(19) OFICINA ESPAÑOLA DE PATENTES Y MARCAS

ESPAÑA

(11) Número de publicación: **2 653 799**

(51) Int. CI.:

G10L 19/08 (2013.01)
G10L 19/09 (2013.01)
G10L 25/90 (2013.01)

(12) TRADUCCIÓN DE PATENTE EUROPEA T3

(96) Fecha de presentación y número de la solicitud europea: **29.07.2011 E 15201425 (4)**

(97) Fecha y número de publicación de la concesión europea: **04.10.2017 EP 3021322**

(54) Título: **Sistemas, procedimientos, aparatos y medios legibles por ordenador para la decodificación de señales armónicas**

(30) Prioridad:

30.07.2010 US 201013192956
31.07.2010 US 369705 P
01.08.2010 US 369751 P
17.08.2010 US 374565 P
17.09.2010 US 384237 P
31.03.2011 US 201161470438 P
28.07.2011 US 201113192956

(45) Fecha de publicación y mención en BOPI de la traducción de la patente:
08.02.2018

(73) Titular/es:

QUALCOMM INCORPORATED (100.0%)
5775 Morehouse Drive
San Diego, CA 92121-1714, US

(72) Inventor/es:

RAJENDRAN, VIVEK;
DUNI, ROBERT ETHAN;
KRISHNAN, VENKATESH y
TAWARI, KUMAR ASHISH

(74) Agente/Representante:

FORTEA LAGUNA, Juan José

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

Sistemas, procedimientos, aparatos y medios legibles por ordenador para la decodificación de señales armónicas

ANTECEDENTES

Campo

Esta divulgación se refiere al campo del procesamiento de señales de audio.

Antecedentes

Normalmente se utilizan esquemas de codificación basados en la transformación de coseno discreta modificada (MDCT) para la codificación de señales de audio generalizadas, que pueden incluir contenido vocal y/o no vocal, como, por ejemplo, música. Ejemplos de códecs de audio existentes que utilizan la codificación MDCT incluyen la Capa 3 de Audio de MPEG-1 (MP3), Dolby Digital (Dolby Labs., Londres, Reino Unido; denominado también AC-3 y normalizado como ATSC A/52), Vorbis (Fundación Xiph.Org, Somerville, MA), Audio de Windows Media (WMA, Microsoft Corp., Redmond, WA), Codificación acústica de transformación adaptativa (Adaptative Transform Acoustic Coding) (ATRAC, Sony Corp., Tokio, JP) y Codificación de audio avanzada (Advanced Audio Coding) (AAC, como se ha normalizado recientemente en ISO/IEC 14496-3:2009). La codificación MDCT es también un componente de algunas normas de telecomunicaciones, tales como el Códec de velocidad variable mejorado (Enhanced Variable Rate Codec) (EVRC, como se ha normalizado en el documento del Proyecto 2 de Colaboración de Tercera Generación (3GPP2) C.S0014-D v2.0, 25 de enero de 2010). El códec G.718 ("Codificación de voz y audio a velocidad binaria variable entre 8 y 32 kbits/s, integrada, de banda estrecha y banda ancha y robusta en cuanto a errores en las tramas", Sector de Normalización de las Telecomunicaciones (UIT-T), Ginebra, CH, junio de 2008, corregido en noviembre 2008 y agosto de 2009, modificado en marzo de 2009 y marzo de 2010) es un ejemplo de un códec multicapa que utiliza la codificación MDCT.

En el documento US 2008/0162149 A1, titulado APARATO DE CODIFICACIÓN Y DECODIFICACIÓN DE AUDIO, Y PROCEDIMIENTO DEL MISMO, se describe un procedimiento de codificación de audio que incluye las etapas de: detectar al menos una onda sinusoidal a partir de una señal de audio de entrada; calcular componentes de vectores de base adicionales, basándose en señales de audio residuales y en los vectores de base adicionales de la onda sinusoidal; determinar la transmisión de componentes de los vectores de base adicionales; y al menos uno entre (a) codificar frecuencias y (b) fases y amplitudes de las ondas sinusoidales cuando está determinada la transmisión de los componentes de los vectores de base adicionales, en donde las señales de audio residuales se obtienen excluyendo las ondas sinusoidales detectadas de la señal de audio de entrada.

RESUMEN

De acuerdo a un primer aspecto, se proporciona un procedimiento de construcción de una trama de audio decodificada, utilizando una pluralidad de vectores de sub-banda decodificada, comprendiendo cada vector de sub-banda valores para una pluralidad de recipientes de frecuencia, comprendiendo dicho procedimiento: colocar, en un dominio de frecuencia, un primer vector de una pluralidad de vectores de sub-banda decodificada, de acuerdo a un valor de frecuencia fundamental; colocar, en el dominio de frecuencia, el resto de la pluralidad de vectores de sub-banda decodificada, de acuerdo al valor de frecuencia fundamental y a un valor de separación armónica; e insertar, en el dominio de frecuencia, una señal residual decodificada en ubicaciones de la trama que no estén ocupadas por la pluralidad de vectores de sub-banda decodificada.

De acuerdo con un aspecto adicional, se proporciona un medio de almacenamiento legible por ordenador (por ejemplo, medios no transitorios) con características tangibles que provocan que una máquina lea las características para realizar dicho procedimiento.

De acuerdo a un aspecto adicional, se proporciona un aparato de procesamiento de señales de audio para construir una trama de audio decodificado, utilizando una pluralidad de vectores de sub-banda decodificada, comprendiendo cada vector de sub-banda valores para una pluralidad de recipientes de frecuencia, comprendiendo el aparato: medios para colocar, en un dominio de frecuencia, un primer vector de una pluralidad de vectores de sub-banda decodificada, de acuerdo a un valor de frecuencia fundamental; medios para colocar, en el dominio de frecuencia, el resto de la pluralidad de vectores de sub-banda decodificada, de acuerdo al valor de frecuencia fundamental y a un valor de separación armónica; y medios para insertar, en el dominio de la frecuencia, una señal residual decodificada en ubicaciones de la trama que no estén ocupadas por la pluralidad de vectores de sub-banda decodificada.

BREVE DESCRIPCIÓN DE LOS DIBUJOS

La FIG. 1A muestra un diagrama de flujo para un procedimiento MA100 de procesamiento de una señal de audio de acuerdo con una configuración general.

La FIG. 1B muestra un diagrama de flujo para una implementación TA602 de la tarea TA600.

La FIG. 2A ilustra un ejemplo de una ventana de selección de picos.

La FIG. 2B muestra un ejemplo de una aplicación de la tarea T430.

La FIG. 3A muestra un diagrama de flujo de una implementación MA110 del procedimiento MA100.

La FIG. 3B muestra un diagrama de flujo de un procedimiento MD100 de decodificación de una señal codificada.

La FIG. 4 muestra un gráfico de un ejemplo de una señal armónica y conjuntos alternativos de sub-bandas seleccionadas.

La FIG. 5 muestra un diagrama de flujo de una implementación T402 de la tarea T400.

La FIG. 6 muestra un ejemplo de un conjunto de sub-bandas dispuesto de acuerdo con una implementación del procedimiento MA100.

La FIG. 7 muestra un ejemplo de un enfoque para compensar la falta de información de arritmia (fluctuación de fase).

La FIG. 8 muestra un ejemplo de expansión de una región de una señal residual.

La FIG. 9 muestra un ejemplo de codificación de una parte de una señal residual como un número de pulsos unitarios.

La FIG. 10A muestra un diagrama de flujo para un procedimiento MB100 de procesamiento de una señal de audio de acuerdo con una configuración general.

La FIG. 10B muestra un diagrama de flujo de una implementación MB110 del procedimiento MB100.

La FIG. 11 muestra un gráfico de magnitud frente a frecuencia para un ejemplo en el que la señal de audio de destino es una señal UB-MDCT.

La FIG. 12A muestra un diagrama de bloques de un aparato MF100 para el procesamiento de una señal de audio de acuerdo con una configuración general.

La FIG. 12B muestra un diagrama de bloques de un aparato A100 para el procesamiento de una señal de audio de acuerdo con una configuración general.

La FIG. 13A muestra un diagrama de bloques de una implementación MF110 del aparato MF100.

La FIG. 13B muestra un diagrama de bloques de una implementación A110 del aparato A100.

La FIG. 14 muestra un diagrama de bloques de un aparato MF210 para el procesamiento de una señal de audio de acuerdo con una configuración general.

Las FIGs. 15A y 15B ilustran ejemplos de aplicaciones del procedimiento MB110 para codificar señales de destino.

Las FIGs. 16A a 16E muestran una serie de aplicaciones para diversas implementaciones del aparato A110, MF110 o MF210.

La FIG. 17A muestra un diagrama de bloques de un procedimiento MC100 de clasificación de señales.

La FIG. 17B muestra un diagrama de bloques de un dispositivo de comunicaciones D10.

La FIG. 18 muestra vistas delantera, trasera y lateral de un equipo manual H100. La FIG. 19 muestra un ejemplo de una aplicación del procedimiento MA100.

DESCRIPCIÓN DETALLADA

Puede ser deseable identificar regiones de energía significativa dentro de una señal a codificar. La separación de dichas regiones del resto de la señal permite la codificación deseada de estas regiones para una mayor eficacia de codificación. Por ejemplo, puede ser deseable aumentar la eficacia de codificación utilizando relativamente más bits para codificar dichas regiones y relativamente menos bits (o incluso ningún bit) para codificar otras regiones de la

señal.

Para las señales de audio que tienen un alto contenido armónico (por ejemplo, señales de música, señales de habla vocalizada), las localizaciones de las regiones de energía significativa en el dominio de frecuencia pueden estar relacionadas. Puede ser deseable llevar a cabo una codificación eficaz en el dominio de la transformación de una señal de audio explotando dicha armonicidad.

Un esquema como el que se describe en el presente documento, para la codificación de un conjunto de coeficientes de transformación, que representan un intervalo de frecuencias de audio de una señal, explota la armonicidad en el espectro de la señal mediante el uso de un modelo armónico para parametrizar una relación entre las localizaciones de las regiones de energía significativa en el dominio de la frecuencia. Los parámetros de este modelo armónico pueden incluir la localización de la primera de estas regiones (por ejemplo, en orden creciente de frecuencia) y una separación entre regiones sucesivas. La estimación de los parámetros del modelo armónico puede incluir la generación de un banco de conjuntos de candidatos de valores de parámetros y la selección de un conjunto de valores de parámetros del modelo entre el banco generado. En una aplicación particular, se utiliza un esquema de este tipo para codificar los coeficientes de la transformación MDCT correspondientes al intervalo entre 0 y 4 kHz (en adelante denominada MDCT de banda baja, o LB-MDCT) de una señal de audio, tal como un residuo de una operación de codificación predictiva lineal.

La separación de las localizaciones de regiones de energía significativa de su contenido permite transmitir al decodificador una representación de una relación armónica entre las localizaciones de estas regiones, utilizando información complementaria mínima (por ejemplo, los valores de los parámetros del modelo armónico). Dicha eficacia puede ser especialmente importante para aplicaciones de baja velocidad de bits, tales como la telefonía móvil.

Salvo que esté expresamente limitado por su contexto, el término "señal" se utiliza en esta memoria para indicar cualquiera de sus significados comunes, incluyendo un estado de una posición de memoria (o un conjunto de posiciones de memoria) tal como se expresa en un cable, un bus u otro medio de transmisión. Salvo que esté expresamente limitado por su contexto, el término "generación" se utiliza en la presente memoria para indicar cualquiera de sus significados comunes, tales como el cálculo o, en otro sentido, la fabricación. Salvo que esté expresamente limitado por su contexto, el término "cálculo" se utiliza en la presente memoria para indicar cualquiera de sus significados comunes, tales como el cómputo, la evaluación, el allanamiento y/o la selección a partir de una pluralidad de valores. Salvo que esté expresamente limitado por su contexto, el término "obtención" se utiliza para indicar cualquiera de sus significados comunes, tales como el cálculo, la deducción, la recepción (por ejemplo, desde un dispositivo externo) y/o la recuperación (por ejemplo, a partir de una formación de elementos de almacenamiento). Salvo que esté expresamente limitado por su contexto, el término "selección" se utiliza para indicar cualquiera de sus significados comunes, tales como la identificación, la indicación, la aplicación y/o el uso de al menos uno, y menos de la totalidad, de un conjunto de dos o más. Allí donde el término "comprende" se usa en la presente descripción y reivindicaciones, no excluye otros elementos u operaciones. El término "basado en" (como en "A está basado en B") se utiliza para indicar cualquiera de sus significados comunes, incluyendo los casos (i) "deducido de" (por ejemplo, "B es un precursor de A"), (ii) "basado en al menos" (por ejemplo, "A está basado en al menos B") y, si corresponde en el contexto particular, (iii) "igual a" (por ejemplo, "A es igual a B"). De manera similar, el término "en respuesta a" se utiliza para indicar cualquiera de sus significados comunes, incluyendo "en respuesta a al menos".

Salvo que se indique lo contrario, el término "serie" se usa para indicar una secuencia de dos o más elementos. El término "logaritmo" se utiliza para indicar el logaritmo en base diez, aunque las extensiones de dicha operación a otras bases están dentro del alcance de esta divulgación. El término "componente de frecuencia" se utiliza para indicar una entre un conjunto de frecuencias o bandas de frecuencias de una señal, tal como una muestra de una representación en el dominio de la frecuencia de la señal (por ejemplo, como la generada por una transformación rápida de Fourier) o de una sub-banda de la señal (por ejemplo, una sub-banda de la escala de Bark o de la escala de Mel).

Salvo que se indique otra cosa, cualquier divulgación de una operación de un aparato que tenga una característica particular también pretende expresamente divulgar un procedimiento que tenga una característica análoga (y viceversa), y cualquier divulgación de una operación de un aparato de acuerdo con una configuración particular también pretende expresamente divulgar un procedimiento de acuerdo con una configuración análoga (y viceversa). El término "configuración" se puede utilizar en referencia a un procedimiento, aparato y/o sistema, según lo indicado por su contexto particular. Los términos "método", "proceso", "procedimiento" y "técnica" se utilizan de forma genérica e indistintamente, salvo que esté indicado lo contrario por el contexto particular. Los términos "aparato" y "dispositivo" también se utilizan genéricamente e indistintamente salvo que esté indicado lo contrario por el contexto particular. Los términos "elemento" y "módulo" se utilizan normalmente para indicar una parte de una configuración mayor. Salvo que esté expresamente limitado por su contexto, el término "sistema" se utiliza en la presente memoria para indicar cualquiera de sus significados comunes, incluyendo "un grupo de elementos que interactúan para servir a un objetivo común". También se entenderá que cualquier incorporación por referencia de una parte de un documento incorpora definiciones de términos o variables que están mencionados dentro de la parte, donde tales

definiciones aparecen en otra parte en el documento, así como figuras cualesquiera mencionadas en la parte incorporada.

Los sistemas, procedimientos y aparatos descritos en este documento son aplicables en general a representaciones de la codificación de señales de audio en un dominio de la frecuencia. Un ejemplo típico de dicha representación es una serie de coeficientes de transformación en un dominio de transformación. Los ejemplos de transformaciones adecuadas incluyen transformaciones ortogonales discretas, tales como las transformaciones sinusoidales unitarias. Los ejemplos de transformaciones sinusoidales unitarias adecuadas incluyen las transformaciones trigonométricas discretas, que incluyen sin limitación las transformaciones de coseno discretas (DCT), las transformaciones de seno discretas (DST) y la transformación discreta de Fourier (DFT). Otros ejemplos de transformaciones adecuadas incluyen versiones solapadas de dichas transformaciones. Un ejemplo particular de una transformación adecuada es la DCT modificada (MDCT) presentada anteriormente.

En toda esta divulgación se hace referencia a una "banda baja" y a una "banda alta" (de manera equivalente, "banda superior") de un intervalo de frecuencias de audio, y al ejemplo particular de una banda baja de cero a cuatro kilohercios (kHz) y una banda alta de 3,5 a siete kHz. Se hace notar expresamente que los principios expuestos en el presente documento no están limitados a este ejemplo particular de ninguna manera, salvo que se indique explícitamente dicho límite. Otros ejemplos (de nuevo, sin limitación) de intervalos de frecuencias para los que se contempla expresamente y se divulga por la presente la aplicación de estos principios de codificación, decodificación, asignación, cuantización y/u otro procesamiento, incluyen una banda baja que tiene un límite inferior en cualquiera entre 0, 25, 50, 100, 150 y 200 Hz, y un límite superior en cualquiera entre 3000, 3500, 4000 y 4500 Hz, y una banda alta que tiene un límite inferior en cualquiera entre 3000, 3500, 4000, 4500 y 5000 Hz, y un límite superior en cualquiera entre 6000, 6500, 7000, 7500, 8000, 8500 y 9000 Hz. La aplicación de dichos principios (de nuevo, sin limitación) a una banda alta que tiene un límite inferior en cualquiera entre 3000, 3500, 4000, 4500, 5000, 5500, 6000, 6500, 7000, 7500, 8000, 8500 y 9000 Hz, y un límite superior en cualquiera entre 10, 10,5, 11, 11,5, 12, 12,5, 13, 13,5, 14, 14,5, 15, 15,5 y 16 kHz, también está expresamente contemplada y se divulga por la presente. También se hace constar expresamente que, aunque una señal de banda alta típicamente se convertirá en una velocidad de muestreo más baja en una etapa anterior del proceso de codificación (por ejemplo, mediante el re-muestreo y/o el diezmado), sigue siendo una señal de banda alta y la información que transporta continúa representando el intervalo de frecuencias de audio de banda alta. Para un caso en el que la banda baja y la banda alta se solapen en frecuencia, puede ser deseable reducir a cero la parte de solapamiento de la banda baja, reducir a cero la parte de solapamiento de la banda alta, o realizar una atenuación cruzada sobre la parte de solapamiento desde la banda baja hasta la banda alta.

Un esquema de codificación tal como el descrito en el presente documento se puede aplicar para codificar cualquier señal de audio (por ejemplo, incluyendo la voz). Alternativamente, puede ser deseable utilizar un esquema de codificación de este tipo sólo para el audio no vocal (por ejemplo, la música). En tal caso, el esquema de codificación se puede utilizar con un esquema de clasificación que determine el tipo de contenido de cada trama de la señal de audio y seleccione un esquema de codificación adecuado.

Un esquema de codificación tal como el descrito en este documento se puede utilizar como un códec principal o como una capa o etapa en un códec multi-capa o multi-etapa. En un ejemplo de este tipo, dicho esquema de codificación se utiliza para codificar una parte del contenido en frecuencia de una señal de audio (por ejemplo, una banda baja o una banda alta), y se utiliza otro esquema de codificación para codificar otra parte del contenido en frecuencia de la señal. En otro ejemplo de este tipo, dicho esquema de codificación se utiliza para codificar un residuo (es decir, un error entre las señales original y codificada) de otra capa de codificación.

La FIG. 1A muestra un diagrama de flujo para un procedimiento MA100 de procesamiento de una señal de audio de acuerdo con una configuración general que incluye las tareas TA100, TA200, TA300, TA400, TA500 y TA600. El procedimiento MA100 puede estar configurado para procesar la señal de audio como una serie de segmentos (por ejemplo, realizando una instancia de cada una de las tareas TA100, TA200, TA300, TA400, TA500 y TA600 para cada segmento). Un segmento (o "trama") puede ser un bloque de coeficientes de transformación que corresponde a un segmento en el dominio del tiempo, con una longitud que habitualmente está en el intervalo entre aproximadamente cinco o diez milisegundos y aproximadamente cuarenta o cincuenta milisegundos. Los segmentos en el dominio del tiempo pueden estar solapados (por ejemplo, con un solapamiento con segmentos adyacentes de un 25% o 50%) o pueden no estar solapados.

En un codificador de audio puede ser deseable obtener tanto una alta calidad como un bajo retardo. Un codificador de audio puede utilizar un tamaño de trama grande para obtener una alta calidad, pero desafortunadamente un tamaño de trama grande normalmente provoca un retardo más largo. Las ventajas potenciales de un codificador de audio como el descrito en el presente documento incluyen una codificación de alta calidad con tamaños de trama cortos (por ejemplo, un tamaño de trama de veinte milisegundos, con una anticipación de diez milisegundos). En un ejemplo particular, la señal en el dominio del tiempo se divide en una serie de segmentos no solapados de veinte milisegundos, y la MDCT para cada trama se toma sobre una ventana de cuarenta milisegundos que se solapa con cada una de las tramas adyacentes en diez milisegundos.

Un segmento como el procesado por el procedimiento MA100 también puede ser una parte (por ejemplo, una banda baja o una banda alta) de un bloque como el producido por la transformación, o una parte de un bloque como el producido por una operación anterior sobre un bloque de este tipo. En un ejemplo particular, cada uno de una serie de segmentos procesados por el procedimiento MA100 contiene un conjunto de 160 coeficientes de la MDCT que representan un intervalo de frecuencias de la banda baja de 0 a 4 kHz. En otro ejemplo particular, cada uno de una serie de segmentos procesados por el procedimiento MA100 contiene un conjunto de 140 coeficientes de la MDCT que representan un intervalo de frecuencias de la banda alta de 3,5 a 7 kHz.

La tarea TA100 localiza una pluralidad de picos en la señal de audio en un dominio de la frecuencia. Dicha operación también se puede denominar como "detección de picos". La tarea TA100 puede estar configurada para seleccionar un número particular de los picos más altos a partir de todo el intervalo de frecuencias de la señal. Alternativamente, la tarea TA100 puede estar configurada para seleccionar picos de un intervalo de frecuencias especificado de la señal (por ejemplo, un intervalo de baja frecuencia) o puede estar configurada para aplicar diferentes criterios de selección en diferentes intervalos de frecuencias de la señal. En un ejemplo particular tal como el descrito en el presente documento, la tarea TA100 está configurada para localizar al menos un primer número (Nd+1) de los picos más altos en la trama, incluyendo al menos un segundo número Nf de los picos más altos en un intervalo de baja frecuencia de la trama.

La tarea TA100 puede estar configurada para identificar un pico como una muestra de la señal en el dominio de la frecuencia (también denominada un "recipiente") que tiene el valor máximo dentro de una cierta distancia mínima a cada lado de la muestra. En un ejemplo de este tipo, la tarea TA100 está configurada para identificar un pico como la muestra que tiene el valor máximo dentro de una ventana de tamaño ($2d_{min}+1$) que está centrada en la muestra, donde d_{min} es una separación mínima permitida entre picos. El valor de d_{min} se puede seleccionar de acuerdo con un número máximo deseado de regiones de energía significativa (también denominadas "sub-bandas") que se van a localizar. Los ejemplos de d_{min} incluyen ocho, nueve, diez, doce y quince muestras (alternativamente, 100, 125, 150, 175, 200 o 250 Hz), aunque se puede utilizar cualquier valor adecuado para la aplicación deseada. La FIG. 2A ilustra un ejemplo de una ventana de selección de picos de tamaño ($2d_{min}+1$), centrada en una localización potencial de un pico de la señal, para un caso en que el valor de d_{min} es ocho.

Basándose en las localizaciones en el dominio de frecuencia de al menos algunos (por ejemplo, al menos tres) de los picos localizados por la tarea TA100, la tarea TA200 calcula un número Nd de candidatos de separación armónica (también denominados candidatos de "distancia", o candidatos d). Los ejemplos de valores de Nd incluyen cinco, seis y siete. La tarea TA200 puede estar configurada para calcular estos candidatos de separación como las distancias (por ejemplo, en términos del número de recipientes de frecuencia) entre picos adyacentes de los (Nd+1) picos más altos localizados por la tarea TA100.

Basándose en las localizaciones en el dominio de la frecuencia de al menos algunos (por ejemplo, al menos dos) de los picos localizados por la tarea TA100, la tarea TA300 identifica un número Nf de candidatos para la localización de la primera sub-banda (también denominados candidatos de "frecuencia fundamental" o de F0). Los ejemplos de valores para Nf incluyen cinco, seis y siete. La tarea TA300 puede estar configurada para identificar estos candidatos como las localizaciones de los Nf picos más altos en la señal. Alternativamente, la tarea TA300 puede estar configurada para identificar estos candidatos como las localizaciones de los Nf picos más altos en una parte de baja frecuencia (por ejemplo, el 30, 35, 40, 45 o 50 por ciento inferior) del intervalo de frecuencias que se está examinando. En un ejemplo de este tipo, la tarea TA300 identifica el número Nf de candidatos de F0 entre las localizaciones de los picos localizados por la tarea TA100 en el intervalo de 0 a 1.250 Hz. En otro ejemplo de este tipo, la tarea TA300 identifica el número Nf de candidatos de F0 entre las localizaciones de los picos localizados por la tarea TA100 en el intervalo de 0 a 1.600 Hz.

Se hace notar expresamente que el alcance de las implementaciones descritas del procedimiento MA100 incluye el caso en el que sólo se calcula un candidato de separación armónica (por ejemplo, como la distancia entre los dos picos más altos, o la distancia entre los dos picos más altos en un intervalo de frecuencias especificado) y el caso independiente en el que sólo se identifica un candidato de F0 (por ejemplo, como la localización del pico más alto, o la localización del pico más alto en un intervalo de frecuencias especificado).

