

19



OFICINA ESPAÑOLA DE  
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 453 098**

51 Int. Cl.:

**G10L 19/083** (2013.01)

**G10L 19/20** (2013.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

96 Fecha de presentación y número de la solicitud europea: **19.10.2010 E 10766284 (3)**

97 Fecha y número de publicación de la concesión europea: **05.03.2014 EP 2491555**

54 Título: **Códec multimodo de audio**

30 Prioridad:

**20.10.2009 US 253440 P**

45 Fecha de publicación y mención en BOPI de la traducción de la patente:  
**04.04.2014**

73 Titular/es:

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

72 Inventor/es:

**GEIGER, RALF;  
FUCHS, GUILLAUME;  
MULTRUS, MARKUS y  
GRILL, BERNHARD**

74 Agente/Representante:

**PONTI SALES, Adelaida**

**ES 2 453 098 T3**

Aviso: En el plazo de nueve meses a contar desde la fecha de publicación en el Boletín europeo de patentes, de la mención de concesión de la patente europea, cualquier persona podrá oponerse ante la Oficina Europea de Patentes a la patente concedida. La oposición deberá formularse por escrito y estar motivada; sólo se considerará como formulada una vez que se haya realizado el pago de la tasa de oposición (art. 99.1 del Convenio sobre concesión de Patentes Europeas).

## DESCRIPCIÓN

Códec multimodo de audio

- 5 **[0001]** La presente invención se refiere a una codificación multimodo de audio tal como un códec unificado de voz y audio o un códec adaptado para señales de audio generales tales como señales de música, voz, mixtas y otras señales, y un esquema de codificación de predicción lineal con excitación por código (CELP, según sus siglas en inglés) adaptada a éste.
- 10 **[0002]** Es conveniente mezclar distintos modos de codificación a fin de codificar señales de audio generales que representan una mezcla de señales de audio de distintos tipos tales como voz, música, o similares. Los modos individuales de codificación pueden estar adaptados para tipos particulares de audio, y por lo tanto, un codificador multimodo de audio puede sacar provecho del cambio de modo de codificación en el transcurso del tiempo correspondiente al cambio del tipo de contenido de audio. En otras palabras, el codificador multimodo de audio puede decidir, por ejemplo, codificar porciones de la señal de audio que tiene contenido de voz utilizando un modo de codificación especialmente dedicado para codificar voz, y para utilizar otro(s) modo(s) de codificación a fin de codificar distintas porciones del contenido de audio que representan contenido de no voz tal como música. Los modos de codificación de predicción lineal tienden a ser más adecuados para codificar contenidos de voz, en tanto que los modos de codificación de dominio de frecuencia tienden a superar el desempeño de los modos de codificación de predicción lineal en lo que respecta a la codificación de música.
- 15 **[0003]** Sin embargo, la utilización de distintos modos de codificación dificulta ajustar globalmente la ganancia dentro de una corriente de bits codificada o, para ser más precisos, la ganancia de la representación decodificada del contenido de audio de una corriente de bits codificada sin tener que decodificar realmente la corriente de bits codificada y luego recodificar la representación decodificada con ajuste de ganancia nuevamente, cuyo desvío disminuiría necesariamente la calidad de la corriente de bits con ajuste de ganancia debido a recuantificaciones realizadas en la recodificación de la representación decodificada y con ajuste de ganancia.
- 20 **[0004]** Por ejemplo, en la Codificación Avanzada de Audio (AAC, por sus siglas en inglés), un ajuste del nivel de salida se puede lograr fácilmente en el nivel de corriente de bits cambiando el valor de la “ganancia global” de campos de 8 bits. Este elemento de corriente de bits puede ser aprobado y editado simplemente, sin necesidad de decodificar y recodificar completamente. Por lo tanto, este proceso no introduce ninguna degradación de calidad y puede deshacerse sin pérdidas. Existen aplicaciones que realmente hacen uso de esta opción. Por ejemplo, existe un software gratuito denominado “ganancia AAC” [AAC gain] que aplica exactamente el procedimiento recientemente descrito. Este software proviene del software gratuito “ganancia MP3”, que aplica la misma técnica para el grupo de expertos en imágenes en movimiento (MPEG, según sus siglas en inglés) 1/2 capa 3.
- 25 **[0005]** En el códec del sistema unificado de codificación de voz y audio (USAC, según sus siglas en inglés) de reciente aparición, el modo de codificación en el dominio de la frecuencia (FD, según sus siglas en inglés) ha heredado la ganancia global de 8 bits de la AAC. Por lo tanto, si USAC se ejecuta en el único modo en FD, tal como para tasas de bits superiores, la funcionalidad del ajuste de nivel se conservaría totalmente, en comparación con AAC. Sin embargo, en cuanto se admiten transiciones de modo, esta posibilidad ya no está presente. En el modo de transformación con código de excitación (TCX, según sus siglas en inglés), por ejemplo, existe también un elemento de corriente de bits que tiene la misma funcionalidad también denominada “ganancia global”, que tiene una longitud de tan sólo 7-bits. En otras palabras, el número de bits para codificar los elementos individuales de ganancia de los modos individuales está adaptado básicamente al modo de codificación respectivo, a fin de lograr un mejor equilibrio entre el menor consumo de bits para el control de ganancia por una parte, y por otra parte evitar una degradación de la calidad debido a una cuantificación demasiado gruesa de la ajustabilidad de ganancia. Evidentemente, este equilibrio tuvo como resultado un número distinto de bits en comparación con el modo TCX y el modo FD. En el modo de predicción lineal excitada por código algebraico (ACELP, según sus siglas en inglés) del estándar USAC que surge actualmente, el nivel puede ser controlado mediante un elemento de corriente de bits “energía media”, que tiene una longitud de 2-bits. Nuevamente, es evidente que el equilibrio entre demasiados bits para la energía media y muy pocos bits para la energía media tuvo como resultado un número distinto de bits en comparación con los otros modos de codificación, a saber el modo de codificación TCX y el modo de codificación FD.
- 30 **[0006]** Por lo tanto, hasta ahora resulta complicado el ajuste global de la ganancia de una representación decodificada, de una corriente de bits codificada por la codificación multimodo y además tiende a disminuir la calidad. En cualquiera de los casos, se ha de realizar la decodificación seguida del ajuste de ganancia y la recodificación, o se ha de realizar el ajuste del nivel de sonido de forma heurística solamente adaptando los elementos respectivos de la corriente de bits de los distintos modos que influyen la ganancia de las distintas porciones respectivas del modo de codificación de la corriente de bits. Sin embargo, es muy probable que la última posibilidad introduzca artefactos en la representación decodificada con ajuste de ganancia.
- 35 **[0006]** Por lo tanto, hasta ahora resulta complicado el ajuste global de la ganancia de una representación decodificada, de una corriente de bits codificada por la codificación multimodo y además tiende a disminuir la calidad. En cualquiera de los casos, se ha de realizar la decodificación seguida del ajuste de ganancia y la recodificación, o se ha de realizar el ajuste del nivel de sonido de forma heurística solamente adaptando los elementos respectivos de la corriente de bits de los distintos modos que influyen la ganancia de las distintas porciones respectivas del modo de codificación de la corriente de bits. Sin embargo, es muy probable que la última posibilidad introduzca artefactos en la representación decodificada con ajuste de ganancia.
- 40 **[0005]** En el códec del sistema unificado de codificación de voz y audio (USAC, según sus siglas en inglés) de reciente aparición, el modo de codificación en el dominio de la frecuencia (FD, según sus siglas en inglés) ha heredado la ganancia global de 8 bits de la AAC. Por lo tanto, si USAC se ejecuta en el único modo en FD, tal como para tasas de bits superiores, la funcionalidad del ajuste de nivel se conservaría totalmente, en comparación con AAC. Sin embargo, en cuanto se admiten transiciones de modo, esta posibilidad ya no está presente. En el modo de transformación con código de excitación (TCX, según sus siglas en inglés), por ejemplo, existe también un elemento de corriente de bits que tiene la misma funcionalidad también denominada “ganancia global”, que tiene una longitud de tan sólo 7-bits. En otras palabras, el número de bits para codificar los elementos individuales de ganancia de los modos individuales está adaptado básicamente al modo de codificación respectivo, a fin de lograr un mejor equilibrio entre el menor consumo de bits para el control de ganancia por una parte, y por otra parte evitar una degradación de la calidad debido a una cuantificación demasiado gruesa de la ajustabilidad de ganancia. Evidentemente, este equilibrio tuvo como resultado un número distinto de bits en comparación con el modo TCX y el modo FD. En el modo de predicción lineal excitada por código algebraico (ACELP, según sus siglas en inglés) del estándar USAC que surge actualmente, el nivel puede ser controlado mediante un elemento de corriente de bits “energía media”, que tiene una longitud de 2-bits. Nuevamente, es evidente que el equilibrio entre demasiados bits para la energía media y muy pocos bits para la energía media tuvo como resultado un número distinto de bits en comparación con los otros modos de codificación, a saber el modo de codificación TCX y el modo de codificación FD.
- 45 **[0006]** Por lo tanto, hasta ahora resulta complicado el ajuste global de la ganancia de una representación decodificada, de una corriente de bits codificada por la codificación multimodo y además tiende a disminuir la calidad. En cualquiera de los casos, se ha de realizar la decodificación seguida del ajuste de ganancia y la recodificación, o se ha de realizar el ajuste del nivel de sonido de forma heurística solamente adaptando los elementos respectivos de la corriente de bits de los distintos modos que influyen la ganancia de las distintas porciones respectivas del modo de codificación de la corriente de bits. Sin embargo, es muy probable que la última posibilidad introduzca artefactos en la representación decodificada con ajuste de ganancia.
- 50 **[0005]** En el códec del sistema unificado de codificación de voz y audio (USAC, según sus siglas en inglés) de reciente aparición, el modo de codificación en el dominio de la frecuencia (FD, según sus siglas en inglés) ha heredado la ganancia global de 8 bits de la AAC. Por lo tanto, si USAC se ejecuta en el único modo en FD, tal como para tasas de bits superiores, la funcionalidad del ajuste de nivel se conservaría totalmente, en comparación con AAC. Sin embargo, en cuanto se admiten transiciones de modo, esta posibilidad ya no está presente. En el modo de transformación con código de excitación (TCX, según sus siglas en inglés), por ejemplo, existe también un elemento de corriente de bits que tiene la misma funcionalidad también denominada “ganancia global”, que tiene una longitud de tan sólo 7-bits. En otras palabras, el número de bits para codificar los elementos individuales de ganancia de los modos individuales está adaptado básicamente al modo de codificación respectivo, a fin de lograr un mejor equilibrio entre el menor consumo de bits para el control de ganancia por una parte, y por otra parte evitar una degradación de la calidad debido a una cuantificación demasiado gruesa de la ajustabilidad de ganancia. Evidentemente, este equilibrio tuvo como resultado un número distinto de bits en comparación con el modo TCX y el modo FD. En el modo de predicción lineal excitada por código algebraico (ACELP, según sus siglas en inglés) del estándar USAC que surge actualmente, el nivel puede ser controlado mediante un elemento de corriente de bits “energía media”, que tiene una longitud de 2-bits. Nuevamente, es evidente que el equilibrio entre demasiados bits para la energía media y muy pocos bits para la energía media tuvo como resultado un número distinto de bits en comparación con los otros modos de codificación, a saber el modo de codificación TCX y el modo de codificación FD.
- 55 **[0006]** Por lo tanto, hasta ahora resulta complicado el ajuste global de la ganancia de una representación decodificada, de una corriente de bits codificada por la codificación multimodo y además tiende a disminuir la calidad. En cualquiera de los casos, se ha de realizar la decodificación seguida del ajuste de ganancia y la recodificación, o se ha de realizar el ajuste del nivel de sonido de forma heurística solamente adaptando los elementos respectivos de la corriente de bits de los distintos modos que influyen la ganancia de las distintas porciones respectivas del modo de codificación de la corriente de bits. Sin embargo, es muy probable que la última posibilidad introduzca artefactos en la representación decodificada con ajuste de ganancia.
- 60 **[0006]** Por lo tanto, hasta ahora resulta complicado el ajuste global de la ganancia de una representación decodificada, de una corriente de bits codificada por la codificación multimodo y además tiende a disminuir la calidad. En cualquiera de los casos, se ha de realizar la decodificación seguida del ajuste de ganancia y la recodificación, o se ha de realizar el ajuste del nivel de sonido de forma heurística solamente adaptando los elementos respectivos de la corriente de bits de los distintos modos que influyen la ganancia de las distintas porciones respectivas del modo de codificación de la corriente de bits. Sin embargo, es muy probable que la última posibilidad introduzca artefactos en la representación decodificada con ajuste de ganancia.

**[0007]** Un ejemplo de un codec multi-modo conocido para codificar contenidos de voz y audio empleando un algoritmo que cambia entre los modos ACELP y TCX se describe en BESSETTE ET AL: "A wideband speech and audio codec at 16/24/32 kbit/s using hybrid ACELP/TCX techniques", SPEECH CODING PROCEEDINGS, 1999 IEEE WORKSHOP ON PORVOO, FINLANDIA 20-23 JUNIO DE 1999, PISCATAWAY, NJ, USA, IEEE, US, 20 de Junio 1999, páginas 7-9, XP010345581.

**[0008]** Por lo tanto, un objetivo de la presente invención consiste en proporcionar un códec multimodo de audio que permita el ajuste global de la ganancia sin un desvío de decodificación y recodificación con penas moderadas en términos de calidad y tasa de compresión, y un códec de CELP adecuado para ser incorporado en la codificación multimodo de audio logrando propiedades similares.

**[0009]** Este objetivo es logrado por el contenido en cuestión de las reivindicaciones adjuntas al presente documento.

**[0010]** De acuerdo con un primer aspecto de la presente invención, los inventores de la presente solicitud comprendieron que un problema que aparece cuando se intenta armonizar el ajuste de ganancia global a través de distintos modos de codificación proviene del hecho de que los distintos modos de codificación tienen distintos tamaños de marco y se descomponen de manera diferente en submarcos. De acuerdo con el primer aspecto de la presente solicitud, esta dificultad se supera codificando los elementos de corriente de bits de submarcos en forma diferencial al valor de ganancia global de modo tal que un cambio del valor de ganancia global de los marcos tiene como resultado un ajuste de un nivel de salida de la representación decodificada del contenido de audio. Simultáneamente, la codificación diferencial ahorra bits que de otro modo ocurrirían cuando se introduce un nuevo elemento de sintaxis en una corriente de bits codificada. Aún más, la codificación diferencial permite la reducción de la carga del ajuste global de la ganancia de una corriente de bits codificada, permitiendo que la resolución de tiempo al fijar el valor de ganancia global sea menor que la resolución de tiempo a la cual el elemento de corriente de bits mencionado anteriormente y codificado en el valor de ganancia global ajusta la ganancia del submarco respectivo.

**[0011]** Por consiguiente, de acuerdo con un primer aspecto de la presente solicitud, un decodificador multimodo de audio para proporcionar una representación de decodificador de un contenido de audio sobre la base de una corriente de bits codificada está configurado para decodificar un valor de ganancia global por marco de la corriente de bits codificada, en el cual un primer subconjunto de los marcos está codificado en un primer modo de codificación y en el cual un segundo subconjunto de marcos está codificado en un segundo modo de codificación, donde cada marco del segundo subconjunto está compuesto de más de un submarco, para decodificar, por submarco de por lo menos un subconjunto de submarcos del segundo subconjunto de marcos, un elemento de corriente de bits correspondiente, diferencial al valor de ganancia global del marco respectivo, y completar la decodificación de la corriente de bits utilizando el valor de ganancia global y el elemento de corriente de bits correspondiente y decodificar los submarcos de por lo menos un subconjunto de submarcos del segundo subconjunto de los marcos y el valor de ganancia global al decodificar el primer subconjunto de marcos, en el cual el decodificador multicódigo para audio está configurado de modo tal que un cambio del valor de ganancia global de los marcos dentro de la corriente de bits codificada tiene como resultado un ajuste de un nivel de salida de la representación de decodificador del contenido de audio. Un codificador multimodo de audio está configurado, de acuerdo con dicho primer aspecto, para codificar un contenido de audio en una corriente de bits codificada con una codificación de un primer subconjunto de submarcos en un primer modo de codificación y un segundo subconjunto de marcos en el segundo modo de codificación, cuando el segundo subconjunto de marcos está compuesto de uno o más submarcos, cuando el codificador multimodo de audio está configurado para determinar y codificar un valor de ganancia global por marco, y determinar y codificar, los submarcos de por lo menos un subconjunto de submarcos del segundo subconjunto, un elemento de corriente de bits correspondiente, diferencial al valor de ganancia global del marco respectivo, en el cual el codificador multimodo de audio está configurado de modo tal que un cambio del valor de ganancia global de los marcos dentro de la corriente de bits codificada tiene como resultado un ajuste de un nivel de salida de una representación decodificada del contenido de audio en el lado decodificador.

**[0012]** De acuerdo con un segundo aspecto de la presente solicitud, los inventores de la presente solicitud descubrieron que un control de ganancia global a través de marcos con codificación de CELP y marcos con codificación transformada se puede lograr manteniendo las ventajas delineadas anteriormente, si la ganancia de la excitación del libro de código del códec CELP está conjuntamente controlada con un nivel de la transformada o transformada inversa de los marcos con codificación transformada. Obviamente, dicho uso conjunto se puede llevar a cabo mediante codificación diferencial.

**[0013]** Por consiguiente, un decodificador multimodo de audio para proporcionar una representación decodificada de un contenido de audio sobre la base de una corriente de bits codificada, un primer subconjunto de marcos el cual está codificado con CELP y un segundo subconjunto de marcos el cual está codificado con transformada, comprende, de acuerdo con el segundo aspecto, un decodificador de CELP configurado para decodificar un marco actual del primer subconjunto, en el cual el decodificador de CELP comprende un generador de excitación configurado para generar una excitación actual de un marco actual del primer subconjunto construyendo una excitación de libro de código, basada en una excitación pasada e índice de libro de código del marco actual del

primer subconjunto dentro de la corriente de bits codificada, y fijar una ganancia de la excitación del libro de código basada en el valor de ganancia global dentro de la corriente de bits codificada; y un filtro de síntesis de predicción lineal configurado para filtrar la excitación actual basada en coeficientes de filtro de predicción lineal para el marco actual del primer subconjunto dentro de la corriente de bits codificada, y un decodificador de transformada configurado para decodificar un marco actual del segundo subconjunto construyendo información espectral para el marco actual del segundo subconjunto a partir de la corriente de bits codificada y formando una transformación espectral al dominio de tiempo sobre la transformación espectral para obtener una señal de dominio de tiempo de modo tal que un nivel de la señal de dominio de tiempo dependa del valor de ganancia global.

**[0014]** De la misma manera, un codificador multimodo de audio para codificar un contenido de audio en una corriente codificada codificando por CELP un primer subconjunto de marcos del contenido de audio y codificando por transformada un segundo subconjunto de marcos comprende, de acuerdo con el segundo aspecto, un codificador de CELP configurado para codificar el marco actual del primer subconjunto, en el cual el codificador de CELP comprende un analizador de predicción lineal configurado para generar coeficientes de filtro de predicción lineal para el marco actual del primer subconjunto y codificar los mismos en la corriente de bits codificada, y un generador de excitación configurado para determinar una excitación actual del marco actual del primer subconjunto que, cuando es filtrado por un filtro de síntesis de predicción lineal basado en los coeficientes de filtro de predicción lineal dentro de la corriente de bits codificada recupera el marco actual del primer subconjunto, construyendo la excitación del libro de código basada en una excitación pasada y un índice de libro de código para el marco actual del primer subconjunto, y una transformada codificada configurada para codificar un marco actual del segundo subconjunto llevando a cabo una transformación de tiempo al dominio espectral sobre una señal del dominio de tiempo para el marco actual para el segundo subconjunto para obtener información espectral y codificar la información espectral en la corriente de bits codificada, en el cual el codificador multimodo de audio está configurado para codificar un valor de ganancia global en la corriente de bits codificada, en el cual el valor de ganancia global depende de una energía de una versión del contenido de audio del marco actual del primer subconjunto filtrado con un filtro de análisis de predicción lineal dependiendo de los coeficientes de predicción lineal, o una energía de la señal del dominio de tiempo.

**[0015]** De acuerdo con un tercer aspecto de la presente solicitud, los presentes inventores descubrieron que la variación de la intensidad de sonido de una corriente de bits codificada con CELP al cambiar el valor de ganancia global respectivo se adapta mejor al comportamiento de los ajustes de nivel codificados por transformada, si el valor de ganancia global en la codificación de CELP se computa y aplica en el dominio ponderado de la señal de excitación, en vez de la señal de excitación plana directamente. Además, el cómputo y la aplicación del valor de ganancia global en el dominio ponderado de la señal de excitación también representa una ventaja cuando se considera el modo de codificación de CELP exclusivamente ya que las otras ganancias en CELP tales como la ganancia de código y la ganancia de predicción de largo plazo (LTP, según sus siglas en inglés), también se computan en el dominio ponderado.

**[0016]** Por consiguiente, de acuerdo con el tercer aspecto, un decodificador de CELP comprende un generador de excitación configurado para generar una excitación actual para un marco actual de una corriente de bits construyendo una excitación del libro de código adaptativo basada en una excitación pasada y un índice de libro de código adaptativo para el marco actual dentro de la corriente de bits, construyendo una excitación del libro de código de innovación basada en un índice de libro de código de innovación para el marco actual dentro de la corriente de bits, computando un cálculo aproximado de una energía de la excitación del libro de código de innovación ponderada espectralmente por un filtro de síntesis ponderado de predicción lineal construido a partir de coeficientes de predicción lineal dentro de la corriente de bits, fijar una ganancia de la excitación del libro de código de innovación basada en una relación entre un valor de ganancia dentro de la corriente de bits la energía estimada (sic), y combinar la excitación del libro de código adaptativo y la excitación del libro de código de innovación para obtener la excitación actual; y un filtro de síntesis de predicción lineal configurado para filtrar la excitación actual basado en los coeficientes de filtro de predicción lineal.

