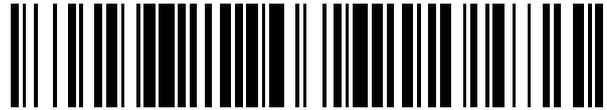


19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 396 686**

51 Int. Cl.:

G10L 21/02 (2006.01)

G10L 19/14 (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

96 Fecha de presentación y número de la solicitud europea: **01.04.2010 E 10712944 (7)**

97 Fecha y número de publicación de la concesión europea: **14.11.2012 EP 2351025**

54 Título: **Aparato y procedimiento para generar una señal de audio de síntesis y para codificar una señal de audio**

30 Prioridad:

09.04.2009 US 168068 P
30.12.2009 EP 09181008

45 Fecha de publicación y mención en BOPI de la traducción de la patente:
25.02.2013

73 Titular/es:

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

72 Inventor/es:

NAGEL, FREDERIK;
MULTRUS, MARKUS;
LECOMTE, JÉRÉMIE;
BAYER, STEFAN;
FUCHS, GUILLAUME;
HILPERT, JOHANNES y
ROBILLIARD, JULIEN

74 Agente/Representante:

PONTI SALES, Adelaida

ES 2 396 686 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

Aparato y procedimiento para generar una señal de audio de síntesis y para codificar una señal de audio

[0001] La presente invención se relaciona con el procesamiento de señales de audio y, en particular, con un aparato y un procedimiento para generar una señal de audio de síntesis, un aparato y un procedimiento para codificar una señal de audio y una señal de audio codificada.

[0002] El almacenamiento o transmisión de señales de audio con frecuencia es objeto de rigurosas restricciones de velocidad de transmisión de bits. Estas restricciones con frecuencia se superan mediante la codificación intermedia de la señal. En el pasado, se forzaba a los codificadores a reducir drásticamente el ancho de banda del audio transmitido cuando sólo se disponía de una velocidad de transmisión de bits muy baja. Los códecs (codificador-decodificador) de audio modernos pueden codificar las señales de banda ancha utilizando procedimientos de extensión del ancho de banda (BWE), de acuerdo con lo descrito por M Dietz, L. Liljeryd, K. Kjörling y O. Kunz, en "Spectral Band Replication, a novel approach in audio coding" en la 112a Convención de AES, Munich, Mayo de 2002; S. Meltzer, R. Böhm y F. Henn, "SBR enhanced audio codecs for digital broadcasting such as "Digital Radio Mondiale" (DRM)," en la 112a Convención de AES, Munich, Mayo de 2002; T. Ziegler, A. Ehret, P. Ekstrand y M. Lutzky, "Enhancing mp3 with SBR: Features and Capabilities of the new mp3PRO Algorithm," en la 112a Convención de AES, Munich, Mayo de 2002; International Standard ISO/IEC 14496-3:2001/FPDAM 1, "Bandwidth Extension," ISO/IEC, 2002. Speech bandwidth extension method and apparatus Vasu lyengar et al. Patente de Estados Unidos 5.455.888; E. Larsen, R. M. Aarts, y M. Danessis. Efficient high-frequency bandwidth extension of music and speech. en la 112a Convención de AES, Munich, Alemania, Mayo de 2002; R.M. Aarts, E. Larsen, y O. Ouweltjes. A unified approach to low-and high frequency bandwidth extension. En la 115a Convención de AES, New York, Estados Unidos, Octubre de 2003; K. Käyhkö. A Robust Wideband Enhancement for Narrowband Speech Signal. Research Report, Helsinki University of Technology, Laboratory of Acoustics and Audio Signal Processing, 2001; E. Larsen and R.M. Aarts. Audio Bandwidth Extension – Application to psychoacoustics, Signal Processing and Loudspeaker Design. John Wiley & Sons, Ltd, 2004; E. Larsen, R.M. Aarts y M. Danessis. Efficient high-frequency bandwidth extension of music and speech. En la 112a Convención de AES, Munich, Alemania, Mayo de 2002; J. Makhoul. Spectral Analysis of Speech by Linear Prediction. IEEE Transactions of Audio and Electroacoustics, AU-21(3), Junio de 1973; Solicitud de patente de los Estados Unidos 08/951,029, Ohmori, et al. Audio band width extending system and method; Patente de Estados Unidos 6895375, Malah, D & Cox, R.V.: System for bandwidth extension of Narrow-band speech y Frederik Nagel, Sascha Disch, "A harmonic bandwidth extension method for audio codecs," ICASSP International Conference on Acoustics, Speech and Signal Processing, IEEE CNF, Taipei, Taiwan, Abril de 2009.

[0003] Estos algoritmos se basan en una representación paramétrica del contenido de alta frecuencia (HF). Esta representación se genera desde la parte de baja frecuencia (LF) de la señal decodificada por medio de la transposición a la región espectral de HF ("patching") y la aplicación de un post procesamiento impulsado por los parámetros.

[0004] En la técnica, se utilizan procedimientos de extensión del ancho de banda tales como replicación de la banda espectral (SBR) como procedimiento eficiente para generar señales de alta frecuencia en un códec basado en HFR (reconstrucción de alta frecuencia).

[0005] La replicación de la banda espectral (SBR), según lo descrito por M Dietz, L. Liljeryd, K. Kjörling y O. Kunz, "Spectral Band Replication, a novel approach in audio coding" en la 112a Convención de AES, Munich, mayo de 2002, utiliza un banco de filtros espejo en cuadratura (QMF) para generar la información de HF (alta frecuencia). Con el denominado "patching", las señales de banda QMF más altas se copian en bandas QMF más altas, lo que lleva a la replicación de la información de la parte de LF en la parte de HF. La parte de HF generada se adapta posteriormente a la parte original de HF con la ayuda de parámetros que ajustan la envolvente espectral y la tonalidad.

[0006] En la SBR (replicación de la banda espectral), normalizada en HE-AAC, todas las operaciones, que incluyen la introducción de parches por medio de simple copiado, se llevan a cabo frecuentemente dentro del dominio de QMF. Sin embargo, se pueden ejecutar otros procedimientos diferentes de introducción de parches en diferentes dominios, tales como el dominio FFT o el dominio del tiempo. Uno se podría imaginar la habilitación de la SBR para elegir alternativamente un algoritmo de parches que funcione en el dominio FFT o en el dominio del tiempo y necesite una transformación adicional para alimentar el paso de análisis de QMF.

[0007] En la SBR básica, sólo se dispone de un algoritmo de emparche que no toma en cuenta ni las necesidades de hardware ni software especial ni las características de la señal. Por ende, la SBR no puede adaptar el algoritmo de parches. Uno se podría imaginar simplemente la elección entre dos algoritmos diferenciados de introducción de parches ("patching"). Dado que los dos procedimientos de introducción de parches ("patching") funcionan en diferentes dominios, las áreas de transición son proclives a producir artefactos de bloqueo, lo que prácticamente imposibilita la conmutación de grano fino entre ambos procedimientos.

[0008] WO 98/57436 describe procedimientos de transposición utilizados en la replicación de la banda espectral, que se combinan con el ajuste de la envoltura espectral.

[0009] WO 02/052545 plantea que las señales se pueden clasificar en forma de tren de pulsos o en forma no de tren de pulsos y basándose en esta clasificación se propone un transpositor. El transpositor de conmutación ejecuta dos algoritmos de introducción de parche en paralelo y la unidad mezcladora combina ambas señales emparchadas dependiendo de la clasificación (tren de pulsos o no tren de pulsos). La conmutación real entre los transpositores o su mezcla se ejecuta en un banco de filtros de ajuste de la envoltura en respuesta a los datos de envoltura y control. Más aun, en el caso de las señales en forma de tren de pulsos, la señal base se transforma en un dominio de banco de filtros, se ejecuta una operación de traslado de frecuencias y se ejecuta un ajuste a la envoltura del resultado de la traslación de frecuencias. Este es un procedimiento combinado de introducción de parche/procesamiento adicional. En el caso de las señales que no tienen forma de tren de pulsos, se incluye un transpositor de dominio de la frecuencia (transpositor FD) y el resultado del transpositor al dominio de la frecuencia se transforma entonces en el dominio de bancos de filtro, en el cual se ejecuta el ajuste de la envoltura. Por consiguiente, la implementación y flexibilidad de este procedimiento, que, en una alternativa, presenta una estrategia de parche/procesamiento adicional y que, en la otra alternativa, incluye un transpositor del dominio de la frecuencia que está situado fuera del banco de filtros en el cual tiene lugar el ajuste de la envoltura, son problemáticas con respecto a las posibilidades de flexibilidad e implementación.

[0010] En el estado de la técnica, los siguientes documentos describen soluciones para realizar un patching de componentes espectrales que faltan: WO 02/052545 A1, US 2002/016698 A1 y WO 03/107329 A1.

[0011] Es un objetivo de la presente invención dar a conocer un concepto para generar una señal de audio de síntesis que dé lugar a una calidad mejorada y permita una implementación eficiente.

[0012] Este objetivo se logra mediante un aparato para generar una señal de audio de síntesis de acuerdo con la reivindicación 1, un aparato para codificar una señal de audio de acuerdo con la reivindicación 10, un procedimiento para generar de acuerdo con la reivindicación 12, un procedimiento para codificar de acuerdo con la reivindicación 13, una señal de audio codificada de acuerdo con la reivindicación 14 o un programa de computación de acuerdo con la reivindicación 15.

[0013] La presente invención se basa en la idea básica de que se puede obtener la calidad mejorada recién mencionada y/o una implementación eficiente cuando una porción de tiempo de una señal de audio se convierte a una representación espectral antes de ejecutar una pluralidad de algoritmos diferentes de introducción de parche en el dominio espectral, donde cada algoritmo de parches genera una representación espectral modificada que comprende componentes espectrales en una banda de frecuencia superior derivada de los componentes espectrales correspondientes en una banda de frecuencia del núcleo de la señal de audio y seleccionando un primer algoritmo de parches en el dominio espectral de una pluralidad de algoritmos de introducción de parche para una primera porción de tiempo y un segundo algoritmo de parches en el dominio espectral de la pluralidad de algoritmos de introducción de parche para una segunda porción de tiempo diferente de acuerdo con una señal de control de introducción de parche para obtener la representación espectral modificada. Con esta medida, se puede prevenir la reducción de la calidad y/o flexibilidad debida a una conmutación entre dos algoritmos de introducción de parche en diferentes dominios y, por lo tanto, el procesamiento puede ser menos complejo y mantener al mismo tiempo la calidad perceptual.

[0014] De acuerdo con una realización de la presente invención, un aparato para generar una señal de audio de síntesis que utiliza una señal de control de parche comprende un primer convertidor, un generador de parches en el dominio espectral, un manipulador de reconstrucción de alta frecuencia y un combinador. El primer convertidor está configurado para convertir una porción de tiempo de una señal de audio en una representación espectral. El generador de parche en el dominio espectral está configurado para ejecutar una pluralidad de algoritmos diferentes en el dominio espectral, donde cada algoritmo de parches genera una representación espectral modificada que comprende componentes espectrales en una banda de frecuencia superior derivados de los correspondientes componentes espectrales de una banda de frecuencia del núcleo de la señal de audio. El generador de parche en el dominio espectral está configurado además para seleccionar un primer algoritmo de parches en el dominio espectral entre la pluralidad de algoritmos de introducción de parche para una primera porción de tiempo y un segundo algoritmo de parches en el dominio espectral entre la pluralidad de algoritmos de introducción de parche para una segunda porción de tiempo diferente de acuerdo con la señal de control de parche para obtener la representación espectral modificada. El manipulador de reconstrucción de alta frecuencia está configurado para manipular la representación espectral modificada o una señal derivada de la representación espectral modificada de acuerdo con un parámetro de replicación de bandas espectrales para obtener una señal de extensión del ancho de banda. El combinador está configurado para combinar la señal de audio que consta de componentes espectrales en la banda de frecuencia del núcleo o una señal derivada de la señal de audio con la señal de extensión de ancho de banda para obtener la señal de audio de síntesis.

[0015] De acuerdo con otra realización de la presente invención, un aparato para codificar una señal de audio comprende un codificador del núcleo, un extractor de parámetros y una calculadora de parámetros. La señal de audio comprende una banda de frecuencia del núcleo y una banda de frecuencia superior. El codificador del núcleo

está configurado para codificar la señal de audio dentro de la banda de frecuencia del núcleo. El extractor de parámetros está configurado para extraer una señal de control de parches de la señal de audio, donde la señal de control de parche indica un algoritmo de parches seleccionado entre una pluralidad de algoritmos de introducción de parche en el dominio espectral diferentes y donde el algoritmo de parche seleccionado se ha de ejecutar en un dominio espectral para generar una señal de audio de síntesis en un decodificador de extensión de ancho de banda. La calculadora de parámetros está configurada para calcular el parámetro de replicación de la banda espectral de la banda de frecuencia superior.

[0016] De acuerdo con otra realización, un flujo de datos de señal de audio codificada comprende una señal de audio codificada dentro de una banda de frecuencia del núcleo, una señal de control de introducción de parche, donde la señal de control de introducción de parche indica un algoritmo de parches seleccionado entre una pluralidad de algoritmos de introducción de parche en el dominio espectral diferentes y donde el algoritmo de parche seleccionado se ha de ejecutar en un dominio espectral para generar una señal de audio de síntesis en un decodificador de extensión de ancho de banda y un parámetro de replicación de la banda espectral calculado de una banda de frecuencia superior de la señal de audio.

[0017] Por lo tanto, las realizaciones de la presente invención se relacionan con un concepto para conmutar entre por lo menos por lo menos dos algoritmos diferentes de introducción de parche en el dominio espectral de un grupo de algoritmos de introducción de parche en el dominio espectral. El grupo de algoritmos de introducción de parche puede comprender un primer algoritmo de parches que comprende una transposición armónica basada en un vocodificador de fase única y funcionalidades de SBR por copiado no armónica, un segundo algoritmo de parches que comprende una transposición armónica basada en un vocodificador de fases múltiples, un tercer algoritmo de parches que comprende funcionalidades de SBR por copiado no armónica y un cuarto algoritmo de parches que comprende una distorsión no lineal. Más aun, la extensión del ancho de banda se puede ejecutar de tal manera que la señal de extensión de ancho de banda comprende la banda de frecuencia superior que tiene una frecuencia máxima de por lo menos cuatro veces la frecuencia de cruce en la banda de frecuencia del núcleo.