Para cada uno entre una pluralidad de pares activos de los candidatos de F0 y d, la tarea TA400 selecciona un conjunto de al menos una sub-banda de la señal de audio, en el que una localización en el dominio de la frecuencia de cada sub-banda en el conjunto está basada en el par (F0, d). En un ejemplo, la tarea TA400 está configurada para seleccionar las sub-bandas de cada conjunto de tal manera que la primera sub-banda esté centrada en la localización de F0 correspondiente, estando el centro de cada sub-banda siguiente separado del centro de la sub-banda anterior por una distancia igual al valor correspondiente de d.

La tarea TA400 puede estar configurada para seleccionar cada conjunto para incluir todas las sub-bandas indicadas por el par (F0, d) correspondiente que se encuentren dentro del intervalo de entrada. Alternativamente, la tarea TA400 puede estar configurada para seleccionar menos que todas estas sub-bandas para al menos uno de los conjuntos. Por ejemplo, la tarea TA400 puede estar configurada para seleccionar no más de un número máximo de sub-bandas para el conjunto. Alternativamente, o adicionalmente, la tarea TA400 puede estar configurada para

seleccionar sólo las sub-bandas que se encuentren dentro de un intervalo particular. Las sub-bandas a frecuencias más bajas tienden a ser más importantes perceptivamente, por ejemplo, de tal manera que puede ser deseable configurar la tarea TA400 para seleccionar no más de un número particular de una o más (por ejemplo, cuatro, cinco o seis) de las sub-bandas de frecuencia más baja en el intervalo de entrada y/o sólo las sub-bandas cuyas localizaciones no estén por encima de una frecuencia particular dentro del intervalo de entrada (por ejemplo, 1.000, 1.500 o 2.000 Hz).

La tarea TA400 puede estar implementada para seleccionar sub-bandas de longitud fija e igual. En un ejemplo particular, cada sub-banda tiene una anchura de siete recipientes de frecuencia (por ejemplo, 175 Hz, para una separación de recipientes de veinticinco Hz). Sin embargo, se contempla expresamente y por la presente se divulga que los principios descritos en el presente documento también se pueden aplicar a casos en los que las longitudes de las sub-bandas pueden variar de una trama a otra, y/o en los que las longitudes de dos o más de (posiblemente todas) las sub-bandas dentro de una trama pueden ser diferentes.

En un ejemplo, se considera que todos los pares diferentes de valores de F0 y d son activos, de manera que la tarea TA400 está configurada para seleccionar un conjunto correspondiente de una o más sub-bandas para cada posible par (F0, d). Por ejemplo, para un caso en el que Nf y Nd son ambos iguales a siete, la tarea TA400 puede estar configurada para considerar cada uno de los cuarenta y nueve pares posibles. Para un caso en el que Nf es igual a cinco y Nd es igual a seis, la tarea TA400 puede estar configurada para considerar cada uno de los treinta pares posibles. Alternativamente, la tarea TA400 puede estar configurada para imponer un criterio de actividad que algunos de los posibles pares (F0, d) pueden no cumplir. En tal caso, por ejemplo, la tarea TA400 puede estar configurada para ignorar los pares que producirían más de un número máximo admisible de sub-bandas (por ejemplo, combinaciones de valores bajos de F0 y d) y/o pares que producirían menos de un número mínimo deseado de sub-bandas (por ejemplo, combinaciones de valores altos de F0 y d).

Para cada uno entre una pluralidad de pares de candidatos de F0 y d, la tarea TA500 calcula al menos un valor de energía a partir del conjunto correspondiente de una o más sub-bandas de la señal de audio. En un ejemplo de este tipo, la tarea TA500 calcula un valor de energía de cada conjunto de una o más sub-bandas como la energía total del conjunto de sub-bandas (por ejemplo, como una suma de las magnitudes al cuadrado de los valores de muestreo en el dominio de la frecuencia en las sub-bandas). Alternativamente, o adicionalmente, la tarea TA500 puede estar configurada para calcular valores de energía a partir de cada conjunto de sub-bandas como las energías de cada sub-banda individual y/o para calcular un valor de energía a partir de cada conjunto de sub-bandas como una energía media por sub-banda (por ejemplo, la energía total normalizada sobre el número de sub-bandas) para el conjunto de sub-bandas. La tarea TA500 puede estar configurada para ejecutarse para cada uno entre la misma pluralidad de pares que la tarea TA400, o para menos que esta pluralidad. Por ejemplo, para un caso en el que la tarea TA400 está configurada para seleccionar un conjunto de sub-bandas para cada posible par (F0, d), la tarea TA500 puede estar configurada para calcular valores de energía sólo para los pares que satisfagan un criterio especificado de actividad (por ejemplo, ignorar los pares que producirían demasiadas sub-bandas y/o los pares que producirían demasiado pocas sub-bandas, tal como se ha descrito anteriormente). En otro ejemplo, la tarea TA400 está configurada para ignorar los pares que producirían demasiadas sub-bandas y la tarea TA500 está configurada para ignorar también los pares que producirían demasiado pocas sub-bandas.

Aunque la FIG. 1A muestra la ejecución de las tareas TA400 y TA500 en serie, se entenderá que la tarea TA500 también puede estar implementada para comenzar a calcular energías para conjuntos de sub-bandas antes de que haya finalizado la tarea TA400. Por ejemplo, la tarea TA500 puede estar implementada para comenzar a calcular (o incluso para terminar de calcular) un valor de energía a partir de un conjunto de sub-bandas antes de que la tarea TA400 comience a seleccionar el siguiente conjunto de sub-bandas. En un ejemplo de este tipo, las tareas TA400 y TA500 están configuradas para alternarse para cada uno entre la pluralidad de pares activos de candidatos de F0 y d. Del mismo modo, la tarea TA400 también puede estar implementada para comenzar la ejecución antes de que hayan finalizado las tareas TA200 y TA300.

Basándose en los valores de energía calculados a partir de al menos algunos de los conjuntos de una o más sub-bandas, la tarea TA600 selecciona un par de candidatos entre los pares de candidatos (F0, d). En un ejemplo, la tarea TA600 selecciona el par correspondiente al conjunto de sub-bandas que tienen la energía total más alta. En otro ejemplo, la tarea TA600 selecciona el par de candidatos correspondiente al conjunto de sub-bandas que tienen la energía media por sub-banda más alta.

La FIG. 1B muestra un diagrama de flujo para una implementación adicional TA602 de la tarea TA600. La tarea TA620 incluye una tarea TA610 que ordena la pluralidad de pares de candidatos activos según la energía media por sub-banda de los conjuntos de sub-bandas correspondientes (por ejemplo, en orden descendente). Esta operación ayuda a inhibir la selección de pares de candidatos que produzcan conjuntos de sub-bandas que tengan una energía total alta, pero en los que una o más sub-bandas puedan tener demasiado poca energía como para ser significativas perceptivamente. Dicha condición puede indicar un número excesivo de sub-bandas.

La tarea TA602 también incluye una tarea TA620 que selecciona, entre los Pv pares de candidatos que producen los conjuntos de sub-bandas que tienen las energías medias por sub-banda más altas, el par de candidatos asociado al

conjunto de sub-bandas que captura la mayor parte de la energía total. Esta operación ayuda a inhibir la selección de pares de candidatos que produzcan conjuntos de sub-bandas que tengan una alta energía media por sub-banda, pero demasiado pocas sub-bandas. Dicha condición puede indicar que el conjunto de sub-bandas no logra incluir las regiones de la señal que tengan una energía más baja, pero que no obstante pueden ser significativas perceptivamente.

La tarea TA620 puede estar configurada para utilizar un valor fijo para Pv, tal como cuatro, cinco, seis, siete, ocho, nueve o diez. Alternativamente, la tarea TA620 tarea puede estar configurada para usar un valor de Pv que esté relacionado con el número total de pares de candidatos activos (por ejemplo, igual a, o no más de, el diez, veinte o veinticinco por ciento del número total de pares de candidatos activos).

Los valores seleccionados de F0 y d comprenden información complementaria del modelo, que son valores enteros y se pueden transmitir al decodificador utilizando un número finito de bits. La FIG. 3 muestra un diagrama de flujo de una implementación MA110 del procedimiento MA100 que incluye una tarea TA700. La tarea TA700 produce una señal codificada que incluye indicaciones de los valores del par de candidatos seleccionado. La tarea TA700 puede estar configurada para codificar el valor seleccionado de F0, o para codificar un desplazamiento del valor seleccionado de F0 a partir de una localización mínima (o máxima). De manera similar, la tarea TA700 puede estar configurada para codificar el valor seleccionado de d, o para codificar un desplazamiento del valor seleccionado de d a partir de una distancia mínima o máxima. En un ejemplo particular, la tarea TA700 utiliza seis bits para codificar el valor de F0 seleccionado y seis bits para codificar el valor de d seleccionado. En ejemplos adicionales, la tarea TA700 puede estar implementada para codificar el valor actual de F0 y/o d de manera diferencial (por ejemplo, como un desplazamiento relativo a un valor anterior del parámetro).

Puede ser deseable implementar la tarea TA700 para utilizar un esquema de codificación de cuantización vectorial (VQ) que codifique como vectores los contenidos de las regiones de energía significativa identificadas por el par de candidatos seleccionado (es decir, los valores dentro de cada uno de los conjuntos de sub-bandas seleccionados). Un esquema de VQ codifica un vector correlacionándolo con una entrada en cada uno de uno o más libros de códigos (que también son conocidos para el decodificador) y utilizando el índice o índices de estas entradas para representar el vector. La longitud de un índice del libro de códigos, que determina el número máximo de entradas en el libro de códigos, puede ser cualquier número entero arbitrario que se considere adecuado para la aplicación.

Un ejemplo de un esquema de VQ adecuado es la VQ de ganancia-forma (GSVQ), en la que el contenido de cada sub-banda se descompone en un vector de formas normalizado (que describe, por ejemplo, la forma de la sub-banda a lo largo del eje de frecuencias) y un factor de ganancia correspondiente, de manera que el vector de formas y el factor de ganancia se cuanticen por separado. El número de bits asignados para la codificación de los vectores de formas puede estar distribuido de manera uniforme entre los vectores de formas de las distintas sub-bandas. Alternativamente, puede ser deseable asignar más de los bits disponibles para la codificación de los vectores de formas que capturan más energía que otros, como los vectores de formas cuyos factores de ganancia correspondientes tienen valores relativamente altos en comparación con los factores de ganancia de los vectores de formas de otras sub-bandas.

Puede ser deseable utilizar un esquema de GSVQ que incluya una codificación predictiva de la ganancia, de tal manera que los factores de ganancia para cada conjunto de sub-bandas se codifiquen de forma independiente entre sí y de manera diferencial con respecto al factor de ganancia correspondiente de la trama anterior. En un ejemplo particular, el procedimiento MA110 está organizado para codificar las regiones de energía significativa en un intervalo de frecuencias de un espectro de la LB-MDCT.

La FIG. 3B muestra un diagrama de flujo de un procedimiento correspondiente MD100 de decodificación de una señal codificada (por ejemplo, como la producida por la tarea TA700) que incluye las tareas TD100, TD200 y TD300. La tarea TD100 decodifica los valores de F0 y d a partir de la señal codificada, y la tarea TD200 descuantiza el conjunto de sub-bandas. La tarea TD300 construye la señal decodificada mediante la disposición de cada sub-banda descuantizada en el dominio de la frecuencia, basándose en los valores decodificados de F0 y d. Por ejemplo, la tarea TD300 puede estar implementada para construir la señal decodificada centrando cada sub-banda m en la localización, en el dominio de la frecuencia, F0 + md, donde 0 <= m < M y M es el número de sub-bandas en el conjunto seleccionado. La tarea TD300 puede estar configurada para asignar valores cero a recipientes no ocupados de la señal decodificada o, alternativamente, para asignar valores de un residuo decodificado, como el descrito en este documento, a recipientes no ocupados de la señal decodificada.

En una modalidad de codificación armónica, la disposición de las regiones en las localizaciones adecuadas puede ser crítica para una codificación eficaz. Puede ser deseable configurar el esquema de codificación para capturar la mayor cantidad de la energía en el intervalo de frecuencias dado utilizando el mínimo número de sub-bandas.

La FIG. 4 muestra un gráfico del valor absoluto de los coeficientes de la transformación frente al índice del recipiente de frecuencia para un ejemplo de una señal armónica en el dominio de la MDCT. La FIG. 4 también muestra las localizaciones en el dominio de la frecuencia para dos posibles conjuntos de sub-bandas para esta señal. Las localizaciones del primer conjunto de sub-bandas se muestran mediante los bloques separados uniformemente, que

se dibujan en gris y también se indican mediante los corchetes bajo el eje x. Este conjunto se corresponde con el par de candidatos (F0, d), según la selección del procedimiento MA100. En este ejemplo se puede ver que, aunque las localizaciones de los picos en la señal parecen regulares, no se ajustan exactamente a la separación uniforme de las sub-bandas del modelo armónico. De hecho, en este caso el modelo casi obvia el pico más alto de la señal. En consecuencia, se puede esperar que un modelo que está configurado estrictamente, incluso de acuerdo al mejor par de candidatos (F0, d), puede no lograr capturar algo de la energía en uno o más de los picos de señal.

Puede ser deseable implementar el procedimiento MA100 para incorporar las faltas de uniformidad en la señal de audio mediante la relajación del modelo armónico. Por ejemplo, puede ser deseable permitir que una o más de las sub-bandas relacionadas armónicamente de un conjunto (es decir, las sub-bandas localizadas en F0, F0+d, F0+2d, etc.) se desplacen un número finito de recipientes en cada dirección. En tal caso, puede ser deseable implementar la tarea TA400 para permitir que la localización de una o más de las sub-bandas se desvíe en una pequeña magnitud (también llamada desplazamiento o "arritmia") de la localización indicada por el par (F0, d). El valor de dicho desplazamiento se puede seleccionar de tal modo que la sub-banda resultante capture más de la energía del pico.

Los ejemplos para la magnitud de arritmia permitida para una sub-banda incluyen el veinticinco, el treinta, el cuarenta y el cincuenta por ciento de la anchura de sub-banda. La magnitud de arritmia permitida en cada dirección del eje de frecuencia no necesariamente tiene que ser igual. En un ejemplo particular, se permite que cada sub-banda de siete recipientes desplace su posición inicial a lo largo del eje de frecuencia, como se indica mediante el par actual de candidatos (F0, d), hasta cuatro recipientes de frecuencia por encima o hasta tres recipientes de frecuencia por debajo. En este ejemplo, el valor de arritmia seleccionado para la sub-banda se puede expresar en tres bits. También es posible que el intervalo de valores de arritmia permitidos sea una función de F0 y/o de d.

El valor del desplazamiento para una sub-banda se puede determinar como el valor que dispone la sub-banda para capturar la mayor cantidad de energía. Alternativamente, el valor del desplazamiento para una sub-banda se puede determinar como el valor que centra el valor máximo de las muestras dentro de la sub-banda. Se puede ver que las localizaciones relajadas de las sub-bandas en la FIG. 4, indicadas por los bloques de líneas negras, se disponen según un criterio de centrado de picos de este tipo (como se muestra con suma claridad con referencia a los picos segundo y último, de izquierda a derecha). Un criterio de centrado de los picos tiende a producir menos varianza entre las formas de las sub-bandas, lo que puede llevar a una mejor codificación de GSVQ. Un criterio de máxima energía puede aumentar la entropía entre las formas, por ejemplo, produciendo formas que no están centradas. En un ejemplo adicional, el valor del desplazamiento para una sub-banda se determina utilizando ambos criterios.

La FIG. 5 muestra un diagrama de flujo de una implementación TA402 de la tarea TA400 que selecciona los conjuntos de sub-bandas de acuerdo con un modelo armónico relajado. La tarea TA402 incluye las tareas TA410, TA420, TA430, TA440, TA450, TA460 y TA470. En este ejemplo, la tarea TA402 está configurada para ejecutarse una vez para cada par de candidatos activos y para tener acceso a una lista ordenada de localizaciones de los picos en el intervalo de frecuencias (por ejemplo, según lo localizado por la tarea TA100). Puede ser deseable que la longitud de la lista de localizaciones de los picos sea al menos tan larga como el número máximo permitido de sub-bandas para la trama de destino (por ejemplo, ocho, diez, doce, catorce, dieciséis o dieciocho picos por trama, para un tamaño de trama de 140 o 160 muestras).

La tarea de inicialización de bucle TA410 establece el valor de un contador de bucles i en un valor mínimo (por ejemplo, uno). La tarea TA420 determina si el i-ésimo pico más alto en la lista está disponible (es decir, todavía no está en una sub-banda activa). Si el i-ésimo pico más alto está disponible, la tarea TA430 determina si cualquier sub-banda no activa se puede disponer, de acuerdo con las localizaciones indicadas por el par actual de candidatos (F0, d) (es decir, F0, F0+d, F0+2d, etc.), según lo relajado mediante el intervalo de arritmia permitido, para incluir la localización del pico. En este contexto, una "sub-banda activa" es una sub-banda que ya ha sido dispuesta sin solaparse con ninguna sub-banda dispuesta previamente y que tiene una energía mayor que (alternativamente, no inferior a) un valor de umbral T, donde T es una función de la energía máxima en las sub-bandas activas (por ejemplo, el quince, el veinte, el veinticinco o el treinta por ciento de la energía de la sub-banda activa de mayor energía ya dispuesta para esta trama). Una sub-banda no activa es una sub-banda que no está activa (es decir, todavía no está dispuesta, está dispuesta pero se solapa con otra sub-banda o no tiene energía suficiente). Si la tarea TA430 no logra encontrar ninguna sub-banda no activa que se pueda disponer para el pico, el control vuelve a la tarea TA410 mediante la tarea de incremento de bucle TA440, para procesar el siguiente pico más alto en la lista (si lo hubiera).

Puede ocurrir que existan dos valores de un número entero j para los cuales se pueda disponer una sub-banda en la localización (F0 + j*d), para incluir el i-ésimo pico (por ejemplo, el pico se encuentra entre las dos localizaciones), y que ninguno de estos valores de j esté asociado todavía a una sub-banda activa. Para dichos casos, puede ser deseable implementar la tarea TA430 para que seleccione entre estas dos sub-bandas. La tarea TA430 se puede implementar, por ejemplo, para seleccionar la sub-banda que de otro modo tendría la menor energía. En tal caso, la tarea TA430 se puede implementar para disponer cada una de las dos sub-bandas sujetas a las limitaciones de excluir el pico y no solaparse con ninguna sub-banda activa. Dentro de estas limitaciones, la tarea T430 se puede implementar para centrar cada sub-banda en la muestra más alta posible (alternativamente, para disponer cada sub-banda de tal manera que capture la máxima energía posible), para calcular la energía resultante en cada una de las

dos sub-bandas y para seleccionar la sub-banda que tenga la energía más baja como la que se debe disponer (por ejemplo, mediante la tarea TA450) para que incluya el pico. Un enfoque de este tipo puede ayudar a maximizar la energía conjunta en las localizaciones finales de las sub-bandas.

La FIG. 2B muestra un ejemplo de una aplicación de la tarea TA430. En este ejemplo, el punto en el centro del eje de frecuencia indica la localización del i-ésimo pico, el corchete en negrita indica la localización de una sub-banda activa existente, la anchura de la sub-banda es de siete muestras y el intervalo de arritmia permitida es de (+ 5, -4). También se indican las localizaciones vecinas izquierda y derecha [F0 + kd], [F0 + (k + 1) d] del i-ésimo pico, y el intervalo de disposiciones admisibles de las sub-bandas para cada una de estas localizaciones. Tal como se ha descrito anteriormente, la tarea TA430 limita el intervalo admisible de disposiciones para cada sub-banda para excluir el pico y para que no haya solapamiento con ninguna sub-banda activa. Dentro de cada intervalo restringido, como se indica en la FIG. 2B, la tarea TA430 dispone la sub-banda correspondiente de tal manera que esté centrada en la muestra más alta posible (o, alternativamente, para capturar la máxima energía posible) y selecciona la sub-banda resultante, que tenga la energía más baja, como la que se debe disponer para incluir el i-ésimo pico.

La tarea TA450 dispone la sub-banda proporcionada por la tarea TA430 y marca la sub-banda como activa o no activa según corresponda. La tarea TA450 puede estar configurada para disponer la sub-banda de tal manera que la sub-banda no se solape con ninguna sub-banda activa existente (por ejemplo, reduciendo el intervalo de arritmia permitida para la sub-banda). La tarea TA450 también puede estar configurada para disponer la sub-banda de tal manera que el i-ésimo pico esté centrado dentro de la sub-banda (es decir, en la medida permitida por el intervalo de arritmia y/o el criterio de solapamiento).

La tarea TA460 devuelve el control a la tarea TA420 mediante la tarea de incremento de bucle TA440 si quedan más sub-bandas para el actual par activo de candidatos. Del mismo modo, la tarea TA430 devuelve el control a la tarea TA420 mediante la tarea de incremento de bucle TA440 al no lograr encontrar una sub-banda no activa que se pueda disponer para el i-ésimo pico.

Si la tarea TA420 falla para cualquier valor de i, la tarea TA470 dispone las sub-bandas restantes para el actual par activo de candidatos. La tarea TA470 puede estar configurada para disponer cada sub-banda de tal manera que el valor más alto de las muestras esté centrado dentro de la sub-banda (es decir, en la medida permitida por el intervalo de arritmia y/o de tal manera que la sub-banda no se solape con ninguna sub-banda activa existente). Por ejemplo, la tarea TA470 puede estar configurada para realizar una instancia de la tarea TA450 para cada una de las sub-bandas restantes para el actual par activo de candidatos.

En este ejemplo, la tarea TA402 también incluye una tarea optativa TA480 que poda las sub-bandas. La tarea TA480 puede estar configurada para rechazar las sub-bandas que no cumplan con un umbral de energía (por ejemplo, T) y/o para rechazar las sub-bandas que se solapen con otra sub-banda que tenga una energía más alta.

La FIG. 6 muestra un ejemplo de un conjunto de sub-bandas, dispuestas de acuerdo con una implementación del procedimiento MA100 que incluye las tareas TA402 y TA602, para el intervalo entre 0 y 3,5 kHz de una señal armónica tal como la mostrada en el dominio de la MDCT. En este ejemplo, el eje e indica el valor absoluto de la MDCT, y las sub-bandas se indican mediante los bloques próximos al eje x o al eje de recipientes de frecuencia.

La tarea TA700 puede estar implementada para empaquetar los valores de arritmia seleccionados en la señal codificada (por ejemplo, para la transmisión al decodificador). Sin embargo, también es posible aplicar un modelo armónico relajado en la tarea TA400 (por ejemplo, como la tarea TA402), pero implementar la instancia correspondiente de la tarea TA700 para omitir los valores de arritmia de la señal codificada. Por ejemplo, incluso para un caso de baja velocidad de bits en el que no esté disponible ningún bit para transmitir la arritmia, aún puede ser deseable aplicar un modelo relajado en el codificador, ya que se puede esperar que la ventaja perceptiva obtenida mediante la codificación de más energía de la señal compensará el error perceptivo causado por la arritmia no corregida. Un ejemplo de una aplicación de este tipo es para la codificación de baja velocidad de bits de señales musicales.

En algunas aplicaciones, puede ser suficiente que la señal codificada incluya sólo las sub-bandas seleccionadas por un modelo armónico, de tal manera que el codificador descarte la energía de señal que esté fuera de las sub-bandas modeladas. En otros casos, puede ser deseable que la señal codificada también incluya dicha información de señal que no es capturada por el modelo armónico.

En un enfoque, se calcula una representación de la información no codificada (también denominada una señal residual) en el codificador, restando las sub-bandas reconstruidas del modelo armónico al espectro de entrada original. Un residuo calculado de esta manera habitualmente tendrá la misma longitud que la señal de entrada.

Para un caso en el que se utiliza un modelo armónico relajado para codificar la señal, los valores de arritmia que se utilizaron para desplazar las localizaciones de las sub-bandas pueden o no estar disponibles en el decodificador. Si los valores de arritmia están disponibles en el decodificador, entonces las sub-bandas decodificadas se pueden disponer en las mismas localizaciones, tanto en el decodificador como en el codificador. Si los valores de arritmia no

están disponibles en el decodificador, las sub-bandas seleccionadas se pueden disponer en el decodificador de acuerdo con una separación uniforme, como se indica mediante el par (F0, d) seleccionado. Sin embargo, para un caso en el que la señal residual fue calculada restando la señal reconstruida a la señal original, las sub-bandas sin arritmia ya no estarán alineadas en fase con la señal residual, y la adición de la señal reconstruida a una señal residual de este tipo puede dar como resultado una interferencia destructiva.

Un enfoque alternativo es calcular la señal residual como una concatenación de las regiones del espectro de la señal de entrada que no fueron capturadas por el modelo armónico (por ejemplo, aquellos recipientes que no se incluyeron en las sub-bandas seleccionadas). Un enfoque de este tipo puede ser deseable especialmente para aplicaciones de codificación en las que no se transmiten al decodificador los valores de los parámetros de arritmia. Un residuo calculado de tal manera tiene una longitud que es menor que la de la señal de entrada y que puede variar de trama a trama (por ejemplo, en función del número de sub-bandas en la trama). La FIG. 19 muestra un ejemplo de una aplicación del procedimiento MA100 para codificar los coeficientes de la MDCT correspondientes a la banda entre 3,5 y 7 kHz de una trama de la señal de audio en la que se etiquetan las regiones de un residuo de este tipo. Tal como se describe en el presente documento, puede ser deseable utilizar un esquema de codificación por pulsos (por ejemplo, la codificación factorial por pulsos) para codificar un residuo de este tipo.

