

19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 904 275**

51 Int. Cl.:

G10L 19/008	(2013.01)	G10L 19/032	(2013.01)
G10L 19/02	(2013.01)		
H04S 1/00	(2006.01)		
G10L 19/002	(2013.01)		
G10L 19/09	(2013.01)		
G10L 19/06	(2013.01)		
G10L 19/24	(2013.01)		
G10L 25/03	(2013.01)		
G10L 25/51	(2013.01)		
G10L 25/21	(2013.01)		

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

- 86 Fecha de presentación y número de la solicitud internacional: **22.09.2016 PCT/CA2016/051108**
- 87 Fecha y número de publicación internacional: **30.03.2017 WO17049399**
- 96 Fecha de presentación y número de la solicitud europea: **22.09.2016 E 16847686 (9)**
- 97 Fecha y número de publicación de la concesión europea: **01.12.2021 EP 3353780**

54 Título: **Método y sistema de decodificación de los canales izquierdo y derecho de una señal sonora estéreo**

30 Prioridad:

25.09.2015 US 201562232589 P
14.07.2016 US 201662362360 P

45 Fecha de publicación y mención en BOPI de la traducción de la patente:
04.04.2022

73 Titular/es:

VOICEAGE CORPORATION (100.0%)
750 Lucerne Road Suite 250
Town of Mount Royal, Quebec H3R 2H6, CA

72 Inventor/es:

VAILLANCOURT, TOMMY y
JELINEK, MILAN

74 Agente/Representante:

SÁEZ MAESO, Ana

ES 2 904 275 T3

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

DESCRIPCIÓN

Método y sistema de decodificación de los canales izquierdo y derecho de una señal sonora estéreo

Campo técnico

5 La presente divulgación se refiere a la codificación de sonido estéreo, en particular, pero no exclusivamente, a la codificación de voz y/o audio estéreo capaz de producir una buena calidad estéreo en una escena de audio compleja a baja tasa de bits y bajo retardo.

Antecedentes

10 Históricamente, la telefonía conversacional se ha implementado con teléfonos que tienen solo un transductor para emitir sonido solo a uno de los oídos del usuario. En la última década, los usuarios han comenzado a utilizar su teléfono portátil junto con un auricular para recibir el sonido en sus dos oídos principalmente para escuchar música pero también, a veces, para escuchar la conversación. Sin embargo, cuando se utiliza un teléfono portátil para transmitir y recibir voz conversacional, el contenido sigue siendo monofónico, pero se presenta a los dos oídos del usuario cuando se usan auriculares.

15 Con el estándar de codificación de voz 3GPP más reciente como se describe en la Referencia [1], se ha mejorado significativamente la calidad del sonido codificado, por ejemplo, voz y/o audio que se transmite y recibe a través de un teléfono portátil. La siguiente etapa natural es transmitir información estéreo de tal modo que el receptor se acerque lo más posible a una escena de audio de la vida real que se captura en el otro extremo del enlace de comunicación.

En los códecs de audio, por ejemplo, como se describe en la Referencia [2], normalmente se utiliza la transmisión de información estéreo.

20 Para los códecs de voz conversacional, la señal monofónica es la norma. Cuando se transmite una señal estereofónica, la tasa de bits a menudo debe duplicarse, ya que los canales izquierdo y derecho se codifican mediante un códec monofónico. Esto funciona bien en la mayoría de los escenarios, pero presenta los inconvenientes de duplicar la tasa de bits y no aprovechar cualquier redundancia potencial entre los dos canales (canales izquierdo y derecho). Además, para mantener la tasa de bits general a un nivel razonable, se utiliza una tasa de bits muy baja para cada canal, lo que afecta la calidad general del sonido.

30 Una posible alternativa es utilizar el llamado estéreo paramétrico como se describe en la Referencia [6]. El estéreo paramétrico envía información tal como por ejemplo, la diferencia de tiempo interaural (ITD) o las diferencias de intensidad interaural (IID). La última información se envía por banda de frecuencia y, a baja tasa de bits, el presupuesto de bits asociado a la transmisión estéreo no es lo suficientemente alto para permitir que estos parámetros funcionen de manera eficiente.

35 La transmisión de un factor de panorámica podría ayudar a crear un efecto estéreo básico a una tasa de bits baja, pero dicha técnica no hace nada para preservar el ambiente y presenta limitaciones inherentes. Una adaptación demasiado rápida del factor de panorámica resulta molesta para el oyente, a la vez que una adaptación demasiado lenta del factor de panorámica no refleja la posición real de los altavoces, lo cual dificulta la obtención de una buena calidad en caso de interferencias de los interlocutores o cuando la fluctuación del ruido de fondo es importante. Actualmente, la codificación de voz estéreo conversacional con una calidad decente para todas las posibles escenas de audio requiere una tasa de bits mínima de alrededor de 24 kb/s para señales de banda ancha (WB); por debajo de esa tasa de bits, la calidad de la voz comienza a sufrir.

40 La WO 02/23528 describe un método para decodificar señales multicanal, tal como estéreo, detectando la correlación entre canales y seleccionando uno de diversos modos de codificación posibles con base en la correlación detectada.

45 Con la globalización cada vez mayor de la fuerza laboral y la división de los equipos de trabajo en todo el mundo, existe la necesidad de mejorar las comunicaciones. Por ejemplo, los participantes de una teleconferencia pueden estar en lugares diferentes y distantes. Algunos participantes podrían estar en sus coches, otros podrían estar en una gran sala anecoica o incluso en su sala de estar. De hecho, todos los participantes desean sentir que tienen una discusión cara a cara. Implementar la voz estéreo, más en general el sonido estéreo en dispositivos portátiles, sería un gran paso en esta dirección.

Resumen

50 De acuerdo con un primer aspecto, la presente divulgación se refiere a un método de decodificación de sonido estéreo para decodificar los canales izquierdo y derecho de una señal de sonido estéreo, que comprende: recibir parámetros de codificación que comprenden parámetros de codificación de un canal primario, parámetros de codificación de un canal secundario, y un factor β , en donde los parámetros de codificación del canal primario comprenden coeficientes de filtro LP del canal primario; decodificar el canal primario en respuesta a los parámetros de codificación del canal primario para proporcionar un canal primario decodificado; decodificar el canal secundario utilizando uno de una pluralidad de modelos de codificación para proporcionar un canal secundario decodificado; en donde al menos uno de

los modelos de codificación utiliza los coeficientes de filtro LP del canal primario para decodificar el canal secundario; y el dominio del tiempo mezclando de manera ascendente los canales primario y secundario decodificados utilizando el factor β para producir los canales izquierdo y derecho decodificados de la señal de sonido estéreo, en donde el factor β determina las contribuciones respectivas de los canales primario y secundario sobre la producción de los canales izquierdo y derecho.

De acuerdo con un segundo aspecto, se proporciona un sistema de decodificación de sonido estéreo para decodificar los canales izquierdo y derecho de una señal de sonido estéreo, que comprende: medios para recibir parámetros de codificación que comprenden parámetros de codificación de un canal primario, parámetros de codificación de un canal secundario y un factor β , en donde los parámetros de codificación del canal primario comprenden coeficientes de filtro LP del canal primario; un decodificador del canal primario configurado para proporcionar un canal primario decodificado en respuesta a los parámetros de codificación del canal primario; un decodificador del canal secundario configurado para proporcionar un canal secundario decodificado utilizando uno de una pluralidad de modelos de codificación, en donde al menos uno de los modelos de codificación utiliza los coeficientes de filtro LP del canal primario para decodificar el canal secundario; y un mezclador ascendente en el dominio del tiempo de los canales primario y secundario decodificados; un mezclador ascendente en el dominio del tiempo de los canales primario y secundario decodificados configurado para utilizar el factor β para producir los canales izquierdo y derecho decodificados de la señal de sonido estéreo, en donde el factor β determina las contribuciones respectivas de los canales primario y secundario sobre la producción de los canales izquierdo y derecho.

y correctos derechos.

La presente divulgación se refiere aún más a una memoria legible por procesador que comprende instrucciones no transitorias que, cuando se ejecutan, hacen que un procesador implemente las operaciones del método descrito anteriormente.

Los anteriores y otros objetos, ventajas y características del método y sistema de decodificación de sonido estéreo para decodificar los canales izquierdo y derecho de una señal de sonido estéreo se harán más evidentes al leer la siguiente descripción no restrictiva de las realizaciones ilustrativas de la misma, dada a modo de ejemplo solo con referencia a los dibujos adjuntos.

Breve descripción de los dibujos

En los dibujos adjuntos:

La Figura 1 es un diagrama de bloques esquemático de un sistema de comunicación y procesamiento de sonido estéreo que representa un posible contexto de implementación del método y sistema de codificación de sonido estéreo como se describe en la siguiente descripción;

La Figura 2 es un diagrama de bloques que ilustra simultáneamente un método y un sistema de codificación de sonido estéreo de acuerdo con un primer modelo, presentado como un diseño estéreo integrado; este aspecto es útil para comprender la invención pero no está cubierto por las reivindicaciones;

La Figura 3 es un diagrama de bloques que ilustra simultáneamente un método y un sistema de codificación de sonido estéreo de acuerdo con un segundo modelo, presentado como un modelo integrado; este aspecto es útil para comprender la invención pero no está cubierto por las reivindicaciones;

La Figura 4 es un diagrama de bloques que muestra simultáneamente suboperaciones de una operación de mezcla descendente en el dominio del tiempo del método de codificación de sonido estéreo de las Figuras 2 y 3, y módulos de un mezclador de canales del sistema de codificación de sonido estéreo de las Figuras 2 y 3; este aspecto es útil para comprender la invención pero no está cubierto por las reivindicaciones;

La Figura 5 es un gráfico que muestra cómo una diferencia de correlación linealizada a largo plazo se asigna a un factor β y a un factor ϵ de normalización de energía;

La Figura 6 es un gráfico de curvas múltiples que muestra una diferencia entre utilizar un esquema *pca/klf* sobre un marco completo y utilizando una función de asignación de "coseno";

La Figura 7 es un gráfico de curvas múltiples que muestra un canal primario, un canal secundario y los espectros de estos canales primario y secundario resultantes de aplicar la mezcla descendente en el dominio del tiempo a una muestra estéreo que se ha grabado en una pequeña sala ecoica utilizando una configuración de micrófonos binaurales con ruido de fondo en la oficina;

La Figura 8 es un diagrama de bloques que ilustra simultáneamente un método y sistema de codificación de sonido estéreo, con una posible implementación de optimización de la codificación de los canales primario Y y X secundario de la señal de sonido estéreo; este aspecto es útil para comprender la invención pero no está cubierto por las reivindicaciones;

La Figura 9 es un diagrama de bloques que ilustra una operación de análisis de coherencia de filtro LP y el analizador de coherencia de filtro LP correspondiente del método y sistema de codificación de sonido estéreo de la Figura 8; este aspecto es útil para comprender la invención pero no está cubierto por las reivindicaciones;

5 La Figura 10 es un diagrama de bloques que ilustra simultáneamente un método de decodificación de sonido estéreo y un sistema de decodificación de sonido estéreo de acuerdo con un aspecto de la invención;

La Figura 11 es un diagrama de bloques que ilustra características adicionales del método y sistema de decodificación de sonido estéreo de la Figura 10;

10 La Figura 12 es un diagrama de bloques simplificado de una configuración de ejemplo de componentes de hardware que forman el sistema de codificación de sonido estéreo y el decodificador de sonido estéreo de la presente divulgación;

15 La Figura 13 es un diagrama de bloques que ilustra simultáneamente otras realizaciones de suboperaciones de la operación de mezcla descendente en el dominio del tiempo del método de codificación de sonido estéreo de las Figuras 2 y 3, y módulos del mezclador de canales del sistema de codificación de sonido estéreo de las Figuras 2 y 3, utilizando un factor de adaptación previa para mejorar la estabilidad de la imagen estéreo; este aspecto es útil para comprender la invención pero no está cubierto por las reivindicaciones;

La Figura 14 es un diagrama de bloques que ilustra simultáneamente las operaciones de una corrección de retardo temporal y los módulos de un corrector de retardo temporal; este aspecto es útil para comprender la invención pero no está cubierto por las reivindicaciones;

20 La Figura 15 es un diagrama de bloques que ilustra simultáneamente un método y un sistema de codificación de sonido estéreo alternativo; este aspecto es útil para comprender la invención pero no está cubierto por las reivindicaciones;

La Figura 16 es un diagrama de bloques que ilustra simultáneamente suboperaciones de un análisis de coherencia de tono y módulos de un analizador de coherencia de tono; este aspecto es útil para comprender la invención pero no está cubierto por las reivindicaciones;

25 La Figura 17 es un diagrama de bloques que ilustra simultáneamente un método y un sistema de codificación estéreo que utiliza una mezcla descendente en el dominio del tiempo con la capacidad de operar en el dominio del tiempo y en el dominio de la frecuencia; y este aspecto es útil para comprender la invención pero no está cubierto por las reivindicaciones;

30 La Figura 18 es un diagrama de bloques que ilustra simultáneamente otro método y sistema de codificación estéreo que utiliza mezcla descendente en el dominio del tiempo con la capacidad de operar en el dominio del tiempo y en el dominio de la frecuencia. Este aspecto es útil para comprender la invención pero no está cubierto por las reivindicaciones;

Descripción detallada

35 La presente divulgación se refiere a la producción y transmisión, con una tasa de bits baja y un retardo bajo, de una representación realista de contenido de sonido estéreo, por ejemplo, contenido de voz y/o audio, a partir de, en particular, pero no exclusivamente, una escena de audio compleja. Una escena de audio compleja incluye situaciones en las cuales (a) la correlación entre las señales de sonido que son grabadas por los micrófonos es baja, (b) hay una fluctuación importante del ruido de fondo, y/o (c) hay un interlocutor presente. Ejemplos de escenas de audio complejas comprenden una gran sala de conferencias anecoica con una configuración de micrófonos A/B, una pequeña sala ecoica con micrófonos binaurales, y una pequeña sala ecoica con una configuración de micrófonos mono/laterales. Todas estas configuraciones de sala podrían incluir fluctuaciones de ruido de fondo y/o interferencias de personas.

40 Los códecs de sonido estéreo conocidos, tales como 3GPP AMR-WB+ como se describe en la Referencia [7], cuyo contenido completo se incorpora en el presente documento como referencia, son ineficaces para codificar el sonido que no se acerca al modelo monofónico, especialmente a baja tasa de bits. Ciertos casos son particularmente difíciles de codificar utilizando técnicas estéreo existentes. Dichos casos incluyen:

- LAAB (Sala anecoica grande con configuración de micrófonos A/B);
- SEBI (Sala ecoica pequeña con configuración de micrófonos binaurales); y
- SEMS (Sala ecoica pequeña con configuración de micrófonos mono/laterales).

50 La adición de un ruido de fondo fluctuante y/o interferencias de los interlocutores hace que estas señales de sonido sean aún más difíciles de codificar a una tasa de bits baja utilizando técnicas dedicadas al estéreo, tal como el estéreo paramétrico. Una alternativa para codificar dichas señales es utilizar dos canales monofónicos, duplicando así la tasa de bits y el ancho de banda de la red que se está usando.

El último estándar de voz conversacional 3GPP EVS proporciona un rango de tasa de bits de 7.2 kb/s a 96 kb/s para operación de banda ancha (WB) y de 9.6 kb/s a 96 kb/s para operación de banda súper ancha (SWB). Esto significa que las tres tasas de bits mono dual más bajas que utilizan EVS son 14.4, 16.0 y 19.2 kb/s para la operación WB y 19.2, 26.3 y 32.8 kb/s para la operación SWB. Aunque la calidad de voz del 3GPP AMR-WB desplegado como se describe en la Referencia [3], cuyo contenido completo se incorpora en el presente documento como referencia, mejora con respecto a su códec predecesor, la calidad de la voz codificada a 7.2 kb/s en un entorno ruidoso es lejos de ser transparente y, por lo tanto, se puede anticipar que la calidad de voz del mono dual a 14.4 kb/s también sería limitada. A dichas tasas de bits bajas, el uso de la tasa de bits se maximiza de tal modo que se obtenga la mejor calidad de voz posible con la mayor frecuencia posible. Con el método y el sistema de codificación de sonido estéreo como se divulga en la siguiente descripción, la tasa de bits total mínima para el contenido de voz estéreo conversacional, incluso en el caso de escenas de audio complejas, debe ser de alrededor de 13 kb/s para WB y 15.0 kb/s para SWB. A tasas de bits inferiores a las utilizadas en un enfoque mono dual, la calidad y la inteligibilidad de la voz estéreo se mejoran enormemente para escenas de audio complejas.

La Figura 1 es un diagrama de bloques esquemático de un sistema 100 de comunicación y procesamiento de sonido estéreo que representa un posible contexto de implementación del método y sistema de codificación de sonido estéreo como se divulga en la siguiente descripción.

El sistema 100 de comunicación y procesamiento de sonido estéreo de la Figura 1 soporta la transmisión de una señal de sonido estéreo a través de un enlace 101 de comunicación. El enlace 101 de comunicación puede comprender, por ejemplo, un cable o un enlace de fibra óptica. De manera alternativa, el enlace 101 de comunicación puede comprender al menos en parte un enlace de radiofrecuencia. El enlace de radiofrecuencia a menudo admite comunicaciones múltiples y simultáneas que requieren recursos de ancho de banda compartido, tales como los que se pueden encontrar con la telefonía celular. Aunque no se muestra, el enlace 101 de comunicación puede ser reemplazado por un dispositivo de almacenamiento en una implementación de dispositivo único del sistema 100 de procesamiento y comunicación que graba y almacena la señal de sonido estéreo codificada para su posterior reproducción.

Aún con referencia a la Figura 1, por ejemplo, un par de micrófonos 102 y 122 produce los canales izquierdo 103 y derecho 123 de una señal de sonido estéreo analógica original detectada, por ejemplo, en una escena de audio compleja. Como se indica en la descripción anterior, la señal de sonido puede comprender, en particular, pero no exclusivamente, voz y/o audio. Los micrófonos 102 y 122 pueden disponerse de acuerdo con una configuración A/B, binaural o mono/lateral.