[0018] Como resultado, al conmutar entre los por lo menos dos algoritmos de parches diferentes en el dominio espectral, se puede obtener una complejidad reducida en la misma calidad perceptual, como por ejemplo dentro de una situación de extensión del ancho de banda.

[0019] Otras realizaciones de la presente invención se relacionan con un aparato que no comprende un transformador de tiempo/ frecuencia para transformar una señal en el dominio del tiempo derivada de de la representación espectral modificada al dominio espectral. Por lo tanto, las realizaciones dan lugar a que el manipulador de reconstrucción de frecuencia pueda ser operativo en la representación espectral modificada directamente sin requerir más transformación (por ejemplo, un análisis de QMF) del dominio del tiempo al dominio espectral, como en el caso de una técnica de parche/procesamiento adicional que es operativa en diferentes dominios.

[0020] Otras realizaciones de la presente invención se relacionan con un extractor de parámetros que está configurado para determinar, de la pluralidad de diferentes algoritmos de parches en el dominio espectral, un algoritmo de parche seleccionado. En este caso, el algoritmo de parche seleccionado se basa en una comparación de la señal de audio, o una señal derivada de la señal de audio, con una pluralidad de señales de extensión de ancho de banda que fueran obtenidas mediante la ejecución de la pluralidad de algoritmos de parche en el dominio espectral y la manipulación de una representación modificada de una porción de tiempo de la señal de audio. Por lo tanto, las realizaciones presentan un procedimiento para seleccionar el algoritmo de parche óptimo para generar una señal de audio de síntesis en un decodificador de extensión del ancho de banda.

[0021] Se pueden utilizar parámetros de control para decidir qué parche es el más apropiado. Para obtenerlo, se puede emplear una etapa de análisis por síntesis; es decir, se pueden aplicar todos los parches y se elige el mejor de acuerdo con el objetivo. En la modalidad preferida de la invención, el objetivo consiste en obtener la mejor calidad perceptual de la restitución. En modalidades alternativas, se tiene que optimizar una función objetivo. Por ejemplo, la función objetivo puede ser conservar lo plano de las HF originales al máximo posible.

[0022] Por un lado, la selección del parche se puede realizar sólo en el codificador tomando en cuenta la señal original, la señal sintetizada o ambas. La decisión (señal de control de parches) es transmitida entonces al decodificador. Por otro lado, la selección se puede ejecutar de manera sincrónica en los lados del codificador y decodificador teniendo en cuenta sólo el ancho de banda del núcleo de la señal sintetizada. Este último procedimiento no necesita generar información secundaria adicional.

[0023] A continuación, se explican las realizaciones de la presente invención con referencia a los dibujos que las acompañan, en los cuales:

La Fig. 1a ilustra un diagrama de bloques de una realización de un aparato para generar una señal de audio de síntesis que utiliza una señal de control de parches;

La Fig. 1b ilustra un diagrama de bloques de una implementación de un generador de parches en el dominio espectral de la Fig. 1a;

La Fig. 2a ilustra un diagrama de bloques de otra realización de un aparato para generar una señal de audio de síntesis;

5 La Fig. 2b presenta una ilustración esquemática de un esquema de extensión de ancho de banda;

La Fig. 3 presenta una ilustración esquemática de un primer ejemplo de algoritmo de parches;

La Fig. 4 presenta una ilustración esquemática de un segundo ejemplo de algoritmo de parches;

La Fig. 5 presenta una ilustración esquemática de un tercer ejemplo de algoritmo de parches;

La Fig. 6 presenta una ilustración esquemática de un cuarto ejemplo de algoritmo de parches;

10 La Fig. 7 ilustra un diagrama de bloques de una realización de la Fig. 1a sin transformador de tiempo/ frecuencia colocado después del generador de parches en el dominio espectral;

La Fig. 8 ilustra un diagrama de bloques de una realización de Fig. 1a con un segundo convertidor (convertidor de frecuencia/ tiempo);

La Fig. 9 ilustra un diagrama de bloques de una realización de un aparato para codificar una señal de audio;

15 La Fig. 10 ilustra un diagrama de bloques de otra realización de un aparato para codificar una señal de audio y

La Fig. 11 ilustra una vista general de una realización correspondiente a un esquema de parches en el dominio de la frecuencia.

[0024] La Fig. 1a ilustra un diagrama de bloques de un aparato 100 para generar una señal de audio de síntesis 145 utilizando una señal de control de parches 119 de acuerdo con una realización de la presente invención. El aparato 20 100 comprende un primer convertidor 110, un generador de parches en el dominio espectral 120, un manipulador de reconstrucción de alta frecuencia 130 y un combinador 140. El primer convertidor 110 está configurado para convertir una porción de tiempo de una señal de audio 105 en una representación espectral 115. El generador de parches en el dominio espectral 120 está configurado para ejecutar una pluralidad 117-1 de diferentes algoritmos de parches en el dominio espectral, donde cada algoritmo de parche genera una representación espectral modificada 25 125 que comprende componentes espectrales en una banda de frecuencia superior derivada de los componentes espectrales correspondientes de una banda de frecuencia del núcleo de la señal de audio 105. Como se ilustra en la Fig. 1b, el generador de parches en el dominio espectral 120 puede estar configurado para seleccionar un primer algoritmo de parches en el dominio espectral 117-2 de la pluralidad 117-1 de algoritmos de parches correspondiente a una primera porción de tiempo 107-1 y un segundo algoritmo de parches en el dominio espectral 30 117-3 de la pluralidad 117-1 de algoritmos de parches correspondiente a una segunda porción de tiempo diferente 107-2 de acuerdo con la señal de control de parches 119 para obtener la representación espectral modificada 125.

[0025] El manipulador de reconstrucción de alta frecuencia 130 está configurado para manipular la representación espectral modificada 125 o una señal derivada de la representación espectral modificada 125 de acuerdo con un parámetro de replicación de la banda espectral 127 para obtener una señal de extensión del ancho de banda 135. La señal derivada de la representación espectral modificada 125 puede ser, por ejemplo, una señal en el dominio QMF obtenida después de aplicar un análisis de QMF a una señal modificada en el dominio del tiempo basado en la representación espectral modificada 125. El combinador 140 está configurado para combinar la señal de audio 105 que tiene componentes espectrales en la banda de la frecuencia del núcleo o una señal derivada de la señal de audio 105 con la señal de ancho de banda extendido 135 para obtener la señal de audio de síntesis 145. En este caso, la señal derivada de la señal de audio 105 puede ser, por ejemplo, una señal de baja frecuencia decodificada que fuera obtenida después de decodificar una señal de audio codificada dentro de la banda de la frecuencia del núcleo.

[0026] Como se puede apreciar en la Fig. 1a, el generador de parches en el dominio espectral 120 del aparato 100 está implementado de manera que sea operativo en un dominio espectral y no en un dominio del tiempo.

45 **[0027]** La Fig. 2a ilustra un diagrama de bloques de otra realización de un aparato 200 para generar la señal de audio de síntesis 145. En este caso, se omiten los componentes del aparato 200 de la Fig. 2a, que son los mismos que en el aparato 100 de la Fig. 1a, y no se los ilustra ni describe nuevamente. En la realización ilustrada en la Fig. 2a, el generador de parches en el dominio espectral 120 del aparato 200 está configurado para ejecutar por lo menos dos algoritmos de parche diferentes en el dominio espectral de un grupo 203 de algoritmos de parche en el dominio espectral. El grupo 203 de algoritmos de parches comprende un primer algoritmo de parche 205-1 que comprende una transposición armónica basada en un vocodificador de fase única y funcionalidades de SBR por copiado no armónico, un segundo algoritmo de parches 205-2 que comprende una transposición armónica basada en un vocodificador de fases múltiples, un tercer algoritmo de parches 205-3 que comprende funcionalidades de SBR por copiado no armónico y un cuarto algoritmo de parches 205-4 que comprende una distorsión no lineal.

[0028] Como se ilustra en la Fig. 2b, el aparato 200 puede estar adaptado para ejecutar una extensión del ancho de banda de manera tal que la señal de ancho de banda extendido 135 comprende la banda de frecuencia superior 220 que tiene una frecuencia máxima 225 de por lo menos cuatro veces la frecuencia de cruce 215 en la banda de la frecuencia del núcleo 210. En el contexto de la SBR, el valor típico de la frecuencia de cruce 215 que se define como la frecuencia más alta de la banda de la frecuencia del núcleo 210 puede estar, por ejemplo, en un rango inferior a 4 kHz, 5 kHz o 6 kHz. En consecuencia, la frecuencia máxima 225 de la banda de frecuencia superior 220 puede ser, por ejemplo, de aproximadamente 16 kHz, 20 kHz o 24 kHz.

[0029] La Fig. 3 presenta una ilustración esquemática de un primer ejemplo de algoritmo de parches 205-1. En particular, el generador de parches en el dominio espectral 120 está configurado para ejecutar un algoritmo de parches seleccionado entre por lo menos dos algoritmos de parche diferentes en el dominio espectral, donde el algoritmo de parche seleccionado comprende el primer algoritmo de parches 205-1. El primer algoritmo de parches 205-1 comprende una transposición armónica basada en un vocodificador de fase única 305 que comprende un factor de extensión de ancho de banda (σ) de dos que controla una transformada de una banda de la frecuencia de origen 310 extraída de la banda de la frecuencia del núcleo 210 en una primera banda de frecuencia objetivo 310'. En este caso, las fases de los componentes espectrales en la banda de la frecuencia de origen 310 se multiplican por el factor de extensión del ancho de banda (σ) de manera tal que la primera banda de la frecuencia objetivo 310 tiene frecuencias en el rango de la frecuencia de cruce (f_x) al doble de la frecuencia de cruce (f_x). El primer algoritmo de parches 205-1 comprende además funcionalidades de SBR de copiado no armónico 315 para transformar los componentes espectrales de la primera banda de frecuencia objetivo 310' en una segunda banda de frecuencia objetivo 320' mediante un primer copiado, de tal manera que la segunda banda de frecuencia objetivo 320' tiene frecuencias en el rango del doble de la frecuencia de cruce (f_x) a tres veces la frecuencia de cruce (f_x) y para transformar a su vez los componentes espectrales de la segunda banda de frecuencia objetivo 320' en una tercera banda de frecuencia objetivo 330' mediante una segunda copia, de tal manera que la tercera banda de frecuencia objetivo 330' tiene frecuencias en el rango de tres veces la frecuencia de cruce (f_x) a cuatro veces la frecuencia de cruce (f_x) incluida en la banda de frecuencia superior 220, donde la banda de frecuencia superior 220 comprende la primera 310', segunda 320' y tercera 330' bandas de frecuencia objetivo. En particular, como se ilustra en la Fig. 3, la señal de ancho de banda extendido 135 comprende la banda de frecuencia superior 220 generada a partir de la banda de frecuencia del núcleo 210, donde la banda de frecuencia superior 220 tiene una frecuencia máxima de cuatro veces la frecuencia de cruce (f_x).

[0030] La Fig. 4 presenta una ilustración esquemática de un segundo ejemplo de algoritmo de parches 205-2. En este caso en particular, el generador de parches en el dominio espectral 120 está configurado para ejecutar un algoritmo de parches seleccionado entre los por lo menos dos algoritmos de parche diferentes en el dominio espectral, donde el algoritmo de parches seleccionado comprende el segundo algoritmo de parches 205-2. El segundo algoritmo de parches 205-2 comprende una transposición armónica basada en un vocodificador de fases múltiples 405 que comprende un primer factor de extensión del ancho de banda (σ_1) de 2 que controlan una transformada de una primera banda de frecuencia de origen 410 extraída de la banda de frecuencia del núcleo 210 en una primera banda de frecuencia objetivo 410'. En este caso, las fases de los componentes espectrales de la primera banda de frecuencia de origen 410 se multiplican por el primer factor de extensión del ancho de banda (σ_1) por lo que la primera banda de frecuencia objetivo 410' tiene frecuencias en el rango que va de la frecuencia de cruce (f_x) al doble de la frecuencia de cruce (f_x). El segundo algoritmo de parches 205-2 comprende además un segundo factor de extensión del ancho de banda (σ_2) de 3 que controlan una transformada de una segunda banda de frecuencia de origen 420-1, 420-2 extraída de la banda de frecuencia del núcleo 210 en una segunda banda de frecuencia objetivo 420', 420''. En este caso, las fases de los componentes espectrales de la segunda banda de frecuencia de origen 420-1, 420-2 se multiplican por el segundo factor de extensión del ancho de banda (σ_2) por lo que la segunda banda de frecuencia objetivo 420', 420'' tiene frecuencias en el rango del doble de la frecuencia de cruce (f_x) a tres veces la frecuencia de cruce (f_x) o en el rango comprendido desde la frecuencia de cruce (f_x) a tres veces la frecuencia de cruce (f_x), respectivamente. Por último, el segundo algoritmo de parches 205-2 comprende además un tercer factor de extensión del ancho de banda (σ_3) de 4 que controlan una transformada de una tercera banda de frecuencia de origen 430-1, 430-2 extraída de la banda de frecuencia del núcleo 210 en una tercera banda de frecuencia objetivo 430', 430''. En este caso, las fases de los componentes espectrales de la tercera banda de frecuencia de origen 430-1, 430-2 se multiplican por el tercer factor de extensión del ancho de banda (σ_3) de manera tal que la tercera banda de frecuencia objetivo 430', 430'' tiene frecuencias en el rango de tres veces la frecuencia de cruce (f_x) a cuatro veces la frecuencia de cruce (f_x) o en el rango de la frecuencia de cruce (f_x) a cuatro veces la frecuencia de cruce (f_x) incluida en la banda de frecuencia superior 220, respectivamente. Como en el caso del primer algoritmo de parches 205-1 expuesto en la Fig. 3, la banda de frecuencia superior 220 de la señal de ancho de banda extendido 135 comprende la primera 410', segunda 420', 420'' y tercera 430', 430'' bandas de frecuencia objetivo con una frecuencia máxima de cuatro veces la frecuencia de cruce (f_x).

