

19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 796 552**

51 Int. Cl.:

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: **19.06.2009 PCT/EP2009/004451**

87 Fecha y número de publicación internacional: **14.01.2010 WO10003539**

96 Fecha de presentación y número de la solicitud europea: **19.06.2009 E 09776790 (9)**

97 Fecha y número de publicación de la concesión europea: **04.03.2020 EP 2301026**

54 Título: **Sintetizador de señales de audio y codificador de señales de audio**

30 Prioridad:

11.07.2008 US 79839 P
08.10.2008 US 103820 P

45 Fecha de publicación y mención en BOPI de la traducción de la patente:
27.11.2020

73 Titular/es:

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

72 Inventor/es:

NAGEL, FREDERIK;
DISCH, SASCHA;
RETTTELBACH, NIKOLAUS;
NEUENDORF, MAX;
GRILL, BERNHARD;
KRÄMER, ULRICH y
WABNIK, STEFAN

74 Agente/Representante:

SALVÀ FERRER, Joan

ES 2 796 552 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

Sintetizador de señales de audio y codificador de señales de audio

5 **[0001]** La presente invención se refiere a un sintetizador de señales de audio para generar una señal de audio de síntesis, un codificador de señales de audio y un flujo de datos que comprende una señal de audio codificada.

[0002] La codificación de voz y la codificación de audio natural son dos clases principales de códecs para señales de audio. Los codificadores de audio natural se utilizan comúnmente para música o señales arbitrarias a tasas de bits medias y ofrecen, en general, anchos de banda de audiofrecuencia amplios. Los codificadores de voz están básicamente limitados a reproducción de voz y pueden utilizarse a una tasa de bits muy baja. La voz de banda ancha proporciona una mejora de calidad subjetiva principal sobre la voz de banda estrecha. El aumento del ancho de banda no solamente mejora la naturalidad de voz, sino además el reconocimiento e inteligibilidad del hablante. Así, la codificación de voz de banda ancha es un tema importante en la próxima generación de sistemas telefónicos. 10 Asimismo, debido al gran crecimiento del campo multimedia, la transmisión de música y otras señales no vocales a una alta calidad sobre sistemas telefónicos, así como el almacenamiento y, por ejemplo, la transmisión para radio/TV u otros sistemas de difusión, es una característica conveniente.

[0003] Para reducir radicalmente la tasa de bits, puede llevarse a cabo una codificación fuente mediante el uso de códecs perceptuales de audio de banda dividida. Estos códecs de audio natural explotan la irrelevancia perceptual y la redundancia estadística en la señal. En caso de que la explotación de lo anterior por sí sola no resulte suficiente con respecto a las restricciones de tasa de bits dadas, se reduce la tasa de muestreo. También es frecuente la reducción de la cantidad de niveles de composición, lo cual permite una distorsión de cuantificación audible ocasional, y el uso de degradación del campo de estéreo por medio de codificación de estéreo conjunta o codificación paramétrica de dos o más canales. El uso excesivo de tales procedimientos tiene como resultado una degradación perceptual molesta. Con el fin de mejorar el desempeño de la codificación, se utilizan procedimientos de ampliación de ancho de banda tales como la replicación de banda espectral (SBR), como un procedimiento eficiente para generar señales de alta frecuencia en un códec basado en HFR (reconstrucción de alta frecuencia). 20

[0004] En el procedimiento de replicación de las señales de alta frecuencia, puede aplicarse una cierta transformación, por ejemplo, a las señales de baja frecuencia, y las señales transformadas se insertan a continuación como señales de alta frecuencia. Este procedimiento también se conoce como "parche" ("patching") y pueden utilizarse diferentes transformaciones. Los estándares de audio MPEG-4 utilizan solamente un algoritmo de parche ("patching") para todas las señales de audio. Por lo tanto, carece de flexibilidad para adaptar el parche ("patching") en diferentes 25 señales o esquemas de codificación.

[0005] Por un lado, el estándar MPEG-4 proporciona un procesamiento sofisticado de banda alta regenerada, en el cual se aplica una gran cantidad de parámetros de SBR importantes. Estos parámetros de SBR importantes son los datos en la envolvente espectral, los datos sobre el umbral mínimo de ruido que deben sumarse a la porción 30 espectral regenerada, información sobre la herramienta de filtrado inverso con el fin de adaptar la tonalidad de la banda alta regenerada a la tonalidad de la banda alta original, y datos adicionales de procesamiento de replicación de banda espectral tales como datos sobre armónicos que faltan, etc. Se demuestra que este procesamiento establecido de manera adecuada del espectro replicado que es proporcionado por un parche ("patching") de señales de paso de banda consecutivas dentro del dominio de banco de filtros resulta eficiente para proporcionar una alta calidad y para 35 ser susceptible de ser implementado con recursos razonables referidos a potencia de procesamiento, requisitos de memoria y requisitos de energía.

[0006] Por otro lado, el parche ("patching") tiene lugar en el mismo banco de filtros en el que tiene lugar el procesamiento adicional de la señal de parche, de forma que hay una estrecha relación entre la operación de parche 40 ("patching") y el procesamiento adicional del resultado de la operación de parche ("patching"). Por lo tanto, la implementación de diferentes algoritmos de parche ("patching") resulta problemática en esta estrategia combinada.

[0007] El documento WO 98/57436 describe procedimientos de transposición utilizados en replicación de banda espectral, que se combinan con un ajuste de envolvente espectral. 45

[0008] El documento WO 02/052545 enseña que las señales pueden clasificarse ya sea en tipo tren de impulsos o bien que no son del tipo tren de impulsos y, sobre la base de esta clasificación, se propone un instrumento de transposición conmutado adaptativo. El instrumento de transposición conmutado lleva a cabo dos algoritmos de parche ("patching") en paralelo y una unidad mezcladora combina ambas señales de parche que dependen de la 50 clasificación (tren de impulsos o no de tren de impulsos). La conmutación real entre los repetidores o su mezcla se lleva a cabo en un banco de filtros de ajuste por envolvente en respuesta a datos de control y envolvente. Asimismo, para señales del tipo tren de impulsos, la señal de banda base se transforma en un dominio de banco de filtros, se lleva a cabo una operación de traducción de frecuencia y se efectúa un ajuste de envolvente del resultado de la traducción de frecuencia. Éste es un procedimiento de parche/procesamiento adicional combinado. Para señales que 55 no son del tipo tren de impulsos, se proporciona un instrumento de transposición de dominio frecuencial (I.T. de D.F)

y el resultado del I.T. de D.F. se transforma a continuación en el dominio de banco de filtros, en el cual se lleva a cabo el ajuste de envolvente. Así, la implementación y flexibilidad de este procedimiento que presenta, en una alternativa, una estrategia de parche/procesamiento adicional combinado y que tiene, en la otra alternativa, un I.T. de D.F. que está posicionado fuera del banco de filtros en el cual tiene lugar el ajuste de envolvente, resulta problemática con respecto a las posibilidades de flexibilidad e implementación.

[0009] Un objetivo de la presente invención es proporcionar un sintetizador que proporcione una calidad mejorada y permita una implementación eficaz.

10 **[0010]** Este objetivo se logra por medio del sintetizador según la reivindicación 1, un codificador según la reivindicación 8, un procedimiento para generar una señal de audio de síntesis según la reivindicación 10, un procedimiento para generar un flujo de datos según la reivindicación 11 o un programa informático según la reivindicación 12.

15 **[0011]** La presente invención se basa en la conclusión de que, por un lado, la operación de parche ("patching") y, por otro, el procesamiento adicional de la salida de la operación de parche ("patching") deben llevarse a cabo por completo en dominios independientes. Esto proporciona, por un lado, la flexibilidad para optimizar diferentes algoritmos de parche ("patching") dentro de un generador de parche ("patching") y, por otro, para utilizar siempre el mismo ajuste de envolvente, independientemente del algoritmo de parche ("patching") básico. Por lo tanto, la creación de cualquier
20 señal de parche fuera del dominio espectral, en la cual tiene lugar el ajuste de envolvente, permite una aplicación flexible de diferentes algoritmos de parche ("patching") a diferentes porciones de señal completamente independientes del procesamiento adicional de SBR posterior, y el diseñador no debe preocuparse por los detalles para algoritmos de parche ("patching") procedentes del ajuste de envolvente o no debe preocuparse por los detalles de los algoritmos de parche ("patching") para un cierto ajuste de envolvente. En lugar de ello, los diferentes componentes de la replicación
25 de banda espectral, es decir, por un lado, la operación de parche ("patching") y, por otro, el procesamiento adicional del resultado de parche ("patching"), pueden llevarse a cabo de forma independiente uno de otro. Esto significa que en la replicación de banda espectral completa, el algoritmo de parche ("patching") se lleva a cabo por separado, lo cual tiene como consecuencia que el parche ("patching") y las operaciones de SBR restantes pueden optimizarse independientemente unas de otras y, por lo tanto, son flexibles con respecto a futuros algoritmos de parche
30 ("patching"), etc., que simplemente pueden aplicarse sin tener que cambiar ninguno de los parámetros del procesamiento adicional del resultado de parche ("patching") que se lleva a cabo en un dominio espectral en el cual no tiene lugar ningún parche ("patching").

[0012] La presente invención proporciona una calidad mejorada, dado que permite una aplicación sencilla de
35 diferentes algoritmos de parche ("patching") a porciones de señales de forma que cada porción de la señal de banda base se emparcha al algoritmo de parche ("patching") que se ajusta de la mejor manera a esta porción de señal. Asimismo, aún puede utilizarse la herramienta de ajuste de envolvente directa, eficaz y de alta calidad que opera en el banco de filtros y que está establecida de manera adecuada y ya existe en varias aplicaciones tales como HE-AAC del MPEG-4. Al separar los algoritmos de parche ("patching") del procesamiento adicional, de forma que no se aplique
40 algoritmo de parche ("patching") alguno en el dominio de banco de filtros, en el cual se lleva a cabo el procesamiento adicional del resultado de parche ("patching"), el procesamiento adicional establecido de manera adecuada del resultado de parche ("patching") puede aplicarse a todos los algoritmos de parche ("patching") disponibles. No obstante, el parche ("patching") también puede ser llevado a cabo, de forma opcional, en el banco de filtros, así como en otros dominios.

45 **[0013]** Asimismo, esta característica proporciona una capacidad de ampliaciones futuras, dado que, para aplicaciones de bajo nivel, pueden utilizarse los algoritmos de parche ("patching") que requieran una menor cantidad de recursos, mientras que, para aplicaciones de alto nivel, pueden utilizarse algoritmos de parche ("patching") que requieran una mayor cantidad de recursos, lo cual tiene como resultado una mejor calidad de audio. De forma
50 alternativa, los algoritmos de parche ("patching") pueden no presentar cambios, pero la complejidad del procesamiento adicional del resultado de parche ("patching") puede adaptarse a diferentes necesidades. Para aplicaciones de bajo nivel, por ejemplo, puede aplicarse una resolución de frecuencia reducida para el ajuste de envolvente espectral mientras que, para aplicaciones de nivel más alto, puede aplicarse una resolución de frecuencia más fina que proporciona una mejor calidad, pero que requiere, además, recursos aumentados de memoria, procesador y consumo
55 de energía específicamente en un dispositivo móvil. Todo esto puede llevarse a cabo sin implicaciones en la otra herramienta correspondiente, dado que la herramienta de parche ("patching") no depende de la herramienta de ajuste de envolvente espectral y viceversa. En lugar de ello, se ha demostrado que la separación de la generación de parche ("patching") y el procesamiento de los datos no procesados de parche por una transformada en una representación espectral tal como por medio de un banco de filtros, es una característica óptima.

60 **[0014]** Por lo tanto, la presente invención se refiere a un procedimiento para conmutación entre diferentes algoritmos de parche ("patching") en replicación de banda espectral, en el que el algoritmo de parche ("patching") utilizado depende, del lado del codificador, de una decisión tomada en el codificador, y, del lado del decodificador, de información transmitida en el flujo de bits. Por medio del uso de una replicación de banda espectral (SBR), la
65 generación de los componentes de alta frecuencia puede llevarse a cabo, por ejemplo, copiando los componentes de

señal de baja frecuencia en un banco de filtro de QMF (QMF = filtro espejo en cuadratura) sobre las bandas de alta frecuencia. Este copiado también se conoce como parche (“patching”) y, según las realizaciones de la presente invención, este parche (“patching”) es reemplazado o suplementado por procedimientos alternativos, que también pueden llevarse a cabo en el dominio del tiempo. Ejemplos de los algoritmos de parche (“patching”) alternativos son:

- 5 (1) sobremuestreo (por ejemplo, por reflexión en espejo del espectro);
- (2) vocoder de fase;
- (3) distorsión no lineal;
- 10 (4) reflexión en espejo del espectro en el dominio de QMF por intercambio del orden de banda de QMF;
- (5) accionado por modelo (en particular, para voz); y
- (6) modulación

[0015] Los algoritmos de parche (“patching”) alternativos pueden llevarse a cabo, además, dentro del codificador, con el fin de obtener los parámetros de replicación de banda espectral, que son utilizados, por ejemplo, por herramientas de SBR, como relleno de ruido, filtrado inverso, armónicos que faltan, etc. Según las realizaciones, se reemplaza el algoritmo de parche (“patching”) dentro de un generador de parche (“patching”), utilizando aún las herramientas de replicación de banda espectral restantes.

[0016] La elección concreta del algoritmo de parche (“patching”) depende de la señal de audio aplicada. Por ejemplo, el vocoder de fase altera seriamente la característica de señales de voz y, por lo tanto, el vocoder de fase no proporciona un algoritmo de parche (“patching”) adecuado, por ejemplo, para señales de voz o tipo voz. Por lo tanto, dependiendo del tipo de señal de audio, un generador de parche (“patching”) selecciona un algoritmo de parche (“patching”) entre diferentes posibilidades para generar parches para la banda de alta frecuencia. Por ejemplo, el generador de parche puede conmutar entre la herramienta de SBR convencional (copia de bandas de QMF) y el vocoder de fase o cualquiera de los otros algoritmos de parche (“patching”).

[0017] A diferencia de la implementación de SBR convencional (por ejemplo, implementada en el MPEG-4), las realizaciones de la presente invención utilizan, así, el generador de parche (“patching”) para generar la señal de alta frecuencia. El generador de parche (“patching”) puede operar no solamente en el dominio frecuencial, sino también en el dominio del tiempo e implementa algoritmos de parche (“patching”) como por ejemplo: reflexión en espejo y/o sobremuestreo y/o vocoder de fase y/o distorsión no lineal. El hecho de que la replicación de banda espectral se lleve a cabo en el dominio frecuencial o en el dominio del tiempo depende de la señal concreta (es decir, que es adaptativa con respecto a la señal), lo cual se explicará de forma más detallada a continuación.

[0018] La replicación de banda espectral se basa en el hecho de que para una gran cantidad de propósitos es suficiente transmitir una señal de audio solamente dentro de una banda de frecuencia de núcleo y generar los componentes de señal en la banda de frecuencia superior en el decodificador. La señal de audio resultante mantendrá aún una alta calidad perceptual, dado que para voz y música, por ejemplo, los componentes de alta frecuencia presentan con asiduidad una correlación con respecto a los componentes de baja frecuencia en la banda de frecuencia de núcleo. Por lo tanto, por medio del uso de un algoritmo de parche (“patching”) adaptado, que genera los componentes que faltan de alta frecuencia, es posible obtener una señal de audio de alta calidad perceptual. Al mismo tiempo, la generación dirigida por parámetro de las bandas superiores tiene como resultado una reducción significativa de la tasa de bits para codificar una señal de audio, ya que solamente la señal de audio dentro de la banda de frecuencia de núcleo es codificada, comprimida y transmitida al decodificador. Para los componentes de frecuencia restantes, solamente se transmiten información de control y parámetros de replicación de banda espectral, que controlan el decodificador en el procedimiento de generación de un estimado de la señal de banda alta original. Por lo tanto, en sentido estricto, este procedimiento involucra tres aspectos: (i) la estimación de banda de HF paramétrica (cálculo de parámetro de SBR), (ii) la generación de parche no procesada (parche (“patching”) real) y (iii) disposiciones para procesamiento adicional (por ejemplo, ajuste de umbral mínimo de ruido).