Para un caso en el que los valores de los parámetros de arritmia no estén disponibles en el decodificador, la señal residual se puede insertar entre las sub-bandas decodificadas utilizando uno entre varios procedimientos diferentes. Un procedimiento de decodificación de este tipo consiste en reducir a cero cada intervalo de arritmia en la señal residual antes de añadirla a la señal reconstruida sin arritmia. Por ejemplo, para el intervalo de arritmia de (+4, -3), según se ha mencionado anteriormente, un procedimiento de este tipo incluiría la reducción a cero de las muestras de la señal residual, desde cuatro recipientes a la derecha hasta tres recipientes a la izquierda de cada una de las sub-bandas indicadas por el par (F0, d). Sin embargo, aunque un enfoque de este tipo puede eliminar la interferencia entre el residuo y las sub-bandas sin arritmia, también provoca una pérdida de información que puede ser significativa.

Otro procedimiento de decodificación es insertar el residuo para llenar los recipientes no ocupados por la señal reconstruida sin arritmia (por ejemplo, los recipientes antes, después y entre las sub-bandas reconstruidas sin arritmia). Un enfoque de este tipo desplaza efectivamente la energía del residuo para incorporar las disposiciones sin arritmia de las sub-bandas reconstruidas. La FIG. 7 muestra un ejemplo de un enfoque de este tipo, estando todos los tres gráficos de amplitud frente a frecuencia, de A a C, alineados verticalmente en la misma escala horizontal de recipientes y frecuencia. El gráfico A muestra una parte del espectro de la señal que incluye la disposición original con arritmia de una sub-banda seleccionada (puntos con relleno dentro de las líneas discontinuas) y parte del residuo circundante (puntos sin relleno). En el diagrama B, que muestra la disposición de la sub-banda sin arritmia, se puede ver que los dos primeros recipientes de la sub-banda ahora se solapan con una serie de muestras del residuo original que contiene energía (las muestras encerradas en un círculo en el gráfico A). El diagrama C muestra un ejemplo de uso del residuo concatenado para llenar los recipientes no ocupados en orden creciente de frecuencia, lo que dispone esta serie de muestras del residuo en el otro lado de la sub-banda sin arritmia.

Un procedimiento adicional de decodificación es insertar el residuo de tal manera que se mantenga la continuidad del espectro de la MDCT en los límites entre las sub-bandas sin arritmia y la señal residual. Por ejemplo, un procedimiento de este tipo puede incluir la compresión de una región del residuo que esté entre dos sub-bandas sin arritmia (o que esté antes de la primera, o después de la última, sub-banda) con el fin de evitar un solapamiento en cualquiera de los dos extremos, o en ambos. Por ejemplo, dicha compresión se puede realizar mediante la distorsión de frecuencia de la región para ocupar la zona entre las sub-bandas (o entre la sub-banda y el límite del intervalo). De manera similar, un procedimiento de este tipo puede incluir la expansión de una región del residuo que esté entre dos sub-bandas sin arritmia (o que esté antes de la primera, o después de la última, sub-banda) con el fin de llenar una brecha en cualquiera de los dos extremos, o en ambos. La FIG. 8 muestra un ejemplo de este tipo en el que se expande la parte del residuo entre las líneas discontinuas en el gráfico A de amplitud frente a frecuencia (por ejemplo, se interpola linealmente) para llenar una brecha entre sub-bandas sin arritmia, tal como se muestra en el gráfico B de amplitud frente a frecuencia.

Puede ser deseable utilizar un esquema de codificación por pulsos para codificar la señal residual, que codifica un vector correlacionándolo con un patrón de pulsos unitarios y utilizando un índice que identifica ese patrón para representar el vector. Por ejemplo, un esquema de este tipo puede estar configurado para codificar el número, las posiciones y los signos de los pulsos unitarios en la señal residual. La FIG. 9 muestra un ejemplo de un procedimiento de este tipo en el que una parte de una señal residual se codifica como un número de pulsos unitarios. En este ejemplo, un vector de dimensión treinta, cuyo valor en cada dimensión se indica mediante la línea continua, está representado por el patrón de pulsos,(0, 0, -1, -1, +1, +2, -1, 0, 0, +1, -1, -1, +1, -1, +1, -1, -1, +2, -1, 0, 0, 0, 0, -1, +1, +1, 0, 0, 0, 0), tal como se indica por los puntos (en las localizaciones de los pulsos) y los cuadrados (en las localizaciones de valor cero).

Las posiciones y los signos de un número determinado de pulsos unitarios se pueden representar como un índice de un libro de códigos. Por ejemplo, un patrón de pulsos como el mostrado en la FIG. 9 normalmente puede ser

representado por un índice de un libro de códigos cuya longitud es mucho menor que treinta bits. Los ejemplos de esquemas de codificación por pulsos incluyen los esquemas de codificación factorial por pulsos y los esquemas de codificación combinatoria por pulsos.

Puede ser deseable configurar un códec de audio para codificar diferentes bandas de frecuencia de la misma señal por separado. Por ejemplo, puede ser deseable configurar un códec de este tipo para producir una primera señal codificada que codifique una parte de banda baja de una señal de audio y una segunda señal codificada que codifique una parte de banda alta de la misma señal de audio. Las aplicaciones en las que puede ser deseable dicha codificación de banda dividida incluyen sistemas de codificación de banda ancha que deben seguir siendo compatibles con los sistemas de decodificación de banda estrecha. Dichas aplicaciones incluyen también esquemas de codificación de audio generalizados que logran una codificación eficaz de una gama de diferentes tipos de señales de entrada de audio (por ejemplo, tanto voz como música) dando soporte al uso de esquemas de codificación diferentes para bandas de frecuencia diferentes.

Para un caso en el que diferentes bandas de frecuencia de una señal se codifican por separado, puede ser posible en algunos casos aumentar la eficacia de la codificación en una banda mediante el uso de información codificada (por ejemplo, cuantizada) de otra banda, ya que esta información codificada ya será conocida en el decodificador. Por ejemplo, los principios de aplicación de un modelo armónico como el descrito en el presente documento (por ejemplo, un modelo armónico relajado) se pueden ampliar para utilizar información de una representación decodificada de los coeficientes de transformación de una primera banda de una trama de señal de audio (también denominada la señal de "referencia") para codificar los coeficientes de transformación de una segunda banda de la misma trama de señal de audio (también denominada la señal "de destino"). Para un caso de este tipo en el que el modelo armónico sea relevante, la eficacia de la codificación se puede aumentar, porque la representación decodificada de la primera banda ya está disponible en el decodificador.

Un procedimiento ampliado de este tipo puede incluir la determinación de sub-bandas de la segunda banda que estén relacionadas armónicamente con la primera banda codificada. En algoritmos de codificación de baja velocidad de bits para señales de audio (por ejemplo, señales musicales complejas), puede ser deseable dividir una trama de la señal en varias bandas (por ejemplo, una banda baja y una banda alta) y explotar una correlación entre estas bandas para codificar de manera eficaz la representación de las bandas en el dominio de la transformación.

En un ejemplo particular de dicha ampliación, los coeficientes de la MDCT correspondientes a la banda entre 3,5 y 7 kHz de una trama de señal de audio (en adelante denominada MDCT de banda superior o UB-MDCT) se codifican basándose en el espectro cuantizado de la MDCT de banda baja (de 0 a 4 kHz) de la trama. Se observa explícitamente que, en otros ejemplos de dicha ampliación, los dos intervalos de frecuencia no necesariamente tienen que solaparse y que incluso pueden estar separados (por ejemplo, codificando una banda entre 7 y 14 kHz de una trama basándose en la información de una representación decodificada de la banda entre 0 y 4 kHz). Dado que las MDCT de banda baja codificada se utilizan como una referencia para la codificación de las UB-MDCT, muchos parámetros del modelo de codificación de la banda alta se pueden obtener en el decodificador sin requerir explícitamente su transmisión.

La FIG. 10A muestra un diagrama de flujo para un procedimiento MB100 de procesamiento de señales de audio de acuerdo con una configuración general que incluye las tareas TB100, TB200, TB300, TB400, TB500, TB600 y TB700. La tarea TB100 localiza una pluralidad de picos en una señal de audio de referencia (por ejemplo, una representación descuantizada de un primer intervalo de frecuencias de una señal de frecuencias de audio). La tarea TB100 puede estar implementada como una instancia de la tarea TA100, tal como se describe en el presente documento. Para un caso en el que la señal de audio de referencia fue codificada utilizando una implementación del procedimiento MA100, puede ser deseable configurar las tareas TA100 y TB100 para que utilicen el mismo valor de d_{min}, aunque también es posible configurar las dos tareas para que utilicen valores diferentes de d_{min}. (Sin embargo, es importante señalar que el procedimiento MB100 generalmente se puede aplicar independientemente del esquema de codificación particular que fue utilizado para producir la señal de audio de referencia decodificada).

Basándose en las localizaciones en el dominio de la frecuencia de al menos algunos (es decir, al menos tres) de los picos localizados por la tarea TB100, la tarea TB200 calcula un número Nd2 de candidatos de separación armónica en la señal de audio de referencia. Los ejemplos de valores para Nd2 incluyen tres, cuatro y cinco. La tarea TB200 puede estar configurada para calcular estos candidatos de separación como las distancias (por ejemplo, en términos del número de recipientes de frecuencia) entre los picos adyacentes de los (Nd2+1) picos más altos localizados por la tarea TB100.

Basándose en las localizaciones en el dominio de la frecuencia de al menos algunos (es decir, al menos dos) de los picos localizados por la tarea TB100, la tarea TB300 identifica un número Nf2 de candidatos de F0 en la señal de audio de referencia. Los ejemplos de valores para Nf2 incluyen tres, cuatro y cinco. La tarea TB300 puede estar configurada para identificar estos candidatos como las localizaciones de los Nf2 picos más altos en la señal de audio de referencia. Alternativamente, la tarea TB300 puede estar configurada para identificar estos candidatos como las localizaciones de los Nf2 picos más altos en una parte de baja frecuencia (por ejemplo, el 30, 35, 40, 45 o 50 por ciento inferior) del intervalo de frecuencias de referencia. En un ejemplo de este tipo, la tarea TB300 identifica el

número Nf2 de candidatos de F0 entre las localizaciones de los picos localizados por la tarea TB100 en el intervalo de 0 a 1.250 Hz. En otro ejemplo de este tipo, la tarea TB300 identifica el número Nf2 de candidatos de F0 entre las localizaciones de los picos localizados por la tarea TB100 en el intervalo de 0 a 1.600 Hz.

Se hace constar expresamente que el alcance de las implementaciones descritas del procedimiento MB100 incluye el caso en el que sólo se calcula un candidato de separación armónica (por ejemplo, como la distancia entre los dos picos más altos, o la distancia entre los dos picos más altos en un intervalo de frecuencias especificado) y el caso independiente en el que sólo se identifica un candidato de F0 (por ejemplo, como la localización del pico más alto, o la localización del pico más alto en un intervalo de frecuencias especificado).

Para cada uno entre una pluralidad de pares activos de candidatos de F0 y d, la tarea TB400 tarea selecciona un conjunto de al menos una sub-banda de una señal de audio de destino (por ejemplo, una representación de un segundo intervalo de frecuencias de la señal de frecuencias de audio), en el que una localización en el dominio de la frecuencia de cada sub-banda del conjunto está basada en el par (F0, d). Sin embargo, a diferencia de la tarea TA400, en este caso las sub-bandas se disponen en relación con las localizaciones F0m, F0m+d, F0m+2d, etc., donde el valor de F0m se calcula correlacionando F0 con el intervalo de frecuencias de la señal de audio de destino. Una correlación de este tipo se puede realizar de acuerdo con una expresión tal como F0m = F0 + Ld, donde L es el entero más pequeño tal que F0m esté dentro del intervalo de frecuencias de la señal de audio de destino. En tal caso, el decodificador puede calcular el mismo valor de L sin más información del codificador, ya que el intervalo de frecuencias de la señal de audio de destino y los valores de F0 y d ya son conocidos en el decodificador.

La tarea TB400 puede estar configurada para seleccionar cada conjunto para incluir todas las sub-bandas indicadas por el par (F0, d) correspondiente que se encuentren dentro del intervalo de entrada. Alternativamente, la tarea TB400 puede estar configurada para seleccionar menos que todas estas sub-bandas para al menos uno de los conjuntos. Por ejemplo, la tarea TB400 puede estar configurada para seleccionar no más de un número máximo de sub-bandas para el conjunto. Alternativamente, o adicionalmente, la tarea TB400 puede estar configurada para seleccionar sólo las sub-bandas que se encuentren dentro de un intervalo particular. Por ejemplo, puede ser deseable configurar la tarea TB400 para seleccionar no más de un número particular de una o más (por ejemplo, cuatro, cinco o seis) de las sub-bandas de frecuencia más baja en el intervalo de entrada y/o sólo las sub-bandas cuyas localizaciones no estén por encima de una frecuencia particular dentro del intervalo de entrada (por ejemplo, 5.000, 5.500 o 6.000 Hz).

En un ejemplo, la tarea TB400 está configurada para seleccionar las sub-bandas de cada conjunto de tal manera que la primera sub-banda esté centrada en la localización de F0m correspondiente, estando el centro de cada sub-banda siguiente separado del centro de la sub-banda anterior por una distancia igual al valor correspondiente de d.

Se puede considerar que todos los pares diferentes de valores de F0 y d son activos, de manera que la tarea TB400 esté configurada para seleccionar un conjunto correspondiente de una o más sub-bandas para cada posible par (F0, d). Para un caso en el que Nf2 y Nd2 son ambos iguales a cuatro, por ejemplo, la tarea TB400 puede estar configurada para considerar cada uno de los dieciséis pares posibles. Alternativamente, la tarea TB400 puede estar configurada para imponer un criterio de actividad que algunos de los posibles pares (F0, d) pueden no cumplir. En tal caso, por ejemplo, la tarea TB400 puede estar configurada para desechar los pares que producirían más de un número máximo admisible de sub-bandas (por ejemplo, combinaciones de valores bajos de F0 y d) y/o pares que producirían menos de un número mínimo deseado de sub-bandas (por ejemplo, combinaciones de valores altos de F0 y d).

Para cada uno entre una pluralidad de pares de los candidatos de F0 y d, la tarea TB500 tarea calcula al menos un valor de energía a partir del conjunto correspondiente de una o más sub-bandas de la señal de audio de destino. En un ejemplo de este tipo, la tarea TB500 calcula un valor de energía de cada conjunto de una o más sub-bandas como la energía total del conjunto de sub-bandas (por ejemplo, como una suma de las magnitudes al cuadrado de los valores de muestreo en el dominio de frecuencia en las sub-bandas). Alternativamente, o adicionalmente, la tarea TB500 puede estar configurada para calcular valores de energía a partir de cada conjunto de sub-bandas como las energías de cada sub-banda individual y/o para calcular un valor de energía a partir de cada conjunto de sub-bandas como una energía media por sub-banda (por ejemplo, la energía total normalizada sobre el número de sub-bandas) para el conjunto de sub-bandas. La tarea TB500 puede estar configurada para ejecutarse para cada uno entre la misma pluralidad de pares que la tarea TB400, o para menos que esta pluralidad. Para un caso en el que la tarea TB400 está configurada para seleccionar un conjunto de sub-bandas para cada posible par (F0, d), por ejemplo, la tarea TB500 puede estar configurada para calcular valores de energía sólo para los pares que satisfagan un criterio especificado de actividad (por ejemplo, desechar los pares que producirían demasiadas sub-bandas y/o los pares que producirían demasiado pocas sub-bandas, tal como se ha descrito anteriormente). En otro ejemplo, la tarea TB400 está configurada para desechar los pares que producirían demasiadas sub-bandas y la tarea TB500 está configurada para desechar también los pares que producirían demasiado pocas sub-bandas.

Aunque la FIG. 10A muestra la ejecución de las tareas TB400 y TB500 en serie, se entenderá que la tarea TB500 también puede estar implementada para comenzar a calcular energías para conjuntos de sub-bandas antes de que haya finalizado la tarea TB400. Por ejemplo, la tarea TB500 puede estar implementada para comenzar a calcular (o

incluso para terminar de calcular) un valor de energía a partir de un conjunto de sub-bandas antes de que la tarea TB400 comience a seleccionar el siguiente conjunto de sub-bandas. En un ejemplo de este tipo, las tareas TB400 y TB500 están configuradas para alternarse para cada uno entre la pluralidad de pares activos de los candidatos de F0 y d. Del mismo modo, la tarea TB400 también puede estar implementada para comenzar la ejecución antes de que hayan finalizado las tareas TB200 y TB300.

Basándose en los valores de energía calculados a partir de al menos algunos de los conjuntos de al menos una sub-banda, la tarea TB600 selecciona un par de candidatos de entre los pares de candidatos de (F0, d). En un ejemplo, la tarea TB600 selecciona el par correspondiente al conjunto de sub-bandas que tienen la energía total más alta. En otro ejemplo, la tarea TB600 selecciona el par de candidatos correspondiente al conjunto de sub-bandas que tienen la energía media por sub-banda más alta. En un ejemplo adicional, la tarea TB600 está implementada como una instancia de la tarea TA602 (por ejemplo, como se muestra en la FIG. 1B).

La FIG. 10B muestra un diagrama de flujo de una implementación MB110 del procedimiento MB100 que incluye una tarea TB700. La tarea TB700 produce una señal codificada que incluye indicaciones de los valores del par de candidatos seleccionado. La tarea TB700 puede estar configurada para codificar el valor seleccionado de F0, o para codificar un desplazamiento del valor seleccionado de F0 a partir de una localización mínima (o máxima). De manera similar, la tarea TB700 puede estar configurada para codificar el valor seleccionado de d, o para codificar un desplazamiento del valor seleccionado de d a partir de una distancia mínima o máxima. En un ejemplo particular, la tarea TB700 utiliza seis bits para codificar el valor de F0 seleccionado y seis bits para codificar el valor de d seleccionado. En ejemplos adicionales, la tarea TB700 puede estar implementada para codificar el valor actual de F0 y/o d de manera diferencial (por ejemplo, como un desplazamiento relativo a un valor anterior del parámetro).

Puede ser deseable implementar la tarea TB700 para utilizar un esquema de codificación VQ (por ejemplo, GSVQ) para codificar el conjunto seleccionado de sub-bandas como vectores. Puede ser deseable utilizar un esquema de GSVQ que incluya una codificación predictiva de la ganancia, de tal manera que los factores de ganancia para cada conjunto de sub-bandas se codifiquen de forma independiente entre sí y de manera diferencial con respecto al factor de ganancia correspondiente de la trama anterior. En un ejemplo particular, el procedimiento MB110 está organizado para codificar las regiones de energía significativa en un intervalo de frecuencias de un espectro de la UB-MDCT.

Debido a que la señal de audio de referencia está disponible en el decodificador, las tareas TB100, TB200 y TB300 también pueden realizarse en el decodificador para obtener el mismo número (o "libro de códigos") Nf2 de candidatos de F0 y el mismo número ("libro de códigos") Nd2 de candidatos de d a partir de la misma señal de audio de referencia. Los valores en cada libro de códigos se pueden ordenar, por ejemplo, en orden de valores crecientes. En consecuencia, es suficiente que el codificador transmita un índice para cada una de estas pluralidades ordenadas, en lugar de codificar los valores reales del par (F0, d) seleccionado. Para un ejemplo particular en el que Nf2 y Nd2 son ambos iguales a cuatro, la tarea TB700 puede estar implementada para utilizar un índice del libro de códigos de dos bits, para indicar el valor de d seleccionado y otro índice del libro de códigos de dos bits para indicar el valor de F0 seleccionado.

Un procedimiento para decodificar una señal de audio de destino codificada, producida por la tarea TB700, también puede incluir la selección de los valores de F0 y d indicados por los índices, la descuantización del conjunto de sub-bandas seleccionado, el cálculo del valor de correlación m y la construcción de una señal de audio de destino decodificada, disponiendo (por ejemplo, centrando) cada sub-banda p en la localización en el dominio de la frecuencia F0m + pd, donde 0 <= p < P y P es el número de sub-bandas en el conjunto seleccionado. A los recipientes no ocupados de la señal decodificada de destino se les pueden asignar valores cero o, alternativamente, valores de un residuo decodificado como el descrito en el presente documento.

Como la tarea TA400, la tarea TB400 puede estar implementada como instancias iteradas de la tarea TA402, tal como se ha descrito anteriormente, con la excepción de que cada valor de F0 se correlaciona primero con F0m, tal como se ha descrito anteriormente. En este caso, la tarea TA402 está configurada para ejecutarse una vez para cada par de candidatos a evaluar y para tener acceso a una lista de localizaciones de los picos en la señal de destino, donde la lista está ordenada en orden decreciente de los valores de las muestras. Para producir una lista de este tipo, el procedimiento MB100 también puede incluir una tarea de detección de picos análoga a la tarea TB100 (por ejemplo, otra instancia de la tarea TB100) que esté configurada para operar sobre la señal de destino en lugar de sobre la señal de referencia.

La FIG. 11 muestra un gráfico de magnitud frente a frecuencia para un ejemplo en el que la señal de audio de destino es una señal de UB-MDCT de 140 coeficientes de transformación que representan el espectro de frecuencias de audio entre 3,5 y 7 kHz. Esta figura muestra la señal de audio de destino (línea gris), un conjunto de cinco sub-bandas separadas uniformemente, seleccionadas de acuerdo con un par de candidatos (F0, d) (indicado por los bloques dibujados en gris y por los corchetes), y un conjunto de cinco sub-bandas con arritmia, seleccionadas de acuerdo al par (F0, d) y a un criterio de centrado de picos (indicado por los bloques dibujados en negro). Como se muestra en este ejemplo, el espectro de la UB-MDCT se puede calcular a partir de una señal de banda alta que ha sido convertida hacia una velocidad de muestreo más baja, o desplazada de otro modo, con fines de codificación, para comenzar en un recipiente de frecuencias cero o uno. En tal caso, cada correlación de F0m

también incluye un desplazamiento para indicar la frecuencia adecuada del espectro desplazado. En un ejemplo particular, el primer recipiente de frecuencias del espectro de la UB-MDCT de la señal de audio de destino corresponde al recipiente 140 del espectro de la LB-MDCT de la señal de audio de referencia (por ejemplo, representando contenido acústico a 3,5 kHz), de tal manera que la tarea TA400 pueda ser implementada para correlacionar cada F0 con un F0m correspondiente de acuerdo con una expresión tal como F0m = F0 + Ld - 140.

Para un caso en el que la señal de audio de referencia fue codificada utilizando un modelo armónico relajado tal como el descrito en el presente documento, se pueden utilizar los mismos límites de arritmia (por ejemplo, hasta cuatro recipientes a la derecha y hasta tres recipientes a la izquierda) para la codificación de la señal de destino, utilizando un modelo armónico relajado, o se puede utilizar un límite de arritmia diferente en un lado o en ambos. Para cada sub-banda, puede ser deseable seleccionar el valor de arritmia que centre el pico dentro de la sub-banda si es posible o, si no se dispone de dicho valor de arritmia, el valor de arritmia que centre parcialmente el pico o, si no se dispone de dicho valor de arritmia, el valor de arritmia que maximice la energía capturada por la sub-banda.

En un ejemplo, la tarea TB400 está configurada para seleccionar el par (F0, d) que compacte la energía máxima por sub-banda en la señal de destino (por ejemplo, el espectro de la UB-MDCT). La compactación de la energía también se puede utilizar como una medida para decidir entre dos o más candidatos de arritmia que centran, o centran parcialmente (por ejemplo, como se ha descrito anteriormente con referencia a la tarea TA430).

Los valores de los parámetros de arritmia (por ejemplo, uno para cada sub-banda) se pueden transmitir al decodificador. Si los valores de arritmia no se transmiten al decodificador, entonces puede aparecer un error en las localizaciones de frecuencia de las sub-bandas del modelo armónico. Sin embargo, para las señales de destino que representan un intervalo de frecuencias de audio de banda alta (por ejemplo, el intervalo entre 3,5 y 7 kHz), este error habitualmente no es perceptible, de tal manera que puede ser deseable codificar las sub-bandas de acuerdo a los valores de arritmia seleccionados, pero no enviar esos valores de arritmia al decodificador, y las sub-bandas pueden estar uniformemente separadas (por ejemplo, basándose únicamente en el par (F0, d) seleccionado) en el decodificador. Por ejemplo, para la codificación de velocidad binaria muy baja de señales musicales (por ejemplo, aproximadamente veinte kilobits por segundo), puede ser deseable no transmitir los valores de los parámetros de arritmia y permitir un error en las localizaciones de las sub-bandas en el decodificador.