**[0017]** De la misma manera, un codificador de CELP comprende, de acuerdo con el tercer aspecto, un analizador de predicción lineal configurado para generar coeficientes de filtro de predicción lineal para un marco actual de un contenido de audio y codificar un coeficiente de filtro de predicción lineal en una corriente de bits; un generador de excitación configurado para determinar una excitación actual del marco actual como una combinación de una excitación del libro de código adaptativo y una excitación del libro de código de innovación que, cuando es filtrado por un filtro de síntesis de predicción lineal basado en los coeficientes de filtro de predicción lineal, recupera el marco actual, construyendo la excitación del libro de código adaptativo definida por una excitación pasada y un índice de libro de código adaptativo para el marco actual y codificando el índice de libro de código adaptativo en la corriente de bits, y construyendo la excitación del libro de código de innovación definida por un índice de libro de código de innovación para el marco de corriente y codificando el índice de libro de código de innovación en la corriente de bits; y un determinador de energía configurado para determinar una energía de una versión de un contenido de audio del marco actual filtrado con un filtro de síntesis de predicción lineal dependiendo de los coeficientes de filtro de predicción lineal y un filtro de ponderación perceptual para obtener un valor de ganancia y

una codificación del valor de ganancia en la corriente de bits, en el cual el filtro de ponderación está construido a partir de los coeficientes de filtro de predicción lineal.

5 **[0018]** Si algún objeto en la descripción y / o los dibujos que no está cubierto por las reivindicaciones se presenta como una forma de realización de la invención, tal objeto se limita a entenderse como un ejemplo que es útil para la comprensión de la invención

Breve Descripción de los Dibujos

10 **[0019]** Las formas de realización preferidas de la presente solicitud constituyen el tema de las reivindicaciones subordinadas adjuntas al presente documento. Asimismo, las formas de realización preferidas de la presente solicitud se describen a continuación con respecto a las figuras, entre las cuales:

15 La figura 1 ilustra un diagrama en bloque de un codificador multimodo de audio de acuerdo con una forma de realización;

La figura 2 ilustra un diagrama en bloque de la porción de cómputo de energía del codificador de la figura 1 de acuerdo con una primera alternativa;

20 La figura 3 ilustra un diagrama en bloque de la porción de cómputo de energía del codificador de la figura 1 de acuerdo con una segunda alternativa;

La figura 4 ilustra un decodificador multimodo de audio de acuerdo con una forma de realización y adaptado para decodificar corrientes de bits codificadas por el codificador de la figura 1;

25 La figura 5a y la figura 5b ilustran un codificador multimodo de audio y un decodificador multimodo de audio de acuerdo con una forma de realización adicional de la presente invención;

30 La figura 6a y la figura 6b ilustran un codificador multimodo de audio y un decodificador multimodo de audio de acuerdo con una forma de realización adicional de la presente invención; y

La figura 7a y la figura 7b ilustran un codificador de CELP y un decodificador de CELP de acuerdo con una forma de realización adicional de la presente invención.

35 **[0020]** La figura 1 ilustra una forma de realización de un codificador multimodo de audio de acuerdo con una forma de realización de la presente solicitud. El codificador multimodo de audio de la figura 1 es adecuado para codificar señales de audio de tipo mixto tales como de una mezcla de voz y música, o similares. A fin de obtener un compromiso óptimo de velocidad/distorsión, el codificador multimodo de audio está configurado para conmutar entre diversos modos de codificación a fin de adaptar las propiedades de codificación a las necesidades actuales del contenido de audio a codificar. En particular, de acuerdo con la forma de realización de la figura 1, el codificador  
40 multimodo de audio generalmente utiliza tres modos distintos de codificación, a saber codificación FD (dominio de frecuencia), y codificación LP (predicción lineal), que a su vez, está dividida en codificación de TCX (transformada con código de excitación) y CELP (predicción lineal con excitación por código). En el modo de codificación FD, el contenido de audio a codificar está dividido en ventanas, descompuesto espectralmente, y la descomposición  
45 espectral está cuantificada y ajustada a escala de acuerdo con la psicoacústica a fin de ocultar el ruido de cuantificación debajo del umbral de enmascaramiento. En los modos de codificación de TCX y CELP, el contenido de audio está sujeto a análisis de predicción lineal a fin de obtener coeficientes de predicción lineal, y dichos coeficientes de predicción lineal son transmitidos dentro de la corriente de bits junto con una señal de excitación que, cuando es filtrada con un filtro de síntesis de predicción lineal correspondiente que utiliza los coeficientes de predicción lineal dentro de la corriente de bits produce la representación decodificada del contenido de audio. En el caso de la codificación de TCX, la señal de excitación es codificada por transformada, en tanto que en el caso de la codificación de CELP, la señal de excitación es codificada indexando entradas dentro de un libro de código o de otro modo construyendo sintéticamente un vector de libro de código de muestras a filtrar. En el modo de predicción lineal  
50 excitada por código algebraico (ACELP, según sus siglas en inglés), que se utiliza de acuerdo con la presente forma de realización, la excitación está compuesta de una excitación del libro de código adaptativo y una excitación del libro de código de innovación. Según se describirá con más detalle a continuación, en TCX, los coeficientes de predicción lineal pueden ser aprovechados en el lado decodificador también directamente en el dominio de frecuencia para modelar la cuantificación de ruido deduciendo factores de escala. En este caso, la TCX se configura para transformar la señal original y aplicar el resultado de la codificación de predicción lineal (LPC, según sus siglas en inglés) sólo en el dominio de frecuencia.  
60

**[0021]** A pesar de los distintos modos de codificación, el codificador de la figura 1 genera la corriente de bits de modo tal que un determinado elemento de sintaxis asociado a todos los marcos de la corriente de bits codificada – con ejemplificaciones asociadas a los marcos individualmente o en grupos de marcos–, permite una adaptación de

ganancia global a través de todos los modos de codificación aumentando o reduciendo, por ejemplo, dichos valores globales en la misma cantidad tal como, por ejemplo, el mismo número de dígitos (que iguala un ajuste de escala con un factor (o divisor) de los tiempos de base logarítmica el número de dígitos (sic)).

5 **[0022]** En particular, de acuerdo con los diversos modos de codificación admitidos por el codificador multimodo de audio 10 de la figura 1, el mismo comprende un codificador FD 12 y un codificador LPC (codificación de predicción lineal) 14. El codificador LPC 14, a su vez, está compuesto de una porción de codificación TCX 16, una porción de codificación de CELP 18, y un conmutador de modo de codificación 20. Más precisamente, un conmutador de modo de codificación adicional incluido en el codificador 10 está ilustrado generalmente en 22 como designador de modo.

10 El designador de modo está configurado para analizar el contenido de audio 24 a codificar a fin de asociar porciones de tiempo consecutivas del mismo a distintos modos de codificación. En particular, en el caso de la figura 1, el designador de modo 22 asigna distintas porciones de tiempo consecutivas del contenido de audio 24 a cualquiera del modo de codificación FD y el modo de codificación LPC. En el ejemplo ilustrativo de la figura 1, por ejemplo, el designador de modo 22 ha asignado la porción 26 del contenido de audio 24 al modo de codificación FD, en tanto

15 que la porción inmediatamente siguiente 28 es asignada al modo de codificación LPC. Dependiendo del modo de codificación asignado por el designador de modo 22, el contenido de audio 24 puede estar subdividido en marcos consecutivos distintos. Por ejemplo, en la forma de realización de la figura 1, el contenido de audio 24 dentro de la porción 26 está codificado en marcos 30 de igual longitud y con un solapamiento entre sí de, por ejemplo, 50%. En otras palabras, el codificador FD 12 está configurado para codificar la porción FD 26 del contenido de audio 24 en dichas unidades 30. De acuerdo con la forma de realización de la figura 1, el codificador LPC 14 también está configurado para codificar su porción asociada 28 del contenido de audio 24 en unidades de marcos 32 con dichos marcos, sin embargo, no necesariamente tienen el mismo tamaño que los marcos 30. En el caso de la figura 1, por ejemplo, el tamaño de los marcos 32 es menor que el tamaño de los marcos 30. En particular, de acuerdo con una forma de realización específica, la longitud de los marcos 30 es de 2048 muestras del contenido de audio 24, en tanto que la longitud de los marcos 32 es de 1024 muestras cada uno. Cabe la posibilidad de que el último marco se solape con el primer marco en un borde entre el modo de codificación LPC y el modo de codificación FD. Sin embargo, en la forma de realización de la figura 1, y según se ilustra en forma ejemplar en la figura 1, también cabe la posibilidad de que no exista solapamiento de marcos en el caso de transiciones desde el modo de codificación FD al modo de codificación de LPC, y viceversa.

30 **[0023]** Como se indica en la figura 1, el codificador FD 12 recibe los marcos 30 y los codifica por codificación de transformada en el dominio de frecuencia en los respectivos marcos 34 de la corriente de bits codificada 36. A tal fin, el codificador FD 12 comprende un divisor en ventanas 38, un transformador 40, un módulo de cuantificación y ajuste de escala 42, y un codificador sin pérdidas 44, así como también un controlador psicoacústico 46. En principio, el codificador FD 12 puede estar implementado de acuerdo con la norma AAC siempre que la siguiente descripción no divulgue un comportamiento diferente del codificador FD 12. En particular, el divisor en ventanas 38, el transformador 40, el módulo de cuantificación y ajuste de escala 42 y el codificador sin pérdidas 44, están conectados en serie entre una entrada 48 y una salida 50 del codificador FD 12 y el controlador psicoacústico 46 tiene una entrada conectada a la entrada 48 y una salida conectada a otra entrada del módulo de cuantificación y ajuste de escala 42. Cabe señalar que el codificador FD 12 puede comprender otros módulos para otras opciones de codificación que, sin embargo, no son críticas en el presente documento.

45 **[0024]** El divisor en ventanas 38 puede utilizar distintas ventanas para dividir en ventanas un marco actual que ingresa a la entrada 48. El marco dividido en ventanas está sujeto a una transformación de tiempo al dominio espectral en el transformador 40, tal como el uso de una transformada coseno discreta modificada (MDCT, según sus siglas en inglés) o similar. El transformador 40 puede utilizar distintas longitudes de transformada a fin de transformar los marcos divididos en ventanas.

50 **[0025]** En particular, el divisor en ventanas 38 puede soportar ventanas cuya longitud coincida con la longitud de los marcos 30 utilizando el transformador 40 la misma longitud de transformada a fin de producir un número de coeficientes de transformada que puedan corresponder, por ejemplo, en el caso de la MDCT, a la mitad del número de muestras del marco 30. El divisor en ventanas 38 sin embargo, también puede estar configurado para soportar opciones de codificación de acuerdo con las cuales diversas ventanas más cortas, tales como las ocho ventanas cuya longitud es la mitad de los marcos 30 que están descentradas unas con otras en tiempo, son aplicadas a un marco actual transformando el transformador 40 dichas versiones divididas en ventanas del marco actual utilizando una longitud de transformada que cumpla con los requisitos de la división en ventanas produciendo, de ese modo, ocho espectros para que dicho marco realice un muestreo del contenido de audio en diferentes momentos durante dicho marco. Las ventanas utilizadas por el divisor en ventanas 38 pueden ser simétricas o asimétricas y pueden tener un extremo delantero cero y/o un extremo trasero cero. En el caso de aplicar diversas ventanas cortas a un marco actual, la porción no cero de dichas ventanas cortas está desplazada entre sí, sin embargo se solapa unas con otras. Obviamente se pueden utilizar otras opciones de codificación para las ventanas y las longitudes de transformada para el divisor en ventanas 38 y el transformador 40 de acuerdo con una forma de realización alternativa.

[0026] La salida de los coeficientes de transformada por el transformador 40 es cuantificada y ajustada a escala en el módulo 42. En particular, el controlador psicoacústico 46 analiza la señal de entrada en la entrada 48 a fin de determinar un umbral de enmascaramiento 48 de acuerdo con el cual se forma el ruido de cuantificación introducido por la cuantificación y el ajuste de escala para que esté por debajo del umbral de enmascaramiento. En particular, módulo de ajuste de escala 42 puede operar en bandas de factores de escala que juntas abarcan el dominio espectral del transformador 40 en el cual el dominio espectral está subdividido. Por consiguiente, los grupos de coeficientes de transformada consecutivos son asignados a distintas bandas de factores de escala. El módulo 42 determina un factor de escala por banda de factor de escala, el cual cuando es multiplicado por los valores de coeficientes de transformada respectivos, asignados a las bandas de factores de escala respectivas, produce la versión reconstruida de la salida de los coeficientes de transformada por el transformador 40. Adicionalmente, el módulo 42 fija un valor de ganancia que ajusta a escala en forma uniforme y espectral el espectro. Por lo tanto, un coeficiente de transformada reconstruido es igual a los tiempos de valores de coeficientes de transformada los tiempos de factores de escala asociados (sic) el valor de ganancia  $g_i$  del respectivo marco  $i$ . Los valores de coeficientes de transformada, los factores de escala y el valor de ganancia están sujetos a codificación sin pérdidas en el codificador sin pérdidas 44, tal como a modo de codificación por entropía tal como la codificación de Huffman o aritmética, junto con otros elementos de sintaxis relacionados con, por ejemplo, las decisiones de ventana y longitud de transformada mencionadas anteriormente y otros elementos de sintaxis que permiten otras opciones de codificación. Para obtener más detalles al respecto, se hace referencia a la norma AAC en cuanto a otras opciones de codificación.

[0027] A fin de ser un poco más precisos, el módulo de cuantificación y ajuste de escala 42 puede estar configurado para transmitir un valor de coeficiente de transformada cuantificado por línea espectral  $k$ , que produce, cuando está reajustado a escala, el coeficiente de transformada reconstruido en la línea espectral respectiva  $k$ , a saber  $x_{rescal}$ , cuando se multiplica con:

$$ganancia = 2^{0,25 \cdot (sf - sf\_offset)}$$

en el cual  $sf$  es el factor de escala de la banda de factor de escala respectivo a la cual pertenece el coeficiente de transformada cuantificado respectivo, y  $sf\_offset$  es una constante que puede ser fijada, por ejemplo, a 100.

[0028] Por lo tanto, los factores de escala están definidos en el dominio logarítmico. Los factores de escala pueden estar codificados dentro de la corriente de bits 36 en forma diferencial entre sí a lo largo del acceso espectral, es decir, solamente la diferencia entre factores de escala espectralmente vecinos  $sf$  puede ser transmitida dentro de la corriente de bits. El primer factor de escala  $sf$  puede ser transmitido dentro de la corriente de bits codificado diferencialmente con respecto al valor de ganancia global ( $global\_gain$ ) mencionado anteriormente. Este elemento de sintaxis de ganancia global ( $global\_gain$ ) resultará importante en la siguiente descripción.

[0029] El valor de ganancia global ( $global\_gain$ ) puede ser transmitido dentro de la corriente de bits en el dominio logarítmico. Es decir, el módulo 42 podría estar configurado para adoptar un primer factor de escala  $sf$  de un espectro de corriente, como  $ganancia\_global$ . Luego, dicho valor  $sf$  puede ser transmitido diferencialmente con un cero y los siguientes valores  $sf$  en forma diferencial al predecesor respectivo.

[0030] Evidentemente, el cambio de ganancia global ( $global\_gain$ ) cambia la energía de la transformada reconstruida, y por lo tanto se traduce en un cambio de intensidad del sonido de la porción codificada por FD 26, cuando se lleva a cabo uniformemente en todos los marcos 30.

[0031] En particular, la ganancia global ( $global\_gain$ ) de los marcos FD es transmitida dentro de la corriente de bits de modo tal que la  $ganancia\_global$  ( $global\_gain$ ) depende logarítmicamente del medio de ejecución de muestras de tiempo de audio reconstruidas o, vice versa, el medio de ejecución de las muestras de tiempo de audio reconstruidas depende exponencialmente de la ganancia global ( $global\_gain$ ).

[0032] De manera similar a los marcos 30, todos los marcos asignados al modo de codificación LPC, a saber los marcos 32, ingresan al codificador LPC 14. Dentro del codificador LPC 14, el conmutador 20 subdivide cada marco 32 en uno o más submarcos 52. Cada uno de dichos submarcos 52 puede ser asignado al modo de codificación de TCX o al modo de codificación de CELP. Los submarcos 52 asignados al modo de codificación de TCX son enviados a una entrada 54 del codificador TCX 16, en tanto que los submarcos asociados al modo de codificación de CELP son enviados por el conmutador 20 a una entrada 56 del codificador de CELP 18.

[0033] Cabe señalar que el arreglo del conmutador 20 entre la entrada 58 del codificador LPC 14 y las entradas 54 y 56 del codificador TCX 16 y del codificador de CELP 18, respectivamente, está ilustrado en la figura 1 sólo con fines ilustrativos y que, en realidad, la decisión de codificación con respecto a la subdivisión de los marcos 32 en los submarcos 52 con modos de codificación respectivos asociados entre TCX y CELP en los submarcos individuales se

puede llevar a cabo en forma interactiva entre los elementos internos del codificador TCX 16 y del codificador de CELP 18 a fin de maximizar una determinada medida de peso/distorsión.

5 **[0034]** En cualquiera de los casos, el codificador TCX 16 comprende un generador de excitación 60, un analizador de LP 62 y un determinador de energía 64, en el cual el analizador de LP 62 y el determinador de energía 64 son utilizados conjuntamente por el codificador de CELP 18 (y pertenecen conjuntamente al mismo) que además comprende un generador propio de excitación 66. Las entradas respectivas del generador de excitación 60, analizador de LP 62 y determinador de energía 64 están conectadas a la entrada 54 del codificador TCX 16. De la misma manera, las entradas respectivas del analizador de LP 62, determinador de energía 64 y generador de excitación 66 están conectadas a la entrada 56 del codificador de CELP 18. El analizador de LP 62 está configurado para analizar el contenido de audio dentro del marco actual, es decir, el marco TCX o el marco de CELP, a fin de determinar los coeficientes de predicción lineal, y está conectado a las entradas de coeficientes respectivas del generador de excitación 60, determinador de energía 64 y generador de excitación 66 a fin de enviar los coeficientes de predicción lineal a dichos elementos. Según se describirá en mayor detalle a continuación, el analizador de LP puede operar sobre una versión pre-enfatizada del contenido de audio original, y el filtro de pre-énfasis respectivo puede formar parte de una porción de entrada respectiva del analizador de LP, o puede estar conectado en frente de la entrada del mismo. Esto mismo se aplica al determinador de energía 66 según se describirá en mayor detalle a continuación. En lo que respecta al generador de excitación 60, sin embargo, el mismo puede operar sobre la señal original directamente. Las salidas respectivas del generador de excitación 60, analizador de LP 62, determinador de energía 64, y generador de excitación 66, así como también la salida 50, están conectadas a las entradas respectivas de un multiplexor 68 del codificador 10 que está configurado para multiplexar los elementos de sintaxis recibidos en la corriente de bits 36 en la salida 70.

25 **[0035]** Tal como se indicó anteriormente, el analizador de LPC 62 está configurado para determinar coeficientes de predicción lineal para los marcos LPC entrantes 32. Para proporcionar mayor cantidad de detalles con respecto a una probable funcionalidad del analizador de LP 62, se hace referencia al estándar ACELP. En general, el analizador de LP 62 puede utilizar un procedimiento de auto-correlación o varianza conjunta a fin de determinar los coeficientes LPC. Por ejemplo, cuando se utiliza el procedimiento de auto-correlación, el analizador de LP 62 puede producir una matriz de auto-correlación resolviendo los coeficientes LPC utilizando un algoritmo de Levinson-Durban. Como es sabido en la técnica, los coeficientes LPC definen un filtro de síntesis que modela aproximadamente el tracto vocal humano, y cuando es accionado por una señal de excitación, básicamente modela el flujo de aire a través de las cuerdas vocales. Dicho filtro de síntesis se modela utilizando predicción lineal mediante el analizador de LP 62. La tasa a la cual cambia la forma del tracto vocal es limitada y, por consiguiente, el analizador de LP 62 puede utilizar una tasa actualizada adaptada a la limitación y diferente de la tasa del marco de los marcos 32 para actualizar los coeficientes de predicción lineal. El análisis de LP llevado a cabo por el analizador 35 62 proporciona información sobre determinados filtros para los elementos 60, 64 y 66, tal como:

- el filtro de síntesis de predicción lineal  $H(z)$ ;
- 40 • el filtro inverso del mismo, a saber el filtro de análisis de predicción lineal o filtro blanqueador

45 •  $A(z)$  con 
$$H(z) = \frac{1}{A(z)}$$

- un filtro de ponderación perceptual tal como  $W(z)=A(z/\lambda)$ , en el cual  $\lambda$  es un factor de ponderación

50 **[0036]** El analizador de LP 62 transmite información sobre los coeficientes LPC al multiplexor 68 para ser insertada en la corriente de bits 36. Dicha información 72 puede representar los coeficientes de predicción lineal cuantificados en un dominio adecuado tal como un dominio de pares espectrales, o similar. Incluso la cuantificación de los coeficientes de predicción lineal se puede llevar a cabo en este dominio. Asimismo, el analizador de LPC 62 puede transmitir los coeficientes LPC o la información 72 sobre los mismos, a una tasa mayor que la tasa a la cual los coeficientes LPC son realmente reconstruidos en el lado decodificador. La última tasa de actualización se logra, por ejemplo, mediante interpolación entre los tiempos de transmisión de LPC. Evidentemente, el decodificador solo tiene acceso a los coeficientes LPC cuantificados, y por consiguiente, los filtros anteriormente mencionados y definidos por las correspondientes predicciones lineales reconstruidas se indican como  $\hat{H}(z)$ ,  $\hat{A}(z)$  y  $\hat{W}(z)$ .

60 **[0037]** Según ya se indicó anteriormente, el analizador de LP 62 define un filtro de síntesis LP  $H(z)$  y  $\hat{H}(z)$ , respectivamente, el cual, cuando se aplica a una excitación respectiva, recupera o reconstruye el contenido de audio original además de un cierto procesamiento posterior, el cual sin embargo, no es considerado en la presente para facilitar la explicación.