[0019] La banda de frecuencia de núcleo puede ser definida por la denominada frecuencia de cruce, que define un umbral dentro de la banda de frecuencia hasta el cual se lleva a cabo una codificación de la señal de audio. El codificador de núcleo codifica la señal de audio dentro de la banda de frecuencia de núcleo limitada por la frecuencia de cruce. Comenzando por la frecuencia de cruce, los componentes de señal serán generados por la replicación de banda espectral. Al utilizar procedimientos convencionales para la replicación de banda espectral, con frecuencia ocurre que algunas señales comprenden degradaciones no convenientes en la frecuencia de cruce del codificador de núcleo.

[0020] Por medio del uso de las realizaciones de la presente invención, es posible determinar un algoritmo de parche (“patching”), que evita estas degradaciones o al menos modifica estas degradaciones de forma que no presenten un efecto perceptual. Por ejemplo, por medio del uso de reflexión en espejo como algoritmo de parche (“patching”) en el dominio del tiempo, la replicación de banda espectral se lleva a cabo de forma similar a la ampliación de ancho de banda (BWE) dentro de AMR-WB+ (códec de banda ancha de múltiples velocidades adaptativas ampliadas). De forma adicional, la posibilidad de cambiar el algoritmo de parche (patching) dependiendo de la señal ofrece la posibilidad de que puedan utilizarse, por ejemplo, diferentes extensiones de ancho de banda para voz y para

música. No obstante, incluso para una señal que no puede identificarse claramente como música o voz (es decir, una señal mezclada), el algoritmo de parche ("patching") puede cambiarse dentro de períodos cortos de tiempo. Por ejemplo, para cualquier período de tiempo dado, puede utilizarse un algoritmo de parche ("patching") preferido para el parche ("patching"). Este algoritmo de parche ("patching") preferido puede ser determinado por el codificador que puede 5 comparar, por ejemplo, para cada bloque procesado de datos de entrada, los resultados de parche ("patching") con la señal de audio original. Esto mejora significativamente la calidad perceptiva de la señal de audio resultante generada por el sintetizador de señales de audio.

[0021] Otras ventajas de la presente invención se deben a la separación del generador de parche ("patching") del procesador de señales no procesadas, que puede comprender herramientas de SBR estándar. Debido a esta separación, pueden emplearse las herramientas de SBR comunes, que pueden comprender un filtrado inverso, suma de un umbral mínimo de ruido o armónicos que faltan u otros. Por lo tanto, las herramientas de SBR estándar aún pueden utilizarse mientras el parche ("patching") puede ajustarse de forma flexible. De forma adicional, dado que las herramientas de SBR estándar son utilizadas en el dominio frecuencial, la separación del generador de parche 15 de las herramientas de SBR permite calcular el parche ("patching") ya sea en el dominio frecuencial o bien en el dominio del tiempo.

Breve descripción de los dibujos

20 **[0022]** La presente invención se describirá ahora a modo de ejemplos ilustrados. Puede tenerse una apreciación más sencilla y una mejor comprensión de las características de la invención con referencia a la siguiente descripción detallada, que debe ser considerada con referencia a los dibujos adjunto, en los cuales:

25 La figura 1 muestra un diagrama de bloque de un procesamiento de señal de audio según las realizaciones de la presente invención;

La figura 2 muestra un diagrama de bloque para el generador de parche según las realizaciones;

30 La figura 3 muestra un diagrama de bloque para el combinador que opera en el dominio del tiempo;

Las figuras 4a a 4d ilustran, de forma esquemática, ejemplos para diferentes algoritmos de parche ("patching");

Las figuras 5a a 5b ilustran el vocoder de fase y el parche ("patching") por copiado;

35 Las figuras 6a a 6d muestran diagramas de bloque para el procesamiento del flujo de audio codificado en muestras de PCM de salida; y

Las figuras 7a a 7c muestran diagramas de bloque para un codificador de audio según otras realizaciones.

40 Descripción detallada de la invención

[0023] Las realizaciones descritas a continuación son meramente ilustrativas para los principios de la presente invención con el fin de mejorar la replicación de banda espectral, por ejemplo, utilizada con un decodificador de audio. Se comprende que las modificaciones y las variantes de las disposiciones y los detalles descritos en esta invención 45 serán evidentes para los expertos en la materia. Por lo tanto, están destinadas a no estar limitadas por los detalles específicos presentados a modo de descripción y explicación de las realizaciones en esta invención.

[0024] Todas las siguientes apariciones de la palabra "realización (realizaciones)", si se refiere a combinaciones de características diferentes de las definidas por las reivindicaciones independientes, consulte los ejemplos que fueron presentados originalmente pero que no representan realizaciones de la invención reivindicada 50 actualmente; estos ejemplos todavía se muestran con fines ilustrativos únicamente.

[0025] La figura 1 muestra un sintetizador de señales de audio para generar una señal de audio de síntesis 105 que tiene una primera banda de frecuencia y una segunda banda de frecuencia replicada derivada de la primera banda de frecuencia. El sintetizador de señales de audio comprende un generador de parche 110 para llevar a cabo al menos dos algoritmos de parche ("patching") diferentes, donde cada algoritmo de parche ("patching") genera una señal no procesada 115 que tiene componentes de señal en la segunda banda de frecuencia replicada utilizando la señal de audio 105 que tiene componentes de señal en la primera banda de frecuencia. El generador de parche 110 está adaptado para seleccionar uno de los al menos dos algoritmos de parche ("patching") diferentes en respuesta a 60 información de control 112 para una primera parte de tiempo y el otro de los al menos dos algoritmos de parche ("patching") diferentes en respuesta a la información de control 112 para una segunda parte de tiempo que se diferencia de la primera parte de tiempo con el fin de obtener la señal no procesada 115 para la primera parte de tiempo y la segunda parte de tiempo. El sintetizador de señales de audio comprende, además, un conversor espectral 120 para convertir la señal no procesada 115 en una representación espectral no procesada 125 que comprende 65 componentes en una primera sub-banda, una segunda sub-banda, y así sucesivamente. El sintetizador de señales de

audio comprende, además, el procesador de señales no procesadas 130 para procesar la representación espectral no procesada 125 en respuesta a parámetros de replicación de banda espectral de dominio espectral 132 con el fin de obtener una representación espectral de señal no procesada ajustada 135. El sintetizador de señales de audio comprende, además, un combinador 140 para combinar la señal de audio 105 que tiene componentes de señal en la primera banda o una señal derivada de la señal de audio 105 con la representación espectral de señal no procesada ajustada 135 o con otra señal derivada de la representación espectral de señal no procesada ajustada 135 con el fin de obtener una señal de audio de síntesis 145.

[0026] En otras realizaciones, el combinador 140 está adaptado para utilizar como la señal derivada de la señal de audio 105, la representación espectral de señal no procesada 125. La señal derivada de la señal de audio utilizada por el combinador puede ser, además, la señal de audio procesada por un conversor de tiempo/espectral, tal como un banco de filtros de análisis o una señal de banda baja de la forma generada por medio de un generador de parche que opera en el dominio del tiempo o en el dominio espectral o una señal de audio diferida o la señal de audio procesada por una operación de sobremuestreo de forma que las señales que deben combinarse cuenten con la misma tasa de muestreo básica.

[0027] Incluso en otra realización, el sintetizador de señales de audio comprende, además, un analizador para analizar una característica de la señal de audio 105 que tiene componentes de señal en la primera banda de frecuencia 201 y para proporcionar la información de control 112, que identifica el primer algoritmo de parche ("patching") o el segundo algoritmo de parche ("patching").

[0028] En otras realizaciones, el analizador está adaptado para identificar un algoritmo de parche no armónico para una parte de tiempo con un grado de voz o un algoritmo de parche armónico para una parte de tiempo específica en la señal de audio 105.

[0029] Incluso en otras realizaciones, la señal de audio 105 está codificada junto con metadatos en un flujo de datos, y en la que el generador de parche 110 está adaptado para obtener la información de control 112 a partir de los metadatos en el flujo de datos.

[0030] Incluso en otras realizaciones, el conversor espectral 120 comprende un banco de filtros de análisis o los al menos dos algoritmos de parche ("patching") diferentes comprenden un algoritmo de vocoder de fase o un algoritmo de parche ("patching") de sobremuestreo o un algoritmo de parche ("patching") de distorsión lineal o un algoritmo de copiado.

[0031] Incluso en otras realizaciones, el procesador de señales no procesadas 130 está adaptado para llevar a cabo un ajuste de energía de las bandas espectrales o un filtrado inverso en las bandas espectrales o para sumar un umbral mínimo de ruido a la banda espectral o para sumar armónicos que faltan a la banda espectral.

[0032] La figura 2 muestra un diagrama de bloque que proporciona más detalles del generador de parche 110 que comprende un controlador, que recibe la información de control 112 y la señal de audio 105, y medios de parche ("patching") 113. El controlador 111 está adaptado para seleccionar un algoritmo de parche basado en la información de control 112. El generador de parche 110 comprende un primer medio de parche ("patching") 113a que lleva a cabo un primer algoritmo 1, un segundo medio de parche ("patching") 113b que lleva a cabo un segundo algoritmo 2, y así, sucesivamente. En general, el generador de parche 110 comprende tantos medios de parche ("patching") 113 como algoritmos de parche ("patching") haya disponibles. Por ejemplo, el generador de parche ("patching") 110 puede comprender dos, tres, cuatro o más de cuatro medios de parche ("patching") 113. Después de que el controlador 111 haya seleccionado, sobre la base de la información de control 112, uno de los medios de parche ("patching") 113, el controlador 111 envía la señal de audio 105 a uno de los medios de parche ("patching") 113, que lleva a cabo el algoritmo de parche ("patching") y da salida a la señal no procesada 115, que comprende componentes de señal en las bandas de frecuencia replicadas 202, 203.

[0033] La figura 3 muestra un diagrama de bloque que proporciona más detalles para el combinador 140, en el que el combinador 140 comprende un banco de filtros de síntesis 141, un retardador 143 y un sumador 147. Se da entrada a la señal no procesada ajustada 135 al banco de filtros de síntesis 141, que genera, a partir de la señal no procesada ajustada 135 (por ejemplo, en la representación espectral), una señal no procesada ajustada dentro del dominio del tiempo 135t (señal no procesada de dominio del tiempo). La señal de audio de banda base 105 se introduce en el retardador 143, que está adaptado para diferir la señal de banda base 105 por un cierto período de tiempo y emite la señal de banda base diferida 105d. La señal de banda base diferida 105d y la señal no procesada ajustada de dominio del tiempo 135t son sumadas por el sumador 147, lo cual produce la señal de audio de síntesis 145, que es emitida fuera del combinador 140. El retardo en el retardador 143 depende del algoritmo de procesamiento del sintetizador de señales de audio con el fin de lograr que la señal no procesada ajustada de dominio del tiempo 135t corresponda al mismo tiempo que la señal de banda base diferida 105d (sincronización).

[0034] Las figuras 4a a 4d muestran diferentes algoritmos de parche ("patching") utilizados en el generador de parche 110 por los medios de parche ("patching") 113. Según lo explicado anteriormente, el algoritmo de parche

("patching") genera una señal parcheada en la banda de frecuencia replicada. En las realizaciones según lo mostrado en la figura 4, una primera banda de frecuencia 201 se extiende a la frecuencia de cruce $f_{m\acute{a}x}$ en la que una segunda banda de frecuencia 202 (o una segunda banda de frecuencia replicada) se inicia y extiende al doble de la frecuencia de cruce $2*f_{m\acute{a}x}$. Más allá de esta frecuencia, comienza una tercera banda de frecuencia 203 (o tercera banda de frecuencia replicada). La primera banda de frecuencia 201 puede comprender la banda de frecuencia de núcleo mencionada anteriormente.

[0035] En la figura 4, se muestran cuatro algoritmos de parche ("patching") a modo de ejemplos. El primer algoritmo de parche ("patching") en la figura 4a comprende una reflexión o muestreo ascendente, un segundo algoritmo de parche ("patching") comprende un copiado o modulado y se muestra en la figura 4b, un tercer algoritmo de parche ("patching") comprende un vocoder de fase que se muestra en la figura 4c, y un cuarto algoritmo de parche ("patching") que comprende una distorsión se muestra en la figura 4d.

[0036] La reflexión de espejo según lo mostrado en la figura 4a se lleva a cabo de forma que la señal parcheada en la segunda banda de frecuencia 202 sea obtenida por reflexión de la primera banda de frecuencia 201 en la frecuencia de cruce $f_{m\acute{a}x}$. La señal parcheada en la tercera banda de frecuencia 203 se obtiene, a su vez, por reflexión de la señal en la segunda banda de frecuencia 202. Dado que la señal en la segunda banda de frecuencia 202 ya era una señal reflejada, la señal en la tercera banda de frecuencia 203 también puede obtenerse simplemente por desplazamiento de la señal de audio 105 en la primera banda de frecuencia 201 a la tercera banda de frecuencia 203.

[0037] Un segundo algoritmo de parche ("patching") según lo mostrado en la figura 4 implementa el copiado (o modulado) de la señal. En esta realización, la señal en la segunda banda de frecuencia 202 se obtiene por desplazamiento (copiado) de la señal en la primera banda de frecuencia 201 a la segunda banda de frecuencia 202. De forma similar, también la señal en la tercera banda de frecuencia 203 se obtiene por desplazamiento de la señal en la primera banda de frecuencia 201 a la tercera banda de frecuencia 203.

[0038] La figura 4c muestra una realización que utiliza un vocoder de fase como algoritmo de parche ("patching"). La señal parcheada es generada por etapas posteriores, en las que una primera etapa genera componentes de señal hasta el doble de la frecuencia máxima $2*f_{m\acute{a}x}$ y una segunda etapa genera componentes de señal hasta el triple de la frecuencia máxima $3*f_{m\acute{a}x}$ y así, sucesivamente. Un vocoder de fase multiplica las frecuencias de muestras con un factor n ($n = 2, 3, 4, \dots$), lo cual produce una dispersión de los valores de muestra sobre un intervalo de frecuencia de n veces de la banda de frecuencia de núcleo (primera banda de frecuencia 201).

[0039] El algoritmo de parche ("patching") que utiliza distorsión (por ejemplo, por encuadre de la señal) se muestra en la figura 4d. Las distorsiones pueden obtenerse de diferentes formas. Una forma sencilla es encuadrando el nivel de señal que genera componentes de frecuencia más alta. Otra posibilidad de distorsión se obtiene por recorte (por ejemplo, cortando la señal por encima de un cierto umbral). Además, en este caso, se generarán componentes de alta frecuencia. Básicamente, aquí puede utilizarse cualquier distorsión conocida en procedimientos convencionales.

[0040] La figura 5a muestra, de forma más detallada, los algoritmos de parche ("patching") de un vocoder de fase. La primera banda de frecuencia 201 se extiende una vez más hasta la frecuencia máxima $f_{m\acute{a}x}$ (frecuencia de cruce) en la cual comienza la segunda banda de frecuencia 202, que finaliza, por ejemplo, en el doble de la máxima frecuencia $2*f_{m\acute{a}x}$. Después de la segunda banda de frecuencia 202, se inicia la tercera banda de frecuencia 203 y puede extenderse, por ejemplo, hasta el triple de la frecuencia máxima $3*f_{m\acute{a}x}$.

[0041] Para mayor simplicidad, la figura 5a muestra un espectro (nivel P como función de la frecuencia f) con ocho líneas de frecuencia 105a, 105b, ..., 105h para la señal de audio 105. A partir de estas ocho líneas 105a, ..., 105h, el vocoder de fase genera una nueva señal por medio de un desplazamiento de las líneas según las flechas mostradas. El desplazamiento corresponde a la multiplicación mencionada anteriormente. En detalle, la primera línea 105a se desplaza a la segunda línea 105b, la segunda línea se desplaza a la cuarta línea, y así sucesivamente, hasta la octava línea 105h, que se desplaza a la línea 16 (última línea en el segundo dominio frecuencial 202). Esto corresponde a la multiplicación por dos. Con el fin de generar líneas hasta el triple de la frecuencia máxima, $3*f_{m\acute{a}x}$, todas las frecuencias de las líneas pueden ser multiplicadas por tres, es decir, la primera línea 105a se desplaza a la tercera línea 105c, la segunda línea 105b se desplaza a la sexta línea, y así sucesivamente, hasta la octava línea 105h, que se desplaza hasta la línea 24 (la última línea en la tercera banda de frecuencia 203). Es obvio que, por medio de este vocoder de fase, las líneas ya no son equidistantes, pero se dispersan para mayores frecuencias.

