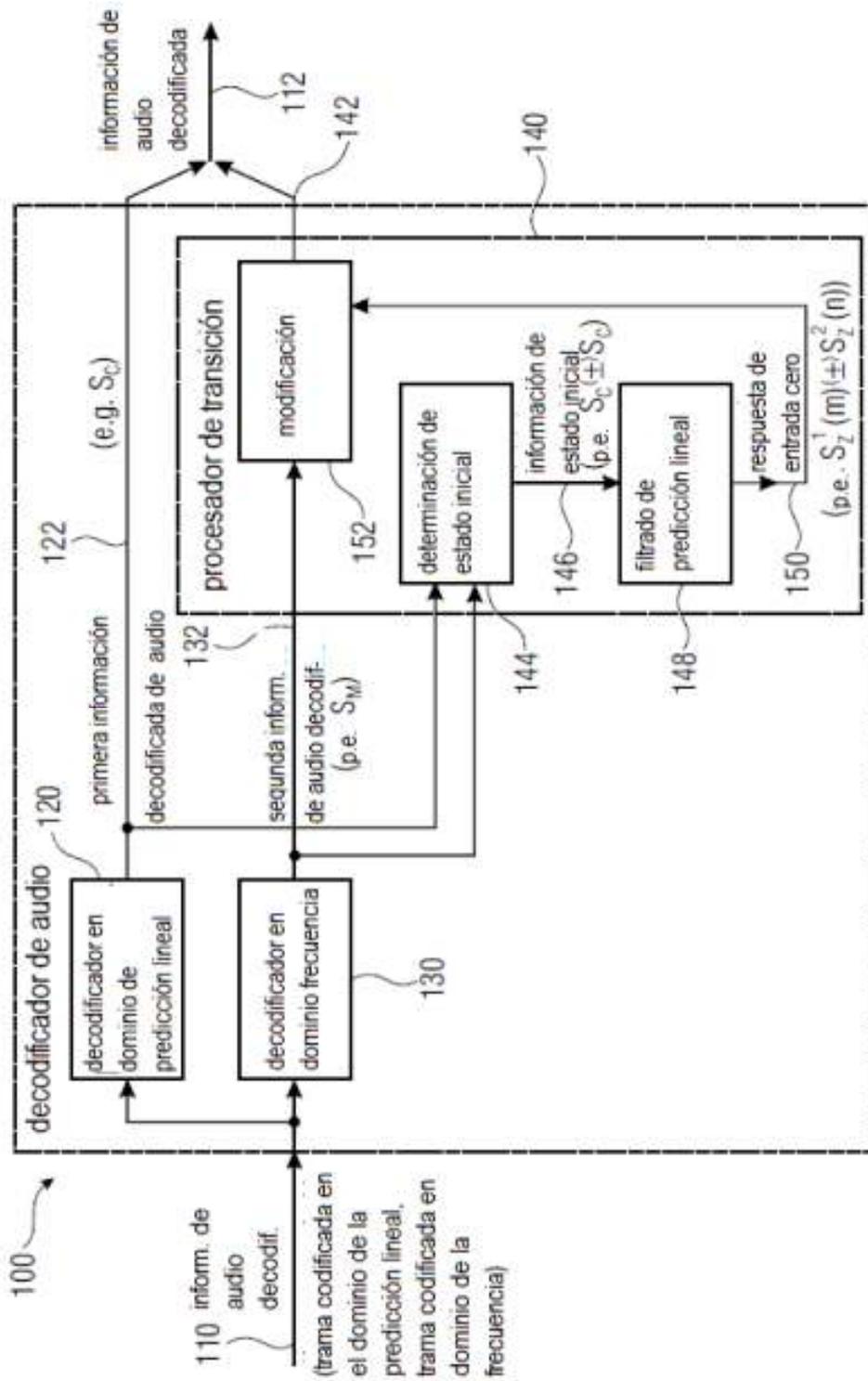


FIG. 1

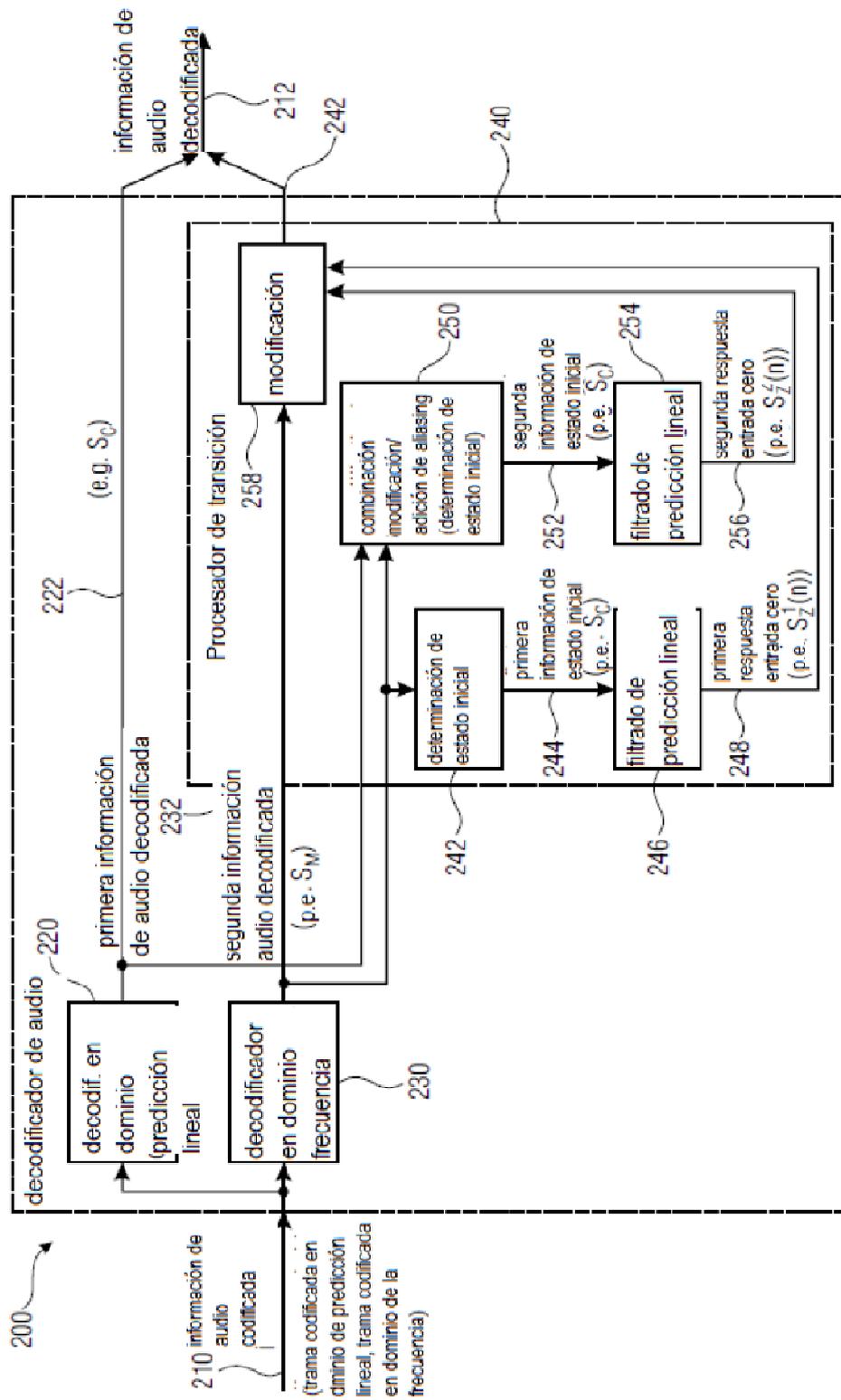


FIG. 2

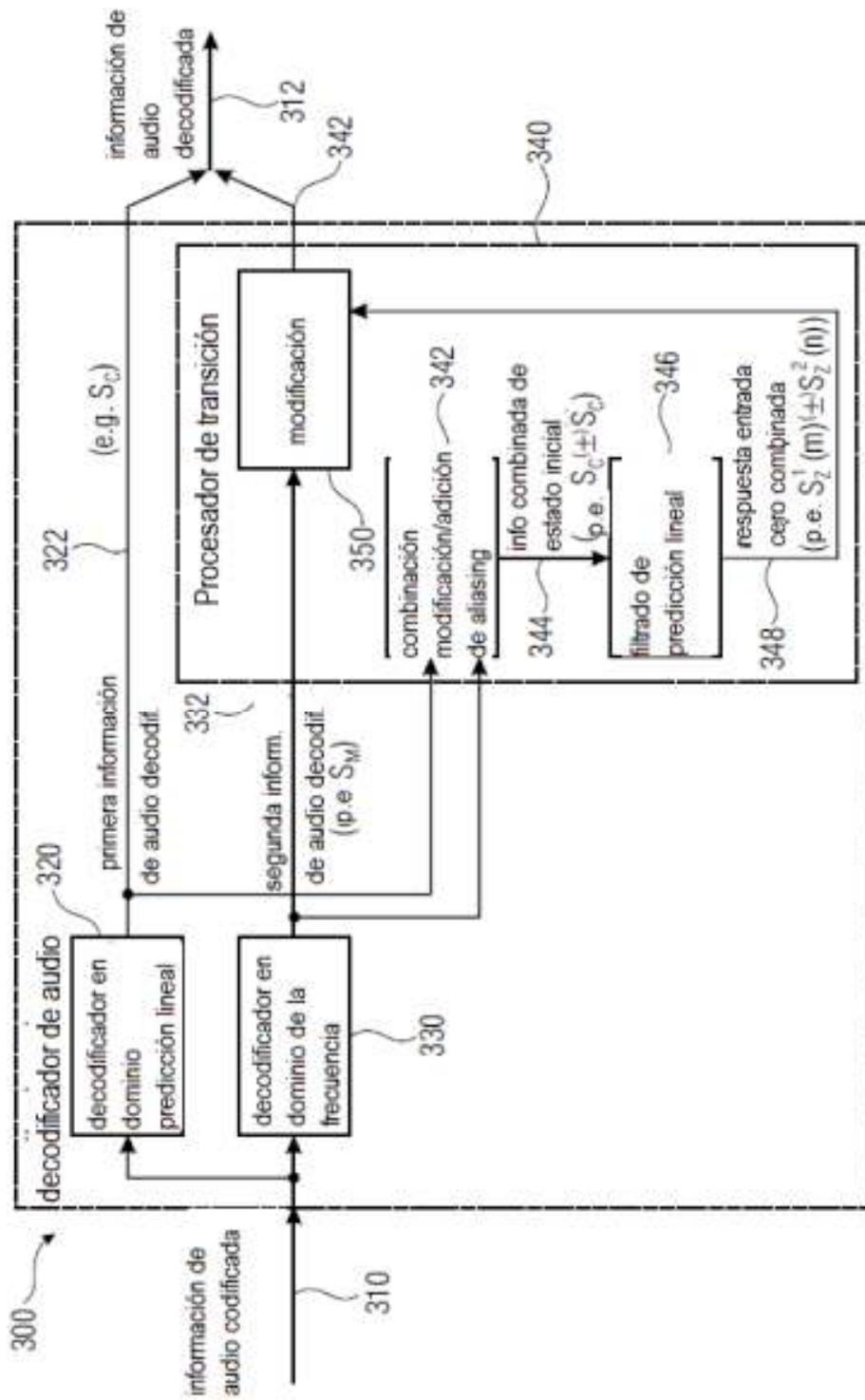