[0031] La Fig. 5 presenta una ilustración esquemática de un tercer ejemplo de algoritmo de parches 205-3. En la realización de la Fig. 5, el generador de parches en el dominio espectral 120 está configurado para ejecutar un algoritmo de parches selecto de los por lo menos dos algoritmos de parche diferentes en el dominio espectral, donde el algoritmo de parches seleccionado comprende el tercer algoritmo de parches 205-3. El tercer algoritmo de parches 205-3 comprende funcionalidades de SBR de copiado no armónico 505 para transformar los componentes espectrales de una banda de frecuencia de origen 510 que es la banda de frecuencia del núcleo 210 en una banda

de frecuencia objetivo 510' mediante una primera copia, por lo que la primera banda de frecuencia objetivo 510' tiene frecuencias en el rango de la frecuencia de cruce (f_x) al doble de la frecuencia de cruce (f_x). Los componentes espectrales de la primera banda de frecuencia objetivo 510' se transforman a su vez en una segunda banda de frecuencia objetivo 520' mediante un segundo copiado, por lo que la segunda banda de frecuencia objetivo 520' tiene frecuencias en el rango del doble de la frecuencia de cruce (f_x) a tres veces la frecuencia de cruce (f_x). Por último, los componentes espectrales de la segunda banda de frecuencia objetivo 520' se transforman a su vez en una tercera banda de frecuencia objetivo 530' mediante una tercera copia, por lo que la tercera banda de frecuencia objetivo 530' tiene frecuencias en el rango de tres veces la frecuencia de cruce (f_x) a cuatro veces la frecuencia de cruce (f_x) incluida en la banda de frecuencia superior 220. Una vez más, la banda de frecuencia superior 220 de la señal de ancho de banda extendido 135 comprende la primera 510', segunda 520' y tercera 530' bandas de frecuencia objetivo que tienen una frecuencia máxima de cuatro veces la frecuencia de cruce (f_x).

[0032] La Fig. 6 presenta una ilustración esquemática de un cuarto ejemplo de algoritmo de parches 205–4. En la realización de la Fig. 6, el generador de parches en el dominio espectral 120 está configurado para ejecutar un algoritmo de parches seleccionado entre los por lo menos dos algoritmos de parche diferentes en el dominio espectral, donde el algoritmo de parches seleccionado comprende el cuarto algoritmo de parches 205–4. En este caso, el cuarto algoritmo de parches 205–4 comprende una distorsión no lineal para generar los componentes espectrales de la banda de frecuencia superior 220 con frecuencias en el rango de la frecuencia de cruce (f_x) a cuatro veces la frecuencia de cruce (f_x).

[0033] En general, en las realizaciones de las Figs. 3–6 antes descriptas, los algoritmos de parches en el dominio espectral 205–1; 205–2; 205–3; 205–4 se ejecutan con el generador de parches en el dominio espectral 120 configurado para transformar un componente espectral en una banda inicial 310, 310', 320'; 410, 420–1, 420–2, 430–1, 430–2; 510, 510', 520' derivada de la banda de frecuencia del núcleo 210 o una banda de frecuencia superior no incluida en la banda de la frecuencia del núcleo 210 en un componente espectral objetivo de la banda de frecuencia superior 220 por lo que el componente espectral objetivo es diferente para cada algoritmo de parches en el dominio espectral.

[0034] En particular, el generador de parches en el dominio espectral 120 puede comprender un filtro paso banda para extraer la banda inicial de la banda de frecuencia del núcleo 210 o la banda de frecuencia superior 220, donde se puede seleccionar una característica de paso banda del filtro paso banda de tal manera que la banda inicial se transforme en una correspondiente banda de frecuencia objetivo 310', 320', 330'; 410', 420', 420'', 430', 430''; 510', 520', 530' como se ilustra en la Figs. 3–6.

[0035] Los diferentes algoritmos espectrales en el dominio espectral 205–1; 205–2; 205–3; 205–4 se pueden ejecutar de acuerdo con una eficiencia necesaria tal como siguiendo el esquema de ancho de banda de la Fig. 2b.

[0036] Específicamente, empleando un vocodificador de una o múltiples fases como se ilustra, por ejemplo en la Fig. 3 o la Fig. 4, respectivamente, la estructura de frecuencias se extiende de manera correctamente armónica al dominio de alta frecuencia, puesto que la banda base (por ej., la banda de frecuencia del núcleo 210) se propaga espectralmente en un múltiplo par (por ej. $\sigma_1 = 2$, $\sigma_2 = 3$, $\sigma_3 = 4$), y porque los componentes espectrales de la banda base se combinan con componentes espectrales generados adicionales.

[0037] Un algoritmo de parches basado en un vocodificador de fases puede ser ventajoso si la banda base ya tiene un ancho de banda fuertemente limitado, por ejemplo, utilizando sólo una velocidad de transmisión de datos muy baja. Por ende, la reconstrucción de los componentes de frecuencia superior se inicia a una frecuencia relativamente baja. La frecuencia de cruce típica es, en este caso, inferior a aproximadamente 5 KHz (o incluso inferior a 4 KHz). En esta región, el oído humano es muy sensible a las disonancias debidas a los armónicos incorrectamente situados. Esto puede dar origen a la impresión de tonos “anormales”. Además, los tonos espectralmente espaciados a poca distancia (con la disonancia espectral de aproximadamente 30 Hz a 300 Hz) se perciben en forma de tonos ásperos. La continuación armónica de la estructura de frecuencias de la banda base evita estas impresiones auditivas incorrectas y desagradables.

[0038] Más aun, al emplear funcionalidades de SBR por copiado no armónico como se ilustra, por ejemplo en la Fig. 5, se pueden copiar las regiones espectrales en forma de subbandas a una región de frecuencia más alta o en la región de frecuencias a replicar. Además, el copiado se basa en la observación, como sucede con todos los procedimientos de introducción de parches, de que las propiedades espectrales de las señales de frecuencia más alta son similares, en muchos aspectos, a las propiedades de las señales de la banda base. Hay sólo muy pocas desviaciones entre sí. Además, el oído humano por lo general no es muy sensible a alta frecuencia (típicamente a partir de aproximadamente 5 KHz), especialmente con respecto a un mapeo espectral no preciso. De hecho, esta es una idea clave de la replicación de bandas espectrales en general. El copiado, en particular, tiene la ventaja de que es de fácil y rápida implementación. Este algoritmo de parches tiene además una gran flexibilidad con respecto a los bordes del parche, puesto que la copia del espectro se puede ejecutar en cualquier borde sub-banda.

[0039] Por último, el algoritmo de parches de distorsión no lineal (véase, por ej. la Fig. 6) puede comprender una generación de armónicos mediante corte, limitación, cuadratura, etc. Si, por ejemplo, una señal de difusión está ocupada espectralmente en forma muy rala (por ej. después de aplicar el algoritmo de parches de vocodificador

antes mencionado) es posible que el espectro de difusión pueda estar suplementado aditivamente, opcionalmente, mediante una señal distorsionada a fin de evitar los desagradables huecos de frecuencia.

[0040] Se ha de notar que, además de los algoritmos de parches antes mencionados del grupo 203 de algoritmos de parches (ver Fig. 2a), se pueden ejecutar otros algoritmos de parches dentro del dominio espectral, tal como la producción de espejos espectrales.

[0041] En la realización de la Fig. 7, se ilustra un aparato 700 que no comprende un transformador de tiempo/frecuencia, como se indica por el bloque de guiones 710 para transformar una señal en el dominio del tiempo 705 derivada de la representación espectral modificada 125 a una en el dominio espectral. Esto significa que, en este caso, el manipulador de reconstrucción de alta frecuencia 130 recibe como entrada la representación espectral modificada 125 y no una señal en el dominio de la frecuencia 715, presente a la salida de dicho transformador de tiempo/frecuencia 710.

[0042] La configuración descrita puede ser ventajosa, puesto que, en este caso, el procesamiento adicional de la representación espectral modificada 125 ejecutado por el manipulador de reconstrucción de alta frecuencia 130 puede tener lugar fácilmente en el mismo dominio (por ej. el dominio FFT o QMF) en que opera que el algoritmo de parches ejecutado por el generador de parches en el dominio espectral 120. Por lo tanto, no es necesario otra transformación entre diferentes dominios, como por ejemplo la transformación del dominio del tiempo al dominio espectral (por ej. un análisis QMF), lo que permite una implementación más fácil.

[0043] En la realización de la Fig. 8, se ilustra un aparato 800 que además comprende un segundo convertidor 810 para convertir la representación espectral modificada 125 al dominio del tiempo. Una vez más, se omiten los componentes del aparato 800 de la Fig. 8, que pueden corresponder a los del aparato 100 de la Fig. 1a. Como se ilustra en la Fig. 8, el segundo convertidor 810 puede estar adaptado para aplicar una síntesis compatible con un análisis aplicado por el primer convertidor 110. En este caso, el primer convertidor 110 está configurado para ejecutar una conversión que tiene una primera longitud de conversión 111, en tanto que el segundo convertidor 810 está configurado para ejecutar una conversión que tiene una segunda longitud de conversión. En particular, la segunda longitud de conversión puede depender de una característica de extensión del ancho de banda por el hecho de que se tiene en cuenta la relación de la frecuencia máxima (F_{max}) en la banda de frecuencia superior 220 y la frecuencia de cruce (f_x) en la banda de la frecuencia del núcleo 210 y la primera longitud de conversión 111.

[0044] En realizaciones de la presente invención, el primer convertidor 110 puede ser implementado, por ejemplo, para ejecutar una transformada rápida de Fourier (FFT), una transformada corta de Fourier (STFT), una transformada discreta de Fourier (DFT) o un análisis QMF, en tanto que el segundo convertidor 810 puede estar implementado, por ejemplo, para ejecutar una transformada rápida de Fourier inversa (IFFT), una transformada corta de Fourier inversa (ISTFT), una transformada discreta de Fourier inversa (IDFT) o una síntesis por QMF.

[0045] Específicamente, la segunda longitud de conversión puede ser elegida de manera que sea igual a la relación f_{max}/f_x multiplicada por la primera longitud de conversión 111. De esta manera, la segunda longitud de conversión o resolución de frecuencia aplicada por el segundo convertidor 810 se adapta fácilmente a la característica de extensión del ancho de banda del esquema de extensión de ancho de banda ilustrado en la Fig. 2b. Esto es porque la característica de extensión del ancho de banda está gobernada esencialmente por la relación antes expuesta (f_{max}/f_x) que corresponde a una velocidad de muestreo efectivo más elevada de acuerdo con el principio de Nyquist.

[0046] La Fig. 9 ilustra un diagrama de bloques de una realización de un aparato 900 para codificar una señal de audio 105. La señal de audio 105 comprende una banda de frecuencia del núcleo 210 y una banda de frecuencia superior 220. En particular, el aparato 900 para codificar comprende un codificador de núcleo 910, un extractor de parámetros 920 y una calculadora de parámetros 930. El codificador de núcleo 910 está configurado para codificar la señal de audio 105 dentro de la banda de la frecuencia del núcleo 210 para obtener una señal de audio codificada 915 codificada dentro de la banda de la frecuencia del núcleo 210. Más aun, el extractor de parámetros 920 está configurado para extraer una señal de control de parches 119 de la señal de audio 105, donde la señal de control de parches 119 indica un algoritmo de parches seleccionado de una pluralidad 117-1 de algoritmos espectrales diferentes en el dominio espectral. Específicamente, el algoritmo de parches seleccionado se puede ejecutar en un dominio espectral para generar la señal de audio de síntesis en un decodificador de extensión de ancho de banda. Por último, la calculadora de parámetros 930 está configurada para calcular un parámetro de SBR 127 de la banda de frecuencia superior 220. El parámetro de SBR 127 calculado de la banda de frecuencia superior 220, donde la señal de control de parches 119 indica el algoritmo de parches seleccionado y la señal de audio codificada 915 codificada dentro de la banda de la frecuencia del núcleo 210 puede constituir una señal de audio codificada 935 que se ha de almacenar o transmitir dentro de un flujo de bits.

[0047] En la realización de la Fig. 9, el extractor de parámetros 920 puede estar configurado para analizar la señal de audio 105 o una señal derivada de la señal de audio 105 para determinar la señal de control de parches 119 sobre la base de una característica de señal de la señal analizada. Por ejemplo, la señal de control de parches 119 puede indicar un primer algoritmo de parches correspondiente a una primera porción de tiempo 107-1 de la señal

analizada que se caracteriza como “voz” y un segundo algoritmo de parches correspondiente a una segunda porción de tiempo 107–2 de la señal analizada, que se caracteriza como “música estacionaria”.

5 **[0048]** En consecuencia, en el caso de una señal de voz, se puede emplear un procesamiento basado en un modelo de origen de voz o un modelo de generación de información, como por ejemplo dentro de un dominio LPC (codificación predictiva lineal), en tanto que, en el caso de la música estacionaria, se puede utilizar un modelo de origen estacionario o un modelo de destino de la información. Mientras que en el caso anterior, se describe el sistema de generación de voz humana/sonido que genera sonido, en el último caso se describe el sistema auditivo humano que recibe sonido.

10 **[0049]** Además, se puede implementar un esquema de procesamiento dependiente de la señal conmutando entre una transposición armónica correspondiente a una porción de tiempo que comprende un evento transitorio y una operación por copiado no armónico correspondiente a una porción de tiempo que no comprende un evento transitorio.

15 **[0050]** El procedimiento citado que corresponde a un bucle abierto se basa en un análisis directo de la señal de audio 105 o una señal derivada de la señal de audio 105 con respecto a su característica de señal. Por otro lado, el extractor de parámetros 920 también puede ser operativo en un bucle cerrado correspondiente a una implementación de “análisis por síntesis”.

20 **[0051]** En la realización de la Fig. 10, se ilustra un aparato 1000 para codificar una señal de audio 105 dentro de esa implementación de análisis por síntesis. Específicamente, el extractor de parámetros 920 del aparato 1000 para codificar puede estar configurado para determinar, de la pluralidad 117–1 de diferentes algoritmos de parches en el dominio espectral, el algoritmo de parches seleccionado. En este caso, el algoritmo de parches seleccionado se puede basar en una comparación de la señal de audio 105 o una señal derivada de la señal de audio 105 con la pluralidad 1005 de señales de ancho de banda extendido que se han obtenido al ejecutar la pluralidad 117–1 de algoritmos de parche en el dominio espectral y manipulando una representación espectral modificada 125 de una porción de tiempo de la señal de audio 105. La comparación se puede llevar a cabo, por ejemplo, por medio de una
25 unidad de selección de parches 1010 calculando los parámetros (SFM1005) de medida de la suavidad (o lo plano) espectral (SFM) de la pluralidad 1005 de señales de ancho de banda extendido y la señal de audio 105 (SFMref), comparando los parámetros SFM calculados SFM1005 y SFMref y seleccionando, de la pluralidad 117–1 de algoritmos de parches, un algoritmo de parche específico (óptimo) por el cual la desviación de los parámetros SFM comparados sea mínima. Por último, el algoritmo de parche óptimo seleccionado puede estar indicado por la señal de control de parches 119 presente a la salida del extractor de parámetros 920.

