

19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 733 207**

51 Int. Cl.:

G10L 19/18 (2013.01)
G10L 19/02 (2013.01)
G10L 19/028 (2013.01)
G10L 19/04 (2013.01)
G10L 19/24 (2013.01)
G10L 21/038 (2013.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

86 Fecha de presentación y número de la solicitud internacional: **24.07.2015** **PCT/EP2015/067003**

87 Fecha y número de publicación internacional: **04.02.2016** **WO16016123**

96 Fecha de presentación y número de la solicitud europea: **24.07.2015** **E 15739300 (0)**

97 Fecha y número de publicación de la concesión europea: **24.04.2019** **EP 3186809**

54 Título: **Codificación y decodificación de audio en los dominios de la frecuencia y del tiempo**

30 Prioridad:

28.07.2014 EP 14178817

45 Fecha de publicación y mención en BOPI de la traducción de la patente:
28.11.2019

73 Titular/es:

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

72 Inventor/es:

**DISCH, SASCHA;
DIETZ, MARTIN;
MULTRUS, MARKUS;
FUCHS, GUILLAUME;
RAVELLI, EMMANUEL;
NEUSINGER, MATTHIAS;
SCHNELL, MARKUS;
SCHUBERT, BENJAMIN y
GRILL, BERNHARD**

74 Agente/Representante:

SALVÀ FERRER, Joan

ES 2 733 207 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

Codificación y decodificación de audio en los dominios de la frecuencia y del tiempo

5 **[0001]** La presente invención se refiere a la codificación y decodificación de audio, y en particular al procesamiento de señales de audio usando procesadores de codificación/decodificación en el dominio de la frecuencia y en el dominio del tiempo en paralelo.

[0002] La codificación perceptual de las señales de audio para el fin de la reducción de datos para un
10 almacenamiento o transmisión eficiente de estas señales es una práctica usada ampliamente. En particular, cuando deban lograrse velocidades binarias lo más bajas posible, la codificación empleada conduce a una reducción de la calidad del audio que a menudo es causada principalmente por la limitación, en el lado del codificador, de la anchura de banda de la señal de audio que se transmitirá. En este caso, normalmente la señal de audio es filtrada en paso bajo de manera que el contenido de forma de onda espectral no se mantenga por encima de una cierta frecuencia de
15 corte predeterminada.

[0003] En los códecs contemporáneos existen procedimientos bien conocidos para la restauración de la señal en el lado del decodificador mediante la Extensión de la Anchura de Banda (BWE, Bandwidth Extension) de la señal de audio, por ejemplo, la Replicación de Banda Espectral (SBR, Spectral Band Replication) que opera en el dominio
20 de la frecuencia o la denominada Extensión de la Anchura de Banda en el Dominio del Tiempo (TD-BWE, Time Domain Bandwidth Extension) que es un posprocesador en los códecs de voz que opera en el dominio del tiempo.

[0004] Adicionalmente existen varios conceptos de codificación en el dominio del tiempo/dominio de la frecuencia combinados tales como los conceptos conocidos por el nombre de AMR-WB+ o USAC.

25 **[0005]** Todos estos conceptos de codificación en el dominio del tiempo combinados tienen en común que el codificador en el dominio de la frecuencia se basa en tecnologías de extensión de la anchura de banda que incurren en una limitación de banda en la señal de audio introducida y la porción por encima de una frecuencia de cruce, o frecuencia de límite, es codificada con un concepto de codificación de baja resolución y sintetizada en el lado del
30 decodificador. Por ello, dichos conceptos se basan principalmente en una tecnología de preprocesador en el lado del codificador, y en una funcionalidad de posprocesamiento correspondiente en el lado del decodificador.

[0006] Normalmente, el codificador en el dominio del tiempo es seleccionado de manera que se codifiquen señales útiles en el dominio del tiempo tales como señales de voz, y el codificador en el dominio de la frecuencia se
35 selecciona para señales que no son de voz, señales de música, etc. Sin embargo, específicamente para las señales que no son de voz que tienen armónicos destacados en la banda de frecuencia alta, los codificadores en el dominio de la frecuencia de la técnica anterior tienen una precisión reducida por lo que presentan una calidad de audio reducida debido al hecho de que dichos armónicos destacados pueden codificarse sólo paramétricamente por separado o son eliminados por completo en el proceso de codificación/decodificación.

40 **[0007]** Además, existen conceptos en los que la rama de codificación/decodificación en el dominio del tiempo se basa adicionalmente en la extensión de la anchura de banda que también codifica paramétricamente un intervalo de frecuencias superior mientras que un intervalo de frecuencias inferior es normalmente codificado mediante un codificador ACELP o cualquier otro CELP relacionado, por ejemplo un codificador de voz. Esta funcionalidad de
45 extensión de la anchura de banda incrementa la eficiencia de la velocidad binaria, pero por otra parte introduce una mayor inflexibilidad debido al hecho de que las dos ramas de codificación, es decir, la rama de codificación en el dominio de la frecuencia y la rama de codificación en el dominio del tiempo, presentan una limitación de banda debido al procedimiento de extensión de la anchura de banda o al procedimiento de replicación por encima de una determinada frecuencia de cruce que es sustancialmente inferior a la frecuencia máxima incluida en la señal de audio
50 introducida.

[0008] Los asuntos relevantes en el estado de la técnica comprenden:

- SBR como un posprocesador para la decodificación de forma de onda [1–3]
- 55 – Conmutación de núcleo MPEG-D USAC [4]
- MPEG-H 3D IGF [5]

[0009] En los siguientes documentos y patentes se describen procedimientos que se consideran constitutivos de la técnica anterior para la solicitud:

- 60 [1] M. Dietz, L. Liljeryd, K. Kjörning and O. Kunz, "Spectral Band Replication, a novel approach in audio coding", en 112ª Convención AES, Múnich, Alemania, 2002.
- [2] S. Meltzer, R. Bohm and F. Henn, "SBR enhanced audio codecs for digital broadcasting such as "Digital Radio Mondiale" (DRM)", en 112ª Convención AES, Múnich, Alemania, 2002.
- 65 [3] T. Ziegler, A. Ehret, P. Ekstrand and M. Lutzky, "Enhancing mp3 with SBR: Features and Capabilities of the new

mp3PRO Algorithm", en 112ª Convención AES, Múnich, Alemania, 2002.

[4] Norma MPEG-D USAC.

[5] PCT/EP2014/065109.

5 **[0010]** En MPEG-D USAC se describe un codificador de núcleo conmutable. Sin embargo, en USAC, el núcleo limitado en anchura de banda se limita a transmitir siempre una señal filtrada en paso bajo. Por tanto, determinadas señales de música que tienen un contenido destacado en alta frecuencia, por ejemplo barridos de banda completa, sonidos en triángulo, etc., no pueden reproducirse de manera fiel.

10 **[0011]** La publicación "ISO/IEC MPEG-2 Advanced Audio Coding", Bosie M. y col., Journal of the Audio Engineering Society, AES, Nueva York, Volumen 45, nº 10, 1 de octubre de 1997, páginas 789-812 describe herramientas básicas incluidas en la norma MPEG-AAC.

[0012] El documento EP 0653846 A1 describe un aparato y un procedimiento para codificar o decodificar 15 señales. El procedimiento comprende las etapas de convertir la señal acústica de entrada en el dominio de la frecuencia, separando la salida en componentes tonales y otros componentes que codifican los componentes tonales, que codifican los componentes de ruido y que generan una cadena de código a partir de la salida de la etapa de codificación.

20 **[0013]** El documento WO 2015/010948 A1 describe un aparato y un procedimiento para codificar o decodificar una señal de audio con un relleno inteligente de los intervalos en el dominio espectral. Un decodificador comprende un decodificador de audio en el dominio espectral, un decodificador paramétrico, un regenerador de frecuencias y un convertidor de tiempo espectral para convertir una representación decodificada y una representación reconstruida en una representación en el dominio del tiempo.

25 **[0014]** Un objeto de la presente invención es proporcionar un concepto mejorado para codificación de audio.

[0015] Este objeto se consigue mediante un codificador de audio según la reivindicación 1, un decodificador de 30 audio según la reivindicación 10, un procedimiento de codificación de audio según la reivindicación 18, un procedimiento de decodificación de audio según la reivindicación 19 o un programa informático según la reivindicación 20.

[0016] La presente invención se basa en el descubrimiento de que un procesador de codificación/decodificación en el dominio del tiempo puede combinarse con un procesador de codificación/decodificación en el dominio de la 35 frecuencia que tenga una funcionalidad de relleno de intervalos, pero esta funcionalidad de relleno de intervalos espectrales se aplica en la totalidad de la banda de la señal de audio o al menos por encima de una determinada frecuencia de relleno de intervalos. Es importante observar que el procesador de codificación/decodificación en el dominio de la frecuencia es particularmente adecuado para realizar una codificación/decodificación precisa de forma de onda o de valor espectral hasta la máxima frecuencia y no solamente hasta una frecuencia de cruce. Además, la 40 capacidad de banda completa del codificador en el dominio de la frecuencia para codificar con la alta resolución permite una integración de la funcionalidad de relleno de intervalos en el codificador en el dominio de la frecuencia. Por lo tanto, según la presente invención, mediante el uso de un procesador codificador/decodificador espectral de banda completa, los problemas relacionados con la separación de la extensión de la anchura de banda por una parte y la codificación de núcleo por otra parte pueden ser abordados y superados mediante la realización de la extensión de la 45 anchura de banda en el mismo dominio espectral en el que opera el decodificador de núcleo. Por lo tanto, se proporciona un decodificador de núcleo de velocidad completa que codifica y decodifica el intervalo de señales de audio completo. Esto no requiere la necesidad de un submuestreador en el lado del codificador y de un sobremuestreador en el lado del decodificador. En cambio, la totalidad del procesamiento es llevada a cabo con la velocidad de muestreo completa o con el dominio completo de la anchura de banda. Para obtener una alta ganancia 50 de codificación, la señal de audio es analizada con el fin de encontrar un primer conjunto de primeras porciones espectrales que ha de ser codificado con una alta resolución, en el que este primer conjunto de primeras porciones espectrales puede incluir, en una realización, porciones tonales de la señal de audio. Por otra parte, los componentes no tonales o ruidosos en la señal de audio que constituyen un segundo conjunto de segundas porciones espectrales son codificados paramétricamente con una baja resolución espectral. En tal caso, la señal de audio codificada 55 solamente requiere que el primer conjunto de primeras porciones espectrales sea codificado de manera que se conserve la forma de onda con una alta resolución espectral, y, adicionalmente, que el segundo conjunto de segundas porciones espectrales sea codificado paramétricamente con una baja frecuencia usando "teselas" de frecuencia obtenidas del primer conjunto como fuente. En el lado del decodificador, el decodificador de núcleo, que es un decodificador de banda completa, reconstruye el primer conjunto de primeras porciones espectrales de manera que 60 se conserve la forma de onda, es decir, sin ningún conocimiento de que exista alguna regeneración adicional de la frecuencia. Sin embargo, el espectro así generado tiene muchos intervalos espectrales. Estos intervalos son rellenados posteriormente con la tecnología de la invención IGF (Intelligent Gap Filling, relleno de intervalos inteligente) mediante el uso de una regeneración de frecuencias para lo cual se usan datos paramétricos por una parte y se usa un intervalo espectral fuente, es decir, primeras porciones espectrales reconstruidas por el decodificador de audio de 65 velocidad completa por otra parte.

[0017] En otras realizaciones, las porciones espectrales, que son reconstruidas por relleno de ruido solamente en lugar de replicación de anchura de banda o relleno de intervalos de frecuencia, constituyen un tercer conjunto de terceras porciones espectrales. Debido al hecho de que el concepto de la codificación opera en un único dominio para la codificación/decodificación por una parte y la regeneración de las frecuencias por otra parte, el IGF no se limita sólo a rellenar un intervalo de frecuencias más altas sino que puede rellenar intervalos de frecuencias más bajas, ya sea mediante el relleno de ruidos sin regeneración de las frecuencias o mediante la regeneración de frecuencias usando una tesela de frecuencias con un intervalo de frecuencias diferente.

[0018] Además, se resalta que una información sobre energías espectrales, una información sobre energías individuales o una información sobre energía individual, una información sobre una energía de supervivencia o una información de energía de supervivencia, una información sobre la energía de tesela o una información de energía de tesela, o una información sobre una energía ausente o una información de energía ausente, puede comprender no solamente un valor de energía, sino también un valor de amplitud (por ejemplo, absoluto), un valor de nivel o cualquier otro valor, a partir del cual es posible derivar un valor final de la energía. Por ello, la información acerca de una energía puede comprender por ejemplo el valor de la energía como tal, y/o un valor de un nivel y/o de una amplitud y/o de una amplitud absoluta.

[0019] Otro aspecto se basa en el descubrimiento de que la situación de correlación no solamente es importante para el intervalo fuente sino que también es importante para el intervalo objeto. Además, la presente invención reconoce la situación en que pueden presentarse diferentes situaciones de correlación en el intervalo fuente y en el intervalo objeto. Si, por ejemplo, se considera una señal de voz con un ruido de alta frecuencia, puede darse la situación de que la banda de baja frecuencia que comprende la señal de voz con un bajo número de sobretonos está estrechamente correlacionada en el canal izquierdo y en el canal derecho, cuando el altavoz se coloca en el medio. Sin embargo, la porción de alta frecuencia, puede estar fuertemente no correlacionada debido al hecho de que puede haber un ruido de alta frecuencia diferente en el lado izquierdo en comparación con otro ruido de alta frecuencia o de un ruido sin alta frecuencia en el lado derecho. Por lo tanto, cuando deba llevarse a cabo una operación de relleno directa que ignore esta situación, entonces la porción de alta frecuencia también estaría correlacionada, y esto podría generar serios artefactos de segregación espacial en la señal reconstruida. Con el fin de abordar este problema, se calculan los datos paramétricos para una banda de reconstrucción, o en términos generales, para el segundo conjunto de segundas porciones espectrales que han de ser reconstruidas, usando un primer conjunto de primeras porciones espectrales con el fin de identificar bien una primera representación de dos canales o bien una segunda representación de canales, diferente para la banda de reconstrucción. En el lado del codificador, se calcula por lo tanto una identificación de dos canales para las segundas porciones espectrales, es decir, para las porciones para las que adicionalmente se calcula información de energía adicional para las bandas de reconstrucción. Un regenerador de frecuencias en el lado del decodificador regenera seguidamente una segunda porción espectral dependiendo de una primera porción del primer conjunto de porciones espectrales, es decir, el intervalo fuente y los datos paramétricos para la segunda porción tales como información de energía de envolvente espectral o cualquier otro dato de envolvente espectral y, adicionalmente, dependiendo de la identificación de los dos canales para la segunda porción, es decir, para esta banda de reconstrucción en consideración.

[0020] La identificación de dos canales se transmite preferentemente como un indicador para cada banda de reconstrucción, y estos datos son transmitidos desde un codificador a un decodificador, y el decodificador seguidamente decodifica la señal de núcleo como se indica mediante indicadores calculados preferentemente para la banda de núcleo. A continuación, en una implementación, la señal de núcleo es almacenada en las dos representaciones estéreo (por ejemplo, izquierdo/derecho o centro/lateral) y, para el relleno de teselas de frecuencia por IGF, se elige la representación de tesela fuente para que concuerde con la representación de tesela objeto como se indica mediante los indicadores de identificación de los dos canales para el relleno inteligente de los intervalos o de las bandas de reconstrucción, es decir, para el intervalo objeto.

[0021] Se resalta que este procedimiento no solamente funciona para señales estéreo, es decir, para un canal izquierdo y un canal derecho, sino que también funciona para señales de múltiples canales. En el caso de las señales de múltiples canales, es posible procesar varios pares de diferentes canales de esta manera tal como un canal izquierdo y un canal derecho como primer par, un canal ambiente izquierdo y un canal ambiente derecho como segundo par y un canal de centro y un canal LFE como tercer par. Es posible determinar otros emparejamientos para formatos de canales de salida más altos tales como 7.1, 11.1, etc.

[0022] Otro aspecto más está basado en el descubrimiento de que la calidad de audio de la señal reconstruida puede ser mejorada mediante IGF ya que la totalidad del espectro es accesible para el codificador de núcleo de manera que, por ejemplo, las porciones tonales importantes desde el punto de vista perceptual en un intervalo espectral alto todavía pueden ser codificadas por el codificador de núcleo en lugar de sustitución paramétrica. Adicionalmente se lleva a cabo una operación de relleno de intervalos usando teselas de frecuencia de un primer conjunto de primeras porciones espectrales que es, por ejemplo, un conjunto de porciones tonales normalmente de un intervalo de menor frecuencia, pero también de un intervalo de frecuencias más alto, si estuviera disponible. Sin embargo, para el ajuste de la envolvente espectral en el lado del decodificador, las porciones espectrales del primer conjunto de porciones

espectrales situadas en la banda de reconstrucción no son posprocesadas ulteriormente, por ejemplo, por el ajuste de la envolvente espectral. Solamente los valores espectrales remanentes en la banda de reconstrucción que no se originan en el codificador de núcleo han de ser ajustados por envolvente usando información de envolvente. Preferentemente la información de envolvente es información de envolvente de banda completa que tiene en cuenta la energía del primer conjunto de primeras porciones espectrales en la banda de reconstrucción y el segundo conjunto de segundas porciones espectrales en la misma banda de reconstrucción, en el que los valores espectrales mencionados en último término en el segundo conjunto de porciones espectrales están indicados para ser cero, y por lo tanto, no son codificados por el codificador de núcleo, sino que son codificados paramétricamente con información de energía de baja resolución.

[0023] Se ha descubierto que los valores de energía absoluta, ya sean normalizados con respecto a la anchura de banda de la banda correspondiente o no normalizados, son útiles y eficientes en una aplicación en el lado del decodificador. Esto se aplica en especial cuando deben calcularse los factores de ganancia basándose en una energía residual en la banda de reconstrucción, la energía ausente en la banda de reconstrucción y la información de tesela de frecuencia en la banda de reconstrucción.

[0024] Además, se prefiere que el tren de bits codificado no solamente cubra información de energía para la banda de reconstrucción, sino adicionalmente factores de escala para bandas de factores de escala que se extienden hasta la máxima frecuencia. Esto asegura que para cada banda de reconstrucción, para la que se encuentra disponible una determinada porción tonal, es decir, una primera porción espectral, este primer conjunto de primeras porciones espectrales puede ser realmente decodificado con la amplitud correcta. Por otra parte, además del factor de escala para cada banda de reconstrucción, se genera una energía para esta banda de reconstrucción en un codificador y se transmite a un decodificador. Además, se prefiere que las bandas de reconstrucción coincidan con las bandas de los factores de escala o, en el caso de una agrupación de energía, que al menos los límites de una banda de reconstrucción coincidan con los límites de las bandas de factores de escala.

[0025] Un aspecto adicional se basa en el descubrimiento de que ciertas alteraciones en la calidad de audio se pueden remediar mediante la aplicación de un esquema de relleno de teselas de frecuencia adaptativa de señal. Para este fin, se realiza un análisis en el lado del codificador con el fin de encontrar la mejor región candidata fuente coincidente para una cierta región objeto. Se genera una información coincidente para identificar una región objeto de una región determinada fuente junto con opcionalmente alguna información adicional y se transmite como información lateral al decodificador. A continuación, el decodificador aplica una operación de relleno de teselas de frecuencia usando la información coincidente. Para este fin, el decodificador lee la información coincidente del tren de datos transmitidos o archivo de datos y accede a la región de la fuente identificada para una determinada banda de reconstrucción y, si está indicado en la información coincidente, además, realiza algún procesamiento de estos datos de la región fuente para generar datos espectrales en bruto para la banda de reconstrucción. A continuación, este resultado de la operación de relleno de teselas de frecuencia, es decir, los datos espectrales en bruto para la banda de reconstrucción, se conforma usando la información de envolvente espectral con el fin de obtener finalmente una banda de reconstrucción que comprende también las primeras porciones espectrales tales como las porciones tonales. Sin embargo, estas porciones tonales no se generan por el esquema de relleno de teselas adaptativo, sino que estas primeras porciones espectrales son producidas por el decodificador de audio o decodificador de núcleo directamente.

[0026] El esquema de selección de teselas espectrales adaptativas puede funcionar con una granularidad baja. En esta implementación, una región fuente se subdivide en regiones fuente y la región objeto normalmente superpuestas o las bandas de reconstrucción se proporcionan en las regiones objeto con frecuencia no superpuestas. Entonces, las semejanzas entre cada región fuente y cada región objeto se determinan en el lado del codificador y el mejor par coincidente de una región fuente y una región objeto se identifican mediante la información coincidente y, en el lado del decodificador, la región fuente identificada en la información coincidente se usa para la generación de los datos espectrales en bruto para la banda de reconstrucción.

[0027] Para el propósito de obtener una granularidad más alta, a cada región fuente se le permite cambiar con el fin de obtener un cierto retardo en el que las semejanzas son máximas. Este retardo puede ser tan fino como un segmento de frecuencia y permite una mejor adaptación entre una región fuente y la región objeto.

[0028] Por otra parte, además de identificar solamente un mejor par coincidente, este retardo de correlación también se puede transmitir dentro de la información coincidente y, adicionalmente, se puede transmitir incluso una señal. Cuando se determina que la señal es negativa en el lado del codificador, a continuación se transmite también un indicador del signo correspondiente dentro de la información coincidente y, en el lado del decodificador, los valores espectrales de la región fuente se multiplican por -1 o, en una representación compleja, se "hacen girar" 180 grados.

[0029] Se aplica otra implementación de esta invención para la operación de blanqueo de teselas. El blanqueo de un espectro elimina la información de envolvente espectral en bruto y resalta en la estructura fina espectral que es de sumo interés para evaluar la similitud de las teselas. Por tanto, una tesela de frecuencias por una parte y/o la señal fuente por otra parte son blanqueadas antes de calcular una medida de correlación cruzada. Cuando se blanquee solamente la tesela usando un procedimiento predefinido, se transmite un indicador de blanqueo que indica al

decodificador que el mismo proceso de blanqueo predefinido será aplicado a la tesela de frecuencia dentro del IGF.

[0030] En cuanto a la selección de las teselas, se prefiere usar el retardo de la correlación para desplazar espectralmente el espectro regenerado en un número entero de segmentos de transformada. En función de la transformada subyacente, el desplazamiento espectral puede requerir correcciones de suma. En el caso de los retardos impares, la tesela es modulada adicionalmente mediante multiplicación por una secuencia temporal alternante de $-1/1$ para compensar la representación de frecuencia inversa de cualquier otra banda dentro del MDCT. Además, se aplica el signo del resultado de correlación cuando se genera la tesela de frecuencia.

[0031] Además, se prefiere usar el podado y estabilización de las teselas con el fin de asegurar que se eviten los artefactos creados por las regiones fuente de rápido cambio para la misma región de reconstrucción o región objeto. Con este fin, se lleva a cabo un análisis de similitud entre las diferentes regiones fuente identificadas, y cuando una tesela fuente es similar a otras teselas fuente con una similitud superior a un umbral, entonces esta tesela fuente puede ser descartada del conjunto de teselas fuente potenciales por cuanto que está estrechamente correlacionada con otras teselas fuente. Además, como un tipo de estabilización de la selección, se prefiere mantener el orden de las teselas de la trama precedente si ninguna de las teselas fuente en la trama actual se correlaciona (mejor que en un umbral predeterminado) con las teselas objeto en la trama actual.

[0032] Otro aspecto se basa en el descubrimiento de que se obtiene una calidad mejorada y una velocidad binaria reducida específicamente para señales que comprenden porciones transitorias ya que se presentan muy frecuentemente en las señales de audio mediante la combinación de la tecnología de TSN (conformado temporal del ruido) o de TTS (conformado temporal de teselas) con una reconstrucción de alta frecuencia. El procesamiento TNS/TTS en el lado del codificador implementado mediante una predicción sobre las frecuencias reconstruye la envolvente temporal de la señal de audio. Dependiendo de la implementación, es decir, cuando se determina que el filtro de conformado de ruido temporal dentro de un intervalo de frecuencias no solamente recubre el intervalo de frecuencias fuente sino también el intervalo de frecuencias objeto para su reconstrucción en un decodificador de regeneración de frecuencias, la envolvente temporal no solamente se aplica a la señal de audio de núcleo hasta una frecuencia de inicio de relleno de intervalos, sino que la envolvente temporal también se aplica a los intervalos espectrales de segundas porciones espectrales reconstruidas. De esta manera se reducen o eliminan los pre-ecos o post-ecos que se presentarían sin el conformado temporal de las teselas. Esto se lleva a cabo aplicando una predicción inversa sobre las frecuencias no solamente dentro del intervalo de frecuencias de núcleo hasta una determinada frecuencia de inicio de relleno de intervalos, sino también dentro del intervalo de frecuencias superior al intervalo de frecuencias de núcleo. Con este fin, la regeneración de frecuencias o regeneración de teselas de frecuencia se lleva a cabo en el lado del decodificador antes de aplicarse una predicción sobre las frecuencias. Sin embargo, la predicción en cuanto a las frecuencias puede ser aplicada antes o después del conformado de la envolvente espectral dependiendo de si el cálculo de la información de la energía ha sido llevado a cabo sobre los valores espectrales residuales posteriormente al filtrado o a los valores espectrales (completos) antes del conformado de la envolvente.

[0033] El procesamiento TTS sobre una o más teselas de frecuencias establece una continuidad de la correlación entre el intervalo fuente y el intervalo de reconstrucción o en dos intervalos de reconstrucción adyacentes o teselas.

[0034] En una implementación, se prefiere usar el filtrado complejo TSN/TTS. De esta manera, se evitan los artefactos de superposición (temporales) de una representación real normalmente muestreada, tal como MDCT. Es posible calcular un filtro TNS complejo en el lado del codificador mediante la aplicación de no solamente una transformada de coseno discreta modificada sino también de una transformada de seno discreta modificada adicional con el fin de obtener una transformada modificada compleja. Sin embargo, solamente se transmiten los valores de transformada de coseno discreta modificada, es decir, la parte real de la transformada compleja. Sin embargo, en el lado del decodificador, es posible estimar la parte imaginaria de la transformada usando los espectros de MDCT de tramas precedentes o posteriores de manera que, en el lado del decodificador, es posible aplicar el filtro complejo nuevamente en la predicción inversa sobre las frecuencias, y específicamente, la predicción sobre el límite entre el intervalo fuente y el intervalo de reconstrucción y también sobre el límite entre las teselas de frecuencia de frecuencias adyacentes dentro del intervalo de reconstrucción.

[0035] El sistema de codificación de audio de la invención codifica de manera eficiente señales de audio arbitrarias en un amplio intervalo de velocidad binaria. Mientras, para velocidades binarias altas, el sistema de la invención converge hasta su transparencia, y para las velocidades binarias más bajas, se minimiza la molestia perceptiva. Por tanto, la parte principal de la velocidad binaria disponible se usa para codificar la estructura perceptualmente más relevante de la señal en el codificador, y los intervalos espectrales resultantes son rellenados en el decodificador con un contenido de señales que se aproxima al espectro original. Se consume un presupuesto muy limitado de bits para controlar el relleno inteligente de intervalos (IFG, Intelligent Gap Filling) mediante información secundaria dedicada transmitida desde el codificador al decodificador.