FIG. 3

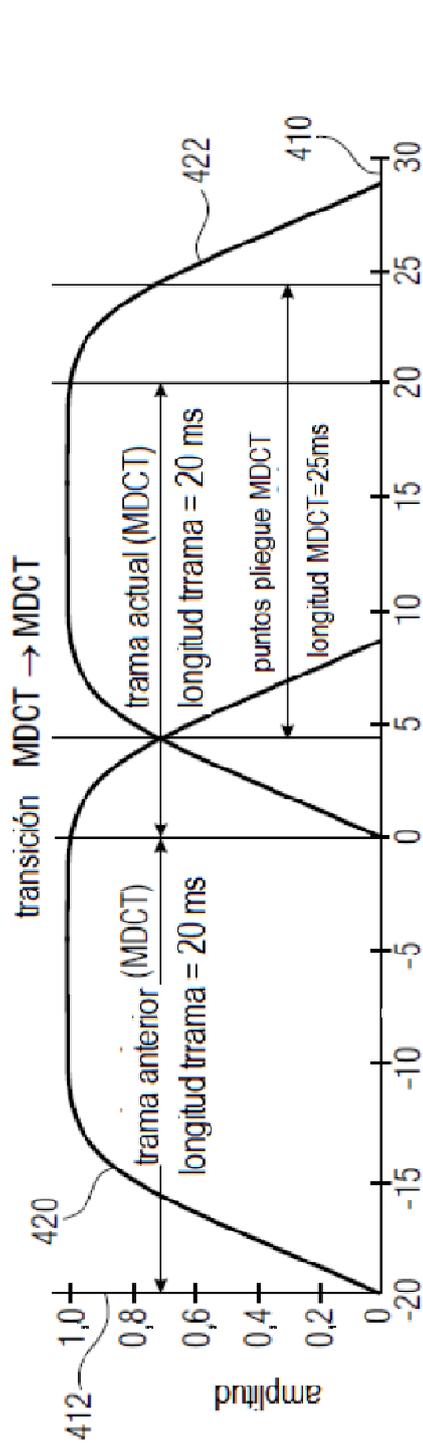


FIG. 4A

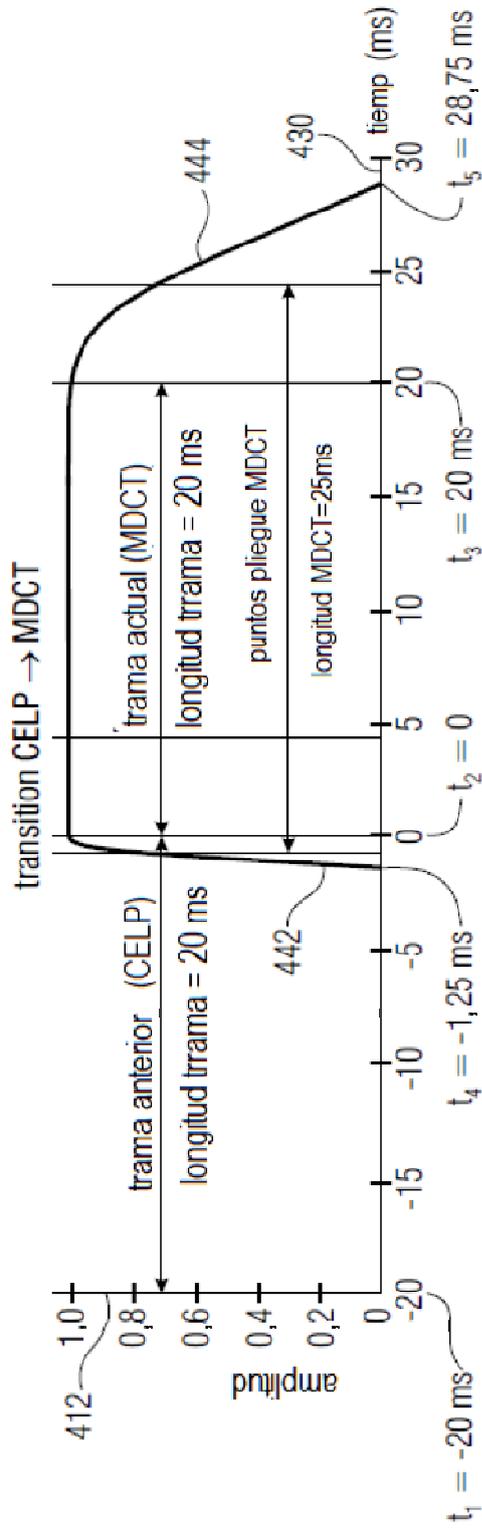