Tras haberse identificado el conjunto de sub-bandas seleccionadas, se puede calcular una señal residual en el codificador restando la señal de destino reconstruida al espectro de la señal de destino original (por ejemplo, como la diferencia entre el espectro de la señal de destino original y las sub-bandas reconstruidas del modelo armónico). Alternativamente, la señal residual se puede calcular como una concatenación de las regiones del espectro de la señal de destino que no fueron capturadas por el modelado armónico (por ejemplo, aquellos recipientes que no se incluyeron en las sub-bandas seleccionadas). Para un caso en el que la señal de audio de destino es un espectro de la UB-MDCT y la señal de audio de referencia es un espectro de la LB-MDCT reconstruido, puede ser deseable obtener el residuo mediante la concatenación de las regiones no capturadas, especialmente para un caso en que los valores de arritmia utilizados para codificar la señal de audio de destino no estarán disponibles en el decodificador. Las sub-bandas seleccionadas se pueden codificar utilizando un esquema de cuantización vectorial (por ejemplo, un esquema de GSVQ), y la señal residual se puede codificar utilizando un esquema de codificación factorial por pulsos o un esquema de codificación combinatoria por pulsos.

Si los valores de los parámetros de arritmia están disponibles en el decodificador, entonces la señal residual se puede volver a colocar en el decodificador en los mismos recipientes que en el codificador. Si los valores de los parámetros de arritmia no están disponibles en el decodificador (por ejemplo, para la codificación de baja velocidad binaria de señales musicales), las sub-bandas seleccionadas se pueden disponer en el decodificador de acuerdo con una separación uniforme basada en el par (F0, d) seleccionado, tal como se ha descrito anteriormente. En este caso, la señal residual se puede insertar entre las sub-bandas seleccionadas utilizando uno de varios procedimientos diferentes, tales como los descritos anteriormente (por ejemplo, reduciendo a cero cada intervalo de arritmia en el residuo antes de añadirlo a la señal sin arritmia reconstruida, utilizando el residuo para llenar recipientes no ocupados al mismo tiempo que se desplaza la energía residual que podría solaparse con una sub-banda seleccionada, o distorsionando en frecuencia el residuo).

La FIG. 12A muestra un diagrama de bloques de un aparato para el procesamiento de señales de audio MF100 según una configuración general. El aparato MF100 incluye medios FA100 para la localización de una pluralidad de picos en la señal de audio en un dominio de la frecuencia (por ejemplo, como se ha descrito en el presente documento con referencia a la tarea TA100). El aparato MF100 también incluye medios FA200 para el cálculo de un número Nd de candidatos de separación armónica (d) (por ejemplo, como se ha descrito en el presente documento con referencia a la tarea TA200). El aparato MF100 también incluye medios FA300 para identificar un número Nf de candidatos de frecuencia fundamental (F0) (por ejemplo, como se ha descrito en el presente documento con referencia a la tarea TA300). El aparato MF100 también incluye medios FA400 para seleccionar, para cada uno entre una pluralidad de pares (F0, d) diferentes, un conjunto de sub-bandas de la señal de audio cuyas localizaciones están basadas en el par (por ejemplo, como se ha descrito en el presente documento con referencia a la tarea TA400). El aparato MF100 también incluye medios FA500 para calcular, para cada uno entre la pluralidad de pares (F0, d) diferentes, una energía del conjunto correspondiente de sub-bandas (por ejemplo, como se ha descrito

en el presente documento con referencia a la tarea TA500). El aparato MF100 también incluye medios FA600 para la selección de un par de candidatos basándose en las energías calculadas (por ejemplo, como se ha descrito en el presente documento con referencia a la tarea TA600). La FIG. 13A muestra un diagrama de bloques de una implementación MF110 del aparato MF100 que incluye medios FA700 para producir una señal codificada que incluye indicaciones de los valores del par de candidatos seleccionado (por ejemplo, como se ha descrito en el presente documento con referencia a la tarea TA700).

La FIG. 12B muestra un diagrama de bloques de un aparato para el procesamiento de señales de audio A100 según otra configuración general. El aparato A100 incluye un localizador de picos en el dominio de la frecuencia 100, configurado para localizar una pluralidad de picos en la señal de audio en un dominio de la frecuencia (por ejemplo, como se ha descrito en el presente documento con referencia a la tarea TA100). El aparato A100 también incluye un calculador de distancia 200 configurado para calcular un número Nd de candidatos de separación armónica (d) (por ejemplo, como se ha descrito en el presente documento con referencia a la tarea TA200). El aparato A100 también incluye un selector de candidatos de frecuencia fundamental 300, configurado para identificar un número Nf de candidatos de frecuencia fundamental (F0) (por ejemplo, como se ha descrito en el presente documento con referencia a la tarea TA300). El aparato A100 también incluye un selector de disposición de sub-bandas 400, configurado para seleccionar, para cada uno entre una pluralidad de pares (F0, d) diferentes, un conjunto de sub-bandas de la señal de audio cuyas localizaciones están basadas en el par (por ejemplo, como se ha descrito en el presente documento con referencia a la tarea TA400). El aparato A100 también incluye un calculador de energía 500 configurado para calcular, para cada uno entre la pluralidad de pares (F0, d) diferentes, una energía del conjunto correspondiente de sub-bandas (por ejemplo, como se ha descrito en el presente documento con referencia a la tarea TA500). El aparato A100 también incluye un selector de pares de candidatos 600, configurado para seleccionar un par de candidatos basándose en las energías calculadas (por ejemplo, como se ha descrito en el presente documento con referencia a la tarea TA600). Se hace constar expresamente que el aparato A100 también puede estar implementado de tal manera que sus diversos elementos estén configurados para llevar a cabo las tareas correspondientes del procedimiento MB100 tal como se ha descrito en el presente documento.

La FIG. 13B muestra un diagrama de bloques de una implementación A110 de un aparato A100 que incluye un cuantizador 710 y un empaquetador de bits 720. El cuantizador 710 está configurado para codificar el conjunto seleccionado de sub-bandas (por ejemplo, como se ha descrito en el presente documento con referencia a la tarea TA700). Por ejemplo, el cuantizador 710 puede estar configurado para codificar las sub-bandas como vectores utilizando un esquema de GSVQ u otro esquema de VQ. El empaquetador de bis 720 está configurado para codificar los valores del par de candidatos seleccionado (por ejemplo, como se ha descrito en el presente documento con referencia a la tarea TA700) y para empaquetar estas indicaciones de los valores de los candidatos seleccionados con las sub-bandas cuantizadas, para producir una señal codificada. Un decodificador correspondiente puede incluir un desempaquetador de bits configurado para desempaquetar las sub-bandas cuantizadas y decodificar los valores de los candidatos, un descuantizador configurado para producir un conjunto descuantizado de sub-bandas y un disponedor de sub-bandas configurado para disponer las sub-bandas descuantizadas en el dominio de la frecuencia en localizaciones que están basadas en los valores de los candidatos decodificados (por ejemplo, como se ha descrito en el presente documento con referencia a la tarea TD300), y posiblemente también para disponer un residuo correspondiente, para producir una señal decodificada. Se hace constar expresamente que el aparato A110 también puede estar implementado de tal manera que sus diversos elementos estén configurados para llevar a cabo las tareas correspondientes del procedimiento MB110 tal como se ha descrito en el presente documento.

La FIG. 14 muestra un diagrama de bloques de un aparato para el procesamiento de señales de audio MF210 según una configuración general. El aparato MF210 incluye medios FB100 para la localización de una pluralidad de picos en una señal de audio de referencia en un dominio de la frecuencia (por ejemplo, como se ha descrito en el presente documento con referencia a la tarea TB100). El aparato MF210 también incluye medios FB200 para el cálculo de un número Nd2 de candidatos de separación armónica (d) (por ejemplo, como se ha descrito en el presente documento con referencia a la tarea TB200). El aparato MF210 también incluye medios FB300 para identificar un número Nf2 de candidatos de frecuencia fundamental (F0) (por ejemplo, como se ha descrito en el presente documento con referencia a la tarea TB300). El aparato MF210 también incluye medios FB400 para seleccionar, para cada uno entre una pluralidad de pares (F0, d) diferentes, un conjunto de sub-bandas de una señal de audio de destino cuyas localizaciones están basadas en el par (por ejemplo, como se ha descrito en el presente documento con referencia a la tarea TB400). El aparato MF210 también incluye medios FB500 para calcular, para cada uno entre la pluralidad de pares (F0, d) diferentes, una energía del conjunto correspondiente de sub-bandas (por ejemplo, como se ha descrito en el presente documento con referencia a la tarea TB500). El aparato MF210 también incluye medios FB600 para la selección de un par de candidatos basándose en las energías calculadas (por ejemplo, como se ha descrito en el presente documento con referencia a la tarea TB600). El aparato MF210 también incluye medios FB700 para producir una señal codificada que incluye indicaciones de los valores del par de candidatos seleccionado (por ejemplo, como se ha descrito en el presente documento con referencia a la tarea TB700).

Para un caso en el que se codifica la señal de referencia (por ejemplo, un espectro de banda baja) usando un modelo armónico (por ejemplo, una instancia del procedimiento MA100), puede ser deseable llevar a cabo una instancia de MA100 en la señal de destino (por ejemplo, un espectro de banda alta) en lugar de una instancia del procedimiento MB100. En otras palabras, puede ser deseable estimar los valores de banda alta para F0 y d

independientemente del espectro de banda alta, en lugar de correlacionar F0 a partir de los valores de banda baja como en el procedimiento MB100. En tal caso, puede ser deseable transmitir al decodificador los valores de banda superior para F0 y d o, alternativamente, transmitir la diferencia entre los valores de banda baja y de banda alta para F0 y la diferencia entre los valores de banda baja y de banda alta para d (también denominado "predicción a nivel de parámetros" de los parámetros del modelo de banda alta).

Dicha estimación independiente de los parámetros de banda alta puede tener una ventaja en términos de la capacidad de recuperación de errores, en comparación con la predicción de los parámetros a partir del espectro de banda baja decodificado (también denominado "predicción a nivel de señal"). En un ejemplo, las ganancias para las sub-bandas armónicas de banda baja se codifican utilizando un esquema de modulación diferencial adaptativo por pulsos codificados (ADPCM), que utiliza la información de las dos tramas anteriores. En consecuencia, si se pierde cualquiera de las anteriores tramas armónicas de banda baja consecutivas, la ganancia de la sub-banda en el decodificador puede ser diferente a la del codificador. Si en tal caso se utilizara la predicción a nivel de señal de los parámetros del modelo armónico de banda alta a partir del espectro decodificado de banda baja, los picos más altos pueden ser diferentes en el codificador y en el decodificador. Dicha diferencia puede dar lugar a estimaciones incorrectas para F0 y d en el decodificador, lo que podría producir eventualmente un resultado decodificado de banda alta completamente erróneo.

La FIG. 15A ilustra un ejemplo de una aplicación del procedimiento MB110 para codificar una señal de destino, que puede estar en un dominio residual de la LPC (Codificación Lineal Predictiva). En el trayecto izquierdo, la tarea S100 realiza la codificación por pulsos de todo el espectro de la señal de destino (lo que puede incluir la realización de una implementación del procedimiento MA100 o MB100 sobre un residuo de la operación de codificación por pulsos). En el trayecto derecho, se utiliza una implementación del procedimiento MB110 para codificar la señal de destino. En este caso, la tarea TB700 puede estar configurada para utilizar un esquema de VQ (por ejemplo, GSVQ) para codificar las sub-bandas seleccionadas y un procedimiento de codificación por pulsos para codificar el residuo. La tarea S200 evalúa los resultados de las operaciones de codificación (por ejemplo, mediante la decodificación de las dos señales codificadas y la comparación de las señales decodificadas con la señal de destino original) e indica qué modalidad de codificación es más adecuada actualmente.

La FIG. 15B muestra un diagrama de bloques de un sistema de codificación del modelo armónico en el que la señal de entrada es la banda alta (banda superior, "UB") de un espectro de la MDCT, que puede estar en un dominio residual de la LPC, y la señal de referencia es un espectro de la LB-MDCT reconstruido. En este ejemplo, una implementación S110 de la tarea S100 codifica la señal de destino utilizando un procedimiento de codificación por pulsos (por ejemplo, un procedimiento de codificación factorial por pulsos (FPC) o un procedimiento de codificación combinatoria por pulsos). La señal de referencia se obtiene a partir de un espectro de LB-MDCT cuantizado de la trama que puede haber sido codificada utilizando un modelo armónico, un modelo de codificación que depende de la trama codificada anterior, un esquema de codificación que utiliza sub-bandas fijas o algún otro esquema de codificación. En otras palabras, el funcionamiento del procedimiento MB110 es independiente del procedimiento particular que fue utilizado para codificar la señal de referencia. En este caso, el procedimiento MB110 puede estar implementado para codificar las ganancias de las sub-bandas utilizando un código de transformación, y el número de bits asignados para cuantizar los vectores de forma se puede calcular basándose en las ganancias codificadas y en los resultados de un análisis de la LPC. La señal codificada producida por el procedimiento MB110 (por ejemplo, utilizando GSVQ para codificar las sub-bandas seleccionadas por el modelo armónico) se compara con la señal codificada producida por la tarea S110 (por ejemplo, utilizando únicamente la codificación por pulsos, tal como la FPC), y una implementación S210 de la tarea S200 selecciona la mejor modalidad de codificación para la trama de acuerdo a una métrica perceptiva (por ejemplo, una métrica de razón entre señal y ruido con ponderación de LPC). En este caso, el procedimiento MB100 puede estar implementado para calcular las asignaciones de bits para la GSVQ y las codificaciones residuales basándose en las ganancias de sub-banda y residuales.

La selección de la modalidad de codificación (por ejemplo, como se muestra en las FIGs. 15A y 15B) se puede ampliar a un caso de múltiples bandas. En un ejemplo de este tipo, cada una entre la banda baja y la banda alta está codificada utilizando tanto una modalidad de codificación independiente (por ejemplo, una modalidad de GSVQ o de codificación por pulsos) como una modalidad de codificación armónica (por ejemplo, el procedimiento MA100 o MB100), de manera que inicialmente se consideren cuatro combinaciones de modalidades diferentes para la trama. En tal caso, puede ser deseable calcular el residuo para la modalidad de codificación armónica de banda baja restando las sub-bandas decodificadas a la señal original, tal como se ha descrito en el presente documento. A continuación, para cada una de las modalidades de banda baja, se selecciona la mejor modalidad de banda alta correspondiente (por ejemplo, de acuerdo con una comparación entre las dos opciones utilizando una métrica perceptiva en la banda alta, tal como una métrica de ponderación de LPC). De las dos opciones restantes (es decir, la modalidad independiente de banda baja con la mejor modalidad de banda alta correspondiente, y la modalidad armónica de banda baja con la mejor modalidad de banda alta correspondiente), la selección entre estas opciones se realiza con referencia a una métrica perceptiva (por ejemplo, una métrica perceptiva de ponderación de LPC) que abarca tanto la banda baja como la banda alta. En un ejemplo de un caso de múltiples bandas de este tipo, la modalidad independiente de banda baja utiliza la GSVQ para codificar un conjunto de sub-bandas fijas, y la modalidad independiente de banda alta utiliza un esquema de codificación por pulsos (por ejemplo, la codificación factorial por pulsos) para codificar la señal de banda alta.

Las FIGs. 16A a 16E muestran una serie de aplicaciones para las diversas implementaciones del aparato A110 (o MF110 o MF210) tal como se ha descrito en el presente documento. La FIG. 16A muestra un diagrama de bloques de un trayecto de procesamiento de audio que incluye un módulo de transformación MM1 (por ejemplo, un módulo de transformación rápida de Fourier o de MDCT) y una instancia del aparato A110 (o MF110 o MF210) que está dispuesto para recibir las tramas de audio SA10 como muestras en el dominio de la transformación (es decir, como coeficientes en el dominio de la transformación) y para producir las correspondientes tramas codificadas SE10.

La FIG. 16B muestra un diagrama de bloques de una implementación del trayecto de la FIG. 16A en la que el módulo de transformación MM1 está implementado utilizando un módulo de la transformación MDCT. El módulo de la DCT modificada MM10 realiza una operación de MDCT sobre cada trama de audio para producir un conjunto de coeficientes en el dominio de la MDCT.

La FIG. 16C muestra un diagrama de bloques de una implementación del trayecto de la FIG. 16A que incluye un módulo de análisis de una codificación predictiva lineal AM10. El módulo de análisis de codificación predictiva lineal (LPC) AM10 realiza una operación de análisis de LPC sobre la trama clasificada para producir un conjunto de parámetros de LPC (por ejemplo, coeficientes de filtro) y una señal residual de LPC. En un ejemplo, el módulo de análisis de LPC AM10 está configurado para realizar un análisis de LPC de décimo orden sobre una trama que tiene un ancho de banda desde cero a 4.000 Hz. En otro ejemplo, el módulo de análisis de LPC AM10 está configurado para realizar un análisis de LPC de sexto orden sobre una trama que representa un intervalo de frecuencias de banda alta desde 3.500 a 7.000 Hz. El módulo de la DCT modificada MM10 realiza una operación de MDCT sobre la señal residual de LPC para producir un conjunto de coeficientes en el dominio de la transformación. Un trayecto de decodificación correspondiente puede estar configurado para decodificar tramas codificadas SE10 y para realizar una transformación inversa de la MDCT sobre las tramas decodificadas, para obtener una señal de excitación de entrada para un filtro de síntesis de LPC.

La FIG. 16D muestra un diagrama de bloques de un trayecto de procesamiento que incluye un clasificador de señales SC10. El clasificador de señales SC10 recibe tramas SA10 de una señal de audio y clasifica cada trama en una de al menos dos categorías. Por ejemplo, el clasificador de señales SC10 puede estar configurado para clasificar una trama SA10 como voz o música, de tal manera que si la trama se clasifica como música, entonces el resto del trayecto que se muestra en la FIG. 16D se utiliza para codificarla, y si la trama se clasifica como voz, entonces se utiliza un trayecto de procesamiento diferente para codificarla. Dicha clasificación puede incluir la detección de actividad de las señales, la detección de ruido, la detección de periodicidad, la detección de dispersión en el dominio del tiempo y/o la detección de dispersión en el dominio de la frecuencia.

La FIG. 17A muestra un diagrama de bloques de un procedimiento MC100 de clasificación de señales que puede ser realizado por el clasificador de señales SC10 (por ejemplo, sobre cada una de las tramas de audio SA10). El procedimiento MC100 incluye las tareas TC100, TC200, TC300, TC400, TC500 y TC600. La tarea TC100 cuantifica el nivel de actividad en la señal. Si el nivel de actividad es inferior a un umbral, la tarea TC200 codifica la señal como silencio (por ejemplo, utilizando un esquema de predicción lineal excitada por ruido (NELP) de baja velocidad de bits y/o un esquema de transmisión discontinua (DTX)). Si el nivel de actividad es suficientemente alto (por ejemplo, superior al umbral), la tarea TC300 cuantifica un grado de periodicidad de la señal. Si la tarea TC300 determina que la señal no es periódica, la tarea TC400 codifica la señal utilizando un esquema de NELP. Si la tarea TC300 determina que la señal es periódica, la tarea TC500 cuantifica un grado de dispersión de la señal en el dominio del tiempo y/o de la frecuencia. Si la tarea TC500 determina que la señal está dispersa en el dominio del tiempo, la tarea TC600 codifica la señal utilizando un esquema de predicción lineal excitada por código (CELP), como por ejemplo la CELP relajada (RCELP) o la CELP algebraica (ACELP). Si la tarea TC500 determina que la señal está dispersa en el dominio de la frecuencia, la tarea TC700 codifica la señal utilizando un modelo armónico (por ejemplo, haciendo pasar la señal al resto del trayecto de procesamiento en la FIG. 16D).

Como se muestra en la FIG. 16D, el trayecto de procesamiento puede incluir un módulo perceptivo de poda PM10 que está configurado para simplificar la señal en el dominio de la MDCT (por ejemplo, para reducir el número de coeficientes a codificar en el dominio de la transformación) mediante la aplicación de criterios psico-acústicos tales como el enmascaramiento temporal, el enmascaramiento de frecuencia y/o un umbral de audición. El módulo PM10 puede estar implementado para calcular los valores para dichos criterios mediante la aplicación de un modelo perceptivo a las tramas de audio originales SA10. En este ejemplo, el aparato A110 (o MF110 o MF210) está dispuesto para codificar las tramas podadas para producir las correspondientes tramas codificadas SE10.

La FIG. 16E muestra un diagrama de bloques de una implementación de ambos trayectos de las FIGs. A1C y A1D, en el que el aparato A110 (o MF110 o MF210) está dispuesto para codificar el residuo de LPC.

La FIG. 17B muestra un diagrama de bloques de un dispositivo de comunicaciones D10 que incluye una implementación del aparato A100. El dispositivo D10 incluye un chip o conjunto de chips CS10 (por ejemplo, un conjunto de chips de un módem de estación móvil (MSM)) que incorpora los elementos del aparato A100 (o MF100 y/o MF210). El chip/conjunto de chips CS10 puede incluir uno o más procesadores, que pueden estar configurados para ejecutar una parte de software y/o firmware del aparato A100 o MF100 (por ejemplo, como instrucciones).

El chip/conjunto de chips CS10 incluye un receptor, que está configurado para recibir una señal de comunicaciones de radiofrecuencia (RF) y para decodificar y reproducir una señal de audio codificada dentro de la señal de RF, y un transmisor, que está configurado para transmitir una señal de comunicaciones de RF que describe una señal de audio codificada (por ejemplo, tal como la producida por la tarea TA700 o TB700). Un dispositivo de este tipo puede estar configurado para transmitir y recibir datos de comunicaciones de voz de forma inalámbrica mediante uno o más esquemas de codificación y decodificación (también denominados "códecs"). Ejemplos de este tipo de códecs incluyen el Códec de velocidad variable mejorado (Enhanced Variable Rate Codec), como se describe en el documento del Proyecto 2 de Colaboración de Tercera Generación (3GPP2) C.S0014-C, v1.0, titulado "Códec de velocidad variable mejorado, opciones del servicio de voz 3, 68 y 70 para sistemas digitales de banda ancha de espectro ensanchado", febrero de 2007 (disponible en línea en www.3gpp.org); el códec de voz Vocodificador de modalidad seleccionable (Selectable Mode Vocoder), como se describe en el documento del 3GPP2 C.S0030-0, v3.0, titulado "Vocodificador de modalidad seleccionable (SMV), opción de servicio para sistemas de comunicación de banda ancha de espectro ensanchado", enero de 2004 (disponible en línea en www.3gpp.org); el códec de voz adaptativo de múltiples velocidades (Adaptive Multi Rate) (AMR), como se describe en el documento del ETSI TS 126 092 V6.0.0 (Instituto europeo de normas de telecomunicaciones (ETSI), Sophia Antipolis Cedex, FR, diciembre de 2004); y el códec de voz AMR de banda ancha, como se describe en el documento del ETSI TS 126 192 V6.0.0 (ETSI, diciembre de 2004).

El dispositivo D10 está configurado para recibir y transmitir las señales de comunicaciones de RF a través de una antena C30. El dispositivo D10 también puede incluir un diplexor y uno o más amplificadores de potencia en el trayecto hacia la antena C30. El chip/conjunto de chips CS10 también está configurado para recibir la entrada del usuario mediante el panel de teclas C10 y para mostrar información mediante la pantalla C20. En este ejemplo, el dispositivo D10 también incluye una o más antenas C40 para dar soporte a los servicios de localización del Sistema de Localización Global (GPS) y/o a comunicaciones de corto alcance con un dispositivo externo, como por ejemplo un auricular inalámbrico (por ejemplo, Bluetooth™). En otro ejemplo, un dispositivo de comunicaciones de este tipo es en sí mismo un auricular Bluetooth™ y carece del panel de teclas C10, la pantalla C20 y la antena C30.

El dispositivo de comunicaciones D10 se puede integrar en diversos dispositivos de comunicaciones, incluyendo teléfonos inteligentes y ordenadores portátiles y de tableta. La FIG. 18 muestra las vistas delantera, trasera y lateral de un equipo de mano H100 (por ejemplo, un teléfono inteligente) que tiene dos micrófonos de voz MV10-1 y MV10-3 dispuestos en la cara delantera, un micrófono de voz MV10-2 dispuesto en la cara trasera, un micrófono de error ME10 situado en una esquina superior de la cara delantera y un micrófono de referencia de ruido MR10 situado en la cara trasera. Un altavoz LS10 está dispuesto en la parte superior central de la cara delantera, cerca del micrófono de error ME10, y también se proporcionan otros dos altavoces LS20L, LS20R (por ejemplo, para aplicaciones de manos libres). Una distancia máxima entre los micrófonos de un equipo de mano de este tipo es habitualmente de aproximadamente diez o doce centímetros.

Los procedimientos y aparatos divulgados en el presente documento se pueden aplicar en general en cualquier aplicación de transcepción y/o de detección de audio, especialmente instancias móviles, o portátiles de otro modo, de dichas aplicaciones. Por ejemplo, la gama de configuraciones divulgadas en el presente documento incluye dispositivos de comunicaciones que residen en un sistema de comunicación de telefonía inalámbrica, configurado para emplear una interfaz inalámbrica (over-the-air) de Acceso Múltiple por División de Código (CDMA). No obstante, los expertos en la técnica comprenderán que un procedimiento y aparato que tengan las características descritas en el presente documento pueden residir en cualquiera de los diversos sistemas de comunicación que emplean una amplia gama de tecnologías conocidas por los expertos en la técnica, tales como sistemas que emplean voz sobre IP (VoIP) por canales de transmisión por cable y/o inalámbricos (por ejemplo, de CDMA, TDMA, FDMA y/o TD-SCDMA).