[0036] En otras realizaciones, el procesador de codificación/decodificación en el dominio del tiempo se basa en una menor velocidad de muestreo y en una funcionalidad de extensión de banda correspondiente.

[0037] En otras realizaciones, se proporciona un procesador cruzado para inicializar el codificador/decodificador en el dominio del tiempo con datos de inicialización deducidos de la señal de codificador/decodificador en el dominio de la frecuencia actualmente procesado. Esto permite que cuando se procesa la porción de señales de audio actualmente procesada por medio del codificador en el dominio de la frecuencia, se inicializa el codificador en el dominio del tiempo en paralelo, de manera que cuando tiene lugar una conmutación desde el codificador en el dominio de la frecuencia a un codificador en el dominio del tiempo, este codificador en el dominio del tiempo puede empezar el procesamiento ya que todos los datos de inicialización relacionados con señales anteriores ya están preparados debidos al procesador. Este procesador cruzado se aplica preferentemente en el lado del codificador, y adicionalmente en el lado del decodificador y usa preferentemente una transformada de frecuencia-tiempo que adicionalmente lleva a cabo un submuestreo muy eficiente desde la velocidad de muestreo de entrada o salida más alta a la velocidad de muestreo de núcleo correspondiente en el dominio del tiempo más baja mediante la simple selección de una porción de banda baja determinada de la señal del dominio junto con un tamaño de transformada reducido determinado. De esta manera se lleva a cabo de forma muy eficiente una conversión desde la alta velocidad de muestreo a la baja velocidad de muestreo, y esta señal obtenida por la transformada con el tamaño de transformada reducido puede usarse a continuación para inicializar el codificador/decodificador en el dominio del tiempo de manera que el codificador/decodificador en el dominio del tiempo esté listo para realizar para realizar inmediatamente la codificación en el dominio del tiempo cuando esta situación es señalizada por un controlador y la porción de la señal de audio inmediatamente precedente ha sido codificada en el dominio de la frecuencia.

[0038] Por ello, las realizaciones preferidas de la presente invención permiten una conmutación ininterrumpida de un codificador de audio porcentual que comprende el relleno de intervalos espectrales y un codificador en el dominio del tiempo con o sin una extensión de anchura de banda.

[0039] Por ello, la presente invención se basa en procedimientos que no se limitan a eliminar el contenido de alta frecuencia por encima de una frecuencia de corte en el codificador en el dominio de la frecuencia a partir de la señal de audio sino que elimina de una manera adaptativa a las señales las regiones de paso de banda espectrales que dejan intervalos espectrales en el codificador y posteriormente reconstruyen estos intervalos espectrales en el decodificador. Preferentemente se usa una solución integrada tal como el relleno inteligente de los intervalos que combina de manera eficiente la codificación de audio de anchura de banda completa con el relleno de intervalos espectrales particularmente en el dominio de las transformadas MDCT.

[0040] Por ello, la presente invención proporciona un concepto mejorado para combinar la codificación de voz y una posterior extensión de la anchura de banda en el dominio del tiempo con una decodificación de forma de onda de banda completa que comprende el relleno de intervalos espectrales en un codificador/decodificador perceptual conmutable.

[0041] Por ello, y a diferencia de procedimientos ya existentes, el nuevo concepto usa la codificación de forma de onda de señales de audio de banda completa en el codificador en el dominio de las transformadas y al mismo tiempo permite una conmutación ininterrumpida a un codificador de voz seguido preferentemente por una extensión de la anchura de banda en el dominio del tiempo.

[0042] Otras realizaciones de la presente invención evitan los problemas explicados que se presenten debido a una limitación de banda fija. El concepto permite la combinación conmutable de un codificador de forma de onda de banda completa en el dominio de la frecuencia equipado con un relleno de intervalos espectrales y un codificador de voz con una velocidad de muestreo más baja y una extensión de la anchura de banda en el dominio del tiempo. Un codificador de este tipo es apto para la codificación de formas de onda de las señales problemáticas anteriormente mencionadas que proporciona una anchura de banda completa hasta la frecuencia de Nyquist de la señal introducida de audio. Sin embargo, la conmutación ininterrumpida entre las dos estrategias de codificación está asegurada en particular por las realizaciones que tienen el procesador cruzado. Para esta conmutación ininterrumpida, el procesador cruzado representa una conexión cruzada tanto en el codificador como en el decodificador entre el codificador en el dominio de la frecuencia de velocidad completa capaz de banda completa (velocidad de muestreo de entrada) y el codificador ACELP de baja velocidad que tiene una velocidad de muestreo más baja con el fin de inicializar adecuadamente los parámetros ACELP y las memorias intermedias particularmente dentro del libro de código adaptativo, el filtro de LPC o la fase de remuestreo, cuando se conmute desde el codificador en el dominio de la frecuencia tal como TCX al codificador en el dominio del tiempo tal como ACELP.

[0043] A continuación la presente invención se describe con respecto a los dibujos adjuntos en los que:

- la Fig. 1a ilustra un aparato para codificar una señal de audio;
- la Fig. 1b ilustra un decodificador para decodificar una señal de audio codificada que se corresponde con el decodificador de la Fig. 1a;
- la Fig. 2a ilustra una implementación preferida del decodificador;
- la Fig. 2b ilustra una implementación preferida del codificador
- la Fig. 3a ilustra una representación esquemática de un espectro generado por el decodificador en el dominio espectral

de la Fig. 1b;

la Fig. 3b ilustra una tabla que indica la relación entre los factores de escala para las bandas de los factores de escala y la información de relleno de ruido para una banda de relleno de ruido;

la Fig. 4a ilustra la funcionalidad del codificador en el dominio espectral para aplicar la selección de porciones espectrales en los conjuntos primero y segundo de porciones espectrales;

la Fig. 4b ilustra una implementación de la funcionalidad de la Fig. 4a;

la Fig. 5a ilustra una funcionalidad del codificador MDCT;

la Fig. 5b ilustra una funcionalidad del decodificador con una tecnología MDCT;

la Fig. 5c ilustra una implementación del regenerador de frecuencias;

la Fig. 6 ilustra una implementación de un codificador de audio;

la Fig. 7a ilustra un procesador cruzado con el codificador de audio;

la Fig. 7b ilustra una implementación de una transformada de frecuencia–tiempo o inversa que adicionalmente proporciona una reducción de la velocidad de muestreo dentro del procesador cruzado;

la Fig. 8 ilustra una implementación preferida del controlador de la Fig. 6;

la Fig. 9 ilustra una realización adicional del codificador en el dominio del tiempo que tiene funcionalidades de extensión de anchura de banda;

la Fig. 10 ilustra un uso preferido de un preprocesador;

la Fig. 11a ilustra una implementación esquemática del decodificador de audio;

la Fig. 11b ilustra un procesador cruzado dentro del decodificador para proporcionar datos de inicialización para el decodificador en el dominio del tiempo;

la Fig. 12 ilustra una implementación preferida del procesador de decodificación en el dominio del tiempo de la Fig. 11a;

la Fig. 13 ilustra otra implementación de la extensión de la anchura de banda en el dominio del tiempo;

la Fig. 14a ilustra una implementación preferida de un codificador de audio;

la Fig. 14b ilustra una implementación preferida de un decodificador de audio;

la Fig. 14c ilustra una implementación de la invención de un decodificador en el dominio del tiempo con una conversión de la velocidad de muestreo y extensión de anchura de banda.

[0044] La Fig. 6 ilustra un codificador de audio para codificar una señal de audio que comprende un primer procesador de codificación 600 para codificar una primera porción de la señal de audio en un dominio de la frecuencia. El primer procesador de codificación 600 comprende un convertidor de tiempo–frecuencia 602 para convertir la primera porción de la señal de audio en una representación en el dominio de la frecuencia que tiene líneas espectrales hasta una frecuencia máxima de la señal introducida. Por otra parte, el primer procesador de codificación 600 comprende un analizador 604 para analizar la representación en el dominio de la frecuencia hasta la frecuencia máxima con el fin de determinar primeras regiones espectrales que se codificarán con una primera representación espectral y de determinar segundas regiones espectrales que se codificarán con una segunda resolución espectral que es inferior a la primera resolución espectral. En particular, el analizador de banda completa 604 determina qué líneas de frecuencia o valores espectrales en el espectro del convertidor de tiempo–frecuencia han de ser codificadas a modo de espectro y qué otras porciones espectrales han de ser codificadas de una manera paramétrica, y estos últimos valores espectrales son seguidamente reconstruidos en el lado del decodificador mediante el procedimiento de relleno de intervalos. La operación de codificación real es llevada a cabo por un codificador espectral 606 para codificar las primeras regiones espectrales o porciones espectrales con la primera resolución y para codificar paramétricamente las segundas regiones espectrales o sus porciones con la segunda resolución espectral.

[0045] El codificador de audio de la Fig. 6 comprende adicionalmente un segundo procesador de codificación 610 para codificar la porción de la señal de audio en un dominio del tiempo. Adicionalmente, el codificador de audio comprende un controlador 620 configurado para analizar la señal de audio en una señal introducida de audio 601 y para determinar qué porción de la señal de audio es la primera porción de la señal de audio codificada en el dominio de la frecuencia y qué porción de la señal de audio es la segunda porción de la señal de audio codificada en el dominio del tiempo. Por otra parte, se ha proporcionado un conformador de señales codificadas 630 que puede, por ejemplo, estar implementado como un multiplexor del tren de bits que está configurado para formar una señal de audio codificada que comprende una primera porción de la señal codificada para la primera señal de audio y una segunda porción de la señal codificada para la segunda porción de audio. Es importante observar que la señal codificada solamente tiene una representación en el dominio de la frecuencia o una representación en el dominio del tiempo, de una y de la misma porción de la señal de audio.

[0046] Por lo tanto, el controlador 620 asegura que para una única porción de la señal de audio se encuentre solamente una representación en el dominio del tiempo o una representación en el dominio de la frecuencia en la señal codificada. Esto puede lograrse mediante el controlador 620 de diversas maneras. Una manera consistiría en que, para una y la misma porción de la señal de audio, las dos representaciones lleguen al bloque 630 y el controlador 620 controle el conformador de la señal codificada 630 con el fin de introducir solamente una de las dos representaciones en la señal codificada. Sin embargo, como alternativa el controlador 620 puede controlar una entrada en el primer procesador de codificación y una entrada en el segundo procesador de codificación de manera que, basándose en el análisis de la correspondiente porción de la señal, solamente uno de los dos bloques 600 o 610 es activado con el fin de realmente llevar a cabo la operación de codificación completa, y el otro bloque es desactivado.

[0047] Esta desactivación puede ser una desactivación o como se ilustra, por ejemplo con respecto a la Fig. 7a, es solamente un tipo de modo de "inicialización" en el que el otro procesador de codificación está activo solamente para recibir y procesar datos de inicialización con el fin de inicializar las memorias internas pero ninguna operación de codificación específica en absoluto. Esta activación puede efectuarse mediante una determinada conmutación en la entrada que no se ilustra en la Fig. 6, o, preferentemente, mediante las líneas de control 621 y 622. Por lo tanto, en esta realización, el segundo procesador de codificación 610 no emite nada cuando el controlador 620 ha determinado que la porción de la señal de audio vigente debería ser codificada por el primer procesador de codificación pero el segundo procesador de codificación esté provisto sin embargo de datos de inicialización de manera que esté activo para una conmutación instantánea en el futuro. Por otra parte, el primer procesador de codificación está configurado de manera que no necesite ningún dato del pasado para actualizar cualesquiera memorias internas, y por ello, cuando la porción de la señal de audio vigente ha de ser codificada por el segundo procesador de codificación 610, entonces el controlador 620 puede controlar el primer procesador de codificación de fin 600 por medio de la línea de control 621 para que esté completamente inactivo. Esto significa que no es necesario que el primer procesador de codificación 600 se encuentre en un estado de inicialización o estado de espera, sino que puede estar en un estado de desactivación completa. Esto es preferible en particular para dispositivos móviles en los que el consumo de energía eléctrica, y por lo tanto, la vida de la batería es una cuestión que debe tenerse en cuenta.

[0048] En la otra implementación específica del segundo procesador de codificación que opera en el dominio del tiempo, el segundo procesador de codificación comprende un submuestreador 900 o un convertidor de velocidad de muestreo para convertir la porción de la señal de audio en una representación con una velocidad de muestreo más baja, en el que la velocidad de muestreo más baja es inferior a la velocidad de muestreo en la entrada en el primer procesador de codificación. Esto se ilustra en la Fig. 9. En particular, cuando la señal de audio introducida comprende una banda baja y una banda alta, se prefiere que la representación de la velocidad de muestreo inferior en la salida del bloque 900 tenga solamente la banda baja de la porción de la señal de audio introducida y esta banda baja es seguidamente codificada por el codificador de banda baja en el dominio del tiempo 910 que está configurado para la codificación en el dominio del tiempo de la representación de la velocidad de muestreo más baja proporcionada por el bloque 900. Además, se proporciona un codificador de extensión de anchura de banda en el dominio del tiempo 920 para codificar paramétricamente la banda alta. Para este fin, se proporciona el codificador de extensión de anchura de banda en el dominio del tiempo 920 para codificar paramétricamente la señal de audio de banda alta o la banda baja y la banda alta de la señal de audio de entrada.

[0049] En una realización adicional de la presente invención, el codificador de audio comprende adicionalmente, si bien no se ilustra en la Fig. 6 pero sí en la Fig. 10, un preprocesador 1000 conformado para preprocesar la primera porción de la señal de audio y la segunda porción de la señal de audio. En una realización, este preprocesador comprende un analizador de predicción para la determinación de coeficientes de predicción. Este analizador de predicción se puede implementar como un analizador de LPC (codificación de predicción lineal) para determinar los coeficientes LPC. Sin embargo, también se pueden implementar otros analizadores. Además, el preprocesador, que también se ilustra en la Fig. 14a, comprende un cuantificador del coeficiente de predicción 1010, en el que este dispositivo que se ilustra en la Fig. 14a recibe datos de coeficientes de predicción del analizador de predicción que también se ilustran en la Fig. 14a en 1002.

[0050] Además, el preprocesador comprende adicionalmente un codificador de entropía para generar una versión codificada de los coeficientes de predicción cuantificados. Es importante observar que el conformador de señales codificadas 630 o la implementación específica, es decir, el multiplexor del tren de bits 613 asegure que la versión codificada de los coeficientes de predicción cuantificados esté incluida en la señal de audio codificada 632. Preferentemente los coeficientes de LPC no se cuantifican directamente sino que son convertidos en ISF, por ejemplo, o en cualquier otra representación mejor adecuada para la cuantificación. Preferentemente esta conversión es llevada a cabo mediante la determinación del bloque de coeficientes LPC 1002 o bien dentro del bloque 1010 para cuantificar los coeficientes de LPC.

[0051] Además, el preprocesador puede comprender un remuestreador 1004 para remuestrear una señal de audio introducida con una velocidad de muestreo introducida en una velocidad de muestreo más baja para el codificador en el dominio del tiempo. Cuando el codificador en el dominio del tiempo es un codificador ACELP que tiene una determinada velocidad de muestreo ACELP, entonces el submuestreo se lleva a cabo preferentemente a 12,8 kHz o 16 kHz. La velocidad de muestreo de entrada puede ser cualquiera de entre un número particular de velocidad de muestreo tal como 32 Hz o una velocidad de muestreo aún más alta. Por otra parte, la velocidad de muestreo del codificador en el dominio del tiempo estará predeterminada por ciertas restricciones, y el remuestreador 1004 lleva a cabo este remuestreo y emite la representación de la velocidad de muestreo más baja de la señal introducida. Por lo tanto, el remuestreador 1004 puede llevar a cabo una funcionalidad similar y puede aún ser uno y el mismo elemento que el submuestreador 900 ilustrado en el contexto de la Fig. 9.

[0052] Además, se requiere aplicar una preacentuación en el bloque de preacentuación: 1005 en la Fig. 14a. El procesamiento de la preacentuación es bien conocido en la técnica de la codificación en el dominio del tiempo, y se describe en la bibliografía con referencia al procesamiento AMR-WB+ y la preacentuación está configurada

particularmente para compensar una inclinación espectral, y por lo tanto permite un mejor cálculo de los parámetros de LPC con un orden de LPC dado.

[0053] Además, el preprocesador puede comprender adicionalmente una extracción de parámetros de TCX–LTP para controlar el posfiltro de LTP ilustrado en 1420 en la Fig. 14b. Este bloque se ilustra en 1006 en la Fig. 14a. Además, el preprocesador puede comprender adicionalmente otras funcionalidades ilustradas en 1007, y estas otras funcionalidades pueden comprender una funcionalidad de búsqueda de tonos, una funcionalidad de detección de la activación de voz (VAD, voice activity detection) o cualesquiera otras funcionalidades conocidas en la técnica de la codificación en el dominio del tiempo o de la voz.

[0054] Como se ilustra, el resultado del bloque 1006 es introducido en la señal codificada, es decir, en la realización de la Fig. 14a, introducido en el multiplexor del tren de bits 630. Además, si fuera necesario, los datos procedentes del bloque 1007 también pueden ser introducidos en el multiplexor del tren de bits o, como alternativa, pueden usarse con fines de codificación en el dominio del tiempo en el codificador en el dominio del tiempo.

[0055] Por lo tanto, para resumir, es común a las dos trayectorias un preprocesamiento 1000 en el que se llevan a cabo las operaciones de procesamiento de las señales usadas comúnmente. Estas operaciones comprenden un remuestreo con una velocidad de muestreo de ACELP (de 12,8 o 16 kHz) para una trayectoria paralela, y este remuestreo siempre se lleva a cabo. Además, se lleva a cabo una extracción de parámetros de TCX LTP ilustrada en el bloque 1006, y adicionalmente se lleva a cabo una preacentuación y una determinación de los coeficientes de LPC. Tal como se señala a grandes rasgos, la preacentuación compensa la inclinación espectral y por lo tanto hace que el cálculo de los parámetros de LPC sea más eficiente en un orden de LPC dado.

[0056] A continuación, se hace referencia a la Fig. 8 con el fin de ilustrar una implementación preferida del controlador 620. El controlador recibe, en una entrada, la porción de la señal de audio en consideración. Preferentemente, tal como se ilustra en la Fig. 14a, el controlador recibe cualquier señal disponible en el preprocesador 1000 que puede ser la señal introducida original con la velocidad de muestreo introducida o una versión remuestreada con la velocidad de muestreo del codificador en el dominio del tiempo, inferior, o una señal obtenida posteriormente al procesamiento de preacentuación en el bloque 1005.

[0057] Basándose en esta porción de la señal de audio, el controlador 620 se dirige a un simulador del codificador en el dominio de la frecuencia 621 y un simulador del codificador en el dominio del tiempo 622 con el fin de calcular para cada posibilidad de codificador una relación señal-ruido estimada. Posteriormente, el selector 623 selecciona el codificador que haya proporcionado la mejor relación señal-ruido, naturalmente con la consideración de una velocidad binaria predefinida. El selector identifica seguidamente el codificador correspondiente por medio de la salida de control. Cuando se determine que la porción de la señal de audio en consideración debe ser codificada usando el codificador en el dominio de la frecuencia, el codificador en el dominio del tiempo se sitúa en un estado de inicialización, o en otras realizaciones que no requieren una conmutación instantánea, en un estado completamente desactivado. Sin embargo, cuando se determine que la porción de la señal de audio en consideración debe ser codificada por el codificador en el dominio del tiempo, en tal caso se desactiva el codificador en el dominio de la frecuencia.

[0058] A continuación se ilustra una implementación preferida del controlador ilustrado en la Fig. 8. La decisión de si debe elegirse la trayectoria ACELP o TCX se lleva a cabo en la decisión de conmutación mediante la simulación del codificador ACELP y TCX y conmutando hacia la rama de mejor rendimiento. Para ello, se estima la RSR de las ramas de ACELP y de TCX basándose en una simulación de codificador/decodificador de ACELP y TCX. La simulación de codificador/decodificador TCX se realiza sin análisis TNS/TTS, codificador IGF, codificador aritmético/salto de cuantificación o sin ningún decodificador TCX; en su lugar, la RSR del TCX se estima usando una estimación de la distorsión del cuantificador en el dominio MDCT conformado. La simulación del codificador/decodificador de ACELP se lleva a cabo usando solamente una simulación del libro del código adaptativo y del libro de código innovador. La RSR de ACELP se estima simplemente calculando la distorsión introducida por un filtro LTP en el dominio de las señales ponderado (libro de código adaptativo) y cambiando de escala esta distorsión mediante un factor constante (libro de código innovador). De esta manera se reduce considerablemente la complejidad en comparación con un enfoque en el que la codificación de TCX y de ACELP se lleva a cabo en paralelo. La rama con la RSR más alta se elige para la posterior ejecución de codificación completa.

[0059] En el caso en que se elige la rama TCX, se ejecuta un decodificador TCX en cada trama que emite una señal con la velocidad de muestreo ACELP. Esto se usa para actualizar las memorias usadas para la trayectoria de codificación ACELT (LPC residual, Memw0, desacentuación de la memoria), de manera que permita la conmutación instantánea de TCX a ACELP. La actualización de la memoria se lleva a cabo en cada trayectoria de TCX.

[0060] Como alternativa puede llevarse a cabo un análisis completo mediante un proceso de síntesis, es decir, los dos simuladores de codificador 621, 622 implementan las operaciones de codificación reales y los resultados son comparados por el selector 623. Nuevamente, como alternativa, es posible efectuar un cálculo directo de alimentación completa mediante la realización de un análisis de las señales. Por ejemplo, cuando se determina que la señal es una

señal de voz por medio de un clasificador de señales, se selecciona el codificador en el dominio del tiempo y cuando se determina que la señal es una señal de música, entonces se selecciona el codificador en el dominio de la frecuencia. También es posible aplicar otros procedimientos para diferenciar entre los dos codificadores basándose en un análisis de las señales de la porción de audio en consideración.

5

[0061] Preferentemente, el codificador de audio comprende además un procesador cruzado 700 ilustrado en la Fig. 7a. Cuando el codificador en el dominio de la frecuencia 600 está activo, el procesador cruzado 700 proporciona datos de inicialización al codificador en el dominio del tiempo 610 de manera que el codificador en el dominio del tiempo está listo para una conmutación ininterrumpida a una porción de la señal futura. Dicho de otro modo, cuando se determina que la porción de la señal actual ha de ser codificada mediante el codificador en el dominio de la frecuencia, y cuando se determina mediante el codificador que la porción de la señal de audio siguiente debe ser codificada por el codificador en el dominio del tiempo 610, entonces, sin un procesador cruzado, no sería posible una conmutación ininterrumpida inmediata de ese tipo. Sin embargo, el procesador cruzado proporciona una señal deducida del codificador en el dominio de la frecuencia 600 al codificador en el dominio del tiempo 610 con el fin de inicializar las memorias en el codificador en el dominio del tiempo ya que el codificador en el dominio del tiempo 610 depende de una trama actual vigente de la señal introducida o codificada de una trama inmediatamente anterior en el tiempo.

[0062] Por ello, el codificador en el dominio del tiempo 610 está configurado para ser inicializado por los datos de inicialización con el fin de codificar una porción de la señal de audio que sigue a una porción de la señal de audio más temprana codificada por el codificador en el dominio de la frecuencia 600 de una manera eficiente.

[0063] En particular, el procesador cruzado comprende un convertidor de tiempo para convertir una representación en el dominio de la frecuencia en una representación en el dominio del tiempo que puede ser enviada al codificador en el dominio del tiempo directamente o después de algún otro procesamiento adicional. Este convertidor se ilustra en la Fig. 14a como un bloque IMDCT (inverse modified discrete cosine transform, transformada de coseno discreta modificada inversa). Sin embargo, este bloque 702 tiene un tamaño de transformadas diferente en comparación con el bloque del convertidor de tiempo-frecuencia 602 indicado en el bloque de la Fig. 14a (bloque de transformadas de coseno discretas modificadas). Tal como se indica en el bloque 602, el convertidor de tiempo-frecuencia 602 opera con una velocidad de muestreo introducida y la transformada de coseno discreta modificada inversa 702 opera con la velocidad de muestreo ACELP más baja.

[0064] La relación entre la velocidad de muestreo del codificador en el dominio del tiempo o velocidad de muestreo ACELP y la velocidad de muestreo del codificador en el dominio de la frecuencia o velocidad de muestreo de entrada puede calcularse, y es un factor de submuestreo DS (downsampling factor) ilustrado en la Fig. 7b. El bloque 602 tiene un tamaño de transformada grande y el bloque de IMDCT 702 tiene un tamaño de transformada pequeño. Por tanto, y tal como se ilustra en la Fig. 7b, el bloque de IMDCT 702 comprende un selector 726 para seleccionar la porción espectral inferior de una entrada en el bloque de IMDCT 702. La porción del espectro de banda completa se define mediante el factor de submuestreo DS. Por ejemplo, cuando la velocidad de submuestreo es de 16 kHz y la velocidad de muestreo de entrada es de 32 kHz, entonces el factor de submuestreo 0,5, y por lo tanto, el selector 726 selecciona la mitad inferior del espectro de banda completa. Cuando el espectro tiene, por ejemplo, 1.024 líneas de MDCT, entonces el selector selecciona las 512 líneas inferiores.

[0065] Esta porción de baja frecuencia del espectro de banda completa es introducida en una transformada de tamaño pequeña y el bloque de despliegue 720, como se ilustra en la Fig. 7b. El tamaño de la transformada es también seleccionado según el factor de submuestreo y representa el 50 % del tamaño de la transformada en el bloque 602. A continuación se lleva a cabo un efecto de ventana de síntesis con una ventana provista de una pequeña cantidad de coeficientes. El número de coeficientes de la ventana de síntesis es igual al factor de submuestreo multiplicado por el número de coeficientes de la ventana de análisis usada por el bloque 602. Finalmente, se lleva a cabo una operación de superposición-suma con un número menor de operaciones por bloque, y nuevamente el número de operaciones por bloque es el número de operaciones por bloque en una implementación de velocidad completa de MDCT multiplicado por el factor de submuestreo.

[0066] Por lo tanto, es posible aplicar una operación de submuestreo eficiente ya que el submuestreo está incluido en la implementación de MDCT. En este contexto, se resalta que el bloque 702 puede ser implementado por un IMDCT pero también puede ser implementado por cualquier otra transformada o implementación de banco de filtros que puede dimensionarse adecuadamente en un núcleo de transformadas real y en otras operaciones relacionadas con transformadas.

[0067] En una realización adicional ilustrada en la Fig. 14a, el convertidor de tiempo-frecuencia comprende funcionalidades adicionales además del analizador. En la realización de la Fig. 14a, el analizador 604 de la Fig. 6 puede comprender un bloque 604a temporal de análisis de conformado de ruido/conformado de teselas temporal que opera como se expuso en el contexto de la Fig. 2b, bloque 222, para el bloque de análisis TNS/TTS 604a y tal como se ilustra con respecto a la Fig. 2b para la máscara tonal 226 que corresponde al codificador de IGF 604b en la Fig. 14a.