FIG. 4B

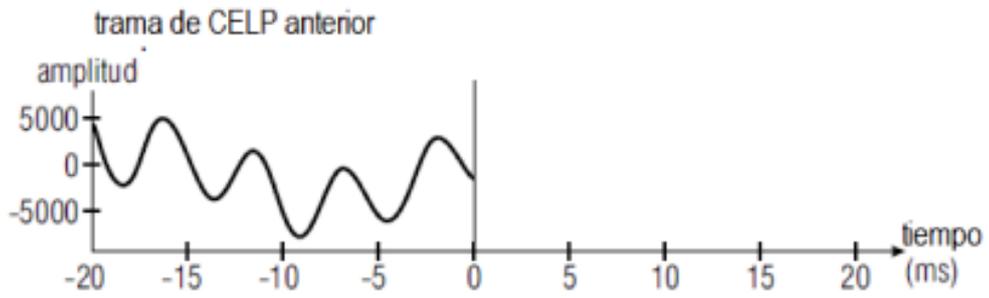


FIG. 5A

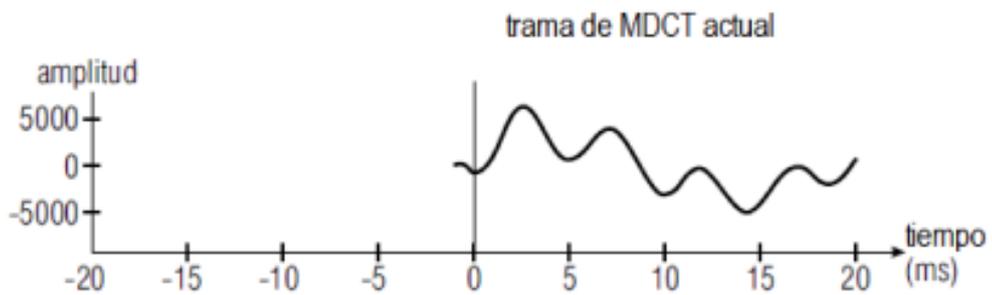


FIG. 5B