30 **[0052]** La Fig. 11 ilustra una vista general de una realización correspondiente a un esquema de introducción de parches en un dominio de la frecuencia. En particular, se ilustra un aparato 1100 para generar una señal de extensión del ancho de banda como el que corresponde al esquema de extensión de ancho de banda de la Fig. 2b. En la realización de la Fig. 11, la señal de audio 105 está representada por los datos de PCM (modulación por pulsos codificados) 1101 que tiene una longitud de cuadro de 1024 muestras (‘cuadro: 1024’). Los datos de PCM 1101 pueden consistir, por ejemplo, en una señal de baja frecuencia decodificada que comprende una banda base derivada de la señal de audio codificada 935, señal de audio codificada 935 que ha sido transmitida desde un aparato para codificar tal como el codificador 900. A continuación se puede utilizar un decrementador (“downsampler”) 1110 para reducir el número de operaciones por muestra de los datos de PCM 1101 en un factor de 2, por ejemplo, para obtener una señal con número reducido de operaciones por muestra 1115. La señal decrementada 1115 es provista a su vez a un generador de ventanas de análisis 1120 indicado por un bloque titulado “ventana” que puede estar configurado para generar una pluralidad de bloques consecutivos de muestras de audio en ventanas superpuestas. En este caso, cada bloque de la pluralidad de bloques consecutivos puede comprender, por ejemplo, 512 muestras de audio. Además, se puede ajustar una primera distancia temporal entre dos bloques consecutivos de muestras de audio, por ejemplo, para que correspondan a las 64 muestras indicadas por “Inc = 64”. La superposición de los bloques consecutivos de muestras de audio puede ser controlada, a su vez, seleccionando una función de ventana de análisis adecuada (óptima) de una pluralidad de funciones de ventana de análisis diferentes aplicadas por el generador de ventanas 1120. Se aplica asimismo una porción de tiempo 1125 de la señal de audio 105, que puede corresponder a un bloque consecutivo de la pluralidad de bloques consecutivos de muestras de audio, al primer convertidor 110 que puede estar implementado, por ejemplo, en forma de procesador FFT 1130 que tiene la primera longitud de conversión 111 de $N = 512$. El procesador FFT 1130 puede estar configurado para convertir la porción de tiempo 1125 en la representación espectral 115, lo que se puede implementar, por ejemplo, en forma polar 1135–1. En particular, esta representación espectral 1135–1 comprende información de magnitud 1135–2 e información de fases 1135–3, que a su vez son procesadas por un generador de parches en el dominio espectral 1141, que puede corresponder al generador de parches en el dominio espectral 120 de la Fig. 2a. El generador de parches en el dominio espectral 1141 de la Fig. 11 puede comprender un primer algoritmo de parches 1141–1 indicado por “vocodificador de fases más copiado” correspondiente al primer algoritmo de parches 205–1, un segundo algoritmo de parches 1143–1 indicado por “vocodificador de fases” que corresponde al segundo algoritmo de parches 205–2, un tercer algoritmo de parches indicado por “función tipo SBR” que corresponde al tercer algoritmo de parches 205–3 y un cuarto algoritmo de parches 1147–1 indicado por “otra función, por ej. distorsión no lineal” que corresponde al cuarto algoritmo de parches 205–4 del grupo 203 de algoritmos de parches ilustrado en la Fig. 2a.

[0053] Como se describiera anteriormente con respecto al contexto de la Fig. 2a, el primer algoritmo de parches 1141-1 comprende un vocodificador de fase única 1141-2 y funcionalidades por copiado no armónico 1141-3, 1141-4. Más aun, el segundo algoritmo de parches 1143-1, que se basa en la operación de un vocodificador de fases múltiples comprende un vocodificador de primera fase 1143-2, un vocodificador de segunda fase 1143-3 y un tercer vocodificador 1143-4. Más aun, el tercer algoritmo de parches 1145-1 comprende funcionalidades de SBR por copiado no armónico que ejecuta una primera operación de copiado 1145-2, una segunda operación de copiado 1145-3 y una tercera operación de copiado 1145-4. Por último, el cuarto algoritmo de parches 1147-1 comprende una funcionalidad de distorsión no lineal.

[0054] Específicamente, en la realización de la Fig. 11, los subcomponentes de los bloques de algoritmos de parches 1141-1, 1143-1, 1145-1, 1147-1 pueden corresponder a los de los bloques 205-1, 205-2, 205-3, 205-4 de la Fig. 2a. Además, el símbolo z ('banda de cruce') puede corresponder a la frecuencia de cruce (f_x).

[0055] Más aun, se puede utilizar un seleccionador de parches 1150 para producir una señal de control de parches 1155 correspondiente a la señal de control de parches 119 para controlar al generador de parches en el dominio espectral 1141 de tal manera que se ejecuten por lo menos dos algoritmos de parche diferentes en el dominio espectral del grupo 1141-1, 1143-1, 1145-1, 1147-1 de algoritmos de parches, llevando a una representación espectral modificada 1149 que corresponde a la representación espectral modificada 125.

[0056] La representación espectral modificada 1149 puede ser procesada (opcionalmente) por un interpolador subsiguiente 1160 para obtener una representación espectral modificada interpolada 1165. A continuación, se puede transferir la representación espectral modificada interpolada 1165 al segundo convertidor 810, que puede estar implementado, por ejemplo, en forma de procesador iFFT 1170 con una segunda longitud de conversión de $N = 2048$. En este caso, como se describe en lo que respecta a la Fig. 8, se ajusta la segunda longitud de conversión de $N = 2048$ de manera que sea exactamente cuatro veces más alta que la primera longitud de conversión de $N = 512$. De esa manera, se puede tomar en cuenta la característica de extensión del ancho de banda del esquema de extensión de ancho de banda ejecutado con los algoritmos espectrales diferentes en el dominio espectral, lo que se ha explicado anteriormente en forma detallada.

[0057] El procesador iFFT 1170 puede estar configurado para convertir la representación espectral modificada interpolada 1165 en una señal modificada en el dominio del tiempo 1175 que corresponde a la señal modificada en el dominio del tiempo 815 de la Fig. 8. Seguidamente, se puede transferir la señal modificada en el dominio del tiempo 1175 a un generador de ventanas de síntesis 1180 para aplicar una función de ventana de síntesis a la señal modificada en el dominio del tiempo 1175 para obtener una señal en el dominio del tiempo en ventana modificada 1185. En este caso, se hace coincidir la función de ventana de síntesis con la función de ventana de análisis de manera que el efecto de la aplicación de la ventana de análisis se compense mediante la aplicación de la función de ventana de síntesis.

[0058] Dado que la señal modificada en el dominio del tiempo expuesta en ventana 1185 debe ser muestreada a una velocidad de muestreo efectivo más elevada (por ej. 32 KHz) en comparación con la velocidad de muestreo original (por ej. 8 KHz) debido a la extensión del ancho de banda, la señal modificada en el dominio del tiempo expuesta en ventana 1185 puede ser agregada de manera superpuesta, en última instancia, en un bloque 1190 indicado por "superposición y adición" por el hecho de que una relación de segunda distancia temporal de, por ejemplo 256 muestras, indicada por "Inc = 256" aplicada por el bloque 1190 y la primera distancia temporal de, por ejemplo 64 muestras, aplicada por el generador de ventanas de análisis 1120 (por ej. relación = 4) ha de ser igual a la relación de la velocidad de muestreo efectivo más alta y la velocidad de muestreo original. De esta manera, se puede obtener una señal de salida 1195 que tenga la misma característica de superposición que la señal original (decrementada) 1115. La señal de salida 1195 provista por el aparato 1100 puede ser procesada a su vez a partir del manipulador de reconstrucción de alta frecuencia 130 como se ilustra en la Fig. 1a para obtener, en última instancia, una señal replicada con su ancho de banda extendido.

[0059] Se debe tener en cuenta que, en la realización de la Fig. 11, todos los diferentes algoritmos de parches se ejecutan en el mismo dominio, por ejemplo en el dominio de la frecuencia. El dominio puede ser un dominio QMF y se realiza en la SBR o cualquier otro dominio, tal como por transposición de Fourier. La generación de datos del parche en sí se puede llevar a cabo en un dominio diferente. En ese caso, la totalidad de la introducción de parche se lleva a cabo en el mismo dominio.

[0060] Además, se pueden asociar diferentes modelos de origen con el parche considerado en la selección. Por ejemplo, se pueden asociar diferentes modelos de origen a la introducción del parche considerado en la selección. Por ejemplo, se puede elegir un modelo de origen de voz utilizado en la extensión de ancho de banda de voz, de acuerdo con lo descrito por Frederik Nagel, Sascha Disch, "A harmonic bandwidth extension method for audio codecs," ICASSP International Conference on Acoustics, Speech and Signal Processing, IEEE CNF, Taipei, Taiwan, abril de 2009, para las señales de voz, en tanto que se puede adoptar un modelo de origen estacionario para la música estacionaria. De la misma manera, como se describiera anteriormente, los transitorios pueden tener su propio modelo para la introducción de parches.

[0061] Más aun, por medio del análisis de superposición y ventanas de síntesis para la transposición de tiempo–frecuencia, se garantizan transiciones suaves entre diferentes esquemas de emparche. Por otro lado, se pueden utilizar ventanas especiales para el análisis y síntesis para posibilitar la menor superposición.

5 **[0062]** En síntesis, en la realización de la Fig. 11, se pueden seleccionar procedimientos para la introducción de parches entre una sencilla operación de copia de secciones de frecuencia lindantes, un esquema de transposición armónica basada en vocodificador de fases y que incluye el copiado de secciones de frecuencia vecinas.

10 **[0063]** Si bien se ha descrito la presente invención en el contexto de diagramas de bloques, donde los bloques representan los componentes de hardware reales o lógicos, también se puede implementar la presente invención mediante un procedimiento implementado por computadora. En este último caso, los bloques representan pasos correspondientes del procedimiento, donde estos pasos representan las funcionalidades ejecutar mediante bloques de hardware lógicos o físicos correspondientes.

15 **[0063]** Las realizaciones aquí descritas son meramente ilustrativas de los principios de la presente invención. Se entiende que las modificaciones y variaciones de las disposiciones y detalles descritos en este documento han de ser evidentes para otras personas con capacitación en la técnica. Por lo tanto, se pretende restringirse sólo al alcance de las reivindicaciones de patente adjuntas y no a los detalles específicos presentados a manera de descripción y explicación de las realizaciones aquí presentadas.

20 **[0065]** Dependiendo de ciertos requisitos de implementación del procedimiento de la invención, el procedimiento de la invención puede ser implementado en hardware o en software. La implementación se puede ejecutar empleando un medio de almacenamiento digital, en un disco específico, un DVD o un CD con señales de control legibles en forma electrónica almacenadas en el mismo, que cooperan con sistemas de computación programables para posibilitar la ejecución de los procedimientos de la invención. En general, la presente invención puede ser implementada, por lo tanto, en forma de producto programa de computación, con un programa almacenado en un portador legible por una máquina, donde el código del programa cumple la función de poner en práctica los procedimientos de la invención al ejecutarse el programa de computación en una computadora. En otras palabras, los procedimientos de la invención consisten, por lo tanto, en un programa de computación que consta de un código de programa para poner en práctica por lo menos uno de los procedimientos de la invención al ejecutarse el programa de computación en una computadora. La señal de audio codificada de acuerdo con la invención puede ser almacenada en cualquier medio de almacenamiento legible por una máquina, como por ejemplo un medio de almacenamiento digital.

30 **[0066]** Las realizaciones de la presente invención permiten que la extensión del ancho de banda tome en cuenta las características de sonido, hardware y señales para el proceso de introducción de parches. La decisión del parche más adecuado se puede tomar dentro de bucles abiertos o cerrados. Por lo tanto, se puede controlar y mejorar la calidad de restitución.

35 **[0067]** El concepto presentado ofrece asimismo la ventaja de que se puede obtener fácilmente una transición suave entre los diferentes algoritmos de parches, dando así lugar a una adaptación rápida y precisa de la extensión del ancho de banda basada en la señal.

[0068] Las aplicaciones más prominentes son decodificadores de audio, que con frecuencia se implementan en dispositivos portátiles y, por consiguiente, funcionan con una fuente de abastecimiento de energía por baterías.

REIVINDICACIONES

1. Un aparato (100; 200; 700; 800; 1100) para generar una señal de audio de síntesis (145) utilizando una señal de control de parches ("patching") (119; 1155), donde el aparato (100; 200; 700; 800; 1100) comprende:

un primer convertidor (110; 1130) para convertir una porción de tiempo (107-1; 107-2; 1125) de una señal de audio (105; 1101) en una representación espectral (115; 1135-1);

5 un generador de parches en el dominio espectral (120; 1141) para ejecutar una pluralidad (117-1) de algoritmos espectrales diferentes en el dominio espectral, donde cada algoritmo de parches genera una representación espectral modificada (125; 1149) que comprende componentes espectrales en una banda de frecuencia superior (220) derivados de los correspondientes componentes espectrales de una banda de frecuencia del núcleo (210) de la señal de audio (105; 1101), y donde el generador de parches en el dominio espectral (120; 1141) está configurado
10 para seleccionar un primer algoritmo de parches en el dominio espectral (117-2) de la pluralidad (117-1) de algoritmos de parches para una primera porción de tiempo (107-1) y un segundo algoritmo de parches en el dominio espectral (117-3) de la pluralidad (117-1) de algoritmos de parches para una segunda porción de tiempo diferente (119; 1155) para obtener la representación espectral modificada (125; 1149);

15 un manipulador de reconstrucción de alta frecuencia (130) para manipular la representación espectral modificada (125; 1149) o una señal derivada de la representación espectral modificada (125; 1195) de acuerdo con un parámetro de replicación de la banda espectral (127) para obtener una señal de extensión del ancho de banda (135) y

20 un combinador (140) para combinar la señal de audio (105; 1101) que tiene componentes espectrales en la banda de frecuencia del núcleo (210) o una señal derivada de la señal de audio (105; 1101) con la señal de ancho de banda extendido (135) para obtener la señal de audio de síntesis (145).

2. El aparato (100; 200; 700; 800; 1100) de acuerdo con la reivindicación 1, en el cual el generador de parches en el dominio espectral (120; 1141) es implementado de manera que sea operativo en un dominio espectral y no en un dominio del tiempo.