Los canales izquierdo 103 y derecho 123 de la señal de sonido analógica original se suministran a un convertidor 104 de analógico a digital (A/D) para convertirlos en canales izquierdo 105 y derecho 125 de una señal de sonido estéreo digital original. Los canales izquierdo 105 y derecho 125 de la señal de sonido estéreo digital original también pueden grabarse y suministrarse a partir de un dispositivo de almacenamiento (no se muestra).

Un codificador 106 de sonido estéreo codifica los canales izquierdo 105 y derecho 125 de la señal de sonido estéreo digital produciendo así un conjunto de parámetros de codificación que se multiplexan bajo la forma de un flujo de bits 107 entregado a un codificador 108 de corrección de errores opcional. El codificador 108 de corrección de errores opcional, cuando está presente, agrega redundancia a la representación binaria de los parámetros de codificación en el flujo de bits 107 antes de transmitir el flujo de bits 111 resultante a través del enlace 101 de comunicación.

En el lado del receptor, un decodificador 109 opcional de corrección de errores utiliza la información redundante mencionada anteriormente en el flujo de bits 111 digital recibido para detectar y corregir errores que pueden haber ocurrido durante la transmisión sobre el enlace 101 de comunicación, produciendo un flujo de bits 112 con parámetros de codificación recibidos. Un decodificador 110 de sonido estéreo convierte los parámetros de codificación recibidos en el flujo de bits 112 para crear canales izquierdo 113 y derecho 133 sintetizados de la señal de sonido estéreo digital. Los canales izquierdo 113 y derecho 133 de la señal de sonido estéreo digital reconstruida en el decodificador 110 de sonido estéreo se convierten en los canales izquierdo 114 y derecho 134 sintetizados de la señal de sonido estéreo analógico en un convertidor 115 de digital a analógico (D/A).

Los canales izquierdo 114 y derecho 134 sintetizados de la señal de sonido estéreo analógico se reproducen respectivamente en un par de unidades 116 y 136 de altavoz. De manera alternativa, los canales izquierdo 113 y derecho 133 de la señal de sonido estéreo digital procedente del decodificador 110 de sonido estéreo también se pueden suministrar y grabar en un dispositivo de almacenamiento (no se muestra).

Los canales izquierdo 105 y derecho 125 de la señal de sonido estéreo digital original de la Figura 1 corresponden a los canales izquierdo L y derecho R de las Figuras 2, 3, 4, 8, 9, 13, 14, 15, 17 y 18. Además, el codificador 106 de sonido estéreo de la Figura 1 corresponde al sistema de codificación de sonido estéreo de las Figuras 2, 3, 8, 15, 17 y 18.

El método y sistema de codificación de sonido estéreo de acuerdo con la presente divulgación son dobles; se proporcionan los primeros y segundos modelos.

La Figura 2 es un diagrama de bloques que ilustra simultáneamente el método y el sistema de codificación de sonido estéreo de acuerdo con el primer modelo, presentado como un diseño estéreo integrado con base en el núcleo EVS. Esta figura y su descripción son útiles para comprender la invención pero no están cubiertas por las reivindicaciones.

5 Con referencia a la Figura 2, el método de codificación de sonido estéreo de acuerdo con el primer modelo comprende una operación 201 de mezcla descendente en el dominio del tiempo, una operación 202 de codificación del canal primario, una operación 203 de codificación del canal secundario, y una operación 204 de multiplexación.

Para realizar la operación 201 de mezcla descendente en el dominio del tiempo, un mezclador 251 de canales mezcla los dos canales estéreo de entrada (canal derecho R y canal izquierdo L) para producir un canal primario Y y un canal secundario X.

10 Para llevar a cabo la operación 203 de codificación del canal secundario, un codificador 253 de canal secundario selecciona y utiliza un número mínimo de bits (tasa de bits mínima) para codificar el canal secundario X utilizando uno de los modos de codificación como se define en la siguiente descripción y produce un flujo de bits 206 codificado del canal secundario correspondiente. El presupuesto de bits asociado puede cambiar cada trama dependiendo del contenido de la trama.

15 Para implementar la operación 202 de codificación del canal primario, se utiliza un codificador 252 del canal primario. El codificador 253 de canal secundario señala al codificador 252 de canal primario el número de bits 208 utilizados en la trama actual para codificar el canal secundario X. Se puede utilizar cualquier tipo adecuado de codificador como codificador 252 de canal primario. Como un ejemplo no limitativo, el codificador 252 de canal primario puede ser un codificador de tipo CELP. En esta realización ilustrativa, el codificador de tipo CELP de canal primario es una versión
20 modificada del codificador EVS heredado, donde el codificador EVS se modifica para presentar una mayor escalabilidad de la tasa de bits para permitir una asignación flexible de la tasa de bits entre los canales primario y secundario. De esta manera, el codificador EVS modificado podrá utilizar todos los bits que no se utilizan para codificar el canal secundario X para codificar, con una tasa de bits correspondiente, el canal primario Y y producir un correspondiente flujo de bits 205 codificado del canal primario.

25 Un multiplexor 254 concatena el flujo de bits 205 del canal primario y el flujo de bits 206 del canal secundario para formar un flujo de bits 207 multiplexado, para completar la operación 204 de multiplexación.

En el primer modelo, el número de bits y la tasa de bits correspondiente (en el flujo de bits 206) utilizados para codificar el canal secundario X es menor que el número de bits y la tasa de bits correspondiente (en el flujo de bits 205) utilizados para codificar el canal primario Y. Esto puede verse como dos (2) canales de tasa de bits variable en donde la suma
30 de las tasas de bits de los dos canales X y Y representa una tasa de bits total constante. Este enfoque puede tener diferentes sabores con más o menos énfasis en el canal primario Y. De acuerdo con un primer ejemplo, cuando se pone un énfasis máximo en el canal primario Y, el presupuesto de bits del canal secundario X se fuerza agresivamente a un mínimo. De acuerdo con un segundo ejemplo, si se pone menos énfasis en el canal primario Y, entonces el presupuesto de bits para el canal secundario X puede hacerse más constante, lo que significa que la tasa de bits promedio del canal secundario X es ligeramente mayor en comparación con el primer ejemplo.

35 Se recuerda que los canales derecho R e izquierdo L de la señal de sonido estéreo digital de entrada se procesan mediante tramas sucesivas de una duración determinada la cual puede corresponder a la duración de las tramas utilizadas en el procesamiento EVS. Cada trama comprende un número de muestras de los canales derecho R e izquierdo L dependiendo de la duración dada de la trama y la frecuencia de muestreo que se esté utilizando.

40 La Figura 3 es un diagrama de bloques que ilustra simultáneamente el método y el sistema de codificación de sonido estéreo de acuerdo con el segundo modelo, presentado como un modelo integrado.

Con referencia a la Figura 3, el método de codificación de sonido estéreo de acuerdo con el segundo modelo comprende una operación 301 de mezcla descendente en el dominio del tiempo, una operación 302 de codificación del canal primario, una operación 303 de codificación del canal secundario, y una operación 304 de multiplexación.
45 Esta figura y su descripción son útiles para comprender la invención pero no están cubiertas por las reivindicaciones.

Para completar la operación 301 de mezcla descendente en el dominio del tiempo, un mezclador 351 de canales mezcla los dos canales de entrada derecho R e izquierdo L para formar un canal primario Y y un canal secundario X.

En la operación 302 de codificación de canal primario, un codificador 352 de canal primario codifica el canal primario Y para producir un flujo de bits 305 codificado de canal primario. De nuevo, se puede utilizar cualquier tipo de
50 codificador adecuado como codificador 352 de canal primario. Como un ejemplo no limitativo, el codificador 352 de canal primario puede ser un codificador de tipo CELP. En esta realización ilustrativa, el codificador 352 de canal primario utiliza un estándar de codificación de voz tal como el modo de codificación mono EVS heredado o el modo de codificación AMR-WB-IO, por ejemplo, lo que significa que la porción monofónica del flujo de bits 305 sería interoperable con el EVS heredado, el decodificador AMR-WB-IO o el decodificador AMR-WB heredado cuando la tasa de bits es compatible con dicho decodificador. Dependiendo del modo de codificación que se seleccione, puede
55 ser necesario algún ajuste del canal primario Y para el procesamiento a través del codificador 352 del canal primario.

En la operación 303 de codificación de canal secundario, un codificador 353 de canal secundario codifica el canal secundario X a una tasa de bits más baja utilizando uno de los modos de codificación como se define en la siguiente descripción. El codificador 353 de canal secundario produce un flujo de bits 306 codificado de canal secundario.

5 Para realizar la operación 304 de multiplexación, un multiplexor 354 concatena el flujo de bits 305 codificado del canal primario con el flujo de bits 306 codificado del canal secundario para formar un flujo de bits 307 multiplexado. Esto se denomina modelo integrado, porque el flujo de bits 306 codificado del canal secundario asociado al estéreo se agrega sobre un flujo de bits 305 interoperable. El flujo de bits 306 del canal secundario puede eliminarse del flujo de bits 307 estéreo multiplexado (flujos de bits 305 y 306 concatenados) en cualquier momento, lo que da como resultado un flujo de bits decodificable por un códec heredado como se describe anteriormente en el presente documento, a la vez que
10 un usuario de una versión más reciente del códec todavía podría disfrutar de la decodificación estéreo completa.

De hecho, los primeros y segundos modelos descritos anteriormente están próximos entre sí. La principal diferencia entre los dos modelos es la posibilidad de utilizar una asignación dinámica de bits entre los dos canales Y y X en el primer modelo, a la vez que la asignación de bits es más limitada en el segundo modelo debido a consideraciones de interoperabilidad.

15 En la siguiente descripción se dan ejemplos de implementación y enfoques utilizados para lograr los primeros y segundos modelos descritos anteriormente.

1) *Mezcla descendente en el dominio del tiempo* (El siguiente aspecto completo es útil para comprender la invención, pero no está cubierto por las reivindicaciones)

20 Como se expresa en la descripción anterior, los modelos estéreo conocidos que operan a baja tasa de bits tienen dificultades con la codificación de la voz que no se acerca al modelo monofónico. Los enfoques tradicionales realizan una mezcla descendente en el dominio de la frecuencia, por banda de frecuencia, utilizando, por ejemplo, una correlación por banda de frecuencia asociada con un Análisis de Componentes Principales (*pca*) utilizando, por ejemplo, una Transformada de Karhunen-Loeve (*klt*), para obtener dos vectores, como se describe en las referencias [4] y [5], cuyo contenido completo se incorpora en el presente documento como referencia. Uno de estos dos vectores
25 incorpora todo el contenido altamente correlacionado a la vez que el otro vector define todo el contenido que no está muy correlacionado. El método más conocido para codificar voz a tasas de bits bajas utiliza un códec en el dominio del tiempo, tal como un códec CELP (Predicción Lineal Excitada por Código), en el cual las soluciones conocidas en el dominio de la frecuencia no son directamente aplicables. Por esa razón, a la vez que la idea detrás del *pca/klt* por banda de frecuencia es interesante, cuando el contenido es voz, el canal primario Y debe volver a convertirse al dominio del tiempo y, después de dicha conversión, su contenido ya no se parece a la voz tradicional, especialmente
30 en el caso de las configuraciones descritas anteriormente que utilizan un modelo específico de la voz como CELP. Esto tiene el efecto de reducir el rendimiento del códec de voz. Además, a una tasa de bits baja, la entrada de un códec de voz debe estar lo más cerca posible de las expectativas del modelo interno del códec.

35 Partiendo de la idea de que una entrada de un códec de voz de baja tasa de bits debería estar lo más cerca posible de la señal de voz esperada, se ha desarrollado una primera técnica. La primera técnica se basa en una evolución del esquema tradicional *pca/klt*. A la vez que el esquema tradicional calcula la *pca/klt* por banda de frecuencia, la primera técnica la calcula en toda la trama, directamente en el dominio del tiempo. Esto funciona adecuadamente durante los segmentos de voz activos, siempre que no haya ruido de fondo ni interferencias del hablante. El esquema *pca/klt* determina qué canal (canal izquierdo L o derecho R) contiene la información más útil, este canal se envía al codificador
40 del canal primario. Desafortunadamente, el esquema *pca/klt* con base en la trama no es confiable en presencia de ruido de fondo o cuando dos o más personas están hablando entre sí. El principio del esquema *pca/klt* implica la selección de un canal de entrada (R o L) u otro, lo que a menudo conduce a cambios drásticos en el contenido del canal primario que se va a codificar. Al menos por las razones anteriores, la primera técnica no es suficientemente fiable y, por consiguiente, en el presente documento se presenta una segunda técnica para superar las deficiencias de la primera técnica y permitir una transición más suave entre los canales de entrada. Esta segunda técnica se describirá a continuación con referencia a las Figuras 4-9.

45 Con referencia a la Figura 4, la operación de mezcla 201/301 descendente en el dominio del tiempo (Figuras 2 y 3) comprende las siguientes suboperaciones: una suboperación 401 de análisis de energía, una suboperación 402 de análisis de tendencias de energía, una suboperación 403 de análisis de correlación normalizada del canal L y R, una suboperación 404 que calcula la diferencia de correlación a largo plazo (LT), una suboperación 405 de conversión y cuantificación de diferencia de correlación a largo plazo al factor β y una suboperación 406 de mezcla descendente en el dominio del tiempo.

55 Teniendo en cuenta la idea de que la entrada de un códec de sonido de baja tasa de bits (tal como voz y/o audio) debe ser lo más homogénea posible, la suboperación 401 de análisis de energía se realiza en el mezclador 252/351 de canales mediante un 451 analizador de energía para determinar primero, por trama, el *rms* (Valor Cuadrático Medio) de la energía de cada canal de entrada R y L utilizando las relaciones (1):

$$rms_L(t) = \sqrt{\frac{\sum_{i=0}^{N-1} L(i)^2}{N}}; \quad rms_R(t) = \sqrt{\frac{\sum_{i=0}^{N-1} R(i)^2}{N}}, \quad (1)$$

donde los subíndices L y R representan los canales izquierdo y derecho respectivamente, $L(i)$ representa la muestra i del canal L, $R(i)$ representa la muestra i del canal R, N corresponde al número de muestras por trama, y t representa una trama actual.

5 A continuación, el 451 analizador de energía utiliza los valores rms de las relaciones (1) para determinar a largo plazo rms valores \overline{rms} para cada canal utilizando relaciones (2):

$$\overline{rms}_L(t) = 0.6 \cdot \overline{rms}_L(t_{-1}) + 0.4 \cdot rms_L; \quad \overline{rms}_R(t) = 0.6 \cdot \overline{rms}_R(t_{-1}) + 0.4 \cdot rms_R, \quad (2)$$

donde t representa la trama actual y t_{-1} la trama anterior.

10 Para realizar la suboperación 402 de análisis de tendencia de energía, un analizador 452 de tendencia de energía del mezclador 251/351 de canales utiliza el rms de largo plazo valores \overline{rms} para determinar la tendencia de la energía en cada canal L y R rms_{dt} utilizando relaciones (3):

$$\overline{rms}_{dt}_L = \overline{rms}_L(t) - \overline{rms}_L(t_{-1}); \quad \overline{rms}_{dt}_R = \overline{rms}_R(t) - \overline{rms}_R(t_{-1}). \quad (3)$$

15 La tendencia de los valores rms de largo plazo se utilizan como información que muestra si los eventos temporales capturados por los micrófonos se están desvaneciendo o si están cambiando de canal. Los valores rms de largo plazo y su tendencia también se utilizan para determinar una tasa de convergencia α de una diferencia de correlación a largo plazo como se describirá a continuación en el presente documento.

Para realizar la suboperación 403 de análisis de correlación normalizada de los canales L y R, un analizador 453 de correlación normalizada L y R calcula una correlación G_{LR} para cada uno de los canales izquierdo L y derecho R normalizados frente a una versión de señal monofónica $m(i)$ del sonido, tal como voz y/o audio, en la trama t utilizando relaciones (4):

$$20 \quad G_L(t) = \frac{\sum_{i=0}^{N-1} (L(i) \cdot m(i))}{\sum_{i=0}^{N-1} m(i)^2}, \quad G_R(t) = \frac{\sum_{i=0}^{N-1} (R(i) \cdot m(i))}{\sum_{i=0}^{N-1} m(i)^2}, \quad m(i) = \left(\frac{L(i) + R(i)}{2} \right), \quad (4)$$

25 donde N , como ya se mencionó, corresponde al número de muestras en una trama, y t representa la trama actual. En la realización actual, todas las correlaciones normalizadas y los valores rms determinados por las relaciones 1 a 4 se calculan en el dominio del tiempo, para toda la trama. En otra configuración posible, estos valores se pueden calcular en el dominio de la frecuencia. Por ejemplo, las técnicas descritas en el presente documento, las cuales están adaptadas a señales de sonido que tienen características de voz, pueden ser parte de un marco más amplio el cual puede cambiar entre un método de codificación de audio estéreo genérico en el dominio de frecuencia y el método descrito en la presente divulgación. En este caso, calcular las correlaciones normalizadas y los valores rms en el dominio de la frecuencia pueden presentar alguna ventaja en términos de complejidad o reutilización del código.

30 Para calcular la diferencia de correlación a largo plazo (LT) en la suboperación 404, una calculadora 454 calcula para cada canal L y R en la trama actual correlaciones normalizadas suavizadas utilizando relaciones (5):

$$\overline{G}_L(t) = \alpha \cdot \overline{G}_L(t_{-1}) + (1-\alpha) \cdot G_L(t) \quad \text{y} \quad \overline{G}_R(t) = \alpha \cdot \overline{G}_R(t_{-1}) + (1-\alpha) \cdot G_R(t), \quad (5)$$

donde α es la tasa de convergencia mencionada anteriormente. Finalmente, la calculadora 454 determina la diferencia de correlación a largo plazo (LT) \overline{G}_{LR} utilizando la relación (6):

$$\overline{G}_{LR}(t) = \overline{G}_L(t) - \overline{G}_R(t). \quad (6)$$

35 En una realización de ejemplo, la tasa de convergencia α puede tener un valor de 0.8 o 0.5 dependiendo de las energías a largo plazo calculadas en las relaciones (2) y la tendencia de las energías a largo plazo calculadas en las relaciones (3). Por ejemplo, la tasa de convergencia α puede tener un valor de 0.8 cuando las energías a largo plazo de los canales izquierdo L y derecho R evolucionan en la misma dirección, una diferencia entre la diferencia de correlación a largo plazo \overline{G}_{LR} en la trama t y la diferencia de correlación a largo plazo G_{LR} en la trama t_{-1} es baja (por debajo de 0.31 para esta realización de ejemplo), y al menos uno de los valores rms a largo plazo de los canales izquierdo L y R derecho está por encima de un cierto umbral (2000 en esta realización de ejemplo). Dichos casos significan que ambos canales L y R están evolucionando sin problemas, no hay un cambio rápido de energía de un canal a otro, y al menos un canal contiene un nivel significativo de energía. De lo contrario, cuando las energías a largo plazo de los canales derecho R e izquierdo L evolucionan en diferentes direcciones, cuando la diferencia entre las diferencias de correlación a largo plazo es alta, o cuando los dos canales derecho R e izquierdo L tienen bajas energías, entonces α se establecerá en 0.5 para aumentar la tasa de adaptación de la diferencia de correlación a largo plazo \overline{G}_{LR} .