[0042] La figura 5b muestra el parche ("patching") de copiado con mayor detalle. Una vez más, se muestra el nivel P como función de la frecuencia f , en el que ocho líneas están en la primera banda de frecuencia 201, que se copian en la segunda banda de frecuencia 202 y, asimismo, en la tercera banda de frecuencia 203. Este copiado solamente implica que la primera línea 105a en la primera banda de frecuencia 201 también se vuelve la primera línea en la segunda banda de frecuencia 202 y en la tercera banda de frecuencia 203. Por lo tanto, las primeras líneas de cada una de las bandas de frecuencia replicadas 202 y 203 se copian de la misma línea en la primera banda de frecuencia 201. De forma análoga, esto se aplica, además, a las otras líneas. En consecuencia, se copia toda la banda

de frecuencia.

[0043] Los diferentes algoritmos de parche ("patching") como se muestra en las figuras 4 y 5 pueden aplicarse de manera diferente, ya sea dentro del dominio del tiempo o bien en el dominio frecuencial y comprenden diferentes 5 ventajas o desventajas, que pueden explotarse para diferentes aplicaciones.

[0044] Por ejemplo, la reflexión en el dominio frecuencial se muestra en la figura 4a. En el dominio del tiempo, la reflexión puede llevarse a cabo aumentando la tasa de muestreo por un factor entero, que puede efectuarse por medio de la inserción de muestras adicionales entre cada par de muestras existentes. Estas muestras adicionales no 10 se obtienen de la señal de audio, sino que son introducidas por el sistema y comprende, por ejemplo, valores cercanos a cero o iguales que cero. En el caso más sencillo, si se introduce solamente una muestra adicional entre dos muestras existentes, se logra una duplicación de la cantidad de muestras lo cual implica una duplicación de la tasa de muestreo. Si se introduce más de una muestra adicional (por ejemplo, de forma equidistante), la tasa de muestreo aumentará según ello y, por lo tanto, aumentará el espectro de frecuencia. En general, la cantidad de muestras adicionales entre 15 cada dos muestras existentes puede ser cualquier cantidad n ($n = 2, 3, 4, \dots$) que aumente la tasa de muestreo por el factor $n+1$. La inserción de las muestras adicionales produce la reflexión del espectro de frecuencia en la frecuencia *Nyquist*, que especifica la frecuencia más alta susceptible de ser representada a una tasa de muestreo dada. El dominio frecuencial del espectro de banda base (espectro en la primera banda de frecuencia) se refleja así de forma directa por medio de este procedimiento en la siguiente banda de frecuencia. De forma opcional, esta reflexión puede 20 combinarse con un filtrado de paso bajo y/o formado espectral posibles.

[0045] Las ventajas de este algoritmo de parche ("patching") pueden resumirse de la siguiente manera. Utilizando este procedimiento, la estructura de tiempo de señal se conserva mejor que utilizando procedimientos similares en el dominio frecuencial. Asimismo, por reflexión espectral, las líneas de frecuencia cercanas a la frecuencia 25 *Nyquist* son mapeadas sobre líneas, que también están cerca de la frecuencia *Nyquist*. Esta es una ventaja, ya que, después de una reflexión, las regiones espectrales alrededor de la frecuencia de reflexión (es decir, la frecuencia *Nyquist* de la señal de audio original 105) son similares en muchos aspectos como, por ejemplo, con respecto a la propiedad de la uniformidad espectral, la propiedad tonal, la acumulación o definición de puntos de frecuencia, etc. Por medio de este procedimiento, el espectro es continuado hasta la siguiente banda de frecuencia de una forma más 30 moderada como, por ejemplo, utilizando las técnicas de copiado, en las que las regiones de frecuencia finalizan una cerca de otra, que se originan desde regiones completamente diferentes en el espectro original y así, muestran características muy diferentes. En copiado: la primera muestra pasa a ser una vez más la primera muestra en la banda replicada, mientras que, en la reflexión, la última muestra pasa a ser la primera muestra en la banda replicada. Esta continuación más fluida del espectro puede reducir, a su vez, las degradaciones perceptuales, que son provocadas 35 por características no continuas del espectro reconstruido generado por otros algoritmos de parche ("patching").

[0046] Finalmente, hay señales que comprenden una mayor cantidad de armónicos, por ejemplo, en la región de frecuencia más baja (primera banda de frecuencia 201). Estos armónicos aparecen como picos localizados en el espectro. No obstante, en la parte superior del espectro, puede haber solamente muy pocos armónicos presentes o, 40 en otras palabras, la cantidad de armónicos es menor en la parte superior del espectro. Utilizando simplemente un copiado del espectro, esto tendrá como resultado una señal replicada en la que la parte inferior del espectro con una gran cantidad de armónicos es copiada de forma directa en la región de frecuencia superior, donde había solamente muy pocos armónicos en la señal original. Como resultado, la banda de frecuencia superior de la señal original y la señal replicada son muy diferentes con respecto a la cantidad de armónicos, lo cual no se considera conveniente y 45 debería evitarse.

[0047] El algoritmo de parche ("patching") de reflexión puede aplicarse, además, en el dominio frecuencial (por ejemplo, en la región de QMF), en cuyo caso se invierte el orden en las bandas de frecuencia de forma que se presente un reordenamiento desde atrás hacia delante. De forma adicional, para muestras de sub-banda, debe formarse un 50 valor del complejo conjugado de forma que la parte imaginaria de cada muestra cambie su signo. Esto produce una inversión del espectro dentro de la sub-banda.

[0048] Este algoritmo de parche ("patching") comprende una alta flexibilidad con respecto a los bordes del parche ("patching"), dado que no necesariamente debe llevarse a cabo una reflexión del espectro en la frecuencia 55 *Nyquist*, pero puede llevarse a cabo, además, en cualquier borde de sub-banda.

[0049] No obstante, la cancelación de la generación de señal ajena (aliasing) entre bandas de QMF vecinas en los límites de parches puede no presentarse, lo cual puede ser tolerable o no.

60 **[0050]** Por medio de dispersión o del uso del vocoder de fase (véase la figura 4c o 5a), la estructura de frecuencia se extiende de forma armónica y correcta en el dominio de alta frecuencia, ya que la banda base 201 se dispersa de forma espectral por medio de un múltiple par llevado a cabo por uno o más vocoders de fase, y dado que los componentes espectrales en la banda base 201 se combinan con los componentes espectrales generados 65 adicionales.

[0051] Este algoritmo de parche ("patching") se prefiere si la banda base 201 ya está altamente limitada con respecto al ancho de banda, por ejemplo, utilizando solamente una tasa de bits muy baja. Por lo tanto, la reconstrucción de los componentes de frecuencia superior ya se inicia en una frecuencia relativamente baja. En este caso, una frecuencia de cruce típica es menor que 5 kHz aproximadamente (o incluso menor que 4 kHz). En esta región, el oído humano es muy sensible a disonancias que se deben a armónicos posicionados de forma incorrecta. Esto puede tener como resultado la impresión de tonos "no naturales". De forma adicional, los tonos espaciados unos cerca de otros en el espectro (con una distancia espectral de entre 30 Hz y 300 Hz aproximadamente) se perciben como tonos ásperos. Una continuación armónica de la estructura de frecuencia de la banda base 201 evita estas impresiones auditivas incorrectas y desagradables.

10

[0052] En el tercer algoritmo de parche ("patching") de copiado (véase la figura 4a o 5b), las regiones espectrales son copiadas por sub-banda en una región de frecuencia más alta o en la región de frecuencia que debe replicarse. Asimismo, el copiado se basa en la observación de que las propiedades espectrales de las señales de frecuencia más alta son similares en diversos aspectos a las propiedades de las señales de banda base, lo cual es así para todos los procedimientos de parche ("patching"). Hay solamente muy pocas desviaciones entre sí. De forma adicional, el oído humano típicamente no es muy sensible a una alta frecuencia (que comienza típicamente a 5 kHz aproximadamente), en especial, con respecto a un mapeo espectral no preciso. En realidad, esta es la idea clave de la replicación de la banda espectral en general. El copiado comprende, en particular, la ventaja de que su implementación es sencilla y rápida.

20

[0053] Este algoritmo de parche ("patching") presenta, además, una alta flexibilidad con respecto a los bordes del parche ("patching"), dado que el copiado del espectro puede llevarse a cabo en cualquier borde de sub-banda.

[0054] Finalmente, el algoritmo de parche ("patching") de distorsión (véase la figura 4d) puede comprender la generación de armónicos por recorte, limitación, encuadre, etc. Por ejemplo, si una señal dispersa es ocupada de manera muy tenue de forma espectral (por ejemplo, después de aplicar el algoritmo de parche ("patching") de vocoder de fase mencionado anteriormente), es posible que el espectro disperso pueda ser opcionalmente complementado de manera adicional por una señal distorsionada con el fin de evitar huecos de frecuencia no convenientes.

[0055] Las figuras 6a a 6d muestran diferentes realizaciones para el sintetizador de señales de audio implementadas en un decodificador de audio.

[0056] En la realización mostrada en la figura 6a, un flujo de audio codificado 345 se introduce en un desformatador de carga útil de flujo de bits 350, que separa, por un lado, una señal de audio codificada 355 y, por otro, información adicional 375. Por ejemplo, la señal de audio codificada 355 se introduce en un decodificador de núcleo de AAC 360, que genera la señal de audio decodificada 105 en la primera banda de frecuencia 201. La señal de audio 105 se introduce en un banco de QMF de 32 bandas de análisis 370, que comprende, por ejemplo, 32 bandas de frecuencia y que genera la señal de audio 105₃₂ en el dominio frecuencial. Se prefiere que el generador de parche solamente emita una señal de banda alta como la señal no procesada y no de salida a la señal de banda baja. De forma alternativa, si el algoritmo de parche ("patching") en un bloque 110 también genera la señal de banda baja, se prefiere someter la señal de entrada a filtrado de paso alto en un bloque 130a.

[0057] La señal de audio de dominio frecuencial 105₃₂ se introduce en el generador de parche 110, que, en esta realización, genera el parche dentro del dominio frecuencial (dominio de QMF). La representación espectral de señal no procesada resultante 125 se introduce en una herramienta de SBR 130a, que puede generar, por ejemplo, un umbral mínimo de ruido, reconstruir armónicos que faltan o llevar a cabo un filtrado inverso.

[0058] Por otro lado, la información adicional 375 se introduce en un analizador sintáctico de flujo de bits 380, que analiza la información adicional para obtener subinformación diferente 385 e introducirla, por ejemplo, en una unidad de descuantificación y decodificación Huffman 390 que, por ejemplo, extrae la información de control 112 y los parámetros de replicación de banda espectral 132. La información de control 112 se introduce en la herramienta de SBR y los parámetros de replicación de banda espectral 132 se introducen en la herramienta de SBR 130a así como en un ajustador de envolvente 130b. El ajustador de envolvente 130b es operativo con el fin de ajustar la envolvente para el parche generado. Como resultado, el ajustador de envolvente 130b genera la señal no procesada ajustada 135 y la introduce en un banco de QMF de síntesis 140, que combina la señal no procesada ajustada 135 con la señal de audio en el dominio frecuencial 105₃₂. El banco de QMF de síntesis puede comprender, por ejemplo 64 bandas de frecuencia y genera, por medio de una combinación de ambas señales (la señal no procesada ajustada 135 y la señal de audio de dominio frecuencial 105₃₂), la señal de audio de síntesis 145 (por ejemplo, una salida de muestras de PCM, PCM = modulación por impulsos codificados).

60

[0059] De forma adicional, la figura 6a muestra las herramientas de SBR 130a, que pueden implementar procedimientos de replicación de banda espectral conocidos para ser utilizados en la salida de datos espectrales de QMF del generador de parche 110. El algoritmo de parche ("patching") utilizado en el dominio frecuencial como se muestra en la figura 6a puede emplear, por ejemplo, la simple reflexión o copiado de los datos espectrales dentro del dominio frecuencial (véase la figura 4a y la figura 4b).

65

- [0060]** Así, esta estructura general está según decodificadores convencionales conocidos en la técnica anterior, pero las realizaciones reemplazan el generador de parche convencional por el generador de parche 110, configurado para llevar a cabo diferentes algoritmos de parche ("patching") adaptados con el fin de mejorar la calidad perceptual de la señal de audio. De forma adicional, las realizaciones pueden utilizar, además, un algoritmo de parche ("patching") dentro del dominio del tiempo y no necesariamente el parche ("patching") en el dominio frecuencial como se muestra en la figura 6a.
- [0061]** La figura 6b muestra realizaciones de la presente invención en las que el generador de parche ("patching") 110 puede utilizar un algoritmo de parche ("patching") dentro del dominio frecuencial, así como dentro del dominio del tiempo. El decodificador como se muestra en la figura 6b comprende, de nuevo, el desformateador de carga útil de flujo de bits 350, el decodificador de núcleo de AAC 360, el analizador sintáctico de flujo de bits (bit stream parser) 380, y la unidad de descuantificación y decodificación Huffman 390. Por lo tanto, en la realización como se muestra en la figura 6b, el flujo de audio codificado 345 se introduce de nuevo en el desformateador de carga útil de flujo de bits 350 que, por otro lado, genera la señal de audio codificada 355 y separa de ella la información adicional 375, que es analizada a continuación de forma sintáctica por el analizador sintáctico de flujo de bits 380 para separar la información diferente 385, que se introduce en la unidad de descuantificación y decodificación Huffman 390. Por otro lado, la señal de audio codificada 355 se introduce en el decodificador de núcleo de AAC 360.
- [0062]** Ahora, las realizaciones diferencian los dos casos: el generador de parche 110 opera ya sea dentro del dominio frecuencial (siguiendo las líneas de puntos de señal) o bien, dentro del dominio del tiempo (siguiendo las líneas de guiones de la señal).
- [0063]** Si el generador de parche opera en el dominio del tiempo, la salida del decodificador de núcleo de AAC 360 se introduce en el generador de parche 110 (línea de puntos para la señal de audio 105) y su salida se transmite al banco de filtros de análisis 370. La salida del banco de filtros de análisis 370 es la representación espectral de señal no procesada 125, que se introduce en las herramientas de SBR 130a (que son parte del ajustador de señal no procesada 130) así como en el banco de QMF de síntesis 140.
- [0064]** Por otro lado, si el algoritmo de parche ("patching") utiliza el dominio frecuencial (como se muestra en la figura 6a), la salida del decodificador de núcleo de AAC 360 se introduce en el banco de QMF de análisis 360 a través de la línea de puntos para la señal de audio 105 que, a su vez, genera una señal de audio de dominio frecuencial 105_{32} y transmite la señal de audio 105_{32} al generador de parche 110 y al banco de QMF de síntesis 140 (líneas de puntos). El generador de parche 110 genera una vez más una representación de señal no procesada 125 y transmite esta señal a las herramientas de SBR 130a.
- [0065]** Por lo tanto, la realización lleva a cabo ya sea un primer modo de procesamiento utilizando las líneas de puntos (parche ("patching") de dominio frecuencial) o bien, un segundo modo de procesamiento utilizando las líneas de guiones (parche ("patching") de dominio del tiempo), donde todas las líneas continuas entre otros elementos funcionales se utilizan en ambos modos de procesamiento.
- [0066]** Se prefiere que el modo de procesamiento de tiempo del generador de parche (líneas de guiones) sea de forma que la salida del generador de parche incluya la señal de banda baja y la señal de banda alta, es decir, que la señal de salida del generador de parche sea una señal difundida que consista en la señal de banda baja y la señal de banda alta. La señal de banda baja se introduce en el bloque 140 y la señal de banda alta se introduce en el bloque 130a. Las separaciones de banda pueden llevarse a cabo en el banco de análisis 370, pero también pueden efectuarse de forma alternativa. Asimismo, la señal de salida del decodificador de AAC puede aplicarse de forma directa al bloque 370 de forma que la porción de banda baja de la señal de salida del generador de parche no sea utilizada en absoluto y la porción de banda baja original sea utilizada en el combinador 140.
- [0067]** En el modo de procesamiento del dominio frecuencial (líneas de puntos), el generador de parche solamente emite, preferentemente, la señal de banda alta, y la señal de banda baja original es aplicada de forma directa al bloque 370 para aplicación del banco de síntesis 140. De forma alternativa, el generador de parche puede generar, además, una señal de salida de ancho de banda completo y aplicar la señal de banda baja al bloque 140.
- [0068]** De nuevo, la unidad de descuantificación y decodificación Huffman 390 genera el parámetro de replicación de banda espectral 132 y la información de control 112, que se introduce en el generador de parche 110. De forma adicional, los parámetros de replicación de banda espectral 132 son transmitidos al ajustador de envolvente 130b, así como a las herramientas de SBR 130a. La salida del ajustador de envolvente 130b es la señal no procesada ajustada 135 que se combina en el combinador 140 (banco de QMF de síntesis) con la señal de audio de banda espectral 105_{32} (para el parche ("patching") de dominio frecuencial) o con una representación espectral de señal no procesada 125 (para el parche ("patching") de dominio del tiempo) con el fin de generar la señal de audio de síntesis 145, que puede comprender, de nuevo, muestras de PCM de salida.
- [0069]** Además, en esta realización, el generador de parche 110 utiliza uno de los algoritmos de parche

("patching") (por ejemplo, como se muestra en las figuras 4a a 4d) con el fin de generar la señal de audio en la segunda banda de frecuencia 202 o la tercera banda de frecuencia 203, utilizando la señal de banda base en la primera banda de frecuencia 201. Solamente las muestras de señal de audio dentro de la primera banda de frecuencia 201 son codificadas en el flujo de salida codificado 345 y las muestras que faltan son generadas utilizando el procedimiento de replicación de banda espectral.