25 3. El aparato (200) de acuerdo con la reivindicación 1 o 2, en el cual el generador de parches en el dominio espectral (120) está configurado para ejecutar por lo menos dos algoritmos de parche diferentes en el dominio espectral de un grupo (203) de algoritmos de parche en el dominio espectral, donde el grupo (203) de algoritmos de parches comprende un primer algoritmo de parches (205-1) que comprende una transposición armónica basada en un vocodificador de fase única y funcionalidades de replicación de la banda espectral con copiado no armónico, un
30 segundo algoritmo de parches (205-2) que comprende una transposición armónica basada en un vocodificador de fases múltiples, un tercer algoritmo de parches (205-3) que comprende funcionalidades de replicación de la banda espectral con copiado no armónico y un cuarto algoritmo de parches (205-4) que comprende una distorsión no lineal, donde el aparato (200) está adaptado para ejecutar una extensión del ancho de banda, por lo que la señal de ancho de banda extendido (135) comprende la banda de frecuencia superior (220) que tiene una frecuencia máxima (225; f_{max}) de por lo menos cuatro veces la frecuencia de cruce (215; f_x) en la frecuencia del núcleo (210).

35 4. El aparato de acuerdo con la reivindicación 3, en el cual el generador de parches en el dominio espectral (120) está configurado para ejecutar un algoritmo de parches seleccionado entre los por lo menos dos algoritmos de parche diferentes en el dominio espectral, donde el algoritmo de parches seleccionado comprende el primer algoritmo de parches (205-1), el primer algoritmo de parches (205-1) que comprende una transposición armónica basada en un vocodificador de fase única (305) que comprende un factor de extensión del ancho de banda (σ) de
40 dos que controlan una transformación de una banda de frecuencia de origen (310) extraída de la banda de frecuencia del núcleo (210) en una primera banda de frecuencia objetivo (310'), donde las fases de los componentes espectrales de la banda de frecuencia de origen (310) se multiplican por el factor de extensión del ancho de banda (σ) por lo que la primera banda de frecuencia objetivo (310) tiene frecuencias en el rango comprendido desde la frecuencia de cruce (f_x) al doble de la frecuencia de cruce (f_x), donde el primer algoritmo de parches (205-1) comprende además funcionalidades de replicación de la banda espectral por copiado no armónico (315) para
45 transformar los componentes espectrales comprendidos en la primera banda de frecuencia objetivo (310') en una segunda banda de frecuencia objetivo (320') mediante un primer copiado, de manera tal que la segunda banda de frecuencia objetivo (320') tiene frecuencias en el rango del doble de la frecuencia de cruce (f_x) a tres veces la frecuencia de cruce (f_x) y para transformar además los componentes espectrales de la segunda banda de frecuencia objetivo (320') en una tercera banda de frecuencia objetivo (330') mediante un segundo copiado, por lo que la
50 tercera banda de frecuencia objetivo (330') tiene frecuencias en el rango de tres veces la frecuencia de cruce (f_x) a cuatro veces la frecuencia de cruce (f_x) incluidas en la banda de frecuencia superior (220), donde la banda de frecuencia superior (220) comprende la primera (310'), segunda (320') y tercera (330') bandas de frecuencia objetivo.

55 5. El aparato de acuerdo con la reivindicación 3, en el cual el generador de parches en el dominio espectral (120) está configurado para ejecutar un algoritmo de parches seleccionado entre los por lo menos dos algoritmos de parche diferentes en el dominio espectral, donde el algoritmo de parches seleccionado comprende el segundo

5 algoritmo de parches (205-2), segundo algoritmo de parches (205-2) que comprende una transposición armónica basada en un vocodificador de fases múltiples (405) que comprende un primer factor de extensión del ancho de banda (σ_1) de dos que controlan una transformación de una primera banda de frecuencia (410) extraída de la banda de frecuencia del núcleo (210) en una primera banda de frecuencia objetivo (410'), donde las fases de los componentes espectrales de la primera banda de frecuencia de origen (410) se multiplican por el primer factor de extensión del ancho de banda (σ_1) de tal manera que la primera banda de frecuencia objetivo (410') tiene frecuencias en el rango de la frecuencia de cruce (fx) al doble de la frecuencia de cruce (fx), donde el segundo algoritmo de parches (205-2) comprende además un segundo factor de extensión del ancho de banda (σ_2) de los tres que controlan una transformación de una segunda banda de frecuencia de origen (420-1, 420-2) extraída de la banda de frecuencia del núcleo (210) en una segunda banda de frecuencia objetivo (420', 420''), donde las fases de los componentes espectrales de la segunda banda de frecuencia de origen (420-1, 420-2) se multiplican por el segundo factor de extensión del ancho de banda (σ_2) por lo que la segunda banda de frecuencia objetivo (420', 420'') tiene frecuencias en el rango del doble de la frecuencia de cruce (fx) a tres veces la frecuencia de cruce (fx) o en el rango de la frecuencia de cruce (fx) a tres veces la frecuencia de cruce (fx), donde el segundo algoritmo de parches (205-2) comprende además un tercer factor de extensión del ancho de banda (σ_3) de cuatro que controlan una transformación de una tercera banda de frecuencia de origen (430-1, 430-2) extraída de la banda de frecuencia del núcleo (210) en una tercera banda de frecuencia objetivo (430', 430''), donde las fases de los componentes espectrales de la tercera banda de frecuencia de origen (430', 430'') se multiplican por el tercer factor de extensión del ancho de banda (σ_3) por lo que la tercera banda de frecuencia objetivo (430', 430'') tiene frecuencias en el rango de tres veces la frecuencia de cruce (fx) a cuatro veces la frecuencia de cruce (fx) o en el rango comprendido desde la frecuencia de cruce (fx) a cuatro veces la frecuencia de cruce (fx) incluida en la banda de frecuencia superior (220), donde la frecuencia superior (220) comprende la primera (410'), segunda (420', 420'') y tercera (430', 430'') bandas de frecuencia objetivo.

6. El aparato de acuerdo con la reivindicación 3, en el cual el generador de parches en el dominio espectral (120) está configurado para ejecutar un algoritmo de parches seleccionado entre los por lo menos dos algoritmos de parche diferentes en el dominio espectral, donde el algoritmo de parches seleccionado comprende el tercer algoritmo de parches (205-3), donde el tercer algoritmo de parches (205-3) comprende funcionalidades de replicación de la banda espectral por copiado no armónico (505) para transformar los componentes espectrales de una banda de frecuencia de origen (510) que es la banda de frecuencia del núcleo (210) en una primera banda de frecuencia objetivo (510') mediante un primer copiado, de tal manera que la primera banda de frecuencia objetivo (510') tiene frecuencias en el rango de la frecuencia de cruce (fx) al doble de la frecuencia de cruce (fx), para transformar, a su vez, los componentes espectrales de la primera banda de frecuencia objetivo (510') en una segunda banda de frecuencia objetivo (520') mediante un segundo copiado, por lo que la segunda banda de frecuencia objetivo (520') tiene frecuencias en el rango del doble de la frecuencia de cruce (fx) a tres veces la frecuencia de cruce (fx) y para transformar además los componentes espectrales de la segunda banda de frecuencia objetivo (520') en una tercera banda de frecuencia objetivo (530') mediante un tercer copiado, por lo que la tercera banda de frecuencia objetivo (530') tiene frecuencias en el rango de tres veces la frecuencia de cruce (fx) a cuatro veces la frecuencia de cruce (fx) incluida en la banda de frecuencia superior (220), donde la banda de frecuencia superior (220) comprende la primera (510'), segunda (520') y tercera (530') bandas de frecuencia objetivo.

7. El aparato de acuerdo con la reivindicación 3, en el cual el generador de parches en el dominio espectral (120) está configurado para ejecutar un algoritmo de parches seleccionado entre los por lo menos dos algoritmos de parche diferentes en el dominio espectral, donde el algoritmo de parches seleccionado comprende el cuarto algoritmo de parches (205-4), cuarto algoritmo de parches (205-4) que comprende una distorsión no lineal para generar los componentes espectrales de la banda de frecuencia superior (220) que tienen frecuencias en el rango comprendido desde la frecuencia de cruce (fx) a cuatro veces la frecuencia de cruce (fx).

8. El aparato (700) de acuerdo con una de las reivindicaciones anteriores, donde el aparato (700) no comprende un transformador de tiempo/ frecuencia (710) para transformar una señal en el dominio del tiempo (705) derivada de la representación espectral modificada (125) al dominio espectral.

9. El aparato (800) de acuerdo con una de las reivindicaciones anteriores, donde el aparato (800) comprende además un segundo convertidor (810) para convertir la representación espectral modificada (125) al dominio del tiempo, donde el segundo convertidor (810) está adaptado para aplicar una síntesis que coincide con un análisis aplicado por el primer convertidor (110), donde el primer convertidor (110) está configurado para ejecutar una conversión que tiene una primera longitud de conversión (111), y donde el segundo convertidor (810) está configurado para ejecutar una conversión que tiene una segunda longitud de conversión, donde la segunda longitud de conversión depende de una característica de extensión del ancho de banda por el hecho de que se toma en cuenta la relación de la frecuencia máxima (f_{max}) de la banda de frecuencia superior (220) y la frecuencia de cruce (fx) de la banda de la frecuencia del núcleo (210) y la primera longitud de conversión (111).

10. Un aparato (900; 1000) para codificar una señal de audio (105), donde la señal de audio (105) comprende una banda de frecuencia del núcleo (210) y una banda de frecuencia superior (220), donde el aparato (900; 1000) comprende:

un codificador del núcleo (910) para codificar la señal de audio (105) dentro de la banda de la frecuencia del núcleo (210);

5 un extractor de parámetros (920) para extraer una señal de control de parches (119) de la señal de audio (105), donde la señal de control de parches (119) indica un algoritmo de parches seleccionado de una pluralidad (117-1) de algoritmos espectrales diferentes en el dominio espectral, donde el algoritmo de parches seleccionado se ha de ejecutar sobre una representación espectral obtenida por la conversión de una porción de tiempo de una señal de audio para generar una representación espectral modificada para generar una señal de audio de síntesis en un decodificador de extensión de ancho de banda y

10 una calculadora de parámetros (930) para calcular un parámetro de replicación de la banda espectral (127) de la banda de frecuencia superior (220).

11. El aparato (1000) para codificar de acuerdo con la reivindicación 10, en el cual el extractor de parámetros (920) está configurado para determinar, de la pluralidad (117-1) de algoritmos espectrales diferentes en el dominio espectral, el algoritmo de parches seleccionado, donde el algoritmo de parches seleccionado se basa en una comparación de la señal de audio (105) o una señal derivada de la señal de audio (105) con una pluralidad (1005) de señales de ancho de banda extendido que han sido obtenidas mediante la ejecución de la pluralidad (117-1) de algoritmos de parche en el dominio espectral y la manipulación de una representación espectral modificada (125) de una porción de tiempo de la señal de audio (105).

12. Un procedimiento (100; 200; 700; 800; 1100) para generar una señal de audio de síntesis (145) utilizando una señal de control de parches (190; 1155), procedimiento (100; 200; 700; 800; 1100) que comprende:
20 convertir (110; 1130) una porción de tiempo (107-1; 107-2; 1125) de una señal de audio (105; 1101) en una representación espectral (115; 1135-1);

25 ejecutar (120; 1141) una pluralidad (117-1) de algoritmos espectrales diferentes en el dominio espectral, donde cada algoritmo de parche genera una representación espectral modificada (125; 1149) que comprende componentes espectrales de una banda de frecuencia superior (220) derivados de los correspondientes componentes espectrales de una banda de frecuencia del núcleo (210) de la señal de audio (105; 1101), y seleccionar (120; 1141) un primer algoritmo de parches en el dominio espectral (117-2) de la pluralidad (117-1) de algoritmos de parches para una primera porción de tiempo (107-1) y un segundo algoritmo de parches en el dominio espectral (117-3) de la pluralidad (117-1) de algoritmos de parches para una segunda porción de tiempo diferente (107-2) de acuerdo con la señal de control de parches (119; 1155) para obtener la representación espectral modificada (125; 1149);

30 manipular (130) la representación espectral modificada (125; 1149) o una señal derivada de la representación espectral modificada (125; 1195) de acuerdo con un parámetro de replicación de la banda espectral (127) para obtener una señal de extensión del ancho de banda (135) y

35 combinar (140) la señal de audio (105; 1101) que tiene componentes espectrales en la banda de la frecuencia del núcleo (210) o una señal derivada de la señal de audio (105; 1101) con la señal de ancho de banda extendido (135) para obtener la señal de audio de síntesis (145).

13. Un procedimiento (900; 1000) para codificar una señal de audio (105), donde la señal de audio (105) comprende una banda de frecuencia del núcleo (210) y una banda de frecuencia superior (220), procedimiento (900; 1000) que comprende:

codificar (910) la señal de audio (105) dentro de la banda de la frecuencia del núcleo (210);

40 extraer (920) una señal de control de parches (119) de la señal de audio (105), donde la señal de control de parches (119) indica un algoritmo de parches seleccionado de una pluralidad (117-1) de algoritmos espectrales diferentes en el dominio espectral, donde el algoritmo de parches seleccionado se ha de ejecutar sobre una representación espectral obtenida por la conversión de una porción de tiempo de una señal de audio para generar una representación espectral modificada para generar una señal de audio de síntesis en un decodificador de extensión de ancho de banda y

calcular (930) un parámetro de replicación de la banda espectral (127) de la banda de frecuencia superior (220).

14. Una señal de audio codificada (935) que comprende:

una señal de audio codificada (915) codificada dentro de una banda de frecuencia del núcleo (210);

50 una señal de control de parches (119), donde la señal de control de parches (119) indica un algoritmo de parches seleccionado de una pluralidad (117-1) de algoritmos espectrales diferentes en el dominio espectral, algoritmo de parches seleccionado que se ha de ejecutar sobre una representación espectral obtenida por la conversión de una porción de tiempo de una señal de audio para generar una representación espectral modificada para generar una señal de audio de síntesis (145) en un decodificador de extensión de ancho de banda y

un parámetro de replicación de la banda espectral (127) calculado a partir de una banda de frecuencia superior (220) de la señal de audio (105).

15. Un programa de computación que consta de un código de programa para poner en práctica el procedimiento de acuerdo con la reivindicación 12 o 13, al ejecutarse el programa de computación en una computadora.