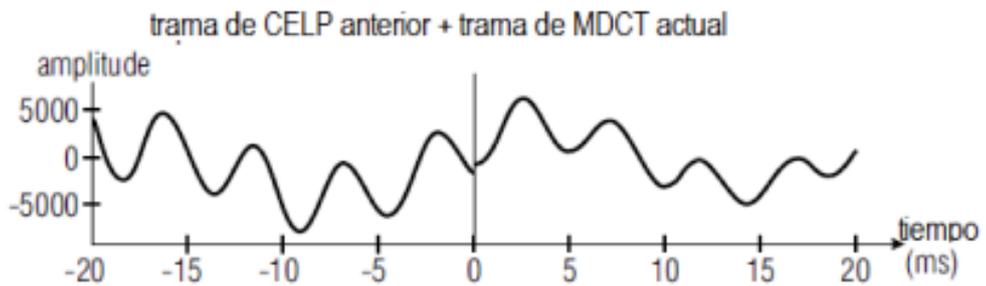


FIG. 5C

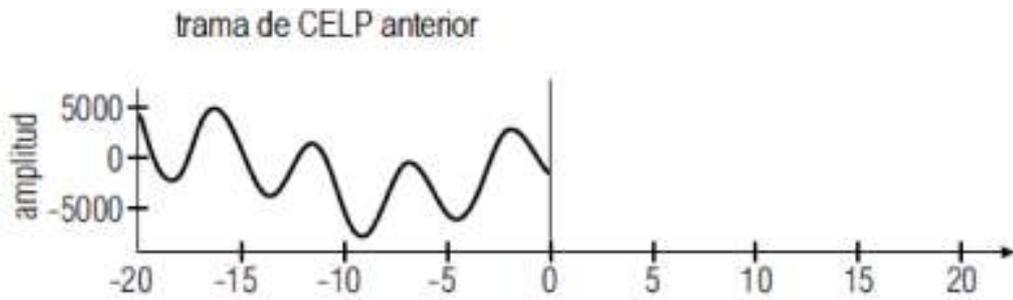


FIG. 6A

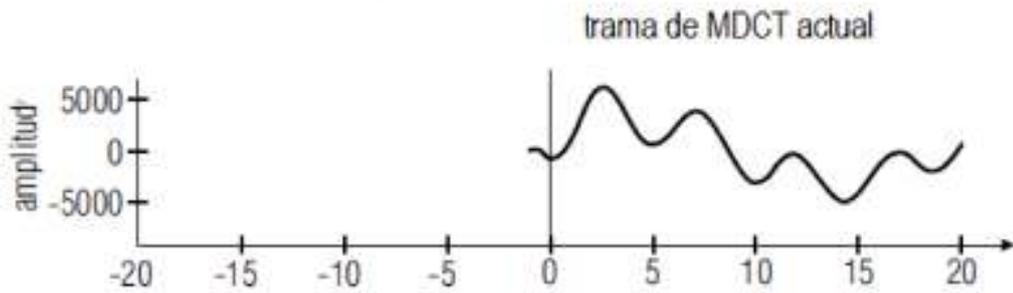