[0068] Por otra parte, el codificador en el dominio de la frecuencia comprende preferentemente un bloque de conformado de ruido 606a. El bloque de conformado de ruido 606a es controlado por coeficientes LPC cuantificados tales como los generados por el bloque 1010. Los coeficientes LPC cuantificados usados para el conformado de ruido 606a llevan a cabo un conformado espectral de los valores espectrales de alta resolución o de las líneas espectrales codificadas directamente (en lugar de codificadas paramétricamente), y el resultado del bloque 606a es similar al espectro de una señal posterior a una fase de filtrado de LPC que opera en el dominio del tiempo tal como un bloque de filtrado de análisis de LPC 704 que será descrito más adelante. Por otra parte, a continuación el resultado del bloque del conformado del ruido 606a es cuantificado y sometido a codificación de entropía tal como se indica mediante el bloque 606b. El resultado del bloque 606b corresponde a la primera porción de la señal de audio codificada o a una porción de la señal de audio codificada en el dominio de la frecuencia (junto con otra información secundaria).

[0069] El procesador cruzado 700 comprende un decodificador espectral para calcular una versión decodificada de la primera porción de la señal codificada. En la realización de la Fig. 14a, el decodificador espectral 701 comprende un bloque de conformado de ruido inverso 703, un decodificador de relleno de intervalos 704, un bloque de síntesis de TNS/TTS 705 y el bloque de IMDCT 702 anteriormente mencionados. Estos bloques deshacen las operaciones específicas llevadas a cabo por los bloques 602 a 606b. En particular, un bloque de conformado de ruidos 703 deshace el conformado de ruido llevado a cabo por el bloque 606a basándose en los coeficientes de LPC cuantificados 1010. El decodificador IGF 704 opera como se expuso con respecto a la Fig. 2A, bloques 202 y 206, y el bloque de síntesis TNS/TTS 705 opera como se expuso en el contexto del bloque 210 de la Fig. 2A y el decodificador espectral comprende adicionalmente el bloque de IMDCT 702. Además, adicionalmente o como alternativa, el procesador cruzado 700 en la Fig. 14a comprende una fase de retardo 707 para suministrar una versión retardada de la versión decodificada obtenida por el decodificador espectral 701 en una fase de desacentuación 617 del segundo procesador de codificación con el fin de inicializar la fase de desacentuación 617.

[0070] Además, como alternativa o adicionalmente, el procesador cruzado 17 puede comprender una fase de filtrado de análisis de los coeficientes de predicción ponderada 708 para filtrar la versión decodificada y para suministrar una versión decodificada filtrada a un determinador de libro de código 613 indicado como "MMSE" en la Fig. 14a del segundo procesador de codificación, para inicializar este bloque. Adicionalmente o como alternativa, el procesador cruzado comprende la fase de filtrado de análisis de LPC para filtrar la versión decodificada de la primera porción de la señal codificada emitida por el decodificador 700 a una fase de libro de código adaptativo 712 para la inicialización del bloque 612. Además o como alternativa, el procesador cruzado también comprende una fase de preacentuación 709 para realizar un procesamiento de preacentuación a la versión decodificada emitida por un decodificador espectral 701 antes del filtrado de LPC. La fase de preacentuación emitida también puede ser introducida en una fase de retardo adicional 710 con el fin de inicializar un bloque de filtrado de síntesis de LPC 616 dentro del codificador en el dominio del tiempo 610, para el fin de inicializar este bloque de filtrado de análisis LPC 611.

[0071] El procesador codificador en el dominio del tiempo 610 comprende, tal como se ilustra en la Fig. 14a, una preacentuación que opera sobre la velocidad de muestreo ACELP inferior. Tal como se ilustra, esta preacentuación es la preacentuación realizada en la fase de preprocesamiento 1000, y tiene el número de referencia 1005. Los datos de preacentuación son introducidos en una fase de filtrado de análisis de LPC 611 que opera en el dominio del tiempo, y este filtro es controlado por los coeficientes de LPC cuantificados 1010 obtenidos por la fase de preprocesamiento 1000. Como es sabido para los codificadores AMR-WB+ o USAC u otros codificadores CELP, la señal residual generada por el bloque 611 es proporcionada a un libro de código adaptativo 612, y además, el libro de código adaptativo 612 está conectado con una fase de libro de código innovador 614, y los datos del libro de código 612 y del libro de código innovador son introducidos en el multiplexor del tren de bits, tal como se ilustra.

[0072] Por otra parte, se ha proporcionado una fase de ganancia/codificación ACELP 612 en serie a la fase del libro de código innovador 614 y el resultado de este bloque es introducido en un determinador del libro de código 613 indicado como MMSE en la Fig. 14a. Este bloque coopera con el bloque del libro de código innovador 614. Además, el codificador en el dominio del tiempo comprende adicionalmente una porción de decodificador que tiene un bloque de filtrado de síntesis de LPC 616, un bloque de desacentuación 617 y una fase de posfiltrado de bajos adaptativo 618 para calcular los parámetros para un posfiltro adaptativo que, sin embargo, se aplica en el lado del decodificador. Sin ningún posfiltrado de bajos adaptativo en el lado del decodificador, los bloques 616, 617, 618 no serían necesarios para el codificador en el dominio del tiempo 610.

[0073] Como se ilustra, varios bloques del decodificador en el dominio del tiempo dependen de señales previas y estos bloques son el bloque del libro de código adaptativo, el determinador de libro del código 613, el bloque de filtrado de síntesis PC L616 y el bloque del desacentuación 617. A estos bloques se les proporcionan datos procedentes del procesador cruzado deducido del procesador de codificación en el dominio de la frecuencia con el fin de inicializar estos bloques para que estén listos para una conmutación instantánea desde el codificador en el dominio de la frecuencia al codificador en el dominio del tiempo. Como puede observarse en la Fig. 14a, no es necesaria ninguna dependencia con respecto a los datos anteriores para el codificador en el dominio de la frecuencia. Por tanto, el procesador cruzado 700 no proporciona ningún dato de inicialización de memoria procedente del codificador en el dominio del tiempo para el codificador en el dominio de la frecuencia. Sin embargo, para otras implementaciones del

codificador en el dominio de la frecuencia, en el que existen dependencias del pasado y en el que se requieren datos para la inicialización de la memoria, el procesador cruzado 700 está configurado de manera que funcione en las dos direcciones.

5 **[0074]** Una realización preferida de un codificador de audio comprende por tanto las siguientes partes:

A continuación se describe el decodificador de audio preferido: La parte del decodificador de forma de onda consiste en una trayectoria de decodificador TCX de banda completa con un IGF y los dos operan con la velocidad de muestreo introducida del códec. En paralelo, existe una trayectoria de decodificador ACELP alternativa con una velocidad de muestreo que es baja y que es reforzada adicionalmente en sentido descendente por un TD-BWE.

10 **[0075]** Para la inicialización de ACELP cuando se conmute de TCX a ACELP, existe una trayectoria cruzada (consistente en un extremo frontal del decodificador TCX compartido que proporciona adicionalmente una salida con la velocidad de muestreo más baja y algún posprocesamiento) que lleva a cabo la inicialización de ACELP de la invención. La compartición de la misma velocidad de muestreo y orden de filtrado entre TCX y ACELP en los LPC permite una inicialización más fácil y más eficiente del ACELP.

[0076] Para visualizar la conmutación, en la Fig. 14a se han esquematizado dos conmutadores. Mientras el segundo conmutador en sentido descendente elige entre las salidas TCX/IGF o ACELP/TD-BWE, el primer conmutador preactualiza las memorias intermedias en la fase QMF del muestreo en sentido descendente de la trayectoria de ACELP por la salida de la trayectoria cruzada o simplemente pasa a la salida de ACELP.

[0077] A continuación, y en el contexto de las Fig. 11a–14c se exponen implementaciones del decodificador de audio según aspectos de la presente invención.

25 **[0078]** Un decodificador de audio para decodificar una señal de audio codificada 1101 comprende un primer procesador de decodificación 1120 para decodificar una primera porción de la señal de audio en un dominio de la frecuencia. El primer procesador de decodificación 1120 comprende un decodificador espectral 1122 para decodificar primeras porciones espectrales con una alta resolución espectral y para sintetizar segundas porciones espectrales usando una representación paramétrica de las segundas regiones espectrales y al menos una primera región espectral decodificada con el fin de obtener una representación espectral decodificada. La representación espectral decodificada es una representación espectral decodificada de banda completa, como se expone el contexto de la Fig. 6, y como también se expone en el contexto de la Fig. 1a. Por tanto, y en términos generales, el primer procesador de decodificación comprende una implementación de banda completa con un procedimiento de relleno de intervalos en el dominio de la frecuencia. El primer procesador de decodificación 1120 comprende además un convertidor de frecuencia–tiempo 1124 para convertir la representación espectral decodificada en un dominio del tiempo con el fin de obtener una primera porción decodificada de la señal de audio.

[0079] Además, el decodificador de audio comprende un segundo procesador de decodificación 1140 para decodificar la segunda porción de la señal de audio codificada en el dominio del tiempo con el fin de obtener una segunda porción de audio decodificada. Además, el decodificador de audio comprende un combinador 1160 para combinar la primera porción de la señal decodificada y la segunda porción de la señal decodificada con el fin de obtener una señal de audio decodificada. Las porciones de la señal decodificada se combinan en secuencia lo que también se ilustra en la Fig. 14b mediante una implementación de conmutación 1160 que representa una realización del combinador 1160 de la Fig. 11a.

[0080] Preferentemente, el segundo procesador de decodificación 1140 es un procesador de extensión de anchura de banda en el dominio del tiempo y comprende, tal como se ilustra en la Fig. 12, un decodificador de banda baja en el dominio del tiempo 1200 para decodificar una señal en el dominio del tiempo de banda baja. Esta implementación comprende además un sobremuestreador 1210 para sobremuestrear la señal en el dominio del tiempo de banda baja. Adicionalmente, se proporciona un decodificador de extensión de anchura de banda en el dominio del tiempo de 1220 para sintetizar una banda alta de la señal de audio emitida. Además, se proporciona un mezclador 1230 para mezclar una banda alta sintetizada de la señal introducida en el dominio del tiempo y una señal en el dominio del tiempo de banda baja sobremuestreada con el fin de obtener la salida del decodificador en el dominio del tiempo. Por ello, en una realización preferida el bloque 1140 en la Fig. 11a puede ser implementado por la funcionalidad de la Fig. 12.

[0081] La Fig. 13 ilustra una realización preferida del decodificador de extensión de anchura de banda en el dominio del tiempo 1220 de la Fig. 12. Preferentemente, se proporciona un sobremuestreador en el dominio del tiempo 1221 que recibe, como una entrada, una señal residual procedente de un decodificador de banda baja en el dominio del tiempo incluido dentro del bloque 1140 e ilustrado como 1200 en la Fig. 12 y además ilustrado en el contexto de la Fig. 14b. El sobremuestreador en el dominio del tiempo 1221 genera una versión sobremuestreada de la señal residual de LPC. A continuación, esta versión es introducida en un bloque de distorsión 1222 que genera, basándose en la señal introducida, una señal emitida que tiene valores de frecuencia más altos. Una distorsión no lineal puede ser una operación de copia ascendente, una duplicación en espejo, un desplazamiento de frecuencia o un dispositivo no lineal

tal como un diodo o un transistor operado en la región no lineal. La señal emitida del bloque 1222 es introducida en un bloque de filtrado de síntesis de LPC 1223 que es controlado por los datos de LPC usados para el decodificador de banda baja así como por datos de envolvente específicos generados por el bloque de extensión de anchura de banda en el dominio del tiempo 920 en el lado del decodificador de la Fig. 14a, por ejemplo. La salida del bloque de síntesis de LPC es introducida a continuación en un filtro de paso de banda o de paso alto 1224 con el fin de obtener finalmente la banda alta, que es seguidamente introducida en el mezclador 1230, tal como se ilustra en la Fig. 12.

[0082] A continuación, en el contexto de la Fig. 14b se expone una implementación preferida del sobremuestreador 1210 de la Fig. 12. El sobremuestreador comprende preferentemente un banco de filtros de análisis que opera con una velocidad de muestreo del decodificador de banda baja en el dominio del tiempo. Una implementación específica de un banco de filtros de análisis de este tipo es un banco de filtros de análisis de QMF 1471 ilustrado en la Fig. 14b. Además, el sobremuestreador comprende un banco de filtros de síntesis 1473 que opera con una velocidad de muestreo de salida que es más alta que la primera velocidad de muestreo de banda baja en el dominio del tiempo. Por ello, el banco de filtros de síntesis de QMF 1473 que es una implementación preferida del banco de filtros general opera con la velocidad de muestreo de salida. Cuando el factor de submuestreo T explicado en el contexto de la Fig. 7b es de 0,5, entonces el banco de filtros de análisis de QMF 1471 tiene, por ejemplo, solamente 32 canales del banco de filtros y el banco de filtros de síntesis QMF 1473 tiene por ejemplo 64 canales QMF, pero la mitad superior de los canales del banco de filtros, es decir, los 32 canales superiores del banco de filtros son alimentados con ceros o ruido, mientras que los 32 canales inferiores del banco de filtros son alimentados con las correspondientes señales proporcionadas por el banco de filtros de análisis de QMF 1471. Sin embargo, preferentemente se lleva a cabo un filtrado de paso de banda 1472 en el dominio del banco de filtros de QMF con el fin de asegurar que la salida de síntesis 1473 sea una versión sobremuestreada de la salida del decodificador ACELP, pero sin ningún artefacto por encima de la frecuencia máxima del decodificador ACELP.

[0083] Pueden realizarse operaciones de procesamiento adicionales dentro del dominio de QMF además o en lugar del filtrado de paso de banda 1472. Si no se lleva a cabo ningún procesamiento, entonces el análisis de QMF y la síntesis de QMF constituyen un sobremuestreador eficiente 1210.

[0084] A continuación se expone con mayor detalle el diseño de los elementos individuales de la Fig. 14b.

[0085] El decodificador en el dominio de la frecuencia de banda completa 1120 comprende un primer bloque de decodificación 1122a para decodificar los coeficientes espectrales de alta resolución y para realizar adicionalmente el relleno de los ruidos en la porción de banda baja, como es conocido por ejemplo de la tecnología USAC. Además, el decodificador de banda completa comprende un procesador IGF 1122b para relleno del

intervalo espectral usando valores espectrales sintetizados que han sido codificados sólo paramétricamente y, por lo tanto, con una baja resolución en el lado del codificador. A continuación, en el bloque 1122c, se lleva a cabo un conformado inverso del ruido y el resultado es introducido en un bloque de síntesis TNS/TTS 705 que proporciona, como una salida final, una entrada al convertidor de frecuencia-tiempo 1124, que está preferentemente implementado como una transformada de coseno discreta modificada inversa que opera con alta velocidad de muestreo de salida.

[0086] Por otra parte, se usa un posfiltro de armónicos o de LTP que es controlado por los datos obtenidos por el bloque de extracción de parámetros TCX-LTP 1006 en la Fig. 14b. El resultado es entonces la primera porción de la señal de audio decodificada con la velocidad de muestreo de salida como puede observarse en la Fig. 14b, teniendo dichos datos una alta velocidad de muestreo, y por tanto, no es necesario ningún refuerzo adicional de la frecuencia debido al hecho de que el procesador de decodificación es un decodificador de banda completa en el dominio de la frecuencia que preferentemente usa la tecnología inteligente de relleno de intervalos expuesta en el contexto de las Fig. 1a-5C.

[0087] Varios elementos en la Fig. 14b son bastante similares a los bloques correspondientes en el procesador cruzado 700 de la Fig. 14a, particularmente con respecto al decodificador de IGF 704 correspondiente al procesamiento de IGF 1122 y la operación de conformado inversa del ruido controlada por los coeficientes de LPC cuantificados 1145 corresponde al conformado inverso del ruido 703 de la Fig. 14a y el bloque de síntesis de TNS/TTS 705 en la Fig. 14b corresponde al bloque de TNS/TTS 705 en la Fig. 14a. Sin embargo, es importante observar que el bloque de IMDCT 1124 en la Fig. 14b opera con alta velocidad de muestreo mientras que el bloque de IMDCT 702 en la Fig. 14a opera con una velocidad de muestreo baja. Por ello, el bloque 1124 en la Fig. 14b comprende la transformada de gran tamaño y el bloque de desplegado 710, la ventana de síntesis en el bloque 712 y la fase de superposición-suma 714 con el correspondiente gran número de operaciones, gran número de coeficientes de ventana y un gran tamaño de transformada en comparación con las características 720, 722, 724 correspondientes, que se manejan en el bloque 702, y como se destacará más adelante, en el bloque 1171 del procesador cruzado 1170 en la Fig. 14b también.

[0088] Preferentemente el procesador de decodificación en el dominio del tiempo 1140 comprende el ACELP o el decodificador de banda baja en el dominio del tiempo 1200 que comprende una fase de decodificador de ACELP 1149 para obtener ganancias decodificadas y la información del libro del código innovador. Adicionalmente se

proporciona una fase de libro del código adaptativo de ACELP 1141 y una fase de posprocesamiento de ACELP 1142 posterior y un filtro de síntesis final tal como un filtro de síntesis LPC 1143, que, nuevamente, es controlado por los coeficientes cuantificados 1145 obtenidos desde el demultiplexor de tren de bits 1100 correspondiente al analizador sintáctico de señales codificadas 1100 en la Fig. 11a. La salida del filtro de síntesis de LPC 1143 es introducida en una fase de desacentuación 1144 para cancelar o deshacer el procesamiento introducido por la fase de preacentuación 1005 del preprocesador 1000 de la Fig. 14a. El resultado es la señal de salida en el dominio del tiempo con una baja velocidad de muestreo y una banda baja y si se requiere una salida en el dominio de la frecuencia, el conmutador 1480 se encuentra en la posición indicada y la salida de la fase de desacentuación 1144 es introducida en el sobremuestreador 1210 y a continuación mezclada con las bandas altas procedentes del decodificador de extensión de anchura de banda en el dominio del tiempo 1220.

[0089] De acuerdo con realizaciones de la presente invención, el decodificador de audio comprende adicionalmente el procesador cruzado 1170 ilustrado en la Fig. 11b y en la Fig. 14b para calcular, a partir de la representación espectral decodificada de la primera porción de audio codificada, datos de inicialización del segundo procesador de decodificación de manera que el segundo procesador de decodificación sea inicializado con el fin de decodificar la segunda porción de audio codificada que sigue en el tiempo a la primera porción de audio en la señal de audio codificada, es decir, de manera que el procesador de decodificación en el dominio del tiempo 1140 esté listo para una conmutación instantánea desde una porción de la señal de audio a la siguiente sin ninguna pérdida en calidad ni eficiencia.

[0090] Preferentemente, el procesador cruzado 1170 comprende un convertidor adicional de frecuencia-tiempo 1171 que opera con una velocidad de muestreo más baja que el convertidor de frecuencia del primer procesador de decodificación con el fin de obtener otra primera porción decodificada en el dominio del tiempo que se usará como la señal de inicialización o para el que sea posible deducir cualquier dato de inicialización. Preferentemente, este IMDCT o convertidor de frecuencia-tiempo de baja velocidad de muestreo se implementa como se ilustra en la Fig. 7b, elemento 726 (selector), elemento 720 (transformada de tamaño pequeño y desplegado), formación de ventana de síntesis con un número menor de coeficientes de ventana que lo indicado en 722 y una fase de superposición-suma con un número menor de operaciones que lo indicado en 724. Por lo tanto, el bloque de IMDCT 1124 en el decodificador de banda completa en el dominio de la frecuencia es implementado como se indica mediante los bloques 710, 712, 714, y el bloque de IMDCT 1171 es implementado como se indica en la Fig. 7b mediante los bloques 726, 720, 722, 724. Nuevamente, el factor de muestra descendente es la relación entre la velocidad de muestreo del codificador en el dominio del tiempo o la baja velocidad de muestreo y la velocidad de muestreo en el dominio de la frecuencia, más alta, o la velocidad de muestreo de salida, y este factor de submuestreo es menor que 1 y puede ser cualquier número superior a 0 e inferior a 1.

[0091] Tal como se ilustra en la Fig. 14b, el procesador cruzado 1170 comprende además, sola o añadida a otros elementos, una fase de retardo 1172 para retrasar la primera porción de señales decodificada adicional y para suministrar la primera porción de la señal decodificada en una fase de desacentuación 1144 del segundo procesador de decodificación para la inicialización. Además, el procesador cruzado comprende, de forma adicional o alternativa, un filtro de preacentuación 1173 y una fase de retardo 1175 para filtrar y retrasar una primera porción de la señal decodificada adicional y para proporcionar la salida retardada del bloque 1175 en una fase de filtrado de síntesis de LPC 1143 del decodificador de ACELP con fines de inicialización.

[0092] Además, el procesador cruzado puede comprender de forma alternativa o adicional a los otros elementos mencionados, un filtro de análisis de LPC 1174 para generar una señal residual de predicción procedente de la primera porción de la señal adicional o de una primera porción de la señal adicional de preacentuación y para introducir los datos en un sintetizador de libro de códigos de un segundo proceso de decodificación y, preferentemente, en la fase del libro del código adaptativo 1141. Además, la salida del convertidor de frecuencia-tiempo 1171 con la baja velocidad de muestreo se introduce también en la fase de análisis de QAMF 1471 del sobremuestreador 1210 con fines de inicialización, es decir, cuando la porción de la señal de audio de código actualmente decodificada es suministrada al decodificador de banda completa en el dominio de la frecuencia 1120.

[0093] A continuación se describe el decodificador de audio preferido: La parte del decodificador de forma de onda consiste en una trayectoria de decodificador TCX de banda completa con un IGF y los dos operan con la velocidad de muestreo de entrada del códec. En paralelo, existe una trayectoria de decodificador ACELP alternativa con una velocidad de muestreo menor que es reforzada adicionalmente en sentido descendente por un TD-BWE.

[0094] Para la inicialización de ACELP cuando se conmute de TCX a ACELP, existe una trayectoria cruzada (consistente en un extremo frontal del decodificador TCX compartido que proporciona adicionalmente una salida con la velocidad de muestreo más baja y algún posprocesamiento) que lleva a cabo la inicialización de ACELP de la invención. La compartición de la misma velocidad de muestreo y orden de filtrado entre TCX y ACELP en los LPC permite una inicialización más fácil y más eficiente del ACELP.

[0095] Para visualizar la conmutación, en la Fig. 14a se han esquematizado dos conmutadores. Mientras el segundo conmutador en sentido descendente elige entre las salidas TCX/IGF o ACELP/TD-BWE, el primer

conmutador preactualiza las memorias intermedias en la fase QMF del muestreo en sentido descendente de la trayectoria de ACELP por la salida de la trayectoria cruzada o simplemente pasa a la salida de ACELP.

[0096] Para resumir, los aspectos preferidos de la invención que pueden usarse solos o en combinación se refieren a una combinación de un codificador ACELP y TDS-BWE con una tecnología TCX/IGF capaz de banda completa asociada preferentemente con el uso de una señal cruzada.

[0097] Otra característica específica se refiere a una trayectoria de señal cruzada para la inicialización de ACELP con el fin de habilitar una conmutación ininterrumpida.

[0098] Otro aspecto más es que se suministra un IMDCT corto con una parte inferior de coeficientes de MDCT largos de alta velocidad con el fin de implementar de manera eficiente una conversión de la velocidad binaria de muestras en la trayectoria cruzada.

[0099] Una característica adicional se refiere a una realización eficiente de una trayectoria cruzada parcialmente compartida con una banda completa de TCX/IGF en el decodificador.

[0100] Una característica adicional es la trayectoria de señal cruzada para la inicialización de QMF con el fin de permitir la conmutación ininterrumpida de TCX a ACELP.

[0101] Una característica adicional es una trayectoria de señal cruzada para el QMF que permite compensar el intervalo de retardo entre la salida remuestreada de ACELP y una salida de banco de filtros-TCX/IGF cuando se conmuta de ACELP a TCX.

[0102] Otro aspecto más es que se proporciona un LPC tanto para el codificador de TCX como de ACELP con la misma velocidad de muestreo de muestras y orden de filtro, si bien el codificador/decodificador TCX/IGF es capaz de banda completa.

[0103] A continuación en la Fig. 14c se expone una implementación preferida de un decodificador en el dominio del tiempo que opera como un decodificador autónomo o en combinación con el decodificador en el dominio de la frecuencia capaz de funcionar con banda completa.

[0104] En términos generales, el decodificador en el dominio del tiempo comprende un decodificador ACELP, un remuestreador o sobremuestreador conectados posteriormente y una funcionalidad de extensión de anchura de banda en el dominio del tiempo. En particular, el decodificador ACELP comprende una fase de decodificación ACELP para restaurar las ganancias y el libro del código innovador 1149, una fase de libro de código ACELP-adaptativo 1141, un posprocesador ACELP 1142, un filtro de síntesis de LPC 1143 controlado por coeficientes LPC cuantificados procedentes de un multiplexor de tren de bits o analizador sintáctico de señales codificadas y la fase de desacentuación 1144 conectada posteriormente. Preferentemente, la señal residual en el dominio del tiempo con una velocidad de muestreo ACELP es introducida en un decodificador de extensión de anchura de banda en el dominio del tiempo 1220 que proporciona una banda alta en las salidas.

[0105] Con el fin de sobremuestrear la salida de desacentuación 1144, se proporciona un sobremuestreador que comprende el bloque de análisis 1471 y el bloque de síntesis 1473. Dentro del dominio del banco de filtros definido por los bloques 1471 y 1473, preferentemente se ha aplicado un filtro de paso de banda. En particular, y como ya se expuso anteriormente, pueden usarse las mismas funcionalidades que han sido expuestas con los mismos números de referencia. Además, el decodificador de extensión de anchura de banda en el dominio del tiempo 1220 puede implementarse como se ilustra en la Fig. 13, y en términos generales, comprende un sobremuestreo de la señal residual ACELP o de la señal residual en el dominio del tiempo con la velocidad de muestreo de ACELP finalmente con una velocidad de muestreo emitida de la señal extendida de anchura de banda.

[0106] Posteriormente, y haciendo referencia a las Fig. 1A-5C se exponen más detalles con respecto al codificador y decodificador en el dominio de la frecuencia.

[0107] En la Fig. 1 se ilustra un aparato para codificar una señal de audio 99. La señal de audio 99 es introducida en un convertidor en el espectro del tiempo 100 con el fin de convertir una señal de audio que tiene una velocidad de muestreo en una representación espectral 101 emitida por el convertidor en el espectro del tiempo. El espectro 101 es introducido en un analizador espectral 102 para analizar la representación espectral 101. El analizador espectral 101 está configurado para determinar un primer conjunto de primeras porciones espectrales 103 que se codificarán con una primera resolución espectral y un segundo conjunto, diferente, de porciones espectrales 105 que se codificarán con una segunda resolución espectral. La segunda resolución espectral es menor que la primera resolución espectral. El segundo conjunto de segundas porciones espectrales 105 es introducido en un calculador de parámetros o codificador paramétrico 104 para calcular información de envolvente espectral que tiene la segunda resolución espectral. Además, se proporciona un codificador de audio en el dominio espectral 106 para generar una primera representación codificada 107 del primer conjunto de primeras porciones espectrales que tienen la primera resolución

espectral. Además, el calculador de parámetros/codificador paramétrico 104 está configurado para generar una segunda representación codificada 109 del segundo conjunto de segundas porciones espectrales. La primera representación codificada 107 y la segunda representación codificada 109 son introducidas en un multiplexor de tren de bits o conformador de tren de bits 108, y el bloque 108 finalmente emite la señal de audio codificada para su transmisión o para su almacenamiento en un dispositivo de almacenamiento.