Se contempla expresamente y por la presente se divulga que los dispositivos de comunicaciones descritos en el presente documento se pueden adaptar para su uso en redes que son conmutadas por paquetes (por ejemplo, redes por cable y/o inalámbricas dispuestas para transportar transmisiones de audio de acuerdo con protocolos tales como VoIP) y/o conmutadas por circuitos. También se contempla expresamente y por la presente se divulga que los dispositivos de comunicaciones divulgados en el presente documento se pueden adaptar para su uso en sistemas de codificación de banda estrecha (por ejemplo, sistemas que codifican un intervalo de frecuencias de audio de aproximadamente cuatro o cinco kilohercios) y/o para su uso en sistemas de codificación de banda ancha (por ejemplo, sistemas que codifican frecuencias de audio superiores a cinco kilohercios), incluyendo sistemas de codificación de banda ancha y banda completa y sistemas de codificación de banda ancha y banda dividida.

Se proporciona la presentación de las configuraciones descritas con el fin de permitir que cualquier experto en la técnica realice o use los procedimientos y otras estructuras divulgados en el presente documento. Los diagramas de flujo, los diagramas de bloques y otras estructuras mostradas y descritas en el presente documento son solamente ejemplos, y otras variantes de estas estructuras también están dentro del alcance de la divulgación. Son posibles diversas modificaciones de estas configuraciones, y los principios genéricos presentados en el presente documento se pueden aplicar también a otras configuraciones. Por lo tanto, la presente divulgación no pretende estar limitada a las configuraciones mostradas anteriormente, sino que debe concedérsele el alcance más amplio congruente con los

principios y características novedosas divulgados de cualquier manera en el presente documento, incluso en las reivindicaciones adjuntas como fueron presentadas, que forman una parte de la divulgación original.

Los expertos en la técnica entenderán que la información y las señales se pueden representar usando cualquiera entre varias tecnologías y técnicas diferentes. Por ejemplo, los datos, las instrucciones, los comandos, la información, las señales, los bits y los símbolos que pueden haber sido mencionados a lo largo de la descripción anterior, pueden representarse mediante tensiones, corrientes, ondas electromagnéticas, campos o partículas magnéticos, campos o partículas ópticos, o cualquier combinación de los mismos.

Los requisitos de diseño importantes para la implementación de una configuración, según lo divulgado en el presente documento pueden incluir la minimización del retardo de procesamiento y/o de la complejidad computacional (medida habitualmente en millones de instrucciones por segundo, o MIPS), especialmente para aplicaciones intensivas en cálculos, tales como la reproducción de audio comprimido o de información audiovisual (por ejemplo, un fichero o un flujo codificado de acuerdo a un formato de compresión, tal como uno de los ejemplos identificados en el presente documento) o aplicaciones para comunicaciones de banda ancha (por ejemplo, comunicaciones de voz a velocidades de muestreo superiores a ocho kilohercios, tales como, por ejemplo, 12, 16, 44,1, 48 o 192 kHz).

Un aparato como el descrito en el presente documento (por ejemplo, el aparato A100, A110, MF100, MF110 o MF210) se puede implementar en cualquier combinación de hardware con software, y/o con firmware, que se considere adecuada para la aplicación deseada. Por ejemplo, dichos elementos pueden fabricarse como dispositivos electrónicos y/u ópticos que residen, por ejemplo, en el mismo chip o entre dos o más chips en un conjunto de chips. Un ejemplo de un dispositivo de este tipo es una formación fija o programable de elementos lógicos, tales como transistores o compuertas lógicas, y cualquiera de estos elementos se puede implementar como una o más de dichas formaciones. Dos o más cualesquiera, o incluso la totalidad, de estos elementos se pueden implementar dentro de la(s) misma(s) formación o formaciones. Una formación o formaciones de este tipo se pueden implementar en uno o más chips (por ejemplo, dentro de un conjunto de chips que incluya dos o más chips).

Uno o más elementos de las diversas implementaciones del aparato divulgado en el presente documento (por ejemplo, el aparato A100, A110, MF100, MF110 o MF210) se pueden implementar, en su totalidad o en parte, como uno o más conjuntos de instrucciones dispuestas para ejecutarse en una o más formaciones fijas o programables de elementos lógicos, tales como microprocesadores, procesadores integrados, núcleos de IP, procesadores de señales digitales, FPGA (Formaciones de compuertas programables en el terreno (Field Programmable Gate Arrays)), ASSP (productos estándar específicos para una aplicación (application-specific standard products)) y ASIC (circuitos integrados específicos de la aplicación (application-specific integrated circuits)). Cualquiera de los diversos elementos de una implementación de un aparato, según lo divulgado en el presente documento, también se puede realizar como uno o más ordenadores (por ejemplo, máquinas que incluyen una o más formaciones programadas para ejecutar uno o más conjuntos o secuencias de instrucciones, también denominadas "procesadores"), y dos o más cualesquiera, o incluso la totalidad, de estos elementos se pueden implementar dentro del (de los) mismo(s) ordenador u ordenadores de este tipo.

Un procesador u otros medios para el procesamiento, según lo divulgado en el presente documento, se pueden fabricar como uno o más dispositivos electrónicos y/u ópticos que residen, por ejemplo, en el mismo chip o entre dos o más chips en un conjunto de chips. Un ejemplo de un dispositivo de este tipo es una formación fija o programable de elementos lógicos, tales como transistores o compuertas lógicas, y cualquiera de estos elementos se puede implementar como una o más de dichas formaciones. Una formación, o formaciones, de este tipo se pueden implementar dentro de uno o más chips (por ejemplo, dentro de un conjunto de chips que incluya dos o más chips). Los ejemplos de dichas formaciones incluyen formaciones fijas o programables de elementos lógicos, tales como microprocesadores, procesadores integrados, núcleos de IP, DSP, FPGA, ASSP y ASIC. Un procesador u otros medios para el procesamiento, según lo divulgado en el presente documento, también se pueden realizar como uno o más ordenadores (por ejemplo, máquinas que incluyen una o más formaciones programadas para ejecutar uno o más conjuntos o secuencias de instrucciones) u otros procesadores. Es posible que un procesador como el descrito en el presente documento se utilice para realizar tareas o ejecutar otros conjuntos de instrucciones que no estén directamente relacionadas con un procedimiento de una implementación del procedimiento MA100, MA110, MB100, MB110 o MD100, tal como una tarea relacionada con otra operación de un dispositivo o sistema en el que está integrado el procesador (por ejemplo, un dispositivo de detección de audio). También es posible que parte de un procedimiento, según lo divulgado en el presente documento, sea realizado por un procesador del dispositivo de detección de audio y que otra parte del procedimiento se realice bajo el control de uno o más procesadores distintos.

Los expertos apreciarán que los diversos módulos, bloques lógicos, circuitos y pruebas ilustrativos, y otras operaciones descritas en relación con las configuraciones divulgadas en el presente documento, se pueden implementar como hardware electrónico, software informático o combinaciones de ambos. Dichos módulos, bloques lógicos, circuitos y operaciones se pueden implementar o realizar con un procesador de propósito general, un procesador de señales digitales (DSP), un ASIC o ASSP, una FPGA u otro dispositivo de lógica programable, lógica de compuertas discretas o de transistores, componentes de hardware discretos, o con cualquier combinación de los mismos diseñada para producir la configuración según lo divulgado en el presente documento. Por ejemplo, una

configuración de este tipo se puede implementar, al menos en parte, como un circuito cableado, como una configuración de circuitos fabricada en un circuito integrado específico de la aplicación, o como un programa de firmware cargado en un almacenamiento no volátil o un programa de software cargado desde o en un medio de almacenamiento de datos, como código legible por máquina, siendo dicho código instrucciones ejecutables por una formación de elementos lógicos tales como un procesador de propósito general u otra unidad de procesamiento de señales digitales. Un procesador de propósito general puede ser un microprocesador, pero, como alternativa, el procesador puede ser cualquier procesador, controlador, micro-controlador o máquina de estados convencional. Un procesador también puede implementarse como una combinación de dispositivos informáticos, por ejemplo, una combinación de un DSP y un microprocesador, una pluralidad de microprocesadores, uno o más microprocesadores junto con un núcleo de DSP o cualquier otra configuración de este tipo. Un módulo de software puede residir en un medio de almacenamiento no transitorio tal como una RAM (memoria de acceso aleatorio), una ROM (memoria de sólo lectura), una RAM no volátil (NVRAM), tal como una RAM flash, una ROM programable y borrable (EPROM), una ROM programable y borrable eléctricamente (EEPROM), registros, un disco duro, un disco extraíble o un CD-ROM; o en cualquier otra forma de medio de almacenamiento conocido en la técnica. Un medio de almacenamiento ilustrativo está acoplado al procesador de manera que el procesador pueda leer información de, y escribir información en, el medio de almacenamiento. Como alternativa, el medio de almacenamiento puede estar integrado en el procesador. El procesador y el medio de almacenamiento pueden residir en un ASIC. El ASIC puede residir en un terminal de usuario. Como alternativa, el procesador y el medio de almacenamiento pueden residir como componentes discretos en un terminal de usuario.

Se observa que los diversos procedimientos divulgados en el presente documento (por ejemplo, los procedimientos MA100, MA110, MB100, MB110 o MD100) se pueden realizar mediante una formación de elementos lógicos tales como un procesador, y que los diversos elementos de un aparato, según lo descrito en el presente documento, se pueden implementar como módulos diseñados para ejecutarse en una formación de este tipo. Tal como se utiliza en el presente documento, el término "módulo" o "sub-módulo" puede referirse a cualquier procedimiento, aparato, dispositivo, unidad o medio de almacenamiento de datos legible por ordenador que incluya instrucciones de ordenador (por ejemplo, expresiones lógicas) en forma de software, hardware o firmware. Se ha de entender que varios módulos o sistemas se pueden combinar en un módulo o sistema y un módulo o sistema se puede separar en varios módulos o sistemas para realizar las mismas funciones. Cuando se implementan en software u otras instrucciones ejecutables por ordenador, los elementos de un proceso son esencialmente los segmentos de código que realizan las tareas relacionadas, tal como, por ejemplo, con rutinas, programas, objetos, componentes, estructuras de datos y similares. Se debe comprender que el término "software" incluye el código fuente, el código en lenguaje ensamblador, el código de máquina, el código binario, el firmware, el macro-código, el micro-código, cualquiera entre uno o más conjuntos o secuencias de instrucciones ejecutables por una formación de elementos lógicos y cualquier combinación de dichos ejemplos. Los segmentos de programa o de código se pueden almacenar en un medio legible por procesador, o transmitir por una señal de datos de ordenador incorporada en una onda portadora, por un medio de transmisión o un enlace de comunicación.

Las implementaciones de procedimientos, esquemas y técnicas descritos en el presente documento también se pueden realizar de manera tangible (por ejemplo, en características tangibles, legibles por ordenador, de uno o más medios de almacenamiento legibles por ordenador según lo enumerado en el presente documento) como uno o más conjuntos de instrucciones ejecutables por una máquina que incluye una formación de elementos lógicos (por ejemplo, un procesador, un microprocesador, un micro-controlador u otra máquina de estados finitos). El término "medio legible por ordenador" puede incluir cualquier medio que pueda almacenar o transmitir información, incluyendo medios de almacenamiento volátiles, no volátiles, extraíbles y no extraíbles. Los ejemplos de un medio legible por ordenador incluyen un circuito electrónico, un dispositivo de memoria de semiconductores, una ROM, una memoria flash, una ROM borrable (EROM), un disquete flexible u otro almacenamiento magnético, un CD-ROM/DVD u otro almacenamiento óptico, un disco duro o cualquier otro medio que se pueda utilizar para almacenar la información deseada, un medio de fibra óptica, un enlace de radiofrecuencia (RF) o cualquier otro medio que se pueda utilizar para transportar la información deseada y al que se pueda acceder. La señal de datos de ordenador puede incluir cualquier señal que se pueda propagar a través de un medio de transmisión tal como canales de redes electrónicas, fibras ópticas, aire, campos electromagnéticos, enlaces de RF, etc. Los segmentos de código se pueden descargar mediante redes de ordenadores tales como Internet o una Intranet. En cualquier caso, el alcance de la presente divulgación no debería interpretarse como limitado por dichos modos de realización.

Cada una de las tareas de los procedimientos descritos en el presente documento se puede realizar directamente en hardware, en un módulo de software ejecutado por un procesador o en una combinación de los dos. En una aplicación típica de una implementación de un procedimiento, según lo divulgado en el presente documento, una formación de elementos lógicos (por ejemplo, compuertas lógicas) está configurada para realizar una, más de una, o incluso la totalidad de las diversas tareas del procedimiento. Una o más de (posiblemente todas) las tareas también se pueden implementar como código (por ejemplo, uno o más conjuntos de instrucciones), incorporar en un producto de programa de ordenador (por ejemplo, uno o más medios de almacenamiento de datos tales como discos, memorias flash u otras tarjetas de memoria no volátil, chips de memoria de semiconductores, etc.), que sea legible y/o ejecutable por una máquina (por ejemplo, un ordenador) que incluya una formación de elementos lógicos (por ejemplo, un procesador, un microprocesador, un micro-controlador u otra máquina de estados finitos). Las tareas de una implementación de un procedimiento, según lo descrito en el presente documento, también pueden ser

realizadas por más de una formación o máquina de este tipo. En estas u otras implementaciones, las tareas se pueden llevar a cabo dentro de un dispositivo de comunicaciones inalámbricas, tal como un teléfono celular u otro dispositivo que tenga dicha capacidad de comunicaciones. Un dispositivo de este tipo se puede configurar para comunicarse con redes conmutadas por circuitos y/o conmutadas por paquetes (por ejemplo, utilizando uno o más protocolos, tales como VoIP). Por ejemplo, un dispositivo de este tipo puede incluir circuitos de RF configurados para recibir y/o transmitir tramas codificadas.

Se divulga expresamente que los diversos procedimientos divulgados en el presente documento se pueden realizar mediante un dispositivo de comunicaciones portátil, tal como un equipo de mano, un auricular o un asistente digital portátil (PDA), y que los diversos aparatos descritos en el presente documento se pueden incluir en un dispositivo de este tipo. Una aplicación típica de tiempo real (por ejemplo, en línea) es una conversación telefónica realizada utilizando un dispositivo móvil de este tipo.

En uno o más modos de realización ejemplares, las operaciones descritas en el presente documento se pueden implementar en hardware, software, firmware o cualquier combinación de los mismos. Si se implementan en software, dichas operaciones, como una o más instrucciones o código, se pueden almacenar en, o transmitir por, un medio legible por ordenador. El término "medios legibles por ordenador" incluye tanto medios de almacenamiento legibles por ordenador como medios de comunicación (por ejemplo, de transmisión). A modo de ejemplo, y no de limitación, los medios de almacenamiento legibles por ordenador pueden comprender una formación de elementos de almacenamiento, tales como una memoria de semiconductores (que puede incluir, sin limitación, una RAM dinámica o estática, una ROM, una EEPROM, y/o una RAM flash), o una memoria ferro-eléctrica , magneto-resistiva, ovónica, polimérica o de cambio de fase; un CD-ROM u otro almacenamiento de disco óptico; y/o un almacenamiento en disco magnético u otros dispositivos de almacenamiento magnético. Dichos medios de almacenamiento pueden almacenar información en forma de instrucciones o estructuras de datos a las que se puede acceder mediante un ordenador. Los medios de comunicación pueden comprender cualquier medio que se pueda utilizar para transportar el código de programa deseado en forma de instrucciones o estructuras de datos y al que se pueda acceder mediante un ordenador, incluyendo cualquier medio que facilite la transferencia de un programa informático de un lugar a otro. Además, cualquier conexión puede denominarse debidamente un medio legible por ordenador. Por ejemplo, si el software se transmite desde una sede de la Red, un servidor u otro origen remoto, usando un cable coaxial, un cable de fibra óptica, un par trenzado, una línea de abonado digital (DSL) o tecnología inalámbrica tal como infrarrojos, radio y/o microondas, entonces el cable coaxial, el cable de fibra óptica, el par trenzado, la DSL o la tecnología inalámbrica tal como infrarrojos, radio y/o microondas, se incluyen en la definición de medio. Los discos, como se usan en el presente documento, incluyen el disco compacto (CD), el disco de láser, el disco óptico, el disco versátil digital (DVD), el disco flexible y el disco Blu-ray™ (Asociación de discos Blu-Ray, Universal City, CA), donde algunos discos normalmente reproducen datos de manera magnética, mientras que otros discos reproducen los datos de manera óptica con láser. Las combinaciones de lo que antecede también deberían incluirse dentro del alcance de los medios legibles por ordenador.

Un aparato de procesamiento de señales acústicas como el descrito en el presente documento puede estar incorporado en un dispositivo electrónico que acepte la entrada de voz a fin de controlar ciertas operaciones, o que de otro modo pueda beneficiarse de la separación de ruidos deseados de ruidos de fondo, tal como los dispositivos de comunicaciones. Muchas aplicaciones pueden beneficiarse de la mejora o de la separación del sonido claro deseado de los sonidos de fondo procedentes de múltiples direcciones. Dichas aplicaciones pueden incluir interfaces hombre-máquina en dispositivos electrónicos o informáticos que incorporan capacidades tales como el reconocimiento y la detección de voz, la mejora y la separación de voz, el control activado por voz y similares. Puede ser deseable implementar un aparato de procesamiento de señales acústicas de este tipo para que sea adecuado en dispositivos que sólo proporcionan capacidades de procesamiento limitadas.

Los elementos de las diversas implementaciones de los módulos, elementos y dispositivos descritos en el presente documento pueden fabricarse como dispositivos electrónicos y/u ópticos que residen, por ejemplo, en el mismo chip o entre dos o más chips en un conjunto de chips. Un ejemplo de un dispositivo de este tipo es una formación fija o programable de elementos lógicos, tales como transistores o compuertas. Uno o más elementos de las diversas implementaciones del aparato divulgado en el presente documento también se pueden implementar, en su totalidad o en parte, como uno o más conjuntos de instrucciones dispuestas para ejecutarse en una o más formaciones fijas o programables de elementos lógicos, tales como microprocesadores, procesadores integrados, núcleos de IP, procesadores de señales digitales, FPGA, ASSP y ASIC.

Es posible que uno o más elementos de una implementación de un aparato, según lo descrito en el presente documento, se utilicen para llevar a cabo tareas o ejecutar otros conjuntos de instrucciones que no estén directamente relacionadas con una operación del aparato, tal como una tarea relacionada con otra operación de un dispositivo o sistema en el que está integrado el aparato. También es posible que uno o más elementos de una implementación de un aparato de este tipo tengan una estructura en común (por ejemplo, un procesador utilizado para ejecutar partes de código correspondientes a diferentes elementos en instantes diferentes, un conjunto de instrucciones ejecutadas para llevar a cabo las tareas correspondientes a diferentes elementos en instantes diferentes, o una disposición de dispositivos electrónicos y/u ópticos que realizan operaciones para diferentes elementos en instantes diferentes).

REIVINDICACIONES

1. Un procedimiento (MD100) de construcción de una trama de audio decodificado, utilizando una pluralidad de vectores de sub-bandas decodificadas, comprendiendo cada vector de sub-banda valores para una pluralidad de recipientes de frecuencia, comprendiendo dicho procedimiento:

disponer (TD300), en un dominio de frecuencia, un primer vector de una pluralidad de vectores de sub-bandas decodificadas, de acuerdo a un valor de frecuencia fundamental;

disponer (TD300), en el dominio de la frecuencia, el resto de la pluralidad de vectores de sub-bandas decodificadas, de acuerdo al valor de frecuencia fundamental y a un valor de separación armónica; e

insertar, en el dominio de la frecuencia, una señal residual decodificada en ubicaciones de la trama que no estén ocupadas por la pluralidad de vectores de sub-bandas decodificadas.

2. El procedimiento de acuerdo a la reivindicación 1, en el que, para cada par adyacente de la pluralidad de vectores de sub-bandas decodificadas, una distancia entre los centros de los vectores es igual al valor de separación armónica.

3. El procedimiento de acuerdo a la reivindicación 1 o la reivindicación 2, en el que dicho procedimiento comprende borrar partes de la señal residual decodificada que corresponden a posibles ubicaciones de la pluralidad de vectores de sub-bandas decodificadas.

4. El procedimiento de acuerdo a cualquiera de las reivindicaciones 1 a 3, en el que dicha inserción de una señal residual decodificada incluye la inserción de valores de la señal residual decodificada, en orden desde un primer valor de la señal residual decodificada hasta un último valor de la señal residual decodificada, en las ubicaciones no ocupadas de la trama en orden de frecuencia creciente.

5. El procedimiento de acuerdo a cualquiera de las reivindicaciones 1 a 4, en el que dicha inserción de una señal residual decodificada incluye la distorsión de una parte de la señal residual decodificada con respecto a un eje del dominio de la frecuencia para que quepa entre vectores adyacentes entre la pluralidad de vectores de sub-bandas decodificadas.

6. Un aparato de procesamiento de señales de audio para construir una trama de audio decodificado utilizando una pluralidad de vectores de sub-bandas decodificadas, comprendiendo cada vector de sub-banda valores para una pluralidad de recipientes de frecuencia, comprendiendo el aparato:

medios para disponer, en un dominio de frecuencia, un primer vector de una pluralidad de vectores de sub-bandas decodificadas de acuerdo a un valor de frecuencia fundamental;

medios para disponer, en el dominio de la frecuencia, el resto de la pluralidad de vectores de sub-bandas decodificadas de acuerdo al valor de frecuencia fundamental y a un valor de separación armónica; y

medios para insertar, en el dominio de la frecuencia, una señal residual decodificada en ubicaciones de la trama que no estén ocupadas por la pluralidad de vectores de sub-bandas decodificadas.

7. El aparato de acuerdo a la reivindicación 6, en el que, para cada par adyacente de la pluralidad de vectores de sub-bandas decodificadas, una distancia entre los centros de los vectores es igual al valor de separación armónica.

8. El aparato de acuerdo a la reivindicación 6 o la reivindicación 7, que comprende medios para borrar partes de la señal residual decodificada que corresponden a posibles ubicaciones de la pluralidad de vectores de sub-bandas decodificadas.

9. El aparato de acuerdo a cualquiera de las reivindicaciones 6 a 8, en el que dicho medio para insertar una señal residual decodificada comprende medios para insertar valores de la señal residual decodificada, en orden desde un primer valor de la señal residual decodificada hasta un último valor de la señal residual decodificada, en las ubicaciones no ocupadas de la trama en orden de frecuencia creciente.

10. El aparato de acuerdo a cualquiera de las reivindicaciones 6 a 9, en el que dicho medio para la inserción de una señal residual decodificada comprende medios para distorsionar una parte de la señal residual decodificada con respecto a un eje del dominio de la frecuencia para que quepa entre vectores adyacentes entre la pluralidad de vectores de sub-bandas decodificadas.

11. Un medio de almacenamiento no transitorio legible por ordenador que tiene características tangibles que, cuando son leídas por una máquina, provocan que la máquina lleve a cabo los procedimientos de cualquiera

de las reivindicaciones 1 a 5.