[0070] La figura 6c muestra una realización para el algoritmo de parche ("patching") dentro del dominio del tiempo. En comparación con la figura 6a, la realización como se muestra en la figura 6c se diferencia por la posición del generador de parche 110 y el banco de QMF de análisis 120. Todos los componentes restantes del sistema de decodificación son los mismos que los que se muestran en la figura 6a y, por lo tanto, aquí se omite una descripción repetida.

[0071] El generador de parche 110 recibe la señal de audio 105 desde el decodificador de núcleo de AAC 360 y ahora lleva a cabo el parche ("patching") dentro del dominio del tiempo para generar la señal no procesada 115, que se introduce en el conversor espectral 120 (por ejemplo, un banco de QMF de análisis, que comprende 64 bandas). De varias posibilidades, un algoritmo de parche ("patching") en el dominio del tiempo llevado a cabo por el generador de parche ("patching") 110 tiene como resultado una señal no procesada 115 que comprende la tasa de muestreo duplicada, si el generador de parche ("patching") 110 lleva a cabo el parche ("patching") por medio de la introducción de muestras adicionales entre muestras existentes (que por ejemplo son cercanas a valores cero). La salida del conversor espectral 120 es la representación espectral de señal no procesada 125, que se introduce en el ajustador de señal no procesada 130, que comprende, de nuevo, por un lado, la herramienta de SBR 130a y por otro, el ajustador de envolvente 130b. En lo que respecta a las realizaciones mostradas anteriormente, la salida del ajustador de envolvente comprende la señal no procesada ajustada 135 que se combina con la señal de audio en el dominio frecuencial 105f en el combinador 140 que, de nuevo, comprende, por ejemplo, un banco de QMF de síntesis de 64 bandas de frecuencia.

[0072] Por lo tanto, la principal diferencia es que, por ejemplo, la reflexión se lleva a cabo en el dominio del tiempo y los datos de frecuencia superior ya están reconstruidos antes de introducir la señal 115 en el banco de filtros de análisis de 64 bandas 120, lo cual significa que la señal ya comprende la velocidad sometida a muestreo duplicada (en la SBR de velocidad dual). Después de esta operación de parche ("patching"), puede emplearse una herramienta de SBR común, que puede comprender, de nuevo, un filtrado inverso, una suma de un umbral mínimo de ruido o una suma de armónicos que faltan. Aunque la reconstrucción de la región de alta frecuencia se presenta en el dominio del tiempo, se lleva a cabo un análisis/síntesis en el dominio de QMF de forma que los mecanismos de SBR restantes aún puedan utilizarse.

[0073] En la realización de la figura 6c, el generador de parche emite, preferentemente, una señal de banda completa que comprende la señal de banda baja y la señal de banda alta (señal no procesada). De forma alternativa, el generador de parche solamente emite la porción de banda alta, por ejemplo, obtenida por filtrado de paso alto, y el banco de QMF 120 es aplicado directamente por la salida del decodificador de núcleo de AAC 105.

[0074] En otra realización, el generador de parche 110 comprende una interfaz de entrada del dominio del tiempo y/o una interfaz de salida del dominio del tiempo (interfaz del dominio del tiempo), y el procesamiento dentro de este bloque puede tener lugar en cualquier dominio tal como un dominio de QMF o un dominio frecuencial tal como DFT, FFT, DCT, DST o cualquier otro dominio frecuencial. A continuación, la interfaz de entrada del dominio del tiempo se conecta con un conversor de tiempo/frecuencia o, en general, un conversor, para efectuar una conversión desde el dominio del tiempo hasta una representación espectral. A continuación, la representación espectral es procesada utilizando al menos dos algoritmos de parche ("patching") diferentes que operan sobre datos del dominio frecuencial. De forma alternativa, un primer algoritmo de parche ("patching") opera en el dominio frecuencial y un segundo algoritmo de parche ("patching") opera en el dominio del tiempo. Los datos de dominio frecuencial interconectado se vuelven a convertir en una representación del dominio del tiempo, que a continuación se introduce en un bloque 120 a través de la interfaz de salida del dominio del tiempo. En la realización, en la que la señal sobre la línea 115 no comprende la banda completa, pero comprende solamente la banda baja, el filtrado se lleva a cabo, preferentemente, en el dominio espectral antes de volver a convertir la señal espectral en el dominio del tiempo.

[0075] Preferentemente, la resolución espectral en un bloque 110 es mayor que la resolución espectral obtenida por un bloque 120. En una realización, la resolución espectral en un bloque 110 es al menos dos veces tan alta como en el bloque 120.

[0076] Al aislar el algoritmo de parche ("patching") en un bloque funcional separado, que es implementado por esta realización, es posible aplicar procedimientos de replicación espectral arbitrarios completamente independientes del uso de las herramientas de SBR. En una implementación alternativa, también es posible generar el componente de alta frecuencia por parche ("patching") en el dominio del tiempo en paralelo a la entrada de la señal de decodificador de AAC a un banco de filtros de análisis de 32 bandas. Las señales interconectadas y de banda base se combinarán solamente después del análisis de QMF.

65

- 5 **[0077]** La figura 6d muestra una realización como esta, donde el algoritmo es llevado a cabo dentro del dominio del tiempo. De forma similar a la realización como se muestra en la figura 6c, en esta realización, asimismo, la diferencia con la figura 6a comprende la posición del generador de parche 110 así como los bancos de filtros de análisis. En particular, el decodificador de núcleo de AAC 360, el desformateador de carga útil de flujo de bits 350 así como el analizador sintáctico de flujo de bits 380 y la unidad de descuantificación y decodificación Huffman 390 son los mismos que en la realización como se muestra en la figura 6a y, de nuevo, aquí se omite una descripción repetida.
- 10 **[0078]** La realización como se muestra en la figura 6d deriva la señal de audio 105 emitida por el decodificador 360 e introduce la señal de audio 105 en el generador de parche 110 así como en el banco de QMF de análisis de 32 bandas 370. El banco de QMF de análisis de 32 bandas 370 (asimismo, conversor 370) genera otra representación espectral de señal no procesada 123. El generador de parche 110 vuelve a llevar a cabo un parche ("patching") dentro del dominio del tiempo y genera una señal no procesada 115 que se introduce en el conversor espectral 120 que pueden comprender, de nuevo, un banco de filtros de QMF de análisis de 64 bandas. El conversor espectral 120 genera la representación espectral de señal no procesada 125, que en esta realización comprende componentes de frecuencia en la primera banda de frecuencia 201 y las bandas de frecuencia replicadas en la segunda banda de frecuencia y la tercera banda de frecuencia 202, 203. Esta realización comprende, asimismo, un sumador 124, adaptado para sumar la salida del banco de filtros de análisis de 32 bandas 370 y una representación espectral de señal no procesada 125 con el fin de obtener una representación espectral de señal no procesada combinada 126. En general, el sumador 124 puede ser un combinador 124 configurado, además, para restar los componentes de banda base (componentes en la primera banda de frecuencia 201) a la representación espectral de señal no procesada 125. Por lo tanto, el sumador 124 puede estar configurado para sumar una señal invertida o, de forma alternativa, puede comprender un inversor opcional para invertir la señal de salida desde el banco de filtros de análisis de 32 bandas 370.
- 15 **[0079]** Después de esta sustracción ilustrativa de los componentes de banda de frecuencia en la banda de frecuencia base 201, la salida se introduce de nuevo en la herramienta de replicación de banda espectral 130a que, a su vez, reenvía la señal resultante al ajustador de envolvente 130b. El ajustador de envolvente 130b genera, de nuevo, la señal no procesada ajustada 135 que se combina en el combinador 140 con la salida del banco de filtros de análisis de 32 bandas 370, de forma que el combinador 140 combine los componentes de frecuencia parcheados (en la segunda banda de frecuencia y la tercera banda de frecuencia 202 y 203, por ejemplo) con los componentes de banda base emitidos por el banco de filtros de análisis de 32 bandas 370. De nuevo, el combinador 140 puede comprender un banco de filtros de QMF de síntesis de 64 bandas que produce la señal de audio de síntesis que comprende a su vez, por ejemplo, muestras de PCM de salida.
- 20 **[0080]** En la realización de la figura 6d, el generador de parche emite preferentemente una señal de banda completa que comprende la señal de banda baja y la señal de banda alta (señal no procesada). De forma alternativa, el generador de parche solamente emite la porción de banda alta, por ejemplo, obtenida por filtrado de paso alto para aplicación dentro del bloque 120, y el banco de QMF 370 es aplicado directamente por la salida de AAC como se muestra en la figura 6c. Asimismo, no se requiere el substractor 124, y la salida de un bloque 120 es aplicada directamente dentro del bloque 130a, dado que esta señal comprende la banda alta solamente. De forma adicional, el bloque 370 no necesita la salida para el substractor 124.
- 25 **[0081]** En otra realización, el generador de parche 110 comprende una interfaz de entrada de dominio del tiempo y/o una interfaz de salida del dominio del tiempo (interfaz del dominio del tiempo), y el procesamiento dentro de este bloque puede tener lugar en cualquier dominio tal como un dominio de QMF o un dominio frecuencial tal como DFT, FFT, DCT, MDCT, DST o cualquier otro dominio frecuencial. A continuación, la interfaz de entrada del dominio del tiempo se conecta a un conversor de tiempo/frecuencia o, en general, a un conversor para efectuar una conversión desde el dominio del tiempo hasta una representación espectral. A continuación, la representación espectral es procesada utilizando al menos dos algoritmos de parche ("patching") diferentes que operan sobre datos de dominio frecuencial. De forma alternativa, un primer algoritmo de parche ("patching") opera en el dominio frecuencial y un segundo algoritmo de parche ("patching") opera en el dominio del tiempo. Los datos del dominio frecuencial interconectado se vuelven a convertir en una representación del dominio del tiempo, que es introducida a continuación en el bloque 120 a través de la interfaz de salida del dominio del tiempo.
- 30 **[0082]** Preferentemente, la resolución espectral en un bloque 110 es mayor que la resolución espectral obtenida por un bloque 120. En una realización, la resolución espectral en un bloque 110 es al menos dos veces tan alta como en el bloque 120.
- 35 **[0083]** Las figuras 6a a 6d abarcaron la estructura del decodificador y, en especial, la incorporación del generador de parche 110 dentro de la estructura del decodificador. Con el fin de que el decodificador y, en especial, el generador de parche 110 sea capaz de generar o replicar componentes de alta frecuencia, el codificador puede transmitir información adicional al decodificador, en el que la información adicional 112, por un lado, proporciona información de control que puede utilizarse, por ejemplo, para fijar el algoritmo de parche ("patching") y, de forma adicional, el parámetro de replicación de banda espectral 132 que debe ser utilizado por las herramientas de replicación de la banda espectral 130a.
- 40 **[0080]** En la realización de la figura 6d, el generador de parche emite preferentemente una señal de banda completa que comprende la señal de banda baja y la señal de banda alta (señal no procesada). De forma alternativa, el generador de parche solamente emite la porción de banda alta, por ejemplo, obtenida por filtrado de paso alto para aplicación dentro del bloque 120, y el banco de QMF 370 es aplicado directamente por la salida de AAC como se muestra en la figura 6c. Asimismo, no se requiere el substractor 124, y la salida de un bloque 120 es aplicada directamente dentro del bloque 130a, dado que esta señal comprende la banda alta solamente. De forma adicional, el bloque 370 no necesita la salida para el substractor 124.
- 45 **[0081]** En otra realización, el generador de parche 110 comprende una interfaz de entrada de dominio del tiempo y/o una interfaz de salida del dominio del tiempo (interfaz del dominio del tiempo), y el procesamiento dentro de este bloque puede tener lugar en cualquier dominio tal como un dominio de QMF o un dominio frecuencial tal como DFT, FFT, DCT, MDCT, DST o cualquier otro dominio frecuencial. A continuación, la interfaz de entrada del dominio del tiempo se conecta a un conversor de tiempo/frecuencia o, en general, a un conversor para efectuar una conversión desde el dominio del tiempo hasta una representación espectral. A continuación, la representación espectral es procesada utilizando al menos dos algoritmos de parche ("patching") diferentes que operan sobre datos de dominio frecuencial. De forma alternativa, un primer algoritmo de parche ("patching") opera en el dominio frecuencial y un segundo algoritmo de parche ("patching") opera en el dominio del tiempo. Los datos del dominio frecuencial interconectado se vuelven a convertir en una representación del dominio del tiempo, que es introducida a continuación en el bloque 120 a través de la interfaz de salida del dominio del tiempo.
- 50 **[0082]** Preferentemente, la resolución espectral en un bloque 110 es mayor que la resolución espectral obtenida por un bloque 120. En una realización, la resolución espectral en un bloque 110 es al menos dos veces tan alta como en el bloque 120.
- 55 **[0083]** Las figuras 6a a 6d abarcaron la estructura del decodificador y, en especial, la incorporación del generador de parche 110 dentro de la estructura del decodificador. Con el fin de que el decodificador y, en especial, el generador de parche 110 sea capaz de generar o replicar componentes de alta frecuencia, el codificador puede transmitir información adicional al decodificador, en el que la información adicional 112, por un lado, proporciona información de control que puede utilizarse, por ejemplo, para fijar el algoritmo de parche ("patching") y, de forma adicional, el parámetro de replicación de banda espectral 132 que debe ser utilizado por las herramientas de replicación de la banda espectral 130a.
- 60 **[0083]** Las figuras 6a a 6d abarcaron la estructura del decodificador y, en especial, la incorporación del generador de parche 110 dentro de la estructura del decodificador. Con el fin de que el decodificador y, en especial, el generador de parche 110 sea capaz de generar o replicar componentes de alta frecuencia, el codificador puede transmitir información adicional al decodificador, en el que la información adicional 112, por un lado, proporciona información de control que puede utilizarse, por ejemplo, para fijar el algoritmo de parche ("patching") y, de forma adicional, el parámetro de replicación de banda espectral 132 que debe ser utilizado por las herramientas de replicación de la banda espectral 130a.
- 65 **[0083]** Las figuras 6a a 6d abarcaron la estructura del decodificador y, en especial, la incorporación del generador de parche 110 dentro de la estructura del decodificador. Con el fin de que el decodificador y, en especial, el generador de parche 110 sea capaz de generar o replicar componentes de alta frecuencia, el codificador puede transmitir información adicional al decodificador, en el que la información adicional 112, por un lado, proporciona información de control que puede utilizarse, por ejemplo, para fijar el algoritmo de parche ("patching") y, de forma adicional, el parámetro de replicación de banda espectral 132 que debe ser utilizado por las herramientas de replicación de la banda espectral 130a.