FIG. 6B

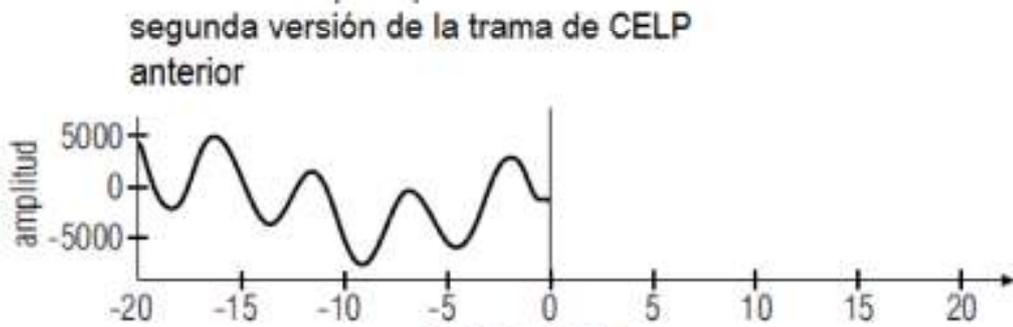


FIG. 6C

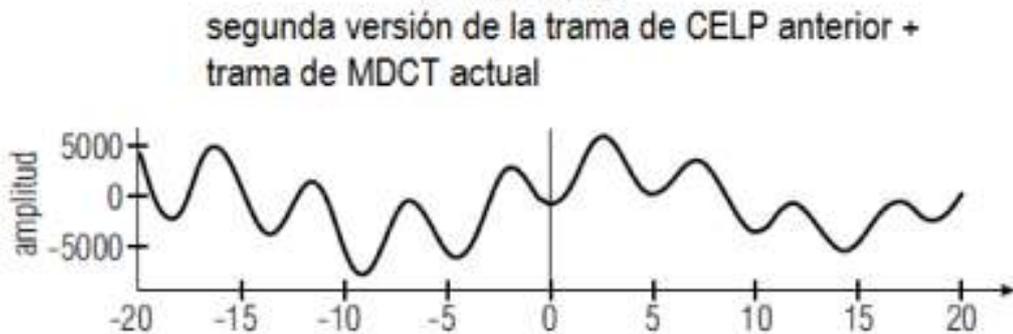


FIG. 6D