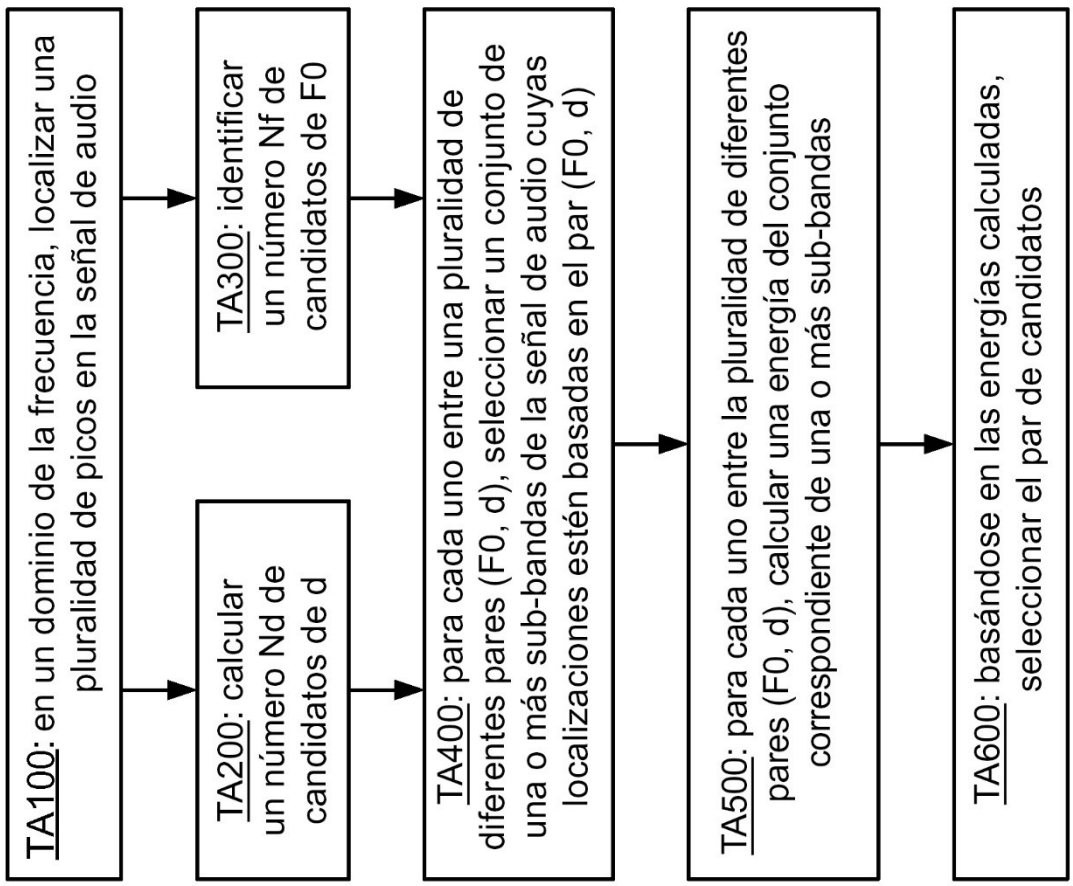

TA100: en un dominio de la frecuencia, localizar una pluralidad de picos en la señal de audio
TA200: calcular un número Nd de candidatos de d
TA300: identificar un número Nf de candidatos de F0
Procedimiento MA100
TA400: para cada uno entre una pluralidad de diferentes pares (F0, d), seleccionar un conjunto de una o más sub-bandas de la señal de audio cuyas localizaciones estén basadas en el par (F0, d)
TA500: para cada uno entre la pluralidad de diferentes pares (F0, d), calcular una energía del conjunto correspondiente de una o más sub-bandas
TA600: basándose en las energías calculadas, seleccionar el par de candidatos

FIG. 1A

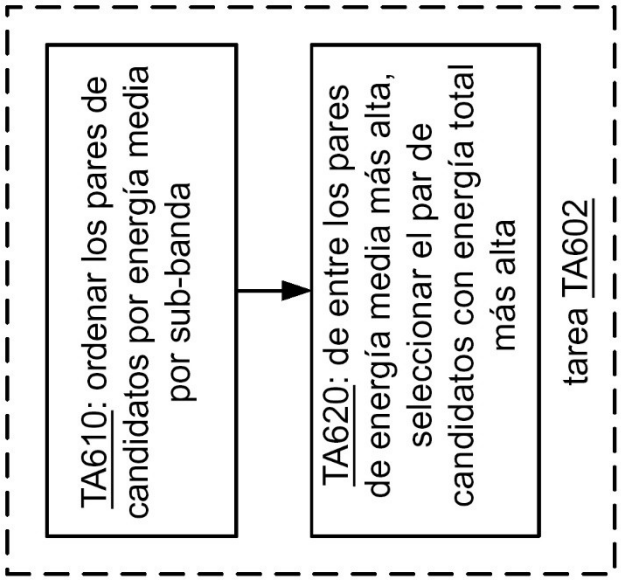

TA610: ordenar los pares de candidatos por energía media por sub-banda
TA620: de entre los pares de energía media más alta, seleccionar el par de candidatos con energía total más alta
tarea TA602

FIG. 1B

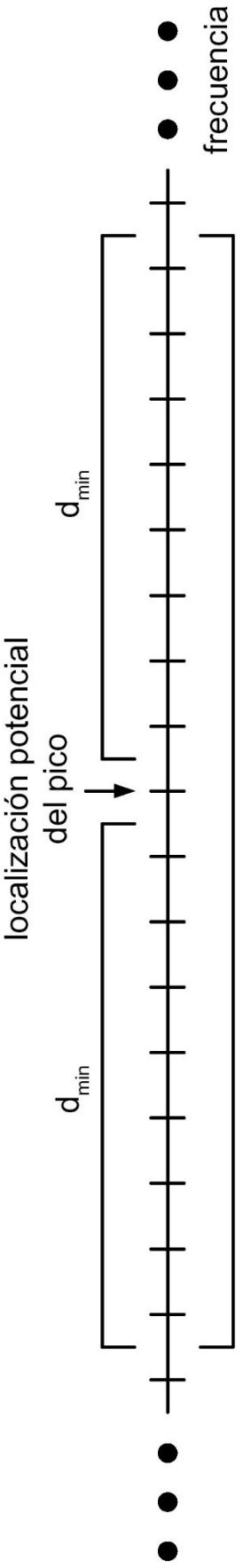

localización potencial
del pico
d_{min}
d_{min}
ventana de selección de picos
frecuencia

FIG. 2A

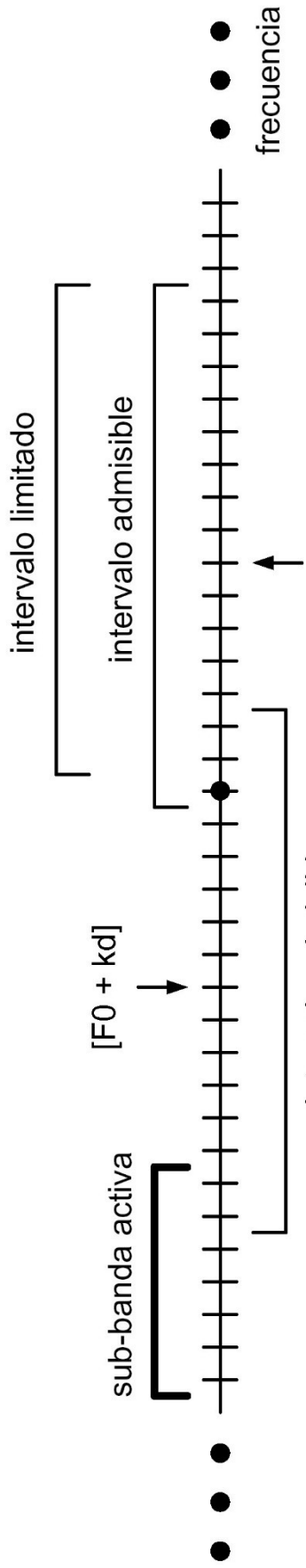

intervalo limitado
intervalo admisible
[F0 + (k+1)d]
frecuencia
[F0 + kd]
sub-banda activa
intervalo admisible
intervalo limitado

FIG. 2B

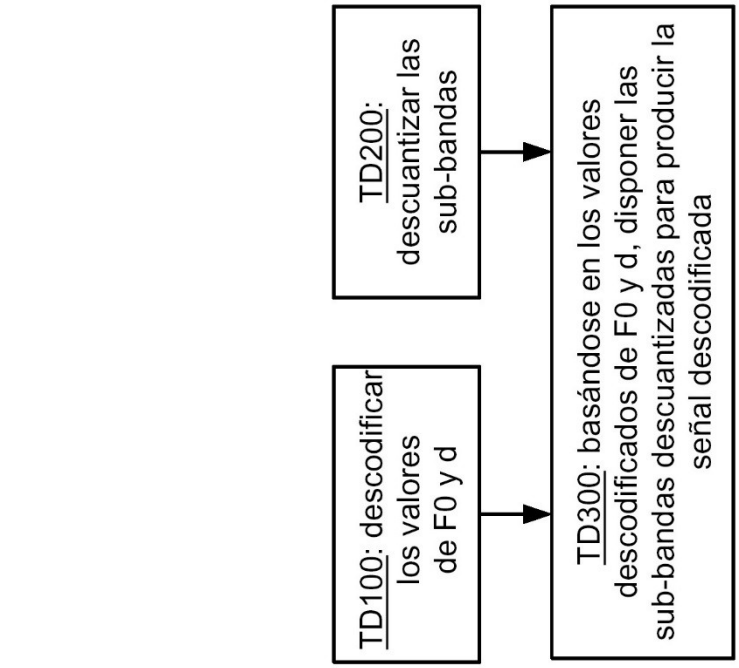

TD200: descuantizar las sub-bandas
TD100: descodificar los valores de F0 y d
TD300: basándose en los valores descodificados de F0 y d, disponer las sub-bandas descuantizadas para producir la señal descodificada
procedimiento MD100

FIG. 3B

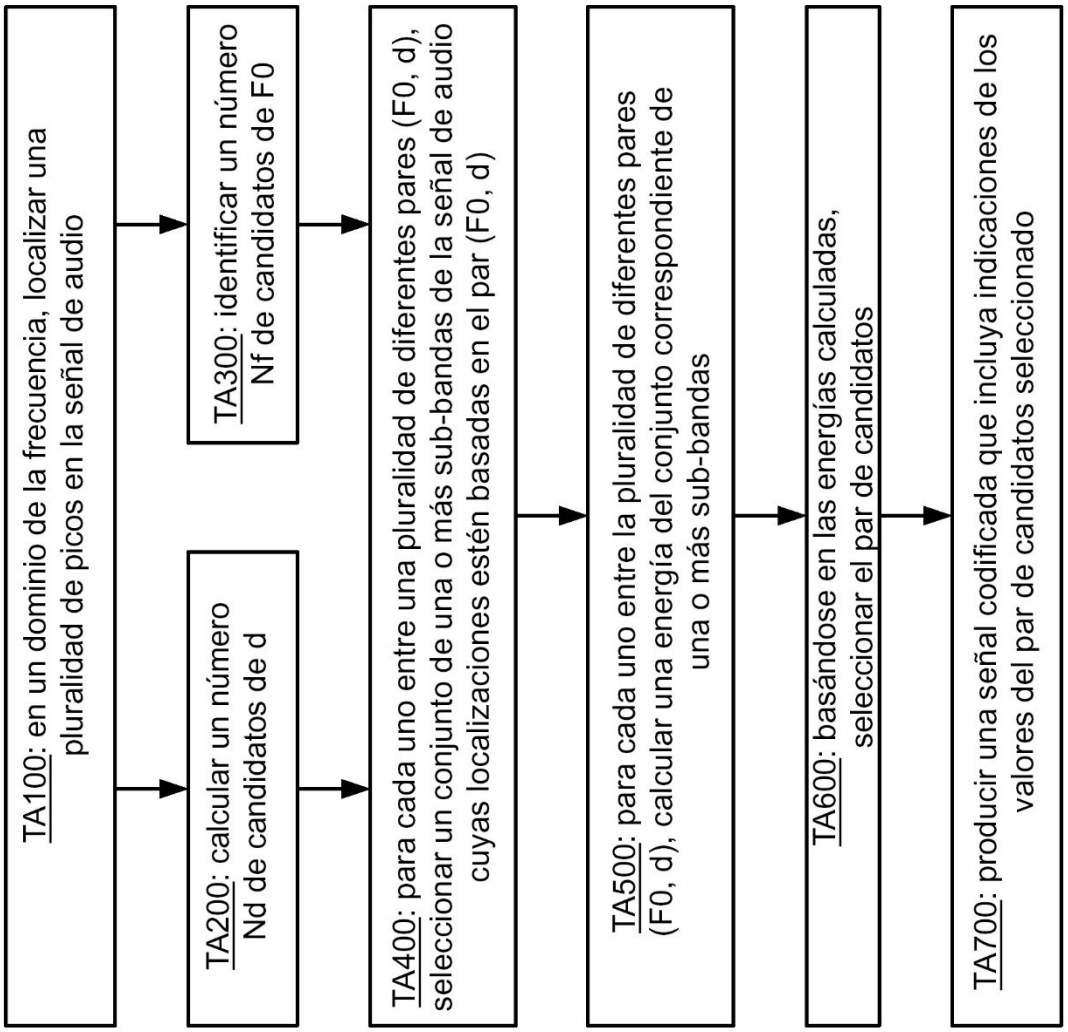
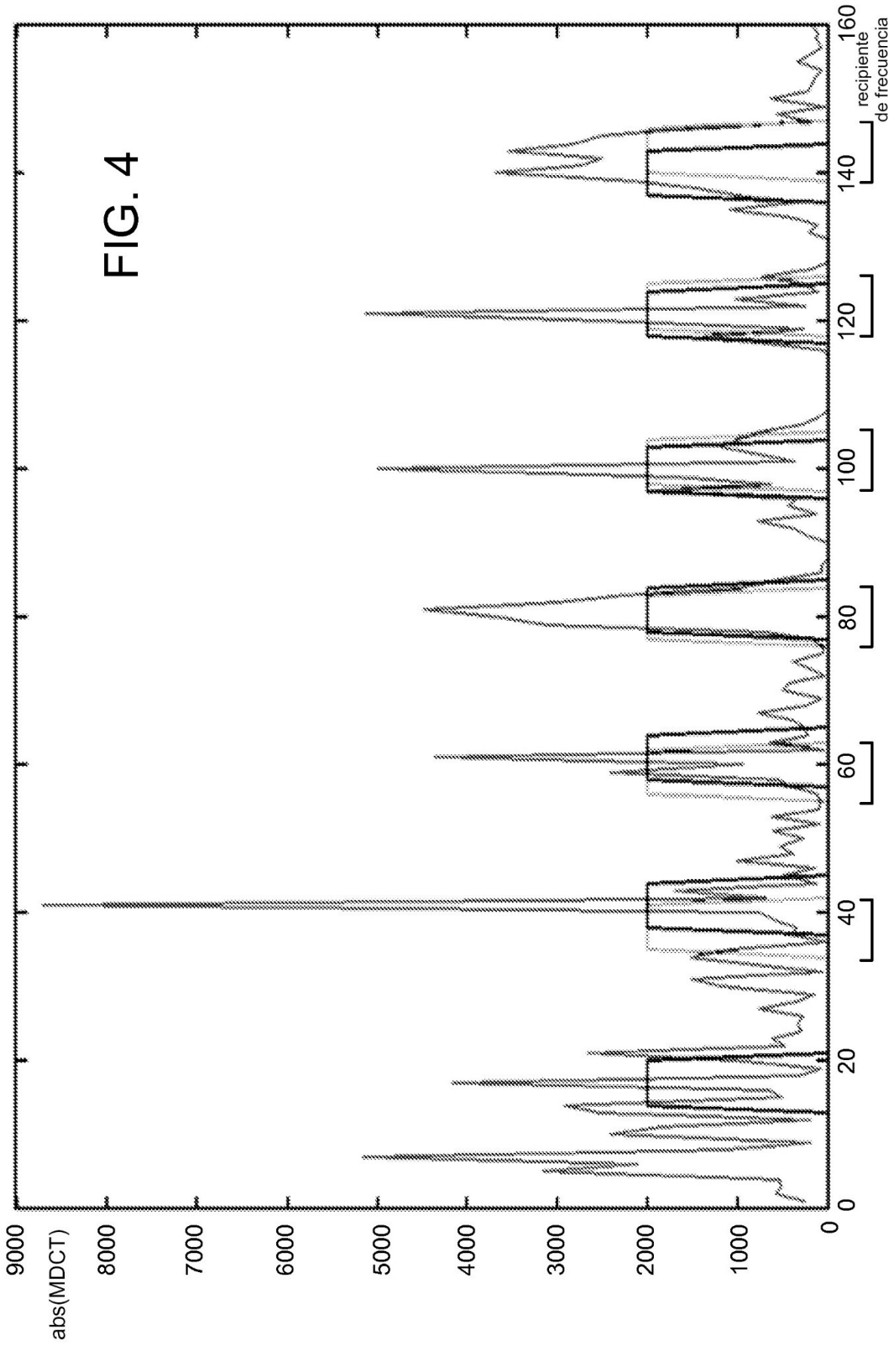

TA100: en un dominio de la frecuencia, localizar una pluralidad de picos en la señal de audio
TA200: calcular un número Nd de candidatos de d
TA300: identificar un número Nf de candidatos de F0
TA400: para cada uno entre una pluralidad de diferentes pares (F0, d), seleccionar un conjunto de una o más sub-bandas de la señal de audio cuyas localizaciones estén basadas en el par (F0, d)
TA500: para cada uno entre la pluralidad de diferentes pares (F0, d), calcular una energía del conjunto correspondiente de una o más sub-bandas
TA600: basándose en las energías calculadas, seleccionar el par de candidatos
TA700: producir una señal codificada que incluya indicaciones de los valores del par de candidatos seleccionado
procedimiento MA110

FIG. 3A

FIG. 4


abs(MDCT)
9000
8000
7000
6000
5000
4000
3000
2000
1000
0
0
20
40
60
80
100
120
140
160
recipiente
de frecuencia

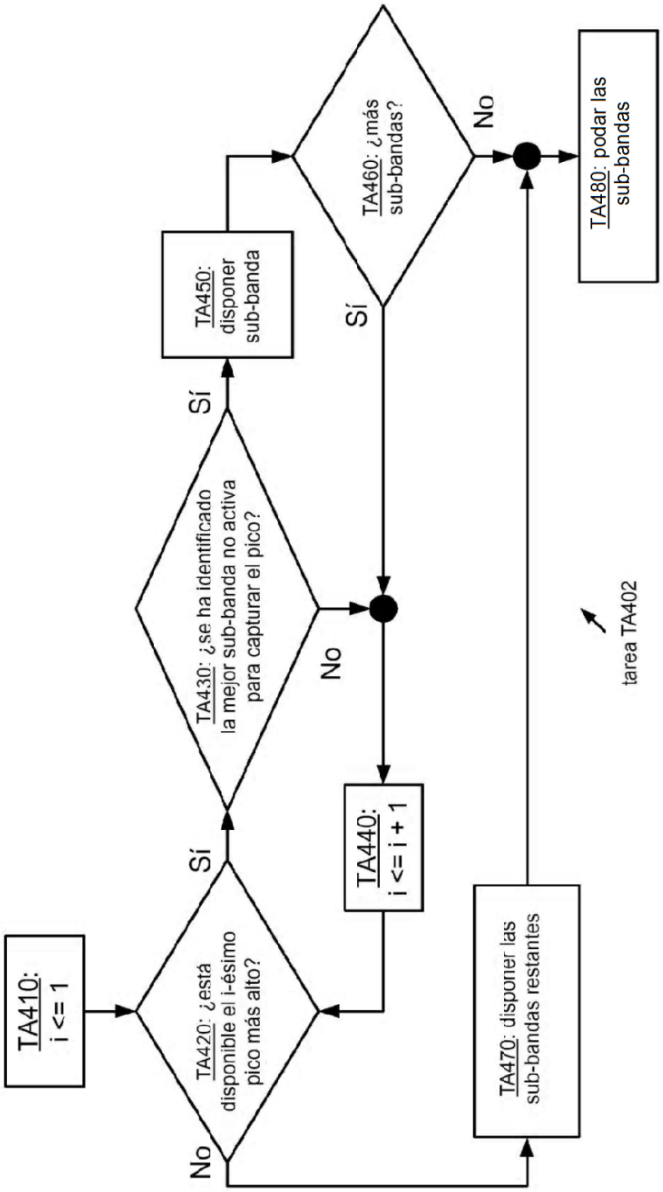

TA410: i <= 1
TA420: ¿está disponible el i-ésimo pico más alto?
No
Sí
TA430: ¿se ha identificado la mejor sub-banda no activa para capturar el pico?
No
Sí
TA440: i <= i + 1
TA450: disponer sub-banda
TA460: ¿más sub-bandas?
Sí
No
TA470: disponer las sub-bandas restantes
TA480: podar las sub-bandas
tarea TA402

FIG. 5

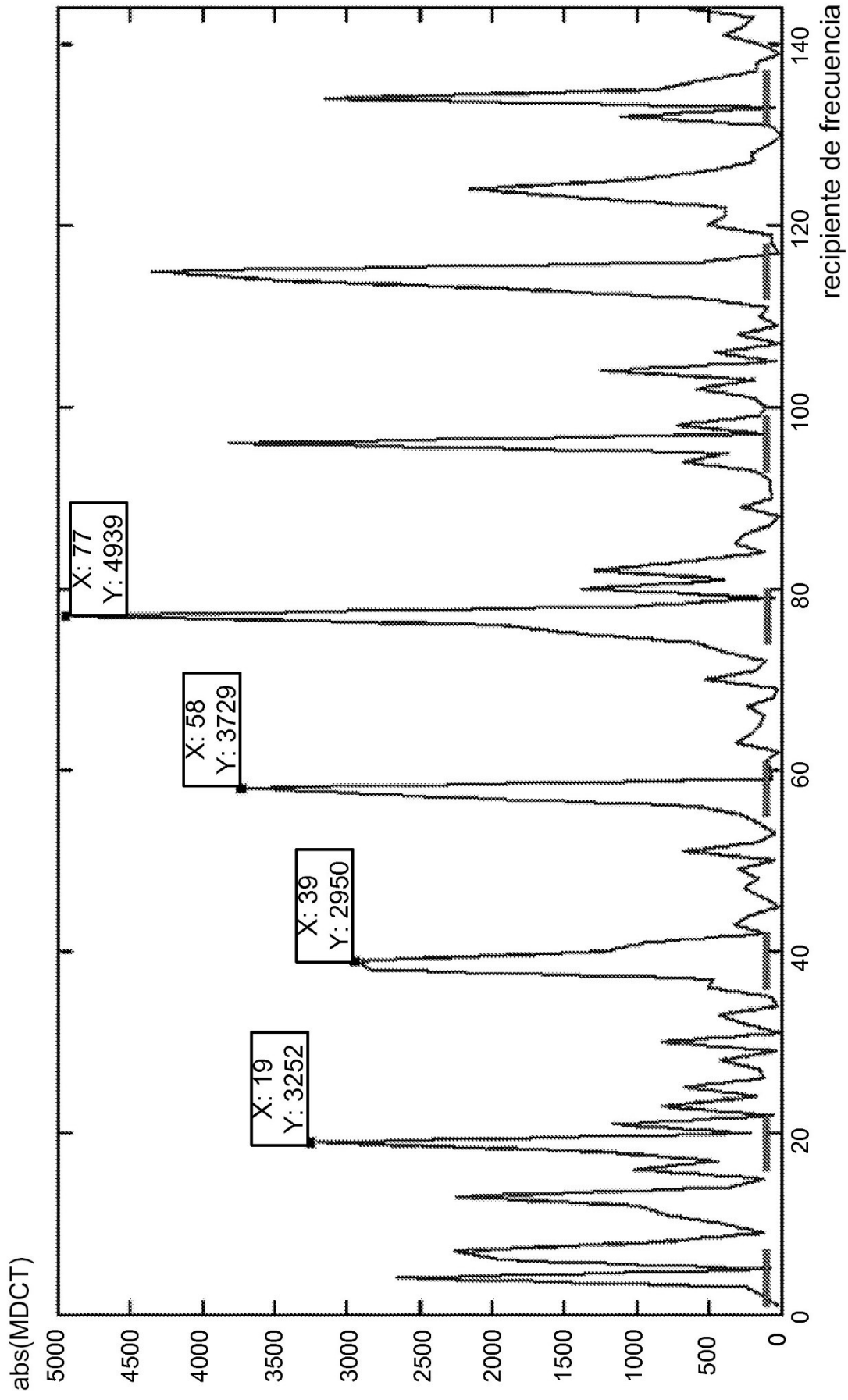
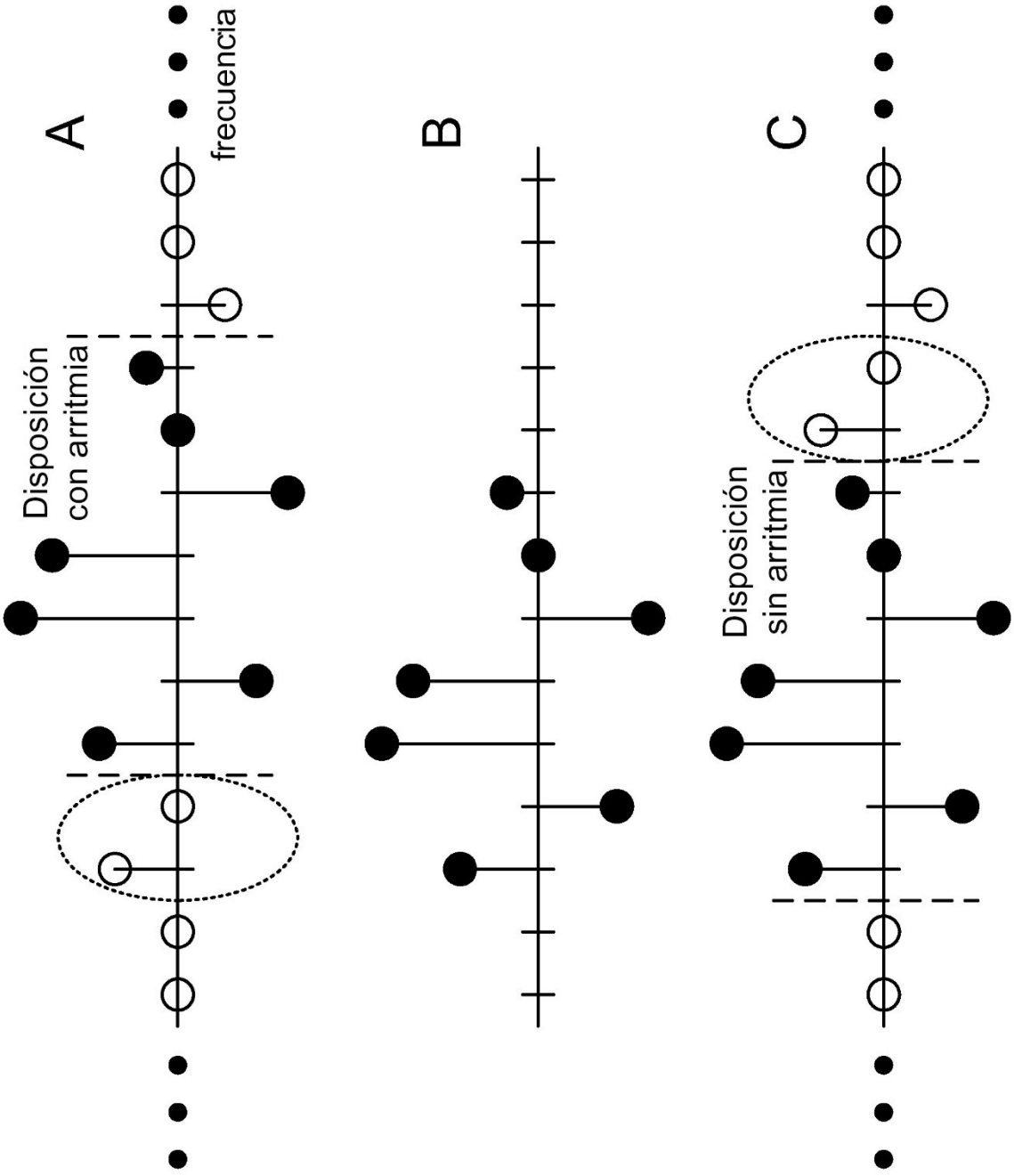

abs(MDCT)
recipiente de frecuencia
X: 77
Y: 4939
X: 58
Y: 3729
X: 39
Y: 2950
X: 19
Y: 3252
0
500
1000
1500
2000
2500
3000
3500
4000
4500
5000
0
20
40
60
80
100
120
140