- 5 [0038] Los generadores de excitaciones 60 y 66 tienen el objeto de definir esta excitación y transmitir la información respectiva sobre los mismos al lado decodificador mediante los multiplexores 68 y la corriente de bits 36, respectivamente. En lo que respecta al generador de excitación 60 del codificador TCX 16, el mismo codifica la excitación actual sometiendo una excitación encontrada adecuada, por ejemplo, por algún esquema de optimización a una transformación de tiempo al dominio espectral a fin de producir una versión espectral de la excitación, en el cual dicha versión espectral de la información espectral 74 es enviada al multiplexor 68 para su inserción en la corriente de bits 36, con la información espectral cuantificada y ajustada a escala, por ejemplo, en forma análoga al espectro en el cual opera el módulo 42 del codificador FD 12.
- 10 [0039] Es decir, la información espectral 74 que define la excitación del codificador TCX 16 del submarco actual 52, puede tener coeficientes de transformada cuantificados asociados a la misma, que son ajustados a escala de acuerdo con un único factor de escala que, a su vez, es transmitido con respecto a un elemento de sintaxis de marco de LPC también denominado ganancia global (`global_gain`) en el siguiente contexto. Al igual que en el caso de la ganancia global (`global_gain`) del codificador FD 12, la ganancia global (`global_gain`) del codificador LPC 14 también puede estar definida en el dominio logarítmico. Un aumento de este valor se traduce directamente en un aumento de intensidad de sonido de la representación decodificada del contenido de audio de los submarcos de TCX respectivos debido a que la representación decodificada se logra procesando los coeficientes de transformada ajustados a escala dentro de la información 74 mediante operaciones lineales que mantienen el ajuste de ganancia. Dichas operaciones lineales son la transformada inversa de tiempo a frecuencia y, eventualmente, el filtrado de síntesis LP. Sin embargo, tal como se explicará en mayor detalle a continuación, el generador de excitación 60 está configurado para codificar la ganancia recién mencionada de la información espectral 74 en la corriente de bits en una resolución de tiempo superior que en unidades de marcos LPC. En particular, el generador de excitación 60 utiliza un elemento de sintaxis denominado ganancia global delta (`global_gain`) a fin de codificar diferencialmente – diferencialmente en la ganancia global (`global_gain`) del elemento de corriente de bits – la ganancia real utilizada para fijar la ganancia del espectro de la excitación. La ganancia global delta (`delta_global_gain`) también puede estar definida en el dominio logarítmico. La codificación diferencial puede ser llevada a cabo de modo tal que la ganancia global delta (`delta_global_gain`) puede estar definida como la corrección multiplicativa de la ganancia global (`global_gain`) – ganancia en el dominio lineal.
- 20 [0040] A diferencia del generador de excitación 60, el generador de excitación 66 del codificador de CELP 18 está configurado para codificar la excitación actual del submarco actual utilizando índices de libro de código. En particular, el generador de excitación 66 está configurado para determinar la excitación actual por una combinación de una excitación del libro de código adaptativo y una excitación del libro de código de innovación. El generador de excitación 66 está configurado para construir la excitación del libro de código adaptativo para un marco actual a fin de estar definida por una excitación pasada, es decir, la excitación utilizada para un submarco de CELP codificado previamente, por ejemplo, y un índice de libro de código adaptativo para el marco actual. El generador de excitación 66 codifica el índice de libro de código adaptativo 76 en la corriente de bits enviando el mismo al multiplexor 68. Asimismo, el generador de excitación 66 construye la excitación del libro de código de innovación definida por un índice de libro de código de innovación para el marco actual y codifica el índice de libro de código de invocación 78 en la corriente de bits enviando el mismo al multiplexor 68 para su inserción en la corriente de bits 36. En realidad, ambos índices pueden estar integrados en un elemento de sintaxis común. En conjunto, los mismos permiten que el decodificador recupere la excitación del libro de código así determinada por el generador de excitación. A fin de garantizar la sincronización de los estados internos del codificador y decodificador, el generador 66 no sólo determina los elementos de sintaxis para permitir que el decodificador recupere la excitación de libro de código de corriente, sino que el mismo también actualiza realmente su estado generando realmente el mismo a fin de utilizar la excitación de libro de código de corriente como un punto de inicio, es decir, la excitación pasada, para codificar el marco de CELP siguiente.
- 30 [0041] El generador de excitación 66 puede estar configurado para, al construir la excitación del libro de código adaptativo y la excitación del libro de código de innovación, minimizar una medida de distorsión de peso perceptual, con respecto al contenido de audio del submarco actual considerando que la excitación resultante está sujeta al filtrado de síntesis LP en el lado decodificador para la reconstrucción. En efecto, los índices 76 y 78 indexan determinadas tablas disponibles en el codificador 10 así como también el lado decodificador a fin de indexar o de otro modo determinar los vectores que sirven como entrada de excitación del filtro de síntesis LP. A diferencia de la excitación del libro de código adaptativo, la excitación del libro de código de innovación es determinada independientemente de la excitación pasada. En efecto, el generador de excitación 66 puede estar configurado para determinar la excitación del libro de código adaptativo para el marco actual utilizando la excitación pasada y reconstruida del submarco de CELP codificado previamente modificando el último utilizando una determinada demora y valor de ganancia y un filtrado predeterminado (interpolación), de modo tal que la excitación resultante del libro de código adaptativo del marco actual minimice una diferencia hasta un determinado blanco para la recuperación de excitación del libro de código adaptativo, cuando es filtrado por el filtro de síntesis, del contenido de audio original. La demora y ganancia y filtrado mencionados anteriormente están indicados por el índice de libro de código adaptativo. La discrepancia restante es compensada por la excitación del libro de código de innovación. Nuevamente, el generador de excitación 66 fija adecuadamente el índice de libro de código para encontrar una

excitación de libro de código de innovación óptima la que, cuando se combina con (se agrega a), la excitación del libro de código adaptativo produce la excitación actual para el marco actual (luego sirviendo como excitación pasada cuando se construye la excitación del libro de código adaptativo del submarco de CELP siguiente). Incluso en otras palabras, la búsqueda de libro de código adaptativo puede ser llevada a cabo sobre una base de submarco y consiste en llevar a cabo una búsqueda de altura de tono en circuito cerrado, luego computando el vector de código adaptativo interpolando la excitación pasada en el retardo de altura de tono fraccional seleccionado. En efecto, la señal de excitación  $u(n)$  está definida por el generador de excitación 66 como una suma ponderada del vector de libro de código adaptativo  $v(n)$  y el vector de libro de código de innovación  $c(n)$  por

$$u(n) = \hat{g}_p v(n) + \hat{g}_c c(n).$$

**[0042]** La ganancia de altura de tono  $\hat{g}_p$  está definida por el índice de libro de código adaptativo 76. La ganancia de libro de código de innovación  $\hat{g}_c$  está determinada por el índice de libro de código de innovación 78 y por el elemento de sintaxis de ganancia global (global\_gain) mencionado anteriormente para los marcos LPC determinados por el determinador de energía 64 según se indicará a continuación.

**[0043]** Es decir, cuando se optimiza el índice de libro de código de innovación 78, el generador de excitación 66 adopta, y mantiene sin cambio, la ganancia de libro de código de innovación  $\hat{g}_c$  optimizando solamente el índice de libro de código de innovación para determinar posiciones y signos de pulsos del vector de libro de código de innovación, así como también el número de dichos pulsos.

**[0044]** Un primer procedimiento (o alternativa) para fijar el elemento de sintaxis de ganancia global (global\_gain) del marco de LPC mencionado anteriormente mediante el determinador de energía 64 se describe a continuación con respecto a la figura 2. De acuerdo con ambas alternativas que se describen a continuación, el elemento de sintaxis de ganancia global (global\_gain) está determinado para cada uno de los marcos LPC 32. Este elemento de sintaxis luego sirve como una referencia para los elementos de sintaxis de ganancia global delta (delta\_global\_gain) mencionados anteriormente de los submarcos TCX que pertenecen al marco 32 respectivo, así como también la ganancia de libro de código de innovación mencionada anteriormente  $\hat{g}_c$  que está determinada por la ganancia global (global\_gain) que se describe a continuación.

**[0045]** Según se ilustra en la figura 2, el determinador de energía 64 puede estar configurado para determinar el elemento de sintaxis de ganancia global (global\_gain) 80, y puede comprender un filtro de análisis de predicción lineal 82 controlado por el analizador de LP 62, un computador de energía 84 y una etapa de cuantificación y codificación 86, así como también una etapa de decodificación 88 para la recuantificación. Según se ilustra en la figura 2, un pre-énfatisador o filtro de pre-énfasis 90 puede pre-énfatisar el contenido de audio original 24 antes de que el último sea procesado posteriormente dentro del determinador de energía 64 según se describe a continuación. A pesar de que no está ilustrado en la figura 1, el filtro de pre-énfasis también puede estar presente en el diagrama en bloque de la figura 1 directamente en frente de ambos, las entradas del analizador de LP 62 y el determinador de energía 64. En otras palabras, los mismos pueden pertenecer conjuntamente o ser utilizados conjuntamente por ambos. El filtro de pre-énfasis 90 puede estar determinado por

$$H_{emph}(z) = 1 - \alpha z^{-1}.$$

**[0046]** Por lo tanto, el filtro de pre-énfasis puede ser un filtro de paso alto. En el presente contexto, es un filtro de paso alto de primer orden, pero más generalmente, el mismo puede ser un filtro de paso alto de enésimo orden  $n$ . En el presente caso, es ejemplarmente un filtro de paso alto de primer orden, que tiene  $\alpha$  fijado a 0,68.

**[0047]** La entrada del determinador de energía 64 de la figura 2 está conectada a la salida del filtro de pre-énfasis 90. Entre la entrada y la salida 80 del determinador de energía 64, el filtro de análisis de LP 82, el computador de energía 84, y la etapa de cuantificación y codificación 86 están conectados en serie en el orden mencionado. La etapa de codificación 88 tiene su entrada conectada a la salida de la etapa de cuantificación y codificación 86 y produce la ganancia cuantificada tal como se obtiene del decodificador.

**[0048]** En particular, el filtro de análisis de predicción lineal 82  $A(z)$  aplicado al contenido de audio pre-énfatisado tiene como resultado una señal de excitación 92. Por lo tanto, la excitación 92 iguala la versión pre-énfatisada del contenido de audio original 24 filtrado por el filtro de análisis de LPC  $A(z)$ , es decir, el contenido de audio original 24 filtrado con

$$H_{emph}(z)A(z).$$

**[0049]** Sobre la base de esta señal de excitación 92, la ganancia global común para el marco actual 32 es deducida computando la energía sobre cada una de las 1024 muestras de dicha señal de excitación 92 dentro del marco actual 32.

5 **[0050]** En particular, el computador de energía 84 promedia la energía de la señal 92 por segmento de 64 muestras en el dominio logarítmico por:

$$10 \quad nrg = \sum_{l=0}^{15} \frac{1}{16} \cdot \log_2 \sum_{n=0}^{64} \sqrt{\frac{exc[l \cdot 64 + n] * exc[l \cdot 64 + n]}{64}}$$

**[0051]** La ganancia  $g_{index}$  ( $g_{indice}$ ) luego es cuantificada por la etapa de cuantificación y codificación 86 en 6 bits en el dominio logarítmico basado en el promedio de energía nrg por:

$$15 \quad g_{index} = \lfloor 4 \cdot nrg + 0.5 \rfloor$$

**[0052]** Este índice luego se transmite dentro de la corriente de bits como elemento de sintaxis 80, es decir, como ganancia global. Está definido en el dominio logarítmico. En otras palabras, el tamaño del paso de cuantificación aumenta exponencialmente. La ganancia cuantificada es obtenida por la etapa de decodificación 88 computando:

$$20 \quad \hat{g} = 2^{\frac{g_{index}}{4}}$$

**[0053]** La cuantificación utilizada en el presente documento tiene la misma granularidad que la cuantificación de la ganancia global del modo FD, y por consiguiente, el ajuste de escala de  $g_{index}$  ( $g_{indice}$ ) ajusta a escala la intensidad de sonido de los marcos LPC 32 de la misma manera que el ajuste de escala del elemento de sintaxis de ganancia global (global\_gain) de los marcos FD 30 logrando, de este modo, una forma fácil de controlar la ganancia de la corriente de bits multimodo codificada 36 sin necesidad de llevar a cabo un desvío de decodificación y re-codificación y manteniendo aún la calidad.

25 **[0054]** Según se describirá con más detalle a continuación con respecto al decodificador, para mantener la sincronía anteriormente mencionada entre el codificador y el decodificador (excitación *nupdate*), el generador de excitación 66 puede, al optimizar o luego de haber optimizado los índices de libro de código,

a) computar, sobre la base de la ganancia global (global\_gain), una ganancia de predicción  $g'_c$  y

35 b) multiplicar la ganancia de predicción  $g'_c$  con el factor de corrección de libro de código de innovación  $\hat{\gamma}$  para proporcionar la ganancia de libro de código de innovación real  $\hat{g}_c$

40 c) generar realmente la excitación del libro de código combinando la excitación del libro de código adaptativo y la excitación del libro de código de innovación ponderando el último con la ganancia de libro de código de innovación real  $\hat{g}_c$ .

**[0055]** En particular, de acuerdo con la presente alternativa, la etapa de cuantificación y codificación 86 transmite el  $g_{indice}$  dentro de la corriente de bits y el generador de excitación 66 acepta la ganancia cuantificada  $\hat{g}$  como una referencia fija predefinida para optimizar la excitación del libro de código de innovación.

**[0056]** En particular, el generador de excitación 66 optimiza la ganancia del libro de código de innovación  $\hat{g}_c$  utilizando (es decir, optimizando) sólo el índice de libro de código de innovación que además define  $\hat{\gamma}$  cuál es el factor de corrección de ganancia del libro de código de innovación. En particular, el factor de corrección de ganancia de libro de código de innovación determina la ganancia del libro de código de innovación  $\hat{g}_c$  que es la siguiente

$$\bar{E} = 20 \cdot \log(\hat{g})$$

$$G'_c = \bar{E}$$

$$g'_c = 10^{0.05G'_c}$$

5

$$\hat{g}_c = \hat{\gamma}_c \cdot g'_c$$

10 **[0057]** Tal como se describirá adicionalmente a continuación, la ganancia TCX se codifica transmitiendo el elemento de **ganancia global delta (delta\_global\_gain)** codificado en 5 bits:

$$15 \quad \overset{(ganancia\_global\_delta)}{\text{delta\_global\_gain}} = \left\lfloor (4 \cdot \log_2 \left( \frac{\overset{(ganancia\_tcx)}{\text{gain\_tcx}}}{\hat{g}} \right) + 10) + 0.5 \right\rfloor$$

20 **[0058]** Se codifica de la siguiente manera:

$$25 \quad \overset{(ganancia\_tcx)}{\text{gain\_tcx}} = 2^{\frac{\overset{(ganancia\_global\_delta-10)}{\text{delta\_global\_gain}-10}}{4}} \cdot \hat{g}$$

30 **[0059]** Luego

$$35 \quad g = \frac{\overset{(ganancia\_tcx)}{\text{gain\_tcx}}}{2.rms}$$

40 **[0060]** A fin de completar la concordancia entre el control de ganancia ofrecido por el elemento de sintaxis  $g_{\text{indice}}$  en lo que respecta a los submarcos de CELP y a los submarcos de TCX, de acuerdo con la primera alternativa descrita con respecto a la figura 2, la ganancia global  $g_{\text{indice}}$  es por lo tanto codificada en 6 bits por marco o supermarco 32. Esto tiene como resultado la misma granularidad de ganancia que la codificación de ganancia global del modo FD. En este caso, la ganancia global del supermarco  $g_{\text{indice}}$  es codificada sólo en 6 bits, si bien la ganancia global en el modo FD es enviada en 8 bits. Por lo tanto, el elemento de ganancia global no es el mismo para los modos LPD (dominio de predicción lineal) y FD. Sin embargo, debido a que la granularidad de ganancia es similar se puede aplicar fácilmente un control unificado de ganancia. En particular, el dominio logarítmico para codificar la ganancia global (global\_gain) en el modo FD y LPD se lleva a cabo convenientemente en la misma base logarítmica 2.

50 **[0061]** A fin de armonizar completamente ambos elementos globales, sería sencillo extender la codificación en 8 bits incluso en lo que respecta a los marcos LPD. En cuanto a los submarcos de CELP, el elemento de sintaxis  $g_{\text{indice}}$  asume completamente la tarea del control de ganancia. Los elementos de ganancia global delta (delta\_global\_gain) mencionados anteriormente de los submarcos de TCX pueden estar codificados en 5 bits en forma diferencial de la ganancia global de los supermarcos. En comparación con el caso en el que el esquema de codificación multimodo anterior se implementaría por AAC, ACELP y TCX normales, el concepto anterior de acuerdo con la alternativa de la figura 2, tendría como resultado 2 bits menos para la codificación en el caso de un supermarco 32 consistiendo solamente de TCX 20 y/o submarcos ACELP, y consumiría 2 o 4 bits adicionales por supermarco en el caso del supermarco respectivo que comprende el submarco TCX 40 y TCX 80, respectivamente.

60 **[0062]** En cuanto al procesamiento de señales, el  $g_{\text{indice}}$  global de ganancia de supermarco representa la energía residual LPC promediada sobre el supermarco 32 y cuantificada en una escala logarítmica. En (A)CELP, se utiliza en lugar del elemento de "energía media" generalmente utilizado en ACELP para estimar la ganancia de libro del código

de innovación. El nuevo cálculo aproximado de la primera alternativa presente de acuerdo con la figura 2 tiene mayor resolución de amplitud que en el estándar ACELP, pero también tiene menor resolución de tiempo ya que el  $g_{\text{indice}}$  se transmite solamente por supermarco, en vez de transmitirse por submarco. Sin embargo, se ha descubierto que la energía residual no ofrece una estimación adecuada y se utiliza como un indicador de causa del rango de ganancia. En consecuencia, la resolución de tiempo es probablemente más importante. Para evitar cualquier problema que pueda surgir durante los transitorios, el generador de excitación 66 puede estar configurado para subestimar sistemáticamente la ganancia de libro de código de innovación y permitir que el ajuste de ganancia recupere el intervalo. Esta estrategia puede compensar la falta de resolución de tiempo.

**[0063]** Asimismo, la ganancia global de supermarco también se utiliza en TCX como un cálculo aproximado del elemento de "ganancia global" que determina la ganancia de ajuste de escala (`scaling_gain`) mencionada anteriormente. Debido a que el  $g_{\text{indice}}$  de ganancia global del supermarco representa la energía del LPC residual y que TCX global representa en forma aproximada la energía de la señal ponderada, la codificación de ganancia diferencial mediante el uso de ganancia global delta (`delta_global_gain`) incluye implícitamente algunas ganancias LP. No obstante, la ganancia diferencial aún ilustra una amplitud mucho menor que la "ganancia global" plana.

**[0064]** Para mono de 12 kbps y 24 kbps se llevaron a cabo algunas pruebas de audición concentrándose principalmente en la calidad de la voz limpia. Se comprobó que la calidad se aproximaba bastante a la calidad de la USAC corriente diferenciándose de la forma de realización anterior porque se había utilizado el control de ganancia normal de los estándares AAC y ACELP/TCX. Sin embargo, para determinados elementos de voz, la calidad tiende a ser ligeramente peor.

**[0065]** Habiendo descrito la forma de realización de la figura 1 de acuerdo con la alternativa de la figura 2, la segunda alternativa se describirá con respecto a las Figs. 1 y 3. De acuerdo con el segundo procedimiento para el modo LPD, mediante dicho procedimiento se resuelven algunas desventajas de la primera alternativa:

- La predicción de la ganancia de innovación ACELP falló para algunos submarcos de los marcos dinámicos de amplitud alta. Esto se debió principalmente a la computación de energía que se promedió geométricamente. Aunque la SNR promedio fue mejor que la ACELP original, el libro de código de ajuste de ganancia se saturó con mayor frecuencia. Se supone que es la razón principal por la cual se observa una ligera degradación para determinados elementos de voz.

- Asimismo, la predicción de ganancia de la innovación ACELP no fue óptima. En efecto, la ganancia se optimiza en el dominio ponderado en tanto que la predicción de ganancia se computa en el dominio LPC residual. La idea de la siguiente alternativa consiste en llevar a cabo la predicción en el dominio ponderado.

- La predicción de ganancias globales TCX individuales no fue óptima ya que la energía transmitida se computó para la LPC residual en tanto que la TCX computa su ganancia en el dominio ponderado.

**[0066]** La diferencia principal del esquema previo es que la ganancia global representa ahora la energía de la señal ponderada en vez de la energía de la excitación.

**[0067]** En cuanto a la corriente de bits, las modificaciones comparadas con el primer procedimiento son las siguientes:

- Una ganancia global codificada en 8 bits con el mismo cuantificador que en el modo FD. Ahora, ambos modos LPD y FD comparten el mismo elemento de corriente de bits. Resultó que la ganancia global en AAC tiene suficientes motivos para codificarse en 8 bits con un cuantificador de dichas características. Definitivamente, 8 bits resultan demasiados para la ganancia global del modo LPD, que se puede codificar sólo en 6 bits. Sin embargo, es el costo que resulta de la unificación.

- Codificar las ganancias globales individuales de TCX con una codificación diferencial, utilizando:

- 1 bit para TCX1024, códigos de longitud fijos.
- 4 bits en promedio para TCX256 y TCX 512, códigos de longitud variable (Huffman)

**[0068]** En cuanto al consumo de bits, el segundo procedimiento difiere del primer procedimiento porque:

- Para ACELP: el mismo consumo de bits que anteriormente
- Para TCX1024: +2 bits
- Para TCX512: +2 bits en promedio

- Para TCX256: el mismo consumo de bits promedio que anteriormente

**[0069]** En cuanto a la calidad, el segundo procedimiento difiere del primer procedimiento porque:

- las porciones de audio de TCX deberían sonar igual ya que toda la granularidad de cuantificación se mantuvo sin cambios.
- se esperaría que las porciones de audio de ACELP fueran ligeramente mejores ya que la predicción ha mejorado. Las estadísticas recolectadas muestran menor cantidad de valores atípicos en el ajuste de ganancia que en la ACELP corriente.

**[0070]** Véase, por ejemplo, la figura 3. La figura 3 ilustra el generador de excitación 66 que comprende un filtro de ponderación  $W(z)$  100, seguido por un computador de energía 102 y una etapa de cuantificación y codificación 104, así como también una etapa de decodificación 106. En efecto, dichos elementos están dispuestos entre sí al igual que los elementos 82 y 88 en la figura 2.

**[0071]** El filtro de ponderación está definido como:

$$W(z) = A(z/\lambda)$$

en el cual  $\lambda$  es un factor de ponderación perceptual que puede estar fijado a 0,92.

**[0072]** Por lo tanto, de acuerdo con el segundo procedimiento, la ganancia global común para los submarcos TCX y CELP 52 se deduce de un cálculo de energía que se lleva a cabo cada 2024 muestras en la señal ponderada, es decir, en unidades de los marcos LPC 32. La señal ponderada se computa en el codificador dentro del filtro 100 filtrando la señal original 24 mediante el filtro de ponderación  $W(z)$  que se deduce de los coeficientes LPC como salida mediante el analizador de LP 62. De este modo, el pre-énfasis mencionado anteriormente no forma parte de  $W(z)$ . Sólo se utiliza antes de computar los coeficientes LPC, es decir, dentro de o en frente del analizador de LP 62, y antes de ACELP, es decir, dentro de o en frente del generador de excitación 66. De alguna forma, el pre-énfasis ya está reflejado en los coeficientes de  $A(z)$ .