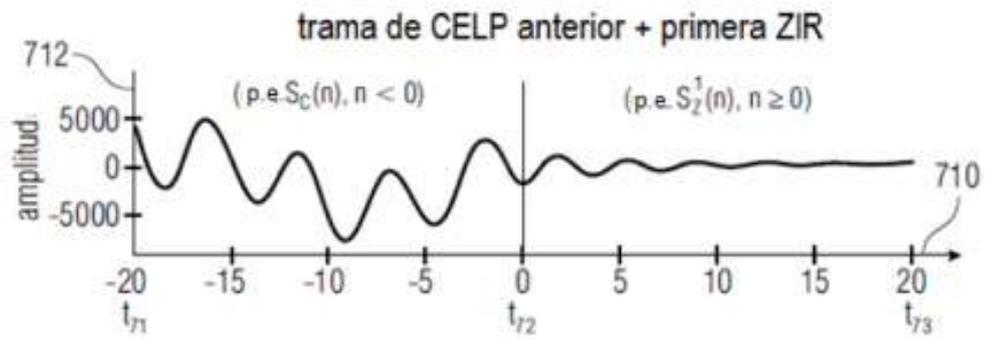


FIG. 7A

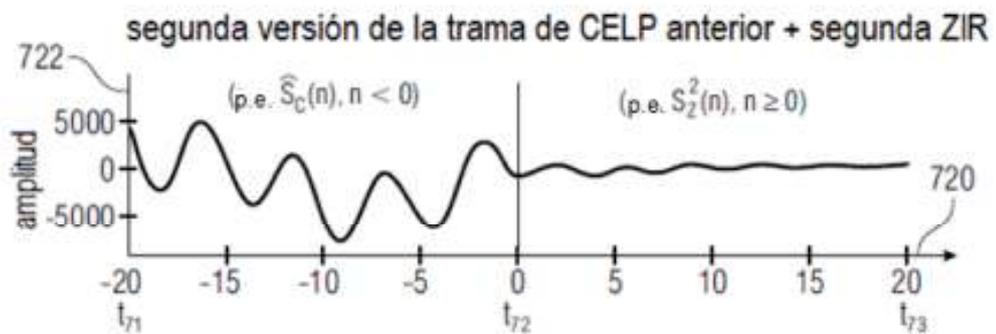


FIG. 7B