[0084] Otras realizaciones comprenden un procedimiento para generar una señal de audio de síntesis 145 con una primera banda de frecuencia y una segunda banda de frecuencia replicada 202 derivada de la primera banda de frecuencia 201. El procedimiento comprende llevar a cabo al menos dos algoritmos de parche ("patching") diferentes, 5 convertir la señal no procesada 115 en una representación espectral de señal no procesada 125 y procesar la representación espectral de señal no procesada 125. Cada algoritmo de parche ("patching") genera una señal no procesada 115 con componentes de señal en la segunda banda de frecuencia replicada 202 utilizando una señal de audio 105 que tiene componentes de señal en la primera banda de frecuencia 201. El parche ("patching") se lleva a cabo de forma que uno de los al menos dos algoritmos de parche ("patching") diferentes sea seleccionado en 10 respuesta a la información de control 112 para una primera parte de tiempo y el otro de los al menos dos algoritmos de parche ("patching") diferentes sea seleccionado en respuesta a la información de control 112 para una segunda parte de tiempo que se diferencia de la primera parte de tiempo con el fin de obtener la señal no procesada 115 para la primera parte de tiempo y la segunda parte de tiempo. El procesamiento de la representación espectral de la señal no procesada 125 se lleva a cabo en respuesta a parámetros de replicación de banda espectral de dominio espectral 15 132 con el fin de obtener una representación espectral de señal no procesada ajustada 135. Finalmente, el procedimiento comprende una combinación de la señal de audio 105 con componentes de señal en la primera banda 201 o una señal derivada de la señal de audio 105 con la representación espectral de señal no procesada ajustada 135 o con otra señal derivada de la representación espectral de señal no procesada ajustada 135 con el fin de obtener la señal de audio de síntesis 145.

20 **[0085]** Las figuras 7a, 7b y 7c comprenden realizaciones del codificador.

[0086] La figura 7a muestra un codificador que codifica una señal de audio 305 para generar la señal de audio codificada 345, que, a su vez, se introduce en los decodificadores como se muestra en las figuras 6a a 6d. El 25 codificador como se muestra en la figura 7a comprende un filtro de paso bajo 310 (o un filtro selectivo con respecto a la frecuencia general) y un filtro de paso alto 320, en el que se introduce la señal de audio 305. El filtro de paso bajo 310 separa el componente de señal de audio dentro de la primera banda de frecuencia 201, mientras que el filtro de paso alto 320 separa los componentes de frecuencia restantes, por ejemplo, los componentes de frecuencia en la segunda banda de frecuencia 202 y otras bandas de frecuencia. Por lo tanto, el filtro de paso bajo 310 genera una 30 señal filtrada de paso bajo 315 y el filtro de paso alto 320 emite una señal de audio filtrada de paso alto 325. La señal de audio filtrada de paso bajo 315 se introduce en un codificador de audio 330, que puede comprender, por ejemplo, un codificador de AAC.

[0087] De forma adicional, la señal de audio filtrada de paso bajo 315 se introduce en un generador de 35 información de control 240, que está adaptado para generar la información de control 112 de forma que pueda identificarse un algoritmo de parche ("patching") preferido, que es seleccionado, a su vez, por el generador de parche 110. La señal de audio filtrada de paso alto 325 se introduce en un generador de datos de banda espectral 328 que genera los parámetros de banda espectral 132, que se introducen, por un lado, en el selector de parche. El codificador de la figura 7a comprende, además, un formateador 343 que recibe la señal de audio codificada desde el codificador 40 de audio 330, el parámetro de replicación de banda espectral 132 desde el generador de datos de replicación de banda espectral 328, y la información de control 112 desde el generador de información de control 340.

[0088] Los parámetros de banda espectral 132 pueden depender del procedimiento de parche ("patching"), es decir, para diferentes algoritmos de parche ("patching"), los parámetros de banda espectral pueden ser diferentes o 45 no, y puede no ser necesario determinar el parámetro de SBR 132 para todos los algoritmos de parche ("patching") (la figura 7c a continuación muestra una realización, donde solamente debe calcularse un conjunto de parámetros de SBR 132). Por lo tanto, el generador de banda espectral 328 puede generar diferentes parámetros de banda espectral 132 para diferentes algoritmos de parche ("patching") y así, el parámetro de banda espectral 132 puede comprender primeros parámetros de SBR 132a adaptados al primer algoritmo de parche ("patching"), segundos parámetros de 50 SBR 132b adaptados al segundo algoritmo de parche ("patching"), terceros parámetros de SBR 132c adaptados al tercer algoritmo de parche ("patching") y así, sucesivamente.

[0089] La figura 7b muestra una realización para el generador de información de control 340 de forma más detallada. El generador de información de control 340 recibe la señal filtrada de paso bajo 315 y los parámetros de 55 SBR 132. La señal filtrada de paso bajo 315 puede ser introducida en una primera unidad de parche ("patching") 342a, en una segunda unidad de parche ("patching") 342b, y en otras unidades de parche ("patching") (que no se muestran). La cantidad de unidades de parche ("patching") 342 puede estar de acuerdo, por ejemplo, con la cantidad de algoritmos de parche ("patching"), que pueden ser llevados a cabo por el generador de parche 110 en el decodificador. La salida de las unidades de parche ("patching") 342 comprende una primera señal de audio parcheada 344a para la primera 60 unidad de parche ("patching") 342a, una segunda señal de audio parcheada 344b para la segunda unidad de parche ("patching") 342b y así, sucesivamente. Las señales de audio parcheadas 344 que comprenden componentes no procesados en la segunda banda de frecuencia 202 se introducen en un bloque de herramientas de replicación de banda espectral 346. De nuevo, la cantidad de bloques de herramientas de replicación de banda espectral 346 puede ser igual, por ejemplo, que la cantidad de algoritmos de parche ("patching") o que la cantidad de unidades de parche 65 ("patching") 342. Los parámetros de replicación de banda espectral 132 se introducen también en los bloques de

herramientas de replicación de banda espectral 346 (bloque de herramientas de SBR) de forma que el primer bloque de herramientas de SBR 246a reciba los primeros parámetros de SBR 132a y la primera señal parcheada 344a. El segundo bloque de herramientas de SBR 346b recibe los segundos parámetros de SBR 132b y la segunda señal de audio parcheada 344b. Los bloques de herramientas de replicación de banda espectral 346 generan la señal de audio replicada 347 que comprende componentes de frecuencia más alta dentro de la segunda banda de frecuencia y/o la tercera banda de frecuencia 202 y 203 sobre la base de los parámetros de replicación 132.

[0090] Finalmente, el generador de información de control 340 comprende unidades de comparación adaptadas para comparar la señal de audio original 305 y, en especial, los componentes de frecuencia más alta de la señal de audio 305 con la señal de audio replicada 347. De nuevo, la comparación puede hacerse para cada algoritmo de parche ("patching") de forma que una primera unidad de comparación 348a compare la señal de audio 305 con una primera señal de audio replicada 347a emitida por el primer bloque de herramientas de SBR 346a. De forma similar, una segunda unidad de comparación 348b compara la señal de audio 305 con una segunda señal de audio replicada 347b desde el segundo bloque de herramientas de SBR 346b. Las unidades de comparación 348 determinan una desviación de las señales de audio replicadas 347 en las bandas de alta frecuencia con respecto a la señal de audio original 305 de forma que, finalmente, una unidad de evaluación 349 pueda comparar la desviación entre la señal de audio original 305 y las señales de audio replicadas 347 utilizando diferentes algoritmos de parche ("patching") y determine un algoritmo de parche ("patching") preferido o una cantidad de algoritmos de parche ("patching") adecuados o no adecuados a partir de ello. La información de control 112 comprende información que permite identificar uno de los algoritmos de parche ("patching") preferidos. La información de control 112 puede comprender, por ejemplo, un número de identificación para el algoritmo de parche ("patching") preferido, que puede determinarse sobre la base de la menor desviación entre la señal de audio original 305 y la señal de audio replicada 347. De forma alternativa, la información de control 112 puede proporcionar una cantidad de algoritmos de parche ("patching") o una jerarquización de algoritmos de parche ("patching"), lo que produce un acuerdo suficiente entre la señal de audio 305 y la señal de audio parcheada 347. La evaluación puede llevarse a cabo, por ejemplo, con respecto a la calidad perceptual de forma que la señal de audio replicada 347 pueda ser diferenciada o estar cerca de poder ser diferenciada de la señal de audio original 305 en una situación ideal para un ser humano.

[0091] La figura 7c muestra otra realización para el codificador en la que, de nuevo, se introduce la señal de audio 305, pero donde también se introducen metadatos 306, de forma opcional, en el codificador. La señal de audio original 305 se introduce de nuevo en un filtro de paso bajo 310 así como en un filtro de paso alto 320. La salida del filtro de paso bajo 310 se introduce de nuevo en un codificador de audio 330 y la salida del filtro de paso alto 320 se introduce en un generador de datos de SBR 328. El codificador comprende, además, una unidad de procesamiento de metadatos 309 y/o una unidad de análisis 307 (o medios de análisis), cuya salida es enviada al generador de información de control 340. La unidad de procesamiento de metadatos 309 está configurada para analizar los metadatos 306 con respecto a un algoritmo de parche ("patching") apropiado. La unidad de análisis 307 puede determinar, por ejemplo, la cantidad e intensidad de segmentos transitorios o de tren de impulsos o no de tren de impulsos dentro de la señal de audio 305. Sobre la base de la salida de la unidad de procesamiento de metadatos 309 y/o la salida de la herramienta de análisis 307, el generador de información de control 340 puede determinar, de nuevo, un algoritmo de interacción preferido o generar una jerarquización de algoritmos de parche ("patching") y codificar esta información dentro de la información de control 112. El formateador 343 volverá a combinar la información de control 112, el parámetro de replicación de banda espectral 132 así como la señal de audio codificada 355 dentro de un flujo de audio codificado 345.

[0092] Los medios de análisis 307 proporcionan, por ejemplo, la característica de la señal de audio y pueden estar adaptados para identificar componentes de señal no armónicos para una parte de tiempo con un grado de voz o un componente de señal armónica para una parte de tiempo diferenciada. Si la señal de audio 305 es puramente discurso o voz, el grado de voz es alto, mientras que para una mezcla de voz y, por ejemplo, música, el grado de voz es más bajo. El cálculo del parámetro de SBR 132 puede llevarse a cabo según esta característica y el algoritmo de parche ("patching") preferido.

[0093] Incluso otra realización comprende un procedimiento para un flujo de datos 345 que comprende componentes de una señal de audio 305 en una primera banda de frecuencia 201, información de control 112 y parámetros de replicación de banda espectral 132. El procedimiento comprende un filtrado selectivo con respecto a la frecuencia de la señal de audio 305 para generar los componentes de la señal de audio 305 en la primera banda de frecuencia 201. El procedimiento comprende, además, una generación del parámetro de replicación de banda espectral 132 a partir de los componentes de la señal de audio 305 en una segunda banda de frecuencia 202. Finalmente, el procedimiento comprende una generación de la información de control 112 que identifica un algoritmo de parche ("patching") preferido a partir de un primer algoritmo de parche ("patching") o un segundo algoritmo de parche ("patching") diferentes, en el que cada algoritmo de parche ("patching") genera una señal no procesada 115 con componentes de señal en la segunda banda de frecuencia replicada 202 utilizando los componentes de la señal de audio 305 en la primera banda de frecuencia 201.

[0094] Aunque algunas realizaciones, específicamente en las figuras 6a a 6d, han sido ilustradas de forma que la combinación entre una banda baja y una banda alta ajustada se lleve a cabo en el dominio frecuencial, cabe destacar

que la combinación puede implementarse, además, en el dominio del tiempo. Con este fin, la señal de salida del decodificador de núcleo puede utilizarse (en la salida de una etapa de retardo potencialmente necesaria para compensar un retardo de procesamiento incurrido por parche ("patching") y ajuste) en el dominio del tiempo, y la banda alta ajustada en el dominio de banco de filtros puede ser convertida en el dominio del tiempo como una señal sin la porción de banda baja y con la porción de banda alta. En la realización de la figura 6, esta señal solamente comprenderá las 32 sub-bandas más altas, y una conversión de esta señal en el dominio del tiempo tiene como resultado una señal de banda alta de dominio del tiempo. A continuación, ambas señales pueden combinarse en el dominio del tiempo tal como por medio de una suma muestra por muestra con el fin de obtener, por ejemplo, muestras de PCM como una señal de salida para ser convertidas de digital a analógica y aplicadas a un altavoz.

10

[0095] Aunque algunos aspectos se han descrito dentro del contexto de un aparato, es evidente que estos aspectos representan, además, una descripción del procedimiento correspondiente, donde un bloque o dispositivo corresponde a una etapa del procedimiento o a una característica de una etapa del procedimiento. De forma análoga, los aspectos descritos dentro del contexto de una etapa del procedimiento representan, además, una descripción de un bloque o elemento o característica correspondiente de un aparato correspondiente.

15

[0096] La señal de audio o flujo de bits codificados de la invención pueden ser almacenados en un medio de almacenamiento digital o pueden transmitirse en un medio de transmisión tal como un medio de transmisión inalámbrica o un medio de transmisión alámbrica tal como Internet.

20

[0097] Dependiendo de ciertos requisitos de implementación, las realizaciones de la invención pueden implementarse en *hardware* o en *software*. La implementación puede llevarse a cabo utilizando un medio de almacenamiento digital, por ejemplo, un disquete, un DVD, un CD, una ROM, una PROM, una EPROM, una EEPROM o una memoria FLASH, con señales de control que pueden leerse de forma electrónica almacenadas en ellos, que cooperan (o son capaces de cooperar) con un sistema informático programable de forma que se lleve a cabo el procedimiento respectivo.

25

[0098] Algunas realizaciones según la invención comprenden un soporte de datos con señales de control que pueden leerse de forma electrónica, que son capaces de cooperar con un sistema informático programable, de forma que se lleve a cabo uno de los procedimientos descritos en esta invención. En general, las realizaciones de la presente invención pueden ser implementadas como un producto de programa informático con un código de programa, código de programa que es operativo para llevar a cabo uno de los procedimientos cuando el producto de programa informático se ejecuta en un ordenador. El código de programa puede almacenarse, por ejemplo, en un soporte que pueda leerse por medio de una máquina. Otras realizaciones comprenden el programa informático para llevar a cabo uno de los procedimientos descritos en esta invención, almacenado en un soporte que pueda leerse por medio de una máquina. Por lo tanto, en otras palabras, una realización del procedimiento de la invención es un programa informático con un código de programa para llevar a cabo uno de los procedimientos descritos en esta invención, cuando el programa informático se ejecuta en un ordenador. Por lo tanto, otra realización de los procedimientos de la invención es un soporte de datos (o un medio de almacenamiento digital, o un medio que puede leerse por medio de un ordenador) que comprende, grabado en él, el programa informático para llevar a cabo uno de los procedimientos descritos en esta invención. Por lo tanto, otra realización del procedimiento de la invención es un flujo de datos o una secuencia de señales que representan el programa informático para llevar a cabo uno de los procedimientos descritos en esta invención. El flujo de datos o la secuencia de señales puede estar configurado, por ejemplo, para ser transferido a través de una conexión de comunicación de datos, por ejemplo, por Internet. Otra realización 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 esta invención. Otra realización comprende un ordenador con el programa informático para llevar a cabo uno de los procedimientos descritos en esta invención instalado en él. En algunas realizaciones, puede utilizarse un dispositivo lógico programable (por ejemplo, una matriz de puertas programable) para llevar a cabo algunas o todas las funcionalidades de los procedimientos descritos en esta invención. En algunas realizaciones, una matriz de puertas programable puede cooperar con un microprocesador para llevar a cabo uno de los procedimientos descritos en esta invención. En general, los procedimientos son llevados a cabo, preferentemente, por cualquier aparato de *hardware*.

30

35

40

45

50

[0099] Las realizaciones descritas anteriormente son meramente ilustrativas de los fundamentos de la presente invención. Se comprende que las modificaciones y variantes de las disposiciones y los detalles descritos en esta invención serán evidentes para los expertos en la materia. Por lo tanto, están destinadas a estar limitadas solamente por el alcance de las reivindicaciones de la patente inminente y no por los detalles específicos presentados a modo de descripción y la explicación de las realizaciones en esta invención.