**[0073]** El computador de energía 102 luego determina la energía que es:

$$nrg = \sum_{n=0}^{1023} w[n]^* w[n]$$

**[0074]** La etapa de cuantificación y codificación 104 luego cuantifica la ganancia  $ganancia\_global$  ( $global\_gain$ ) en 8 bits en el dominio logarítmico sobre la base de la energía media  $nrg$  por:

$$ganancia\_global = \left\lfloor 4 \cdot \log_2 \left( \sqrt{\frac{nrg}{1024}} \right) + 0.5 \right\rfloor$$

**[0075]** La ganancia global cuantificada luego se obtiene mediante la etapa de decodificación 106 por:

$$\hat{g} = 2^{\frac{ganancia\_global}{4}}$$

**[0076]** Según se describirá con más detalle a continuación con respecto al decodificador, para mantener la sincronía anteriormente mencionada entre el codificador y el decodificador (excitación  $nupdate$ ), el generador de excitación 66 puede, al optimizar o luego de haber optimizado los índices de libro de código,

a) realizar un cálculo aproximado de la energía de excitación del libro de código de innovación determinada por una primera información contenida dentro del índice de libro de código de innovación – candidato provisional o finalmente transmitido – a saber, el número, posiciones y signos anteriormente mencionados de los pulsos de vector

de libro de código de innovación, filtrando el vector de libro de código de innovación respectivo con el filtro de síntesis LP, sin embargo ponderado con el filtro de ponderación  $W(z)$  y el filtro de de-énfasis, es decir, la inversa del filtro de énfasis, (filtro  $H2(z)$ , véase a continuación), y determinando la energía del resultado,

5 b) formar una relación entre la energía así derivada y una energía  $\bar{E} = 20 \cdot \log(\hat{g})$  determinada por la ganancia global (global\_gain) a fin de obtener una ganancia de predicción  $g'_c$

c) multiplicar la ganancia de predicción  $g'_c$  con el factor de corrección de libro de código de innovación  $\hat{\gamma}$  para proporcionar la ganancia de libro de código de innovación real  $\hat{g}_c$

10 d) generar realmente la excitación del libro de código combinando la excitación del libro de código adaptativo y la excitación del libro de código de innovación ponderando el último con la ganancia de libro de código de innovación real  $\hat{g}_c$ .

15 **[0077]** En particular, la cuantificación lograda de este modo tiene la misma granularidad que la cuantificación de la ganancia global del modo FD. Nuevamente, el generador de excitación 66 puede adoptar, y tratar como una constante, la ganancia global cuantificada  $\hat{g}$  al optimizar la excitación del libro de código de innovación. En particular, el generador de excitación 66 puede fijar el factor de corrección de la excitación del libro de código de innovación  $\hat{\gamma}$  encontrando el índice óptimo del libro de código de innovación de modo tal que la ganancia de libro de código fija cuantificada óptima tiene como resultado, a saber:

$$\hat{g}_c = \hat{\gamma} \cdot g'_c,$$

de acuerdo con:

$$25 \quad g'_c = 10^{0.05G'_c}$$

$$G'_c = \bar{E} - E_i - 12$$

$$30 \quad \bar{E} = 20 \cdot \log(\hat{g})$$

$$E_i = 10 \cdot \log\left(\frac{1}{64} \sum_{n=0}^{63} c_w^2[n]\right),$$

35 en el cual  $c_w$  es la innovación (sic) es el vector de innovación  $c[n]$  en el dominio ponderado obtenido por una convolución a partir de  $n = 0$  hasta 63 de acuerdo con:

$$c_w[n] = c[n] * h2[n],$$

40 en el cual  $h2$  es la respuesta de impulso del filtro de síntesis ponderado

$$H2(z) = \frac{\hat{W}(z)}{\hat{A}(z)} H_{de\_emph}(z) = \frac{\hat{A}(z/0.92)}{\hat{A}(z)(1 - 0.68z^{-1})}$$

donde  $\gamma = 0.92$  y  $\alpha = 0.68$ , por ejemplo.

45 **[0078]** La ganancia TCX se codifica transmitiendo el elemento **ganancia global delta (delta\_global\_gain)** codificado con Códigos de Longitud Variable.

[0079] Si la TCX tiene un tamaño de 1024 sólo se utiliza 1 bit para el elemento **ganancia global delta** (**delta\_global\_gain**), en tanto que **ganancia global** (**global\_gain**) se recalcula y recuantifica:

$$\begin{matrix} \text{(ganancia global)} & \text{(ganancia tcx)} \\ \text{global\_gain} = \lfloor 4 \cdot \log_2(\text{gain\_tcx}) + 0.5 \rfloor \end{matrix}$$

$$\hat{g} = 2^{\frac{s_{index}}{4}}$$

10

$$\begin{matrix} \text{(ganancia global delta)} & \text{(ganancia tcx)} \\ \text{delta\_global\_gain} = \lfloor 8 \cdot \log_2\left(\frac{\text{gain\_tcx}}{\hat{g}}\right) + 0.5 \rfloor \end{matrix}$$

15

[0080] Se codifica de la siguiente manera:

$$\begin{matrix} \text{(ganancia tcx)} & \text{(ganancia global delta)} \\ \text{gain\_tcx} = 2^{\frac{\text{delta\_global\_gain}}{8}} \cdot \hat{g} \end{matrix}$$

20

[0081] Por otra parte, para los otros tamaños de TCX, la **ganancia global delta** (**delta\_global\_gain**) se codifica de la siguiente manera:

25

$$\begin{matrix} \text{(ganancia global delta)} & \text{(ganancia tcx)} \\ \text{delta\_global\_gain} = \lfloor (28 \cdot \log_2\left(\frac{\text{gain\_tcx}}{\hat{g}}\right) + 64) + 0.5 \rfloor \end{matrix}$$

30

[0082] La ganancia TCX luego se codifica de la siguiente manera:

35

$$\begin{matrix} \text{(ganancia tcx)} & \text{(ganancia global delta-64)} \\ \text{gain\_tcx} = 10^{\frac{\text{delta\_global\_gain}-64}{28}} \cdot \hat{g} \end{matrix}$$

40

la **ganancia global delta** (**delta\_global\_gain**) se puede codificar directamente en 7 bits o utilizando los códigos de Huffman, que pueden producir 4 bits en promedio.

[0083] Finalmente y en ambos casos la ganancia final se deduce:

45

$$\begin{matrix} \text{(ganancia tcx)} \\ g = \frac{\text{gain\_tcx}}{2 \cdot rms} \end{matrix}$$

50

[0084] A continuación, un decodificador multimodo de audio correspondiente a la forma de realización de la figura 1, de acuerdo con las dos alternativas descritas con respecto a la figura 2 y 3, se describe con respecto a la figura 4.

55

[0085] El decodificador multimodo de audio de la figura 4 está indicado generalmente con el signo de referencia 120 y comprende un demultiplexor 122, un decodificador FD 124, y decodificador LPC 126 compuesto por un decodificador de TCX 128 y un decodificador de CELP 130, y un manipulador de solapamiento/transición 132.

[0086] El demultiplexor comprende una entrada 134 que forma simultáneamente la entrada del decodificador multimodo de audio 120. La corriente de bits 36 de la figura 1 ingresa a la entrada 134. El demultiplexor 122 comprende diversas salidas conectadas a los decodificadores 124, 128, y 130, y distribuye elementos de sintaxis comprendidos en la corriente de bits 134 a la máquina decodificadora individual. En efecto, el multiplexor 132

60

distribuye los marcos 34 y 35 de la corriente de bits 36 con el respectivo decodificador 124, 128 y 130, respectivamente.

**[0087]** Cada uno de los decodificadores 124, 128, y 130 comprende una salida de dominio de tiempo conectada a una entrada respectiva del manipulador de solapamiento–transición 132. El manipulador de solapamiento–transición 132 debe llevar a cabo la respectiva manipulación de solapamiento/transición en las transiciones entre marcos consecutivos. Por ejemplo, el manipulador de solapamiento/transición 132 puede llevar a cabo el procedimiento de solapar/agregar relacionado con las ventanas consecutivas de los marcos FD. Esto mismo se aplica a los submarcos de TCX. Si bien no se describe en detalle con respecto a la figura 1, por ejemplo, también el generador de excitación 60 utiliza la división en ventanas seguida por una transformación de tiempo al dominio espectral a fin de obtener los coeficientes de transformada para representar la excitación, y las ventanas se pueden solapar entre sí. Cuando se lleva a cabo la transición a/desde los submarcos de CELP, el manipulador de solapamiento/transición 132 puede llevar a cabo medidas especiales a fin de evitar el solapamiento. A tal fin, el manipulador de solapamiento/transición 132 puede estar controlado por los elementos de sintaxis respectivos transmitidos mediante la corriente de bits 36. Sin embargo, como dichas medidas de transmisión exceden el tema central de la presente solicitud, se hace referencia, por ejemplo, al estándar ACELP W+ para soluciones ejemplares ilustrativas al respecto.

**[0088]** El decodificador FD 124 comprende un decodificador sin pérdidas 134, un módulo de decuantificación y reajuste de escala 136, y un retransformador 138, que están conectados en serie entre el demultiplexor 122 y el manipulador de solapamiento/transición 132 en este orden. El decodificador sin pérdidas 134 recupera, por ejemplo, los factores de escala de la corriente de bits que están, por ejemplo, codificados diferencialmente en la misma. El módulo de cuantificación y reajuste de escala 136 recupera los coeficientes de transformada, por ejemplo, ajustando a escala los valores de coeficientes de transformada para las líneas espectrales individuales con los correspondientes factores de escala de las bandas de factores de escala a las cuales pertenecen dichos valores de coeficientes de transformada. El retransformador 138 lleva a cabo una transformación espectral al dominio de tiempo sobre los coeficientes de transformada así obtenidos tal como una MDCT inversa, a fin de obtener una señal de dominio de tiempo para ser enviada al manipulador de solapamiento/transición 132. Cualquiera del módulo de decuantificación y reajuste de escala 136 o retransformador 138 utiliza el elemento de sintaxis de ganancia global (global\_gain) transmitido dentro de la corriente de bits para cada marco FD, de modo tal que la señal del dominio de tiempo resultante de la transformación es ajustada a escala por el elemento de sintaxis (es decir, ajustado a escala linealmente con alguna función exponencial del mismo). En efecto, el ajuste de escala puede ser llevado a cabo antes de la transformación espectral al dominio de tiempo o posteriormente a la misma.

**[0089]** El decodificador de TCX 128 comprende un generador de excitación 140, un formador espectral 142, y un convertidor de coeficientes LP 144. El generador de excitación 140 y formador espectral 142 están conectados en serie entre el demultiplexor 122 y otra entrada del manipulador de solapamiento/transición 132, y el convertidor de coeficientes LP 144 proporciona una entrada adicional del formador espectral 142 con valores de ponderación espectral obtenidos a partir de los coeficientes LPC transmitidos mediante la corriente de bits. En particular, el decodificador de TCX 128 opera en los submarcos de TCX entre los submarcos 52. El generador de excitación 140 administra la información espectral entrante en forma similar a los componentes 134 y 136 del decodificador FD 124. Es decir, el generador de excitación 140 decuantifica y reajusta a escala los valores de coeficientes de transformada transmitidos dentro de la corriente de bits a fin de representar la excitación en el dominio espectral. Los coeficientes de transformada obtenidos de este modo son ajustados a escala por el generador de excitación 140 con un valor que corresponde a una suma del elemento de sintaxis de ganancia global delta (delta\_global\_gain) transmitido para el submarco de la TCX actual 52 y el elemento de sintaxis de ganancia global (global\_gain) transmitido para el marco actual 32 al que pertenece el submarco de TCX actual 52. Por lo tanto, el generador de excitación 140 produce una representación espectral de la excitación para el submarco corriente ajustado a escala de acuerdo con la ganancia global delta (delta\_global\_gain) y la ganancia global (global\_gain). El convertidor LPC 134 convierte los coeficientes LPC transmitidos dentro de la corriente de bits a modo de, por ejemplo, interpolación y codificación diferencial, o similares, en valores de ponderación espectral, a saber un valor de ponderación espectral por coeficiente de transformada del espectro de la excitación producida por el generador de excitación 140. En particular, el convertidor de coeficientes LP 144 determina dichos valores de ponderación espectral de modo tal que los mismos se asemejen a una función de transferencia del filtro de síntesis de predicción lineal. En otras palabras, ellos se asemejan a una función de transferencia del filtro de síntesis LP  $\hat{H}(z)$ . El formador espectral 140 pondera espectralmente los coeficientes de transformada introducidos por el generador de excitación 140 mediante las ponderaciones espectrales obtenidas por el convertidor de coeficientes LP 144 a fin de obtener coeficientes de transformada ponderados espectralmente que luego son sometidos a una transformación espectral al dominio de tiempo en el retransformador 146 de modo tal que el retransformador 146 produce una versión reconstruida de la representación decodificada del contenido de audio del submarco de TCX actual. Sin embargo, cabe destacar que, tal como se indicó anteriormente, se puede llevar a cabo un post–procesamiento en la salida del retransformador 146 antes de enviar la señal del dominio de tiempo al manipulador de solapamiento/transición 132. En cualquiera de los casos, el nivel de la señal del dominio de tiempo producida por el retransformador 146 es nuevamente controlado por el elemento de sintaxis de ganancia global (global\_gain) del respectivo marco de LPC 32.

**[0090]** El decodificador de CELP 130 de la figura 4 comprende un constructor de libro de código de innovación 148, un constructor de libro de código adaptativo 150, un adaptador de ganancia 152, un combinador 154, y un filtro de síntesis LP 156. El constructor de libro de código de innovación 148, el adaptador de ganancia 152, el combinador 154, y el filtro de síntesis LP 156 están conectados en serie entre el demultiplexor 122 y el manipulador de solapamiento/transición 132. El constructor de libro de código adaptativo 150 tiene una entrada conectada al demultiplexor 122 y una salida conectada a otra entrada del combinador 154, que a su vez, puede estar incorporada como un sumador según se indica en la figura 4. Una entrada adicional del constructor de libro de código adaptativo 150 está conectada a una salida del sumador 154 a fin de obtener la excitación pasada de la misma. El adaptador de ganancia 152 y el filtro de síntesis LP 156 tienen entradas LPC inputs conectadas a una salida determinada del multiplexor 122.

**[0091]** Luego de haber descrito la estructura del decodificador de TCX y del decodificador de CELP, a continuación se describirá con mayor detalle la funcionalidad de los mismos. La descripción comienza en principio con la funcionalidad del decodificador de TCX 128 y luego continúa con la descripción de la funcionalidad del decodificador de CELP 130. Según se describió anteriormente, los marcos LPC 32 están subdivididos en uno o más submarcos 52. En general, los submarcos de CELP 52 tienen una longitud limitada de 256 muestras de audio. Los submarcos de TCX 52 pueden tener distintas longitudes. Los submarcos 52 de TCX 20 o TCX 256, por ejemplo, tienen una longitud de muestra de 256. De la misma manera, los submarcos 52 de TCX 40 (TCX 512) tienen una longitud de 512 muestras de audio, y los submarcos de TCX 80 (TCX 1024) pertenecen a una longitud de muestra de 1024, es decir, pertenecen a todo el marco de LPC 32. Los submarcos de TCX 40 pueden estar simplemente ubicados en los dos cuartos delanteros del marco de LPC actual 32, o en los dos cuartos traseros del mismo. Por lo tanto, en conjunto, existen 26 combinaciones distintas de tipos distintos de submarco en los que puede estar dividido un marco de LPC 32.

**[0092]** Por lo tanto, según se mencionó anteriormente, los submarcos de TCX 52 tienen distinta longitud. Considerando las longitudes de las muestras descritas anteriormente, a saber 256, 512, y 1024, se podría creer que dichos submarcos de TCX no se solapan entre sí. Sin embargo, esto no es correcto en lo que respecta a las longitudes de ventanas y a las longitudes de transformadas medidas en las muestras, y que se utilizan a fin de llevar a cabo la descomposición espectral de la excitación. Las longitudes de transformadas utilizadas por el divisor en ventanas 38 se extienden, por ejemplo, más allá del extremo delantero y trasero de cada submarco de TCX actual y la ventana correspondiente utilizada para la división en ventanas de la excitación está adaptada para extenderse fácilmente en regiones más allá de los extremos delanteros y traseros del respectivo submarco de TCX actual, a fin de comprender porciones no cero que se solapan con submarcos precedentes y sucesivos del submarco actual para permitir, por ejemplo, el solapamiento–cancelación conocido a partir de la codificación FD. Por lo tanto, el generador de excitación 140 recibe coeficientes espectrales cuantificados desde la corriente de bits y reconstruye el espectro de excitación de los mismos. Dicho espectro se ajusta a escala dependiendo de una combinación de ganancia global delta (delta\_global\_gain) del submarco de TCX actual y marco global (global\_frame) del marco actual 32 al cual pertenece el submarco actual. En particular, la combinación puede comprender una multiplicación entre ambos valores en el dominio lineal (correspondiente a una suma en el dominio logarítmico), en la cual ambos elementos de sintaxis de ganancia están definidos. Por consiguiente, el espectro de excitación es de este modo ajustado a escala de acuerdo con el elemento de sintaxis de ganancia global (global\_gain). El formador espectral 142 luego lleva a cabo un modelado de ruido de dominio de frecuencia basado en LPC a los coeficientes espectrales resultantes seguido por una transformación de MDCT inversa llevada a cabo por el retransformador 146 para obtener la señal de síntesis de dominio de tiempo. El manipulador de solapamiento/transición 132 puede llevar a cabo el proceso de solapar y agregar entre los submarcos de TCX consecutivos.

**[0093]** El decodificador de CELP 130 actúa en los submarcos de CELP mencionados anteriormente que tienen, tal como se indicó anteriormente, una longitud de 256 muestras de audio cada uno. Tal como se indicó anteriormente, el decodificador de CELP 130 está configurado para construir la excitación actual como una combinación o adición de vectores ajustados a escala del libro de código adaptativo y del libro de código de innovación. El constructor de libro de código adaptativo 150 utiliza el índice de libro de código adaptativo que es recuperado desde la corriente de bits mediante el demultiplexor 122 para encontrar un número entero y una parte fraccional de un retardo de altura de tono. El constructor de libro de código adaptativo 150 luego puede encontrar un vector inicial de excitación del libro de código adaptativo  $v'(n)$  interpolando la excitación pasada  $u(n)$  en el retardo de altura de tono y fase, es decir, fracción, utilizando un filtro de interpolación FIR. La excitación del libro de código adaptativo es computada para un tamaño de 64 muestras. Dependiendo de un elemento de sintaxis denominado índice de filtro adaptativo recuperado por la corriente de bits, el constructor de libro de código adaptativo puede decidir si el libro de código adaptativo filtrado es

$$v(n) = v'(n) \text{ o}$$

$$v(n) = 0,18 v'(n) + 0,64 v'(n-1) + 0,18 v'(n-2).$$

**[0094]** El constructor de libro de código de innovación 148 utiliza el índice de libro de código de innovación recuperado desde la corriente de bits para extraer posiciones y amplitudes, es decir, signos, de pulsos de excitación dentro de un vector de código algebraico, es decir, el vector de código de innovación  $c(n)$ . Es decir,

5

$$c(n) = \sum_{i=0}^{M-1} s_i \delta(n - m_i)$$

**[0095]** En el cual  $m_i$  y  $s_i$  son las posiciones y signos de pulsos y  $M$  es el número de pulsos. Una vez que el vector de código algebraico  $c(n)$  está decodificado se lleva a cabo un procedimiento de definición de altura de tono. En primer lugar, el  $c(n)$  es filtrado mediante un filtro de pre-énfasis definido a continuación:

15  $F_{emph}(z) = 1 - 0.3 z^{-1}$

**[0096]** El filtro de pre-énfasis cumple la función de reducir la energía de excitación a bajas frecuencias. Naturalmente, el filtro de pre-énfasis puede estar definido de otra manera. A continuación se puede llevar a cabo una periodicidad mediante el constructor de libro de código de innovación 148. Dicha mejora de periodicidad se puede llevar a cabo por medio de un pre-filtro adaptativo con una función de transferencia definida como:

25 
$$F_p(z) = \begin{cases} 1 & \text{si } n < \min(T, 64) \\ (1 + 0.85z^{-T}) & \text{si } T < 64 \text{ y } T \leq n < \min(2T, 64) \\ 1/(1 - 0.85z^{-T}) & \text{si } 2T < 64 \text{ y } 2T \leq n < 64 \end{cases}$$

30 donde  $n$  es la posición real en unidades de grupos inmediatamente consecutivos de 64 muestras de audio, y donde  $T$  es una versión redondeada de la parte entera  $T_0$  y de la parte fraccional  $T_{0, \text{frac}}$  del retardo de altura de tono indicado por:

35 
$$T = \begin{cases} T_0 + 1 & \text{si } T_{0, \text{frac}} > 2 \\ T_0 & \text{de otro modo} \end{cases}$$

**[0097]** El pre-filtro adaptativo  $F_p(z)$  colorea el espectro atenuando las frecuencias inter-armónicas que son desagradables para el oído humano en el caso de señales de voz.

**[0098]** El índice de libro de código adaptativo y de innovación recibido dentro de la corriente de bits proporciona directamente la ganancia de libro de código adaptativo  $\hat{g}_p$  y el factor de corrección de la ganancia de libro de código de innovación  $\hat{\gamma}$ . La ganancia de libro de código de innovación luego se computa multiplicando el factor de corrección de ganancia  $\hat{\gamma}$  por una ganancia estimada de libro de código de innovación  $\gamma'_c$ . Esto es llevado a cabo por el adaptador de ganancia 152.

**[0099]** De acuerdo con la primera alternativa anteriormente mencionada, el adaptador de ganancia 152 lleva a cabo los siguientes pasos:

**[0100]** En primer lugar,  $\bar{E}$  que es transmitido mediante la ganancia global (global\_gain) transmitida y representa la energía de excitación media por supermarco 32, sirve como una ganancia estimada  $G'_c$  en db, es decir,

55 
$$\bar{E} = G'_c$$

**[0101]** La energía de excitación innovativa media en un supermarco 32,  $\bar{E}$ , es codificada de este modo con 6 bits por supermarco por la ganancia global (global\_gain), y  $\bar{E}$  proviene de la ganancia global (global\_gain) mediante su versión cuantificada  $\hat{g}$  por:

5 
$$\bar{E} = 20 \cdot \log(\hat{g})$$

**[0102]** La ganancia de predicción en el dominio lineal es luego obtenida por el adaptador de ganancia 152 por:

10 
$$g'_c = 10^{0.05G'_c}$$

**[0103]** La ganancia de libro de código fija cuantificada es luego computada por el adaptador de ganancia 152 por

15 
$$\hat{g}_c = \hat{\gamma} \cdot g'_c$$

**[0104]** Tal como se describió, el adaptador de ganancia 152 luego ajusta a escala la excitación del libro de código de innovación con  $\hat{g}_c$ , en tanto que el constructor de libro de código adaptativo 150 ajusta a escala la excitación del libro de código adaptativo con  $\hat{g}_p$ , y se forma una suma ponderada de ambas excitaciones de libro de código en el combinador 154.