100

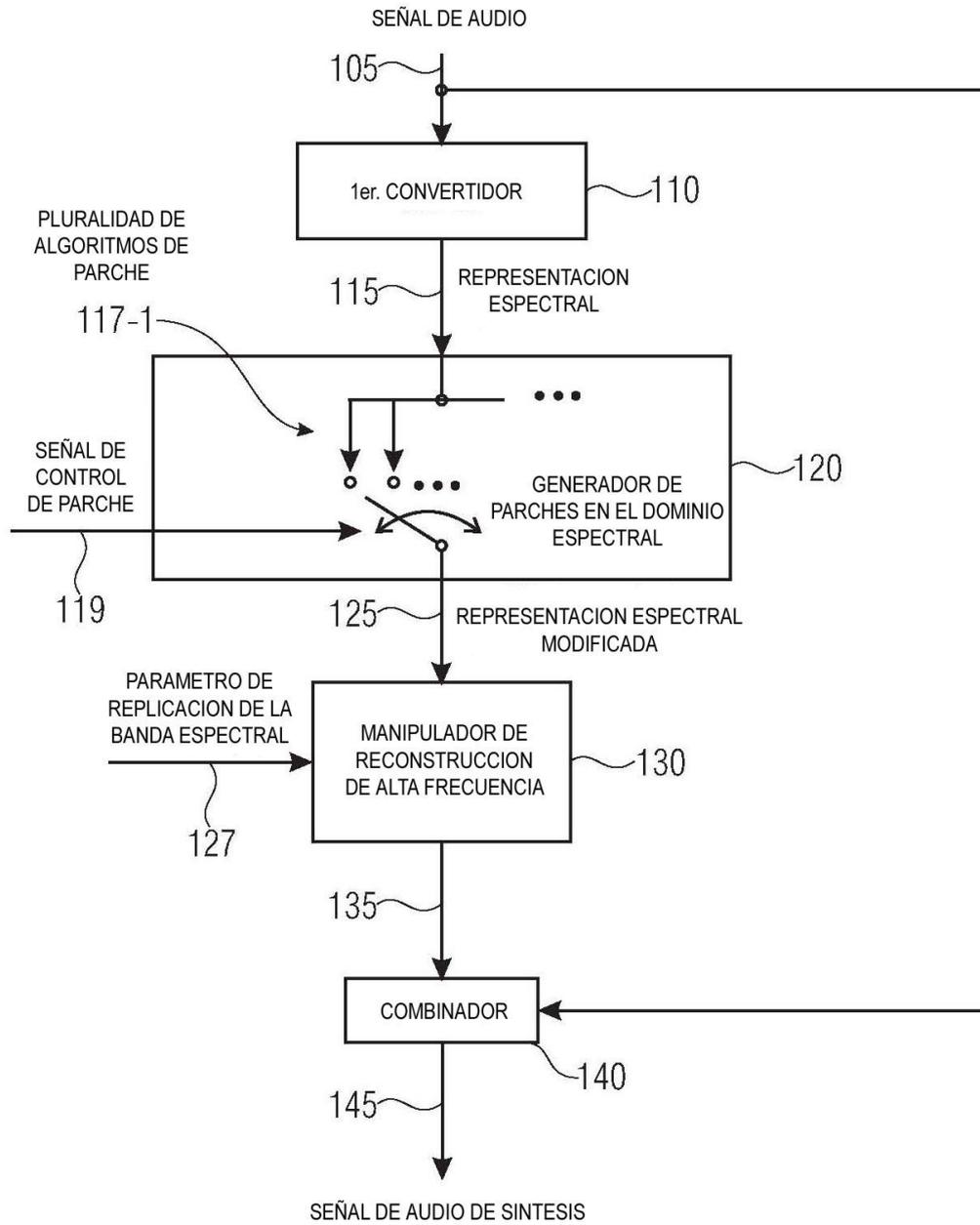


FIG. 1A
(SINTETIZADOR DE SEÑALES DE AUDIO)

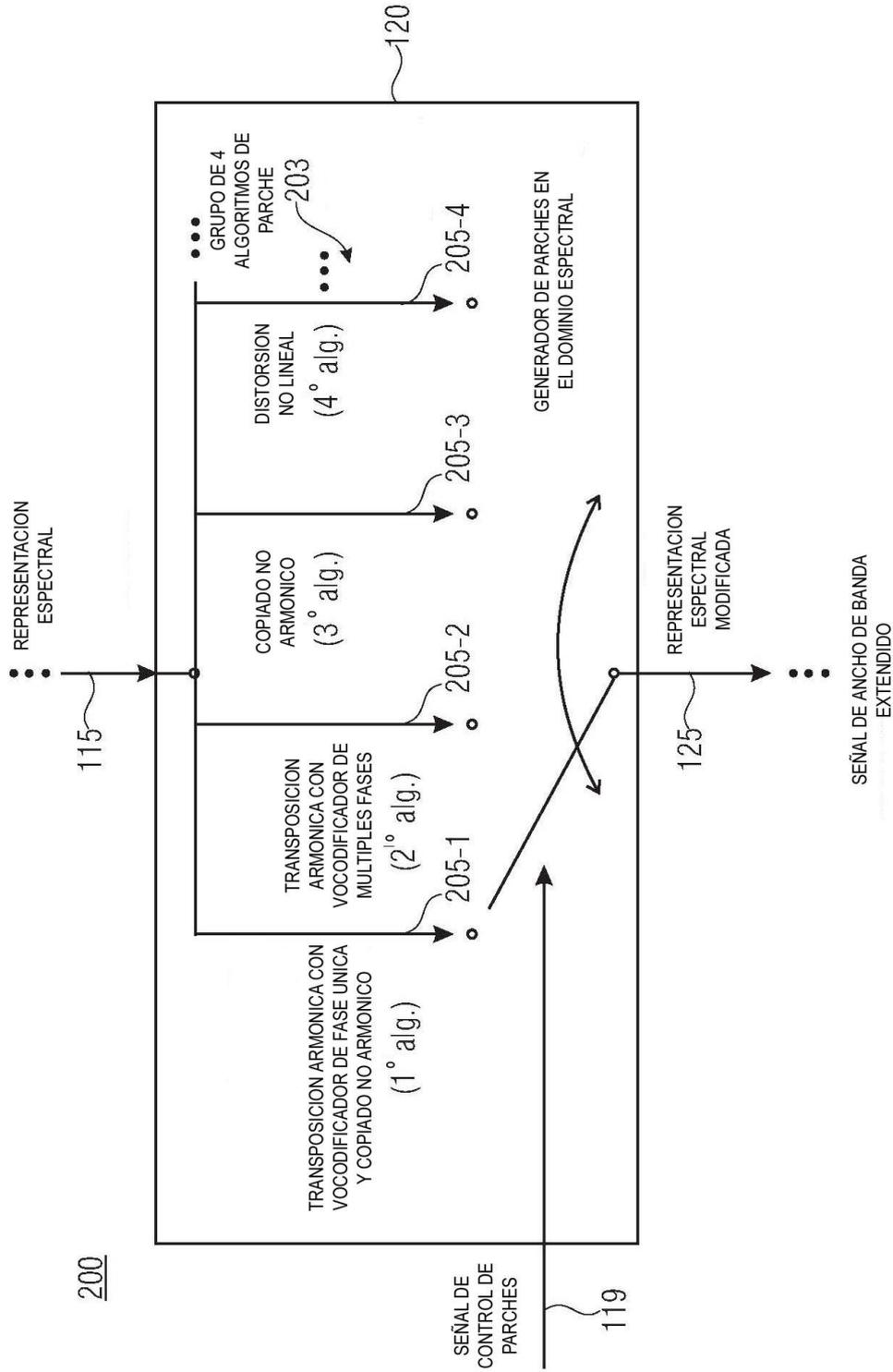


FIG. 2A

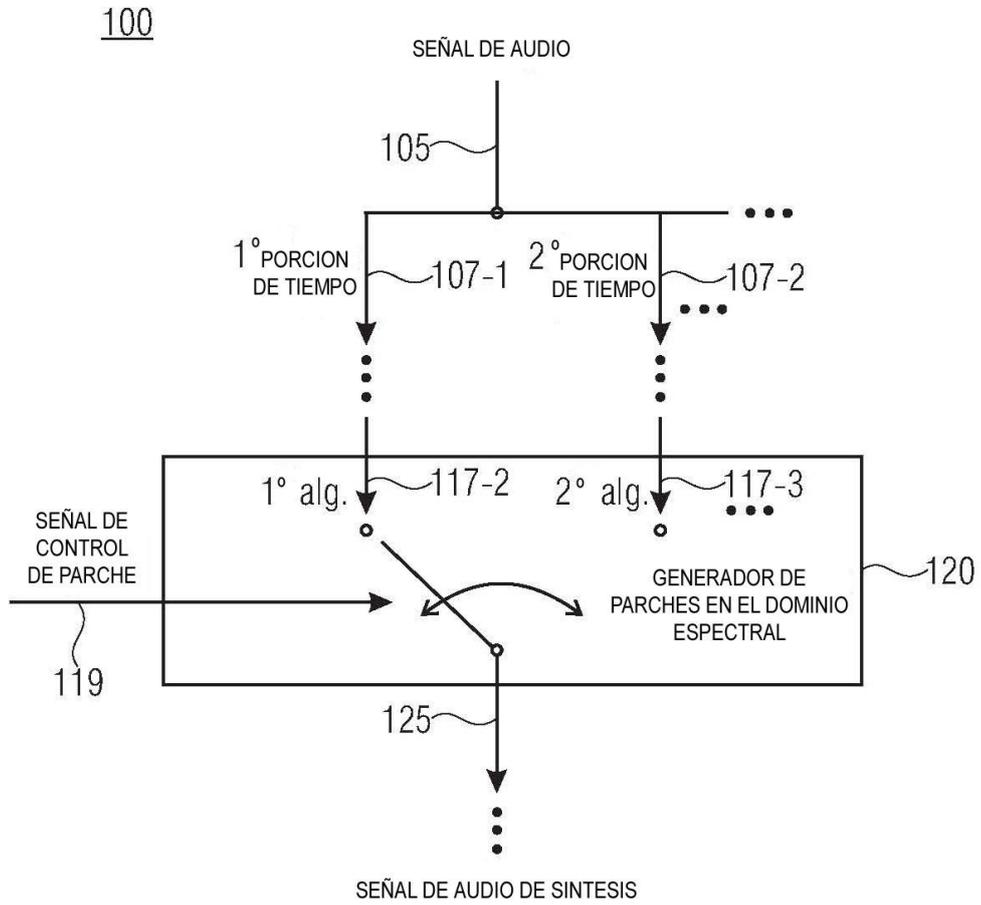


FIG. 1B
(GENERADOR DE PARCHES EN EL DOMINIO ESPECTRAL)

200

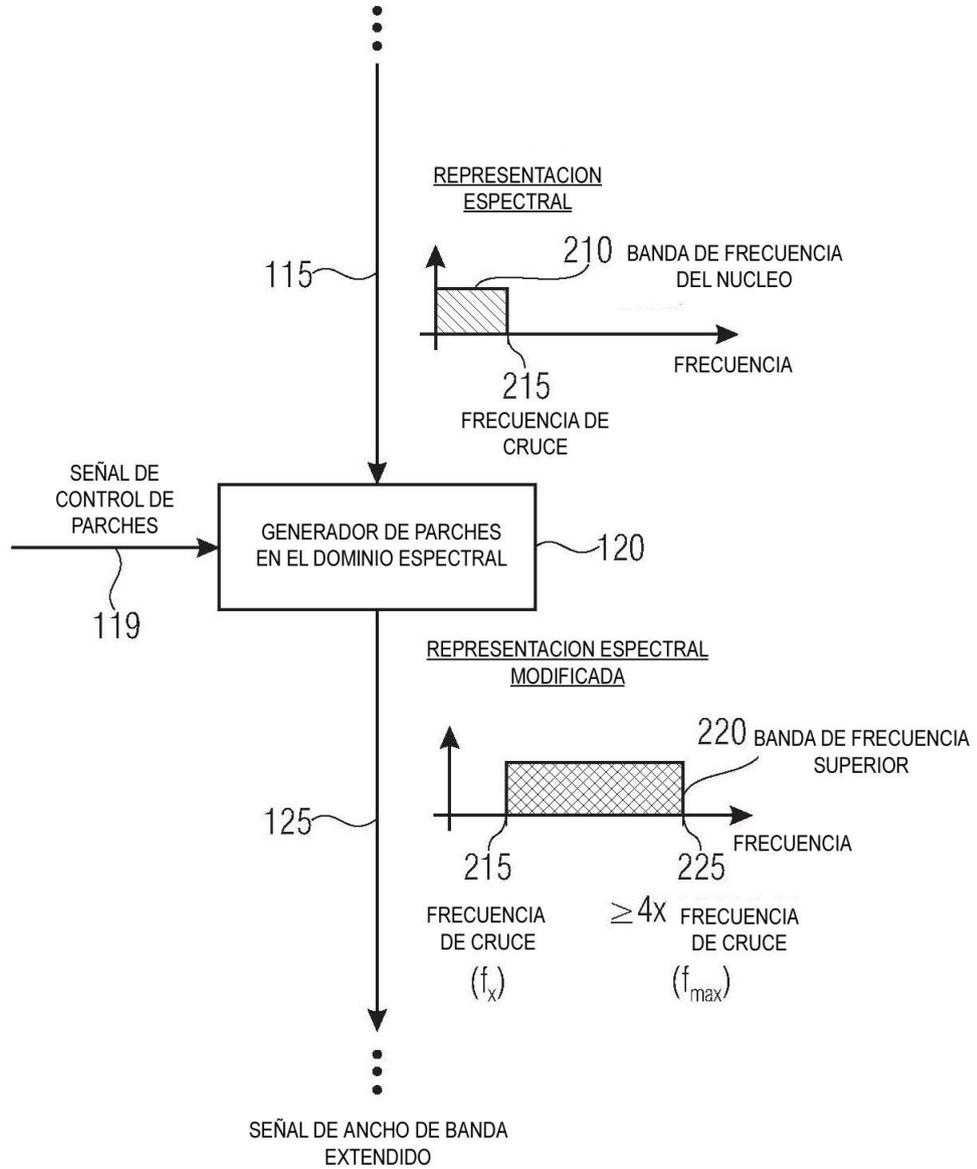


FIG. 2B
(ESQUEMA DE EXTENSION DE ANCHO DE BANDA)