55

REIVINDICACIONES

1. Sintetizador de señales de audio para generar una señal de audio de síntesis (145) que tiene una primera banda de frecuencia (201) y una segunda banda de frecuencia sintetizada (202) derivada de la primera banda de frecuencia (201), que comprende:
- un generador de parche (110) para llevar a cabo al menos dos algoritmos de parche ("patching") diferentes, en el que cada algoritmo de parche ("patching") genera una señal no procesada (115) que tiene componentes de señal en la segunda banda de frecuencia sintetizada (202) mediante el uso de una señal de audio (105) que tiene componentes de señal en la primera banda de frecuencia (201), y en el que el generador de parche (110) está adaptado para seleccionar uno de los al menos dos algoritmos de parche ("patching") diferentes en respuesta a una información de control (112) para una primera parte de tiempo y otro de los al menos dos algoritmos de parche ("patching") diferentes en respuesta a la información de control (112) para una segunda parte de tiempo diferente de la primera parte de tiempo con el fin de obtener la señal no procesada (115) para la primera y la segunda parte de tiempo, en el que el generador de parche (110) está adaptado para operar en un dominio de tiempo para al menos dos algoritmos de parche ("patching") diferentes;
- un conversor espectral (120) para convertir la señal no procesada (115) para la primera y la segunda parte de tiempo desde el dominio del tiempo en el dominio espectral para obtener una representación espectral de señal no procesada (125) para la primera y la segunda parte de tiempo;
- un procesador de señales no procesadas (130) para procesar la representación espectral de señal no procesada (125) para la primera y la segunda parte de tiempo en respuesta a parámetros de replicación de banda espectral de dominio espectral (132) con el fin de obtener una representación espectral de señal no procesada ajustada (135) para la primera y la segunda parte de tiempo; y
- un combinador (140) para combinar la señal de audio (105) que tiene componentes de señal en la primera banda de frecuencia (201) o una señal derivada de la señal de audio (105) con la representación espectral de señal no procesada ajustada (135) o con otra señal derivada de la representación espectral de señal no procesada ajustada (135) con el fin de obtener la señal de audio de síntesis (145).
2. Sintetizador de señales de audio de la reivindicación 1, en el que los al menos dos algoritmos de parche ("patching") se diferencian entre sí en tanto que un componente de señal de la señal de audio (105) en una frecuencia en la primera banda de frecuencia (201) está parcheado a una frecuencia objetivo en la segunda banda de frecuencia (202), y la frecuencia objetivo es diferente para ambos algoritmos de parche ("patching").
3. Sintetizador de señales de audio de la reivindicación 1 o 2, en el que el generador de parche (110) está adaptado para generar la señal no procesada (115) de tal manera que la señal no procesada (115) comprenda componentes de señal adicionales en la primera banda de frecuencia (201) que tengan una tasa de muestreo, que es mayor que una tasa de muestreo de la señal de audio (105) introducida en el generador de parche (110), y en el que el convertidor espectral (120) está adaptado para convertir los componentes de señal en la segunda banda de frecuencia sintetizada (202) y otros componentes de señal en la primera banda de frecuencia (201) en la representación espectral de señal no procesada (125).
4. Sintetizador de señales de audio de la reivindicación 3, que comprende además un conversor espectral adicional (370) y un combinador adicional (124), el conversor espectral adicional (370) está adaptado para convertir la señal de audio (105) que tiene componentes de señal en la primera banda de frecuencia (201) en una representación espectral de señal no procesada adicional (123), y el combinador adicional (124) está adaptado para combinar la representación espectral de señal no procesada (125) y la representación espectral de señal no procesada adicional (123) con el fin de obtener una representación espectral de señal no procesada combinada (126) y en el que el procesador de señales no procesadas (130) está adaptado para procesar la representación espectral de señal no procesada combinada (126).
5. Sintetizador de señales de audio de cualquiera de las reivindicaciones anteriores, en el que el combinador (140) está adaptado para utilizar como señal derivada de la señal de audio (105) la representación espectral de señal no procesada adicional (123).
6. Sintetizador de señales de audio de cualquiera de las reivindicaciones anteriores, en el que la señal de audio (105) y la información de control (112) están codificadas en un flujo de datos, en el que el sintetizador de señales de audio comprende además un desformateador, y en el que el desformateador está configurado para obtener la información de control (112) desde el flujo de datos.
7. Sintetizador de señales de audio de cualquiera de las reivindicaciones anteriores, en el que la señal de audio y los parámetros de replicación de banda espectral (132) están codificados en un flujo de datos, y en el que el procesador de señales no procesadas (130) está adaptado para obtener los parámetros de replicación de banda espectral (132) desde el flujo de datos.
8. Codificador de señales de audio para generar, a partir de una señal de audio (305), un flujo de datos

(345) que comprende componentes de la señal de audio (305) en una primera banda de frecuencia (201), información de control (112) y parámetros de replicación de banda espectral (132), comprendiendo el codificador de señales de audio:

5 un filtro selectivo con respecto a la frecuencia (310) para generar los componentes de la señal de audio (305) en la primera banda de frecuencia (201);
 un generador para generar el parámetro de replicación de banda espectral (132) a partir de los componentes de la señal de audio (305) en una segunda banda de frecuencia (202);
 un generador de información de control (340) para generar la información de control (112), identificando la
 10 información de control (112) un algoritmo de parche ("patching") preferido a partir de un primer algoritmo de parche ("patching") o un segundo algoritmo de parche ("patching") diferentes, en el que cada algoritmo de parche ("patching") genera una señal no procesada (115) que tiene componentes de señal en la segunda banda de frecuencia replicada (202) mediante el uso de los componentes de la señal de audio (305) en la primera banda de frecuencia (201), en el que el primer algoritmo de parche ("patching") y el segundo algoritmo de parche ("patching")
 15 diferente están adaptados para operar en un dominio de tiempo, en el que el generador de información de control (340) está adaptado para identificar el algoritmo de parche ("patching") preferido comparando la señal de audio (305) con las señales de audio parcheadas (347) para el primer y el segundo algoritmo de parche ("patching"), en el que las señales de audio parcheadas de manera diferente (347) se derivan de diferentes señales no procesadas (344) relacionadas con el primer y el segundo
 20 algoritmo de parche ("patching") aplicando el ajuste de la señal no procesada en respuesta a los parámetros de replicación de banda espectral (132) con una herramienta de replicación de banda espectral (346).

9. Codificador de señales de audio de la reivindicación 8, que comprende, además, un medio de análisis (307) de la señal de audio (305) con el fin de proporcionar una característica de la señal de audio, el medio de análisis
 25 (307) está adaptado para identificar componentes de señal no armónicos para una parte de tiempo que tiene un grado de voz o un componente de señal armónico para una parte de tiempo diferenciada.

10. Procedimiento para generar una señal de audio de síntesis (145) que tiene una primera banda de frecuencia y una segunda banda de frecuencia replicada (202) derivada de la primera banda de frecuencia (201), que
 30 comprende:

llevar a cabo al menos dos algoritmos de parche ("patching") diferentes, en el que cada algoritmo de parche ("patching") genera una señal no procesada (115) que tiene componentes de señal en la segunda banda de frecuencia replicada (202) mediante el uso de una señal de audio (105) que tiene componentes de señal en la
 35 primera banda de frecuencia (201), y en el que la realización de al menos dos algoritmos de parche ("patching") diferentes se lleva a cabo de tal forma que uno de los al menos dos algoritmos de parche ("patching") diferentes sea seleccionado en respuesta a una información de control (112) para una primera parte de tiempo y el otro de los al menos dos algoritmos de parche ("patching") diferentes sea seleccionado en respuesta a la información de control (112) para una segunda parte de tiempo diferente de la primera parte de tiempo con el fin de obtener la
 40 señal no procesada (115) para la primera y la segunda parte de tiempo, en el que la realización de los al menos dos algoritmos de parche ("patching") diferentes opera en un dominio de tiempo para los al menos dos algoritmos de parche ("patching") diferentes;
 convertir la señal no procesada (115) para la primera y la segunda parte de tiempo del dominio de tiempo en el dominio espectral con el fin de obtener una representación espectral de señal no procesada (125) para la primera
 45 y la segunda parte de tiempo;
 procesar la representación espectral de señal no procesada (125) para la primera y la segunda parte de tiempo en respuesta a parámetros de replicación de banda espectral del dominio espectral (132) con el fin de obtener una representación espectral de señal no procesada ajustada (135) para la primera y la segunda parte de tiempo; y
 50 combinar la señal de audio (105) que tiene componentes de señal en la primera banda de frecuencia (201) o una señal derivada de la señal de audio (105) con la representación espectral de señal no procesada ajustada (135) o con otra señal derivada de la representación espectral de señal no procesada ajustada (135) con el fin de obtener la señal de audio de síntesis (145).

11. Un procedimiento para generar un flujo de datos (345) que comprende componentes de una señal de
 55 audio (305) en una primera banda de frecuencia (201), información de control (112) y parámetros de replicación de banda espectral (132), comprendiendo el procedimiento:

filtrado selectivo con respecto a la frecuencia de la señal de audio (305) para generar los componentes de la señal de audio (305) en la primera banda de frecuencia (201);
 60 generar los parámetros de replicación de banda espectral (132) a partir de los componentes de la señal de audio (305) en una segunda banda de frecuencia (202);
 generar la información de control (112) que identifica un algoritmo de parche ("patching") preferido a partir de un primer algoritmo de parche ("patching") o un segundo algoritmo de parche ("patching") diferente, en el que cada algoritmo de parche ("patching") genera una señal no procesada (115) que tiene componentes de señal en la
 65 segunda banda de frecuencia (202) utilizando los componentes de la señal de audio (305) en la primera banda de

- frecuencia (201), en el que el primer algoritmo de parche ("patching") y el segundo algoritmo de parche ("patching") diferente están adaptados para operar en un dominio de tiempo, en el que el algoritmo de parche ("patching") preferido se identifica comparando la señal de audio (305) con las señales de audio parcheadas (347) para el primer y el segundo algoritmo de parche ("patching"), en el que las
- 5 señales de audio parcheadas diferentes (347) se derivan de diferentes señales no procesadas (344) relacionadas con el primer y el segundo algoritmo de parche ("patching") mediante la aplicación del ajuste de la señal no procesada en respuesta a los parámetros de replicación de banda espectral (132) con una herramienta de replicación de banda espectral (346).
- 10 12. Programa informático para llevar a cabo, cuando se ejecuta en un procesador, un procedimiento según la reivindicación 10 o la reivindicación 11.