20 **[0105]** De acuerdo con la segunda alternativa de las alternativas anteriormente descritas, la ganancia de libro de código fija estimada  $g_c$  es formada por el adaptador de ganancia 152 de la siguiente manera:

**[0106]** En primer lugar se encuentra la energía media de innovación. La energía media de innovación  $E_i$  representa la energía de innovación en el dominio ponderado. Se calcula convolucionando el código de innovación con la respuesta de impulsos  $h_2$  del filtro de síntesis ponderado siguiente:

25 
$$H_2(z) = \frac{\hat{W}(z)}{\hat{A}(z)} H_{de\_emph}(z) = \frac{\hat{A}(z/0.92)}{\hat{A}(z) \cdot (1 - 0.68z^{-1})}$$

30 **[0107]** La innovación en el dominio ponderado luego se obtiene mediante una convolución a partir de  $n=0$  a 63:

$$c_w[n] = c[n] * h_2[n]$$

35 **[0108]** La energía es entonces:

$$E_i = 10 \cdot \log\left(\frac{1}{64} \sum_{n=0}^{63} c_w^2[n]\right)$$

**[0109]** Luego, la ganancia estimada  $G'_c$  en db es encontrada por

40 
$$G'_c = \bar{E} - E_i - 12$$

donde, nuevamente,  $\bar{E}$  es transmitido mediante la ganancia global (global\_gain) transmitida y representa la energía de excitación media por supermarco 32 en el dominio ponderado. La energía media en un supermarco 32,  $\bar{E}$ , es codificada de este modo con 8 bits por supermarco por ganancia global (global\_gain), y  $\bar{E}$  proviene de la ganancia global (global\_gain) mediante su versión cuantificada  $\hat{g}$  por:

45 
$$\bar{E} = 20 \cdot \log(\hat{g})$$

[0110] La ganancia de predicción en el dominio lineal luego es obtenida por el adaptador de ganancia 152 por:

$$g'_c = 10^{0.05G'_c}$$

5 [0111] La ganancia de libro de código fija cuantificada luego es obtenida por el adaptador de ganancia 152 por

$$\hat{g}_c = \hat{\gamma} \cdot g'_c$$

10 [0112] La descripción anterior no abundó en detalles en lo que respecta a la fijación de la ganancia TCX del espectro de excitación de acuerdo con las dos alternativas descritas anteriormente. La ganancia TCX, por la cual el espectro es ajustado a escala, es – como ya se describió anteriormente – codificada transmitiendo el elemento de delta ganancia global (delta\_global\_gain) codificado en 5 bits en el lado codificador de acuerdo con:

$$15 \quad \begin{matrix} \text{(ganancia global delta)} & \text{(ganancia tcx)} \\ \text{delta\_global\_gain} = \left[ 4 \cdot \log_2 \left( \frac{\text{gain\_tcx}}{\hat{g}} \right) + 10 \right] + 0.5 \end{matrix}$$

20 [0113] Está decodificada por el generador de excitación 140, por ejemplo, de la siguiente manera:

$$25 \quad \begin{matrix} \text{(ganancia global delta-10)} \\ \text{gain\_tcx} = 2^{\frac{\text{delta\_global\_gain}-10}{4}} \cdot \hat{g} \end{matrix},$$

(ganancia tcx)

30 donde  $\hat{g}$  indica la versión cuantificada de ganancia global (global\_gain) de acuerdo con  $\hat{g} = 2^{\frac{\text{global\_gain}}{4}}$  con, a su vez, la ganancia global (global\_gain) presentada dentro de la corriente de bits para el marco de LPC 32 al cual pertenece el marco de TCX actual.

35 [0114] Luego, el generador de excitación 140 ajusta a escala el espectro de excitación multiplicando cada coeficiente de transformada con g donde:

$$40 \quad g = \frac{\text{(ganancia tcx)} \text{ gain\_tcx}}{2.rms}$$

45 [0115] De acuerdo con el segundo procedimiento presentado anteriormente, la ganancia TCX es codificada transmitiendo el elemento de ganancia global delta (delta-global-gain) codificado, por ejemplo, con códigos de longitud variable. Si el submarco de TCX considerado actualmente tiene un tamaño de 1024, sólo se puede utilizar 1-bit para el elemento de ganancia global delta (delta-global-gain), en tanto que la ganancia global (global-gain) se puede recalcular y recuantificar en el lado codificador, de acuerdo con:

$$50 \quad \begin{matrix} \text{(ganancia global)} & \text{(ganancia tcx)} \\ \text{global\_gain} = \left[ 4 \cdot \log_2 (\text{gain\_tcx}) + 0.5 \right] \end{matrix}$$

[0116] El generador de excitación 140 luego obtiene la ganancia TCX por

$$55 \quad \hat{g} = 2^{\frac{g_{index}}{4}}$$

[0117] Luego computa

$$5 \quad \text{gain\_tcx} = 2^{\frac{(\text{ganancia global delta})}{\text{delta\_global\_gain}} \cdot \hat{g}}$$

10 [0118] De otro modo, para los otros tamaños de TCX, la ganancia global delta (delta-global-gain) puede ser computada por el generador de excitación 140 de la siguiente manera:

$$15 \quad \text{delta\_global\_gain} = \left[ \frac{(\text{ganancia tcx})}{(28 \cdot \log(\frac{\text{gain\_tcx}}{\hat{g}}) + 64) + 0.5} \right]$$

20 [0119] La ganancia TCX es luego decodificada por el generador de excitación 140 de la siguiente manera:

$$25 \quad \text{gain\_tcx} = 10^{\frac{(\text{ganancia global delta-64})}{\text{delta\_global\_gain-64}} \cdot \hat{g}}$$

luego computando

$$30 \quad g = \frac{(\text{ganancia tcx})}{2 \cdot \text{rms}}$$

35 [0120] A fin de obtener la ganancia por la cual el generador de excitación 140 ajusta a escala cada coeficiente de transformada.

40 [0121] Por ejemplo, la ganancia global (delta\_global\_gain) puede ser codificada directamente en 7-bits o utilizando los códigos de Huffman que pueden producir 4-bits en promedio. Por lo tanto, de acuerdo con la forma de realización anterior, es posible codificar el contenido de audio utilizando múltiples modos. En la forma de realización anterior se han utilizado tres modos de codificación, a saber FD, TCX y ACELP. A pesar de utilizar tres modos diferentes, es fácil ajustar la intensidad de sonido de la representación decodificada respectiva del contenido de audio codificado en la corriente de bits 36. En particular, de acuerdo con ambos procedimientos descritos anteriormente, es simplemente necesario aumentar/disminuir por igual los elementos de sintaxis de ganancia global (global\_gain) contenidos en cada uno de los marcos 30 y 32, respectivamente. Por ejemplo, todos dichos elementos de sintaxis de ganancia global (global\_gain) pueden ser aumentados por 2 a fin de aumentar de manera uniforme la intensidad de sonido a través de los distintos modos de codificación, o pueden ser disminuidos por 2 a fin de disminuir de manera uniforme la intensidad de sonido a través de las distintas porciones de modo de codificación.

50 [0122] Luego de haber descrito una forma de realización de la presente solicitud, a continuación se describirán otras formas de realización que son más genéricas y que se concentran individualmente en aspectos ventajosos individuales del codificador multimodo de audio y decodificador descritos anteriormente. En otras palabras, la forma de realización descrita anteriormente representa una implementación posible para cada una de las tres formas de realización descritas posteriormente. La forma de realización anterior incorpora todos los aspectos ventajosos a los cuales las formas de realización descritas a continuación sólo se refieren individualmente. Cada una de las formas de realización descritas a continuación se concentra en un aspecto del códec multimodo de audio explicado anteriormente, cuyo aspecto es ventajoso más allá de la implementación específica utilizada por la forma de realización anterior, es decir, que se puede implementar en forma diferente a la realización anterior. Los aspectos a los cuales pertenecen las formas de realización descritas a continuación, se pueden llevar a cabo en forma individual y no necesitan ser implementados en forma simultánea como se describe ilustrativamente con respecto a la forma de realización descrita anteriormente.

60 [0123] Por consiguiente, cuando se describen las formas de realización siguientes, los elementos de las formas de realización respectivas del codificador y decodificador están indicados a través de nuevos signos de referencia. Sin

embargo, detrás de dichos signos de referencia, los números de referencia de los elementos de la Figs. 1 a 4 están presentados entre paréntesis, donde los últimos elementos representan una posible implementación del elemento respectivo dentro de las figuras descritas a continuación. En otras palabras, los elementos en las figuras descritas a continuación pueden estar implementados tal como se describió anteriormente con respecto a los elementos indicados entre paréntesis detrás del número de referencia respectivo del elemento dentro de las figuras descritas a continuación, individualmente o con respecto a todos los elementos de la figura respectiva descrita a continuación.

**[0124]** Las Figs. 5a y 5b ilustran un codificador multimodo de audio y un codificador multimodo de audio de acuerdo con una primera forma de realización. El codificador multimodo de audio de la figura 5a generalmente indicado en 300 está configurado para codificar un contenido de audio 302 en una corriente de bits codificada 304 codificando un primer subconjunto de marcos 306 en un primer modo de codificación 308 y un segundo subconjunto de marcos 310 en un segundo modo de codificación 312, en el cual el segundo subconjunto de marcos 310 está compuesto respectivamente por uno o más submarcos 314, en el cual el codificador multimodo de audio 300 está configurado para determinar y codificar un valor de ganancia global (global\_gain) por marco, y determinar y codificar, por submarco de por lo menos un subconjunto 316 de los submarcos del segundo subconjunto, un elemento de corriente de bits correspondiente (ganancia global delta (delta\_global\_gain)) en forma diferencial al valor de ganancia global 318 del respectivo marco, en el cual el codificador multimodo de audio 300 está configurado de modo tal que un cambio del valor de ganancia global (global\_gain) de los marcos dentro de la corriente de bits codificada 304 tiene como resultado un ajuste de un nivel de salida de una representación decodificada del contenido de audio en el lado decodificador.

**[0125]** El decodificador multimodo de audio correspondiente 320 está ilustrado en la figura 5b. El decodificador 320 está configurado para proporcionar una representación decodificada 322 del contenido de audio 302 sobre la base de una corriente de bits codificada 304. A tal fin, el decodificador multimodo de audio 320 decodifica un valor de ganancia global (global\_gain) por marco 324 y 326 de la corriente de bits codificada 304, un primer subconjunto 324 de los marcos está codificado en un primer modo de codificación y un segundo subconjunto 326 de los marcos está codificado en un segundo modo de codificación, donde cada marco 326 del segundo subconjunto está compuesto de más de un submarco 328 y decodifica, por submarco 328 de por lo menos un subconjunto de submarcos 328 del segundo subconjunto 326 de marcos, un elemento de corriente de bits correspondiente (ganancia global delta (delta\_global\_gain)) en forma diferencial al valor de ganancia global del respectivo marco, y codificando completamente la corriente de bits utilizando el valor de ganancia global (global\_gain) y el elemento de corriente de bits correspondiente (ganancia global delta (delta\_global\_gain)) y decodificando los submarcos de por lo menos el subconjunto de submarcos del segundo subconjunto 326 de marcos y el valor de ganancia global (global\_gain) al decodificar el primer subconjunto de marcos, en el cual el decodificador multimodo de audio 320 está configurado de modo tal que un cambio en el valor de ganancia global (global\_gain) de los marcos 324 y 326 dentro de la corriente de bits codificada 304 tiene como resultado un ajuste 330 de un nivel de salida 332 de la representación decodificada 322 del contenido de audio.

**[0126]** Del mismo modo que en las formas de realización de la figura 1 a 4, el primer modo de codificación puede ser un modo de codificación de dominio de frecuencia, en tanto que el segundo modo de codificación es una codificación de modo de predicción lineal. Sin embargo, la forma de realización de la figura 5a y 5b no se limita a este caso. Sin embargo, los modos de codificación de predicción lineal tienden a necesitar una granularidad de tiempo más precisa en lo que respecta al control de ganancia global, y por consiguiente, es preferible utilizar una codificación de modo de predicción lineal para los marcos 326 y un modo de codificación de dominio de frecuencia para los marcos 324 en comparación con el caso contrario, de acuerdo con el cual el modo de codificación de dominio de frecuencia se utilizó para los marcos 326 y una codificación de modo de predicción lineal para los marcos 324.

**[0127]** Asimismo, la forma de realización de la Figs. 5a y 5b no está limitada al caso en el cual los modos de TCX y ACLEP existen para codificar los submarcos 314. En cambio, la forma de realización de la figura 1 a 4 también puede implementarse, por ejemplo, de acuerdo con la forma de realización de la Figs. 5a y 5b, si falta el modo de codificación de ACELP. En este caso, la codificación diferencial de ambos elementos, a saber ganancia global (global\_gain) y ganancia global delta (delta\_global\_gain) permitiría se justifique una mayor sensibilidad del modo de codificación de TCX contra variantes y la fijación de ganancia evitando, sin embargo, renunciar a las ventajas proporcionadas por un control de ganancia global sin el desvío de decodificar y recodificar, y sin un aumento indebido de información lateral necesaria.

**[0128]** No obstante, el decodificador multimodo de audio 320 puede estar configurado, al completar la decodificación de la corriente de bits codificada 304, para decodificar los submarcos de por lo menos el subconjunto de submarcos del segundo subconjunto 326 de marcos utilizando codificación de predicción lineal excitada por transformada (a saber los cuatro submarcos del marco izquierdo 326 en la figura 5b), y para decodificar un subconjunto desarticulado de los submarcos del segundo subconjunto 326 de los marcos mediante el uso de CELP. En este caso, el decodificador multimodo de audio 220 puede estar configurado para decodificar, por marco del segundo subconjunto

de los marcos, un elemento de corriente de bits adicional que revela una descomposición del respectivo marco en uno o más submarcos. En la forma de realización mencionada anteriormente, por ejemplo, cada marco de LPC puede tener un elemento de sintaxis contenido en el mismo, que identifica una de las veintiséis posibilidades mencionadas anteriormente para descomponer el marco de LPC actual en los marcos de TCX y ACELP. Sin embargo, nuevamente, la forma de realización de la Figs. 5a y 5b no está limitada a ACELP, y a las dos alternativas específicas descritas anteriormente con respecto a la fijación media de energía de acuerdo con el elemento de sintaxis de ganancia global (global\_gain).

**[0129]** En forma análoga a la forma de realización anterior de la Figs. 1 a 4, los marcos 326 pueden corresponder a los marcos 310 que tienen los marcos 326 o que pueden tener una longitud de muestra de 1024 muestras, y el por lo menos subconjunto de submarcos del segundo subconjunto de marcos para el cual es transmitido el elemento de corriente de bits de ganancia global delta (delta\_global\_gain), puede tener una longitud de muestra variable seleccionada entre el grupo integrado por 256, 512, y 1024 muestras, y el subconjunto desarticulado de los submarcos puede tener una longitud de muestra de 256 muestras cada uno. Los marcos 324 del primer subconjunto pueden tener una longitud de muestra igual entre sí. Tal como se describió anteriormente (sic). El decodificador multimodo de audio 320 puede estar configurado para decodificar el valor de ganancia global en 8-bits y el elemento de corriente de bits en el número variable de bits, dependiendo el número de una longitud de muestra del submarco respectivo. De la misma manera, el decodificador multimodo de audio puede estar configurado para decodificar el valor de ganancia global en 6-bits y para decodificar los elementos de corriente de bits en 5-bits. Cabe señalar que existen distintas posibilidades para codificar en forma diferencial los elementos de ganancia global delta (delta\_global\_gain).

**[0130]** Como ocurre con la forma de realización anterior de la Figs. 1 a 4, los elementos de ganancia global (global\_gain) pueden estar definidos en el dominio logarítmico, a saber lineal con la intensidad de muestra de audio. Esto mismo se aplica a la ganancia global delta (delta\_global\_gain). A fin de codificar la ganancia global delta (delta\_global\_gain), el codificador multimodo de audio 300 puede someter una relación de un elemento de ganancia lineal de los submarcos respectivos 316, tal como la ganancia TCX (gain\_TCX) mencionada anteriormente (tal como el primer factor de escala codificado en forma diferencial), y la ganancia global (global\_gain) cuantificada del marco correspondiente 310, es decir, la versión linealizada (aplicada a una función exponencial) de ganancia global (global\_gain), a un logaritmo tal como el logaritmo a la base 2, a fin de obtener el elemento de sintaxis de ganancia global delta (delta\_global\_gain) en el dominio logarítmico. Como se conoce en la técnica se puede obtener el mismo resultado llevando a cabo una sustracción en el dominio logarítmico. Por consiguiente, el decodificador multimodo de audio 320 puede estar configurado para, en primer lugar, retransferir los elementos de sintaxis de ganancia global delta (delta\_global\_gain) y ganancia global (global\_gain) por una función exponencial al dominio de tiempo a fin de multiplicar los resultados en el dominio lineal a fin de obtener la ganancia con la cual el decodificador multimodo de audio tiene que ajustar a escala los submarcos actuales tales como la excitación codificada de TCX y los coeficientes de transformada espectral de los mismos, según se describió anteriormente. Como se conoce en la técnica se puede obtener el mismo resultado agregando ambos elementos de sintaxis en el dominio logarítmico antes de llevar a cabo la transición en el dominio de tiempo.

**[0131]** Asimismo, tal como se describió anteriormente, el códec multimodo de audio de la figura 5a y 5b puede estar configurado de modo tal que el valor de ganancia global está codificado en un número fijo de, por ejemplo, ocho bits y el elemento de corriente de bits en un número variable de bits, dependiendo el número de una longitud de muestra del submarco respectivo. Alternativamente, el valor de ganancia global puede estar codificado en un número fijo de, por ejemplo, seis bits y el elemento de corriente de bits en, por ejemplo, cinco bits.

**[0132]** Por lo tanto, las formas de realización de la Figs. 5a y 5b se concentraron en la ventaja de la codificación diferencial de los elementos de sintaxis de ganancia de los submarcos a fin de tener en cuenta las distintas necesidades de los distintos modos de codificación en lo que respecta al tiempo y a la granularidad de bits en el control de ganancia, a fin de evitar, por una parte, las deficiencias de calidad indeseadas y a lograr, sin embargo, las ventajas comprendidas en el control de ganancia global, es decir, evitar la necesidad de decodificar y recodificar para llevar a cabo un ajuste de escala de la intensidad de sonido.

**[0133]** A continuación, con respecto a la Figs. 6a y 6b se describe otra forma de realización para un códec multimodo de audio y el correspondiente codificador y decodificador. La figura 6a ilustra un codificador multimodo de audio 400 configurado para codificar y contenido de audio 402 en una corriente de bits codificada 404 codificando por CELP un primer subconjunto de marcos del contenido de audio 402 indicado con el número de referencia 406 en la figura 6a, y codificando por transformada un segundo subconjunto de los marcos indicado con el número de referencia 408 en la figura 6a. El codificador multimodo de audio 400 comprende un codificador de CELP 410 y un codificador de transformada 412. El codificador de CELP 410, a su vez, comprende un analizador de LP 414 y un generador de excitación 416. El codificador de CELP está configurado para codificar un marco de corriente del primer subconjunto. A tal fin, el analizador de LP 414 genera coeficientes de filtro LPC 418 para el marco actual y codifica los mismos en la corriente de bits codificada 404. El generador de excitación 416 determina una excitación actual del marco actual del primer subconjunto, que cuando es filtrada por un filtro de síntesis de predicción lineal

basado en los coeficientes de filtro de predicción lineal 418 dentro de la corriente de bits codificada 404, recupera el marco actual del primer subconjunto, definido por una excitación pasada 420 y un índice de libro de código para el marco actual del primer subconjunto y codificando el índice de libro de código 422 en la corriente de bits codificada 404. El codificador de transformada 412 está configurado para codificar un marco actual del segundo subconjunto 408 llevando a cabo una transformación de tiempo al dominio espectral sobre una señal del dominio de tiempo para el marco actual para obtener información espectral y codificar la información espectral 424 en la corriente de bits codificada 404. El codificador multimodo de audio 400 está configurado para codificar un valor de ganancia global 426 en la corriente de bits codificada 404, dependiendo el valor de ganancia global 426 de una energía de una versión del contenido de audio del marco actual del primer subconjunto 406 filtrado con un filtro de análisis de predicción lineal dependiendo de los coeficientes de predicción lineal, o una energía de la señal del dominio de tiempo. En el caso de la forma de realización anterior de la figura 1 a 4, por ejemplo, el codificador de transformada 412 se implementó como un codificador de TCX y la señal del dominio de tiempo fue la excitación del respectivo marco. De la misma manera, el resultado de filtrar el contenido de audio 402 del marco actual del primer subconjunto (CELP) filtrado con el filtro de análisis de predicción lineal – o la versión modificada del mismo en forma del filtro de ponderación  $A(z/\gamma)$  – dependiendo de los coeficientes de predicción lineal 418, tiene como resultado una representación de la excitación. El valor de ganancia global 426 depende, por lo tanto, de ambas energías de excitación de ambos marcos.

**[0134]** Sin embargo, la forma de realización de la Figs. 6a y 6b no está limitada a la codificación de transformada de TCX. Es imaginable que otro esquema de codificación de transformada, tal como AAC, se combine con la codificación de CELP del codificador de CELP 410.

**[0135]** La figura 6b ilustra el decodificador multimodo de audio correspondiente al codificador de la figura 6a. Según la ilustración de la figura, el decodificador de la figura 6b generalmente indicado con el número de referencia 430 está configurado para proporcionar una representación decodificada 432 de un contenido de audio sobre la base de una corriente de bits codificada 434, un primer subconjunto de marcos el cual está codificado con CELP (indicado con "1" en la figura 6b), y un segundo subconjunto de marcos el cual está codificado con transformada (indicado con "2" en la figura 6b). El decodificador 430 comprende un decodificador de CELP 436 y un decodificador de transformada 438. El decodificador de CELP 436 comprende un generador de excitación 440 y un filtro de síntesis de predicción lineal 442.