50 Realizar la suboperación 405 de conversión y cuantificación, una vez que la diferencia de correlación a largo plazo \overline{G}_{LR} se ha estimado correctamente en la calculadora 454, el convertidor y cuantificador 455 convierte esta diferencia en un factor β que se cuantifica, y se suministra a (a) el codificador 252 de canal primario (Figura 2), (b) el codificador

253/353 de canal secundario (Figuras 2 y 3), y (c) el multiplexor 254/354 (Figuras 2 y 3) para su transmisión a un decodificador dentro del flujo de bits 207/307 multiplexado a través de un enlace de comunicación tal como el 101 de la Figura 1.

- 5 El factor β representa dos aspectos de la entrada estéreo combinados en un parámetro. Primero, el factor β representa una proporción o contribución de cada uno de los canales derecho R e izquierdo L que se combinan para crear el canal primario Y y, en segundo lugar, también puede representar un factor de escala de energía para aplicar al canal primario Y para obtener un canal primario que está cerca en el dominio de la energía de cómo se vería una versión de señal monofónica del sonido. Por lo tanto, en el caso de una estructura incorporada, permite decodificar el canal primario Y solo sin la necesidad de recibir el flujo de bits 306 secundario que lleva los parámetros estéreo. Este parámetro de energía también puede usarse para cambiar de escala la energía del canal secundario X antes de codificarla, de tal manera que la energía global del canal secundario X esté más cerca del rango de energía óptimo del codificador del canal secundario. Como se muestra en la Figura 2, la información de energía intrínsecamente presente en el factor β también se puede utilizar para mejorar la asignación de bits entre los canales primario y secundario.
- 10 El factor cuantificado β puede transmitirse al decodificador mediante un índice. Dado que el factor β puede representar tanto (a) contribuciones respectivas de los canales izquierdo y derecho al canal primario y (b) un factor de escala de energía para aplicar al canal primario para obtener una versión de señal monofónica del sonido o una información de correlación/energía que ayuda a asignar más eficientemente los bits entre el canal primario Y y el canal secundario X, el índice transmitido al decodificador transmite dos elementos de información distintos con el mismo número de bits.
- 15 Para obtener una asignación entre la diferencia de correlación a largo plazo $\overline{G_{LR}(t)}$ y el factor β , en esta realización de ejemplo, el convertidor y cuantificador 455 primero limitan la diferencia de correlación a largo plazo $\overline{G_{LR}(t)}$ entre -1.5 a 1.5 y luego linealiza esta diferencia de correlación a largo plazo entre 0 y 2 para obtener una diferencia de correlación a largo plazo linealizada temporal $\overline{G'_{LR}(t)}$ como se muestra en la relación (7):

$$G'_{LR}(t) = \begin{cases} 0, & \overline{G_{LR}(t)} \leq -1.5 \\ \frac{2}{3} \cdot \overline{G_{LR}(t)} + 1.0, & -1.5 < \overline{G_{LR}(t)} < 1.5 \\ 2, & \overline{G_{LR}(t)} \geq 1.5 \end{cases} \quad (7)$$

- 25 En una implementación alternativa, se puede decidir utilizar solo una parte del espacio lleno con la diferencia de correlación linealizada a largo plazo $G'_{LR}(t)$, limitando aún más sus valores entre, por ejemplo, 0.4 y 0.6. Esta limitación adicional tendría el efecto de reducir la localización de la imagen estéreo, pero también ahorraría algunos bits de cuantificación. Dependiendo de la elección de diseño, se puede considerar esta opción.

- Después de la linealización, el convertidor y cuantificador 455 realizan una asignación de la diferencia de correlación linealizada a largo plazo $G'_{LR}(t)$ en el dominio "coseno" utilizando la relación (8):

$$\beta(t) = \frac{1}{2} \cdot \left(1 - \cos \left(\pi \cdot \frac{G'_{LR}(t)}{2} \right) \right) \quad (8)$$

Para realizar la suboperación 406 de mezcla descendente en el dominio del tiempo, un mezclador 456 descendente en el dominio del tiempo produce el canal primario Y y el canal secundario X como una mezcla de los canales derecho R e izquierdo L utilizando las relaciones (9) y (10):

$$35 \quad Y(i) = R(i) \cdot (1 - \beta(t)) + L(i) \cdot \beta(t) \quad (9)$$

$$X(i) = L(i) \cdot (1 - \beta(t)) - R(i) \cdot \beta(t) \quad (10)$$

donde $i = 0, \dots, N-1$ es el índice de la muestra en la trama y t es el índice de la trama.

- La Figura 13 es un diagrama de bloques que muestra simultáneamente otras realizaciones de suboperaciones de la operación 201/301 de mezcla descendente en el dominio del tiempo del método de codificación de sonido estéreo de las Figuras 2 y 3, y módulos del mezclador 251/351 de canales del sistema de codificación de sonido estéreo de las Figuras 2 y 3, utilizando un factor de adaptación previa para mejorar la estabilidad de la imagen estéreo. En una implementación alternativa como se representa en la Figura 13, la operación 201/301 de mezcla descendente en el dominio del tiempo comprende las siguientes suboperaciones: una suboperación 1301 de análisis de energía, una suboperación 1302 de análisis de tendencias de energía, una suboperación 1303 de análisis de correlación normalizada de canal L y R, una suboperación 1304 de cálculo del factor de adaptación previa, una operación 1305 de aplicación del factor de adaptación previa a correlaciones normalizadas, una suboperación 1306 de cálculo de

diferencia de correlación a largo plazo (LT), una suboperación 1307 de conversión y cuantificación de ganancia al factor β , y una suboperación 1308 de mezcla descendente en el dominio del tiempo.

5 Las suboperaciones 1301, 1302 y 1303 son realizadas respectivamente por un analizador 1351 de energía, un analizador 1352 de tendencias de energía y un analizador 1353 de correlación normalizada L y R, sustancialmente de la misma manera que se explica en la descripción anterior en relación con las suboperaciones 401, 402 y 403, y analizadores 451, 452 y 453 de la Figura 4.

10 Para realizar la suboperación 1305, el mezclador 251/351 de canales comprende una calculadora 1355 para aplicar el factor de adaptación previa a_r directamente a las correlaciones $G_{L|R}$ ($G_L(t)$ y $G_R(t)$) a partir de relaciones (4) de tal manera que su evolución se suaviza en función de la energía y las características de ambos canales. Si la energía de la señal es baja o si tiene algunas características sordas, entonces la evolución de la ganancia de correlación puede ser más lenta.

15 Para llevar a cabo la suboperación 1304 de cálculo del factor de adaptación previa, el mezclador 251/351 de canales comprende una calculadora 1354 de factor de adaptación previa, provista de (a) los valores de energía de los canales izquierdo y derecho a largo plazo de las relaciones (2) del analizador 1351 de energía, (b) clasificación de tramas de tramas anteriores y (c) información de actividad de voz de las tramas anteriores. La calculadora 1354 del factor de adaptación previa calcula el factor de adaptación previa a_r , el cual puede linealizarse entre 0.1 y 1 dependiendo de los valores rms mínimos a largo plazo $\overline{rms}_{L|R}$ de los canales izquierdo y derecho del analizador 1351, utilizando la relación (6a):

$$a_r = \max\left(\min\left(M_a \cdot \min\left(\overline{rms}_L(t), \overline{rms}_R(t)\right) + B_a, 1\right), 0.1\right), \quad (11a)$$

20 En una realización, coeficiente M_a puede tener el valor de 0.0009 y el coeficiente B_a el valor de 0.16. En variante, el factor de adaptación previa a_r puede forzarse a 0.15, por ejemplo, si una clasificación previa de los dos canales R y L es indicativa de características sordas y de una señal activa. También puede usarse una bandera de restos de detección de actividad de voz (VAD) para determinar que una parte anterior del contenido de una trama era un segmento activo.

25 La operación 1305 de aplicar el factor de adaptación previa a_r a las correlaciones normalizadas $G_{L|R}$ ($G_L(t)$ y $G_R(t)$) de las relaciones (4) de los canales izquierdo L y derecho R es distinta de la operación 404 de la Figura 4. En lugar de calcular las correlaciones normalizadas suavizadas a largo plazo (LT) aplicándolas a las correlaciones normalizadas $G_{L|R}$ ($G_L(t)$ y $G_R(t)$) un factor $(1-\alpha)$, siendo α la tasa de convergencia definida anteriormente (Relaciones (5)), la calculadora 1355 aplica el factor de adaptación previa a_r directamente a las correlaciones normalizadas $G_{L|R}$ ($G_L(t)$ y $G_R(t)$) de los canales izquierdo L y derecho R utilizando la relación (11b):

$$\tau_L(t) = a_r \cdot G_L(t) + (1 - a_r) \cdot \overline{G}_L(t) \quad \text{y} \quad \tau_R(t) = a_r \cdot G_R(t) + (1 - a_r) \cdot \overline{G}_R(t). \quad (11b)$$

35 La calculadora 1355 genera ganancias de correlación adaptadas $\tau_{L|R}$ que se proporcionan a una calculadora 1356 de diferencias de correlación a largo plazo (LT). La operación de mezcla 201/301 descendente en el dominio del tiempo (Figuras 2 y 3) comprende, en la implementación de la Figura 13, una diferencia de correlación a largo plazo (LT) que calcula la suboperación 1306, una diferencia de correlación a largo plazo con la conversión del factor β y la suboperación 1307 de cuantificación y una suboperación 1358 de mezcla descendente en el dominio del tiempo similar a las suboperaciones 404, 405 y 406, respectivamente, de la Figura 4.

40 La operación 201/301 de mezcla descendente en el dominio del tiempo (Figuras 2 y 3) comprende, en la implementación de la Figura 13, una suboperación 1306 de cálculo de diferencia de correlación a largo plazo (LT), una suboperación 1307 de conversión y cuantificación de diferencia de correlación a largo plazo del factor β y una suboperación 1358 de mezcla descendente en el dominio del tiempo similar a las suboperaciones 404, 405 y 406, respectivamente, de la Figura 4.

45 Las suboperaciones 1306, 1307 y 1308 se realizan respectivamente mediante una calculadora 1356, un convertidor y cuantificador 1357 y un mezclador 1358 descendente en el dominio del tiempo, sustancialmente de la misma manera que se explica en la descripción anterior en relación con las suboperaciones 404, 405 y 406, y la calculadora 454, el convertidor y cuantificador 455 y el mezclador 456 descendente en el dominio del tiempo.

50 La Figura 5 muestra cómo la diferencia de correlación linealizada a largo plazo $G'_{LR}(t)$ se asigna al factor β y el escalamiento de energía. Se puede observar que para una diferencia de correlación linealizada de largo plazo $G'_{LR}(t)$ de 1.0, lo que significa que las energías/correlaciones del canal derecho R e izquierdo L son casi iguales, el factor β es igual a 0.5 y un factor de normalización de energía (cambio de escala) ϵ es 1.0. En esta situación, el contenido del canal primario Y es básicamente una mezcla mono y el canal secundario X forma un canal lateral. El cálculo del factor de normalización de energía (cambio de escala) ϵ se describe a continuación.

Por otro lado, si la diferencia de correlación linealizada a largo plazo $G'_{LR}(t)$ es igual a 2, lo que significa que la mayor parte de la energía está en el canal izquierdo L, entonces el factor β es 1 y el factor de normalización de energía

(cambio de escala) es 0.5, lo que indica que el canal primario Y contiene básicamente el canal izquierdo L en una implementación de diseño integrado o una representación a escala reducida del canal izquierdo L en una implementación de diseño integrado. En este caso, el canal secundario X contiene el canal derecho R. En las realizaciones de ejemplo, el convertidor y cuantificador 455 o 1357 cuantifica el factor β utilizando 31 posibles entradas de cuantificación. La versión cuantificada del factor β se representa utilizando un índice de 5 bits y, como se describió anteriormente, se suministra al multiplexor para su integración en el flujo de bits 207/307 multiplexado, y se transmite al decodificador a través del enlace de comunicación.

En una realización, el factor β también puede usarse como un indicador tanto para el codificador 252/352 de canal primario como para el codificador 253/353 de canal secundario para determinar la asignación de tasa de bits. Por ejemplo, si el factor β está cerca de 0.5, lo que significa que las dos (2) energías/correlación del canal de entrada con el mono están cerca entre sí, se asignarían más bits al canal secundario X y menos bits al canal primario Y, excepto si el contenido de ambos canales está bastante cerca, entonces el contenido del canal secundario será de muy baja energía y probablemente se considerará inactivo, lo que permitirá que muy pocos bits lo codifiquen. Por otro lado, si el factor β está más cerca de 0 o 1, entonces la asignación de la tasa de bits favorecerá al canal primario Y.

La Figura 6 muestra la diferencia entre utilizar el esquema *pca/klt* mencionado anteriormente sobre toda la trama (dos curvas superiores de la Figura 6) con respecto a utilizar la función "coseno" desarrollada en la relación (8) para calcular el factor β (curva inferior de la Figura 6). Por naturaleza el esquema *pca/klt* tiende a buscar un mínimo o un máximo. Esto funciona bien en el caso de voz activa, como se muestra en la curva central de la Figura 6, pero no funciona realmente bien para voz con ruido de fondo, ya que tiende a cambiar continuamente de 0 a 1 como se muestra en la curva central de la Figura 6. El cambio demasiado frecuente a los extremos, 0 y 1, provoca diversos artefactos cuando se codifica a una tasa de bits baja. Una posible solución habría sido suavizar las decisiones del esquema *pca/klt*, pero esto habría impactado negativamente en la detección de ráfagas de voz y sus ubicaciones correctas, a la vez que la función "coseno" de la relación (8) es más eficiente a este respecto.

La Figura 7 muestra el canal primario Y, el canal secundario X y los espectros de estos canales primario Y y secundario X resultantes de aplicar la mezcla descendente en el dominio del tiempo a una muestra estéreo que se ha grabado en una pequeña sala ecoica utilizando una configuración de micrófonos binaurales con ruido de oficina de fondo. Después de la operación de mezcla descendente en el dominio del tiempo, se puede ver que ambos canales todavía tienen formas de espectro similares y el canal secundario X todavía tiene un contenido temporal similar a la voz, lo que permite utilizar un modelo con base en la voz para codificar el canal secundario X.

La mezcla descendente en el dominio del tiempo presentada en la descripción anterior puede mostrar algunos problemas en el caso especial de los canales derecho R e izquierdo L que están invertidos en fase. La suma de los canales derecho R e izquierdo L para obtener una señal monofónica daría como resultado que los canales derecho R e izquierdo L se cancelen entre sí. Para resolver este posible problema, en una realización, el mezclador 251/351 de canales compara la energía de la señal monofónica con la energía de los canales derecho R e izquierdo L. La energía de la señal monofónica debe ser al menos mayor que la energía de uno de los canales derecho R e izquierdo L. De lo contrario, en esta realización, el modelo de mezcla descendente en el dominio del tiempo entra en el caso especial de fase invertida. En presencia de este caso especial, el factor β se fuerza a 1 y el canal secundario X se codifica de manera forzada utilizando el modo genérico o no sonoro, impidiendo así el modo de codificación inactivo y asegurando la codificación adecuada del canal secundario X. Este caso especial, donde no se aplica ningún cambio de escala de energía, se señala al decodificador utilizando la última combinación de bits (valor de índice) disponible para la transmisión del factor β (Básicamente dado que β se cuantifica utilizando 5 bits y 31 entradas (niveles de cuantificación) se utilizan para la cuantificación como se describe anteriormente, la 32ª posible combinación de bits (valor de entrada o índice) se utiliza para señalar este caso especial).

En una implementación alternativa, se puede poner más énfasis en la detección de señales que son subóptimas para las técnicas de codificación y mezcla descendente descritas anteriormente, tal como en los casos de señales fuera de fase o casi fuera de fase. Una vez que se detectan estas señales, las técnicas de codificación subyacentes pueden adaptarse si es necesario.

Por lo general, para la mezcla descendente en el dominio del tiempo como se describe en el presente documento, cuando los canales izquierdo L y R derecho de una señal estéreo de entrada están fuera de fase, puede ocurrir alguna cancelación durante el proceso de mezcla descendente, lo cual podría conducir a una calidad subóptima. En los ejemplos anteriores, la detección de estas señales es simple y la estrategia de codificación comprende codificar ambos canales por separado. Pero a veces, con señales especiales, como señales que están fuera de fase, puede ser más eficiente realizar una mezcla descendente similar a mono/lateral ($\beta = 0.5$), donde se pone un mayor énfasis en el canal lateral. Dado que un tratamiento especial de estas señales puede resultar beneficioso, la detección de dichas señales debe realizarse con cuidado. Además, la transición a partir del modelo de mezcla descendente en el dominio del tiempo normal como se describe en la descripción anterior y el modelo de mezcla descendente en el dominio del tiempo que se ocupa de estas señales especiales puede activarse en una región de muy baja energía o en regiones donde el tono de ambos canales no es estable, de tal modo que el cambio entre los dos modelos tiene un efecto subjetivo mínimo.