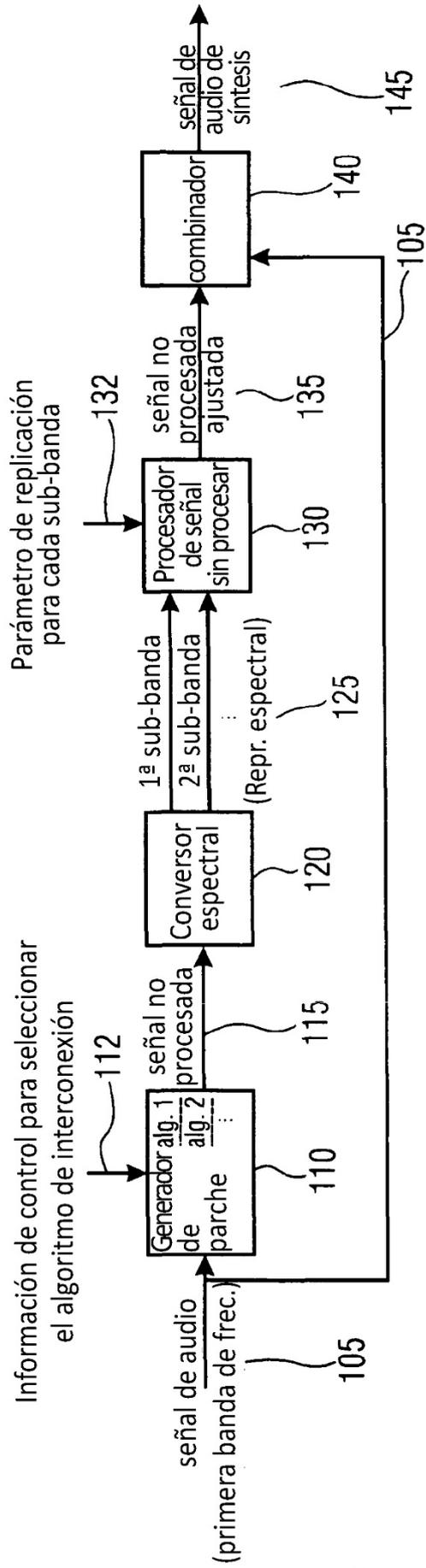


FIG 1

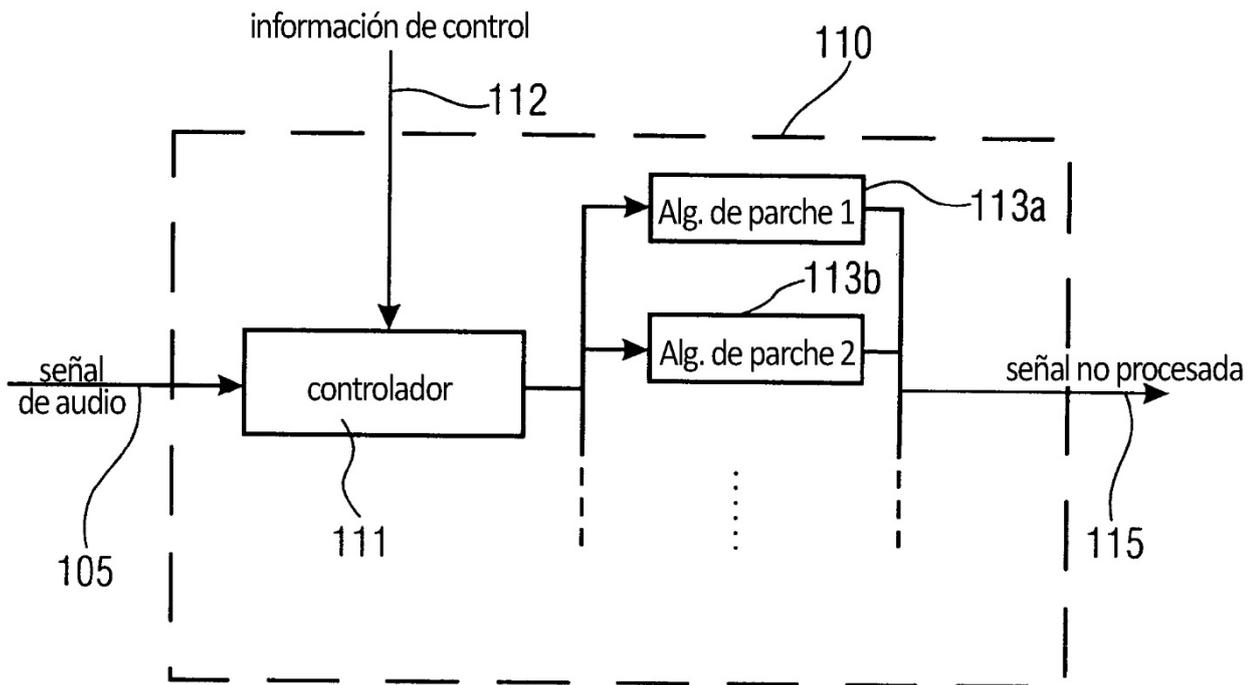


FIG 2

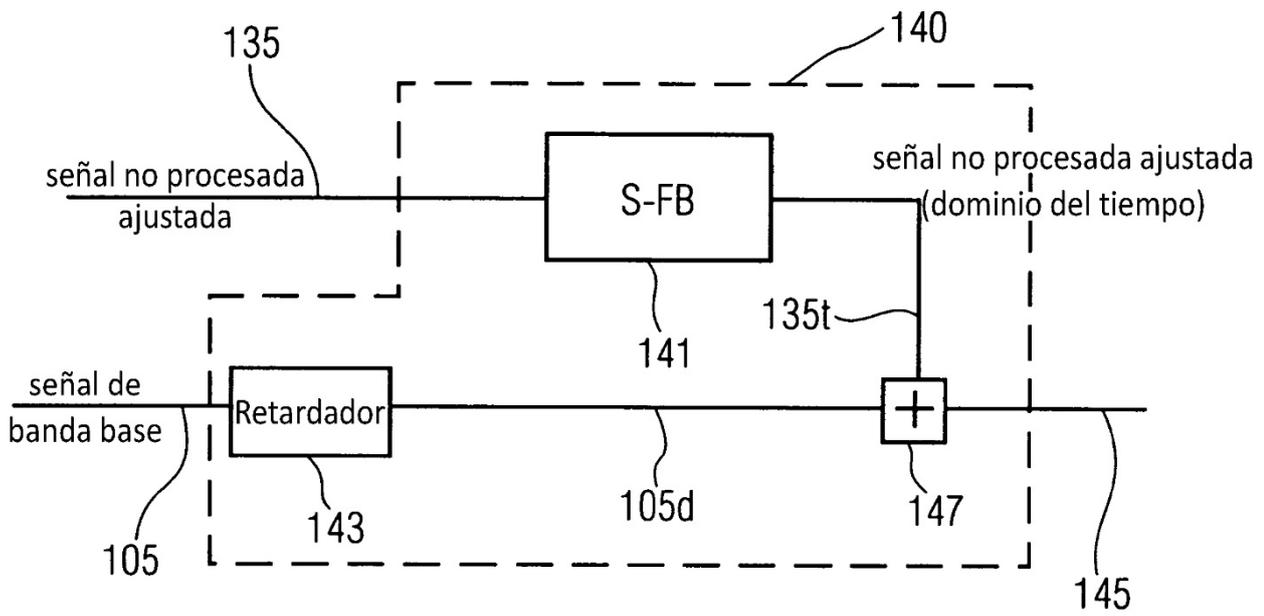


FIG 3

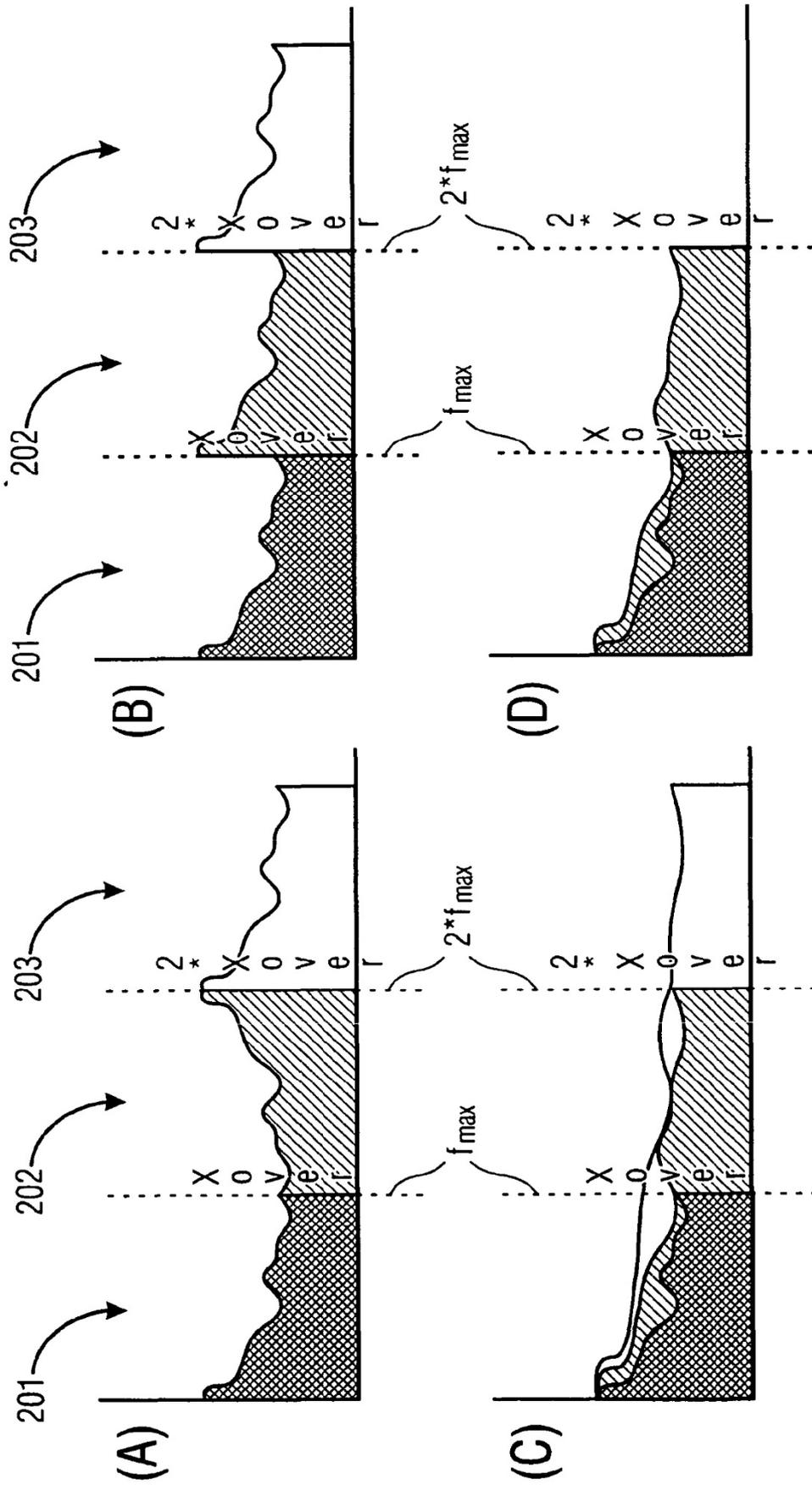


FIG 4

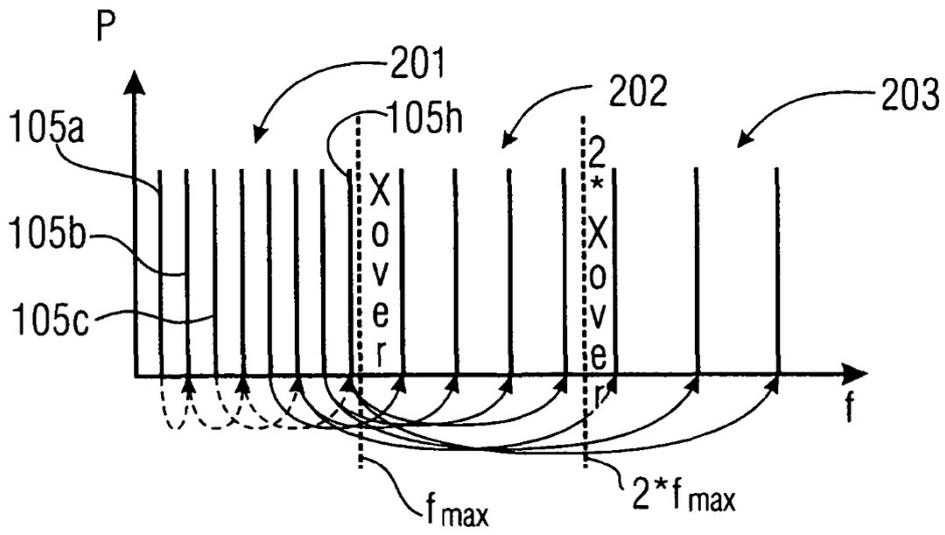


FIG 5A

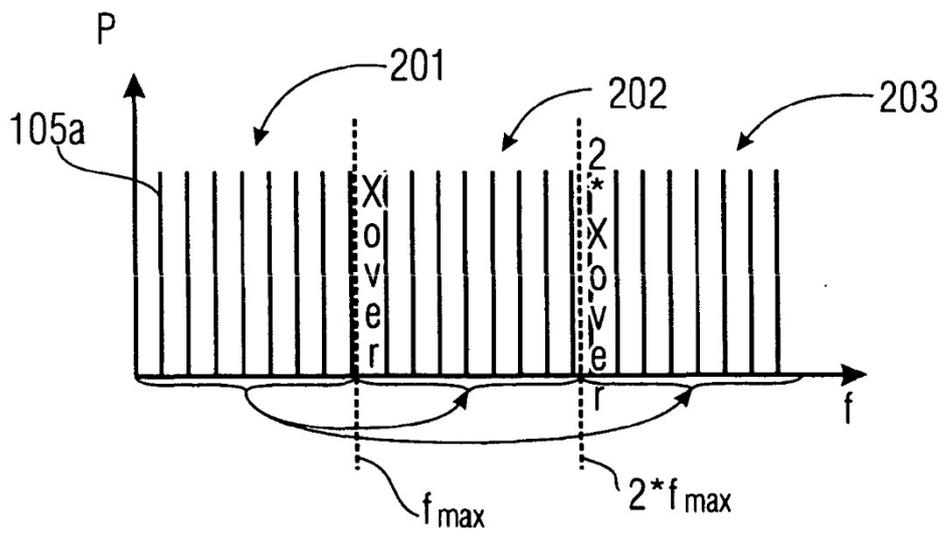


FIG 5B

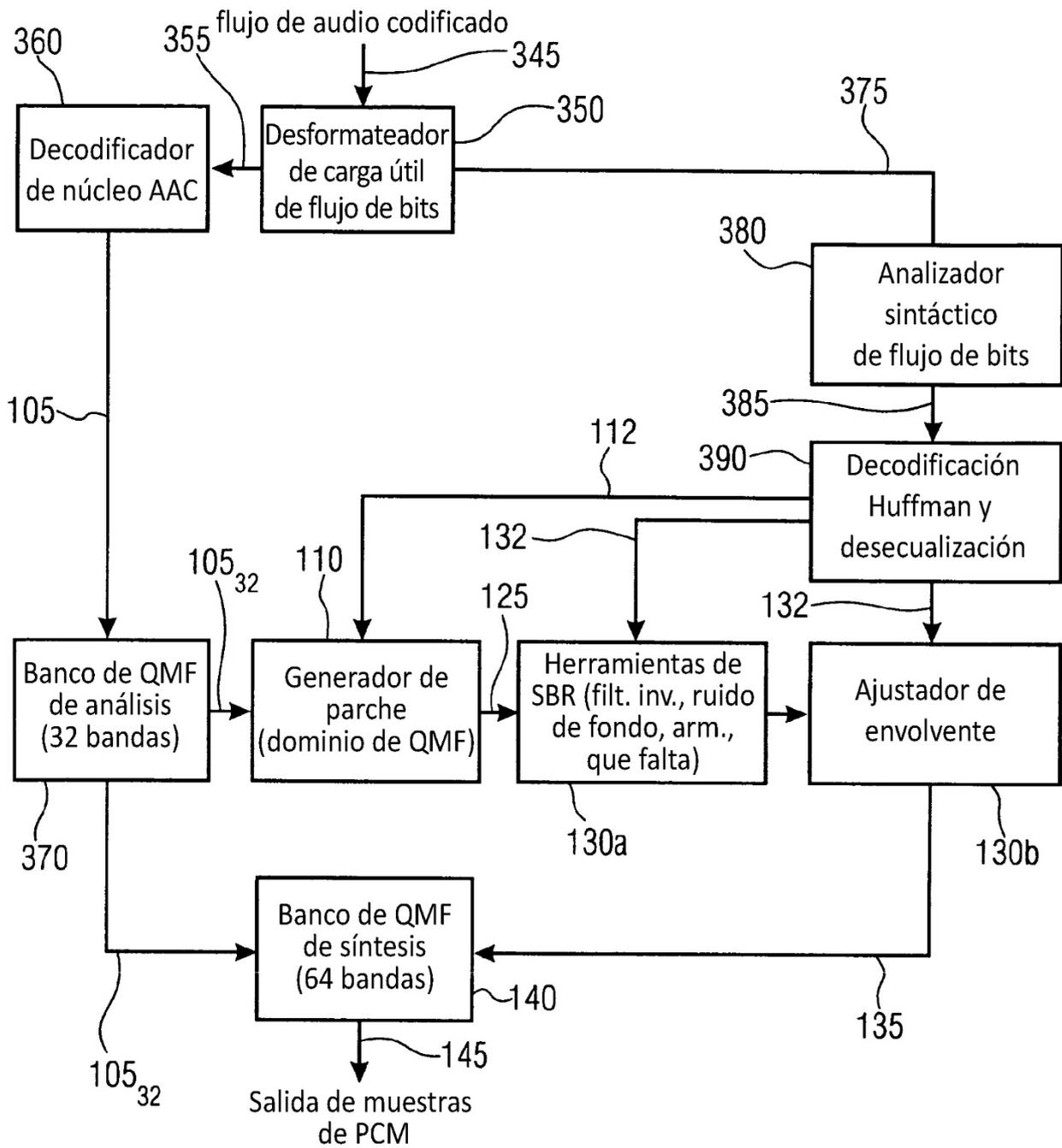