**[0136]** El decodificador de CELP 440 está configurado para decodificar el marco actual del primer subconjunto. A tal fin, el generador de excitación 440 genera una excitación actual 444 del marco actual construyendo una excitación de libro de código basada en una excitación pasada 446, y un índice de libro de código 448 del marco actual del primer subconjunto dentro de la corriente de bits codificada 434, y fijar una ganancia de la excitación del libro de código basada en un valor de ganancia global 450 dentro de la corriente de bits codificada 434. El filtro de síntesis de predicción lineal está configurado para filtrar la excitación actual 444 basada en los coeficientes de filtro de predicción lineal 452 del marco actual dentro de la corriente de bits codificada 434. El resultado del filtrado de síntesis representa, o se utiliza, para obtener la representación decodificada 432 en el marco correspondiente al marco actual dentro de la corriente de bits 434. El decodificador de transformada 438 está configurado para decodificar un marco actual del segundo subconjunto de marcos construyendo la información espectral 454 para el marco actual del segundo subconjunto a partir de la corriente de bits codificada 434 y llevando a cabo una transformación espectral al dominio de tiempo sobre la información espectral para obtener una señal de dominio de tiempo de modo tal que un nivel de la señal de dominio de tiempo dependa del valor de ganancia global 450. Tal como se indicó anteriormente, la información espectral puede ser el espectro de la excitación en el caso del decodificador de transformada que es un decodificador de TCX, o el contenido de audio original en el caso un modo de codificación FD.

**[0137]** El generador de excitación 440 puede estar configurado para, al generar una excitación actual 444 del marco actual del primer subconjunto, construir una excitación del libro de código adaptativo basada en una excitación pasada y un índice de libro de código adaptativo del marco actual del primer subconjunto dentro de la corriente de bits codificada, construir una excitación del libro de código de innovación basada en un índice de libro de código de innovación para el marco actual del primer subconjunto dentro de la corriente de bits codificada, fijar, como la ganancia de la excitación del libro de código, una ganancia de la excitación del libro de código de innovación basada en el valor de ganancia global dentro de la corriente de bits codificada, y combinar la excitación del libro de código adaptativo y la excitación del libro de código de innovación para obtener la excitación actual 444 del marco actual del primer subconjunto. Es decir, un generador de excitación 444 puede estar incorporado tal como se describió anteriormente con respecto a la figura 4, pero no es necesario que esto ocurra.

**[0138]** Asimismo, el decodificador de transformada puede estar configurado de modo tal que la información espectral se refiere a una excitación actual del marco actual, y el decodificador de transformada 438 puede estar configurado para, al decodificar el marco actual del segundo subconjunto, formar espectralmente la excitación actual del marco actual del segundo subconjunto de acuerdo con una función de transferencia del filtro de síntesis de predicción lineal definida por coeficientes de filtro de predicción lineal para el marco actual del segundo subconjunto

dentro de la corriente de bits codificada 434, de modo tal que el desempeño de la transformación espectral al dominio de tiempo sobre la información espectral tiene como resultado la representación del decodificador 432 del contenido de audio. En otras palabras, el decodificador de transformada 438 puede estar incorporado como un codificador de TCX, tal como se describió anteriormente con respecto a la figura 4, pero esto no es obligatorio.

**[0139]** El decodificador de transformada 438 puede estar configurado además para llevar a cabo la información espectral convirtiendo los coeficientes de filtro de predicción lineal en un espectro de predicción lineal y ponderar la información espectral de la excitación actual con el espectro de predicción lineal. Esto se ha descrito anteriormente con respecto al número de referencia 144. Tal como se describió anteriormente, el decodificador de transformada 438 puede estar configurado para ajustar a escala la información espectral con el valor de ganancia global 450. Como tal, el decodificador de transformada 438 puede estar configurado para construir la información espectral para el marco actual del segundo subconjunto mediante el uso de coeficientes espectrales de transformada dentro de la corriente de bits codificada, y factores de escala dentro de la corriente de bits codificada para ajustar a escala los coeficientes espectrales de transformada en una granularidad espectral de bandas de factores de escala, ajustando los factores de escala basados en el valor de ganancia global, a fin de obtener la representación decodificada 432 del contenido de audio.

**[0140]** La forma de realización de la Figs. 6a y 6b resalta los aspectos ventajosos de la forma de realización de la Figs. 1 a 4, de acuerdo con los cuales es la ganancia de la excitación del libro de código de acuerdo con la cual el ajuste de ganancia de la porción codificada por CELP está acoplada a la ajustabilidad de ganancia o capacidad de control de la porción codificada por transformada.

**[0141]** La forma de realización descrita a continuación con respecto a la Figs. 7a y 7b se concentra en las porciones de códec de CELP descritas en las formas de realización mencionadas anteriormente sin necesitar la existencia de otro modo de codificación. En cambio, el concepto de codificación de CELP, descrito con respecto a la Figs. 7a y 7b, se concentra en la segunda alternativa descrita con respecto a la Figs. 1 a 4 de acuerdo con la cual la capacidad de control de ganancia de los datos codificados por CELP se lleva a cabo implementando la capacidad de control de ganancia en el dominio ponderado, a fin de lograr un ajuste de ganancia de la reproducción decodificada con una granularidad precisa posible que no es posible lograr en una CELP convencional. Asimismo, el cómputo de la ganancia anteriormente mencionada en el dominio ponderado puede mejorar la calidad de audio.

**[0142]** Nuevamente, la figura 7a ilustra el codificador y la figura 7b ilustra el decodificador correspondiente. El codificador de CELP de la figura 7a comprende un analizador de LP 502, y un generador de excitación 504, y un determinador de energía 506. El analizador de predicción lineal está configurado para generar coeficientes de predicción lineal 508 para un marco actual 510 de un contenido de audio 512 y para codificar los coeficientes de filtro de predicción lineal 508 en una corriente de bits 514. El generador de excitación 504 está configurado para determinar una excitación actual 516 del marco actual 510 como una combinación 518 de una excitación del libro de código adaptativo 520 y una excitación del libro de código de innovación 522, que cuando es filtrada por un filtro de síntesis de predicción lineal basado en los coeficientes de filtro de predicción lineal 508, recupera el marco actual 510, construyendo la excitación del libro de código adaptativo 520 mediante una excitación pasada 524 y un índice de libro de código adaptativo 526 para el marco actual 510 y codificando el índice de libro de código adaptativo 526 en la corriente de bits 514, y construyendo la excitación del libro de código de innovación definida por un índice de libro de código de innovación 528 para el marco actual 510 y codificando el índice de libro de código de innovación en la corriente de bits 514.

**[0143]** El determinador de energía 506 está configurado para determinar una energía de una versión del contenido de audio 512 del marco actual 510, filtrada por un filtro de ponderación emitido por (o proveniente de) un análisis predictivo lineal para obtener un valor de ganancia 530, y codificando el valor de ganancia 530 en la corriente de bits 514, construyendo el filtro de ponderación a partir de los coeficientes de predicción lineal 508.

**[0144]** De acuerdo con la descripción anterior, el generador de excitación 504 puede estar configurado para, al construir la excitación del libro de código adaptativo 520 y la excitación del libro de código de innovación 522, minimizar una medida de distorsión perceptual con respecto al contenido de audio 512. Asimismo, el analizador de predicción lineal 502 puede estar configurado para determinar los coeficientes de filtro de predicción lineal 508 mediante el análisis de predicción lineal aplicado sobre una versión pre-énfasis dividida en ventanas del contenido de audio y, de acuerdo con un filtro de pre-énfasis predeterminado. El generador de excitación 504 puede estar configurado para, al construir la excitación del libro de código adaptativo y la excitación del libro de código de innovación, minimizar una medida de distorsión ponderada perceptual con respecto al contenido de audio utilizando un filtro de ponderación perceptual:  $W(z) = A(z/\gamma)$ , en el cual  $\gamma$  es un factor de ponderación perceptual y  $A(z)$  es  $1/H(z)$ , en el cual  $H(z)$  es el filtro de síntesis de predicción lineal, y en el cual el determinador de energía está configurado para utilizar el filtro de ponderación perceptual como un filtro de ponderación. En particular, la minimización puede ser llevada a cabo utilizando una medida de distorsión ponderada perceptual con respecto al contenido de audio utilizando el filtro de síntesis de ponderación perceptual:

$$\frac{A(z/\gamma)}{\hat{A}(z) H_{emph}(z)},$$

en el cual  $\gamma$  es un factor de ponderación perceptual,  $\hat{A}(z)$  es una versión cuantificada del filtro de síntesis de predicción lineal  $A(z)$ ,  $H_{emph} = 1 - \alpha z^{-1}$  y  $\alpha$  es un factor de énfasis de alta frecuencia, y en el cual el determinador de energía (506) está configurado para utilizar el filtro de ponderación perceptual  $W(z) = A(z/\gamma)$  como un filtro de ponderación.

**[0145]** Asimismo, a fin de mantener la sincronía entre el codificador y el decodificador, el generador de excitación 504 puede estar configurado para llevar a cabo una actualización de excitación,

a) estimando una energía de excitación del libro de código de innovación determinada por una primera información contenida dentro del índice de libro de código de innovación (transmitido dentro de la corriente de bits), tal como el número, las posiciones y los signos anteriormente mencionados de los pulsos del vector de libro de código de innovación, filtrando el vector de libro de código de innovación respectivo con  $H_2(z)$ , y determinando la energía del resultado,

b) formando una relación entre la energía obtenida de este modo y una energía determinada por la ganancia global (global\_gain) a fin de obtener una ganancia de predicción  $g_c'$

c) multiplicando la ganancia de predicción  $g_c'$  con el factor de corrección de libro de código de innovación, es decir, la segunda información contenida dentro del índice de libro de código de innovación, para proporcionar la ganancia de libro de código de innovación real

d) generando realmente la excitación del libro de código – que sirve como excitación pasada para el próximo marco a codificar por CELP– combinando la excitación del libro de código adaptativo y la excitación del libro de código de innovación ponderando el último con la excitación de libro de código de innovación real.

**[0146]** Fig. 7b ilustra el codificador de CELP correspondiente que tiene un generador de excitación 450 y un filtro de síntesis LP 452. El generador de excitación 440 puede estar configurado para generar una excitación actual 542 para un marco actual 544, construyendo una excitación del libro de código adaptativo 546 basada en una excitación pasada 548 y un índice de libro de código adaptativo 550 para el marco actual 544, dentro de la corriente de bits, construyendo una excitación del libro de código de innovación 552 basada en un índice de libro de código de innovación 554 para el marco actual 544 dentro de la corriente de bits, computando una estimación de una energía de la excitación del libro de código de innovación ponderada espectralmente por un filtro de síntesis ponderado de predicción lineal  $H_2$  construido a partir de coeficientes de filtro de predicción lineal 556 dentro de la corriente de bits, fijando una ganancia 558 de la excitación del libro de código de innovación 552 basada en una relación entre un valor de ganancia 560 dentro de la corriente de bits y la energía estimada, y combinando la excitación del libro de código adaptativo y la excitación del libro de código de innovación para obtener la excitación actual 542. El filtro de síntesis de predicción lineal 542 filtra la excitación actual 542 basada en los coeficientes de filtro de predicción lineal 556.

**[0147]** El generador de excitación 440 puede estar configurado para, al construir la excitación del libro de código adaptativo 546, filtrar la excitación pasada 548 con un filtro dependiendo del índice de libro de código adaptativo 546. Asimismo, el generador de excitación 440 puede estar configurado para, al construir la excitación del libro de código de innovación 554 de modo tal que este último comprenda un vector cero con un número de pulsos no cero, estando indicados el número y las posiciones de los pulsos no cero por el índice de libro de código de innovación 554. El generador de excitación 440 puede estar configurado para computar el cálculo aproximado de la energía de la excitación del libro de código de innovación 554, y filtrar la excitación del libro de código de innovación 554 con

$$\frac{\hat{W}(z)}{\hat{A}(z) H_{emph}(z)},$$

en el cual el filtro de síntesis de predicción lineal está configurado para filtrar la excitación actual 542 de acuerdo con  $1/\hat{A}(z)$ , en el cual  $\hat{W}(z) = \hat{A}(z/\gamma)$  y  $\gamma$  es un factor de ponderación perceptual,  $H_{emph} = 1 - \alpha z^{-1}$  y  $\alpha$  es

un factor de énfasis de alta frecuencia, en el cual el generador de excitación 440 está configurado además para computar una suma cuadrática de muestras de la excitación filtrada del libro de código de innovación para obtener el cálculo aproximado de la energía.

5 **[0148]** El generador de excitación 540 puede estar configurado para, al combinar la excitación del libro de código adaptativo 556 y la excitación del libro de código de innovación 554, formar una suma ponderada de la excitación del libro de código adaptativo 556 ponderada con un factor de ponderación dependiendo del índice de libro de código adaptativo 556, y la excitación del libro de código de innovación 554 ponderada con la ganancia.

10 **[0149]** En la siguiente lista se describen otros aspectos para el modo LPD:

- se podrían lograr mejoras de calidad reentrenando la ganancia VQ en ACELP para que coincida con mayor precisión con las estadísticas del nuevo ajuste de ganancia.

15 • la codificación de ganancia global en AAC podría modificarse

- codificándola en 6/7 bits en lugar de 8 bits como ocurre en TCX. Esto puede funcionar para los puntos operativos actuales pero puede ser una limitación cuando la entrada de audio tiene una resolución mayor que 16 bits.

20 ○ aumentando la resolución de la ganancia global unificada para que coincida con la cuantificación de TCX (esto corresponde al segundo procedimiento descrito anteriormente): la forma en que los factores de escala son aplicados en AAC, no es necesario tener dicha cuantificación exacta. Asimismo, esto implicará una gran cantidad de modificaciones en la estructura de AAC y un mayor consumo de bits para los factores de escala.

25

- Las ganancias globales de TCX pueden ser cuantificadas antes de cuantificar los coeficientes espectrales: se lleva a cabo de esta forma en AAC y esto permite que la cuantificación de los coeficientes espectrales sea la única fuente de error. Este procedimiento parece ser la forma más elegante de hacerlo. No obstante, las ganancias globales codificadas por TCX representan comúnmente una energía, cuya cantidad también es útil en ACELP. Dicha energía se utilizó en los procedimientos de unificación de control de ganancia mencionados anteriormente como un puente entre los dos esquemas de codificación para codificar las ganancias.

30 **[0150]** Las formas de realización anteriores son transferibles a las formas de realización en las que se utiliza SBR. La codificación de envolvente de energía SBR puede ser llevada a cabo de modo tal que las energías de la banda espectral a replicar son transmitidas/codificadas con respecto a/en forma diferencial a la energía de la energía de banda base, es decir, la energía de la banda espectral a la cual se aplican las formas de realización de códec anteriormente mencionadas.

40 **[0151]** En la Replicación Espectral de Ancho de Banda (SBR, según sus siglas en inglés), la envolvente de energía es independiente de la energía de ancho de banda central. La envolvente de energía de la banda extendida luego se reconstruye absolutamente. En otras palabras, cuando el ancho de banda central está ajustado a nivel, esto no afectará la banda extendida que permanecerá sin cambios.

45 **[0152]** En la SBR se pueden utilizar dos esquemas de codificación para transmitir las energías de las distintas bandas de frecuencia. El primer esquema consiste en una codificación diferencial en la dirección de tiempo. Las energías de las distintas bandas se codifican diferencialmente a partir de las bandas correspondientes del marco previo. Mediante el uso de este esquema de codificación, las energías del marco actual se ajustarán automáticamente en caso de que las energías del marco previo ya hayan sido procesadas.

50 **[0153]** El segundo esquema de codificación es una codificación delta de las energías en la dirección de frecuencia. La diferencia entre la energía de banda actual y la energía de la banda previa en la frecuencia se cuantifica y transmite. Sólo la energía de la primera banda está codificada absolutamente. La codificación de dicha primera banda de energía puede ser modificada y realizada con respecto a la energía del ancho de banda central. De esta forma, el ancho de banda extendido es ajustado automáticamente a nivel cuando el ancho de banda central se modifica.

55

60 **[0154]** Otro procedimiento para la codificación de envolvente de energía SBR puede utilizar el cambio de paso de cuantificación de la primera energía de banda cuando utiliza la codificación delta en la dirección de frecuencia a fin de obtener la misma granularidad que para el elemento de ganancia global común del codificador central. De esta forma se podría lograr un ajuste de nivel complete modificando tanto el índice de ganancia global común del codificador central como el índice la primera energía de banda de SBR cuando se utiliza la codificación delta en la dirección de frecuencia.

- 5 **[0155]** Por lo tanto, en otras palabras, un decodificador de SBR puede comprender cualquiera de los decodificadores anteriores como un decodificador central para para decodificar la porción del codificador central de una corriente de bits. El decodificador de SBR luego puede decodificar las energías de envolvente para una banda espectral a replicar, a partir de una porción de SBR de la corriente de bits, determinar una energía de la señal de banda central y ajustar a escala las energías de envolvente de acuerdo con una energía de la señal de banda central. Al hacerlo, la banda espectral replicada de la representación reconstruida del contenido de audio tiene una energía que se ajusta a escala inherentemente con los elementos de sintaxis de ganancia global (global\_gain) anteriormente mencionados.
- 10 **[0156]** Por lo tanto, de acuerdo con las formas de realización anteriores, la unificación de la ganancia global para la USAC puede funcionar de la siguiente manera: comúnmente existe una ganancia global de 7 bits para cada marco de TCX (256, 512 o 1024 muestras de longitud), o correspondientemente un valor de energía media de 2 bits para cada marco de ACELP (256 muestras de longitud). No existe valor global por marco de 1024, en comparación con los marcos de AAC. Para unificar esto, se podría introducir un valor global por marco de 1024 con 8 bits para las partes de TCX/ACELP, y los correspondientes valores por marcos de TCX/ACELP se pueden codificar diferencialmente a dicho valor global. Debido a esta codificación diferencial se puede reducir el número de bits para estas diferencias individuales.
- 15 **[0157]** Si bien se han descrito algunos aspectos en el contexto de un aparato, es evidente que estos aspectos también representan una descripción del procedimiento correspondiente, en donde un bloque o dispositivo corresponde a un paso del procedimiento o a una característica de un paso del procedimiento. En forma análoga, los aspectos descritos en el contexto de un paso del procedimiento también representan una descripción de un bloque o elemento o característica correspondiente de un aparato correspondiente. Algunos o todos los pasos del procedimiento pueden ser ejecutados (o utilizados) por un aparato de hardware, como por ejemplo, un microprocesador, una computadora programable o un circuito electrónico. En algunas formas de realización, algunos pasos o la mayoría de los pasos más importantes del procedimiento pueden ser ejecutados por un aparato de esas características.
- 20 **[0158]** La señal de audio codificada de la invención se puede almacenar en un soporte de almacenamiento digital o se puede transmitir en un medio de transmisión tal como un medio de transmisión inalámbrico o un medio de transmisión por cable tal como Internet.
- 30 **[0159]** Dependiendo de determinados requisitos de aplicación, las formas de realización de la invención se pueden implementar en hardware o en software. La implementación se puede llevar a cabo utilizando un medio de almacenamiento digital, por ejemplo un disco flexible, un DVD, un Blue-Ray, un CD, una memoria ROM, una PROM, una EPROM, una EEPROM o una memoria FLASH, que tienen señales de control electrónicamente legibles que están almacenadas en ellos, que cooperan (o son capaces de cooperar) con un sistema de computadora programable de forma tal que el procedimiento respectivo se lleve a cabo. Por lo tanto, el medio de almacenamiento digital puede ser legible por computadora.
- 35 **[0160]** Algunas formas de realización de acuerdo con la invención comprenden un portador de datos que tiene señales de control electrónicamente legibles, que son capaces de cooperar con un sistema de computadora programable, de modo tal que se lleve a cabo uno de los procedimientos descritos en este documento.
- 40 **[0161]** En general, las formas de realización de la presente invención pueden implementarse como un producto de programa de computadora con un código de programa, cuyo código de programa se opera para llevar a cabo uno de los procedimientos cuando el producto de programa de computadora se ejecuta en una computadora. El código del programa se puede almacenar, por ejemplo, en un portador legible por computadora.
- 45 **[0162]** Otras formas de realización comprenden el programa de computadora para ejecutar uno de los procedimientos descritos en este documento, almacenados en un portador legible por computadora.
- 50 **[0163]** En otras palabras, una forma de realización del procedimiento de la invención es, por lo tanto, un programa de computadora que tiene un código de programa para ejecutar uno de los procedimientos descritos en este documento, cuando el programa de computadora se ejecuta en una computadora.
- 55 **[0164]** Otra forma de realización de los procedimientos de la invención es, por lo tanto, un soporte de datos (o un medio de almacenamiento digital, o un medio legible por computadora), que comprende, grabado en el mismo, el programa de computadora para ejecutar uno de los procedimientos descritos en este documento. El soporte de datos, el medio de almacenamiento digital o el medio grabado son típicamente tangibles y/o no transitorios.
- 60 **[0165]** Otra forma de realización del procedimiento de la invención es, por lo tanto, una corriente de datos o una secuencia de señales que representan el programa de computadora para ejecutar uno de los procedimientos

descritos en este documento. La corriente de datos o la secuencia de señales pueden estar configuradas, por ejemplo, para ser transferidas a través de una conexión de comunicación de datos, por ejemplo, a través de Internet.

5 **[0166]** Otra forma de realización comprende un medio de procesamiento, por ejemplo, una computadora, o un dispositivo lógico programable, configurados o adaptados para ejecutar uno de los procedimientos descritos en este documento.

10 **[0167]** Una forma de realización adicional comprende una computadora que tiene instalado en la misma el programa de computadora para ejecutar uno de los procedimientos descritos en este documento.

15 **[0168]** Una forma de realización adicional de acuerdo con la invención comprende un aparato o un sistema configurado para transferir (por ejemplo, en forma electrónica u óptica) un programa de computadora para ejecutar uno de los procedimientos descritos en el presente documento a un receptor. El receptor puede ser, por ejemplo, una computadora, un dispositivo móvil, un dispositivo de memoria o similares. El aparato o sistema puede comprender, por ejemplo, un procesador de archivos para transferir el programa de computadora al receptor.

20 **[0169]** En algunas formas de realización se puede utilizar un dispositivo lógico programable (por ejemplo, un arreglo de puerta programable de campo) para ejecutar todas o algunas de las funcionalidades de los procedimientos descritos en este documento. En algunas formas de realización, un arreglo de puerta programable de campo puede cooperar con un microprocesador para ejecutar uno de los procedimientos descritos en este documento. En general, los procedimientos son ejecutados preferentemente por cualquier aparato de hardware.