Corrección de retardo temporal (TDC) (véase corrector 1750 de retardo temporal en las Figuras 17 y 18) entre los canales L y R, o una técnica similar a la descrita en la referencia [8], cuyo contenido completo se incorpora en el presente documento como referencia, puede realizarse antes de entrar en el módulo 201/301, 251/351 de mezcla descendente. En dicha realización, el factor β puede terminar teniendo un significado diferente al que se ha descrito anteriormente. Para este tipo de implementación, con la condición de que la corrección del retardo temporal funcione como se esperaba, el factor β puede acercarse a 0.5, lo que significa que la configuración de la mezcla descendente en el dominio del tiempo es cercana a una configuración mono/lateral. Con el funcionamiento adecuado de la corrección de retardo temporal (TDC), el lado puede contener una señal que incluya una cantidad menor de información importante. En ese caso, la tasa de bits del canal secundario X puede ser mínima cuando el factor β está cerca de 0.5. Por otro lado, si el factor β está cerca de 0 o 1, esto significa que la corrección del retardo temporal (TDC) puede no superar adecuadamente la situación de falta de alineación del retardo y es probable que el contenido del canal secundario X sea más complejo, por lo que necesita una tasa de bits más alta. Para ambos tipos de implementación, el factor β y por asociación el factor ϵ de normalización de energía (cambio de escala), pueden usarse para mejorar la asignación de bits entre el canal primario Y y el canal secundario X.

La Figura 14 es un diagrama de bloques que muestra simultáneamente las operaciones de una detección de señal fuera de fase y los módulos de un detector 1450 de señal fuera de fase que forman parte de la operación 201/301 de mezcla descendente y el mezclador 251/351 de canales. Las operaciones de detección de señal fuera de fase incluyen, como se muestra en la Figura 14, una operación 1401 de detección de señal fuera de fase, una operación 1402 de detección de posición de conmutación, y la operación 1403 de selección de mezclador de canal, para elegir entre la operación 201/301 de mezcla descendente en el dominio del tiempo y una operación 1404 de mezcla descendente en el dominio del tiempo específico fuera de fase. Estas operaciones se realizan respectivamente mediante un detector 1451 de señal fuera de fase, un detector 1452 de posición de conmutación, un selector 1453 de mezclador de canal, el mezclador 251/351 de canal descendente en el dominio del tiempo descrito anteriormente, y un mezclador 1454 de canal descendente en el dominio del tiempo específico fuera de fase.

La detección 1401 de señal fuera de fase se basa en una correlación de bucle abierto entre los canales primario y secundario en tramas anteriores. Para ello, el detector 1451 calcula en las tramas anteriores una diferencia de energía $S_m(t)$ entre una señal lateral $s(i)$ y una señal mono $m(i)$ utilizando las relaciones (12a) y (12b):

$$S_m(t) = 10 \cdot \left(\log_{10} \left(\frac{\sqrt{\sum_{i=0}^{N-1} s(i)^2}}{N} \right) - \log_{10} \left(\frac{\sqrt{\sum_{i=0}^{N-1} m(i)^2}}{N} \right) \right), \quad (12a)$$

$$m(i) = \left(\frac{L(i)+R(i)}{2} \right) \quad \text{y} \quad s(i) = \left(\frac{L(i)-R(i)}{2} \right), \quad (12b)$$

Luego, el detector 1451 calcula el lado a largo plazo de la diferencia de energía mono $S_m(t)$ utilizando la relación (12c):

$$\overline{S}_m(t) = \begin{cases} 0.9 \cdot \overline{S}_m(t_{-1}), & \text{para contenido inactivo,} \\ 0.9 \cdot \overline{S}_m(t_{-1}) + 0.1 \cdot S_m(t), & \text{de lo contrario} \end{cases} \quad (12c)$$

donde t indica la trama actual, t_{-1} la trama anterior, y donde el contenido inactivo puede derivarse de la bandera de restos del Detector de Actividad de Voz (VAD) o de un contador de restos de VAD.

Además del lado a largo plazo de la diferencia de energía mono $\overline{S}_m(t)$, la última correlación máxima de bucle abierto de tono C_{FL} de cada canal Y y X, como se define en la cláusula 5.1.10 de la Referencia [1], también se tiene en cuenta para decidir cuándo el modelo actual se considera subóptimo. $C_{P(t-1)}$ representa la correlación máxima de bucle abierto de tono del canal primario Y en una trama anterior y $C_{S(t-1)}$, la correlación máxima de bucle de tono abierto del canal secundario X en la trama anterior. Una bandera F_{sub} subóptima es calculada por el detector 1452 de posición de conmutación de acuerdo con los siguientes criterios:

Si el lado a largo plazo de la diferencia de energía mono $S_m(t)$ está por encima de un cierto umbral, por ejemplo, cuando $\overline{S}_m(t) > 2.0$, si ambas correlaciones máximas de tono en bucle abierto $C_{P(t-1)}$ y $C_{S(t-1)}$ están entre 0.85 y 0.92, lo que significa que las señales tienen una buena correlación, pero no están tan correlacionadas como lo estaría una señal sonora, la bandera de subóptimo F_{sub} se establece en 1, lo que indica una condición fuera de fase entre los canales izquierdo L y derecho R.

De lo contrario, la bandera de subóptimo F_{sub} se establece en 0, lo que indica que no hay ninguna condición fuera de fase entre los canales izquierdo L y derecho R.

Para agregar algo de estabilidad en la decisión de la bandera de subóptimo, el detector 1452 de posición de conmutación implementa un criterio con respecto al contorno de tono de cada canal Y y X. El detector 1452 de posición de conmutación determina que el mezclador 1454 de canales se usará para codificar las subseñales óptimas cuando, en la realización de ejemplo, al menos tres (3) instancias consecutivas de la bandera de subóptimo F_{sub} se establecen en 1 y la estabilidad de tono de la última trama de uno de los canales primario, $P_{pc(t-1)}$, o del canal secundario, $P_{sc(t-1)}$,

es mayor que 64. La estabilidad de tono consiste en la suma de las diferencias absolutas de los tres tonos de bucle abierto. $P_{0|12}$ como se define en 5.1.10 de la Referencia [1], calculado por el detector 1452 de posición de conmutación utilizando la relación (12d):

$$p_{pc} = |p_1 - p_0| + |p_2 - p_1| \quad y \quad p_{sc} = |p_1 - p_0| + |p_2 - p_1| \quad (12d)$$

5 El detector 1452 de posición de conmutación proporciona la decisión al selector 1453 de mezclador de canales que, a su vez, selecciona el mezclador 251/351 de canales o el mezclador 1454 de canales en consecuencia. El selector 1453 de mezclador de canales implementa una histéresis de tal modo que, cuando se selecciona el mezclador 1454 de canales, esta decisión se mantiene hasta que se cumplen las siguientes condiciones: un número de tramas consecutivas, por ejemplo 20 tramas, se consideran óptimas, la estabilidad de tono de la última trama de uno del canal
10 primario $P_{pc(t-1)}$ o secundario $P_{sc(t-1)}$ es mayor que un número predeterminado, por ejemplo 64, y el lado a largo plazo de la diferencia de energía mono $\bar{s}_m(t)$ es inferior o igual a 0.

2) *Codificación dinámica entre canales primario y secundario* (El siguiente aspecto completo es útil para comprender la invención, pero no está cubierto por las reivindicaciones)

15 La Figura 8 es un diagrama de bloques que ilustra simultáneamente el método y el sistema de codificación de sonido estéreo, con una posible implementación de optimización de la codificación de los canales primario Y y X secundario de la señal de sonido estéreo, como voz o audio.

Con referencia a la Figura 8, el método de codificación de sonido estéreo comprende una operación 801 de procesamiento previo de baja complejidad implementada por un procesador 851 previo de baja complejidad, una
20 operación 802 de clasificación de señales implementada por un clasificador 852 de señales, una operación 803 de decisión implementada por un módulo 853 de decisión, una operación 804 de codificación solo genérica de modelo de cuatro (4) subtramas implementada por un módulo 854 de codificación solo genérica del modelo de cuatro (4) subtramas, una operación 805 de codificación de modelo de dos (2) subtramas implementada por un módulo 855 de codificación de modelo de dos (2) subtramas, y una operación 806 de análisis de coherencia de filtro LP implementada por un analizador 856 de coherencia de filtro LP.

25 Después de que el mezclador 351 de canales haya realizado la mezcla 301 descendente en el dominio del tiempo, en el caso del modelo integrado, el canal primario Y se codifica (operación 302 de codificación del canal primario) (a) utilizando como codificador 352 de canal primario un codificador heredado como el codificador EVS heredado o cualquier otro codificador de sonido heredado adecuado (debe tenerse en cuenta que, como se menciona en la descripción anterior, se puede utilizar cualquier tipo de codificador adecuado como codificador 352 de canal primario).
30 En el caso de una estructura integrada, se utiliza un códec de voz dedicado como codificador 252 de canal primario. El codificador 252 de voz dedicado puede ser un codificador con base en tasa de bits variable (VBR), por ejemplo, una versión modificada del codificador EVS heredado, el cual ha sido modificado para tener una mayor escalabilidad de tasa de bits que permite el manejo de una tasa de bits variable en un por nivel de trama (de nuevo debe tenerse en cuenta que, como se menciona en la descripción anterior, se puede utilizar cualquier tipo adecuado de codificador como codificador 252 de canal primario). Esto permite que la cantidad mínima de bits utilizados para codificar el canal secundario X varíe en cada trama y se adapte a las características de la señal de sonido a codificar. Al final, la firma del canal secundario X será lo más homogénea posible.

La codificación del canal secundario X, es decir, la energía/correlación más baja con la entrada mono, se optimiza para utilizar una tasa de bits mínima, en particular, pero no exclusivamente, para contenido similar a la voz. Para ese
40 propósito, la codificación del canal secundario puede aprovechar los parámetros que ya están codificados en el canal primario Y, tales como los coeficientes de filtro LP (LPC) y/o el retardo 807 de tono. Específicamente, se decidirá, como se describe a continuación, si los parámetros calculados durante la codificación del canal primario están lo suficientemente cerca de los parámetros correspondientes calculados durante la codificación del canal secundario para ser reutilizados durante la codificación del canal secundario.

45 Primero, la operación 801 de procesamiento previo de baja complejidad se aplica al canal secundario X utilizando el procesador 851 previo de baja complejidad, en donde un filtro LP, una detección de actividad de voz (VAD) y un tono de bucle abierto se calculan en respuesta al canal secundario X. Los últimos cálculos pueden ser implementados, por ejemplo, por aquellos realizados en el codificador EVS heredado y descritos respectivamente en las cláusulas 5.1.9,
50 5.1.12 y 5.1.10 de la Referencia [1] de los cuales, como se indicó anteriormente, el contenido completo se incorpora en el presente documento como referencia. Dado que, como se menciona en la descripción anterior, puede usarse cualquier tipo de codificador adecuado como codificador 252/352 de canal primario, los cálculos anteriores pueden ser implementados por aquellos realizados en dicho codificador de canal primario.

Luego, las características de la señal del canal secundario X son analizadas por el clasificador 852 de señales para clasificar el canal secundario X como no sonoro, genérico o inactivo utilizando técnicas similares a las de la función
55 de clasificación de señales EVS, cláusula 5.1.13 de la misma Referencia [1]. Estas operaciones son conocidas por los expertos en la técnica y se pueden extraer del Estándar 3GPP TS 26.445, v.12.0.0 para simplificar, pero también se pueden utilizar implementaciones alternativas.

a. Reutilización de los coeficientes de filtro LP del canal primario

Una parte importante del consumo de tasa de bits reside en la cuantificación de los coeficientes de filtro LP (LPC). A una tasa de bits baja, la cuantificación completa de los coeficientes de filtro LP puede ocupar hasta casi el 25% del presupuesto de bits. Dado que el canal secundario X a menudo tiene un contenido de frecuencia cercano al canal primario Y, pero con el nivel de energía más bajo, vale la pena verificar si fuera posible reutilizar los coeficientes de filtro LP del canal primario Y. Para hacerlo, como se muestra en la Figura 8, se ha desarrollado una operación 806 de análisis de coherencia de filtro LP implementada por un analizador 856 de coherencia de filtro LP, en el cual se calculan y comparan pocos parámetros para validar la posibilidad de reutilizar o no los coeficientes de filtro LP (LPC) 807 del canal primario Y.

La Figura 9 es un diagrama de bloques que ilustra la operación 806 de análisis de coherencia de filtro LP y el analizador 856 de coherencia de filtro LP correspondiente del método y sistema de codificación de sonido estéreo de la Figura 8.

La operación 806 de análisis de coherencia de filtro LP y el analizador 856 de coherencia de filtro LP correspondiente del método y sistema de codificación de sonido estéreo de la Figura 8 comprenden, como se ilustra en la Figura 9, una suboperación 903 de análisis de filtro LP (Predicción Lineal) del canal primario implementada por un analizador 953 de filtro LP, una suboperación 904 de ponderación implementada por un filtro 954 de ponderación, una suboperación 912 de análisis de filtro LP de canal secundario implementada por un analizador 962 de filtro LP, una suboperación 901 de ponderación implementada por un filtro 951 de ponderación, una suboperación 902 de análisis de distancia Euclidiana implementada por un analizador 952 de distancia Euclidiana, una suboperación 913 de filtrado residual implementada por un filtro 963 residual, una suboperación 914 de cálculo de energía residual implementada por una calculadora 964 de energía de residual, una sustracción de suboperación 915 implementada por un sustractor 965, una suboperación 910 de cálculo de energía de sonido (tal como voz y/o audio) implementada por una calculadora 960 de energía, una operación 906 de filtrado residual del canal secundario implementada por un filtro 956 residual del canal secundario, una suboperación 907 de cálculo de energía residual implementada por una calculadora de energía del residual 957, una suboperación 908 de sustracción implementada por un sustractor 958, una suboperación 911 de cálculo de relación de ganancia implementada por una calculadora de relación de ganancia, una suboperación 916 de comparación implementada por un comparador 966, una suboperación 917 de comparación implementada por un comparador 967, una suboperación 918 de decisión de uso de filtro LP de canal secundario implementado por un módulo 968 de decisión, y una suboperación 919 de decisión de reutilización de filtro LP de canal primario implementada por un módulo 969 de decisión.

Con referencia a la Figura 9, el analizador 953 de filtro LP realiza un análisis de filtro LP en el canal primario Y a la vez que el analizador 962 de filtro LP realiza un análisis de filtro LP en el canal secundario X. El análisis de filtro LP realizado en cada uno de los canales primario Y y secundario X son similares al análisis descrito en la cláusula 5.1.9 de la Referencia [1].

Entonces, los coeficientes A_y de filtro LP del analizador 953 de filtro LP se suministran al filtro 956 residual para un primer filtrado residual, r_y , del canal secundario X. De la misma manera, los coeficientes A_x óptimos de filtro LP del analizador 962 de filtro LP se suministran al filtro 963 residual para un segundo filtrado residual, r_x , del canal secundario X. El filtrado residual con coeficientes de filtro, A_y o A_x , se realiza utilizando la relación (11):

$$r_{Y|X}(n) = s_X(n) + \sum_{i=0}^{16} (A_{Y|X}(i) \cdot s_X(n-i)), n = 0, \dots, N-1 \quad (13)$$

donde, en este ejemplo, S_X representa el canal secundario, el orden de filtro LP es 16, y N es el número de muestras en la trama (tamaño de la trama) la cual suele ser 256, lo que corresponde a una duración de trama de 20 ms a una frecuencia de muestreo de 12.8 kHz.

La calculadora 910 calcula la energía E_x de la señal de sonido en el canal secundario X utilizando la relación (14):

$$E_x = 10 \cdot \log_{10}(\sum_{i=0}^{N-1} s_x(i)^2), \quad (14)$$

y la calculadora 957 calcula la energía R_{ry} del residual del filtro 956 residual utilizando la relación (15):

$$E_{ry} = 10 \cdot \log_{10}(\sum_{i=0}^{N-1} r_y(i)^2). \quad (15)$$

El sustractor 958 sustrae la energía residual de la calculadora 957 de la energía sonora de la calculadora 960 para producir una ganancia de predicción G_y .

De la misma manera, la calculadora 964 calcula la energía E_{rx} del residual del filtro 963 residual utilizando la relación (16):

$$E_{rx} = 10 \cdot \log_{10}(\sum_{i=0}^{N-1} r_x(i)^2), \quad (16)$$

y el sustractor 965 sustrae esta energía residual de la energía del sonido de la calculadora 960 para producir una ganancia de predicción G_x .

La calculadora 961 calcula la relación de ganancia G_Y/G_X . El comparador 966 compara la relación de ganancia G_Y/G_X hasta un umbral τ , el cual es 0.92 en la realización de ejemplo. Si la relación G_Y/G_X es menor que el umbral τ , el resultado de la comparación se transmite al módulo 968 de decisión que fuerza el uso de los coeficientes de filtro LP del canal secundario para codificar el canal secundario X.

El analizador 952 de distancia Euclidiana realiza una medida de similitud de filtro LP, tal como la distancia Euclidiana entre los pares espectrales de línea lsp_Y calculados por el analizador 953 de filtro LP en respuesta al canal primario Y y los pares espectrales de línea lsp_X calculado por el analizador 962 de filtro LP en respuesta al canal secundario X. Como saben los expertos en la técnica, los pares espectrales de línea lsp_Y y lsp_X representan los coeficientes de filtro LP en un dominio de cuantificación. El analizador 952 utiliza la relación (17) para determinar la distancia Euclidiana $dist$:

$$dist = \sum_{i=0}^{M-1} (lsp_Y(i) - lsp_X(i))^2 \quad (17)$$

donde M representa el orden del filtro, y lsp_Y y lsp_X representan respectivamente los pares espectrales de línea calculados para los canales primario Y y X secundario.

Antes de calcular la distancia Euclidiana en el analizador 952, es posible ponderar ambos conjuntos de pares espectrales de línea lsp_Y y lsp_X a través de factores de ponderación respectivos, de tal modo que se ponga más o menos énfasis en determinadas porciones del espectro. También se pueden utilizar otras representaciones de filtros LP para calcular la medida de similitud de filtro LP.

Una vez que la distancia Euclidiana $dist$ se conoce, se compara con un umbral σ en el comparador 967. En la realización de ejemplo, el umbral σ tiene un valor de 0.08. Cuando el comparador 966 determina que la relación G_Y/G_X es igual o mayor que el umbral τ y el comparador 967 determina que la distancia Euclidiana $dist$ es igual o mayor que el umbral σ , el resultado de las comparaciones se transmite al módulo 968 de decisión el cual fuerza el uso de los coeficientes de filtro LP del canal secundario para codificar el canal secundario X. Cuando el comparador 966 determina que la relación G_Y/G_X es igual o mayor que el umbral τ y el comparador 967 determina que la distancia Euclidiana $dist$ es menor que el umbral σ , el resultado de estas comparaciones se transmite al módulo 969 de decisión el cual fuerza la reutilización de los coeficientes de filtro LP del canal primario para codificar el canal secundario X. En el último caso, los coeficientes de filtro LP del canal primario se reutilizan como parte de la codificación de canal secundario.