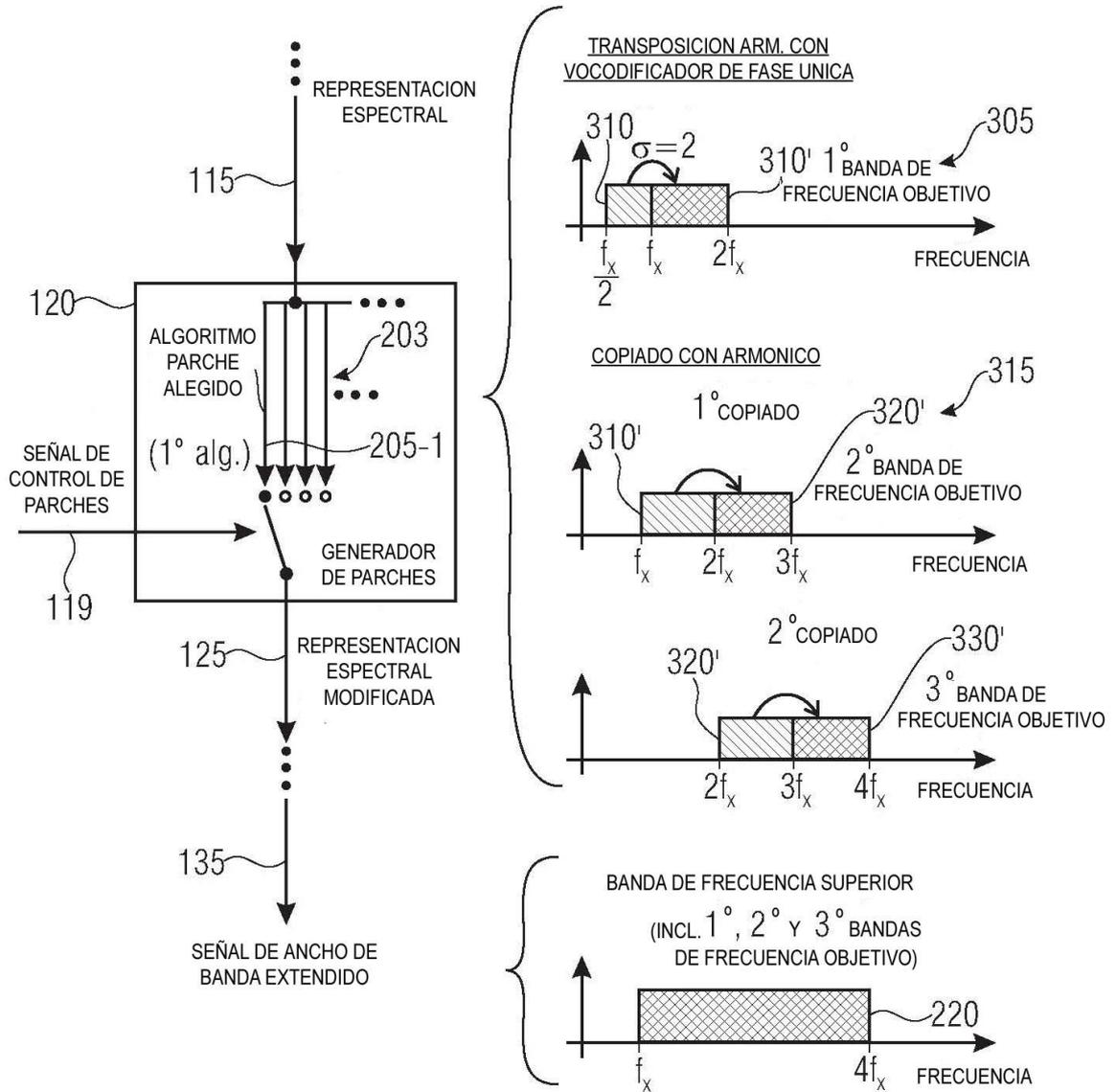


FIG. 3
(1° ALGORITMO DE PARCHES)

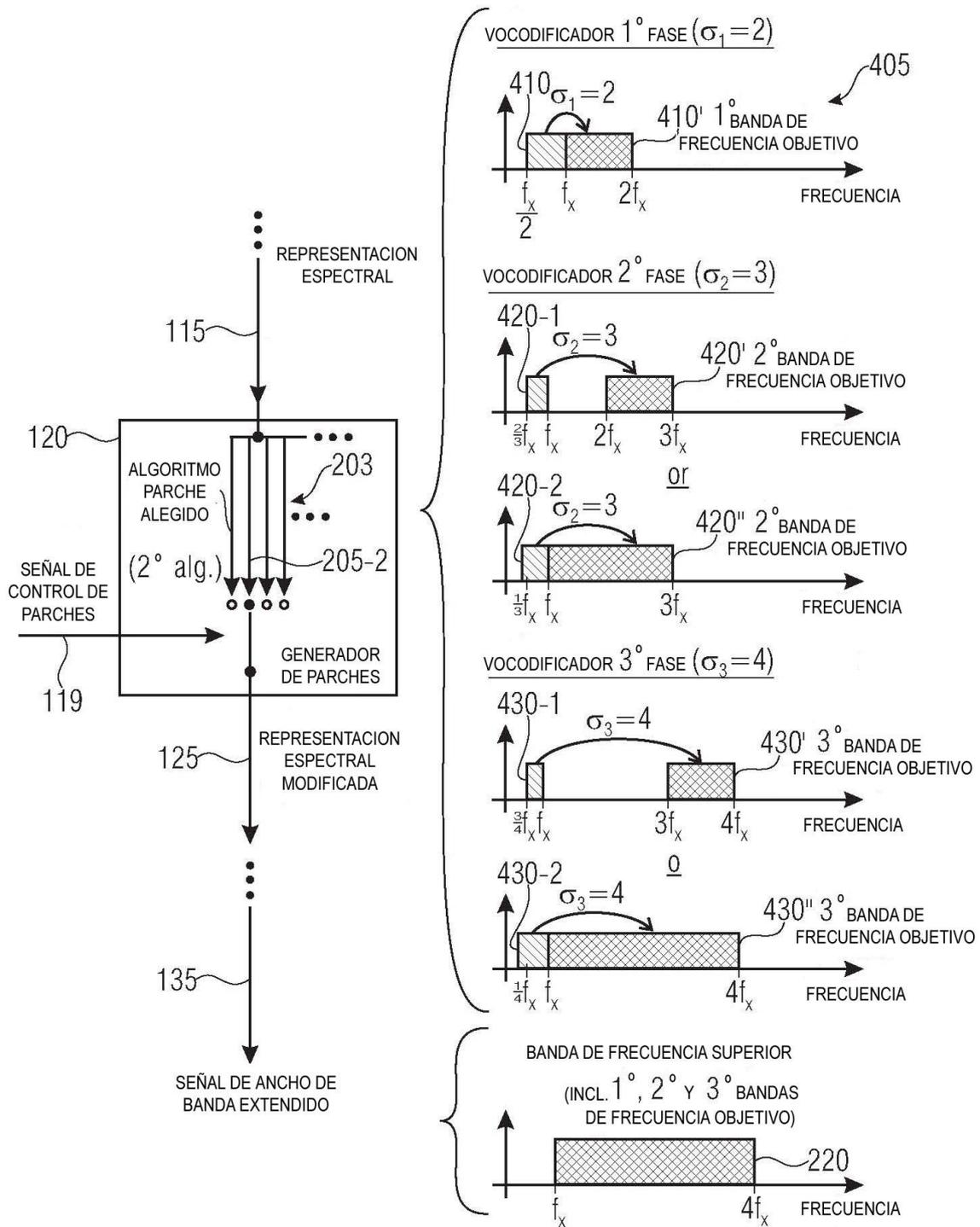


FIG. 4
(2° ALGORITMO DE PARCHES)

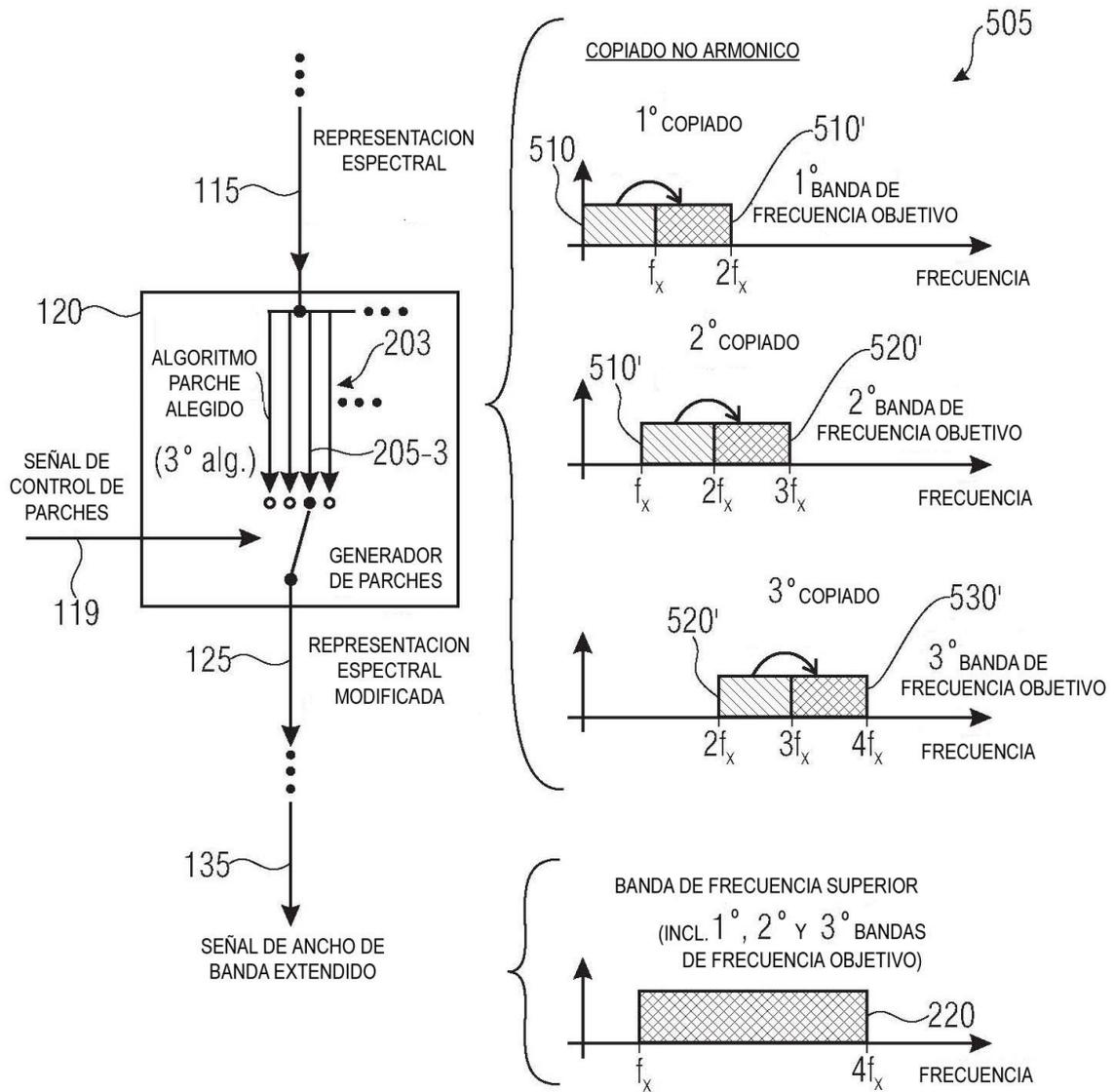


FIG. 5
(3° ALGORITMO DE PARCHES)

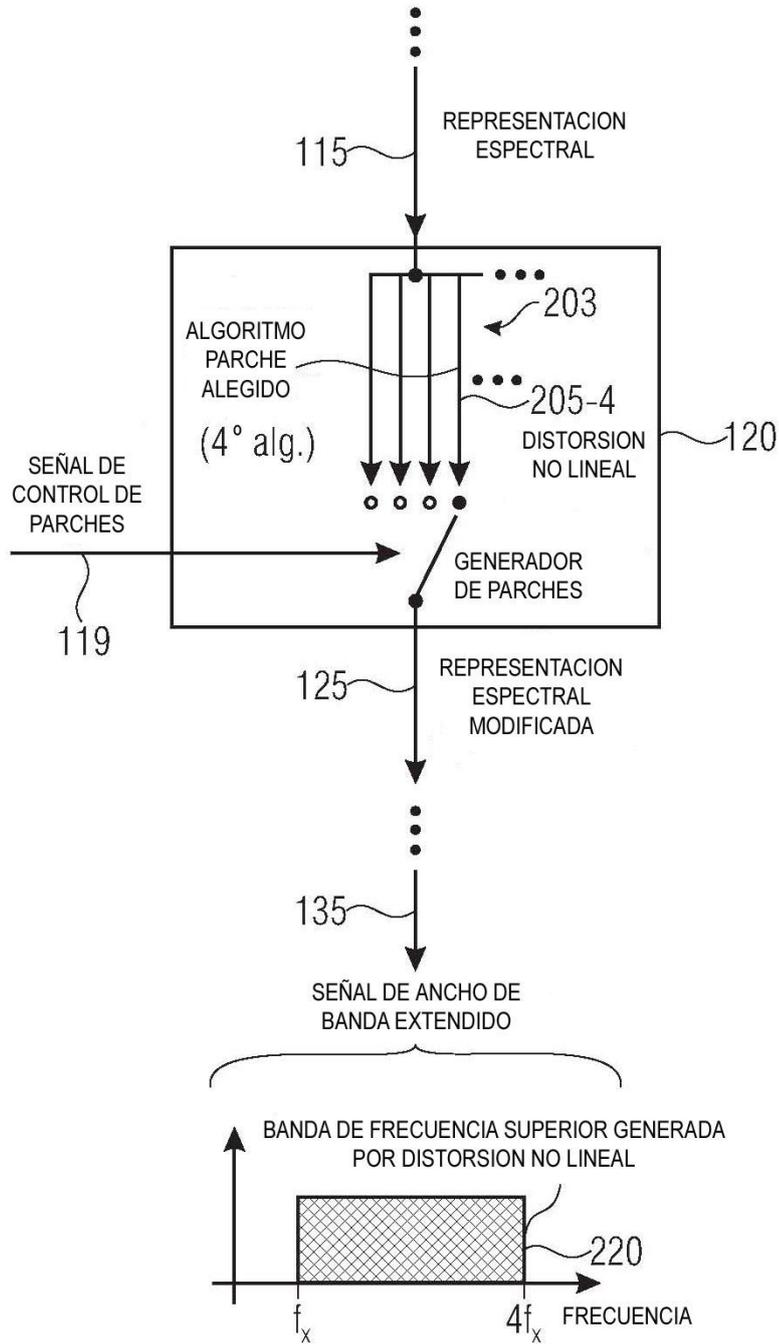


FIG. 6
(4° ALGORITMO DE PARCHES)