[0108] Normalmente, una primera porción espectral tal como 306 de la Fig. 3a estará rodeada por dos porciones espectrales tales como 307a, 307b. No sucede así, por ejemplo, en HE-AAC, en el que el intervalo de frecuencias del codificador de núcleo presenta una anchura de banda limitada.

La Fig. 1b ilustra un decodificador que se corresponde con el codificador de la Fig. 1a. La primera representación codificada 107 es introducida en un decodificador de audio en el dominio espectral 112 para generar una primera representación decodificada de un primer conjunto de primeras porciones espectrales, teniendo la representación decodificada una primera resolución espectral. Además, la segunda representación codificada 109 es introducida en un decodificador paramétrico 114 con el fin de generar una segunda representación decodificada de un segundo conjunto de segundas porciones espectrales que tienen una segunda resolución espectral que es inferior a la primera resolución espectral.

[0109] El decodificador comprende además un regenerador de frecuencias 116 para regenerar una segunda porción espectral reconstruida que tiene la primera resolución espectral que usa una primera porción espectral. El regenerador de frecuencias 116 lleva a cabo una operación de relleno de teselas, es decir, usa una tesela o porción del primer conjunto de primeras porciones espectrales y copia este primer conjunto de primeras porciones espectrales en el intervalo de reconstrucción o banda de reconstrucción que tiene la segunda porción espectral y normalmente lleva a cabo un conformado de envolvente espectral u otra operación indicada por la segunda representación decodificada emitida por el decodificador paramétrico 114, es decir, mediante el uso de la información del segundo conjunto de segundas porciones espectrales. El primer conjunto decodificado de primeras porciones espectrales y el segundo conjunto de porciones espectrales, reconstruido, indicado como salida del regenerador de frecuencias 116 en la línea 117 es introducido en un convertidor de espectro-tiempo 118 conformado para convertir el primer conformado decodificado y la segunda porción espectral reconstruida en una representación en el tiempo 119, teniendo la representación en el tiempo una alta velocidad de muestreo determinada.

[0110] En la Fig. 2b se ilustra una implementación del codificador de la Fig. 1a. Una señal introducida de audio 99 es introducida en un banco de filtros de análisis 220 correspondiente al convertidor de espectros del tiempo 100 de la Fig. 1a. A continuación, en el bloque de TNS 222 se lleva a cabo una operación de conformado temporal del ruido. Por tanto, la entrada en el analizador espectral 102 de la Fig. 1a correspondiente a una máscara tonal de bloque 226 de la Fig. 2b puede consistir en valores espectrales completos, cuando no se aplica el conformado temporal de ruido/operación de conformado temporal de teselas, o puede tratarse de valores residuales espectrales, cuando se aplica la operación de TNS como se ilustra en la Fig. 2b, bloque 222. Para las señales de dos canales o para las señales de múltiples canales, es posible llevar a cabo adicionalmente una codificación de canales conjunta 228, de manera que el codificador en el dominio espectral 106 de la Fig. 1a puede comprender el bloque de codificación de canales conjunto 228. Además, se proporciona un codificador de entropía 232 para realizar una compresión sin pérdida de datos que también es una porción del codificador en el dominio espectral 106 de la Fig. 1a.

[0111] El analizador espectral/máscara tonal 226 separa la salida del bloque TNS 222 en la banda de núcleo y en los componentes tonales correspondientes al primer conjunto de primeras porciones espectrales 103 y los componentes residuales correspondientes al segundo conjunto de segundas porciones espectrales 105 de la Fig. 1a. El bloque 224 indicado como la codificación de la extracción de los parámetros de IGF corresponde al codificador paramétrico 104 de la Fig. 1a, y el multiplexor del tren de bits 230 corresponde al multiplexor del tren de bits 108 de la Fig. 1a.

[0112] Preferentemente, el banco de filtros de análisis 222 se implementa como un MDCT (modified discrete cosine transform filterbank, banco de filtros de transformada de coseno discreta modificada) y el MDCT se utiliza para transformar la señal 99 en el dominio del tiempo-frecuencia, en el que la transformada de coseno discreta modificada actúa como herramienta de análisis de las frecuencias.

[0113] Preferentemente el analizador espectral 226 aplica una máscara de tonalidad. La fase de la estimación de la máscara de tonalidad se usa para separar los componentes tonales de los componentes similares a ruido presentes en la señal. Esto permite que el codificador de núcleo 228 codifique todos los componentes tonales con un módulo psicoacústico. La fase de estimación de la máscara de tonalidad se puede implementar en numerosas vías diferentes y preferentemente se implementa de modo similar en su funcionalidad a la fase de estimación de rastreo sinusoidal usado en el modelo sinusoidal y de ruido para codificación de voz/audio [8, 9] o un codificador de audio basado en el modelo de HILN descrito en [10]. Preferentemente, se usa una implementación que es sencilla de implementar sin la necesidad de mantener trayectorias de nacimiento-muerte, aunque también se puede usar cualquier otro detector de ruidos o tonalidad.

[0114] El módulo de IGF calcula la semejanza que existe entre una región fuente y una región objeto. La región objeto se representará en el espectro de la región fuente. La medida de semejanza entre las regiones fuente y objeto se realiza usando un enfoque de correlación cruzada. La región objeto se divide en teselas de frecuencia que no se

superponen n_{Tar} . Por cada tesela en la región objeto, las teselas fuente n_{Src} se crean a partir de una frecuencia de inicio fija. Estas teselas fuente se superponen en un factor entre 0 y 1, en el que 0 significa el 0 % de superposición y 1 significa el 100 % de superposición. Cada una de estas teselas fuente se correlaciona con la tesela fuente en diversos retardos para encontrar la tesela fuente que mejor coincida con la tesela objeto. El mejor número de teselas

coincidentes se almacena en $tileNum[idx_tar]$, el retardo en que mejor se correlaciona con el objetivo se almacena en $xcorr_lag[idx_tar][idx_src]$ y el signo de la correlación se almacena en

$xcorr_sign[idx_tar][idx_src]$. En el caso en que la correlación es muy negativa, la tesela fuente se debe multiplicar por -1 antes del proceso de relleno de teselas en el decodificador. El módulo de IGF también se encarga de no sobrescribir los componentes tonales en el espectro ya que los componentes tonales se conservan usando la máscara de la tonalidad. Un parámetro de energía en forma de banda se usa para almacenar la energía de la región objeto que permite reconstruir el espectro de precisión.

[0115] El procedimiento presenta ciertas ventajas con respecto al SBR [1] clásico, por cuanto que la retícula de armónicos de una señal de múltiples tonos es conservada por el codificador de núcleo, mientras que solamente los intervalos entre las sinusoides son rellenos con el "ruido conformado" de mejor concordancia procedente de la región fuente. Otra ventaja de este sistema en comparación con ASR (sustitución espectral exacta) [2–4] es la ausencia de una fase de síntesis de señal que crea las porciones importantes de la señal en el decodificador. En su lugar, esta tarea es asumida por el codificador de núcleo, lo que permite la conservación de los componentes importantes del espectro. Otra ventaja del sistema propuesto es la escalabilidad continua que ofrecen las características. Usando solo

$tileNum[idx_tar]$ y $xcorr_lag = 0$, para cada tesela se llama coincidencia de granularidad en bruto y se puede

usar para velocidades binarias bajas mientras que el uso de la variable $xcorr_lag$ para cada tesela permite hacer coincidir mejor los espectros objeto y fuente.

[0116] Además, se propone una técnica de estabilización de la elección de las teselas que elimina los artefactos en el dominio de la frecuencia tales como vibración y ruido musical.

[0117] En el caso de los pares de canales estéreo se aplica un procesamiento de estéreo conjunto adicional. Esto es necesario, dado que para un determinado intervalo de destino la señal puede ser una fuente de sonido en panorámica altamente correlacionada. En el caso en que las regiones fuente elegidas para esta región particular no estén bien correlacionadas entre sí, y si bien las energías se corresponden con las regiones de destino, la imagen espacial puede sufrir debido a las regiones fuente no correlacionadas. El codificador analiza cada banda de energía de la región objeto, normalmente llevando a cabo una correlación cruzada de los valores espectrales y si se supera un determinado umbral, establece un indicador conjunto para esta banda de energía. En el decodificador las bandas de energía de canal izquierdo y derecho son tratadas individualmente si este indicador estéreo conjunto no se ajusta. En el caso en que se ajuste el indicador estéreo conjunto, tanto las energías como los parches se realizan en el dominio estéreo conjunto. La información exterior conjunta para las regiones IGF es señalizada de manera similar a la información estéreo conjunta para la codificación de núcleo, lo que incluye un indicador que indica en el caso de la predicción si la dirección de la predicción discurre desde la submezcla a residual o a la inversa.

[0118] Las energías pueden calcularse a partir de las energías transmitidas en el dominio L/R.

$$midNrg[k] = leftNrg[k] + rightNrg[k];$$

$$sideNrg[k] = leftNrg[k] - rightNrg[k];$$

siendo k el índice de frecuencia en el dominio de la transformada.

[0119] Otra solución consiste en calcular y transmitir las energías directamente en el dominio estéreo conjunto para bandas en las que el estéreo conjunto está activo, por lo que no se necesita información de energía adicional en el lado del decodificador.

[0120] Las teselas fuente siempre se crean según el Mid/Side–Matrix:

$$midTile[k] = 0,5 \cdot (leftTile[k] + rightTile[k])$$

$$sideTile[k] = 0,5 \cdot (leftTile[k] - rightTile[k])$$

[0121] Ajuste de la energía:

$$\begin{aligned} midTile[k] &= midTile[k] * midNrg[k]; \\ sideTile[k] &= sideTile[k] * sideNrg[k]; \end{aligned}$$

[0122] Estéreo conjunto → transformación LR:

10 Si no hay parámetro de predicción adicional codificado:

$$\begin{aligned} leftTile[k] &= midTile[k] + sideTile[k] \\ rightTile[k] &= midTile[k] - sideTile[k] \end{aligned}$$

15 Si se codifica un parámetro de predicción adicional y si la dirección señalada es desde el centro al lateral:

$$\begin{aligned} sideTile[k] &= sideTile[k] - predictionCoeff \cdot midTile[k] \\ leftTile[k] &= midTile[k] + sideTile[k] \\ rightTile[k] &= midTile[k] - sideTile[k] \end{aligned}$$

20

Si la dirección señalizada es desde el lateral al centro:

$$\begin{aligned} midTile[k] &= midTile[k] - predictionCoeff \cdot sideTile[k] \\ leftTile[k] &= midTile[k] - sideTile[k] \\ rightTile[k] &= midTile[k] + sideTile[k] \end{aligned}$$

25

[0123] Este procesamiento asegura que desde las teselas usadas para generar regiones de destino altamente correlacionadas y regiones de destino en panorámica, los canales izquierdo y derecho siguen representando una fuente de sonido correlacionada y en panorámica incluso si las regiones fuente no están correlacionadas, conservándose la imagen estéreo para dichas regiones.

30

[0124] Dicho de otro modo, en el tren de bits se transmiten indicadores estéreo conjuntos que indican si se usará L/R o M/S como ejemplo para la codificación estéreo conjunta general. En primer lugar, en el decodificador la señal de núcleo es decodificada como se indica mediante los indicadores estéreo conjuntos para las bandas de núcleo. En segundo lugar, la señal de núcleo es almacenada tanto en la representación L/R como M/S. Para el relleno de teselas IGF, la representación de teselas fuente se elige de manera que se ajuste a la representación de teselas objeto como se indica mediante la información estéreo conjunta para las bandas IGF.

35

[0125] El conformado temporal del ruido (TNS, Temporal Noise Shaping) es una técnica estándar y forma parte del AAC [11–13]. El TNS puede considerarse una ampliación del esquema básico de un codificador porcentual, al insertar una etapa de procesamiento opcional entre el banco de filtros y la fase de cuantificación. La tarea principal del módulo de TNS consiste en ocultar el ruido de cuantificación producido en la región de enmascaramiento temporal de señales similares a transitorios y por lo tanto conduce a un esquema de codificación más eficiente. En primer lugar, el TNS calcula un conjunto de coeficientes de predicción mediante el uso de la “predicción directa” en el dominio de las transformadas, por ejemplo, MDCT. Estos coeficientes se usan a continuación para aplanar la envolvente temporal de la señal. Dado que la cuantificación afecta al espectro filtrado de TNS, también el ruido de cuantificación es temporalmente plano. Mediante la aplicación del filtrado de TNS inverso en el lado del decodificador, el ruido de la cuantificación se conforma según la envolvente temporal del filtro de TNS y por tanto el ruido de cuantificación llega a ser enmascarado por los transitorios.

50

[0126] El IGF se basa en una representación de MDCT. Para una codificación eficiente deben usarse preferentemente bloques largos de 20 ms aproximadamente. Si la señal con dicho bloque largo contiene transitorios, se presentan pre-ecos y post-ecos audibles en las bandas espectrales de IGF debido al relleno de teselas. La Fig. 7c muestra un efecto pre-eco típico antes de la aparición de transitorios debido al IGF. En el lado izquierdo, se muestra el espectrograma de la señal original y en el lado derecho se muestra el espectrograma de la señal de anchura de banda extendida sin filtrado TNS.

[0127] El efecto del pre-eco se reduce usando TNS en el contexto IGF. En este caso, el TNS se usa como una herramienta de conformado temporal de teselas (TTS, temporal tile shaping) ya que la regeneración espectral en el decodificador se lleva a cabo en la señal residual de TNS. Los coeficientes de predicción de TTS requeridos se calculan y aplican usando el espectro completo en el lado del codificador, como es habitual. Las frecuencias de inicio y parada de TNS/TTS no se ven afectadas por la frecuencia de inicio *freqstart* de IGF de la herramienta de IGF. En comparación con el TNS de legado, la frecuencia de parada de TTS es incrementada hasta la frecuencia de parada de la herramienta IGF, que es mayor que *freqstart*. En el lado del decodificador los coeficientes de TNS/TTS se aplican nuevamente al espectro completo, es decir, el espectro de núcleo más el espectro regenerado más los componentes tonales del mapa de tonalidades (ver la Fig. 7e). La aplicación del TTS es necesaria para formar la envolvente temporal del espectro regenerada para que coincida nuevamente con la envolvente de la señal original. Por lo tanto los pre-ecos mostrados se reducen. Además, así se conforma adicionalmente la cuantificación en la señal inferior *freqstart* como es habitual con TNS.

[0128] En los decodificadores de técnicas anteriores, el parcheo espectral en una señal de audio corrompe la correlación espectral en los bordes del parche y con ello deteriora la envolvente temporal de la señal de audio al introducir una dispersión. Por lo tanto, otro beneficio de la realización de relleno de las teselas de IGF en las señales residuales es que, después de la aplicación del filtro de conformado, los bordes de tesela se correlacionan de manera ininterrumpida, para producir una reproducción temporal más fiel de la señal.

[0129] En un codificador de la invención, el espectro que ha experimentado un filtrado de TNS/TFF, el procesamiento de la máscara de tonalidad y la estimación de los parámetros de IGF carece de señales por encima de la frecuencia de inicio excepto por los componentes tonales. Este espectro disperso es codificado a continuación por el codificador de núcleo usando principio de codificación aritmética y de codificación predictiva. Estos componentes codificados, junto con los bits de señalización, forman el tren de bits del audio.

[0130] En la Fig. 2a se ilustra la correspondiente implementación del decodificador. El tren de bits en la Fig. 2a correspondiente a la señal de audio codificada es introducido en el demultiplexor/decodificador que estaría conectado, con respecto a la Fig. 1b, a los bloques 112 y 114. El demultiplexor del tren de bits separa la señal de audio introducida en la primera representación codificada 107 de la Fig. 1b y la segunda representación codificada 109 de la Fig. 1b.

La primera representación codificada que tiene el primer conjunto de primeras porciones espectrales es introducida en el bloque de decodificación de canales conjuntos 204 correspondiente al decodificador en el dominio espectral 112 de la Fig. 1b. La segunda representación codificada es introducida en el decodificador paramétrico 114 no representado en la Fig. 2a y a continuación se introduce en el bloque de IGF 202 correspondiente al generador de frecuencia 116 de la Fig. 1b. El primer conjunto de porciones espectrales requerido para la regeneración de las frecuencias es introducido en el bloque 202 por medio de la línea 203. Además, posteriormente a la decodificación de los canales conjuntos 204 se aplica la decodificación de núcleo específica al bloque de máscara tonal 206 de manera que la salida de la máscara tonal 206 corresponda a la salida del decodificador en el dominio espectral 112. A continuación, se lleva a cabo una combinación a cargo del combinador 208, es decir, una construcción de tramas en la que la salida del combinador 208 tiene ahora el espectro de rango completo, que se encuentra todavía en el dominio filtrado de TNS/TTS. A continuación, en el bloque 210 se lleva a cabo una operación de TNS/TTS inversa usando la información de filtro de TNS/TTF proporcionada por medio de la línea 109, es decir, la información del lado del TTS está incluida preferentemente en la primera representación codificada generada por el codificador en el dominio espectral 106 que puede, por ejemplo, ser un codificador de núcleo directo AAC o USAC, o también puede estar incluida en la segunda representación codificada. A la salida del bloque 210, se proporciona un espectro completo hasta la frecuencia máxima que es la frecuencia de intervalo completo definida por la velocidad de muestreo de la señal introducida original. A continuación se lleva a cabo una conversión de espectro/tiempo en el banco de filtros de síntesis 212 de manera que se obtenga finalmente la señal emitida de audio.

[0131] En la Fig. 3a se ilustra una representación esquemática del espectro. El espectro está subdividido en bandas de factor de escala SCB en las que hay siete bandas de factor de escala SCB1 a SCB7 en el ejemplo ilustrado de la Fig. 3a. Las bandas de factor de escala pueden ser bandas de factor de escala AAC que están definidas en la norma AAC y que tienen una anchura de banda creciente hasta frecuencias superiores, tal como se ilustra esquemáticamente en la Fig. 3a. Se prefiere llevar a cabo el relleno de intervalos no desde el mismo inicio del espectro, es decir, en frecuencias bajas, sino empezar la operación con una frecuencia de inicio ilustrada en 309. Por tanto, la banda de frecuencias de núcleo se extiende desde la frecuencia más baja a la frecuencia de inicio de IGF. Por encima

de la frecuencia de inicio de IGF, se aplica el análisis del espectro para separar los componentes espectrales de alta resolución 304, 305, 306, 307 (el primer conjunto de primeras porciones espectrales) de los componentes de baja resolución representados por el segundo conjunto de segunda porciones espectrales. La Fig. 3a ilustra un espectro que es un ejemplo de introducción en el codificador del dominio espectral 106 o del codificador de canales conjuntos 5 228, es decir, el codificador de núcleo actúa en el intervalo completo, pero codifica una cantidad significativa de valores espectrales cero, es decir, estos valores espectrales cero son cuantificados en cero o puestos a cero antes de la cuantificación o después de la cuantificación. En cualquier caso, el codificador de núcleo actúa en el intervalo completo, es decir, como si el espectro fuera como se ilustra, es decir, el decodificador de núcleo no debe ser necesariamente consciente de ningún relleno de intervalos o codificación del segundo conjunto de segundas porciones espectrales 10 con una resolución espectral más baja.

[0132] Preferentemente, la alta resolución está definida por una codificación de líneas espectrales tales como las líneas de MDCT, mientras que la segunda resolución o baja resolución se decide, por ejemplo, calculando solamente un único valor espectral por banda de factores de escala, en el que una banda de factores de escala cubre 15 varias líneas de frecuencia. Por lo tanto, la segunda resolución baja es, con respecto a su resolución espectral, mucho más baja que la primera resolución, o resolución alta, definida por la codificación en líneas, normalmente aplicada por el codificador de núcleo tal como un codificador de núcleo AAC o USAC.

[0133] En cuanto al factor de escala o cálculo de energía, la situación se ilustra en la Fig. 3b. Debido al hecho 20 de que el codificador es un codificador de núcleo y a que puede haber, aunque no necesariamente, componentes del primer conjunto de porciones espectrales en cada banda, el codificador de núcleo calcula un factor de escala para cada banda no solamente en el intervalo de núcleo por debajo de la frecuencia de inicio IGF 309, sino también por encima de la frecuencia de inicio IGF hasta la máxima frecuencia $f_{IGFstop}$ que es menor o igual que la mitad de la frecuencia de muestreo, es decir, $f_{s/2}$. Por lo tanto, las porciones tonales codificadas 302, 304, 305, 306, 307 de la Fig. 25 3a, y en esta realización junto con los factores de escala SCB1 a SCB7 corresponden a los datos de alta resolución espectral. Los datos espectrales de baja resolución se calculan partiendo de la frecuencia de inicio de IGF y corresponden a los valores de información de energía E_1 , E_2 , E_3 , E_4 , que son transmitidos conjuntamente con los factores de escala SF4 a SF7.

[0134] En particular, cuando el decodificador de núcleo se encuentra en una condición de baja velocidad binaria, puede aplicarse además una operación de relleno de ruido adicional en la banda de núcleo, es decir, inferior en frecuencia que la frecuencia de inicio de IGF, es decir, en las bandas de factor de escala SCB1 a SCB3. En el relleno de ruido, existen varias líneas espectrales adyacentes que deben ser cuantificadas en cero. En el lado del decodificador, estos valores espectrales cuantificados en cero son resintetizados y los valores espectrales 35 resintetizados son ajustados en su magnitud usando una energía de relleno de ruido tal como NF_2 ilustrada en 308 en la Fig. 3b. La energía de relleno de ruido, que puede indicarse en términos absolutos o en términos relativos particularmente con respecto al factor de escala como en USAC, se corresponde con la energía del conjunto de valores espectrales cuantificados en cero. Estas líneas espectrales de relleno de ruido también pueden ser consideradas como un tercer conjunto de terceras porciones espectrales que son regeneradas mediante síntesis directa de relleno de 40 ruido sin ninguna operación de IGF basada en la regeneración de frecuencias en la que se usan teselas de frecuencia procedentes de otras frecuencias para reconstruir teselas de frecuencia usando valores espectrales tomados de un intervalo fuente y la información de energía E_1 , E_2 , E_3 , E_4 .

[0135] Preferentemente, las bandas para las que se calcula la información de energía coinciden con las bandas 45 de factor de escala. En otras realizaciones, se aplica un agrupamiento de valores de información de energía de manera que, por ejemplo, para las bandas de factor de escala 4 y 5, se transmite un único valor de información de energía, pero incluso en esta realización, los límites de las bandas de reconstrucción agrupadas coinciden con los límites de las bandas de los factores de escala. Si se aplican separaciones de banda diferentes, entonces es posible aplicar ciertos recálculos o cálculos de sincronización, y esto puede tener sentido dependiendo de la implementación de que 50 se trate.

[0136] Preferentemente, el codificador en el dominio espectral 106 de la Fig. 1a es un codificador accionado psicoacústicamente tal como se ilustra en la Fig. 4a. Normalmente, como por ejemplo en el caso ilustrado en la norma MPEG2/4 AAC o en la norma MPEG1/2, Capa 3, la señal de audio que debe ser codificada después de haber sido 55 transformada en el intervalo espectral (401 en la Fig. 4a) es enviada a un calculador de factores de escala 400. El calculador de factores de escala está controlado por un modelo psicoacústico que recibe además la señal de audio que se cuantificará o que recibe, como en la norma MPEG1/2 Capa 3 o MPEG AAC, una representación espectral compleja de la señal de audio. El modelo psicoacústico calcula, para cada banda de factores de escala, un factor de escala que representa el umbral psicoacústico. Además, a continuación los factores de escala se ajustan, mediante la 60 cooperación de los bucles de iteración interiores y exteriores bien conocidos o mediante cualquier otro procedimiento de codificación adecuado, de manera que se satisfagan determinadas condiciones de velocidad binaria. A continuación, los valores espectrales que deben ser cuantificados, por una parte, y los factores de escala calculados, por otra parte, son introducidos en un procesador cuantificador 404. En la operación de codificador de audio directa, los valores espectrales que deben ser cuantificados son ponderados por los factores de escala, y, los valores

espectrales ponderados son introducidos a continuación en un cuantificador fijo que normalmente tiene una función de compresión, a intervalos de amplitud superiores. A continuación, a la salida del procesador cuantificador existen índices de cuantificación que son enviados a continuación a un codificador de entropía que normalmente tiene una codificación específica y muy eficiente para un conjunto de índices de cuantificación en cero para valores de frecuencia adyacentes, o, como también se llama en la especialidad, una “tanda” de valores cero.

[0137] Sin embargo, en el codificador de audio de la Fig. 1a, el procesador cuantificador normalmente recibe información acerca de las segundas porciones espectrales del analizador espectral. Por lo tanto, el procesador cuantificador 404 asegura que, en la salida del procesador cuantificador 404, las segundas porciones espectrales identificadas por el analizador espectral 102 sean cero o tengan una representación reconocida por un codificador o decodificador como una representación cero que puede ser codificada de una manera muy eficiente, específicamente cuando existan “tandas” de valores cero en el espectro.

[0138] La Fig. 4b ilustra una implementación del procesador cuantificador. Los valores espectrales de MDCT pueden ser introducidos en un conjunto de bloques cero 410. A continuación, las segundas porciones espectrales ya están fijadas en cero antes de que se lleve a cabo una ponderación por medio de los factores de escala en el bloque 421. En una implementación adicional, no se proporciona bloque 410, aunque la cooperación de la puesta a cero se realiza en el bloque 418 posterior al bloque de ponderación 412. En una implementación adicional más, la operación de la puesta a cero también puede realizarse en un conjunto de bloques cero 422 posterior a una cuantificación en el bloque cuantificador 420. En esta implementación, los bloques 410 y 418 no estarían presentes. En términos generales, se proporciona al menos uno de los bloques 410, 418 y 422 dependiendo de la implementación específica.

[0139] A continuación, a la salida del bloque 422 se obtiene un espectro cuantificado correspondiente al ilustrado en la Fig. 3a. Este espectro cuantificado es introducido a continuación en un codificador de entropía tal como 232 en la Fig. 2b que puede ser un codificador de Huffman o un codificador aritmético tal como se define por ejemplo en la norma USAC.

[0140] La puesta a cero de los bloques 410, 418, 422, que se proporcionan como alternativa entre sí o en paralelo está controlada por el analizador espectral 424. El analizador espectral comprende preferentemente cualquier implementación de un detector de tonalidad bien conocido o comprende cualquier tipo diferente de detector operable para separar un espectro en componentes que se codificarán con una resolución alta y componentes que se codificarán con una resolución baja. Otros algoritmos de este tipo implementados en el analizador espectral pueden ser un detector de la actividad de voz, un detector de ruidos o un detector de voz o cualquier otro detector que detecte, dependiendo de información espectral o de los metadatos asociados, los requisitos de resolución para diferentes porciones espectrales.