Se pueden realizar algunas pruebas adicionales para limitar la reutilización de los coeficientes de filtro LP del canal primario para codificar el canal secundario X en casos particulares, por ejemplo, en el caso del modo de codificación sorda, donde la señal es lo suficientemente fácil de codificar que todavía hay tasa de bits disponible para codificar también los coeficientes de filtro LP. También es posible forzar la reutilización de los coeficientes de filtro LP del canal primario cuando ya se obtiene una ganancia residual muy baja con los coeficientes de filtro LP del canal secundario o cuando el canal secundario X tiene un nivel de energía muy bajo. Finalmente, las variables τ , σ , el nivel de ganancia residual o el nivel de energía muy bajo al cual se puede forzar la reutilización de los coeficientes de filtro LP se pueden adaptar todos en función del presupuesto de bits disponible y/o en función del tipo de contenido. Por ejemplo, si el contenido del canal secundario se considera inactivo, incluso si la energía es alta, se puede decidir reutilizar los coeficientes de filtro LP del canal primario.

b. Codificación de baja tasa de bits del canal secundario

Dado que los canales primario Y y secundario X pueden ser una mezcla de los canales de entrada derecho R e izquierdo L, esto implica que, incluso si el contenido de energía del canal secundario X es bajo en comparación con el contenido de energía del canal primario Y, se puede percibir un artefacto de codificación una vez que se realiza la mezcla ascendente de los canales. Para limitar dicho posible artefacto, la firma de codificación del canal secundario X se mantiene tan constante como sea posible para limitar cualquier variación de energía no intencionada. Como se muestra en la Figura 7, el contenido del canal secundario X tiene características similares al contenido del canal primario Y y por esa razón se ha desarrollado un modelo de codificación de voz de muy baja tasa de bits.

Volviendo a la Figura 8, el analizador 856 de coherencia de filtro LP envía al módulo 853 de decisión la decisión de reutilizar los coeficientes de filtro LP del canal primario del módulo 969 de decisión o la decisión de utilizar los coeficientes de filtro LP del canal secundario del módulo 968 de decisión. El módulo de decisión 803 decide entonces no cuantificar los coeficientes de filtro LP del canal secundario cuando se reutilizan los coeficientes de filtro LP del canal primario y cuantificar los coeficientes de filtro LP del canal secundario cuando la decisión es utilizar los coeficientes de filtro LP del canal secundario. En el último caso, los coeficientes de filtro LP del canal secundario cuantificados se envían al multiplexor 254/354 para su inclusión en el flujo 207/307 de bits multiplexado.

5 En la operación 804 de codificación solo genérica del modelo de cuatro (4) subtramas y el correspondiente módulo 854 de codificación solo genérica del modelo de cuatro (4) subtramas, para mantener la tasa de bits lo más baja posible, se utiliza una búsqueda ACELP como se describe en la cláusula 5.2.3.1 de la Referencia [1] sólo cuando los coeficientes de filtro LP del canal primario Y se pueden reutilizar, cuando el canal secundario X está clasificado como genérico por el clasificador 852 de señal, y cuando la energía de los canales de entrada derecho R e izquierdo L está cerca del centro, lo que significa que las energías de los canales derecho R e izquierdo L están cerca entre sí. Los parámetros de codificación encontrados durante la búsqueda ACELP en el módulo 854 de codificación genérica del modelo de cuatro (4) subtramas se utilizan luego para construir el flujo 206/306 de bits del canal secundario y se envían al multiplexor 254/354 para su inclusión en el flujo 207/307 de bits multiplexado.

10 De lo contrario, en la operación 805 de codificación del modelo de dos (2) subtramas y el correspondiente módulo 855 de codificación del modelo de dos (2) subtramas, se utiliza un modelo de media banda para codificar el canal secundario X con contenido genérico cuando los coeficientes de filtro LP del canal primario Y no se pueden reutilizar. Para el contenido inactivo y no sonoro, solo se codifica la forma del espectro.

15 En el módulo 855 de codificación, la codificación de contenido inactivo comprende (a) codificación de ganancia de banda espectral en el dominio de frecuencia más relleno de ruido y (b) codificación de los coeficientes de filtro LP del canal secundario cuando sea necesario, como se describe respectivamente en (a) cláusulas 5.2.3.5.7 y 5.2.3.5.11 y (b) cláusula 5.2.2.1 de la Referencia [1]. El contenido inactivo se puede codificar a una tasa de bits tan baja como 1.5 kb/s.

20 En el módulo 855 de codificación, la codificación sorda del canal secundario X es similar a la codificación inactiva del canal secundario X, con la excepción de que la codificación sorda utiliza un número adicional de bits para la cuantificación de los coeficientes de filtro LP del canal secundario los cuales están codificados para el canal secundario no sonoro.

25 El modelo de codificación genérica de media banda se construye de manera similar a ACELP como se describe en la cláusula 5.2.3.1 de la Referencia [1], pero se utiliza con sólo dos (2) subtramas por trama. Por lo tanto, para hacerlo, el residual como se describe en la cláusula 5.2.3.1.1 de la Referencia [1], la memoria del libro de códigos adaptativos como se describe en la cláusula 5.2.3.1.4 de la Referencia [1] y el canal secundario de entrada se muestrean primero por un factor 2. Los coeficientes de filtro LP también se modifican para representar el dominio de muestreo descendente en lugar de la frecuencia de muestreo de 12.8 kHz utilizando una técnica como se describe en la cláusula 5.4.4.2 de la Referencia [1].

30 Después de la búsqueda ACELP, se realiza una extensión de ancho de banda en el dominio de frecuencia de la excitación. La extensión del ancho de banda replica primero las energías de la banda espectral más baja en la banda más alta. Para replicar las energías de la banda espectral, la energía de las primeras nueve (9) bandas espectrales, $G_{bd}(l)$, se encuentran como se describe en la cláusula 5.2.3.5.7 de la Referencia [1] y las últimas bandas se rellenan como se muestra en la relación (18):

35
$$G_{bd}(i) = G_{bd}(16 - i - 1), \quad \text{para } i = 8, \dots, 15. \quad (18)$$

Entonces, el contenido de alta frecuencia del vector de excitación representado en el dominio $F_D(k)$ de frecuencia como se describe en la cláusula 5.2.3.5.9 de la Referencia [1] se rellena utilizando el contenido de frecuencia de banda inferior utilizando la relación (19):

$$f_a(k) = f_a(k - P_b), \quad \text{para } k = 128, \dots, 255, \quad (19)$$

40 donde el desplazamiento de tono, P_b , se basa en un múltiplo de la información de tono como se describe en la cláusula 5.2.3.1.4.1 de la Referencia [1] y se convierte en un desplazamiento de intervalos de frecuencia como se muestra en la relación (20):

$$P_b = \begin{cases} \frac{8 \cdot \left(\frac{F_s}{T}\right)}{F_r}, & \bar{T} > 64 \\ \frac{4 \cdot \left(\frac{F_s}{T}\right)}{F_r}, & \bar{T} \leq 64 \end{cases}, \quad (20)$$

45 donde \bar{T} representa un promedio de la información de tono decodificada por subtrama, F_s es la frecuencia de muestreo interna, 12.8 kHz en esta realización de ejemplo, y F_r es la resolución de frecuencia.

Los parámetros de codificación encontrados durante la codificación inactiva de baja tasa, la codificación sorda de baja tasa o la codificación genérica de media banda realizada en el módulo 855 de codificación del modelo de dos (2) subtramas se utilizan entonces para construir el flujo 206/306 de bits del canal secundario enviado al multiplexor 254/354 para su inclusión en el flujo 207/307 de bits multiplexado.

50 *c. Implementación alternativa de la codificación de baja tasa de bits del canal secundario*

La codificación del canal secundario X se puede lograr de manera diferente, con el mismo objetivo de utilizar un número mínimo de bits a la vez que se logra la mejor calidad posible y se mantiene una firma constante. La codificación del canal secundario X puede ser impulsada en parte por el presupuesto de bits disponible, independientemente de la reutilización potencial de los coeficientes de filtro LP y la información de tono. Además, la codificación del modelo de dos (2) subtramas (operación 805) puede ser de media banda o de banda completa. En esta implementación alternativa de la codificación de baja tasa de bits del canal secundario, los coeficientes de filtro LP y/o la información de tono del canal primario se pueden reutilizar y la codificación del modelo de dos (2) subtramas se puede elegir con base en el presupuesto de bits disponible para codificar el canal secundario X. Además, la codificación del modelo de 2 subtramas que se presenta a continuación se ha creado duplicando la longitud de la subtrama en lugar de muestrear hacia abajo/hacia arriba sus parámetros de entrada/salida.

La Figura 15 es un diagrama de bloques que ilustra de manera simultánea un método de codificación de sonido estéreo alternativo y un sistema de codificación de sonido estéreo alternativo. El método y sistema de codificación de sonido estéreo de la Figura 15 incluyen diversas de las operaciones y módulos del método y sistema de la Figura 8, identificados utilizando los mismos números de referencia y cuya descripción no se repite en el presente documento por brevedad. Además, el método de codificación de sonido estéreo de la Figura 15 comprende una operación 1501 de procesamiento previo aplicada al canal primario Y antes de su codificación en la operación 202/302, una operación 1502 de análisis de coherencia de tono, una operación 1504 de decisión sorda/inactiva, una/operación 1505 de decisión de codificación inactiva, y una operación 1506 de decisión de modelo de 2/4 subtramas.

Las suboperaciones 1501, 1502, 1503, 1504, 1505 y 1506 son realizadas respectivamente por un procesador 1551 previo similar al procesador 851 previo de baja complejidad, un analizador 1552 de coherencia de tono, un estimador 1553 de asignación de bits, un módulo 1554 de decisión sorda/inactiva, un módulo 1555 de decisión de codificación no sonoro/inactivo y un módulo 1556 de decisión de modelo de 2/4 subtramas.

Para realizar la operación 1502 de análisis de coherencia de tono, el analizador 1552 de coherencia de tono es suministrado por los procesadores 851 y 1551 previos con tonos de bucle abierto de los canales primario Y y secundario X, respectivamente $OL_{\text{tono}_{\text{primero}}}$ y $OL_{\text{tono}_{\text{segundo}}}$. El analizador 1552 de coherencia de tono de la Figura 15 se muestra con mayor detalle en la Figura 16, el cual es un diagrama de bloques que ilustra de manera simultánea suboperaciones de la operación 1502 de análisis de coherencia de tono y módulos del analizador 1552 de coherencia de tono.

La operación 1502 de análisis de coherencia de tono realiza una evaluación de la similitud de los tonos de bucle abierto entre el canal primario Y y el canal secundario X para decidir en qué circunstancias se puede reutilizar el tono de bucle abierto primario en la codificación del canal secundario X. Para este fin, la operación 1502 de análisis de coherencia de tono comprende una suboperación 1601 de suma de tonos de bucle abierto de canal primario realizada por un sumador 1651 de tonos de bucle abierto de canal primario, y una suboperación 1602 de suma de tonos de bucle abierto de canal secundario realizada por un sumador 1652 de tonos de bucle abierto de canal secundario. La suma del sumador 1652 se resta (suboperación 1603) de la suma del sumador 1651 utilizando un sustractor 1653. El resultado de la resta de la suboperación 1603 proporciona una coherencia de tono estéreo. Como un ejemplo no limitativo, las sumas en las suboperaciones 1601 y 1602 se basan en tres (3) tonos de bucle abierto consecutivos, previos disponibles para cada canal Y y X. Los tonos de bucle abierto se pueden calcular, por ejemplo, como se define en la cláusula 5.1.10 de la Referencia [1]. La coherencia S_{pc} de tono estéreo se calcula en las suboperaciones 1601, 1602 y 1603 utilizando la relación (21):

$$S_{pc} = \left| \sum_{i=0}^2 p_{p(i)} - \sum_{i=0}^2 p_{s(i)} \right| \quad (21)$$

donde $P_{p(s)(i)}$ representa los tonos de bucle abierto de los canales primario Y y secundario X e i representa la posición de los tonos de bucle abierto.

Cuando la coherencia de tono estéreo está por debajo de un umbral Δ predeterminado, se puede permitir la reutilización de la información de tono del canal primario Y dependiendo de un presupuesto de bits disponible para codificar el canal secundario X. Además, dependiendo del presupuesto de bits disponible, es posible limitar la reutilización de la información de tono para señales que tienen una característica sonora tanto para el canal primario Y como para el canal secundario X.

Con este fin, la operación 1502 de análisis de coherencia de tono comprende una suboperación 1604 de decisión realizada por un módulo 1654 de decisión el cual considera el presupuesto de bits disponible y las características de la señal de sonido (indicadas, por ejemplo, por los modos de codificación de canal primario y secundario). Cuando el módulo 1654 de decisión detecta que el presupuesto de bits disponible es suficiente o que las señales de sonido para los canales primario Y y secundario X no tienen característica sonora, la decisión es codificar la información de tono relacionada con el canal secundario X (1605).

Cuando el módulo 1654 de decisión detecta que el presupuesto de bits disponible es bajo para el propósito de codificar la información de tono del canal secundario X o que las señales de sonido tanto para el canal primario Y como para el secundario X tienen una característica sonora, el módulo de decisión compara la coherencia S_{pc} de tono estéreo hasta el umbral Δ . Cuando el presupuesto de bits es bajo, el umbral Δ se establece en un valor mayor en comparación con

el caso donde el presupuesto de bits es más importante (suficiente para codificar la información de tono del canal secundario X). Cuando el valor absoluto de la coherencia S_{pc} de tono estéreo es menor que o igual que el umbral Δ , el módulo 1654 decide reutilizar la información de tono del canal primario Y para codificar el canal secundario X (1607). Cuando el valor de la coherencia S_{pc} de tono estéreo es más alto que el umbral Δ , el módulo 1654 decide codificar la información de tono del canal secundario X (1605).

Asegurarse de que los canales tengan características sonoras aumenta la probabilidad de una evolución suave de tono, por lo tanto, reduce el riesgo de agregar artefactos al reutilizar el tono del canal primario. Como un ejemplo no limitativo, cuando el presupuesto de bits estéreo está por debajo de 14 kb/s y la coherencia S_{pc} de tono estéreo es inferior o igual a 6 ($\Delta = 6$), la información de tono primario se puede reutilizar en la codificación del canal secundario X. De acuerdo con otro ejemplo no limitativo, si el presupuesto de bits estéreo está por encima de 14 kb/s y por debajo de 26 kb/s, entonces tanto el canal primario Y como el secundario X se consideran sonoros y la coherencia S_{pc} de tono estéreo se compara con un umbral más bajo $\Delta = 3$, lo cual conduce a una menor tasa de reutilización de la información de tono del canal primario Y a una tasa de bits de 22 kb/s.

Volviendo a la Figura 15, el estimador 1553 de asignación de bits se suministra con el factor β a partir del mezclador 251/351 de canales, con la decisión de reutilizar los coeficientes de filtro LP del canal primario o de utilizar y codificar los coeficientes de filtro LP del canal secundario a partir del analizador 856 de coherencia de filtro LP, y con la información de tono determinada por el analizador 1552 de coherencia de tono. Dependiendo de los requisitos de codificación del canal primario y secundario, el estimador 1553 de asignación de bits proporciona un presupuesto de bits para codificar el canal primario Y al codificador 252/352 de canal primario y un presupuesto de bits para codificar el canal secundario X al módulo 1556 de decisión. En una posible implementación, para todo el contenido que no es INACTIVO, una fracción de la tasa de bits total se asigna al canal secundario. Luego, la tasa de bits del canal secundario se incrementará en una cantidad la cual está relacionada con un factor ε de normalización de energía (cambio de escala) descrito anteriormente como:

$$B_x = B_M + (0.25 \cdot \varepsilon - 0.125) \cdot (B_t - 2 \cdot B_M) \quad (21a)$$

donde B_x representa la tasa de bits asignada al canal secundario X, B_t representa la tasa de bits estéreo total disponible, B_M representa la tasa de bits mínima asignada al canal secundario y está en general alrededor del 20% de la tasa de bits estéreo total. Finalmente, ε representa el factor de normalización de energía descrito anteriormente. Por lo tanto, la tasa de bits asignada al canal primario corresponde a la diferencia entre la tasa de bits estéreo total y la tasa de bits estéreo del canal secundario. En una implementación alternativa, la asignación de tasa de bits del canal secundario se puede describir como:

$$B_x = \begin{cases} B_M + ((15 - \varepsilon_{idx}) \cdot (B_t - 2 \cdot B_M)) \cdot 0.05, & \text{si } \varepsilon_{idx} < 15 \\ B_M + ((\varepsilon_{idx} - 15) \cdot (B_t - 2 \cdot B_M)) \cdot 0.05, & \text{si } \varepsilon_{idx} \geq 15 \end{cases} \quad (21b)$$

donde de nuevo B_x representa la tasa de bits asignada al canal secundario X, B_t representa la tasa de bits estéreo total disponible y B_M representa la tasa de bits mínima asignada al canal secundario. Finalmente, ε_{idx} representa un índice transmitido del factor de normalización de energía. Por lo tanto, la tasa de bits asignada al canal primario corresponde a la diferencia entre la tasa de bits estéreo total y la tasa de bits del canal secundario. En todos los casos, para contenido INACTIVO, la tasa de bits del canal secundario se establece en la tasa de bits mínima necesaria para codificar la forma espectral del canal secundario, dando una tasa de bits en general cercana a 2 kb/s.

A la vez, el clasificador 852 de señales proporciona una clasificación de señales del canal secundario X al módulo 1554 de decisión. Si el módulo 1554 de decisión determina que la señal de sonido está inactiva o sorda, el módulo 1555 de codificación sorda/inactiva proporciona la forma espectral del canal secundario X al multiplexor 254/354. De manera alternativa, el módulo 1554 de decisión informa al módulo 1556 de decisión cuando la señal de sonido no está inactiva ni sorda. Para tales señales de sonido, utilizando el presupuesto de bits para codificar el canal secundario X, el módulo 1556 de decisión determina si hay un número suficiente de bits disponibles para codificar el canal secundario X utilizando el módulo 854 de codificación genérica del modelo de cuatro (4) subtramas; de lo contrario, el módulo 1556 de decisión selecciona codificar el canal secundario X utilizando el módulo 855 de codificación del modelo de dos (2) subtramas. Para elegir el módulo de codificación solo genérica del modelo de cuatro subtramas, el presupuesto de bits disponible para el canal secundario debe ser lo suficientemente alto como para asignar al menos 40 bits a los libros de códigos algebraicos, una vez que todo lo demás se cuantifica o reutiliza, incluido el coeficiente LP y la información de tono y ganancias.