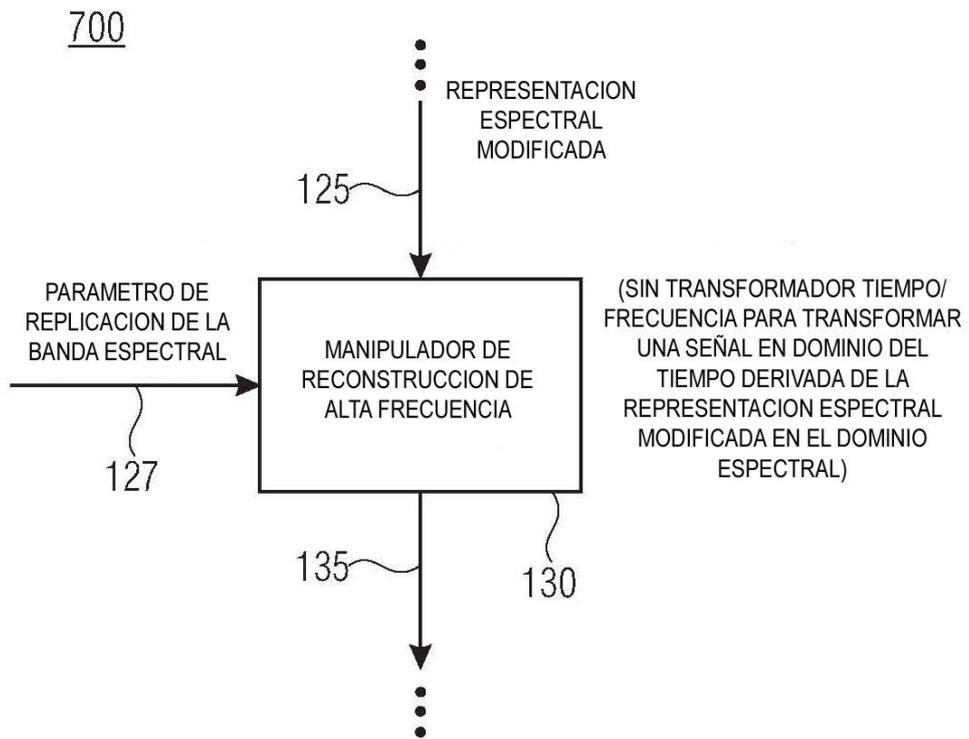


FIG. 7

800

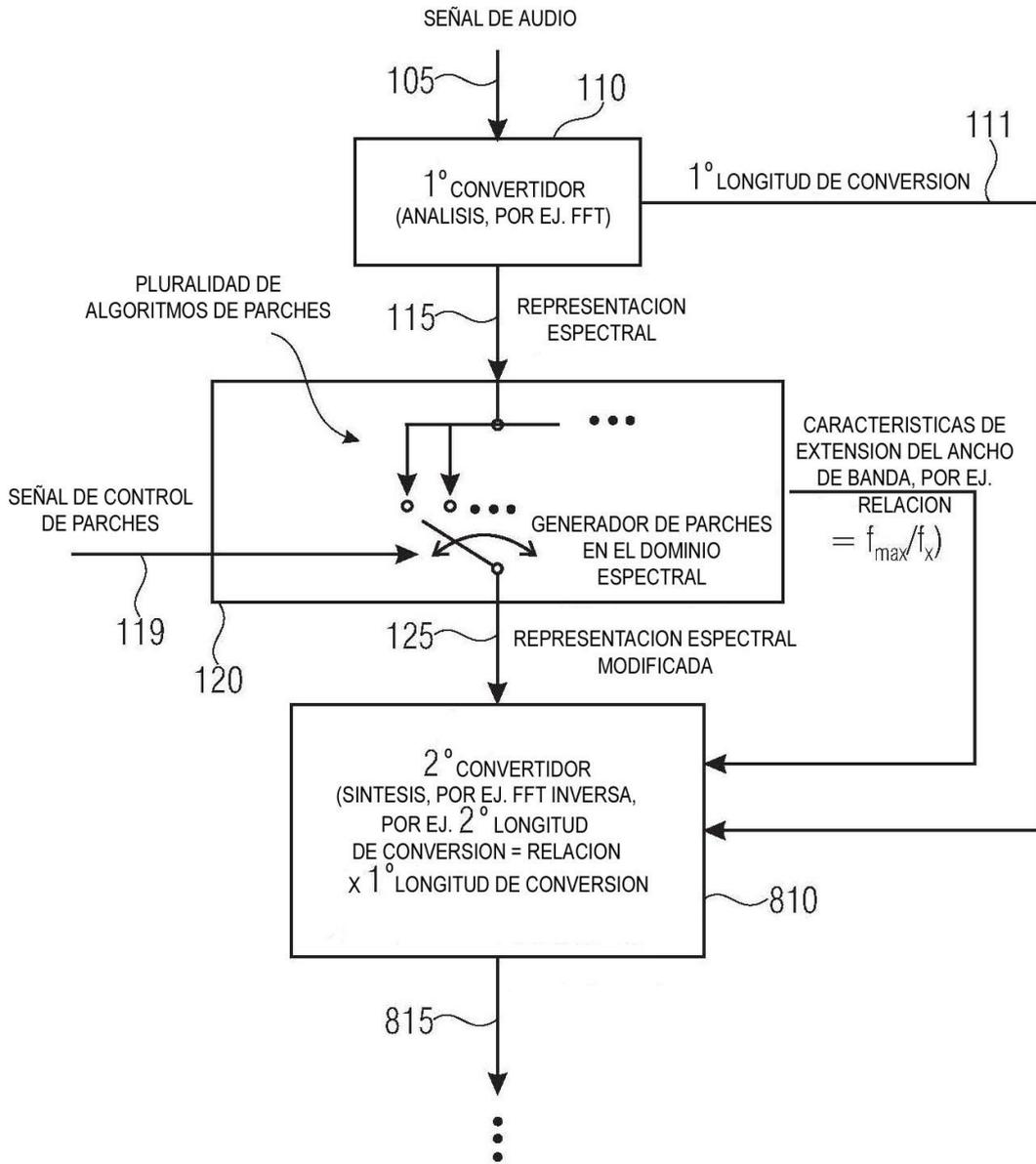


FIG. 8
(SINTETIZADOR DE SEÑALES DE AUDIO)

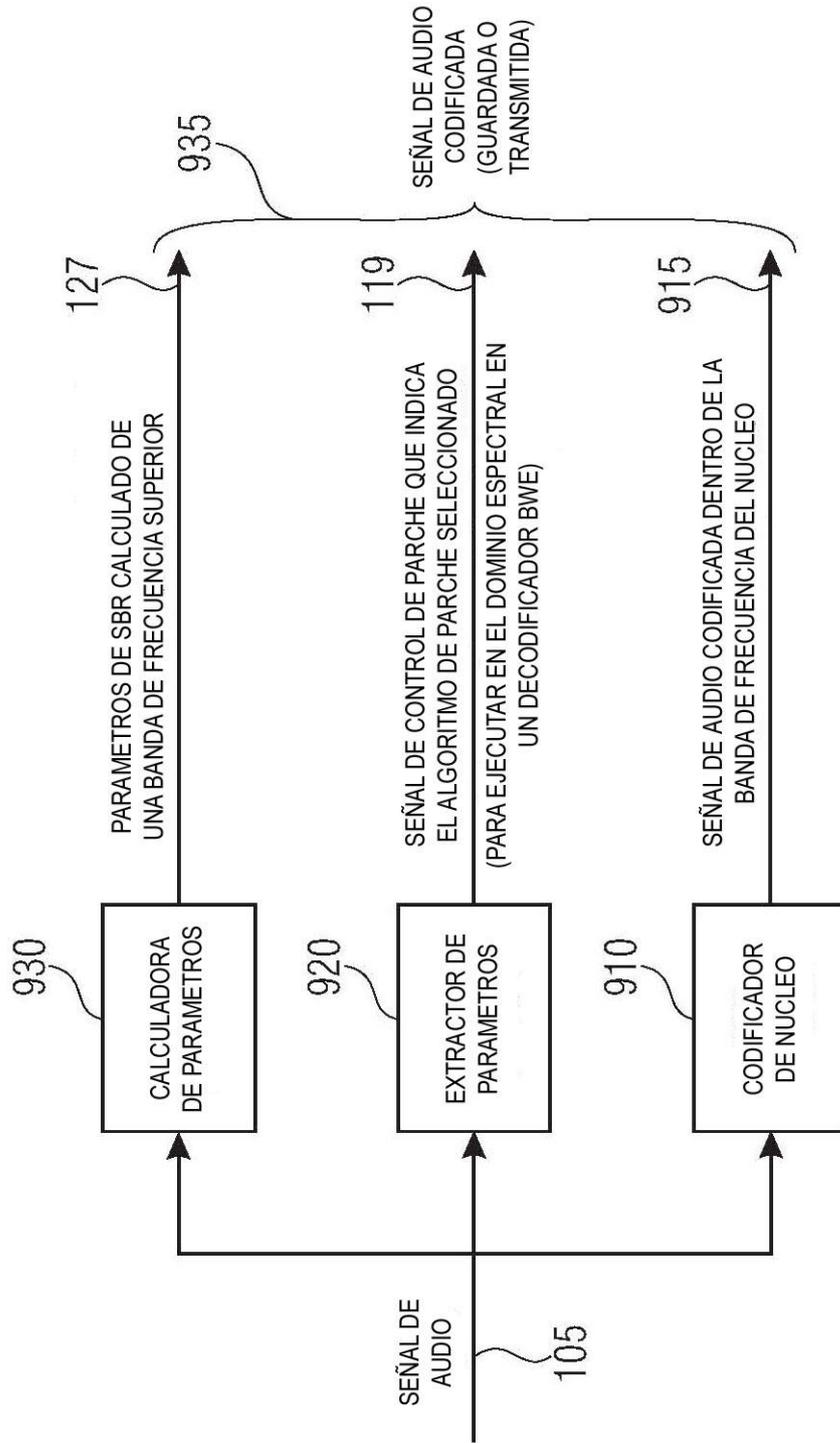


FIG. 9
(CODIFICADOR)

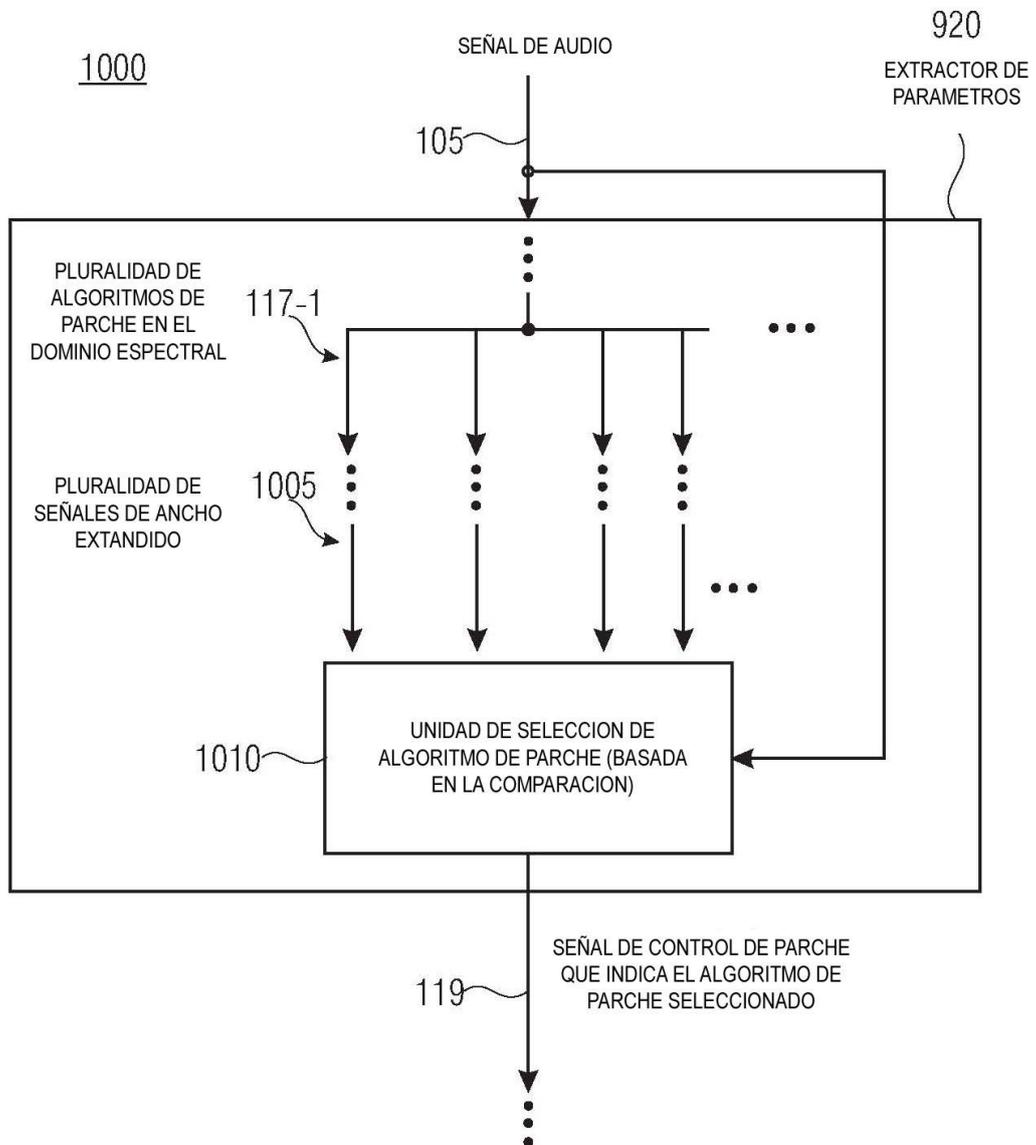


FIG. 10
(CODIFICADOR)

FIG. 11
(ESQUEMA DE INTRODUCCION DE PARCHES
EN EL DOMINIO DE LA FRECUENCIA)