[0141] La Fig. 5a ilustra una implementación preferida del convertidor de espectro-tiempo 100 de la Fig. 1a como por ejemplo se implementa en AAC o USAC. El convertidor de espectro-tiempo 100 comprende un administrador de ventanas 502 controlado por un detector de transitorios 504. Cuando el detector de transitorios 504 detecta un transitorio, entonces se señala la conmutación de ventanas largas a ventanas cortas al administrador de ventanas. A continuación el administrador de ventanas 502 calcula, para los bloques superpuestos, tramas de ventana, en las que cada trama de ventana tiene normalmente N valores, por ejemplo, 2.048 valores. A continuación se lleva a cabo una transformación dentro del transformador de bloque 506, y este transformador de bloque normalmente proporciona además una pérdida, de manera que tiene lugar una pérdida/transformada combinada de manera que obtenga una trama espectral de N valores tales como valores espectrales MDCT. Por lo tanto, para una operación de ventana larga, la trama en la entrada del bloque 506 comprende dos valores N tales como 2.048 valores y una trama espectral tiene entonces 1.024 valores. Sin embargo, entonces se lleva a cabo una conmutación a bloques cortos, cuando se obtienen ocho bloques cortos en los que cada bloque corto representa 1/8 de valores en el dominio del tiempo en ventana en comparación con una ventana larga y cada bloque espectral tiene 1/8 de valores espectrales en comparación con un bloque largo. De esta manera, cuando se combina esta pérdida con una operación de superposición del 50 % del administrador de ventanas, el espectro es una versión muestreada críticamente de la señal de audio en el dominio del tiempo 99.

[0142] A continuación, se hace referencia a la Fig. 5b en la que se ilustra una implementación específica del regenerador de frecuencias 116 y del convertidor de espectro/tiempo 118 de la Fig. 1b, o de la operación combinada de los bloques 208, 212 de la Fig. 2a. En la Fig. 5b, se considera una banda de reconstrucción específica tal como una banda de factores de escala 6 de la Fig. 3a. La primera porción espectral en esta banda de reconstrucción, es decir, la primera porción espectral 306 de la Fig. 3a es introducida en el elemento de construcción de tramas/bloque de elemento de ajuste 510. Por otra parte, también se introduce una segunda porción espectral reconstruida para la banda de factores de escala 6 en el elemento de construcción/ajuste de tramas 510. Además, en el bloque 510 se introduce también la información de energía tal como E_3 de la Fig. 3b para una banda de factores de escala 6. La segunda porción espectral reconstruida en la banda de reconstrucción ya ha sido generada por el relleno de teselas de frecuencia usando un intervalo fuente y a continuación la banda de reconstrucción corresponde al intervalo objeto. Seguidamente se lleva a cabo un ajuste de energía de la trama para finalmente obtener a continuación la trama reconstruida completa que tiene los N valores como por ejemplo los obtenidos a la salida del combinador 208 de la

Fig. 2a. A continuación, en el bloque 512, se lleva a cabo una transformada/interpolación de bloque inverso para obtener 248 valores en el dominio del tiempo, por ejemplo 124 valores espectrales, en la entrada del bloque 512. A continuación se lleva a cabo una operación de gestión de ventana de síntesis en el bloque 514 que nuevamente es controlada por una indicación de ventana larga/ventana corta transmitida como información secundaria en la señal de audio codificada. A continuación, en el bloque 516 se lleva a cabo una operación de superposición/suma con una trama de tiempo anterior. Preferentemente, el MDCT aplica una superposición del 50 % de manera que, para cada nueva trama de tiempo de 2N valores, finalmente se emitan N valores en el dominio del tiempo. Se prefiere ampliamente una superposición del 50 % debido a que proporciona un muestreo crítico y un cruce continuo de una trama con la trama siguiente debido a la operación de superposición/suma en el bloque 516.

[0143] Tal como se ilustra en 301 en la Fig. 3a, es posible aplicar adicionalmente una operación de relleno de ruido no solamente por debajo de la frecuencia de inicio de IGF, sino también por debajo de la frecuencia de inicio de IGF tal como para la banda de reconstrucción considerada que coincide con la banda de factores de escala 6 de la Fig. 3a. A continuación, también es posible introducir valores espectrales de relleno de ruido en el elemento de construcción/ajuste de tramas 510 y también es posible aplicar valores espectrales de relleno de ruido de este bloque o los valores espectrales de relleno de ruido ya pueden estar ajustados usando la energía de relleno de ruido antes de ser introducidos en el elemento de construcción/ajuste de tramas 510.

[0144] Preferentemente se puede aplicar una operación de IGF, es decir, una operación de relleno de teselas de frecuencia usando valores espectrales procedentes de otras porciones en el espectro completo. De esta manera, puede aplicarse una operación de relleno de teselas espectral no sólo en la banda alta por encima de la frecuencia de inicio de IGF sino que también se puede aplicar en la banda baja. Además, el relleno de ruido sin el relleno de teselas de frecuencia también puede ser aplicado no solamente por debajo de la frecuencia de inicio de IGF sino también por encima de la frecuencia de inicio de IGF. Sin embargo, se ha descubierto que es posible obtener una codificación de audio de alta calidad y muy eficiente cuando la operación de relleno de ruido se limita a un intervalo de frecuencia inferior a la frecuencia de inicio de IGF y cuando la operación de relleno de teselas de frecuencia se limita al intervalo de frecuencia por encima del intervalo de frecuencia de inicio de IGF, tal como se ilustra en la Fig. 3a.

[0145] Preferentemente, las teselas objeto (TT, target tiles) (que tienen frecuencias superiores a la frecuencia de inicio de IGF) están vinculadas a límites de bandas de factores de escala del codificador de velocidad completa. Las teselas fuente (ST, source tiles), de las cuales se toma la información, es decir, para frecuencias inferiores a la frecuencia de inicio de IGF, no están vinculadas con límites de bandas de factores de escala. El tamaño de las ST debería corresponder al tamaño de la TT asociado. Esto se ilustra mediante el siguiente ejemplo. TT [0] tiene una longitud de 10 segmentos de MDCT. Esto corresponde exactamente a la longitud de dos SCB posteriores (tales como 4 + 6). Entonces, todas las posibles ST que se van a correlacionar con TT [0], tienen una longitud también de 10 segmentos. Una segunda tesela objeto TT [1] que es adyacente a TT [0] tiene una longitud de 15 segmentos I (SCB tiene una longitud de 7 + 8). Entonces, la ST para esta tiene una longitud de 15 segmentos en lugar de 10 segmentos como en TT [0].

[0146] Si sucediera que no se puede encontrar una TT para una ST con la longitud de la tesela objeto (cuando por ejemplo la longitud de TT es mayor que el intervalo fuente disponible), entonces no se calcula una correlación y el intervalo fuente se copia un número de veces en esta TT (la copia se realiza una después de la otra de modo que una línea de frecuencia para la frecuencia más baja de la segunda copia sigue inmediatamente, en la frecuencia, a la línea de frecuencia para la frecuencia más alta de la primera copia), hasta que la tesela objeto TT está completamente llena.

[0147] A continuación se hace referencia a la Fig. 5c que ilustra una realización adicional preferida del regenerador de frecuencias 116 de la Fig. 1b o del bloque de IGF 202 de la Fig. 2a. El bloque 522 es un generador de teselas de frecuencia que recibe no solamente un ID de banda objeto, sino también un ID de banda fuente. A modo de ejemplo, en el lado del codificador se ha determinado que la banda de factores de escala 3 de la Fig. 3a es adecuado para reconstruir la banda de factores de escala 7. Por lo tanto, el ID de la banda fuente sería 2 y el ID de la banda objeto sería 7. Basándose en esta información, el generador de teselas de frecuencia 522 aplica una copia o una operación de relleno de armónicos de las teselas o cualquier otra operación de relleno de teselas con el fin de generar la segunda porción en bruto con los componentes espectrales 523. La segunda porción en bruto de componentes espectrales tiene una resolución de frecuencia idéntica a la resolución de frecuencias incluida en el primer conjunto de primeras porciones espectrales.

[0148] A continuación, la primera porción espectral de la banda de reconstrucción tal como 307 de la Fig. 3a es introducida en un elemento de construcción de tramas 524 y la segunda porción en bruto 523 es también introducida en el elemento de construcción de tramas 524. A continuación, la trama reconstruida es ajustada por el elemento de ajuste 526 usando un factor de ganancia para la banda de reconstrucción calculada por el calculador de factores de ganancia 528. Sin embargo, es importante señalar que en la primera porción espectral en las tramas no influye el elemento de ajuste 526, sino que solamente la segunda porción en bruto para en la trama de reconstrucción se ve influida por el elemento de ajuste 526. Para este fin, el calculador de factores de ganancia 528 analiza la banda fuente o la segunda porción en bruto 523 y adicionalmente analiza la primera porción espectral en la banda de reconstrucción con el fin de encontrar finalmente el factor de ganancia correcto 527 de manera que la energía de la trama ajustada

emitida por el elemento de ajuste 526 tenga la energía E_4 cuando se considera la banda de factores de escala 7.

[0149] En este contexto, es muy importante evaluar la precisión de la reconstrucción de alta frecuencia de la presente invención en comparación con HE-AAC. Esto se explica con respecto a la banda de factores de escala 7 de la figura 3a. Se supone que un codificador de la técnica anterior tal como se ilustra en la Fig. 13a puede detectar la porción espectral 307 para codificar con una alta resolución como "armónicos ausentes". Entonces, la energía de este componente espectral se transmite junto con una información de envolvente espectral para la banda de reconstrucción tales como banda de factores de escala 7 al decodificador. A continuación, el decodificador puede recrear el armónico ausente. Sin embargo, el valor espectral, en el que el armónico ausente 307 se puede reconstruir mediante el decodificador de la técnica anterior de la Fig. 13b, puede estar en el medio de la banda 7 a una frecuencia indicada por la frecuencia de reconstrucción 390. Por lo tanto, la presente invención evita un error de frecuencia 391 que sería introducido por el decodificador de la técnica anterior de la Fig. 13d.

[0150] En una implementación, el analizador espectral también se implementa para calcular las semejanzas entre las primeras porciones espectrales y las segundas porciones espectrales y para determinar, basándose en las similitudes calculadas, para una segunda porción espectral en un intervalo de reconstrucción, una primera porción espectral que coincide con la segunda porción espectral en la medida de lo posible. Entonces, en esta implementación de la variable intervalo de origen/intervalo de destino, el codificador paramétrico introducirá adicionalmente en la segunda representación codificada una información coincidente que indica para cada intervalo de destino un intervalo de origen correspondiente. En el lado del decodificador, esta información puede ser usada a continuación por un generador de tesela de frecuencia 522 de la Fig. 5c que ilustra una generación de una segunda porción en bruto 523 basada en un ID de banda fuente y un ID de banda objeto.

[0151] Por otra parte, tal como se ilustra en la Fig. 3a, el analizador espectral está configurado para analizar la representación espectral hasta una frecuencia de análisis máxima que es solamente una parte pequeña inferior a la mitad de la frecuencia de muestreo, y preferentemente es de al menos un cuarto de la frecuencia de muestreo o normalmente superior.

[0152] Como se ilustra, el codificador opera sin submuestreo y el decodificador opera sin sobremuestreo. Dicho de otro modo, el codificador de audio en el dominio espectral está configurado para generar una representación espectral que tiene una frecuencia de Nyquist definida por la velocidad de muestreo de la señal de audio introducida originalmente.

[0153] Además, como se ilustra en la Fig. 3a, el analizador espectral está configurado para analizar la representación espectral empezando con una frecuencia de inicio de relleno de intervalos y terminando con una frecuencia máxima representada por una frecuencia máxima incluida en la representación espectral, en el que una porción espectral que se extiende desde una frecuencia mínima hasta la frecuencia de inicio de relleno de intervalos forma parte del primer conjunto de porciones espectrales y en el que otra porción espectral tal como 304, 305, 306, 307 que tiene valores de frecuencias superiores a la frecuencia de relleno de intervalos se incluye adicionalmente en el primer conjunto de primeras porciones espectrales.

[0154] Como se destacó a grandes rasgos, el decodificador de audio en el dominio espectral 112 está configurado de manera que una frecuencia máxima representada por un valor espectral en la primera representación decodificada es igual a una frecuencia máxima incluida en la representación en el tiempo que tiene una velocidad de muestreo en la que el valor espectral para la frecuencia máxima en el primer conjunto de primeras porciones espectrales es cero o diferente de cero. En cualquier caso, para esta frecuencia máxima en el primer conjunto de componentes espectrales existe un factor de escala para la banda de factores de escala, que es generado y transmitido independientemente de si la totalidad de los valores espectrales en esta banda de factores de escala está ajustada en cero o no, como se expuso en el contexto de las Fig. 3a y 3b.

[0155] Por lo tanto, la invención es ventajosa dado que con respecto a otras técnicas paramétricas incrementa la eficiencia de la compresión, por ejemplo, la sustitución del ruido y el relleno de ruido (estas técnicas sirven exclusivamente para la representación eficiente del contenido de señales locales similares de ruido), la invención permite una reproducción exacta de las frecuencias de los componentes tonales. Hasta la fecha, en ninguna técnica del estado de la técnica se aborda la representación paramétrica eficiente de un contenido arbitrario de las señales mediante el relleno de intervalos espectral sin la restricción de una división a priori fija en banda baja (LF, low band) y banda alta (HF, high band).

[0156] Las realizaciones del sistema de la invención mejoran los enfoques del estado de la técnica y por lo tanto proporcionan alta eficiencia de compresión, ninguna o sólo una pequeña molestia perceptual y anchura de banda de audio completa incluso para velocidades binarias bajas.

[0157] El sistema general consiste en codificación de núcleo de banda completa, relleno de intervalos inteligente (relleno de tesela o relleno de ruido), análisis de partes tonales en el núcleo seleccionado por máscara tonal, codificación del par estéreo conjunto para la banda completa, que incluye TNS de relleno de teselas

blanqueamiento espectral de teselas en el intervalo IGF

[0158] Una primera etapa hacia un sistema más eficiente consiste en eliminar la necesidad de transformar los datos espectrales en un segundo dominio de transformación diferente del codificador de núcleo. Como la mayoría de los códecs de audio, tales como AAC, por ejemplo, el uso de la MDCT como transformación básica es útil para realizar también el BWE en el dominio MDCT. Un segundo requisito para el sistema de BWE sería la necesidad de conservar la retícula tonal mediante la cual se conservan incluso los componentes tonales HF y la calidad del audio codificada es por lo tanto superior a los sistemas existentes. Para atender los requisitos mencionados anteriormente para un esquema BWE, se propone un nuevo sistema llamado relleno de intervalos inteligente (IGF). La Fig. 2b muestra el diagrama de bloques del sistema propuesto en el lado del codificador y la Fig. 2a muestra el sistema en el lado del decodificador.

[0159] A continuación se exponen y definen otras características opcionales del primer procesador codificador en el dominio de la frecuencia de banda completa y del procesador decodificador en el dominio de la frecuencia de banda completa que incluyen la operación de relleno de intervalos, que pueden ser implementadas por separado o conjuntamente.

[0160] En particular, el decodificador en el dominio espectral 112 correspondiente al bloque 1122a está configurado para emitir una secuencia de tramas decodificadas de valores espectrales, siendo una trama decodificada la primera representación decodificada, en el que la trama comprende valores espectrales para el primer conjunto de porciones espectrales e indicadores cero para las segundas porciones espectrales. El aparato para decodificar comprende además un combinador 208. Los valores espectrales son generados por un regenerador de frecuencias para el segundo conjunto de segundas porciones espectrales, en el que los dos, el combinador y el generador de frecuencias, están incluidos dentro del bloque 1122b. Por lo tanto, mediante la combinación de las segundas porciones espectrales y de las primeras porciones espectrales se obtiene una trama espectral reconstruida que comprende valores espectrales para el primer conjunto de las primeras porciones espectrales y el segundo conjunto de porciones espectrales, y el convertidor de espectro-tiempo 118 correspondiente al bloque de IMDCT 1124 en la Fig. 14b convierte seguidamente la trama espectral reconstruida en la representación en el tiempo.

[0161] Como se destaca, el convertidor de espectro-tiempo 118 o 1124 está configurado para realizar una transformada de coseno discreta modificada inversa 512, 514, y comprende además una fase de superposición-suma 516 para superponer y añadir tramas posteriores en el dominio del tiempo.

[0162] En particular, el decodificador de audio en el dominio espectral 1122a está configurado para generar la primera representación decodificada de manera que la primera representación decodificada tenga una frecuencia de Nyquist que define una velocidad de muestreo que es igual a una velocidad de muestreo de la representación en el tiempo generada por el convertidor de espectro-tiempo 1124.

[0163] Por otra parte, el decodificador 1112 o 1122a está configurado para generar la primera representación decodificada de manera que una primera porción espectral 306 está colocada con respecto a la frecuencia entre dos segundas porciones espectrales 307a, 307b.

[0164] Una frecuencia máxima representada por un valor espectral para la máxima frecuencia en la primera representación decodificada es igual a la máxima frecuencia incluida en la representación en el tiempo generada por el convertidor de espectro-tiempo, en el que el valor espectral para la frecuencia máxima en la primera representación es cero o diferente de cero.

[0165] Además, tal como se ilustra en la Fig. 3, la primera porción de la señal de audio codificada comprende además una representación codificada de un tercer conjunto de porciones espectrales que se reconstruirá mediante el relleno de ruido, y el primer procesador de decodificación 1120 comprende adicionalmente un relleno de ruido incluido en el bloque 1122b para extraer información de relleno de ruido 308 de una representación codificada del tercer conjunto de terceras porciones espectrales y para aplicar una operación de relleno de ruido en el conjunto de terceras porciones espectrales sin usar una primera porción espectral en un intervalo de frecuencia diferente.

[0166] Además, el decodificador de audio en el dominio espectral 112 está configurado para generar la primera representación decodificada que tiene las primeras porciones espectrales, siendo los valores de frecuencia mayores que la frecuencia que es igual a una frecuencia en el medio del intervalo de frecuencias comprendido por la salida de la representación en el tiempo por el convertidor de espectro-tiempo 118 o 1124.

[0167] Por otra parte, el analizador espectral o el analizador de banda completa 604 están configurados para analizar la representación generada por el convertidor de tiempo-frecuencia 602 con el fin de determinar un primer conjunto de primeras porciones espectrales que deben ser codificadas con la primera alta resolución espectral y el segundo conjunto, diferente, de segundas porciones espectrales que deben ser codificadas con una segunda resolución espectral que es inferior a la primera resolución espectral, y mediante el analizador espectral, se determina una primera porción espectral 306, con respecto a la frecuencia, entre dos porciones espectrales segundas en la Fig.

3 en 307a y 307b.

[0168] En particular, el analizador espectral está configurado para analizar la representación espectral hasta una máxima frecuencia de análisis que es de al menos un cuarto de una frecuencia de muestreo de la señal de audio.

[0169] En particular, el codificador de audio en el dominio espectral está configurado para procesar una secuencia de tramas de valores espectrales para una cuantificación y codificación de la entropía, en el que, en una trama, los valores espectrales del segundo conjunto de segundas porciones se ajustan a cero, o en el que, en la trama, los valores espectrales del primer conjunto de primeras porciones espectrales y el segundo conjunto de la segunda porción de las porciones espectrales están presentes, y en el que, durante el procesamiento posterior, los valores espectrales en el segundo conjunto de porciones espectrales se ajustan a cero, como se ilustra a modo de ejemplo en 410, 418, 422.

[0170] El codificador de audio en el dominio espectral está configurado para generar una representación espectral que tiene una frecuencia de Nyquist definida por la velocidad de muestreo de la señal de audio introducida o por la primera porción de la señal de audio procesada por el primer procesador de codificación que opera en el dominio de la frecuencia.

[0171] El codificador de audio en el dominio espectral 606 está configurado además de manera que proporcione la primera representación codificada de modo que, para una trama de una señal de audio muestreada, la representación codificada comprende el primer conjunto de primeras porciones espectrales y el segundo conjunto de segundas porciones espectrales, en el que los valores espectrales en el segundo conjunto de porciones espectrales están codificados como valores cero o de ruido.

[0172] El analizador de banda completa 604 o 102 está configurado para analizar la representación espectral empezando con la frecuencia de inicio de relleno de intervalos 209 y terminando con una frecuencia máxima $f_{\max}$ representada por una frecuencia máxima incluida en la representación espectral y una porción espectral que se extiende desde una frecuencia mínima hasta la frecuencia de inicio de relleno de intervalos 309 forma parte del primer conjunto de primeras porciones espectrales.

[0173] En particular, el analizador está configurado para aplicar un procesamiento de máscara tonal en al menos una porción de la representación espectral de manera que los componentes tonales y los componentes no tonales estén separados entre sí, en el que el primer conjunto de porciones espectrales comprende los componentes tonales y en el que el segundo conjunto de las segundas porciones espectrales comprende los componentes no tonales.

[0174] Si bien la presente invención ha sido descrita en el contexto de diagramas de bloques en que los bloques presentan componentes de hardware reales o lógicos, la presente invención también puede implementarse mediante un procedimiento implementado por ordenador. En este último caso, los bloques representan etapas de procedimiento correspondientes en las que dichas etapas representan las funcionalidades realizadas por los bloques de hardware lógicos o físicos correspondientes.

[0175] Si bien algunos aspectos han sido descritos en el contexto de un aparato, es evidente que estos aspectos también representan una descripción del procedimiento correspondiente, en el que un bloque o dispositivo corresponde a una etapa de procedimiento o a una característica de una etapa de procedimiento. De manera análoga, los aspectos descritos en el contexto de una etapa de procedimiento también representan una descripción de un bloque o elemento o aspecto correspondiente de un aparato correspondiente. Algunas de las etapas de procedimiento, o todas ellas, pueden ser ejecutadas mediante un aparato de hardware, como por ejemplo un microprocesador, un ordenador programable o un circuito electrónico. En algunas realizaciones, algunas de las etapas de procedimiento más importantes, o varias de ellas, pueden ser ejecutadas mediante un aparato de este tipo.

[0176] La señal transmitida o codificada según la invención puede ser almacenada en un medio de almacenamiento digital o puede ser transmitida en un medio de transmisión tal como un medio de transmisión inalámbrica o un medio de transmisión por cable tal como Internet.

[0177] Dependiendo de ciertos requisitos de implementación, las realizaciones de la invención pueden implementarse en hardware o en software. La implementación puede realizarse usando un medio de almacenamiento digital, por ejemplo, un disquete, un DVD, un Blu-Ray, un CD, un ROM, una PROM, y EPROM, una EEPROM o una memoria flash, que tienen almacenadas señales de control legibles electrónicamente, que cooperan (o pueden cooperar) con un sistema informático programable de manera que se lleve a cabo el procedimiento respectivo. Por tanto, el medio de almacenamiento digital puede ser legible por ordenador.

[0178] Algunas realizaciones según la invención comprenden un soporte de datos que tiene señales de control legibles electrónicamente, que pueden cooperar con un sistema informático programable, de manera que se lleve a cabo uno de los procedimientos descritos en la presente memoria.

- [0179]** En términos generales, las realizaciones de la presente invención pueden ser implementadas como un producto de programa informático con un código de programa, siendo el código de programa operativo para realizar uno de los procedimientos cuando el producto de programa informático es ejecutado en un ordenador. El código de programa puede, por ejemplo, ser almacenado en un soporte legible por máquina.
- [0180]** Otras realizaciones comprenden el programa informático para realizar uno de los procedimientos descritos en la presente, almacenado en un soporte legible por máquina.
- 10 **[0181]** Por lo tanto, dicho de otro modo, una realización del procedimiento de la invención consiste en un programa informático que tiene un código de programa para realizar uno de los procedimientos descritos en la presente memoria, cuando el programa informático es ejecutado en un ordenador.
- [0182]** Por lo tanto, una realización adicional del procedimiento de la invención consiste en un soporte de datos
15 (o en un medio de almacenamiento no transitorio tal como un medio de almacenamiento digital, o un medio legible por ordenador), que comprende, grabado en él, el programa informático para realizar uno de los procedimientos descritos en la presente memoria. El soporte de datos, el medio de almacenamiento digital o el medio grabado son normalmente tangibles y/o no transitorios.
- 20 **[0183]** Una realización adicional del procedimiento de la invención es por lo tanto un tren de datos o una secuencia de señales que representan el programa informático para realizar uno de los procedimientos descritos en la presente memoria. El tren de datos o la secuencia de señales pueden, por ejemplo, estar configurados para ser transferidos por medio de una conexión de datos, por ejemplo, por Internet.
- 25 **[0184]** Una realización adicional comprende un medio de procesamiento, por ejemplo, un ordenador o un dispositivo lógico programable, configurado o adaptado para llevar a cabo uno de los procedimientos descritos en la presente memoria.
- [0185]** Una realización adicional se refiere a un ordenador que tiene instalado el programa informático para
30 ejecutar uno de los procedimientos descritos en la presente memoria.
- [0186]** Una realización adicional según la invención se refiere a un aparato o sistema configurados para transferir (por ejemplo, por medios ópticos o electrónicos) un programa informático para ejecutar uno de los procedimientos descritos en la presente memoria en un receptor. El receptor puede ser, por ejemplo, un ordenador,
35 un dispositivo móvil, un dispositivo de memoria o similar. El aparato o sistema puede comprender, por ejemplo, un servidor de archivos para transferir el programa informático al receptor.
- [0187]** En algunas realizaciones puede usarse un dispositivo lógico programable (por ejemplo, una matriz de puertas programables de campo) para realizar algunas o todas las funcionalidades de los procedimientos descritos en
40 la presente memoria. En algunas realizaciones, la matriz de puertas programables de campo puede cooperar con un microprocesador con el fin de llevar a cabo uno de los procedimientos descritos en la presente memoria. En términos generales, los procedimientos son ejecutados preferentemente mediante cualquier aparato de hardware.
- [0188]** Las realizaciones descritas anteriormente son meramente ilustraciones de los principios de la presente
45 invención. Debe entenderse que para los expertos en la materia serán evidentes modificaciones y variaciones en las disposiciones y detalles descritos en la presente memoria. Por lo tanto, se pretende que la invención esté limitada solamente por el alcance de las reivindicaciones de patente y no por los detalles específicos presentados a modo de descripción y explicación de las realizaciones presentadas en la presente memoria.