25 **[0170]** Las formas de realización descritas anteriormente son simplemente ilustrativas de los principios de la presente invención. Se entiende que las modificaciones y variaciones de los arreglos y los detalles descritos en este documento serán evidentes para otros expertos en la materia. Por lo tanto, la presente invención tiene el propósito de limitarse sólo al alcance de las reivindicaciones de patente inminentes y no a los detalles específicos presentados a modo de descripción y explicación de las formas de realización del presente documento.

## REIVINDICACIONES

- 5 **1.** Decodificador de audio multimodo (120; 320) para proporcionar una representación decodificada (322) del contenido de audio (24; 302) sobre la base de una corriente de bits codificada (36; 304), en el cual el decodificador multimodo de audio (120; 320) está configurado para
- 10 decodificar un valor de ganancia global por marco (324, 326) de la corriente de bits codificada (36; 304), en el cual un primer subconjunto (324) de los marcos está codificado en un primer modo de codificación y un segundo subconjunto (326) de los marcos está codificado en un segundo modo de codificación, donde cada marco del segundo subconjunto está compuesto de más de un submarco (328),
- 15 decodificar, por submarco de por lo menos un subconjunto de submarcos (328) del segundo subconjunto de marcos, un elemento de corriente de bits correspondiente en forma diferencial al valor de ganancia global del respectivo marco, y
- 20 completar la decodificación de la corriente de bits (36; 304) utilizando el valor de ganancia global y el elemento de corriente de bits correspondiente al decodificar los submarcos de por lo menos un subconjunto de submarcos (328) del segundo subconjunto de marcos y el valor de ganancia global al decodificar el primer subconjunto de marcos,
- 25 en el cual el decodificador multimodo de audio está configurado de modo tal que un cambio del valor de ganancia global de los marcos dentro de la corriente de bits codificada (36; 304) tiene como resultado un ajuste (330) de un nivel de salida (332) de la representación decodificada (322) del contenido de audio (24; 302).
- 30 **2.** El decodificador multimodo de audio de acuerdo con la reivindicación 1, en el cual el primer modo de codificación es un modo de codificación de dominio de frecuencia, y el segundo modo de codificación es un modo de codificación de predicción lineal.
- 35 **3.** El decodificador multimodo de audio de acuerdo con la reivindicación 2, en el cual el decodificador multimodo de audio está configurado para, al completar la decodificación de la corriente de bits codificada (36; 304), decodificar los submarcos de por lo menos el subconjunto de submarcos (328) del segundo subconjunto de marcos (310) utilizando decodificación de predicción lineal excitada por transformada, y decodificar un subconjunto desarticulado de los submarcos del segundo subconjunto de marcos mediante el uso de CELP.
- 40 **4.** El decodificador multimodo de audio de acuerdo con cualquiera de las reivindicaciones 1 a 3, en el cual el decodificador multimodo de audio está configurado para decodificar, por marco del segundo subconjunto (326) de los marcos, un elemento de corriente de bits adicional que revela una descomposición del respectivo marco en uno o más submarcos.
- 45 **5.** El decodificador multimodo de audio de acuerdo con cualquiera de las reivindicaciones anteriores, en el cual los marcos del segundo subconjunto tienen igual longitud, y el por lo menos subconjunto de submarcos (328) del segundo subconjunto de marcos tiene una longitud de muestra seleccionada entre el grupo integrado por 256, 512 y 1024 muestras, y un subconjunto desarticulado de los submarcos (328) tiene una longitud de 256 muestras.
- 50 **6.** El decodificador multimodo de audio de acuerdo con cualquiera de las reivindicaciones anteriores, en el cual el decodificador multimodo de audio está configurado para decodificar el valor de ganancia global en un número fijo de bits y el elemento de corriente de bits en un número variable de bits, dependiendo el número de una longitud de muestra del submarco respectivo.
- 55 **7.** El decodificador multimodo de audio de acuerdo con cualquiera de las reivindicaciones 1 a 5, en el cual el decodificador multimodo de audio está configurado para decodificar el valor de ganancia global en un número fijo de bits y para decodificar el elemento de corriente de bits en un número fijo de bits.
- 60 **8.** El decodificador de SBR, el cual comprende un decodificador central para decodificar la porción del codificador central de una corriente de bits para obtener una señal de banda central de acuerdo con cualquiera de las reivindicaciones anteriores, en el cual el decodificador de SBR está configurado para decodificar energías de envolvente para una banda espectral a replicar, a partir de una porción de SBR de la corriente de bits, y ajustar a escala las energías de envolvente de acuerdo con una energía de la señal de banda central.
- 9.** Codificador multimodo de audio configurado para codificar un contenido de audio (302) en una corriente de bits codificada (304) codificando un primer subconjunto de marcos (306) en un primer modo de codificación (308) y un segundo subconjunto de marcos (310) en un segundo modo de codificación (312), en el cual el segundo subconjunto de marcos (310) está respectivamente compuesto por uno o más submarcos (314), en el cual el codificador multimodo de audio está configurado para determinar y codificar un valor de ganancia global por marco, y determinar

- 5 y codificar, por submarcos de por lo menos un subconjunto de submarcos (314) del segundo subconjunto (310), un elemento de corriente de bits correspondiente en forma diferencial al valor de ganancia global del respectivo marco, en el cual el codificador multimodo de audio está configurado de modo tal que un cambio del valor de ganancia global de los marcos dentro de la corriente de bits codificada tiene como resultado un ajuste de un nivel de salida de una representación decodificada del contenido de audio (302) en el lado decodificador.
- 10 **10.** Procedimiento de decodificación multimodo de audio para proporcionar una representación decodificada (322) del contenido de audio (24; 302) sobre la base de una corriente de bits codificada (36; 304), en el cual el procedimiento comprende
- 15 decodificar un valor de ganancia global por marco (324, 326) de la corriente de bits codificada (36; 304), en el cual un primer subconjunto (324) de los marcos está codificado en un primer modo de codificación y un segundo subconjunto (326) de los marcos está codificado en un segundo modo de codificación, donde cada marco del segundo subconjunto está compuesto de más de un submarco (328),
- 20 decodificar, por submarco de por lo menos un subconjunto de submarcos (328) del segundo subconjunto de marcos, un elemento de corriente de bits correspondiente en forma diferencial al valor de ganancia global del respectivo marco, y
- 25 completar la decodificación de la corriente de bits (36; 304) utilizando el valor de ganancia global y el elemento de corriente de bits correspondiente al decodificar los submarcos de por lo menos un subconjunto de submarcos (328) del segundo subconjunto de marcos y el valor de ganancia global al decodificar el primer subconjunto de marcos,
- en el cual el procedimiento de decodificación multimodo de audio se lleva a cabo de modo tal que un cambio del valor de ganancia global de los marcos dentro de la corriente de bits codificada (36; 304) tiene como resultado un ajuste (330) de un nivel de salida (332) de la representación decodificada (322) del contenido de audio (24; 302).
- 30 **11.** Procedimiento de codificación multimodo de audio, el cual comprende codificar un contenido de audio (302) en una corriente de bits codificada (304) codificando un primer subconjunto de marcos (306) en un primer modo de codificación (308) y un segundo subconjunto de marcos (310) en un segundo modo de codificación (312), en el cual el segundo subconjunto de marcos (310) está respectivamente compuesto por uno o más submarcos (314), en el cual el procedimiento de codificación multimodo de audio además comprende determinar y codificar un valor de ganancia global por marco, y determinar y codificar, por submarcos de por lo menos un subconjunto de submarcos (314) del segundo subconjunto (310), un elemento de corriente de bits correspondiente en forma diferencial al valor
- 35 de ganancia global del respectivo marco, en el cual el procedimiento de codificación multimodo de audio se lleva a cabo de modo tal que un cambio del valor de ganancia global de los marcos dentro de la corriente de bits codificada tiene como resultado un ajuste de un nivel de salida de una representación decodificada del contenido de audio (302) en el lado decodificador.
- 40 **12.** El programa de computadora, el cual tiene un código de programa para llevar a cabo, cuando es ejecutado en una computadora, un procedimiento de acuerdo con la reivindicación 11.

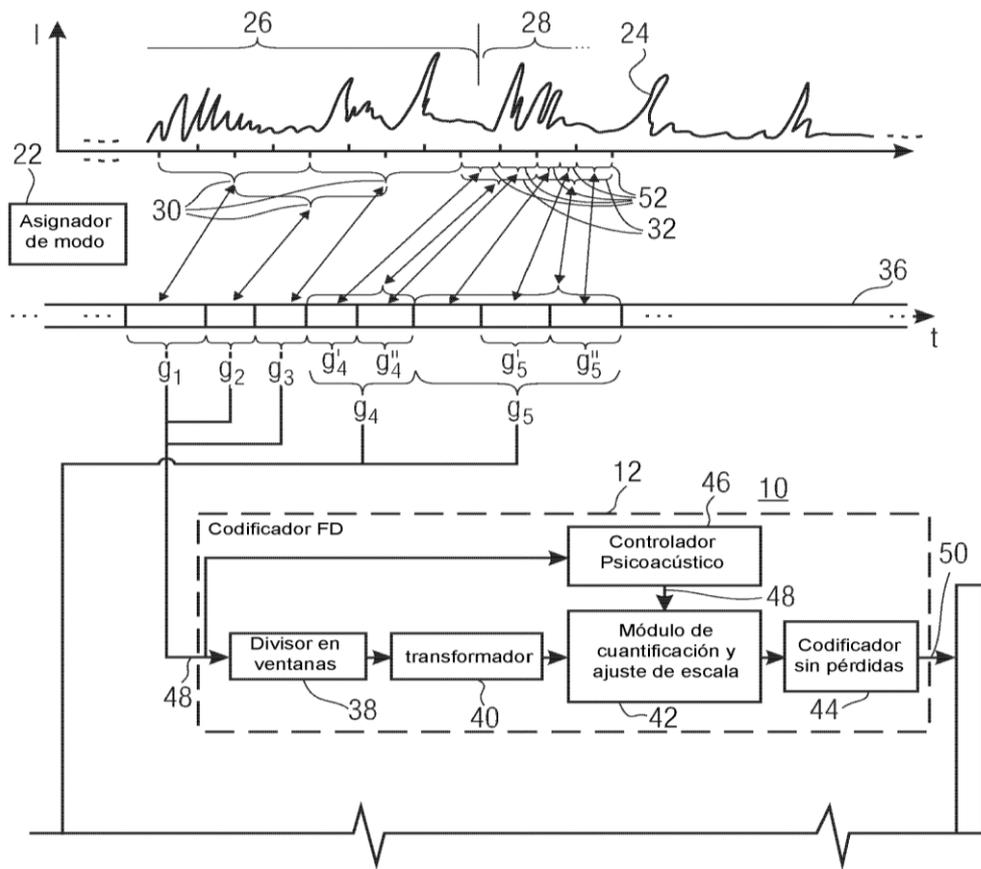
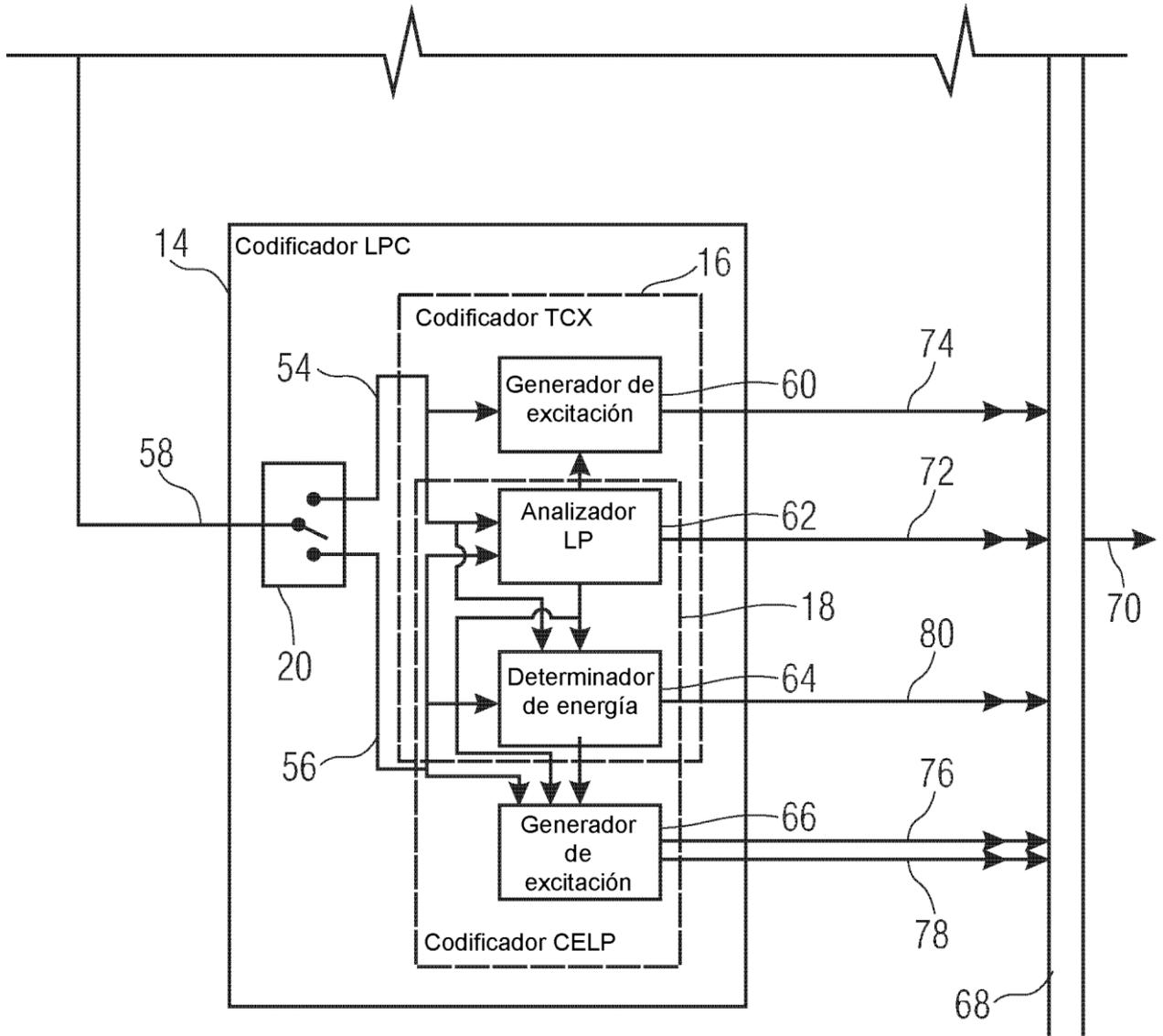