Como se entenderá a partir de la descripción anterior, en la operación 804 de codificación solo genérica del modelo de cuatro (4) subtramas y el correspondiente módulo 854 de codificación solo genérica del modelo de cuatro (4) subtramas, para mantener la tasa de bits lo más baja posible, se utiliza una búsqueda ACELP como se describe en la cláusula 5.2.3.1 de la Referencia [1]. En el modelo de cuatro (4) subtramas de codificación solo genérica, la información de tono se puede reutilizar o no a partir del canal primario. Los parámetros de codificación encontrados durante la búsqueda ACELP en el módulo 854 de codificación genérica del modelo de cuatro (4) subtramas se utilizan

luego para construir el flujo 206/306 de bits del canal secundario y se envían al multiplexor 254/354 para su inclusión en el flujo 207/307 de bits multiplexado.

5 En la operación 805 de codificación del modelo de dos (2) subtramas alternativo y el módulo 855 de codificación del modelo de dos (2) subtramas alternativo correspondiente, el modelo de codificación genérica se construye de manera similar a ACELP como se describe en la cláusula 5.2.3.1 de la Referencia [1], pero se utiliza con solo dos (2) subtramas por trama. Por lo tanto, para hacerlo, la longitud de las subtramas se incrementa de 64 muestras a 128 muestras, manteniendo la tasa de muestreo interna en 12.8 kHz. Si el analizador 1552 de coherencia de tono ha determinado reutilizar la información de tono del canal primario Y para codificar el canal secundario X, entonces se calcula el promedio de los tonos de las dos primeras subtramas del canal primario Y y se utiliza como la estimación de tono para la primera mitad de la trama del canal secundario X. De manera similar, se calcula el promedio de los tonos de las dos últimas subtramas del canal primario Y y se utiliza para la segunda mitad de la trama del canal secundario X. Cuando se reutiliza a partir del canal primario Y, los coeficientes de filtro LP se interpolan y la interpolación de los coeficientes de filtro LP como se describe en la cláusula 5.2.2.1 de la Referencia [1] se modifica para adaptarse a un esquema de dos (2) subtramas reemplazando el primer y tercer factores de interpolación con el segundo y cuarto factores de interpolación.

10 En la realización de la Figura 15, el proceso para decidir entre las cuatro (4) subtramas y el esquema de codificación de dos (2) subtramas es impulsado por el presupuesto de bits disponible para codificar el canal secundario X. Como se mencionó anteriormente, el presupuesto de bits del canal secundario X se deriva a partir de diferentes elementos tales como el presupuesto de bits total disponible, el factor β o el factor ϵ de normalización de energía, la presencia o no de un módulo de corrección de retardo temporal (TDC), la posibilidad o no de reutilizar los coeficientes de filtro LP y/o la información de tono a partir del canal primario Y.

20 La tasa de bits mínima absoluta utilizada por el modelo de codificación de dos (2) subtramas del canal secundario X cuando tanto los coeficientes de filtro LP como la información de tono se reutilizan a partir del canal primario Y es de aproximadamente de 2 kb/s para una señal genérica a la vez que es de aproximadamente 3.6 kb/s para el esquema de codificación de cuatro (4) subtramas. Para un codificador similar a ACELP, que utiliza un modelo de codificación de dos (2) o cuatro (4) subtramas, una gran parte de la calidad proviene del número de bits que se pueden asignar a la búsqueda del libro de códigos algebraicos (ACB) como se define en cláusula 5.2.3.1.5 de la referencia [1].

25 Luego, para maximizar la calidad, la idea es comparar el presupuesto de bits disponible tanto para la búsqueda del libro de códigos algebraico (ACB) de cuatro (4) subtramas y la búsqueda del libro de códigos algebraico (ACB) de dos (2) subtramas, después de eso, todo lo que se codificará es tomado en cuenta. Por ejemplo, si, para una trama específica, hay 4 kb/s (80 bits por trama de 20 ms) disponibles para codificar el canal secundario X y el coeficiente de filtro LP se puede reutilizar a la vez que se debe transmitir la información de tono. Luego se elimina de los 80 bits, la cantidad mínima de bits para codificar la señalización del canal secundario, la información de tono del canal secundario, las ganancias, y el libro de códigos algebraico tanto para las dos (2) subtramas como para las cuatro (4) subtramas, para obtener el presupuesto de bits disponible para codificar el libro de códigos algebraico. Por ejemplo, se elige el modelo de codificación de cuatro (4) subtramas si al menos 40 bits están disponibles para codificar el libro de códigos algebraico de cuatro (4) subtramas; de lo contrario, se utiliza el esquema de dos (2) subtramas.

3) *Aproximación de la señal mono a partir de un flujo de bits parcial* (El siguiente aspecto completo es útil para comprender la invención, pero no está cubierto por las reivindicaciones)

30 Como se describe en la descripción anterior, la mezcla descendente en el dominio del tiempo es compatible con mono, lo que significa que en el caso de una estructura incorporada, donde el canal primario Y está codificado con un códec heredado (debe tenerse en cuenta que, como se menciona en la descripción anterior, se puede utilizar cualquier tipo de codificador adecuado como codificador (252/352) de canal primario) y los bits estéreo se agregan al flujo de bits del canal primario, los bits estéreo se pueden eliminar y un decodificador heredado podría crear una síntesis que sea subjetivamente cercana a una hipotética mono síntesis. Para hacerlo, se necesita una normalización de energía simple en el lado del codificador, antes de codificar el canal primario Y. Al volver a escalar la energía del canal primario Y a un valor lo suficientemente cercano a una energía de una versión de señal monofónica del sonido, la decodificación del canal primario Y con un decodificador heredado puede ser similar a la decodificación por el decodificador heredado de la versión de señal monofónica del sonido. La función de la normalización de energía está directamente relacionada con la diferencia de correlación linealizada a largo plazo.

$$G'_{LR}(t)$$

se calcula utilizando la relación (7) y se calcula utilizando la relación (22):

$$\epsilon = -0.485 \cdot G'_{LR}(t)^2 + 0.9765 \cdot G'_{LR}(t) + 0.5. \quad (22)$$

55 El nivel de normalización se muestra en la Figura 5. En la práctica, en lugar de utilizar la relación (22), se utiliza una tabla de búsqueda que relaciona los valores de normalización ϵ a cada valor posible del factor β (31 valores en esta realización de ejemplo). Incluso si esta etapa adicional no sea necesaria cuando se codifica una señal de sonido

estéreo, por ejemplo, voz y/o audio, con el modelo integrado, esto puede ser útil cuando se decodifica solo la señal mono sin decodificar los bits estéreo.

4) *Decodificación y mezcla estéreo* (El siguiente aspecto ilustra una implementación de la invención reivindicada)

5 La Figura 10 es un diagrama de bloques que ilustra de manera simultánea un método de decodificación de sonido estéreo y un sistema de decodificación de sonido estéreo. La Figura 11 es un diagrama de bloques que ilustra características adicionales del método de decodificación de sonido estéreo y del sistema de decodificación de sonido estéreo de la Figura 10.

10 El método de decodificación de sonido estéreo de las Figuras 10 y 11 comprende una operación 1007 de demultiplexación implementada por un demultiplexor 1057, una operación 1004 de decodificación de canal primario implementada por un decodificador 1054 de canal primario, una operación 1005 de decodificación de canal secundario implementada por un decodificador 1055 de canal secundario, y una operación 1006 de mezcla ascendente en el dominio del tiempo implementada por un mezclador 1056 ascendente de canal en el dominio del tiempo. La operación 1005 de decodificación del canal secundario comprende, como se muestra en la Figura 11, una operación 1101 de decisión implementada por un módulo 1151 de decisión, una operación 1102 de decodificación genérica de cuatro (4) subtramas implementada por un decodificador 1152 genérico de cuatro (4) subtramas, y una operación 1103 de decodificación genérica/sorda/inactiva de dos (2) subtramas implementada por un decodificador 1153 genérico/no sonoro/inactivo de dos (2) subtramas.

20 En el sistema de decodificación de sonido estéreo, se recibe un flujo 1001 de bits a partir de un codificador. El demultiplexor 1057 recibe el flujo 1001 de bits y extrae de él los parámetros de codificación del canal primario Y (flujo 1002 de bits), los parámetros de codificación del canal secundario X (flujo 1003 de bits), y el factor β suministrado al decodificador 1054 de canal primario, al decodificador 1055 de canal secundario y al mezclador 1056 ascendente de canal. Como se mencionó anteriormente, el factor β se utiliza como un indicador tanto para el codificador 252/352 de canal primario como para el codificador 253/353 de canal secundario para determinar la asignación de tasa de bits, por lo tanto, el decodificador 1054 de canal primario y el decodificador 1055 de canal secundario están reutilizando el factor β para decodificar correctamente el flujo de bits.

30 Los parámetros de codificación del canal primario corresponden al modelo de codificación ACELP en la tasa de bits recibida y podrían estar relacionados con un codificador EVS heredado o modificado (debe tenerse en cuenta aquí que, como se menciona en la descripción anterior, cualquier tipo de codificador adecuado se puede utilizar como codificador 252 de canal primario). El decodificador 1054 de canal primario se suministra con el flujo 1002 de bits para decodificar los parámetros de codificación del canal primario (modo₁ de códec, β , LPC₁, tono₁, índices₁ de libro de códigos fijos, y ganancias₁ como se muestra en la Figura 11) utilizando un método similar a la Referencia [1] para producir un canal primario Y' decodificado.

Los parámetros de codificación del canal secundario utilizados por el decodificador 1055 de canal secundario corresponden al modelo utilizado para codificar el canal segundo X y pueden comprender:

35 (a) El modelo de codificación genérico con reutilización de los coeficientes de filtro LP (LPC₁) y/u otros parámetros de codificación (tal como, por ejemplo, el retraso de tono tono₁) a partir del canal primario Y. El decodificador 1152 genérico de cuatro (4) subtramas (Figura 11) del decodificador 1055 de canal secundario se suministra con los coeficientes (LPC₁) de filtro LP y/u otros parámetros de codificación (tal como, por ejemplo, el retraso de tono tono₁) a partir del canal primario Y a partir del decodificador 1054 y/o con el flujo 1003 de bits (β , tono₂, índices₂ de libro de códigos fijos, y ganancias₂ como se muestra en la Figura 11) y utiliza un método inverso al del módulo 854 de codificación (Figura 8) para producir el canal secundario X' decodificado.

40 (b) Otros modelos de codificación pueden o no reutilizar los coeficientes (LPC₁) de filtro LP y/u otros parámetros de codificación (tal como, por ejemplo, el retraso de tono tono₁) a partir del canal primario Y, incluido el modelo de codificación genérica de media banda, el modelo de codificación sorda de baja tasa, y el modelo de codificación inactiva de baja tasa. Como un ejemplo, el modelo de codificación inactivo puede reutilizar los coeficientes LPC₁ de filtro LP del canal primario. El decodificador 1153 genérico/no sonoro/inactivo de dos (2) subtramas (Figura 11) del decodificador 1055 de canal secundario se suministra con los coeficientes (LPC₁) de filtro LP y/u otros parámetros de codificación (tal como, por ejemplo, el retraso de tono tono₁) a partir del canal primario Y y/o con los parámetros de codificación del canal secundario del flujo 1003 de bits (modo₂ códec, β , LPC₂, tono₂, índices₂ de libro de códigos fijos, y ganancias₂ como se muestra en la Figura 11) y utiliza métodos inversos a los del módulo 855 de codificación (Figura 8) para producir el canal secundario X' decodificado.

50 Los parámetros de codificación recibidos correspondientes al canal secundario X (flujo 1003 de bits) contienen información (modo₂ códec) relacionada con el modelo de codificación que se está utilizando. El módulo 1151 de decisión utiliza esta información (modo₂ códec) para determinar e indicar al decodificador 1152 genérico de cuatro (4) subtramas y al decodificador 1153 genérico/no sonoro/inactivo de dos (2) subtramas cual modelo de codificación se va a utilizar.

55 En el caso de una estructura incorporada, el factor β se utiliza para recuperar el índice de escala de energía que está almacenado en una tabla de búsqueda (no se muestra) en el lado del decodificador y se utiliza para volver a escalar

el canal primario Y' antes de realizar la operación 1006 de mezcla ascendente en el dominio del tiempo. Finalmente el factor β se suministra al mezclador 1056 ascendente de canal y se utiliza para mezclar los canales primario Y' y secundario X' decodificados. La operación 1006 de mezcla ascendente en el dominio del tiempo se realiza como la inversa de las relaciones (9) y (10) de mezcla descendente para obtener los canales derecho R' e izquierdo L' decodificados, utilizando las relaciones (23) y (24):

$$L'(n) = \frac{\beta(t) \cdot Y'(n) - \beta(t) \cdot X'(n) + X'(n)}{2 \cdot \beta(t)^2 - 2 \cdot \beta(t) + 1}, \quad (23)$$

$$R'(n) = \frac{-\beta(t) \cdot (Y'(n) + X'(n)) + Y'(n)}{2 \cdot \beta(t)^2 - 2 \cdot \beta(t) + 1}, \quad (24)$$

donde $n = 0, \dots, N-1$ es el índice de la muestra en la trama y t es el índice de la trama.

5) Integración de la codificación en el dominio del tiempo y en el dominio de la frecuencia (El siguiente aspecto completo es útil para comprender la invención, pero no está cubierto por las reivindicaciones)

Para las aplicaciones de la presente técnica donde se utiliza un modo de codificación en el dominio de la frecuencia, también se contempla realizar la mezcla descendente del tiempo en el dominio de la frecuencia para ahorrar algo de complejidad o para simplificar el flujo de datos. En tales casos, se aplica el mismo factor de mezcla a todos los coeficientes espectrales con el fin de mantener las ventajas de la mezcla descendente en el dominio del tiempo. Se puede observar que esto es una desviación de la aplicación de coeficientes espectrales por banda de frecuencia, como en el caso de la mayoría de las aplicaciones de mezcla descendente en el dominio de la frecuencia. El mezclador 456 descendente puede adaptarse para calcular las relaciones (25.1) y (25.2):

$$F_Y(k) = F_R(k) \cdot (1 - \beta(t)) + F_L(k) \cdot \beta(t) \quad (25.1)$$

$$F_X(k) = F_L(k) \cdot (1 - \beta(t)) - F_R(k) \cdot \beta(t), \quad (25.2)$$

donde $F_R(k)$ representa un coeficiente k de frecuencia del canal derecho R y, de manera similar, $F_L(k)$ representa un coeficiente k de frecuencia del canal izquierdo L. Los canales primario Y y secundario X se calculan entonces aplicando una transformada de frecuencia inversa para obtener la representación en el tiempo de las señales mezcladas descendentes.

Las Figuras 17 y 18 muestran posibles implementaciones del método y sistema de codificación estéreo en el dominio del tiempo que utilizan una mezcla descendente en el dominio de la frecuencia capaz de cambiar entre la codificación del dominio del tiempo y la codificación del dominio de la frecuencia de los canales primario Y y secundario X.

En la Figura 17 se muestra una primera variante de dicho método y sistema, la cual es un diagrama de bloques que ilustra el método y el sistema de codificación estéreo de manera concurrente que utiliza un cambio descendente en el dominio del tiempo con una capacidad de operar en el dominio del tiempo y en el dominio de la frecuencia.

En la Figura 17, el método y sistema de codificación estéreo incluye diversas operaciones y módulos descritos anteriormente con referencia a las figuras anteriores e identificados por los mismos números de referencia. Un módulo 1751 de decisión (operación 1701 de decisión) determina si los canales izquierdo L' y derecho R' del corrector 1750 de retardo temporal deben codificarse en el dominio del tiempo o en el dominio de la frecuencia. Si se selecciona la codificación en el dominio del tiempo, el método y sistema de codificación estéreo de la Figura 17 funciona de manera sustancial de la misma manera que el método y sistema de codificación estéreo de las figuras anteriores, por ejemplo y sin limitación como en la realización de la Figura 15.

Si el módulo 1751 de decisión selecciona la codificación de frecuencia, un convertidor 1752 de tiempo a frecuencia (operación 1702 de conversión de tiempo a frecuencia) convierte los canales izquierdo L' y derecho R' al dominio de frecuencia. Un mezclador 1753 descendente de dominio de frecuencia (operación 1703 de mezcla descendente de dominio de frecuencia) genera canales de dominio de frecuencia primario Y y secundario X. El canal primario del dominio de la frecuencia se convierte de nuevo al dominio del tiempo mediante un convertidor 1754 de frecuencia a tiempo (operación 1704 de conversión de frecuencia a tiempo) y el canal primario Y resultante del dominio del tiempo se aplica al codificador 252/352 del canal primario. El canal secundario X de dominio de frecuencia del mezclador 1753 descendente de dominio de frecuencia se procesa a través de un codificador 1755 paramétrico y/o residual convencional (operación 1705 de codificación paramétrica y/o residual).

La Figura 18 es un diagrama de bloques que ilustra de manera simultánea otro método y sistema de codificación estéreo que utiliza mezcla descendente en el dominio de frecuencia con una capacidad de operar en el dominio del tiempo y en el dominio de la frecuencia. En la Figura 18, el método y sistema de codificación estéreo son similares al método y sistema de codificación estéreo de la Figura 17 y solo se describirán las nuevas operaciones y módulos.