REIVINDICACIONES

1. Codificador de audio para codificar una señal de audio, que comprende:

5 un primer procesador de codificación (600) para codificar una primera porción de la señal de audio en un dominio de la frecuencia, teniendo la primera porción de la señal de audio asociada consigo una frecuencia de muestreo, en el que el primer procesador de codificación (600) comprende:

10 un convertidor de frecuencia y de tiempo (602) para convertir la primera porción de la señal de audio en una representación de un dominio de la frecuencia que tiene líneas espectrales hasta una frecuencia máxima de la primera porción de la señal de audio;

15 un analizador (604) para analizar la representación en el dominio de la frecuencia hasta la frecuencia máxima para determinar primeras porciones espectrales para codificar con una primera resolución espectral y segundas porciones espectrales para codificar con una segunda resolución espectral, siendo la segunda resolución espectral inferior a la primera resolución espectral, en el que el analizador (604) está configurado para determinar una primera porción espectral (306) a partir de las primeras porciones espectrales, estando la primera porción espectral situada, con respecto a la frecuencia, entre dos segundas porciones espectrales (307a, 307b) a partir de las segundas porciones espectrales, en el que la frecuencia máxima es inferior o igual a la mitad de la frecuencia de muestreo y al menos una
20 cuarta parte de la frecuencia de muestreo o superior;

25 un codificador espectral (606) para codificar las primeras porciones espectrales con la primera resolución espectral y para codificar las segundas porciones espectrales con la segunda resolución espectral, en el que el codificador espectral comprende un codificador paramétrico para calcular información de envolvente espectral que tiene la segunda resolución espectral a partir de las segundas porciones espectrales;

un segundo procesador de codificación (610) para codificar una segunda porción de la señal de audio diferente en el dominio del tiempo, en el que el segundo procesador codificador (610) comprende:

30 un convertidor de velocidad de muestreo (900) para convertir la segunda porción de la señal de audio en una representación de velocidad de muestreo inferior, siendo la velocidad de muestreo inferior menor que la velocidad de muestreo de la señal de audio, en el que la representación de la velocidad de muestreo inferior no incluye la banda alta de la señal de audio;

35 un codificador de banda baja en el dominio del tiempo (910) para codificar en el dominio del tiempo la representación de la velocidad de muestreo inferior; y

40 un codificador de extensión de anchura de banda en el dominio del tiempo (920) para codificar paramétricamente la banda alta de la señal de ruido;

un controlador (620) configurado para analizar la señal de audio y para determinar qué porción de la señal de audio es la primera porción de la señal de audio codificada en el dominio de la frecuencia y qué porción de la señal de audio es la segunda porción de la señal de audio codificada en el dominio del tiempo; y

45 un conformador de señales codificadas (630) para formar una señal de audio codificada que comprende una primera porción de la señal codificada para la primera porción de la señal de audio y una segunda porción de la señal codificada para la segunda porción de señales de audio.

2. Codificador de audio según la reivindicación 1, que comprende además:

50 un preprocesador (1000) configurado para preprocesar la primera porción de la señal de audio y la segunda porción de la señal de audio diferente,

en el que el preprocesador comprende:

55 un analizador de predicción (1002) para determinar coeficientes de predicción; y

en el que el segundo procesador de codificación comprende:

60 un cuantificador del coeficiente de predicción (1010) para generar una versión cuantificada de los coeficientes de predicción; y

un codificador de entropía para generar una versión codificada de los coeficientes de predicción cuantificados,

65 en el que el conformador de señales codificadas (630) está configurado para introducir la versión codificada en la señal

de audio codificada.

3. Codificador de audio según las reivindicaciones 1 o 2,

5 en el que un preprocesador (1000) comprende un remuestreador (1004) para remuestrear la señal de audio a una velocidad de muestreo del segundo procesador de codificación; y

en el que un analizador de predicción está configurado para determinar los coeficientes de predicción usando una señal de audio remuestreada; o

10

en el que el preprocesador (1000) comprende además una fase de análisis de predicción a largo plazo (1006) para determinar uno o más parámetros de predicción a largo plazo para la primera porción de la señal de audio.

4. Codificador de audio según una de las reivindicaciones precedentes, que comprende además un
15 procesador cruzado (700) para calcular, a partir de la representación espectral codificada de la primera porción de la señal de audio, los datos de inicialización del segundo procesador de codificación (610), de modo que el segundo procesador de codificación (610) se inicializa para codificar la segunda porción de la señal de audio diferente inmediatamente después de la primera porción de la señal de audio en el tiempo en la señal de audio.

20 5. Codificador de audio según la reivindicación 4, en el que el procesador cruzado (700) comprende:

un decodificador espectral (701) para calcular una versión decodificada de la primera porción de la señal codificada;

una fase de retardo (707) para introducir una versión retardada de la versión decodificada en una fase de
25 desacentuación (617) del segundo procesador de codificación para la inicialización;

un bloque de filtrado para el análisis de los coeficientes de predicción ponderados (708) para introducir una salida del filtro en un determinador de libro de código (613) del segundo procesador de codificación (610) para la inicialización;

30 una fase de filtrado de análisis (706) para filtrar la versión decodificada o una versión preacentuada (709) y para introducir un residuo de filtro en un determinador de libro de código adaptativo (612) del segundo procesador de codificación para la inicialización; o

un filtro de preacentuación (709) para filtrar la versión decodificada y para introducir una versión retardada o
35 preacentuada en una fase de filtrado de síntesis (616) del segundo procesador de codificación (610) para la inicialización.

6. Codificador de audio según una de las reivindicaciones precedentes,

40 en el que el analizador (604) está configurado para realizar un conformado de tesela temporal o análisis del conformado de ruido temporal o una operación de ajuste a cero de los valores espectrales en las segundas porciones espectrales,

en el que el primer procesador de codificación (600) está configurado para realizar un conformado (606a) de los valores
45 espectrales de las primeras porciones espectrales usando coeficientes de predicción (1010) deducidos de la primera porción de la señal de audio, y en el que el primer procesador de codificación (600) está configurado además para realizar una cuantificación y una operación de codificación de la entropía (606b) de valores espectrales conformados de las primeras porciones espectrales, y

50 en el que los valores espectrales de las segundas porciones espectrales se ajustan a cero.

7. Codificador de audio según la reivindicación 6, que comprende además un procesador cruzado (700), en el que el procesador cruzado (700) comprende:

55 un conformador de ruidos (703) para conformar valores espectrales cuantificados de las primeras porciones espectrales usando coeficientes LPC (1010) deducidos de la primera porción de la señal de audio;

un decodificador espectral (704, 705) para la decodificación de las porciones espectrales conformadas espectralmente de la primera porción espectral con una resolución espectral alta y para la síntesis de las segundas porciones
60 espectrales usando una representación paramétrica de las segundas porciones espectrales y al menos una primera porción espectral decodificada para obtener una representación espectral decodificada;

un convertidor de frecuencia-tiempo (702) para convertir la representación espectral codificada en un dominio del tiempo con el fin de obtener una primera porción de la señal de audio decodificada, en el que una velocidad de
65 muestreo asociada con la primera porción de la señal de audio decodificada es diferente de una velocidad de muestreo

de la señal de audio, y una velocidad de muestreo asociada con una señal de salida del convertidor de frecuencia-tiempo (702) es diferente de una velocidad de muestreo de la señal de audio introducida en el convertidor de frecuencia-tiempo (602).

5 8. Codificador de audio según una de las reivindicaciones precedentes,

en el que el segundo procesador de codificación comprende al menos un bloque de los siguientes grupos de bloques:

un filtro de análisis de predicción (611);

10 una fase del libro de código adaptativo (612);

una fase del libro de código innovador (614);

un estimador (613) para estimar una entrada del libro de código innovador;

una fase de codificación ACELP/ganancia (615);

una fase de filtrado de síntesis de predicción (616);

15 una fase de desacentuación (617); y

una fase de análisis posfiltro de bajos (618).

9. Codificador de audio según una de las reivindicaciones precedentes,

20 en el que el segundo procesador de codificación tiene una segunda velocidad de muestreo asociada;

en el que el primer procesador de codificación tiene asociada con él una primera velocidad de muestreo que es mayor que la segunda velocidad de muestreo, en el que el codificador de audio comprende además un procesador cruzado (700) para calcular, a partir de la representación espectral codificada de la primera porción de la señal de audio, datos de inicialización del segundo procesador de codificación, en el que el procesador cruzado comprende un convertidor de frecuencia-tiempo (702) para generar una señal en el dominio del tiempo con la segunda velocidad de muestreo;

en el que el convertidor de frecuencia-tiempo (702) comprende:

30 un selector (726) para seleccionar una porción baja de un espectro introducido en el convertidor de frecuencia-tiempo de acuerdo con una relación entre la primera velocidad de muestreo y la segunda velocidad de muestreo, siendo la relación menor que 1;

un procesador de transformadas (720) que tiene una longitud de transformada menor que una longitud de transformada del convertidor de frecuencia-tiempo (602); y

un administrador de ventanas de síntesis (712) para aplicar una ventana usando una ventana que tiene un número menor de coeficientes de ventana en comparación con una ventana usada por el convertidor de tiempo-frecuencia (602).

40

10. Decodificador de audio para decodificar una señal de audio codificada, que comprende:

un primer procesador de decodificación (1120) para decodificar una primera porción de una señal de audio codificada en un dominio de la frecuencia, comprendiendo el primer procesador de decodificación (1120):

45

un decodificador espectral (1122) para decodificar las primeras porciones espectrales con una resolución espectral alta y para sintetizar las segundas porciones espectrales usando una representación paramétrica de las segundas porciones espectrales y al menos una primera porción espectral decodificada para obtener una representación espectral decodificada, en el que el decodificador espectral (1122) está configurado para generar la representación decodificada de modo que una primera porción espectral (306) se coloca con respecto a la frecuencia entre dos segundas porciones espectrales (307a, 307b); y

50

la representación espectral decodificada se extiende hasta una frecuencia máxima de una representación en el tiempo de una señal de audio decodificada, siendo un valor espectral para la frecuencia máxima igual a cero o diferente de

55

cero; y

un convertidor de frecuencia-tiempo (1124) para convertir la representación espectral decodificada en un dominio del tiempo para obtener una primera porción de la señal de audio decodificada;

60 un segundo procesador de decodificación (1140) para decodificar una segunda porción de la señal de audio codificada en el dominio del tiempo para obtener una segunda porción de la señal de audio decodificada, en el que el segundo procesador decodificador comprende:

un decodificador de banda baja en el dominio del tiempo (1200), para decodificar y obtener una señal en el dominio del tiempo de banda baja;

65

un sobremuestreador (1210) para sobremuestrear la señal en el dominio del tiempo de banda baja con el fin de obtener una señal en el dominio del tiempo de banda baja sobremuestreada;

- 5 un decodificador de extensión de anchura de banda en el dominio del tiempo (1220) para sintetizar una banda alta de una señal de salida en el dominio del tiempo; y

un mezclador (1230) para mezclar una banda alta sintetizada de la señal en el dominio del tiempo y una señal en el dominio del tiempo de banda baja sobremuestreada; y

- 10 un combinador (1160) para combinar la primera porción de la señal de audio decodificada y la segunda porción de la señal de audio decodificada para obtener la señal de audio decodificada.

11. Decodificador de audio según la reivindicación 10,

- 15 en el que el sobremuestreador (1210) comprende un banco de filtros de análisis (1471) que opera a una primera velocidad de muestreo del decodificador de banda baja en el dominio del tiempo y un banco de filtros de síntesis (1473) que opera a una segunda velocidad de muestreo de salida que es más alta que la primera velocidad de muestreo de banda baja en el dominio del tiempo.

12. Decodificador de audio según la reivindicación 10 u 11,

- 20 en el que el decodificador de banda baja en el dominio del tiempo (1200) comprende un decodificador de señal residual (1149, 1141, 1142) y un filtro de síntesis (1143) para filtrar una señal residual usando coeficientes del filtro de síntesis (1145),

- 30 en el que el decodificador de extensión de anchura de banda en el dominio del tiempo (1220) está configurado para sobremuestrear la señal residual (1221) y procesar (1222) una señal residual sobremuestreada usando una operación no lineal para obtener una señal residual de banda alta, y para conformar espectralmente (1223) la señal residual de banda alta para obtener la banda alta sintetizada.

13. Decodificador de audio según una de las reivindicaciones 10 a 12,

- 35 en el que el primer procesador de decodificación (1120) comprende un posfiltro de predicción a largo plazo adaptativo (1420) para el posfiltrado de la primera porción de la señal decodificada, en el que el filtro (1420) es controlado por uno o más parámetros de predicción a largo plazo incluidos en la señal de audio codificada.

14. Decodificador de audio según una de las reivindicaciones 10 a 13, que comprende además:

- 40 un procesador cruzado (1170) para calcular, a partir de la representación espectral decodificada de la primera porción de la señal de audio codificada, datos de inicialización del segundo procesador de decodificación (1140), de modo que el segundo procesador de decodificación (1140) se inicializa para decodificar la segunda porción de la señal de audio codificada que sigue en el tiempo a la primera porción de la señal de audio en la señal de audio codificada.

- 45 15. Decodificador de audio según la reivindicación 14, en el que el procesador cruzado (1170) comprende además:

- 50 un convertidor de frecuencia-tiempo (1171) que opera a una velocidad de muestreo menor que el convertidor de frecuencia-tiempo (1124) del primer procesador de decodificación (1120) para obtener una primera porción de la señal decodificada adicional en el dominio del tiempo,

en el que la salida de señal por el convertidor de frecuencia-tiempo (1171) tiene una segunda velocidad de muestreo que es menor que la primera velocidad de muestreo asociada con una salida del convertidor de frecuencia-tiempo (1124) del primer procesador de decodificación (1120),

- 55 en el que el convertidor de frecuencia-tiempo (1171) comprende:

- 60 un selector (726) para seleccionar una porción baja de un espectro introducido en el convertidor de frecuencia-tiempo (1171) de acuerdo con una relación entre la primera velocidad de muestreo y la segunda velocidad de muestreo, siendo la relación menor que 1;

un procesador de transformadas (720) que tiene una longitud de transformada que es menor que una longitud de transformada (710) del convertidor de frecuencia-tiempo (1124) del primer procesador de decodificación (1120); y

- 65 un administrador de ventanas de síntesis (722) que usa una ventana que tiene un número de coeficientes menor que

una ventana

usada por el convertidor de frecuencia-tiempo (1124) del primer procesador de decodificación (1120).

5 16. Decodificador de audio según una de las reivindicaciones 14 o 15,

en el que el procesador cruzado (1170) comprende:

10 una fase de retardo (1172) para retardar la primera porción de la señal de audio decodificada adicional y para introducir una versión retardada de la primera porción de la señal decodificada en una fase de desacentuación (1144) del segundo procesador de decodificación para la inicialización;

15 un filtro de preacentuación (1173) y una fase de retardo (1175) para filtrar y retardar la primera porción de la señal decodificada adicional y para introducir una fase de retardo emitida en un filtro de síntesis de predicción (1143) del segundo procesador de decodificación para la inicialización;

20 un filtro de análisis de predicción (1174) para generar una señal de predicción residual a partir de la primera porción espectral decodificada adicional o de una primera porción de la señal decodificada preacentuada adicional (1173) y para introducir una señal de predicción residual en un sintetizador de libro de código (1141) del segundo procesador de decodificación (1200); o

25 un conmutador (1480) para introducir la primera porción de la señal decodificada adicional o una salida de la fase de desacentuación del segundo procesador de decodificación en una fase de análisis (1471) de un remuestreador (1210) del segundo procesador de decodificación para inicialización.

17. Decodificador de audio según una de las reivindicaciones 10 a 16,

en el que el segundo procesador de decodificación (1200) comprende al menos un bloque del grupo de bloques que comprende:

30 una ACELP para decodificar las ganancias y un libro de código innovador;
una fase de síntesis del libro de código adaptativo (1141);
un posprocesador ACELP (1142);
un filtro de síntesis de predicción (1143); y
35 una fase de desacentuación (1144).

18. Procedimiento para codificar una señal de audio, que comprende:

40 primera codificación (600) de una primera porción de la señal de audio en el dominio de la frecuencia, teniendo la primera porción de la señal de audio asociada con ella una frecuencia de muestreo, en la que la primera codificación (600) comprende:

45 conversión (602) de la primera porción de la señal de audio en una representación en el dominio de la frecuencia que tiene líneas espectrales hasta una frecuencia máxima de la primera porción de la señal de audio;

análisis (604) de la representación en el dominio de la frecuencia hasta la máxima frecuencia para determinar las primeras porciones espectrales que se van a codificar con una primera resolución espectral y segundas porciones espectrales que se van a codificar con una segunda resolución espectral, siendo la segunda resolución espectral menor que la primera resolución espectral, en el que el análisis (604) determina una primera porción espectral (306) de las primeras porciones espectrales, estando la primera porción espectral situada, con respecto a la frecuencia, entre dos segundas porciones espectrales (307a, 307b) de las segundas porciones espectrales, en la que la frecuencia máxima es inferior o igual a la mitad de la frecuencia de muestreo y al menos una cuarta parte de la frecuencia de muestreo o superior;

55 codificación (606) de las primeras porciones espectrales con la primera resolución espectral y codificación de las segundas porciones espectrales con la segunda resolución espectral, en las que la codificación de la segunda porción espectral comprende el cálculo, a partir de las segundas porciones espectrales, de la información de la envolvente espectral que tiene la segunda resolución espectral;

60 segunda codificación (610) de una segunda porción de la señal de audio diferente en el dominio del tiempo, en la que la segunda codificación (610) comprende:

65 conversión (900) de la segunda porción de la señal de audio en una representación de velocidad de muestreo inferior, siendo la velocidad de muestreo inferior menor que la velocidad de muestreo de la señal de audio, en el que la representación de la velocidad de muestreo inferior no incluye la banda alta de la señal de audio;

codificación en el dominio del tiempo (910) de la representación de la velocidad de muestreo inferior; y

codificación paramétrica (920) de la banda alta;

5

análisis (620) de la señal de audio y determinación de qué porción de la señal de audio es la primera porción de la señal de audio codificada en el dominio de la frecuencia y qué porción de la señal de audio es la segunda porción de la señal de audio codificada en el dominio del tiempo; y

10 formación (630) de una señal de audio codificada que comprende una primera porción de la señal codificada para la primera porción de la señal de audio y una segunda porción de la señal codificada para la segunda porción de la señal de audio diferente.

19. Procedimiento para decodificar una señal de audio codificada, que comprende:

15

primera decodificación (1120) de una primera porción de la señal de audio codificada en un dominio de la frecuencia, comprendiendo la primera decodificación (1120):

decodificación (1122) de las primeras porciones espectrales con una alta resolución espectral y síntesis de las
20 segundas porciones espectrales usando una representación paramétrica de las segundas porciones espectrales y al menos una primera porción espectral decodificada para obtener una representación espectral decodificada, en la que la decodificación (1122) comprende la generación de la primera representación decodificada de manera que una primera porción espectral (306) se coloca con respecto a la frecuencia entre dos segundas porciones espectrales (307a, 307b); y que la representación espectral decodificada se extiende hasta una frecuencia máxima de una
25 representación en el tiempo de una señal de audio decodificada, siendo un valor espectral para la frecuencia máxima igual a cero o diferente de cero; y

conversión (1124) de la representación espectral decodificada en un dominio del tiempo para obtener una primera porción de la señal de audio decodificada,

30

segunda decodificación (1140) de una segunda porción de la señal de audio codificada en el dominio del tiempo para obtener una segunda porción de la señal de audio decodificada, en la que la segunda decodificación comprende:

decodificación (1200) para obtener una señal en el dominio del tiempo de banda baja;

35

sobremuestreo (1210) de la señal en el dominio del tiempo de banda baja para obtener una señal en el dominio del tiempo de banda baja sobremuestreada;

síntesis (1220) de una banda alta de una señal de salida en el dominio del tiempo; y

40

mezclado (1230) de una banda alta sintetizada de la señal en el dominio del tiempo y una señal en el dominio del tiempo de banda baja sobremuestreada; y

combinación (1160) de la primera porción de la señal de audio y la segunda porción de la señal de audio decodificada
45 para obtener la señal de audio decodificada.

20. Programa informático para la realización, cuando se ejecuta en un ordenador o un procesador, del procedimiento según la reivindicación 18 o 19.

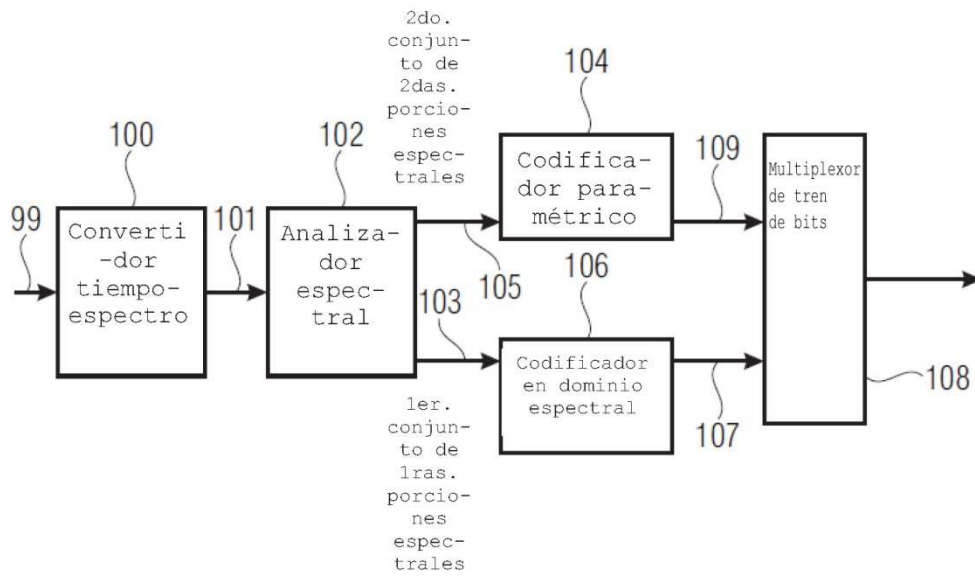


FIG 1A

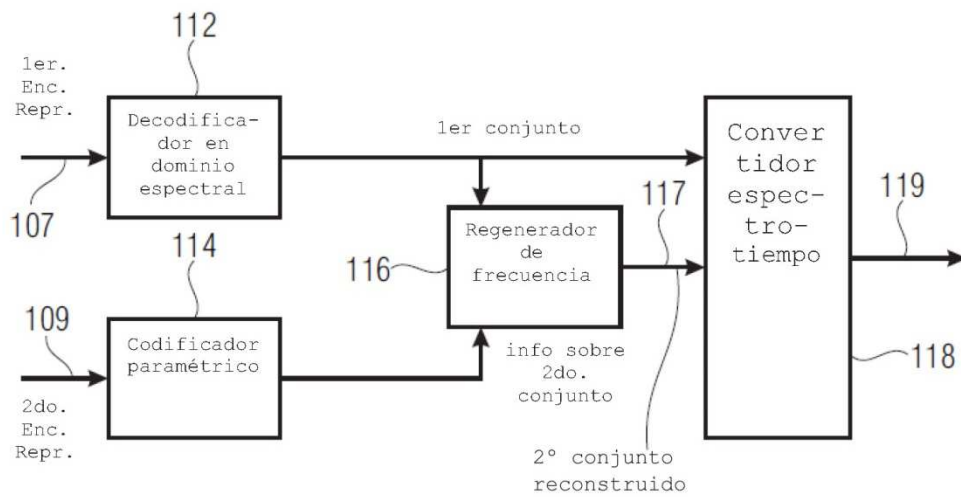


FIG 1B

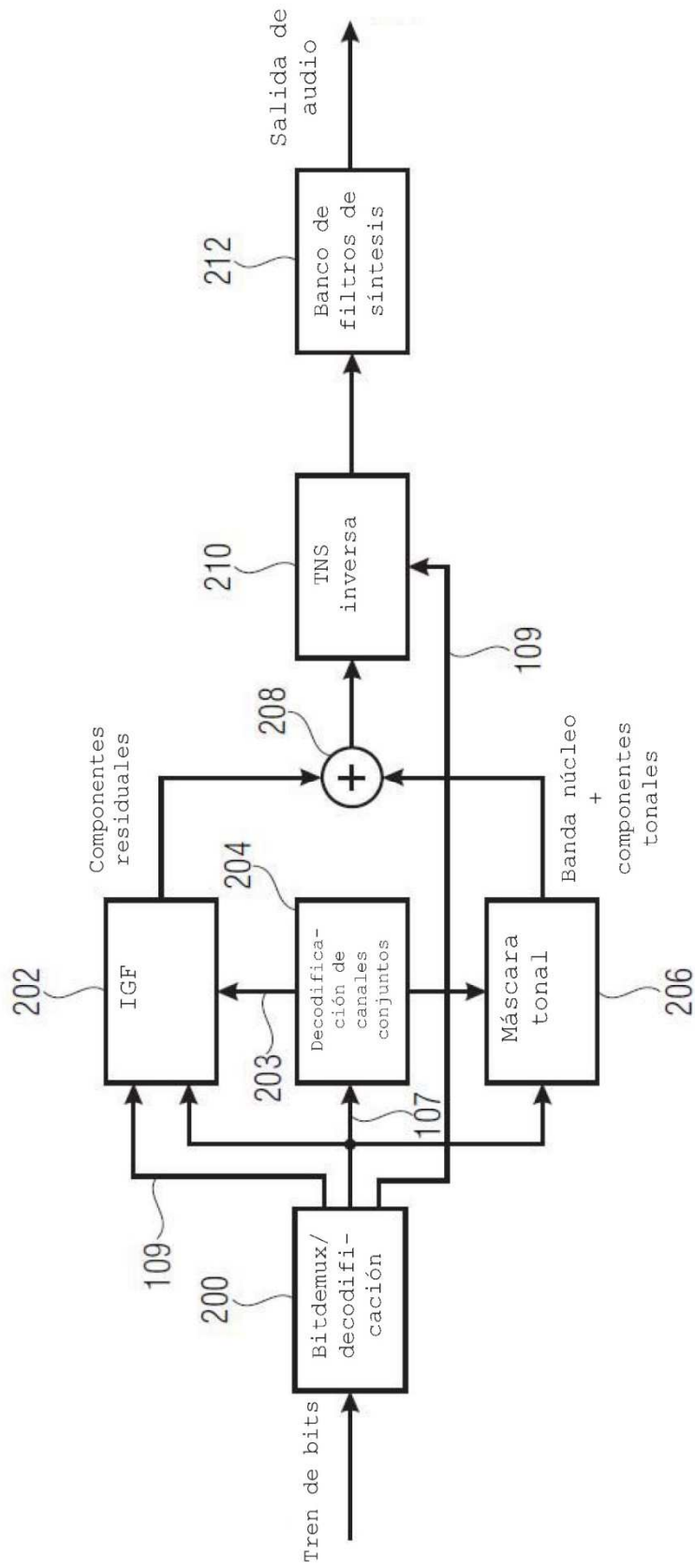


FIG 2A

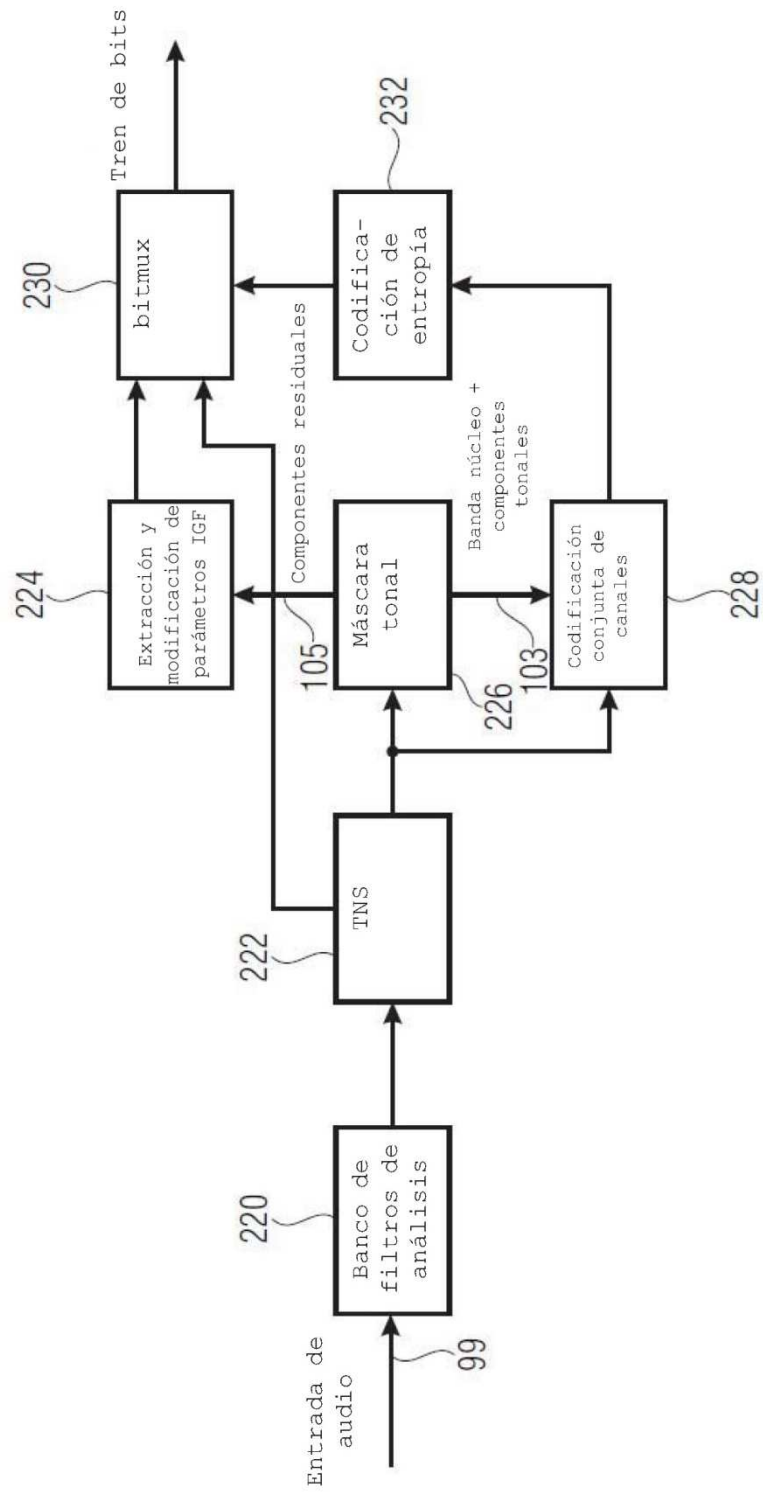


FIG 2B

- 1ra. resolución (elevada resolución) para envoltorio del 1er conjunto (codificación de a líneas)
- 2da. resolución (baja resolución) para envoltorio del 2do conjunto (factor de escala por SCB)