FIG 1 A

FIG 1A	FIG 1
FIG 1B	1



**FIG. 1 B**

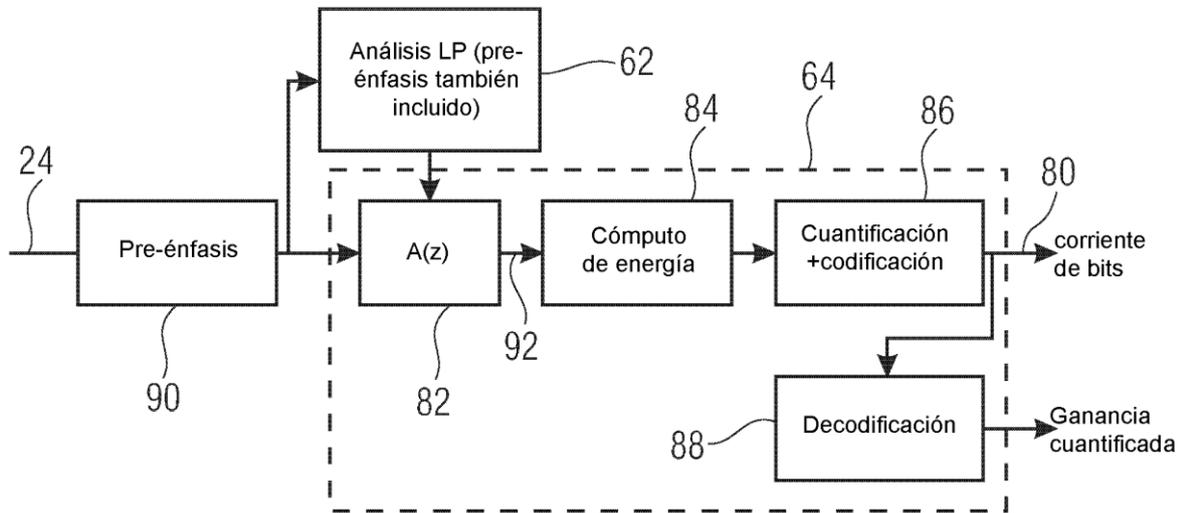


FIG. 2

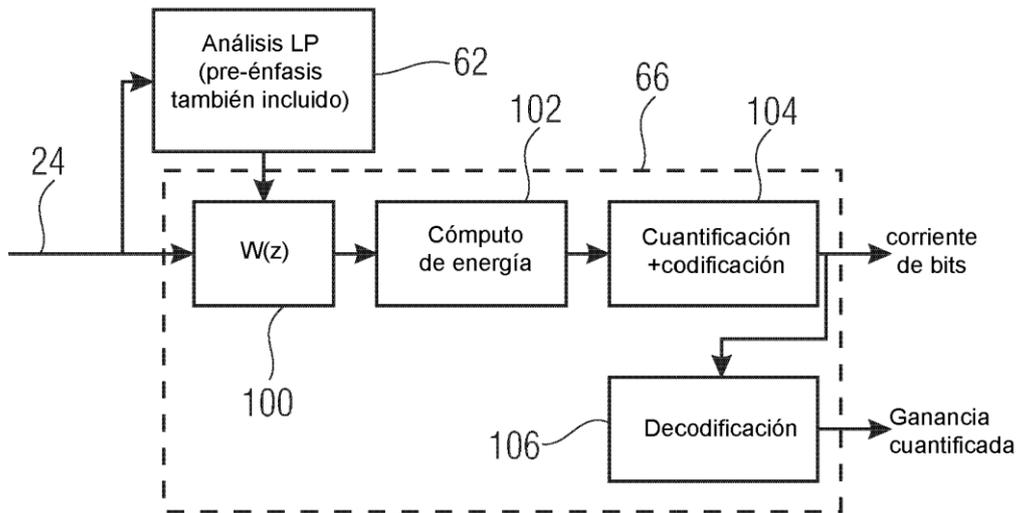


FIG. 3

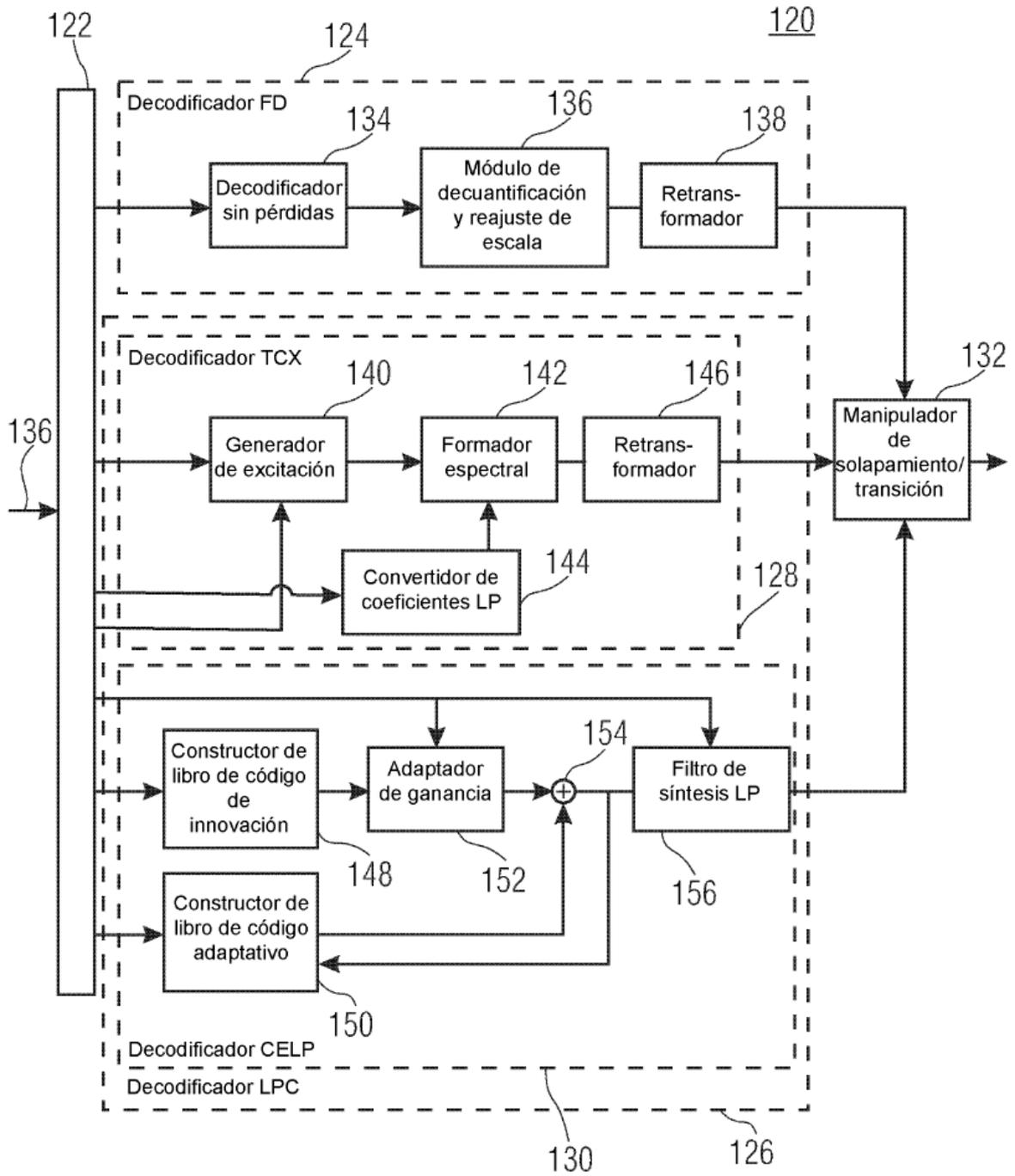
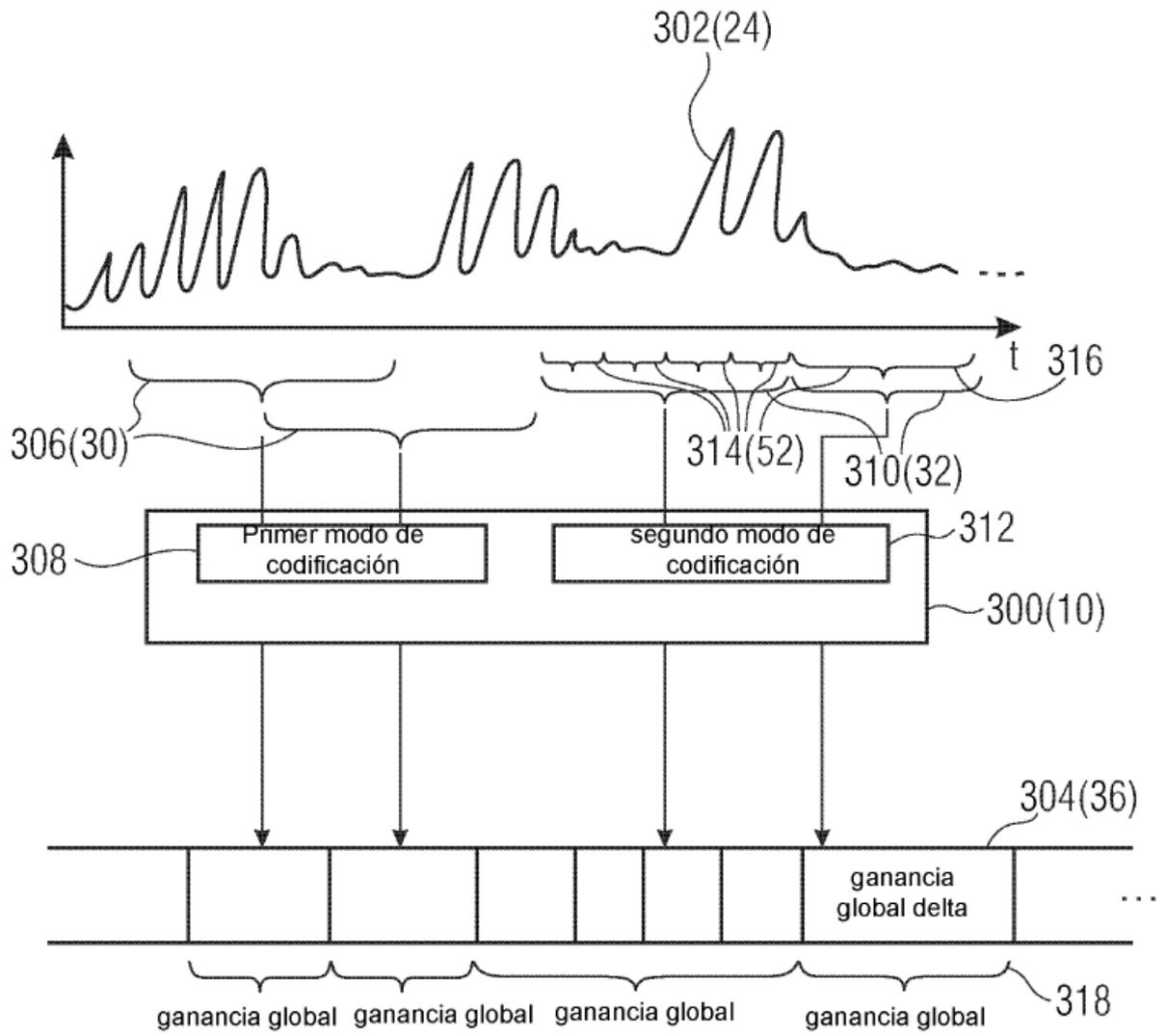
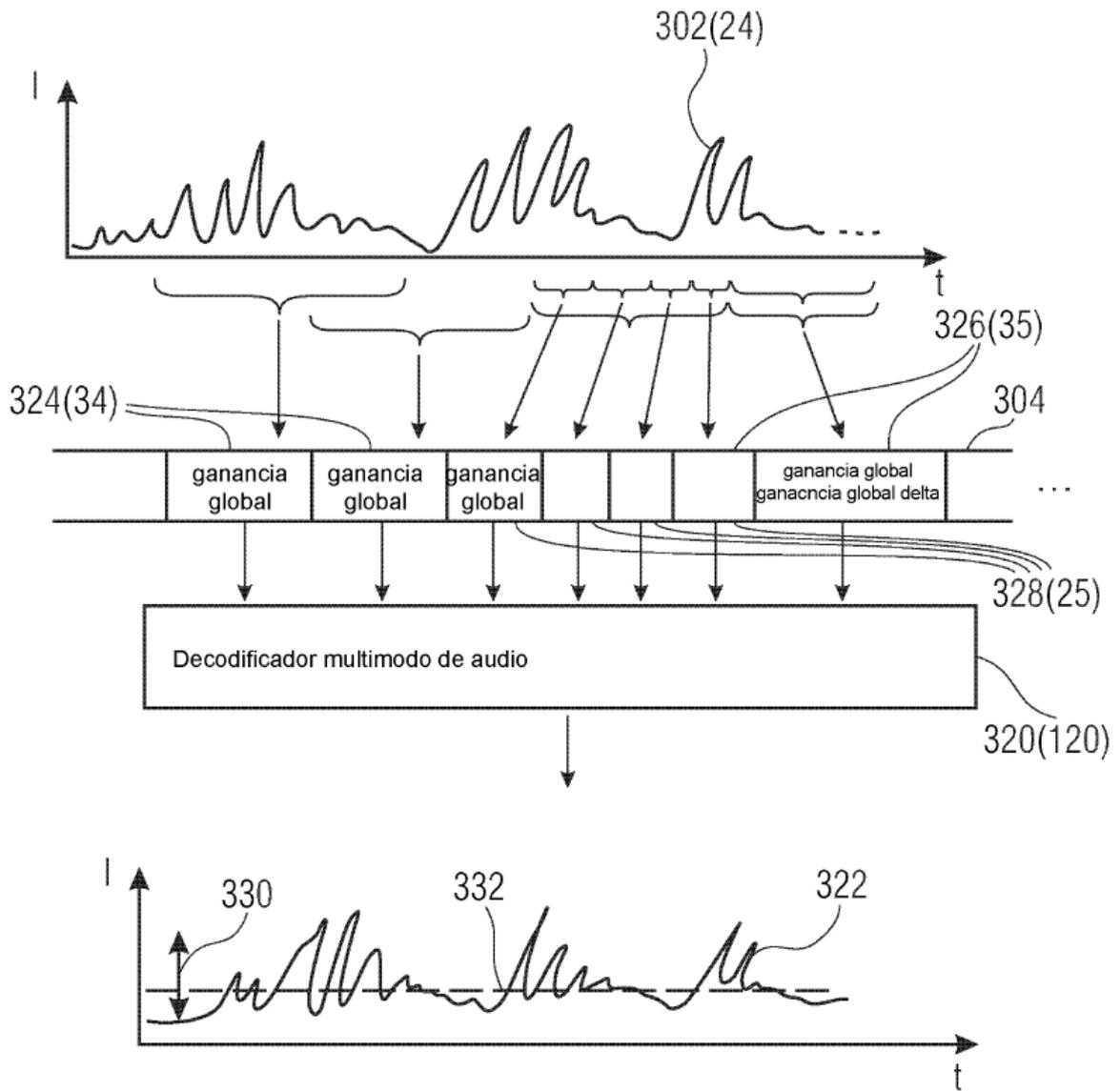


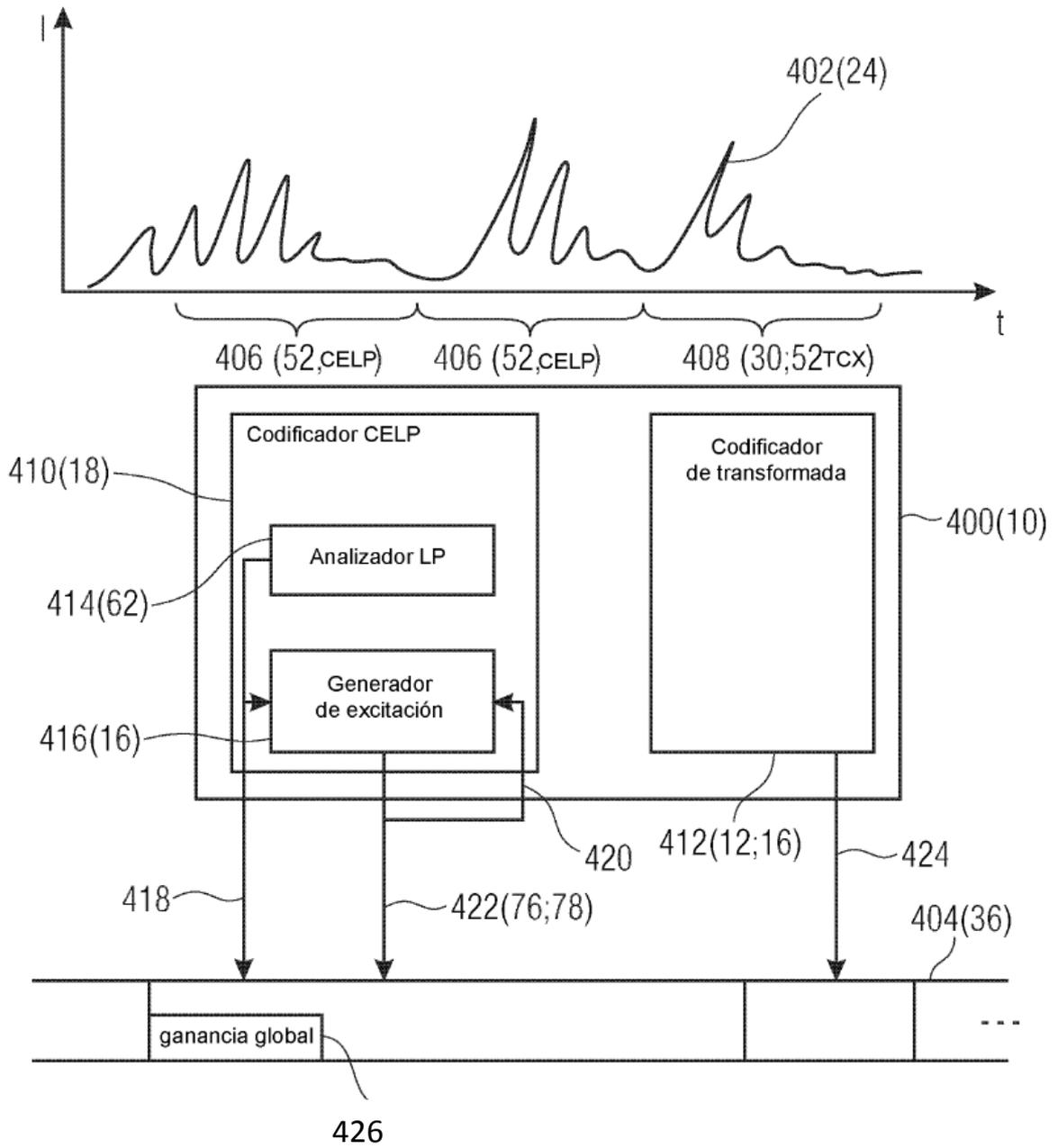
FIG. 4



**FIG. 5 A**



**FIG. 5 B**



**FIG. 6 A**

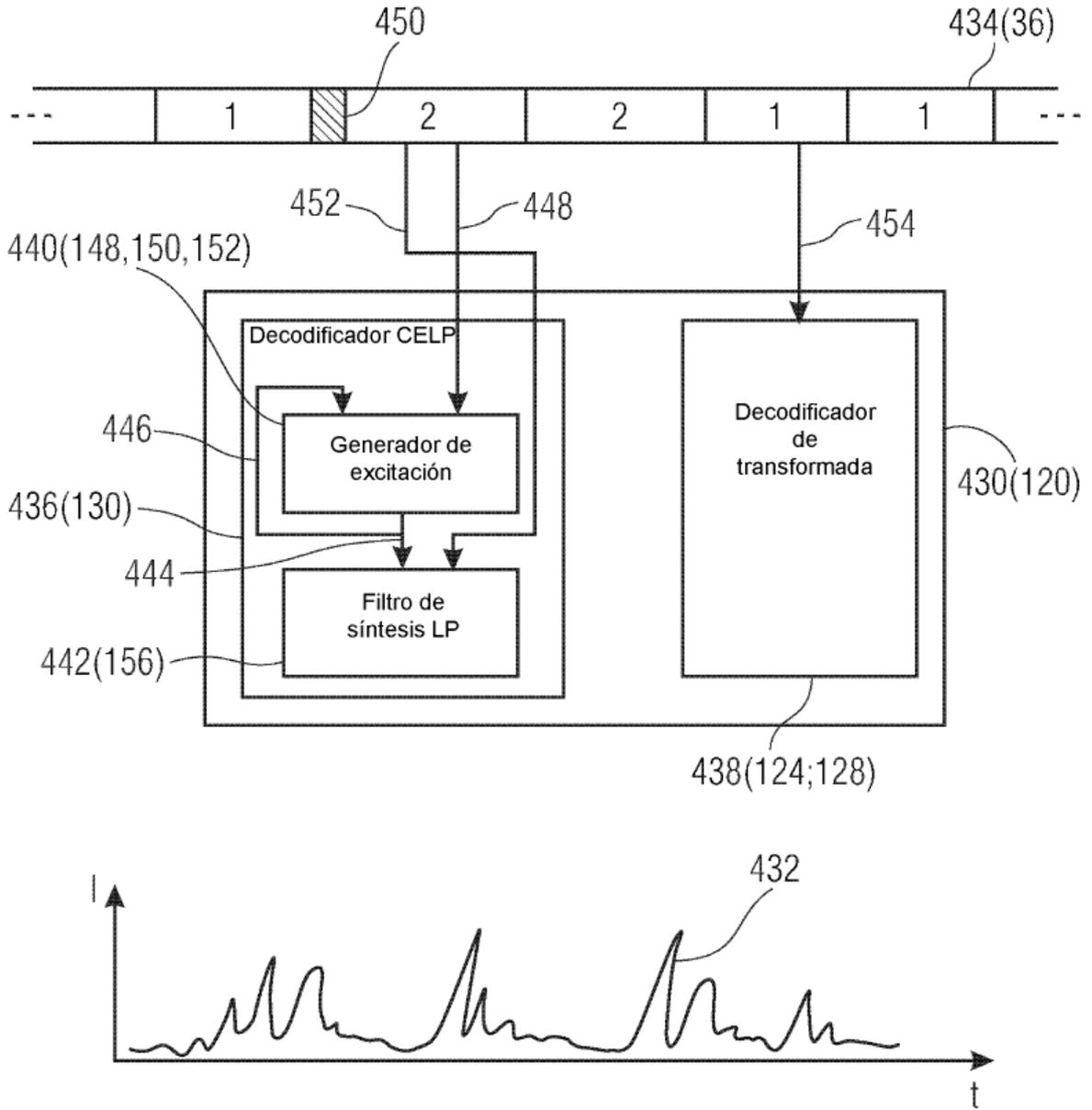


FIG. 6 B

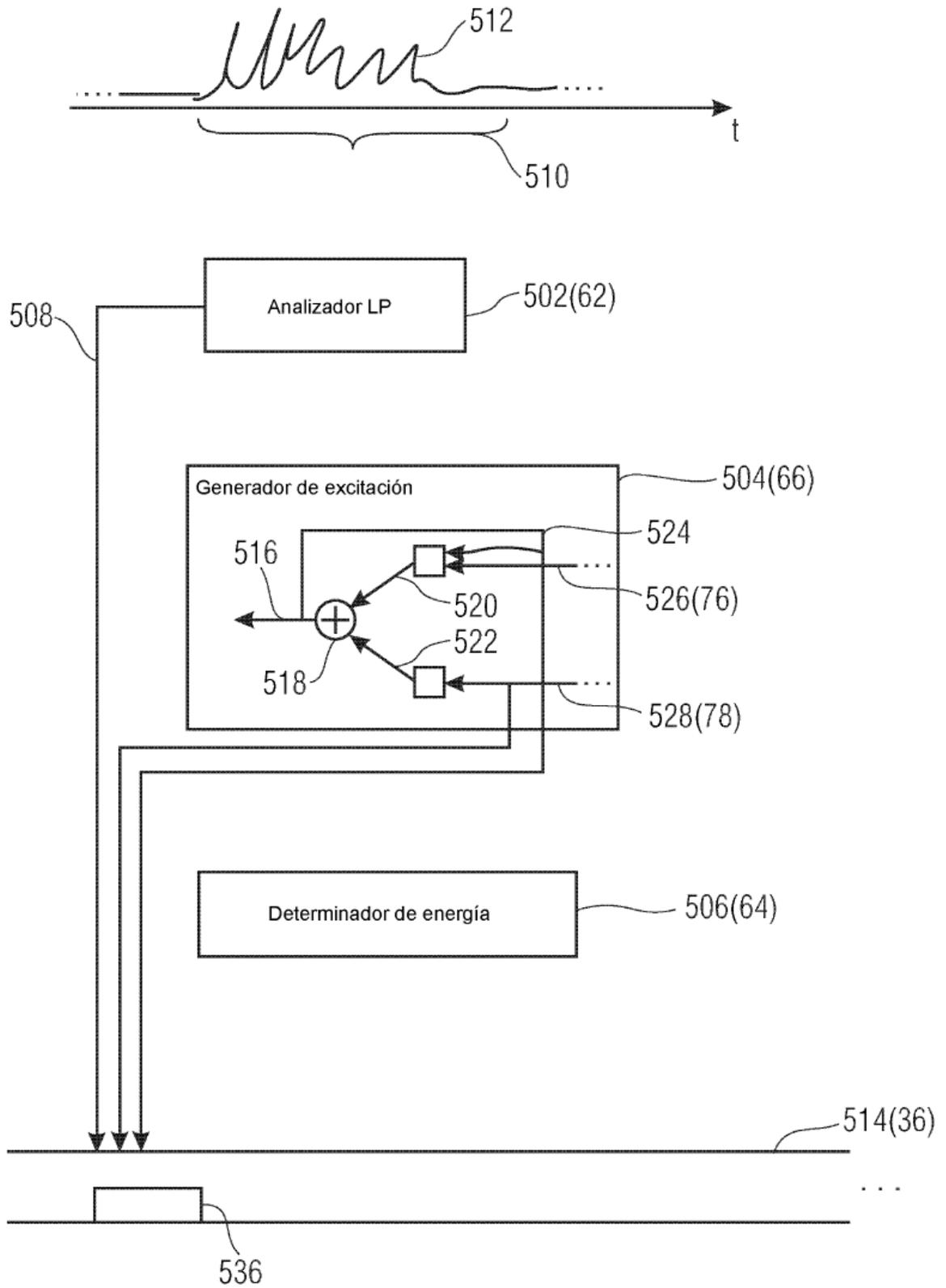
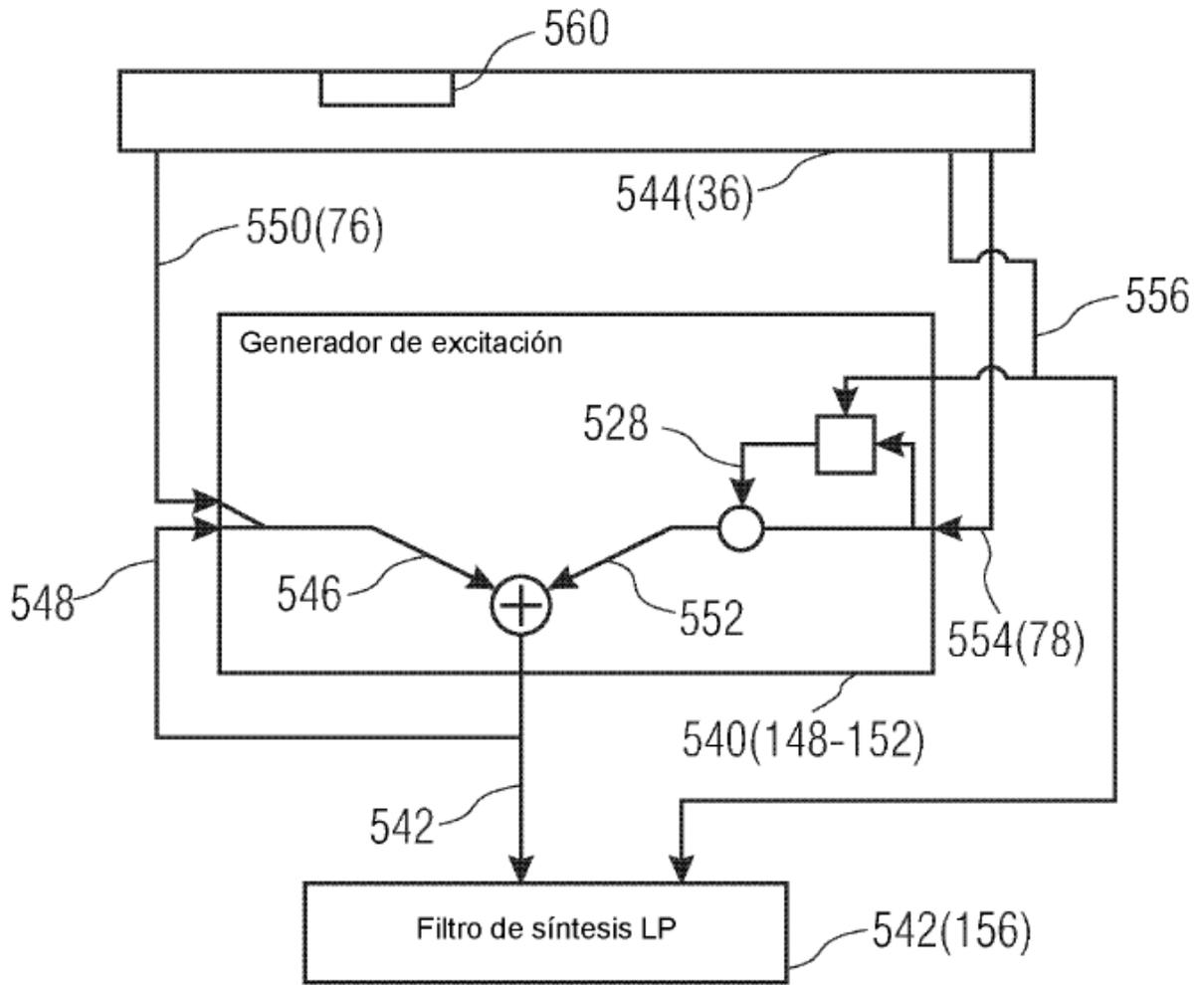


FIG. 7 A



**FIG. 7 B**