Un analizador 1851 en el dominio de tiempo (operación 1801 de análisis en el dominio de tiempo) reemplaza al mezclador 251/351 de canal en el dominio de tiempo descrito anteriormente (operación 201/301 de mezcla descendente en el dominio de tiempo). El analizador 1851 en el dominio del tiempo incluye la mayoría de los módulos de la Figura 4, pero sin el mezclador 456 en el dominio del tiempo. Por lo tanto, su función es, en gran parte, proporcionar un cálculo del factor β . Este factor β se suministra al procesador 851 previo y a los convertidores 1852 y 1853 de dominio de frecuencia a tiempo (operaciones 1802 y 1803 de conversión de dominio de frecuencia a tiempo) que, respectivamente, convierten al dominio de tiempo los canales secundario X y primario Y de dominio de frecuencia recibidos a partir del mezclador 1753 descendente en el dominio de frecuencia para la codificación en el dominio del tiempo. La salida del convertidor 1852 es por lo tanto un canal secundario X en el dominio del tiempo que se proporciona al procesador 851 previo a la vez que la salida del convertidor 1852 es un canal primario Y en el dominio del tiempo que se proporciona tanto al procesador 1551 previo como al codificador 252/352.

6) Ejemplo de configuración de hardware

La Figura 12 es un diagrama de bloques simplificado de una configuración de ejemplo de componentes de hardware que forman cada uno de los sistemas de codificación de sonido estéreo y el sistema de decodificación de sonido estéreo descritos anteriormente.

Cada uno de los sistemas de codificación de sonido estéreo y el sistema de decodificación de sonido estéreo puede implementarse como parte de un terminal móvil, como parte de un reproductor multimedia portátil, o en cualquier dispositivo similar. Cada sistema de codificación de sonido estéreo y sistema de decodificación de sonido estéreo (identificado como 1200 en la Figura 12) comprende una entrada 1202, una salida 1204, un procesador 1206 y una memoria 1208.

La entrada 1202 está configurada para recibir los canales izquierdo L y derecho R de la señal de sonido estéreo de entrada en forma digital o analógica en el caso del sistema de codificación de sonido estéreo, o el flujo 1001 de bits en el caso del sistema de decodificación de sonido estéreo. La salida 1204 está configurada para suministrar el flujo 207/307 de bits multiplexado en el caso del sistema de codificación de sonido estéreo o el canal izquierdo L' decodificado y el canal derecho R' en el caso del sistema de decodificación de sonido estéreo. La entrada 1202 y la salida 1204 pueden implementarse en un módulo común, por ejemplo, un dispositivo de entrada/salida en serie.

El procesador 1206 está conectado de manera operativa a la entrada 1202, a la salida 1204, y a la memoria 1208. El procesador 1206 se realiza como uno o más procesadores para ejecutar instrucciones de código en apoyo de las funciones de los diversos módulos de cada uno de los sistemas de codificación de sonido estéreo como se muestra en la Figura 2, 3, 4, 8, 9, 13, 14, 15, 16, 17 y 18 y el sistema de decodificación de sonido estéreo como se muestra en las Figuras 10 y 11.

La memoria 1208 puede comprender una memoria no transitoria para almacenar instrucciones de código ejecutables por el procesador 1206, específicamente, una memoria legible por el procesador que comprende instrucciones no transitorias que, cuando se ejecutan, hacen que un procesador implemente las operaciones y módulos del método y sistema de codificación de sonido estéreo y el método y sistema de decodificación de sonido estéreo como se describe en la presente divulgación. La memoria 1208 también puede comprender una memoria de acceso aleatorio o búfer(es) para almacenar datos de procesamiento intermedio de las diversas funciones realizadas por el procesador 1206.

Los expertos en la técnica se darán cuenta de que la descripción del método y sistema de codificación de sonido estéreo y el método y sistema de decodificación de sonido estéreo son sólo ilustrativos y no pretenden ser limitativos de ningún modo. Otras realizaciones se les sugerirán fácilmente a aquellos expertos en la técnica que se beneficien de la presente divulgación. Además, el método y sistema de codificación de sonido estéreo divulgado y el método y sistema de decodificación de sonido estéreo pueden personalizarse para ofrecer soluciones valiosas a las necesidades y problemas existentes de codificación y decodificación de sonido estéreo.

Con el fin de dar claridad, no se muestran y describen todas las características de rutina de las implementaciones del método y sistema de codificación de sonido estéreo y el método y sistema de decodificación de sonido estéreo. Por supuesto, se apreciará que en el desarrollo de cualquier implementación real del método y sistema de codificación de sonido estéreo y el método y sistema de decodificación de sonido estéreo, es posible que se deban tomar numerosas decisiones específicas de implementación con el fin de lograr los objetivos específicos del desarrollador, tales como el cumplimiento de las restricciones relacionadas con la aplicación, el sistema, la red, y el negocio, y que estos objetivos específicos variarán de una implementación a otra y de un desarrollador a otro. Además, se apreciará que un esfuerzo de desarrollo puede ser complejo y llevar mucho tiempo, pero no obstante, sería una tarea rutinaria de ingeniería para los expertos en el campo del procesamiento de sonido que se benefician de la presente divulgación.

De acuerdo con la presente divulgación, los módulos, operaciones de procesamiento, y/o estructuras de datos descritos en el presente documento pueden implementarse utilizando diversos tipos de sistemas operativos, plataformas informáticas, dispositivos de red, programas informáticos, y/o máquinas de propósito general. Además, los expertos en la técnica reconocerán que también se pueden utilizar dispositivos de naturaleza menos general, tales como dispositivos cableados, matrices de puertas programables en campo (FPGAs), circuitos integrados específicos de aplicación (ASICs) o similares. Donde un método que comprende una serie de operaciones y suboperaciones es

implementado por un procesador, ordenador o máquina y esas operaciones y suboperaciones pueden almacenarse como una serie de instrucciones de código no transitorias legibles por el procesador, ordenador o máquina, pueden almacenarse en un medio tangible y/o no transitorio.

5 Los módulos del método y sistema de codificación de sonido estéreo y el método y decodificador de decodificación de sonido estéreo como se describe en el presente documento pueden comprender un software, firmware, hardware o cualquier combinación(es) de software, firmware, o hardware adecuado para los fines descritos en el presente documento.

10 En el método de codificación de sonido estéreo y el método de decodificación de sonido estéreo como se describe en el presente documento, las diversas operaciones y suboperaciones pueden realizarse en diversos órdenes y algunas de las operaciones y suboperaciones pueden ser opcionales.

Aunque la presente divulgación se ha descrito anteriormente a través de realizaciones ilustrativas no restrictivas de la misma, estas realizaciones pueden modificarse a voluntad dentro del alcance de las reivindicaciones adjuntas.

Referencias

Las siguientes referencias se mencionan en la presente especificación.

- 15 [1] 3GPP TS 26.445, v.12.0.0, "Codec for Enhanced Voice Services (EVS); Detailed Algorithmic Description", septiembre de 2014.
- [2] M. Neuendorf, M. Multrus, N. Rettelbach, G. Fuchs, J. Robillard, J. Lecompte, S. Wilde, S. Bayer, S. Disch, C. Helmrich, R. Lefebvre, P. Gournay, et al., "The ISO/MPEG Unified Speech and Audio Coding Estándar – Consistent High Quality for All Content Types and at All Bit Rates", J. Audio Eng. Soc., Volumen 61, número 12, páginas 956-977, diciembre de 2013.
- 20 [3] B. Bessette, R. Salami, R. Lefebvre, M. Jelinek, J. Rotola-Pukkila, J. Vainio, H. Mikkola y K. Järvinen, "The Adaptive Multi-Rate Wideband Speech Codec (AMR-WB)", Número especial de IEEE Trans. Proceso de voz y audio, volumen 10, páginas 620-636, noviembre de 2002.
- [4] R.G. van der Waal y RNJ Veldhuis, "Subband coding of stereophonic audio signals", Proceso IEEE ICASSP, volumen 5, páginas 3601-3604, abril de 1991
- 25 [5] Dai Yang, Hongmei Ai, Chris Kyriakakis y C.-C. Jay Kuo, "High-Fidelity Multichannel Audio Coding With Karhunen-Loève Transform", IEEE Trans. Proceso de Voz y Audio, volumen 11, Número 4, páginas 365-379, julio de 2003.
- [6] J. Breebaart, S. van de Par, A. Kohlrausch y E. Schuijers, "Parametric Coding of Stereo Audio", EURASIP Revista sobre Procesamiento de Señales Aplicadas, número 9, páginas 1305-1322, 2005
- 30 [7] 3GPP TS 26.290 V9.0.0, "Extended Adaptive Multi-Rate - Wideband (AMR-WB) codec; Funciones de transcodificación (versión 9)", septiembre de 2009.
- [8] Jonathan A. Gibbs, "Apparatus and Method for encoding a multi-channel audio signal", US 8577045 B2

REIVINDICACIONES

1. Un método de decodificación de sonido estéreo para decodificar los canales izquierdo y derecho de una señal de sonido estéreo, el método es caracterizado por:

5 recibir parámetros de codificación que comprenden parámetros de codificación de un canal primario, parámetros de codificación de un canal secundario, y un factor β , en donde los parámetros de codificación del canal primario comprenden coeficientes de filtro LP del canal primario;

decodificar el canal primario en respuesta a los parámetros de codificación del canal primario para proporcionar un canal primario decodificado;

10 decodificar el canal secundario utilizando uno de una pluralidad de modelos de codificación para proporcionar un canal secundario decodificado, en donde al menos uno de los modelos de codificación utiliza los coeficientes de filtro LP del canal primario para decodificar el canal secundario; y

15 mezcla ascendente en el dominio del tiempo de los canales primario y secundario decodificados utilizando el factor β para producir los canales izquierdo y derecho decodificados de la señal de sonido estéreo, en donde el factor β determina las contribuciones respectivas de los canales primario y secundario sobre la producción de los canales izquierdo y derecho.

2. Un método de decodificación de sonido estéreo como se define en la reivindicación 1, en donde al menos uno de los modelos de codificación utiliza parámetros de codificación de canal primario distintos de los coeficientes de filtro LP para decodificar el canal secundario.

20 3. Un método de decodificación de sonido estéreo como se define en la reivindicación 1 o 2, en donde los modelos de codificación comprenden un modelo de codificación genérico, un modelo de codificación no sonoro y un modelo de codificación inactivo.

4. Un método de decodificación de sonido estéreo como se define en una cualquiera de las reivindicaciones 1 a 3, en donde los parámetros de codificación del canal secundario comprenden información que identifica uno de los modelos de codificación que se utilizará tras decodificar el canal secundario.

25 5. Un método de decodificación de sonido estéreo como se define en una cualquiera de las reivindicaciones 1 a 4, que comprende recuperar un índice de escala de energía utilizando el factor β para volver a escalar el canal primario decodificado antes de realizar la mezcla ascendente en el dominio del tiempo de los canales primario y secundario decodificados.

30 6. Un método de decodificación de sonido estéreo como se define en una cualquiera de las reivindicaciones 1 a 5, en donde la mezcla ascendente en el dominio del tiempo de los canales primario y secundario decodificados utiliza las siguientes relaciones para obtener los canales izquierdo $L'(n)$ y derecho $R'(n)$ decodificados:

$$L'(n) = \frac{\beta(t) \cdot Y'(n) - \beta(t) \cdot X'(n) + X'(n)}{2 \cdot \beta(t)^2 - 2 \cdot \beta(t) + 1},$$

$$R'(n) = \frac{-\beta(t) \cdot (Y'(n) + X'(n)) + Y'(n)}{2 \cdot \beta(t)^2 - 2 \cdot \beta(t) + 1}$$

35 donde el factor $\beta(t)$ representa el factor β , $Y'(n)$ es el canal primario decodificado, $X'(n)$ es el canal secundario decodificado, $n=0, \dots, N-1$ es un índice de muestra en una trama y t es un índice de trama.

7. Un sistema de decodificación de sonido estéreo para decodificar los canales izquierdo y derecho de una señal de sonido estéreo, el sistema es caracterizado porque comprende:

40 medios para recibir parámetros de codificación que comprenden parámetros de codificación de un canal primario, parámetros de codificación de un canal secundario, y un factor β , en donde los parámetros de codificación del canal primario comprenden coeficientes de filtro LP del canal primario;

un decodificador del canal primario configurado para proporcionar un canal primario decodificado en respuesta a los parámetros de codificación del canal primario;

45 un decodificador del canal secundario configurado para proporcionar un canal secundario decodificado utilizando uno de una pluralidad de modelos de codificación, en donde al menos uno de los modelos de codificación utiliza los coeficientes de filtro LP del canal primario para decodificar el canal secundario; y

un mezclador ascendente en el dominio del tiempo de los canales primario y secundario decodificados configurado para utilizar el factor β para producir los canales izquierdo y derecho decodificados de la señal de sonido estéreo, en

donde el factor β determina las contribuciones respectivas de los canales primario y secundario sobre la producción de los canales izquierdo y derecho.

5 8. Un sistema de decodificación de sonido estéreo como se define en la reivindicación 7, en donde al menos uno de los modelos de codificación utiliza parámetros de codificación de canal primario distintos de los coeficientes de filtro LP para decodificar el canal secundario.

9. Un sistema de decodificación de sonido estéreo como se define en la reivindicación 7 u 8, en donde el decodificador de canal secundario comprende un primer decodificador que utiliza un modelo de codificación genérico, y un segundo decodificador que utiliza uno de un modelo de codificación genérico, un modelo de codificación no sonoro y un modelo de codificación inactivo.

10 10. Un sistema de decodificación de sonido estéreo como se define en una cualquiera de las reivindicaciones 7 a 9, en donde los parámetros de codificación del canal secundario comprenden información que identifica uno de los modelos de codificación que se utilizará tras decodificar el canal secundario, y en donde el sistema de decodificación de la señal de sonido estéreo comprende un módulo de decisión para indicar al primer y segundo decodificadores el modelo de codificación que se utilizará tras decodificar el canal secundario.

15 11. Un sistema de decodificación de sonido estéreo como se define en una cualquiera de las reivindicaciones 7 a 10, que comprende una tabla de búsqueda para recuperar un índice de escala de energía utilizando el factor β para volver a escalar el canal primario decodificado antes de realizar la mezcla ascendente en el dominio del tiempo de los canales primario y secundario decodificados.

20 12. Un sistema de decodificación de sonido estéreo definido en una cualquiera de las reivindicaciones 7 a 11, en donde el mezclador ascendente en el dominio del tiempo de los canales primario y secundario decodificados está configurado para utilizar las siguientes relaciones para obtener los canales izquierdo $L'(n)$ y derecho $R'(n)$ decodificados:

$$L'(n) = \frac{\beta(t) \cdot Y'(n) - \beta(t) \cdot X'(n) + X'(n)}{2 \cdot \beta(t)^2 - 2 \cdot \beta(t) + 1},$$

$$R'(n) = \frac{-\beta(t) \cdot (Y'(n) + X'(n)) + Y'(n)}{2 \cdot \beta(t)^2 - 2 \cdot \beta(t) + 1}$$

25 donde el factor $\beta(t)$ representa el factor β , $Y'(n)$ es el canal primario decodificado, $X'(n)$ es el canal secundario decodificado, $n=0, \dots, N-1$ es un índice de muestra en una trama, y t es un índice de trama.

30 13. Un sistema de decodificación de sonido estéreo como se define en una cualquiera de las reivindicaciones 7 a 12, en donde los medios para recibir los parámetros de codificación comprenden un demultiplexor para recibir un flujo de bits de un codificador y para extraer a partir del flujo de bits los parámetros de codificación del canal primario, los parámetros de codificación de la señal secundaria, y el factor β .

14. Una memoria legible por el procesador que comprende instrucciones no transitorias que, cuando se ejecutan, hacen que un procesador implemente las operaciones del método como se indica en una cualquiera de las reivindicaciones 1 a 6.