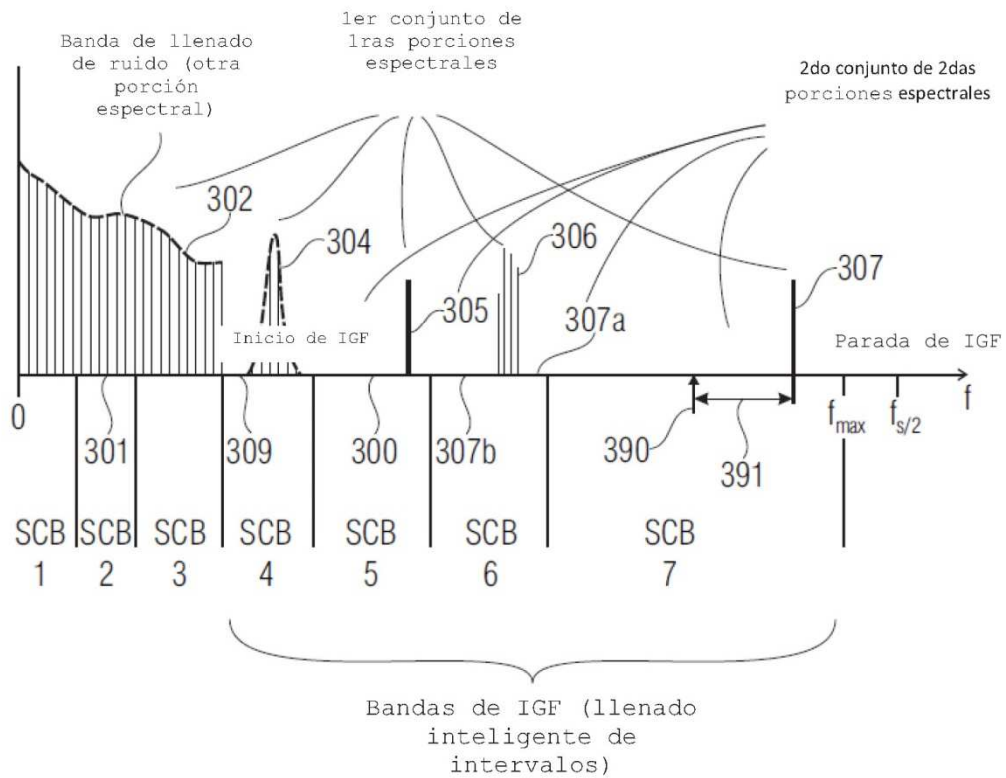


FIG 3A

SCB1	SCB2	SCB3	SCB4	SCB5	SCB6	SCB7
SF1	SF2	SF3	SF4	SF5	SF6	SF7
			E ₁	E ₂	E ₃	E ₄
	NF ₂					

308
310
312

FIG 3B

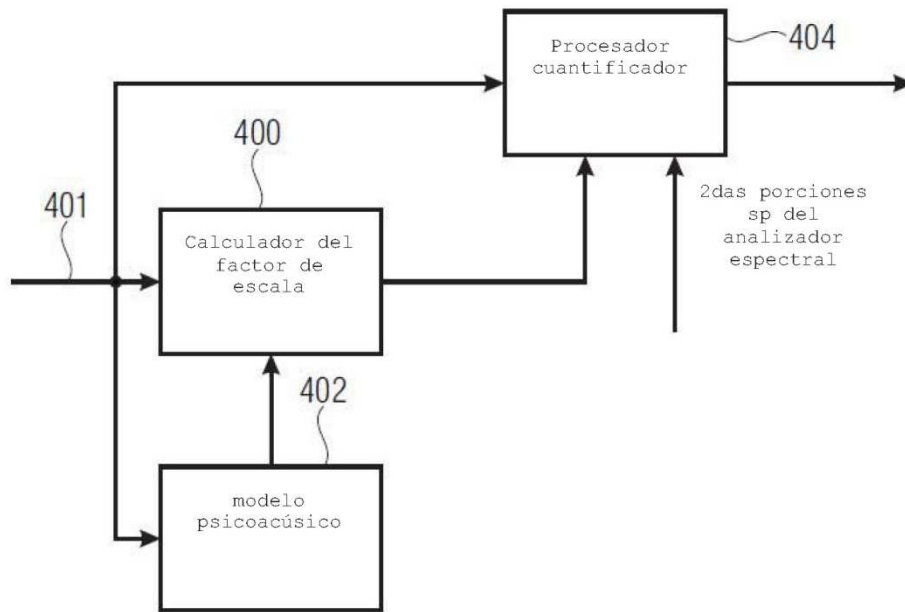


FIG 4A

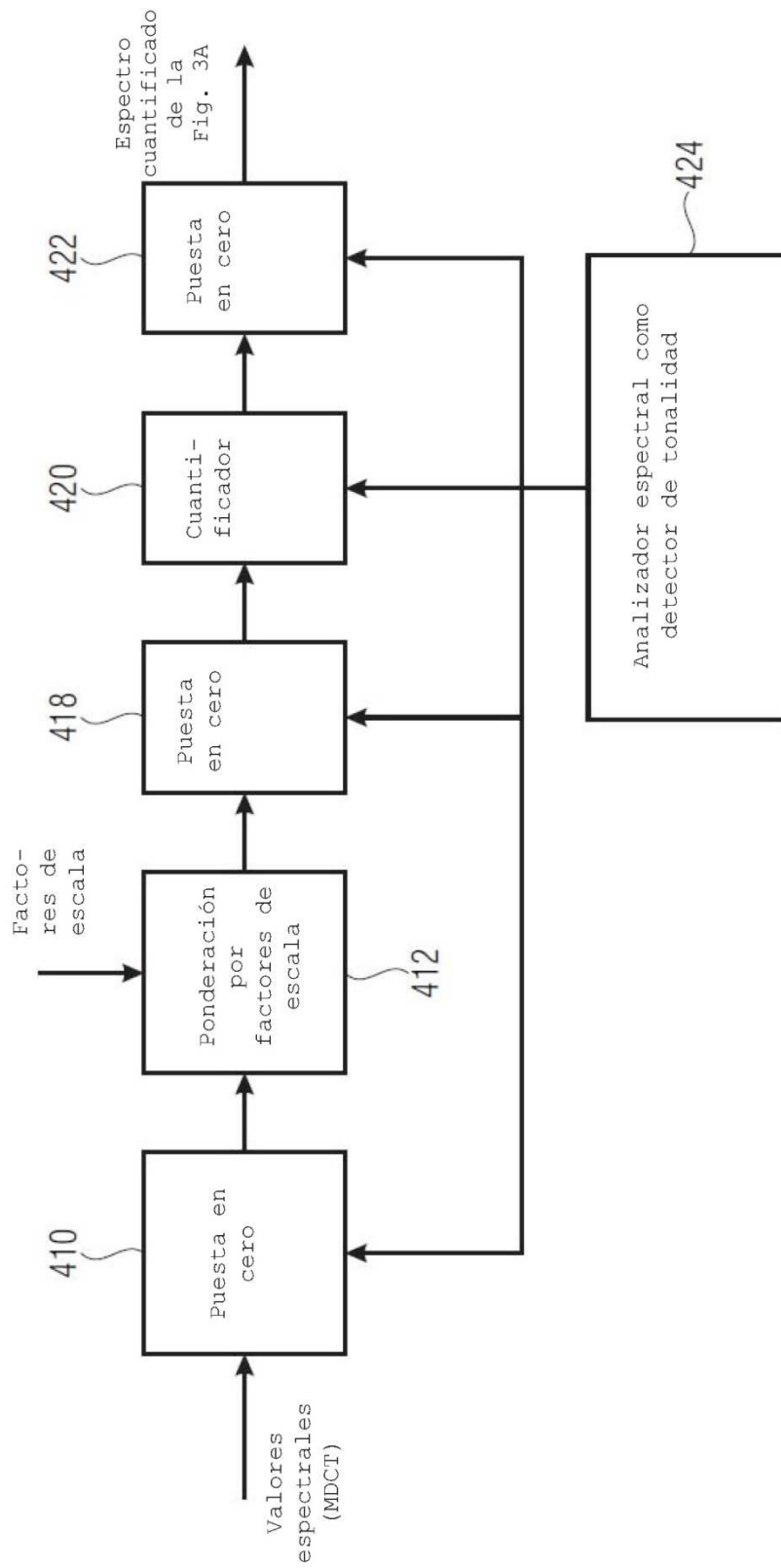


FIG 4B

PROCESADOR CUANTIFICADOR

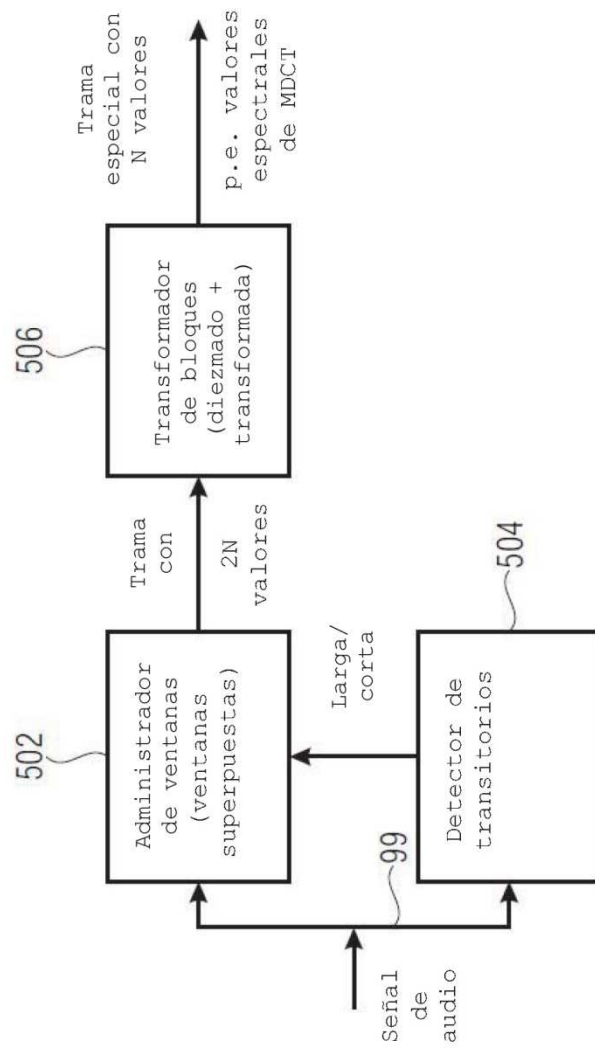


FIG 5A

(otras porciones espectrales)

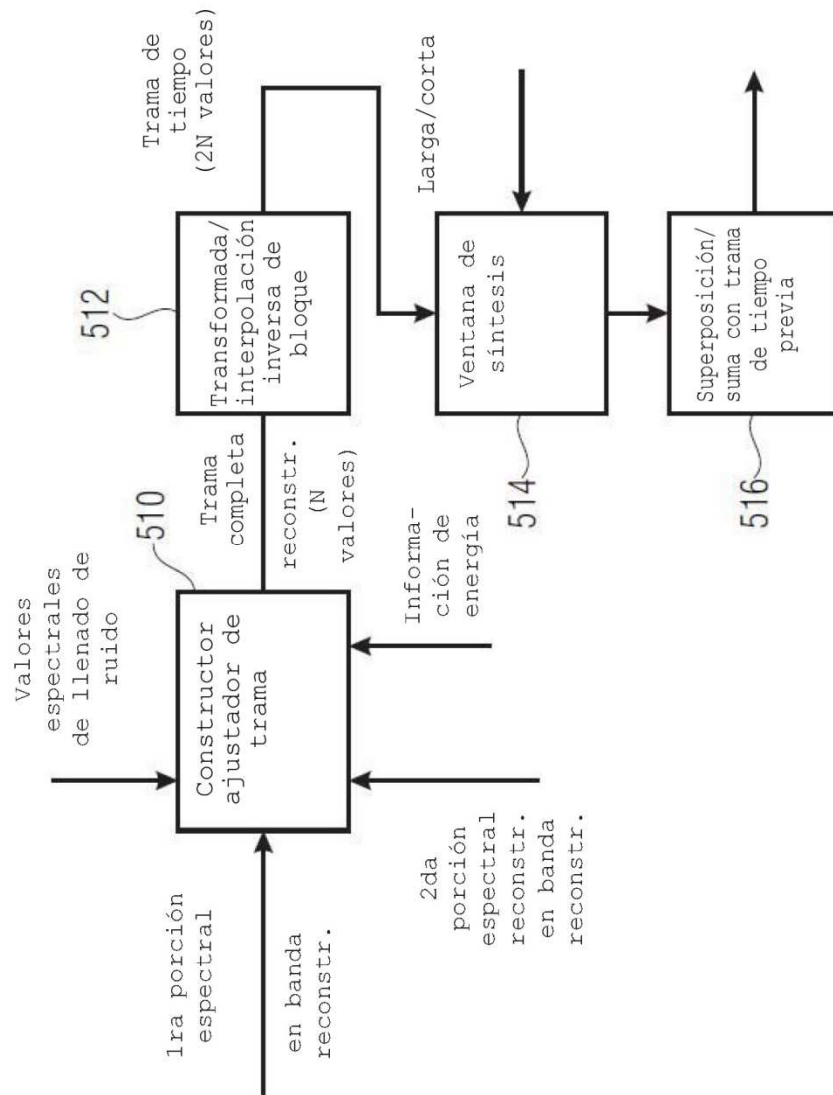


FIG 5B

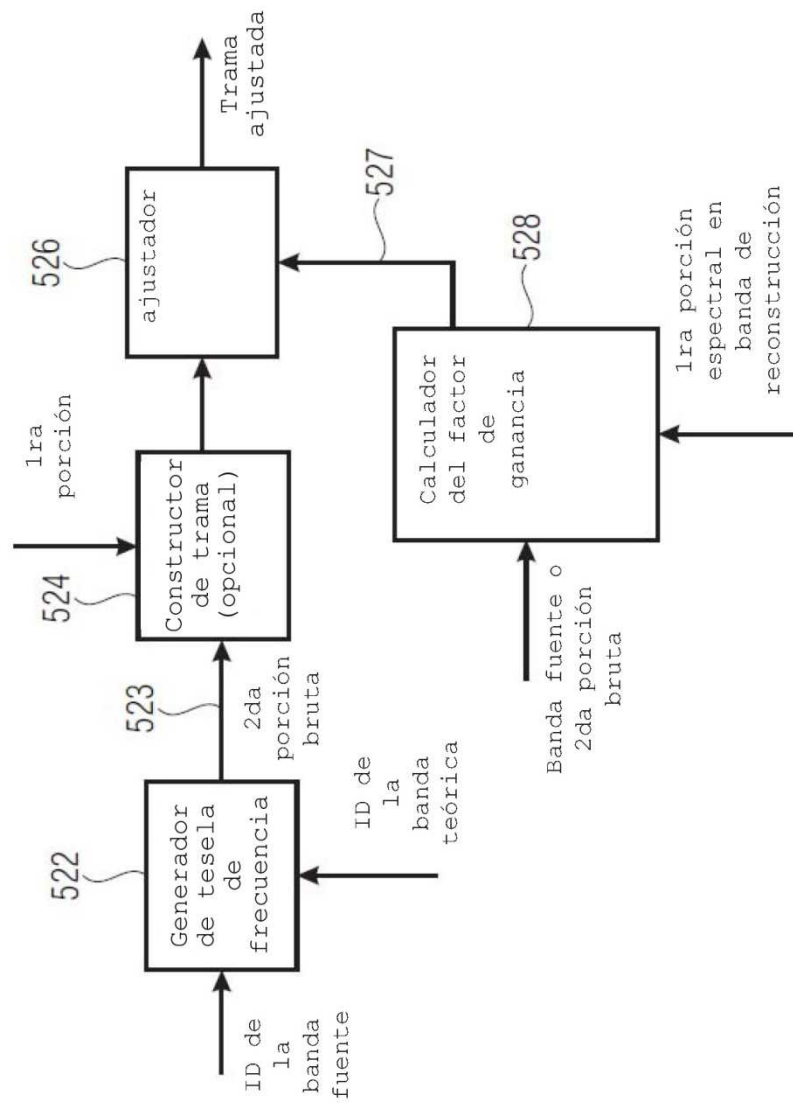


FIG 5C

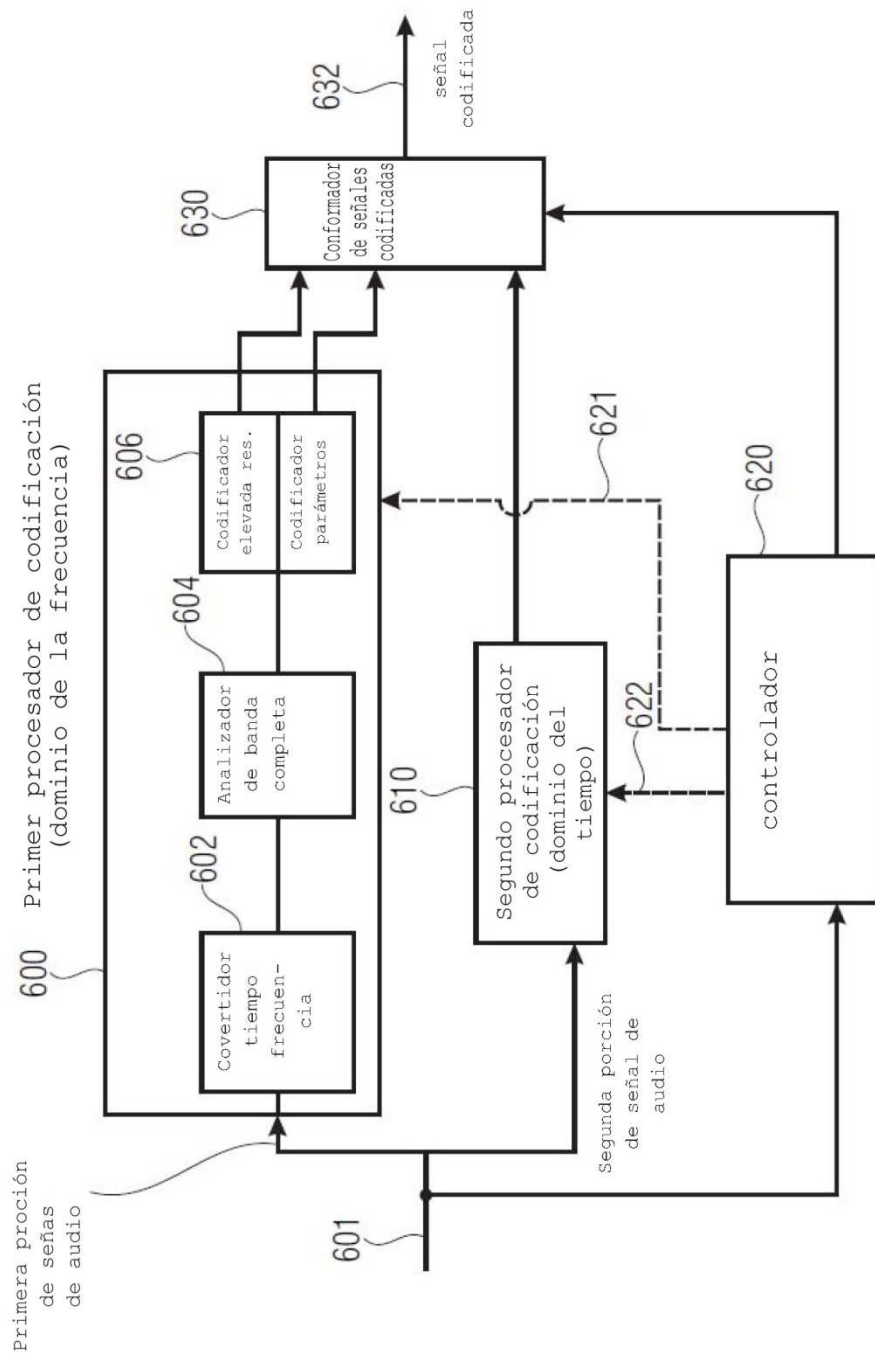


FIG 6

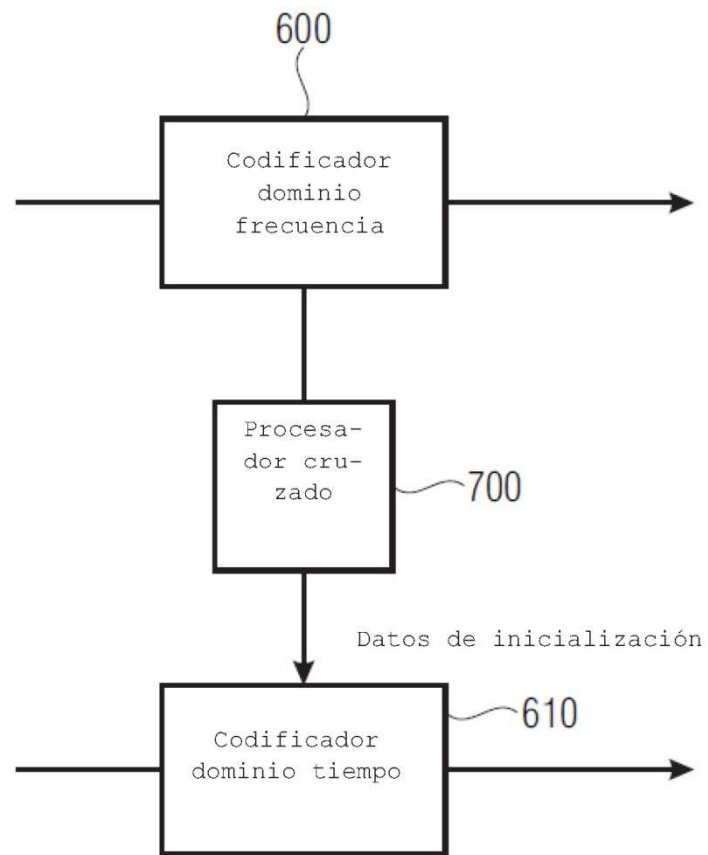
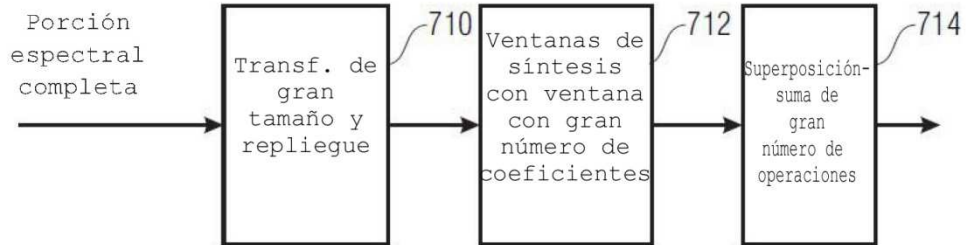
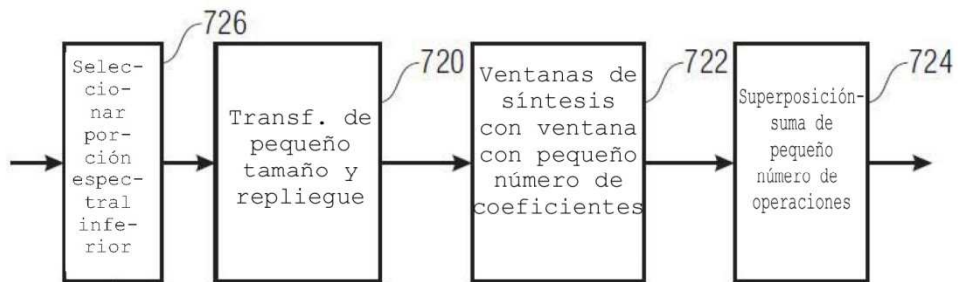


FIG 7A

VELOCIDAD DE MUESTREO DE SALIDA



VELOCIDAD DE MUESTREO DE CODIFICACIÓN EN EL DOMINIO DEL TIEMPO



DS
factor de
submuestreo

$$\frac{1}{DS} = \frac{\text{T.D Codificador S.R.}}{\text{F.D Codificador S.R.}}$$

$$DS * \text{Pequeño tamaño} = \text{Gran tamaño}$$

$$DS * \text{pequeño número de coeficientes} = \text{Gran número de coeficientes}$$

$$DS * \text{pequeño número de coeficientes} = \text{Gran número de coeficientes}$$