FIG 6A

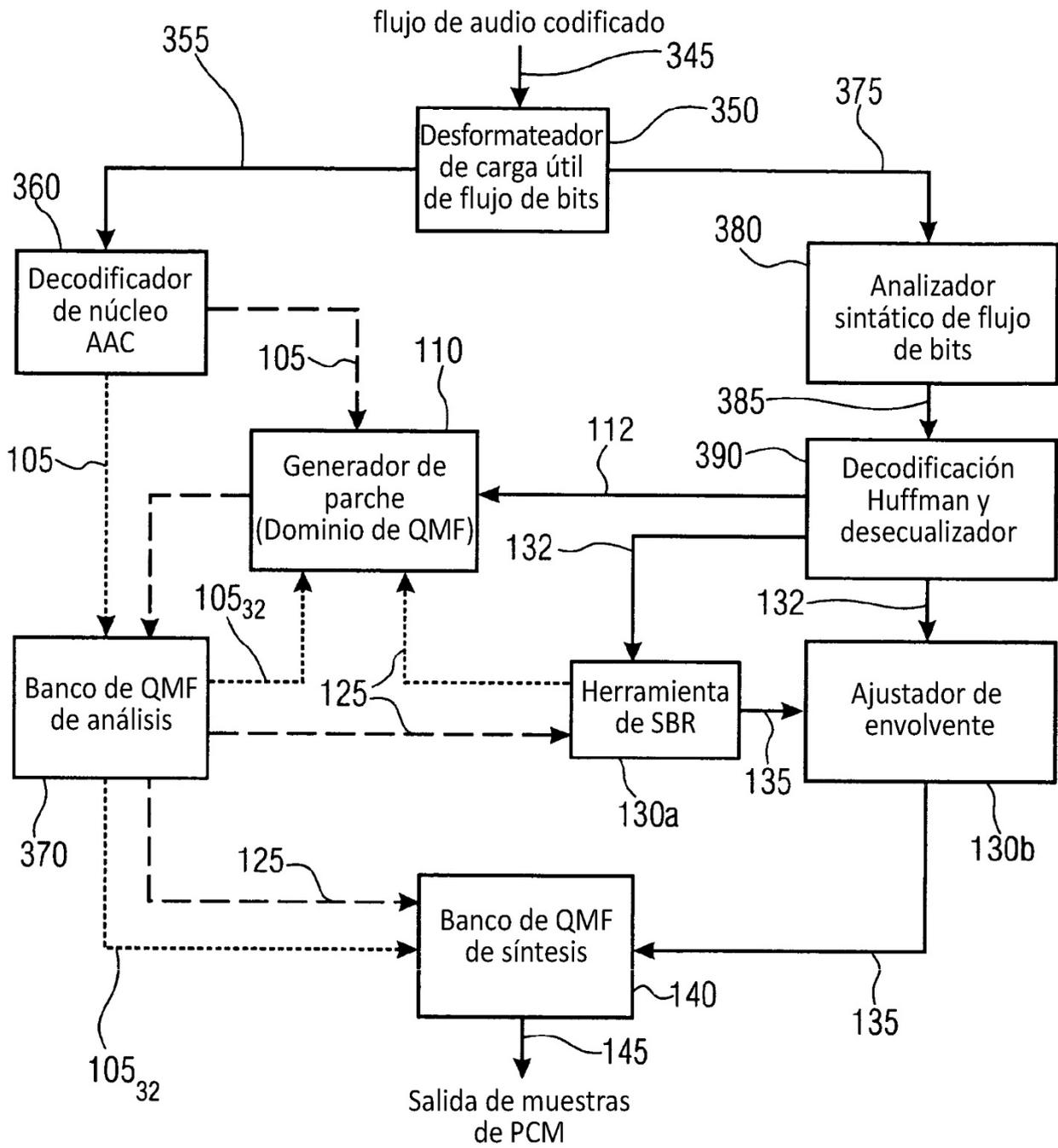


FIG 6B

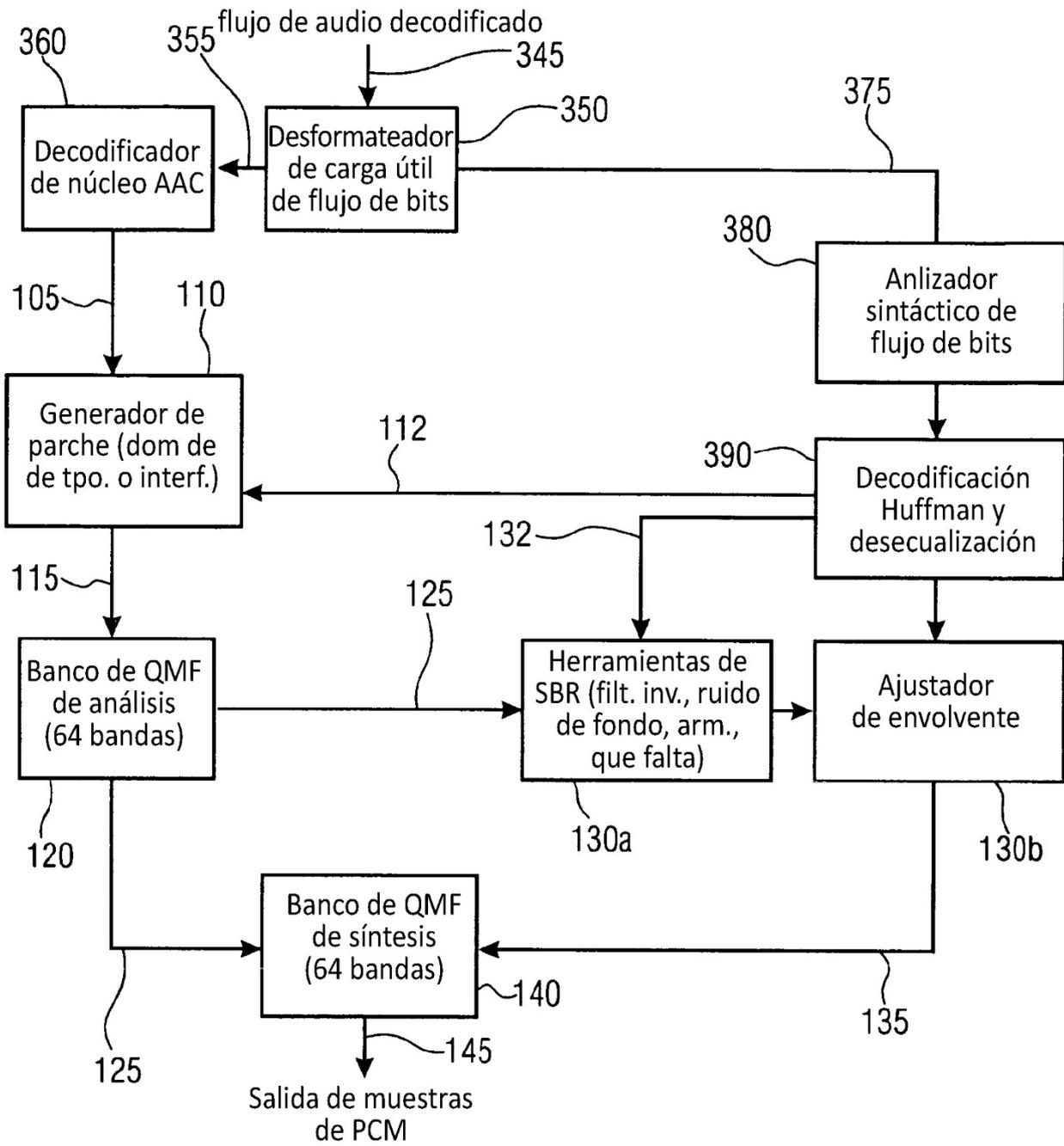


FIG 6C

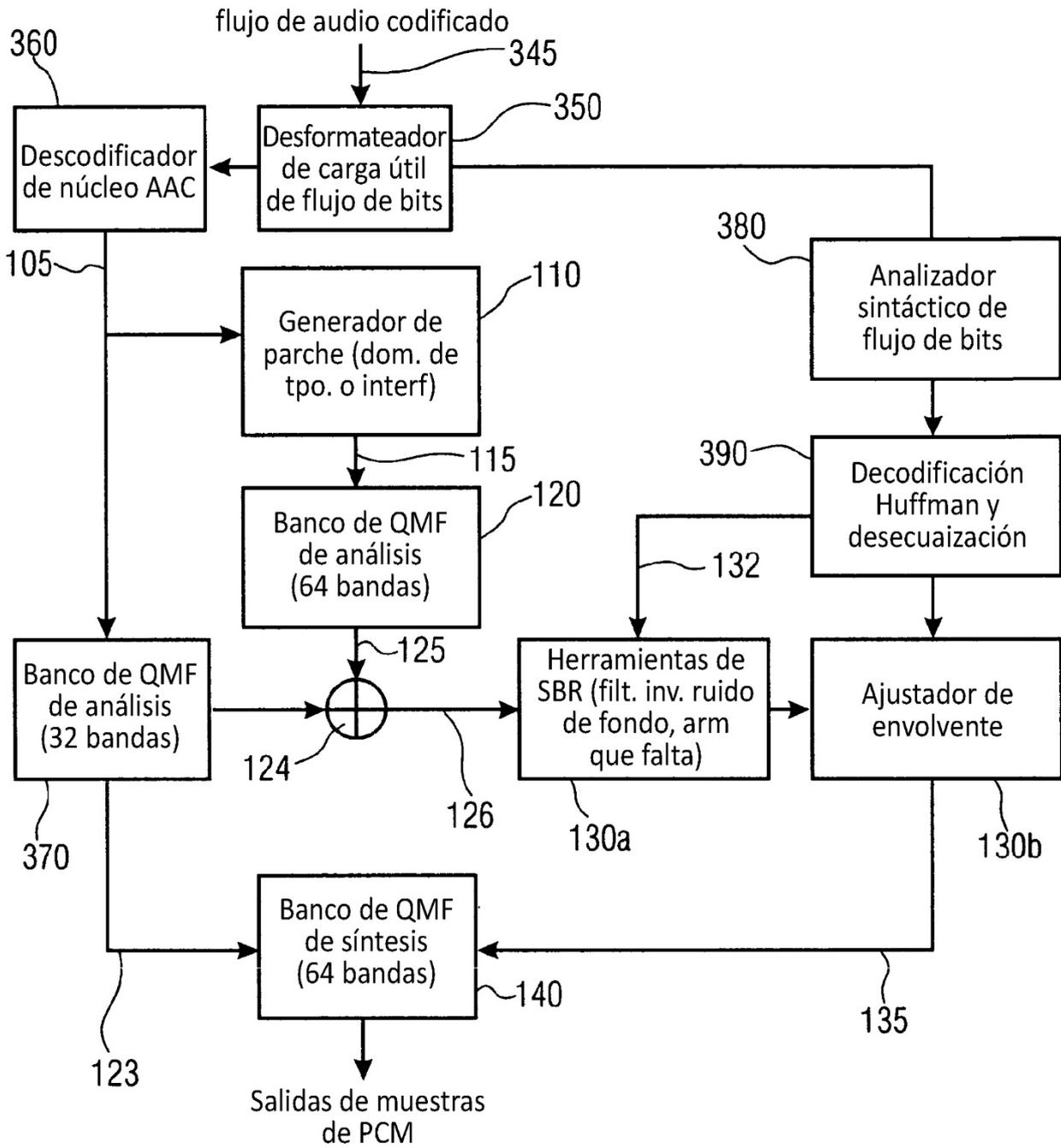


FIG 6D

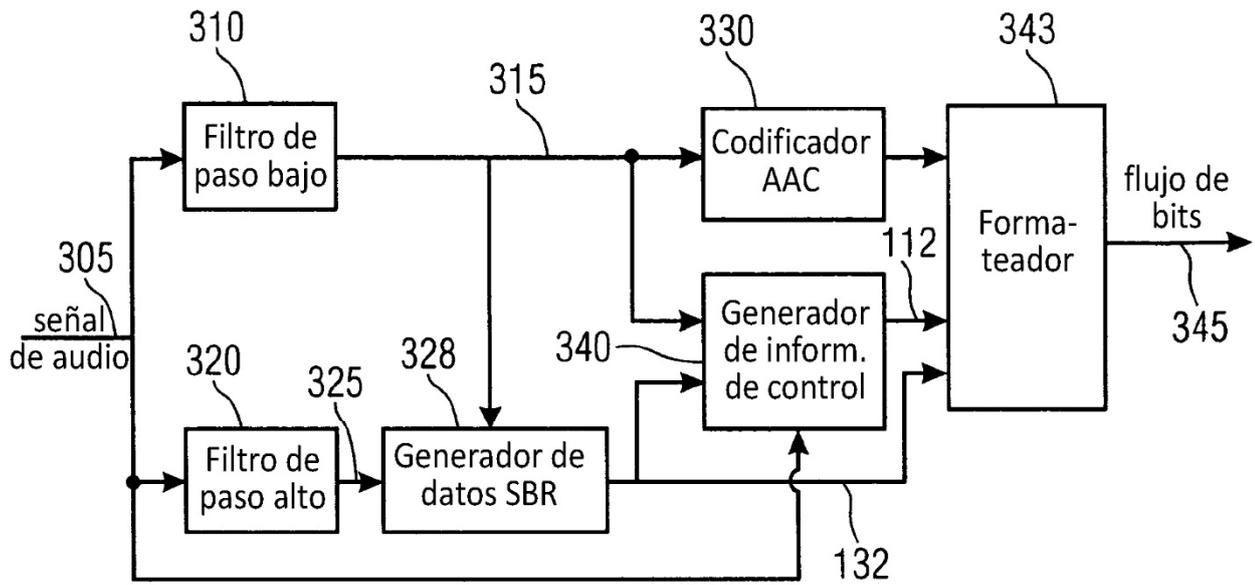


FIG 7A

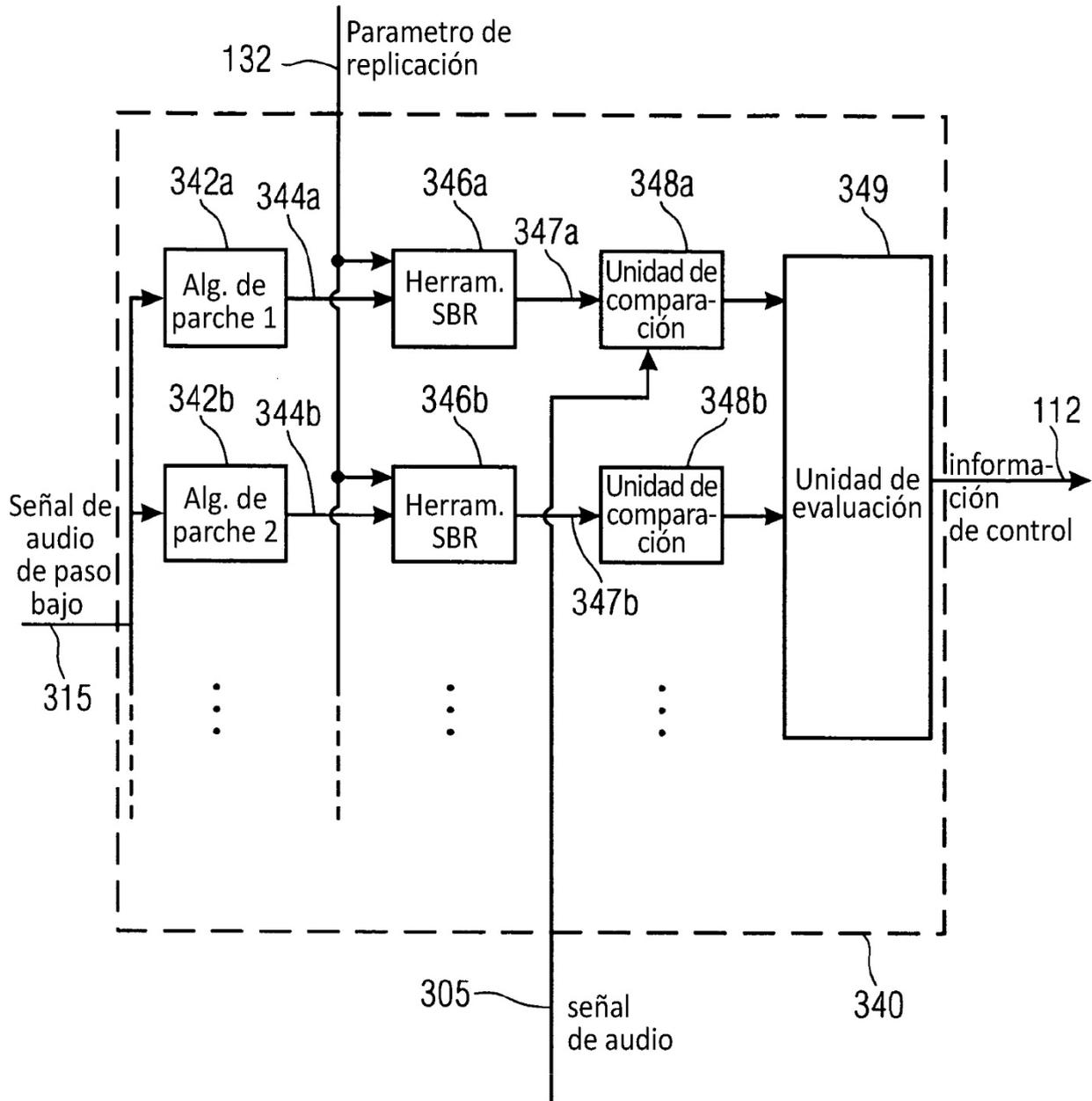


FIG 7B

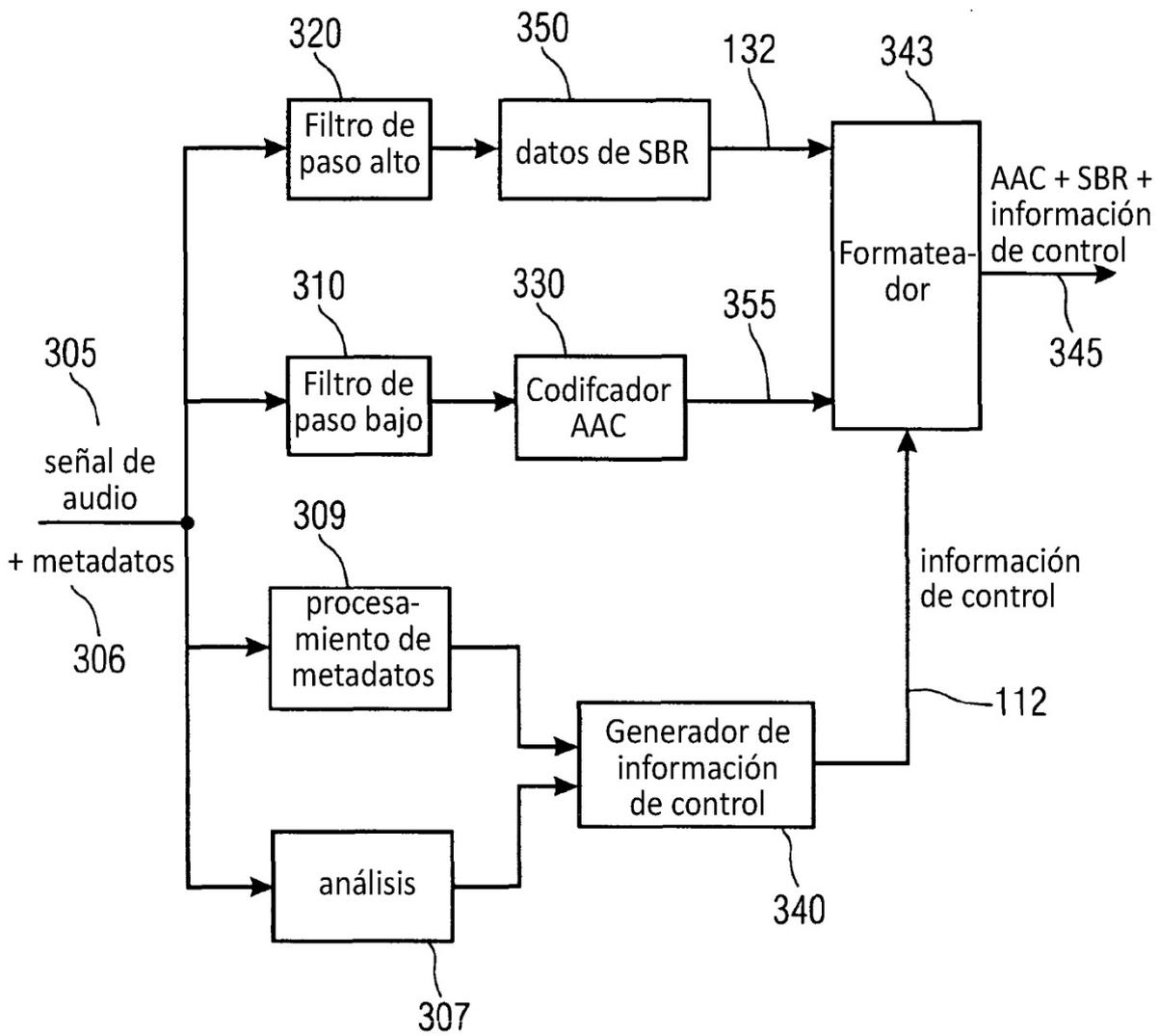


FIG 7C