FIG. 6

FIG. 7


A
frecuencia
Disposición
con arritmia
B
C
Disposición
sin arritmia

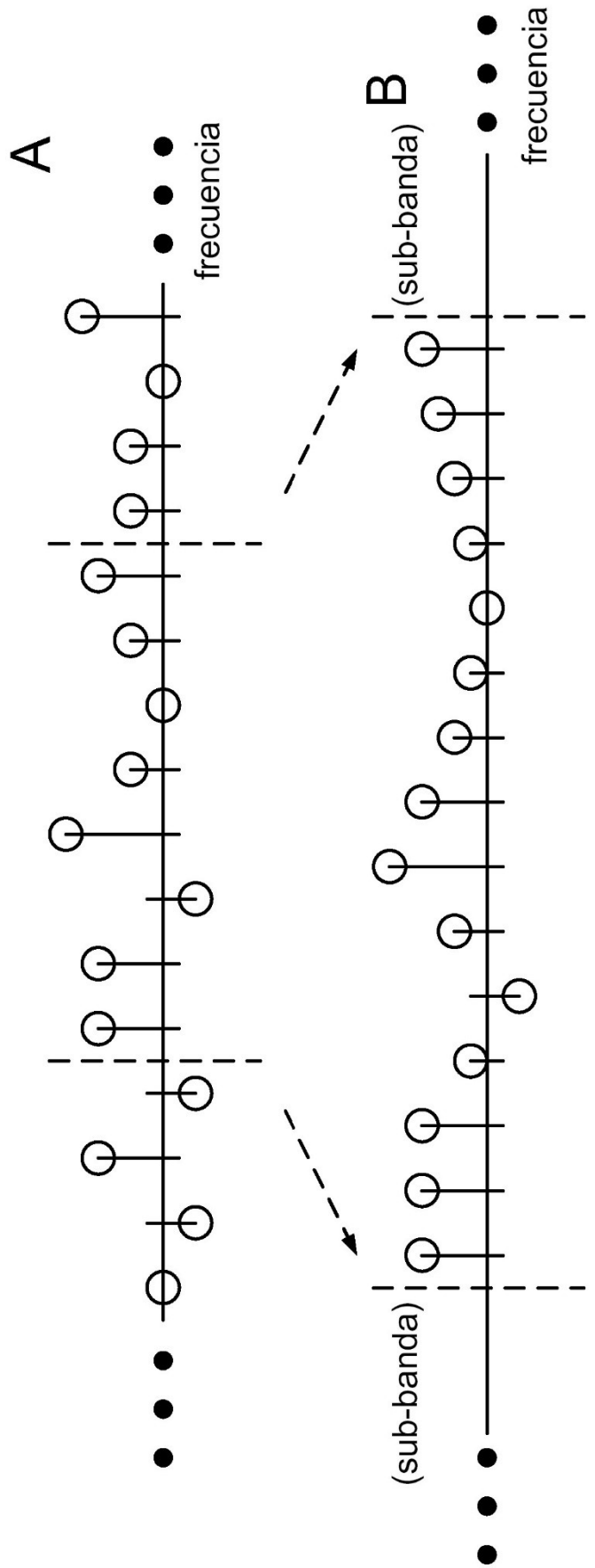


FIG. 8

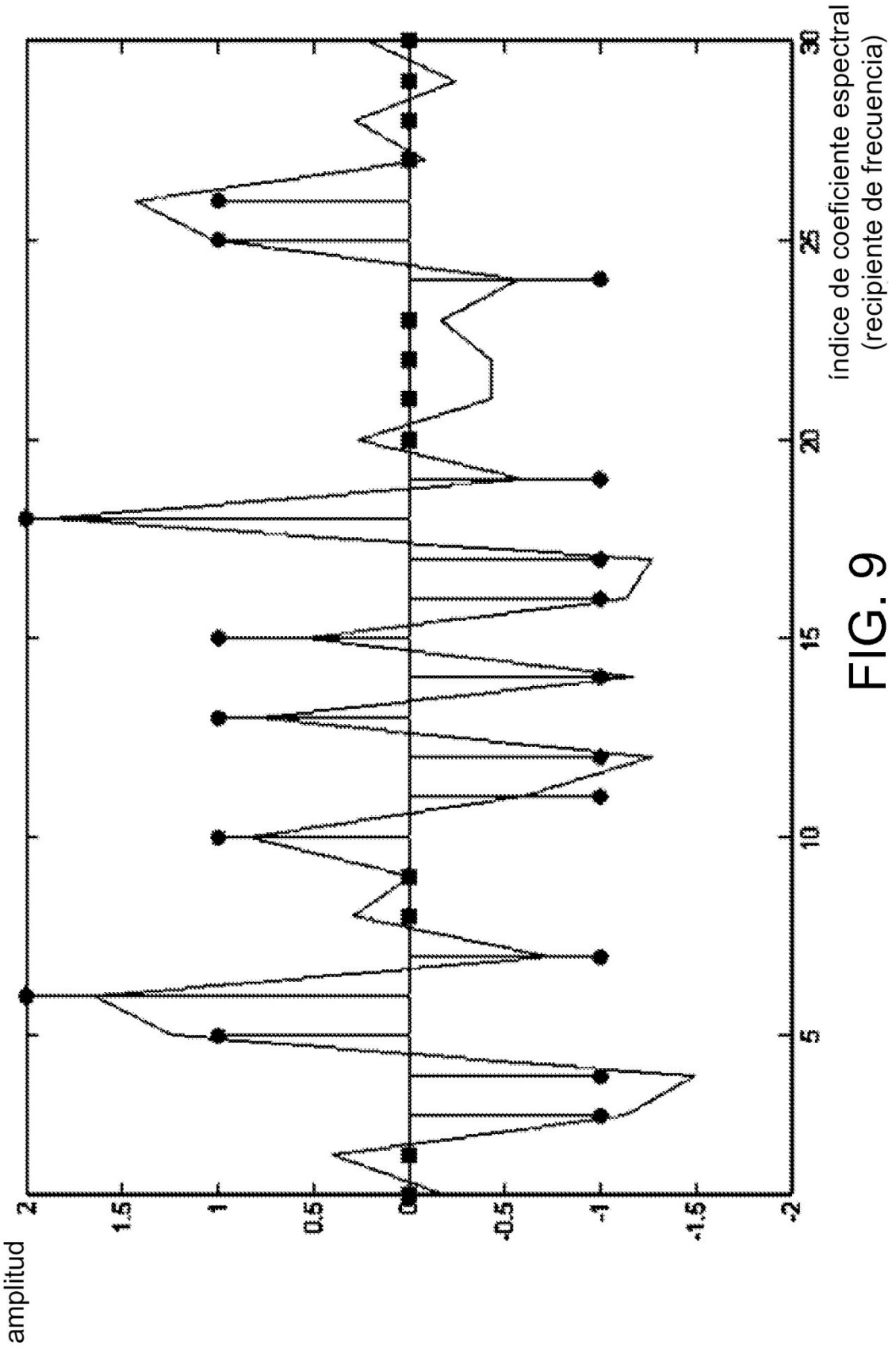

amplitud
2
1.5
1
0.5
0
-0.5
-1
-1.5
-2
5
10
15
20
25
30
índice de coeficiente espectral
(recipiente de frecuencia)

FIG. 9

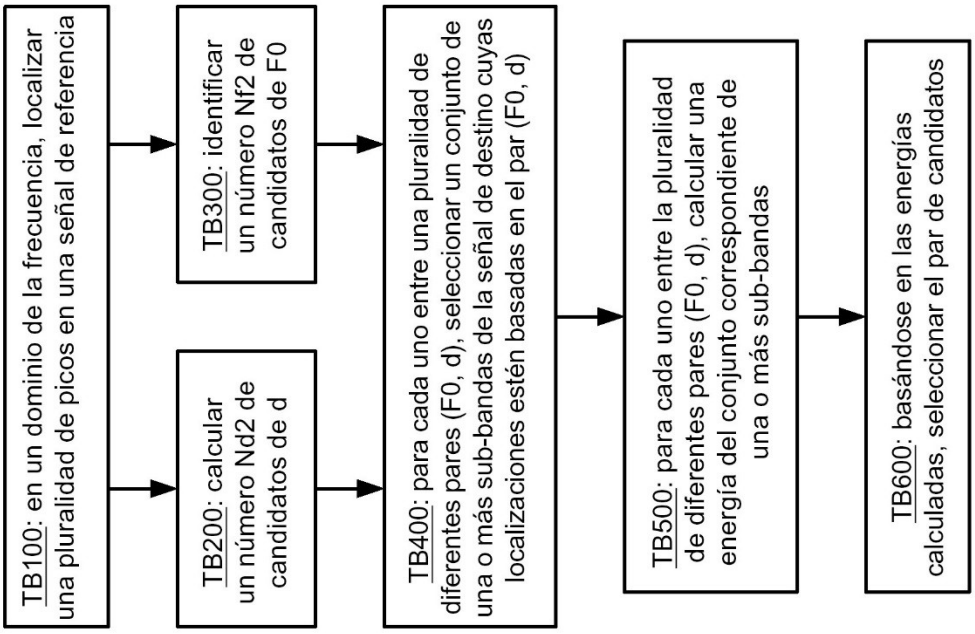


FIG. 10A

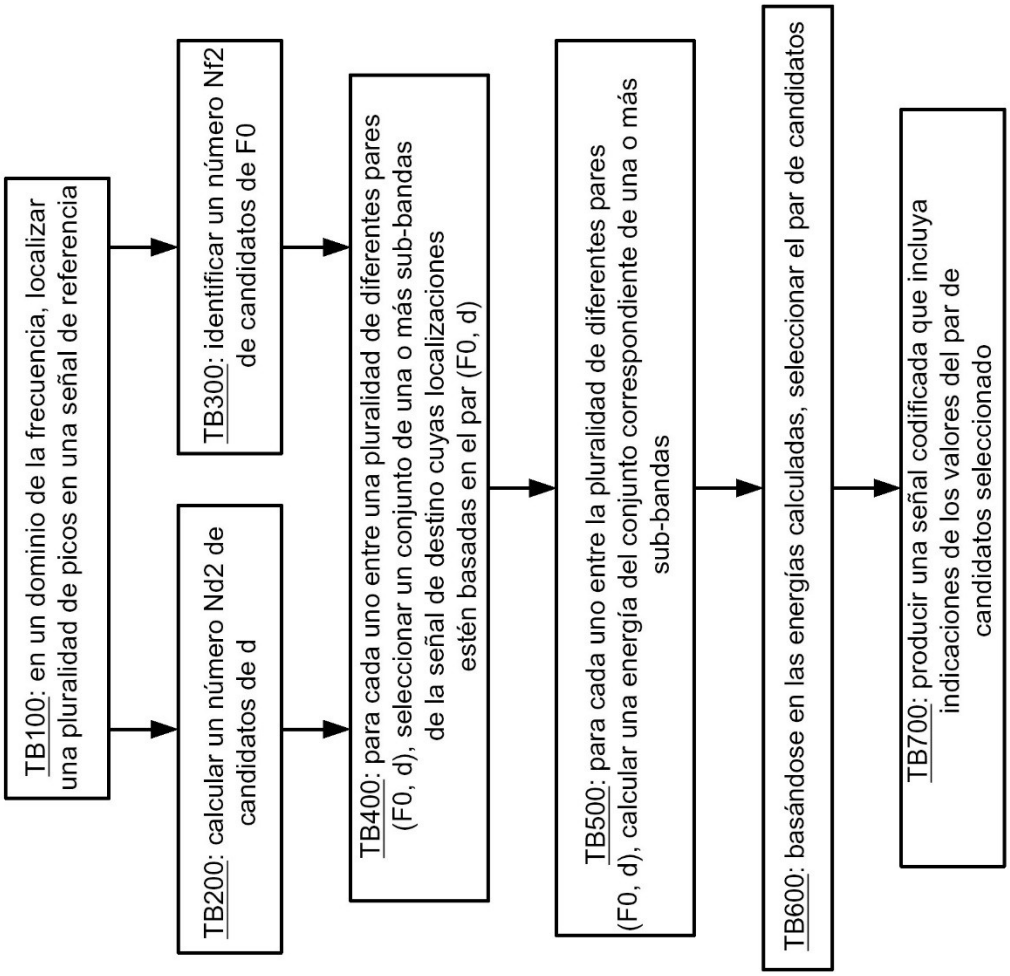


FIG. 10B

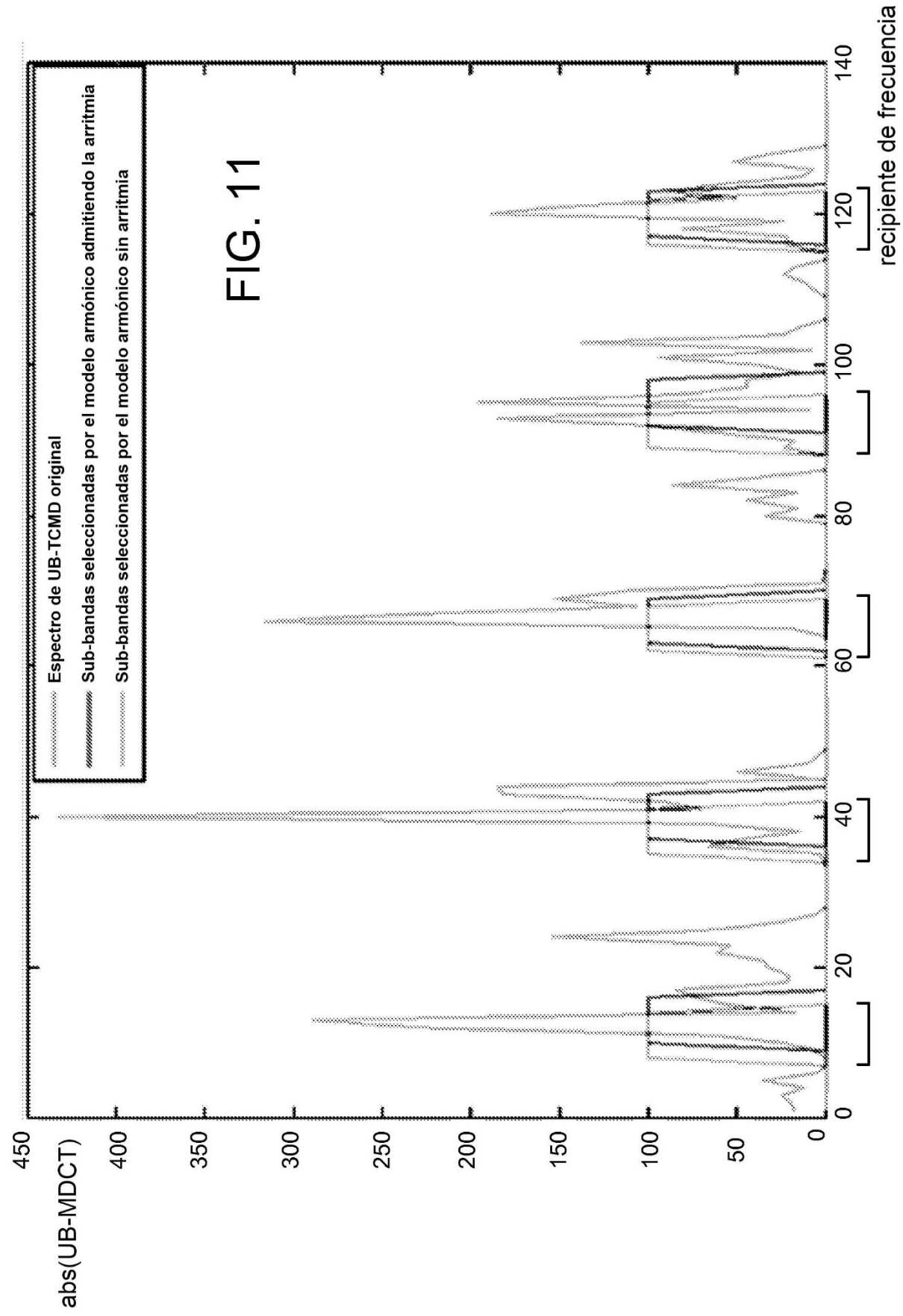

Espectro de UB-TCMD original
Sub-bandas seleccionadas por el modelo armónico admitiendo la arritmia
Sub-bandas seleccionadas por el modelo armónico sin arritmia
abs(UB-MDCT)
recipiente de frecuencia
0
50
100
150
200
250
300
350
400
450
0
20
40
60
80
100
120
140