FIG 7B

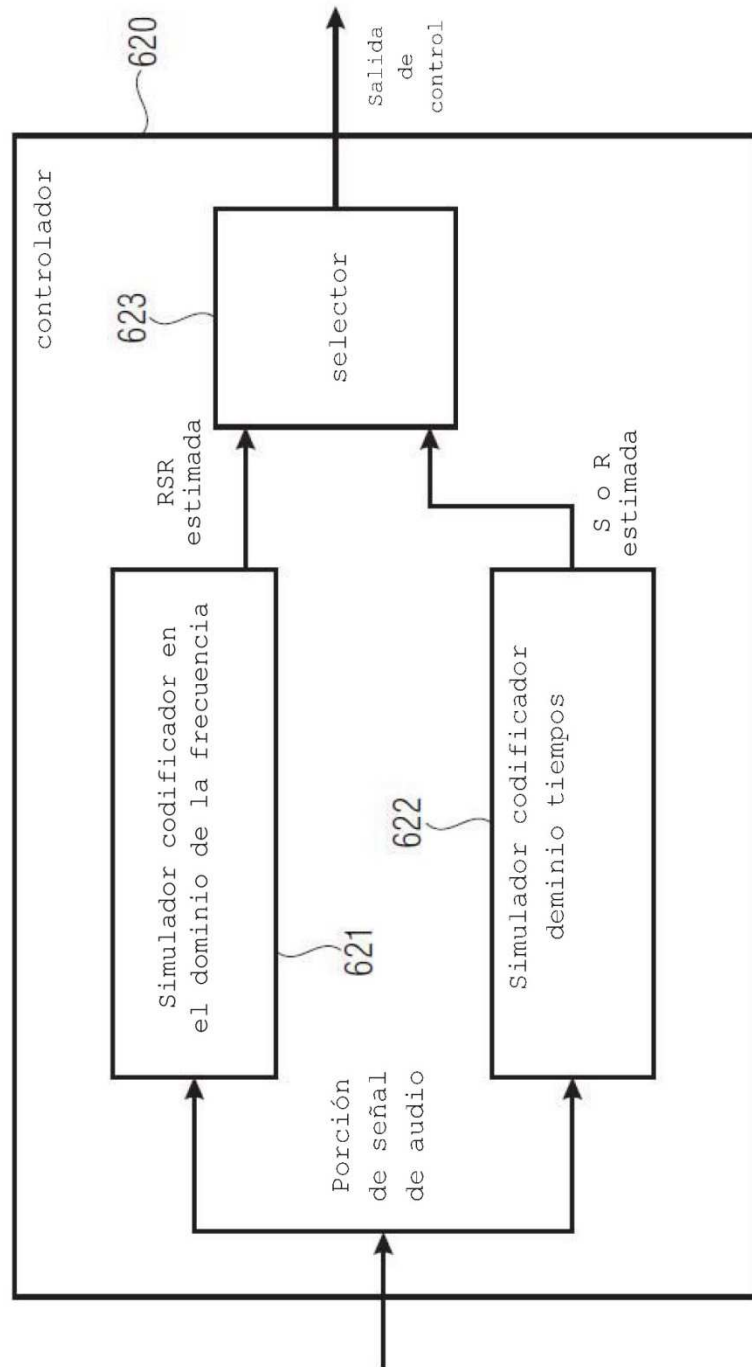


FIG 8

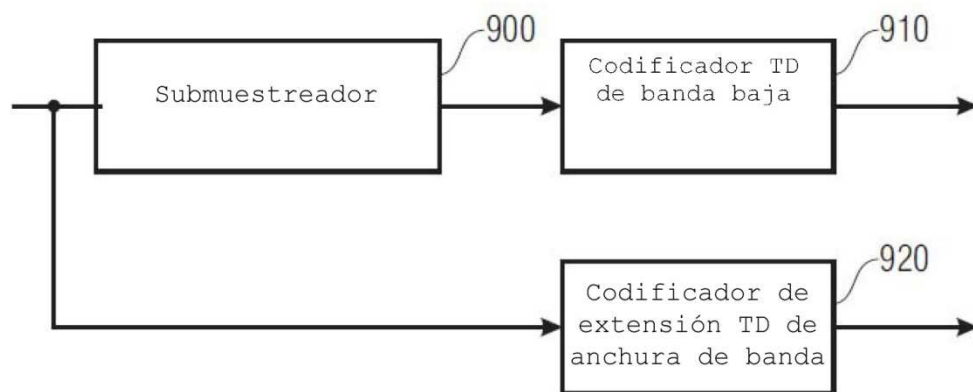


FIG 9

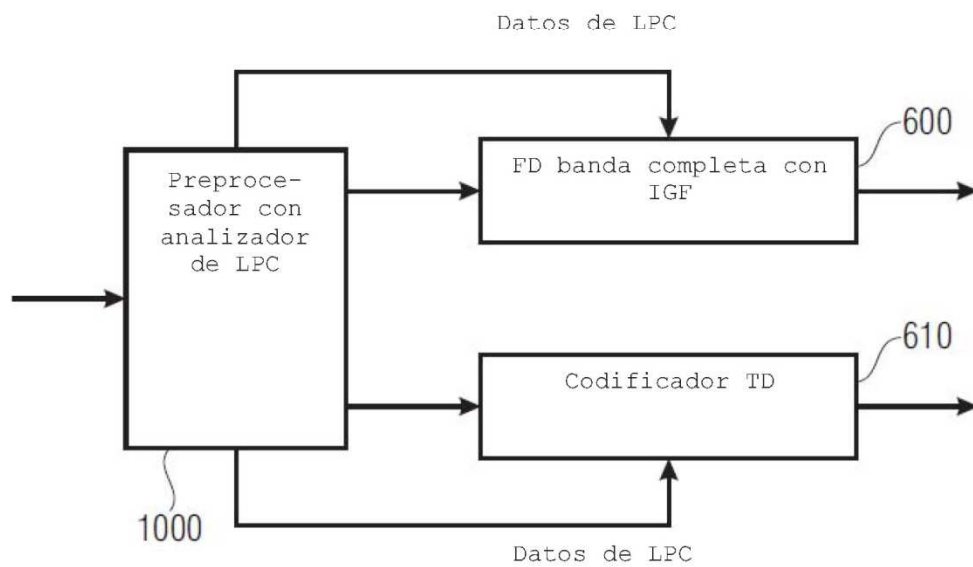


FIG 10

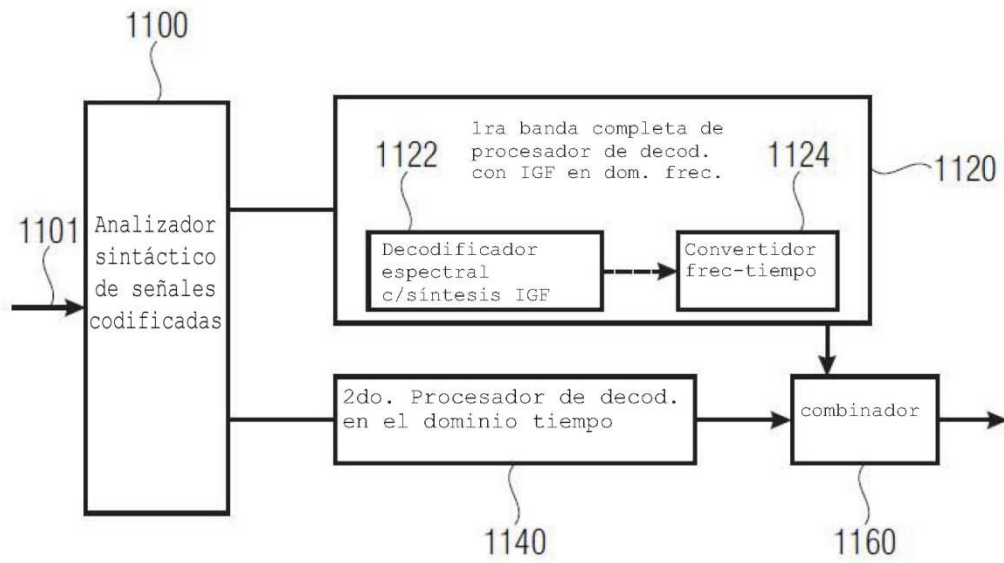


FIG 11A

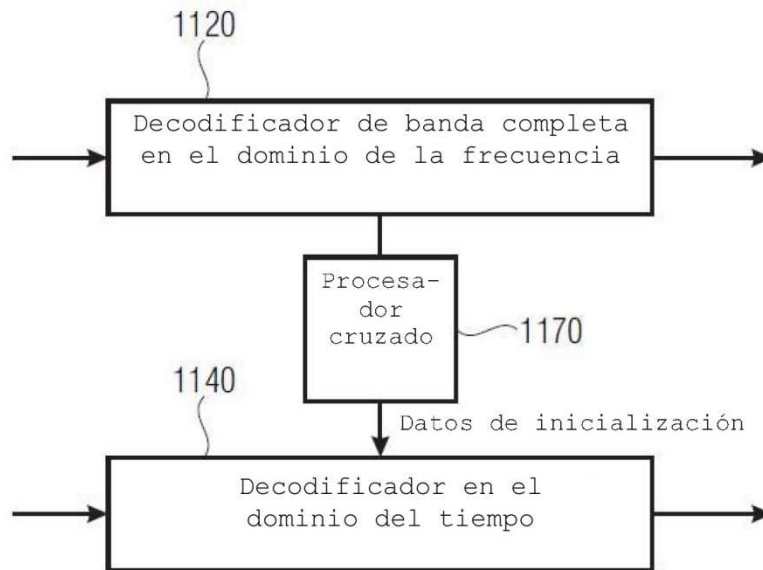


FIG 11B

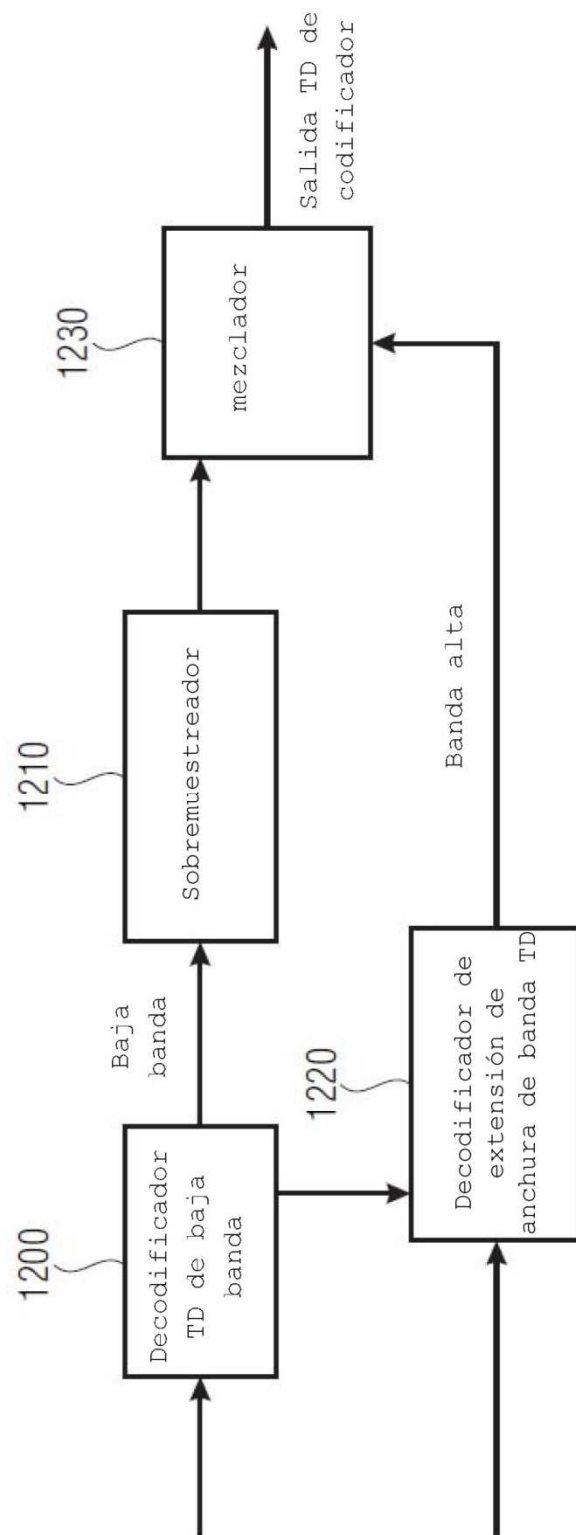


FIG 12

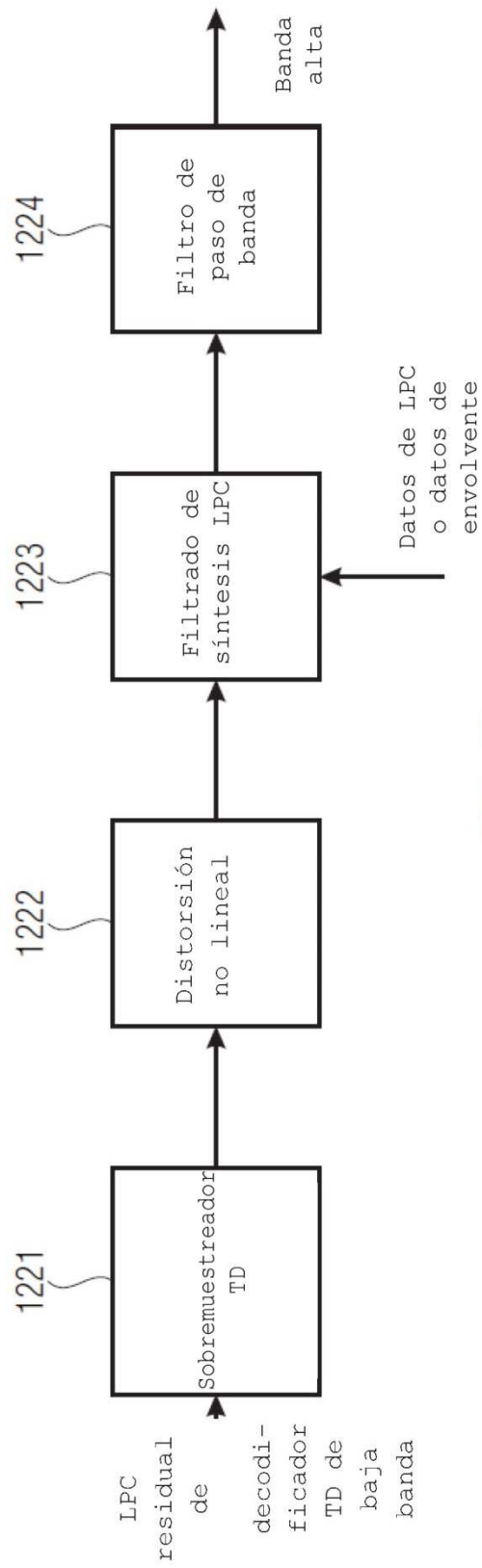


FIG 13

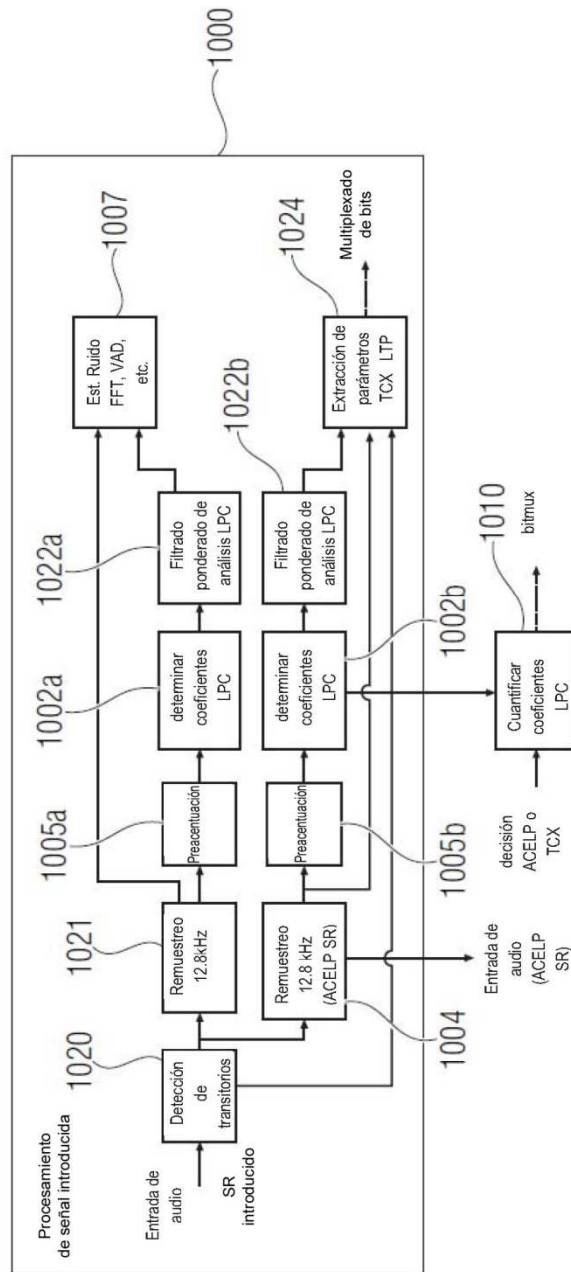


FIG 14A-1

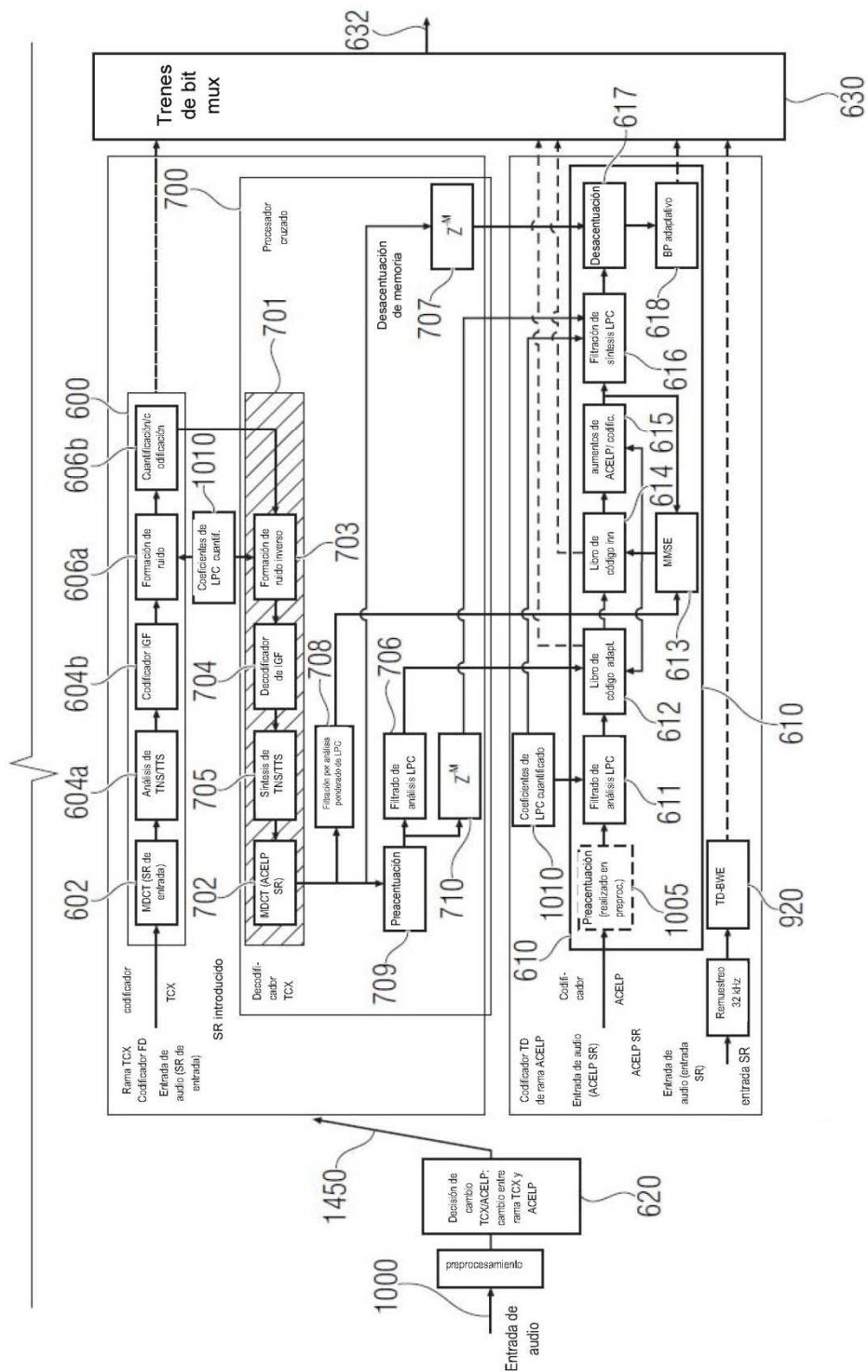


FIG 14A-2

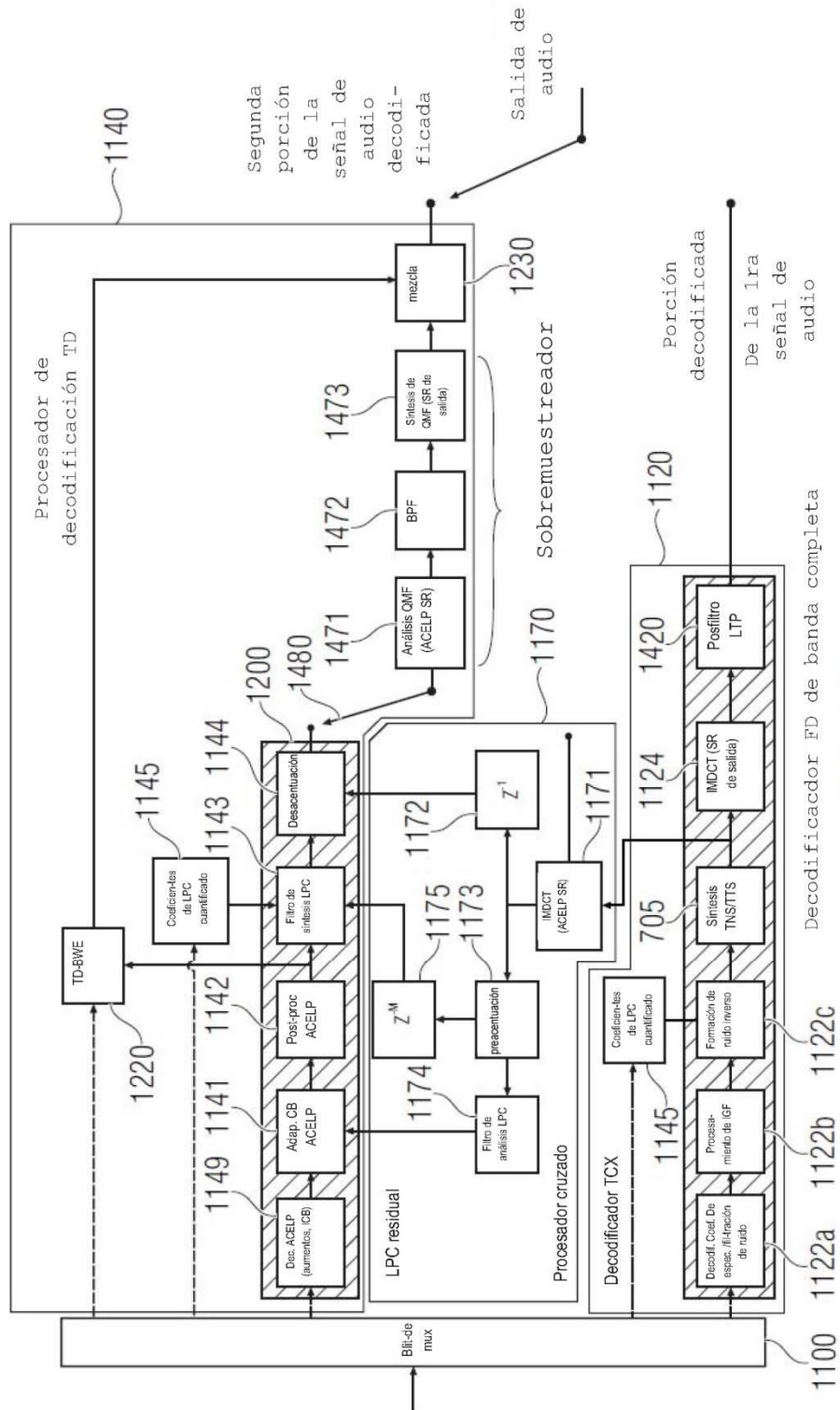


FIG 14B

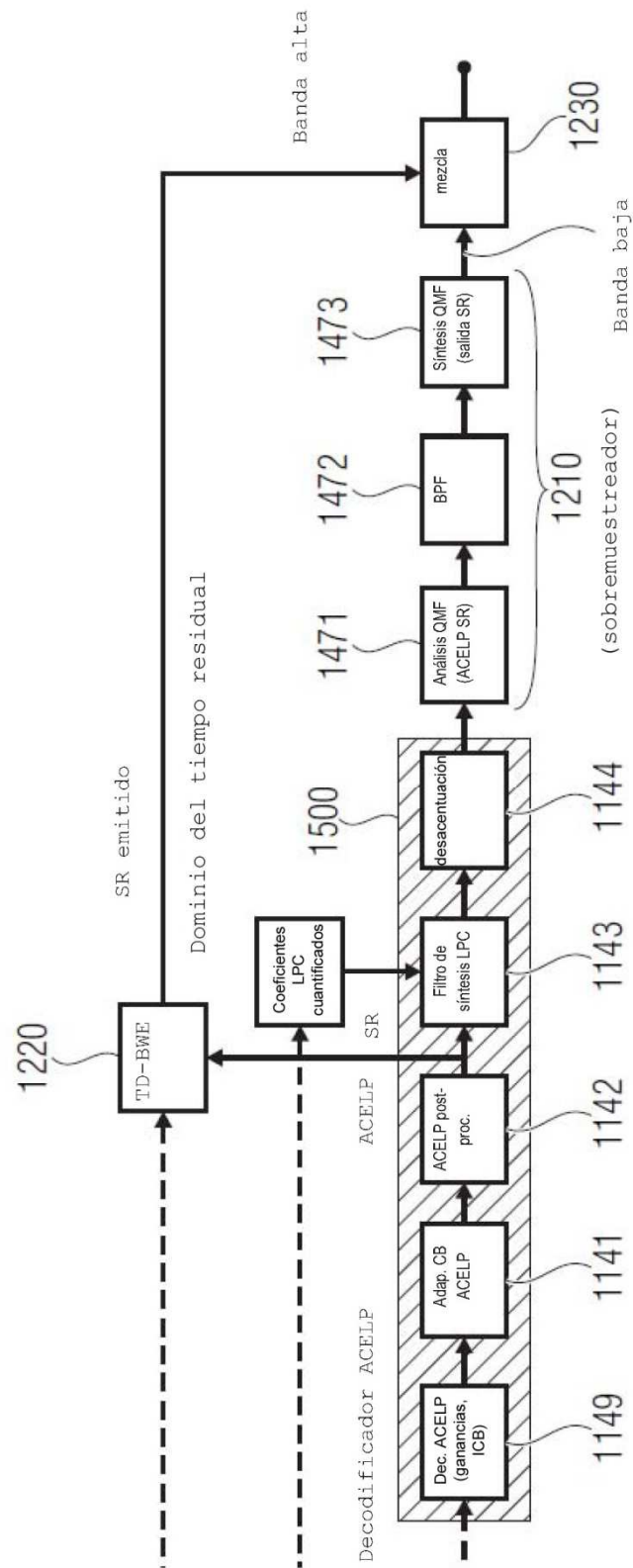


FIG 14C