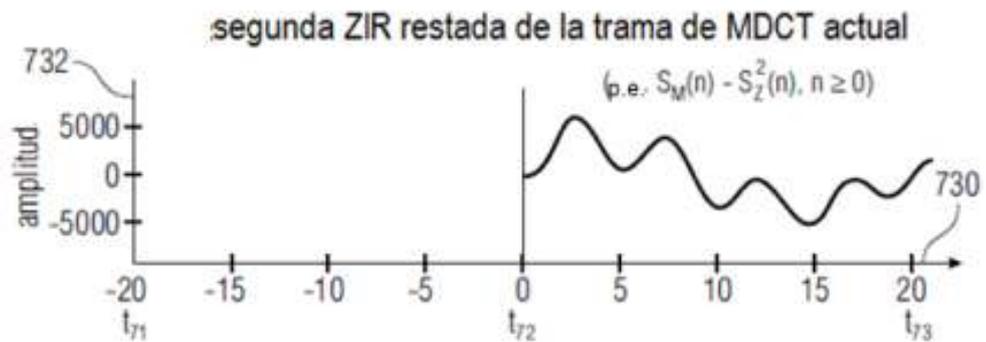


FIG. 7C

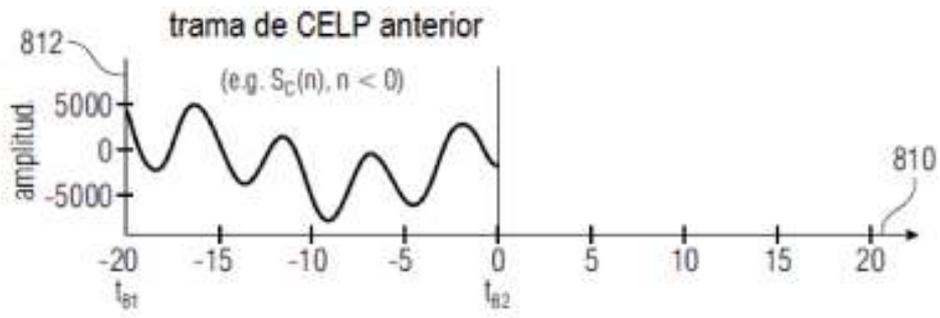


FIG. 8A



FIG. 8B

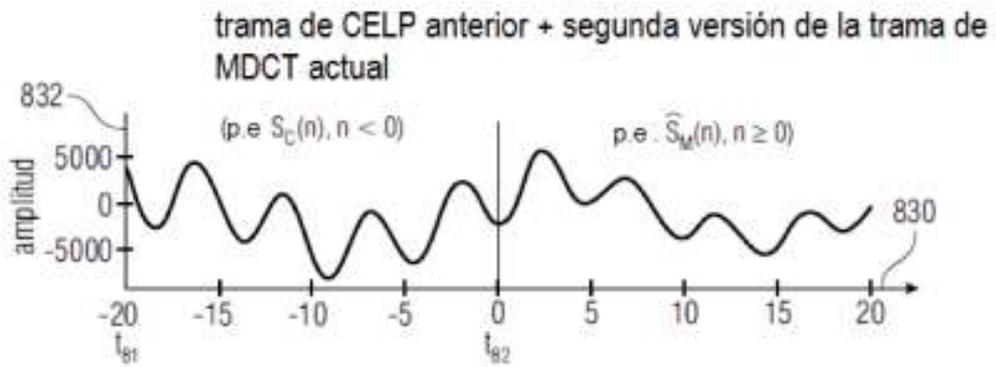