FIG. 11

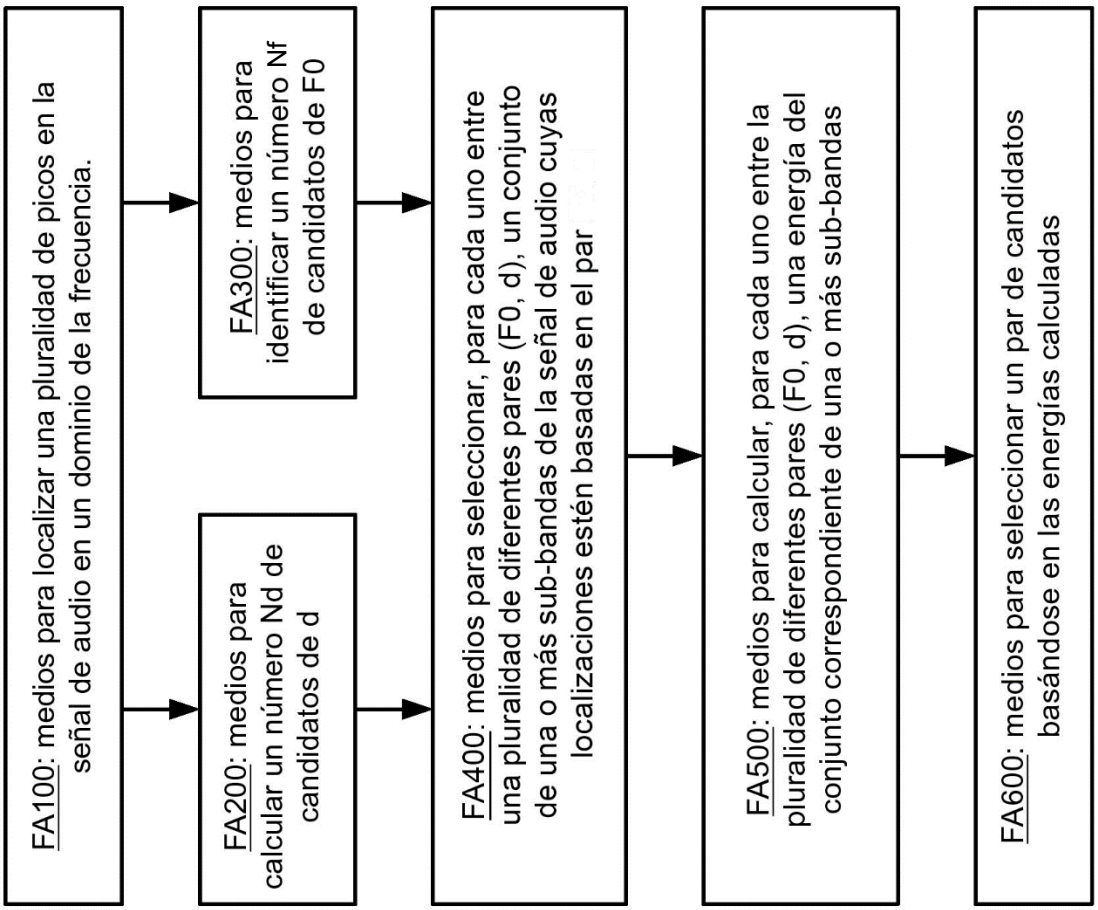


FIG. 12A

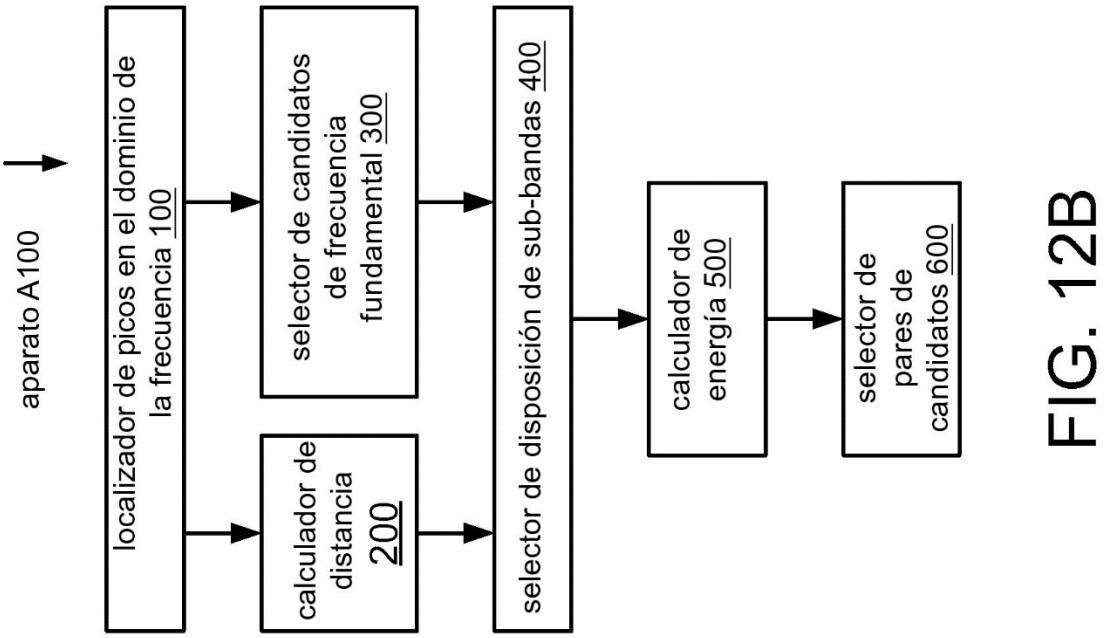


FIG. 12B

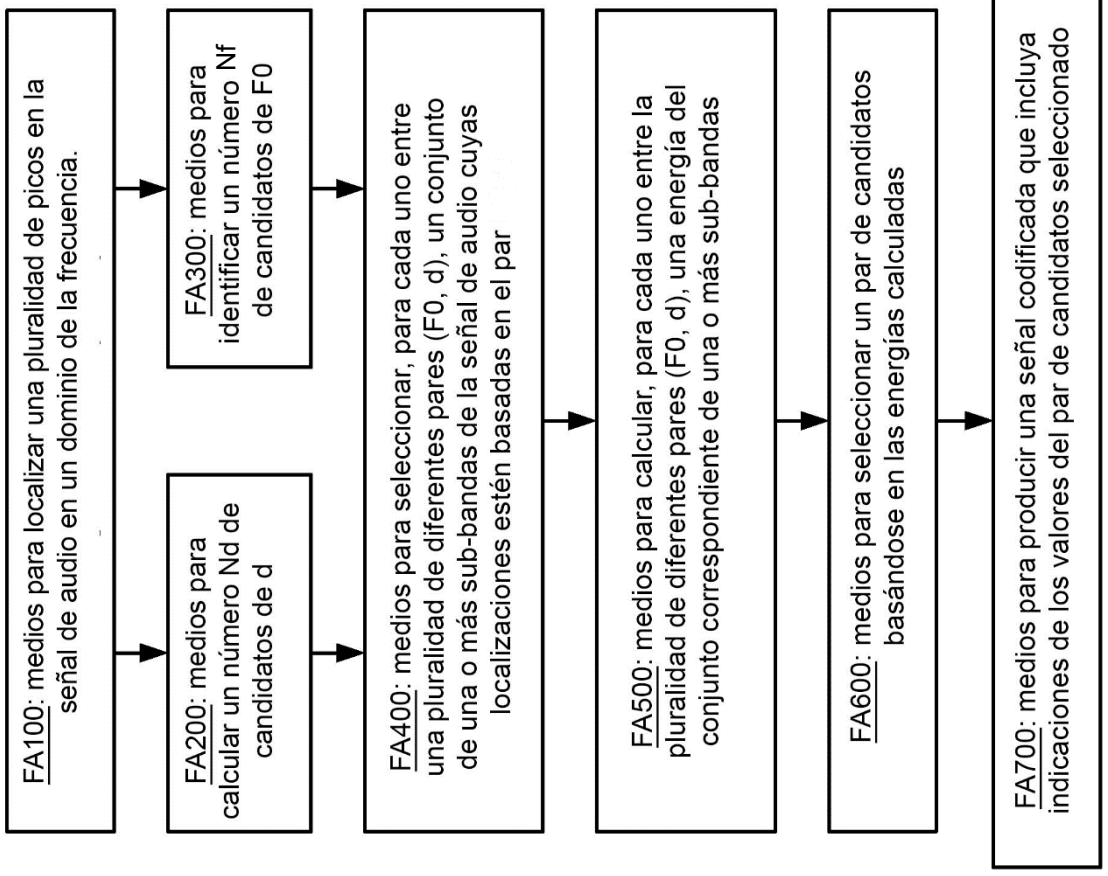


FIG. 13A

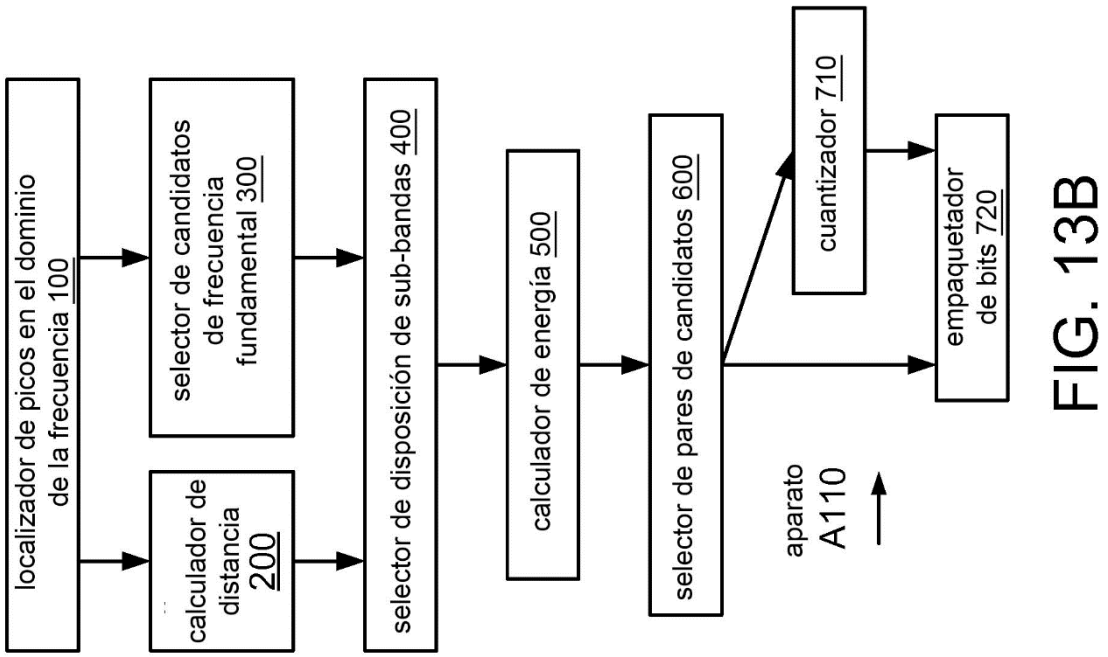
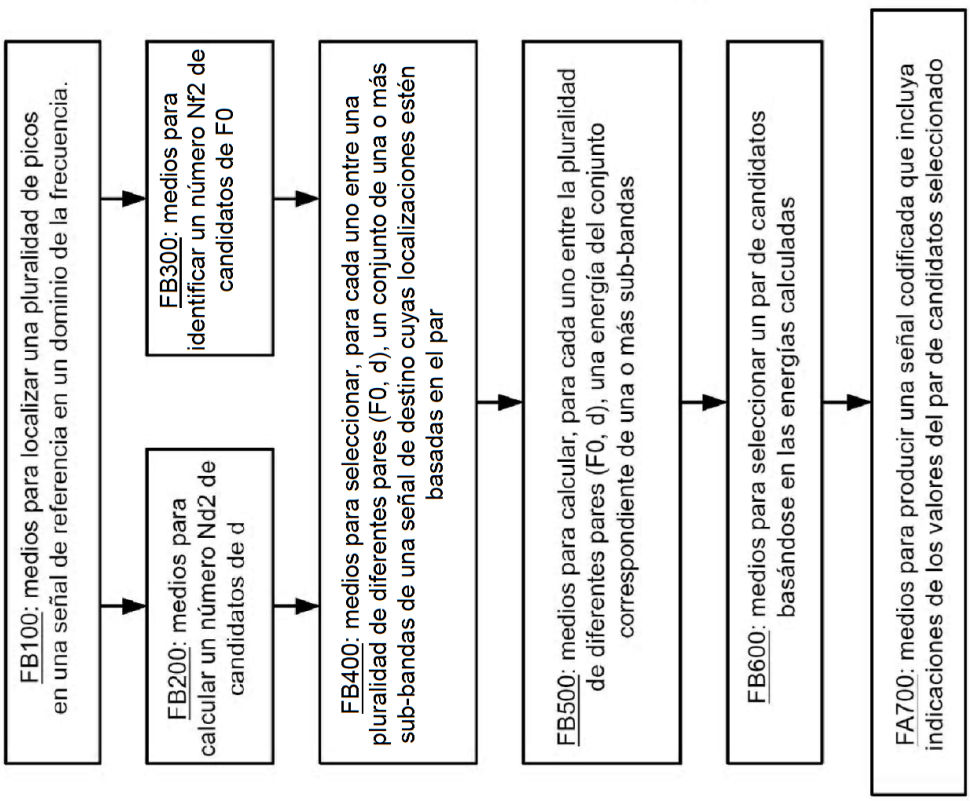


FIG. 13B

FIG. 14


FB100: medios para localizar una pluralidad de picos en una señal de referencia en un dominio de la frecuencia.
FB200: medios para calcular un número Nd2 de candidatos de d
FB300: medios para identificar un número Nf2 de candidatos de F0
FB400: medios para seleccionar, para cada uno entre una pluralidad de diferentes pares (F0, d), un conjunto de una o más sub-bandas de una señal de destino cuyas localizaciones estén basadas en el par
FB500: medios para calcular, para cada uno entre la pluralidad de diferentes pares (F0, d), una energía del conjunto correspondiente de una o más sub-bandas
aparato MF210
FB600: medios para seleccionar un par de candidatos basándose en las energías calculadas
FA700: medios para producir una señal codificada que incluya indicaciones de los valores del par de candidatos seleccionado

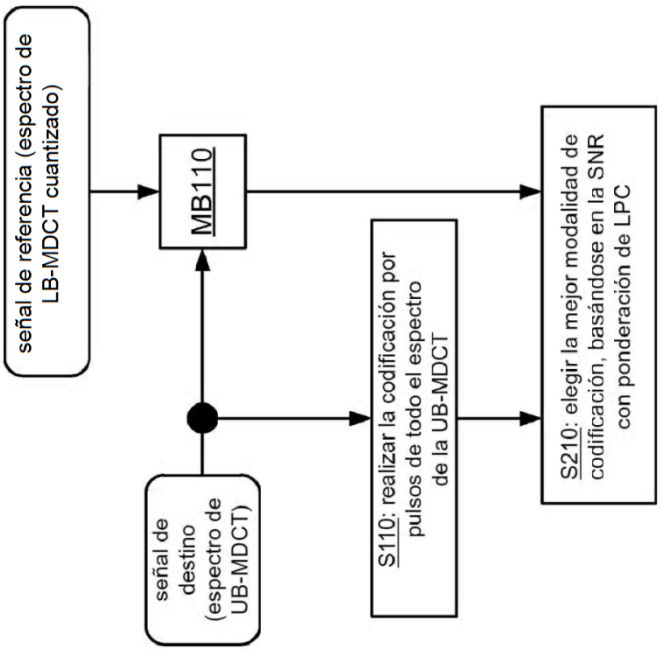


FIG. 15B

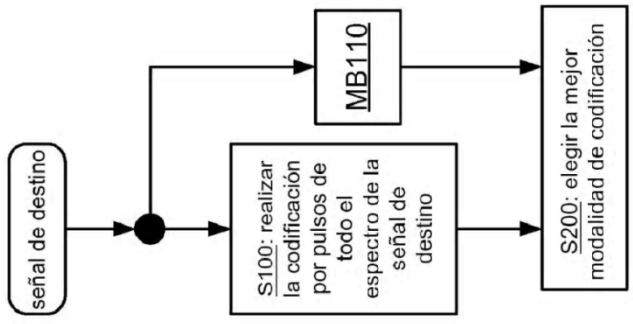


FIG. 15A

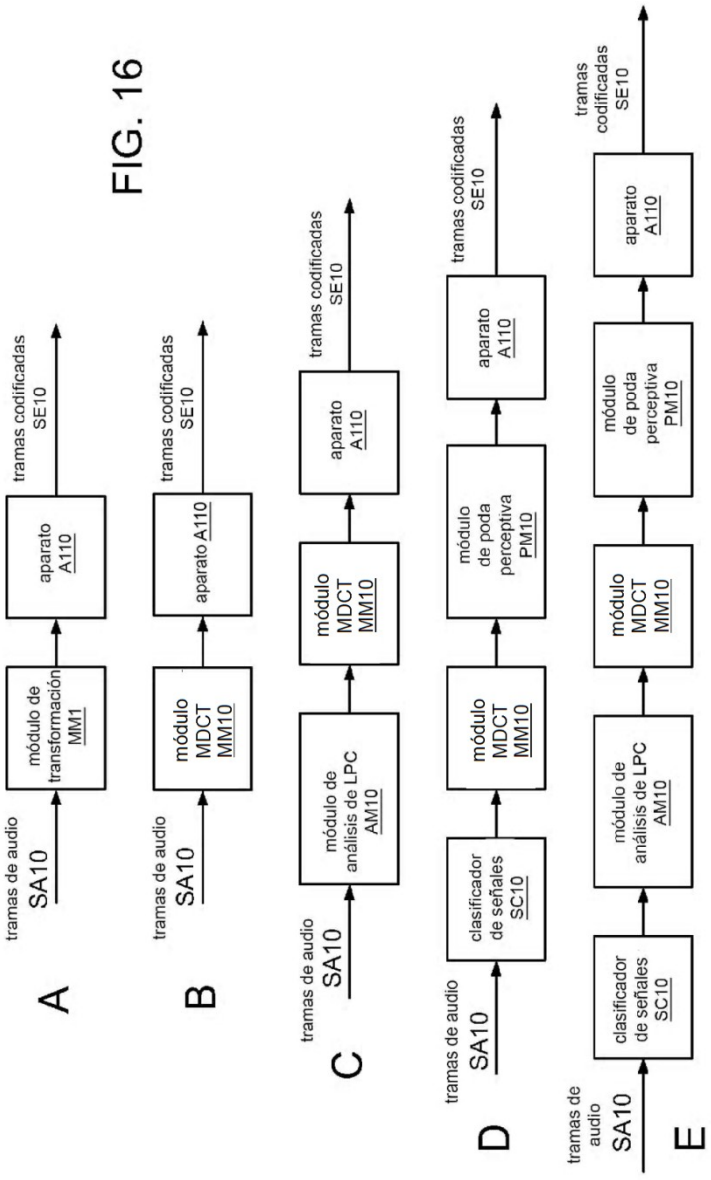

A
tramas de audio SA10
módulo de transformación MM1
aparato A110
tramas codificadas SE10
B
tramas de audio SA10
módulo MDCT MM10
aparato A110
tramas codificadas SE10
C
tramas de audio SA10
módulo de análisis de LPC AM10
módulo MDCT MM10
aparato A110
tramas codificadas SE10
D
tramas de audio SA10
clasificador de señales SC10
módulo MDCT MM10
módulo de poda perceptiva PM10
aparato A110
tramas codificadas SE10
E
tramas de audio SA10
clasificador de señales SC10
módulo de análisis de LPC AM10
módulo MDCT MM10
módulo de poda perceptiva PM10
aparato A110
tramas codificadas SE10

FIG. 16

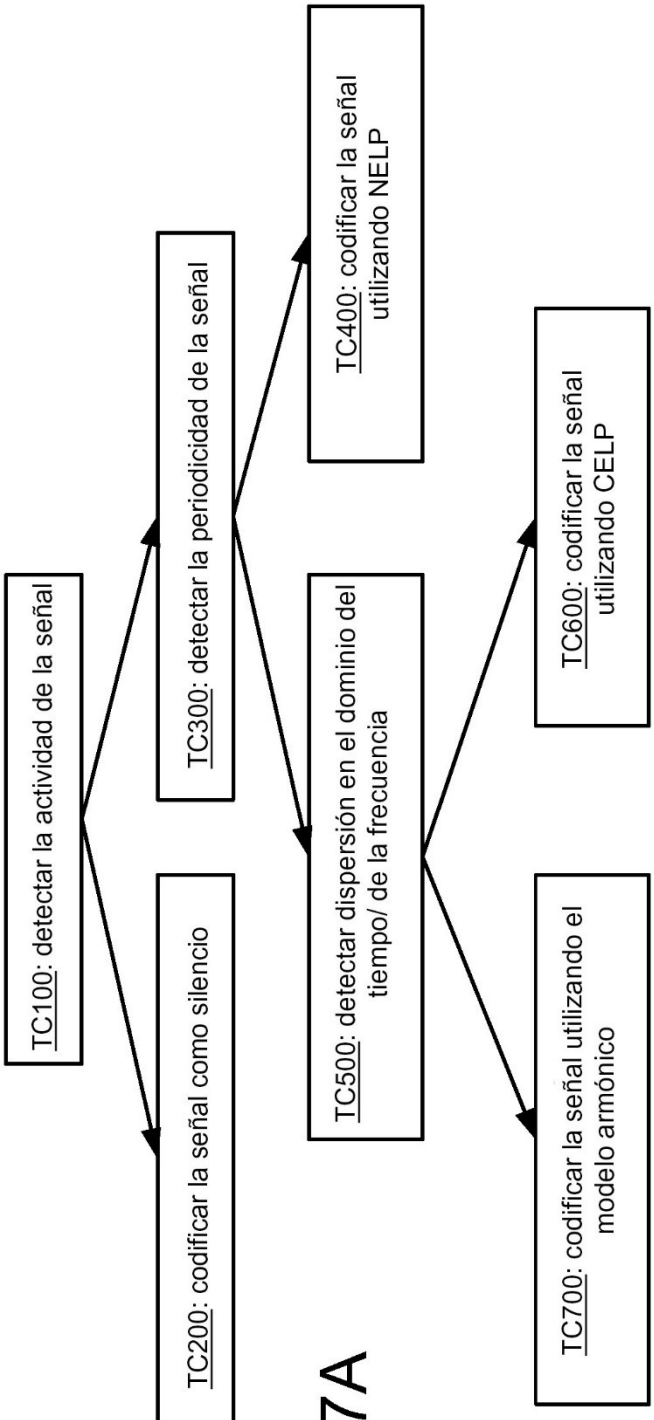


FIG. 17A

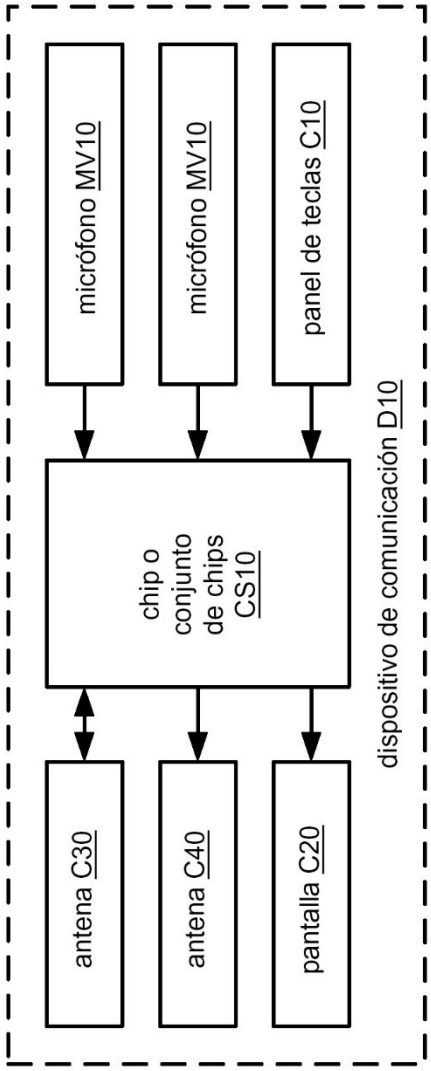


FIG. 17B

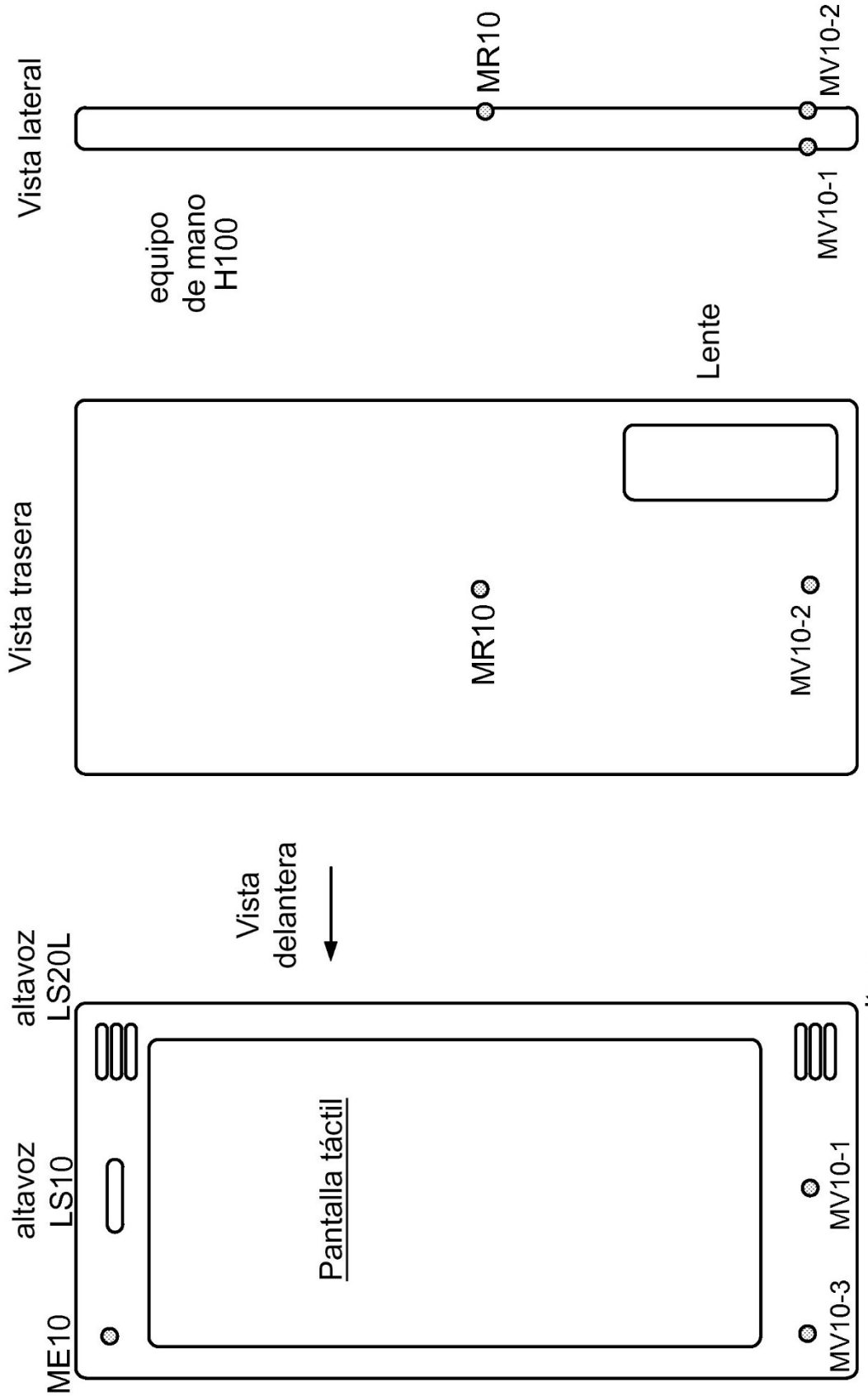
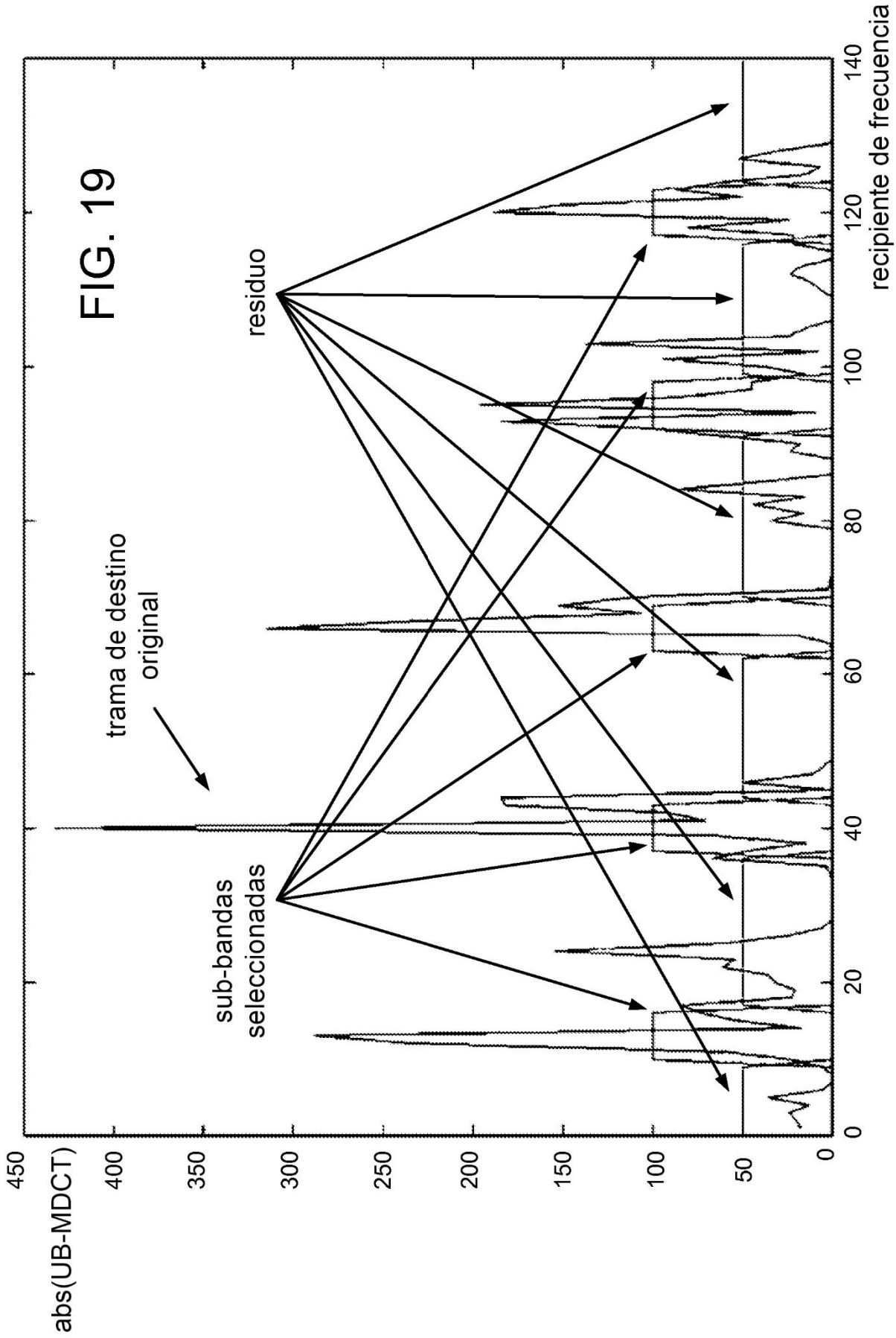

Vista lateral
equipo de mano H100
MR10
MV10-2
MV10-1
Lente
Vista trasera
MR10
MV10-2
Vista delantera
altavoz LS20L
altavoz LS10
Pantalla táctil
ME10
MV10-3
MV10-1
altavoz LS20R

FIG. 18

FIG. 19


abs(UB-MDCT)
recipiente de frecuencia
trama de destino original
sub-bandas seleccionadas
residuo