FIG. 8C

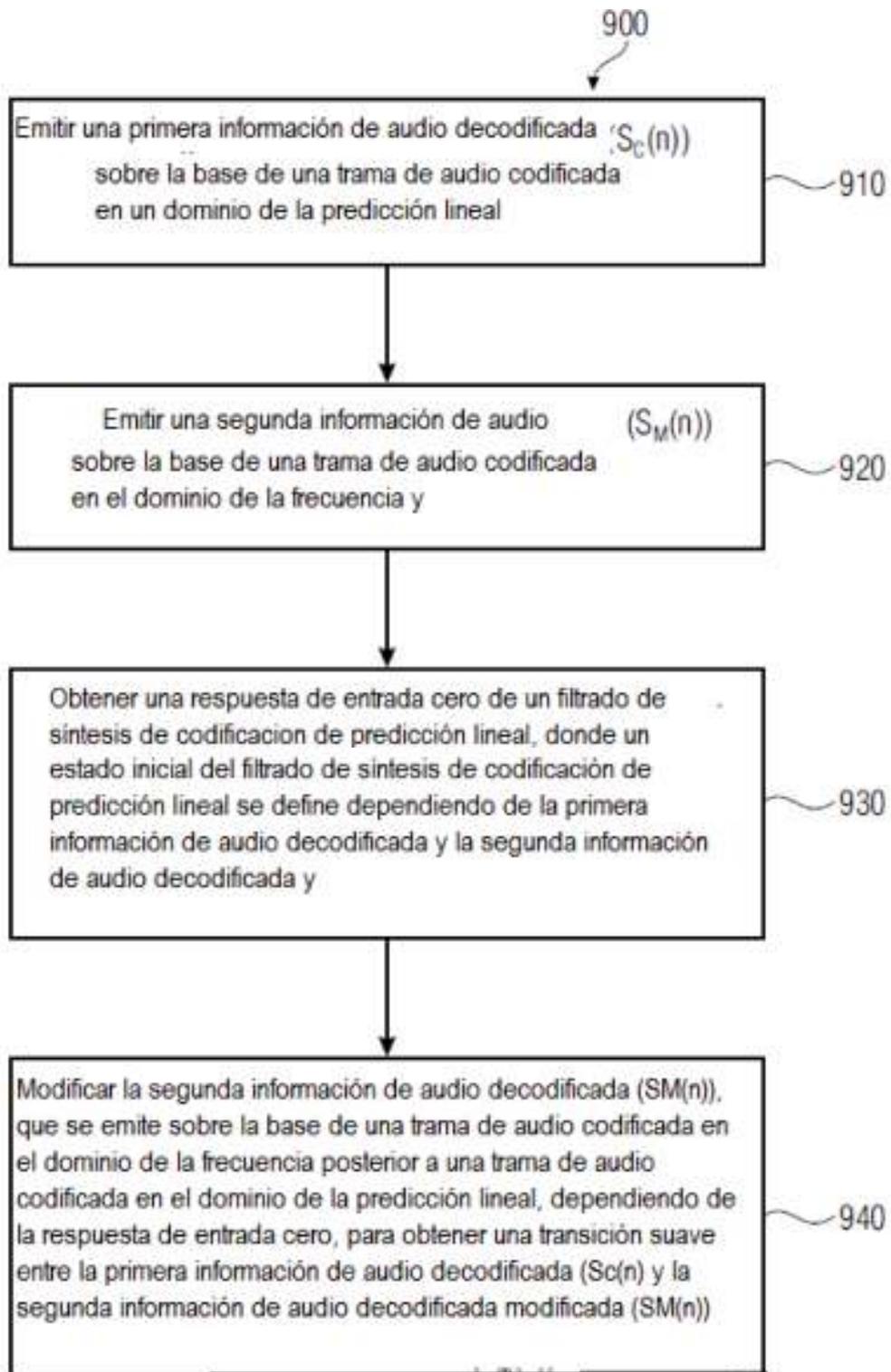


FIG. 9

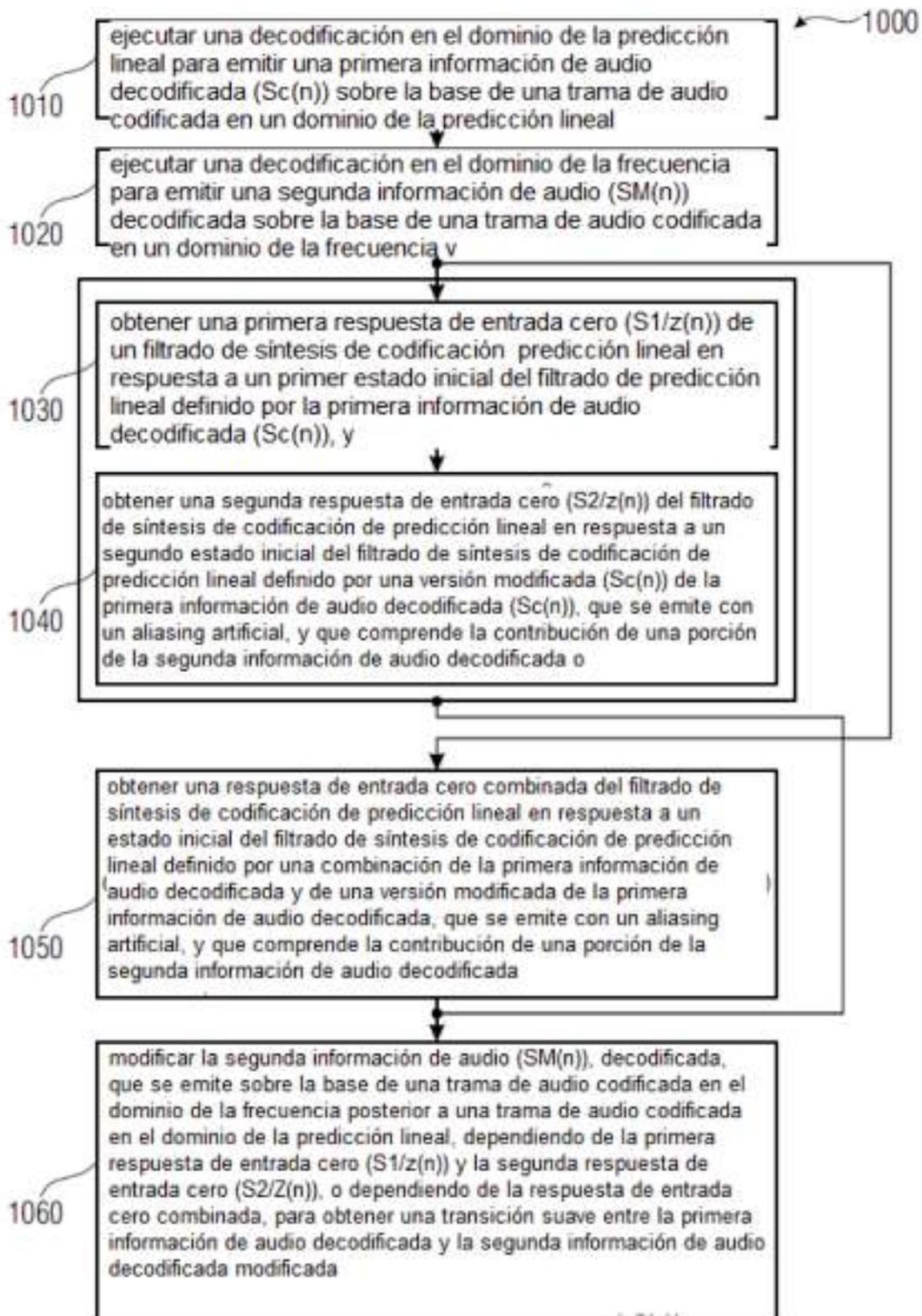


FIG. 10