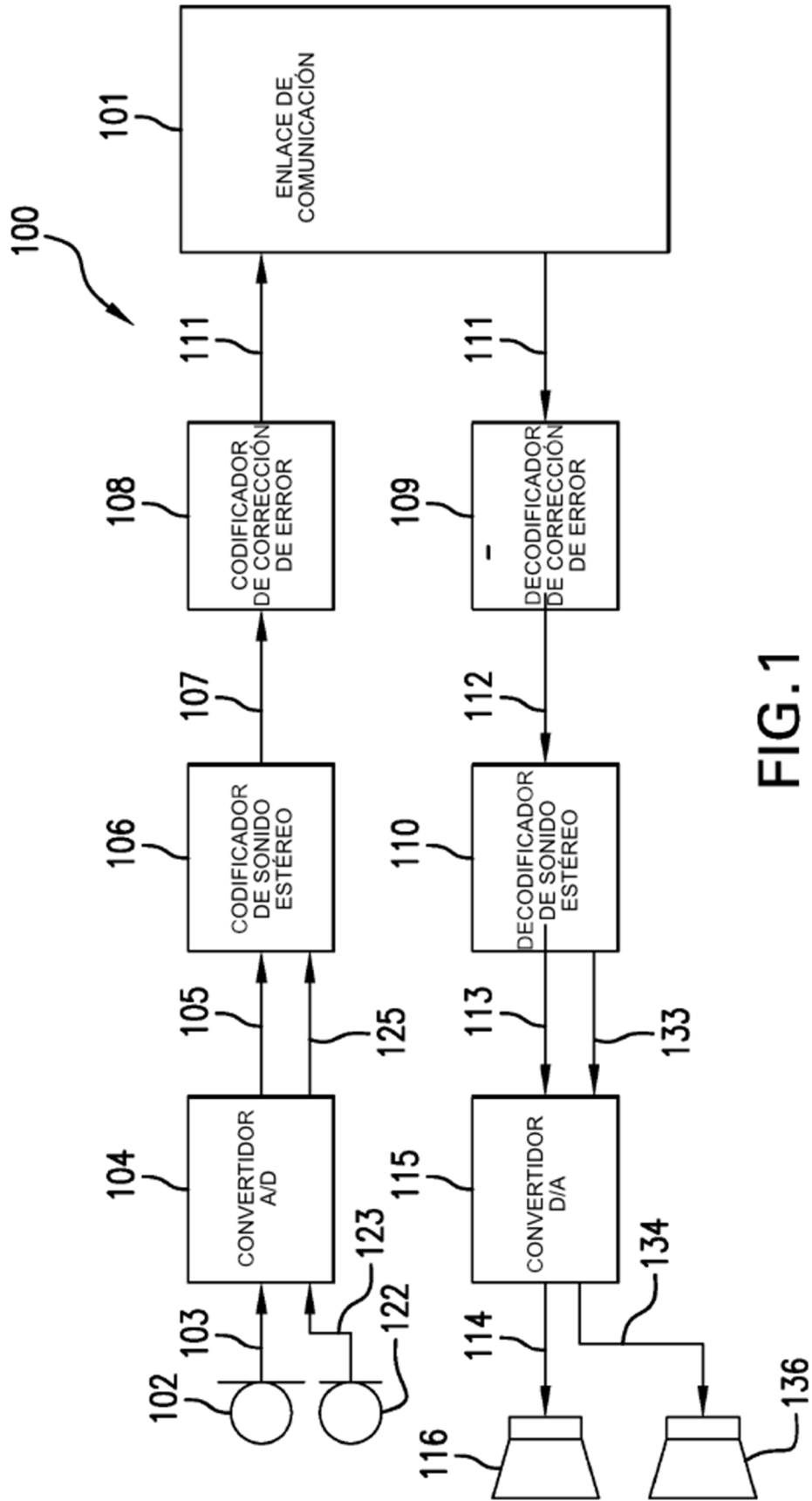


FIG.1

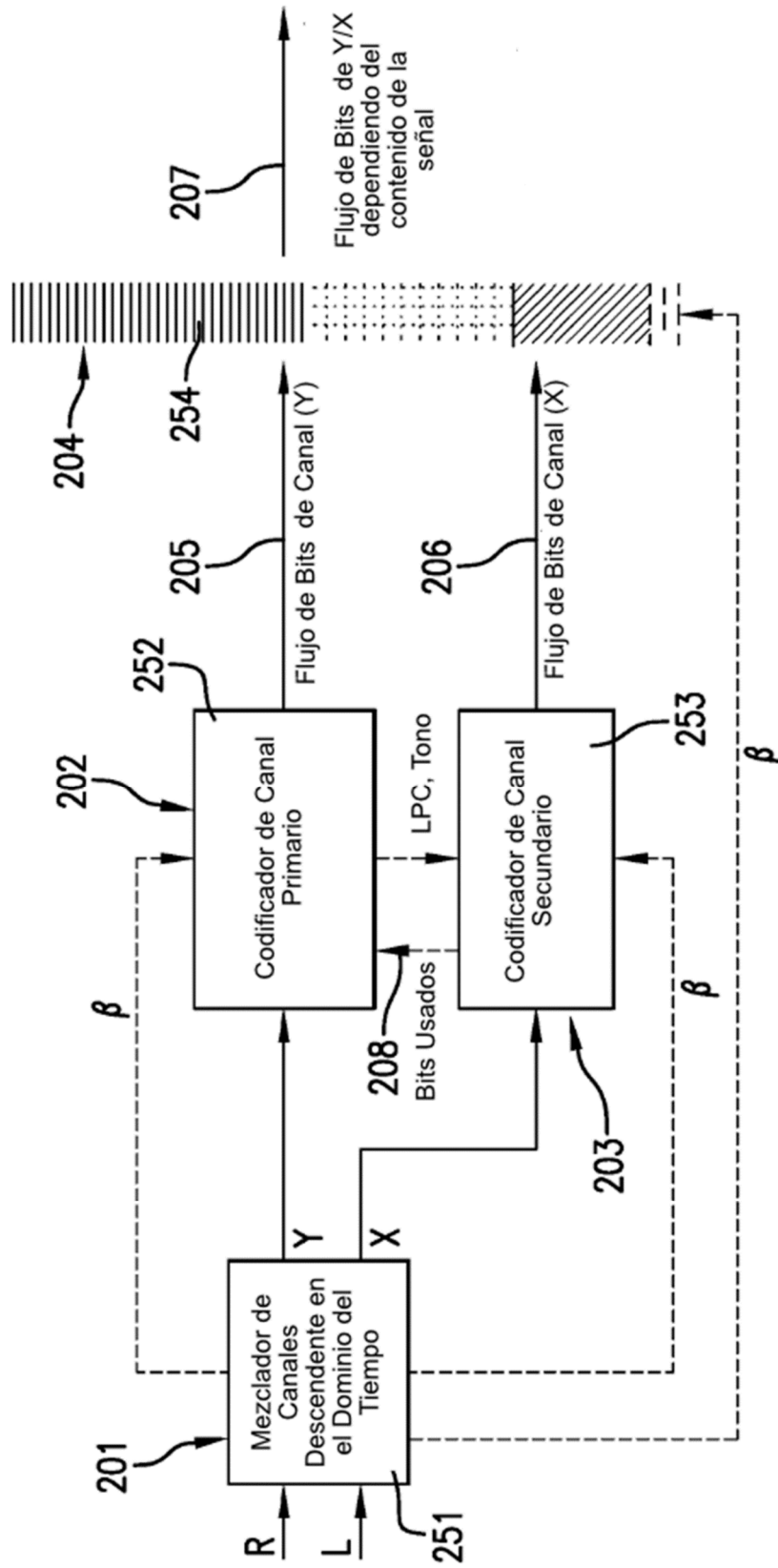


FIG.2

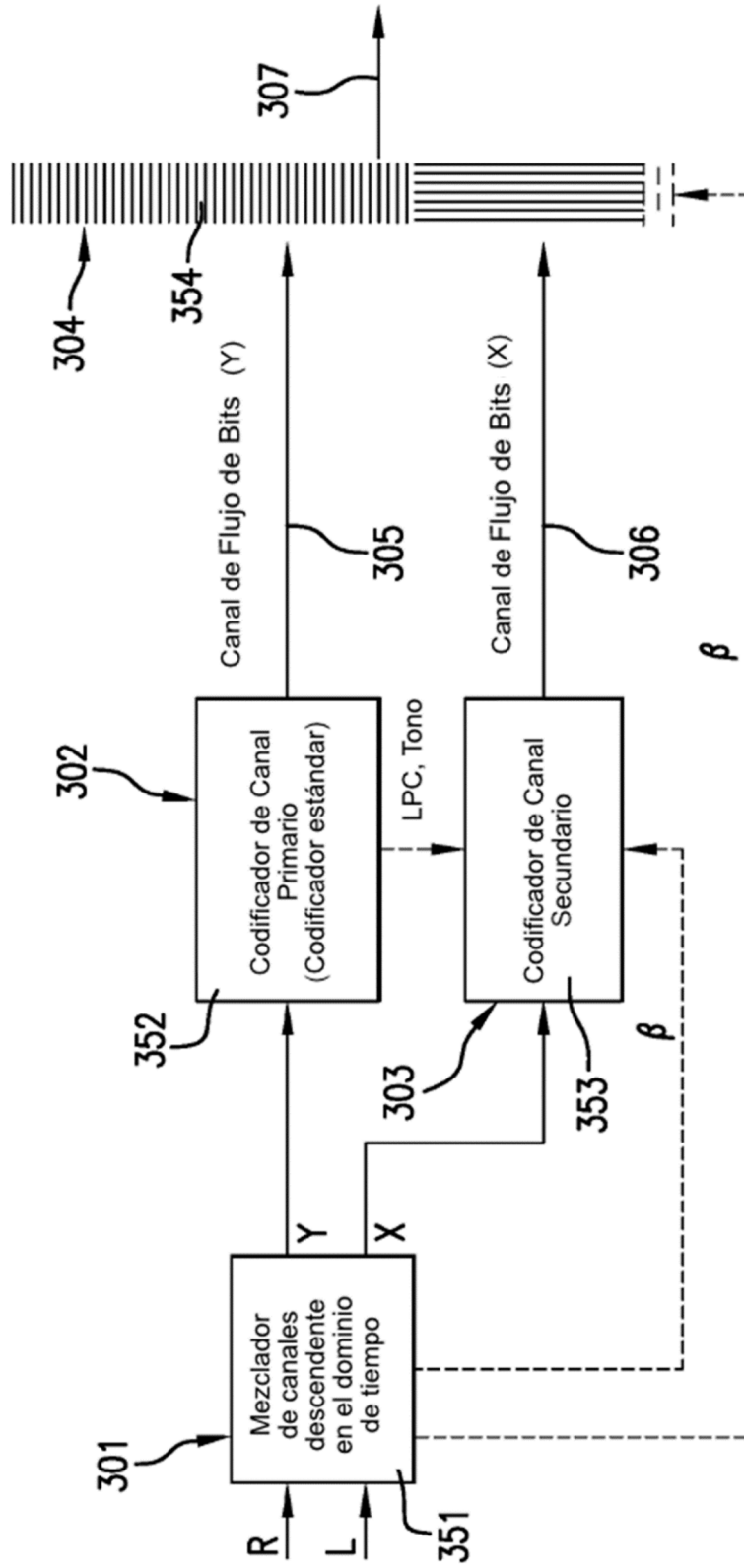


FIG.3

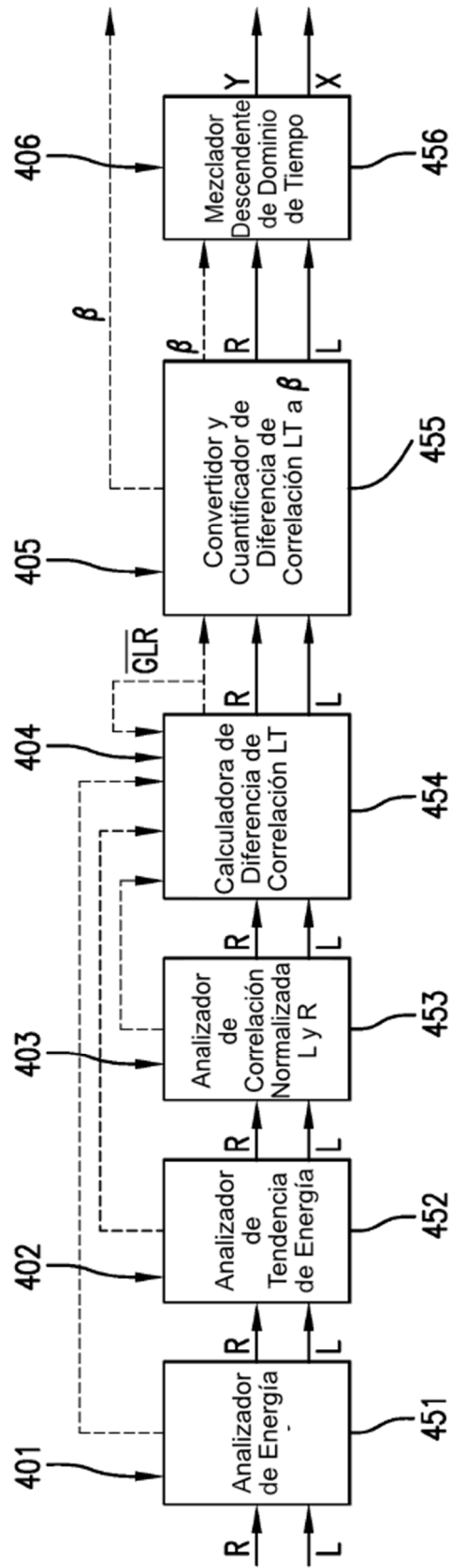


FIG.4

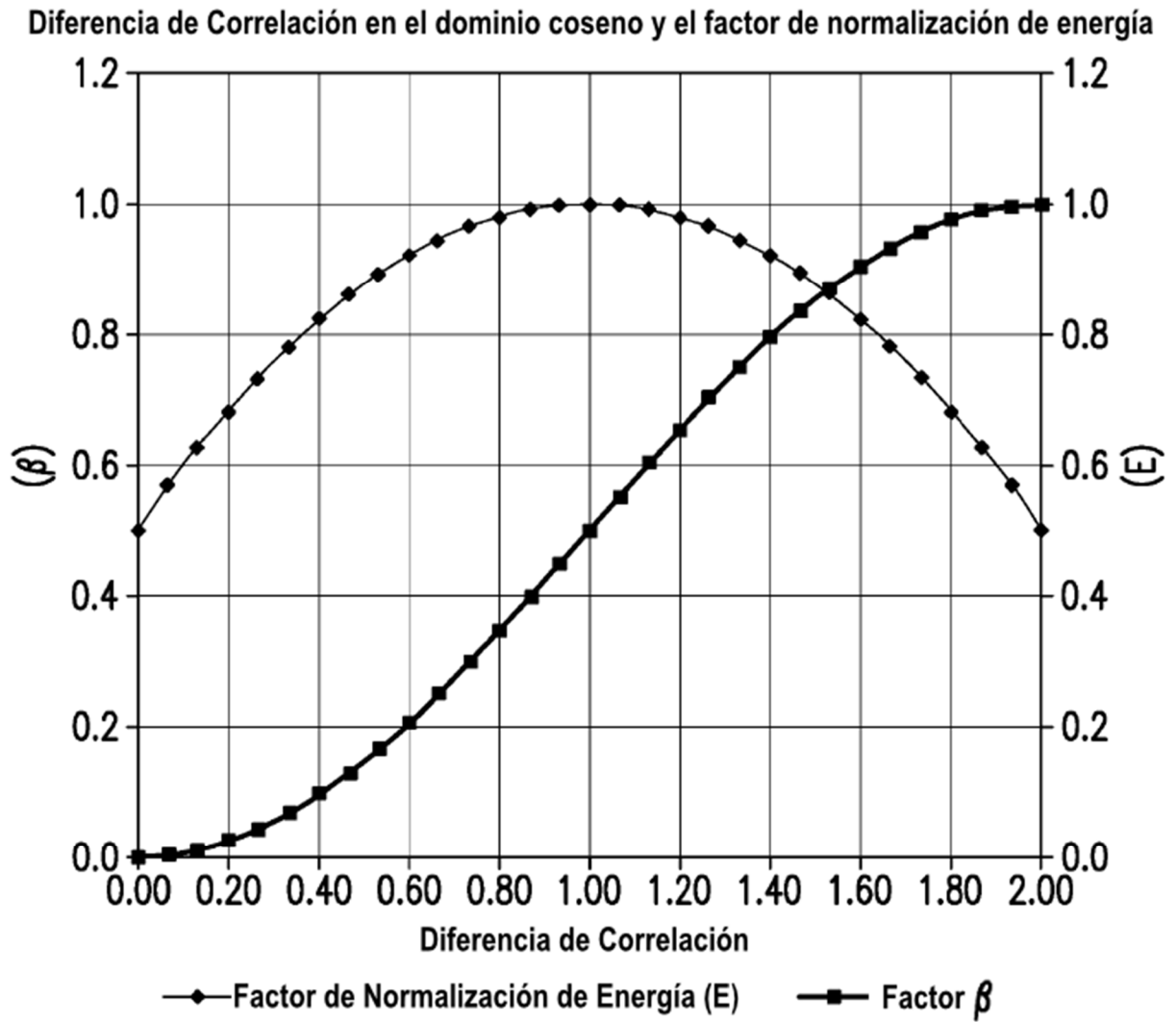


FIG.5

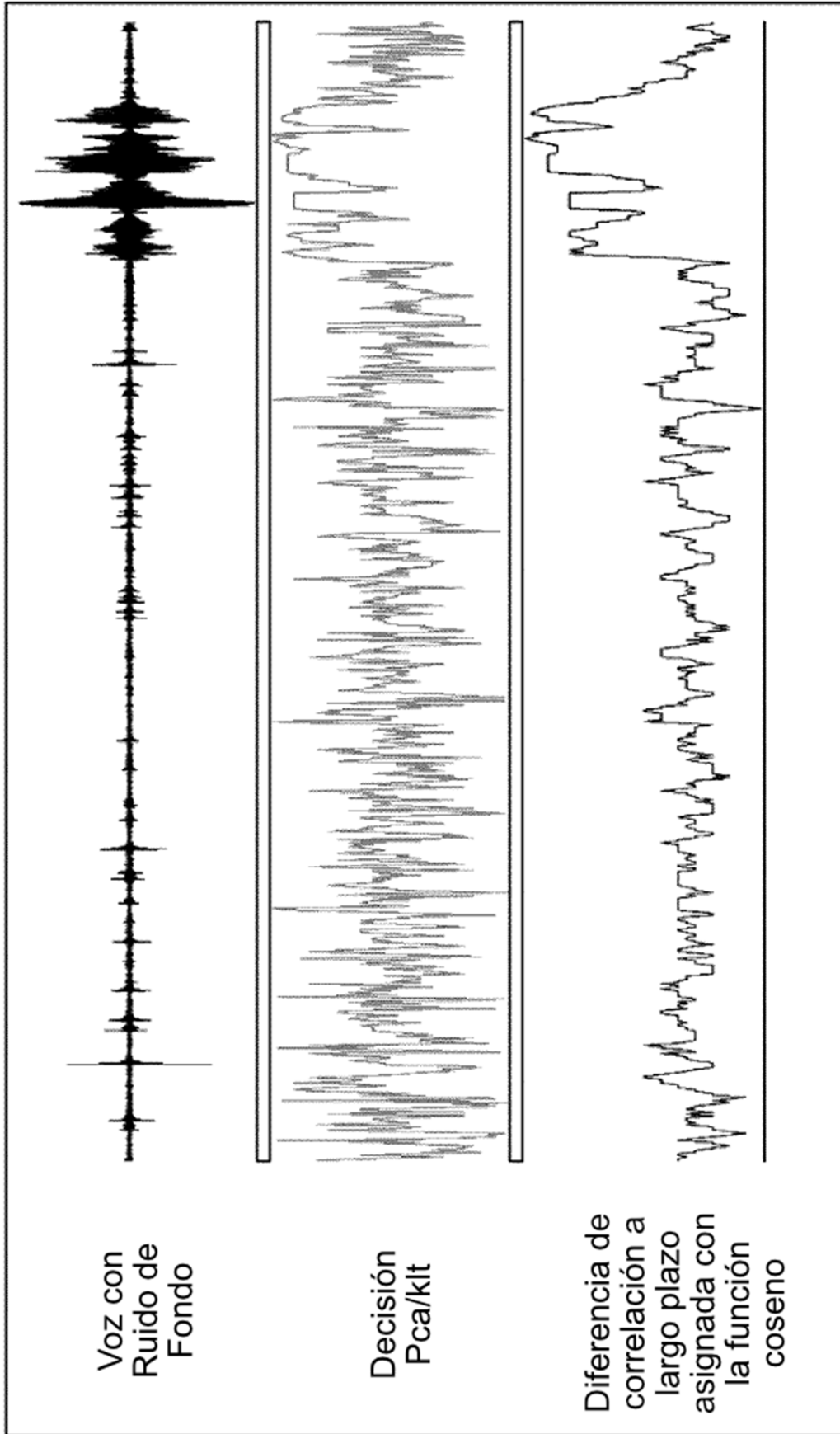


FIG.6

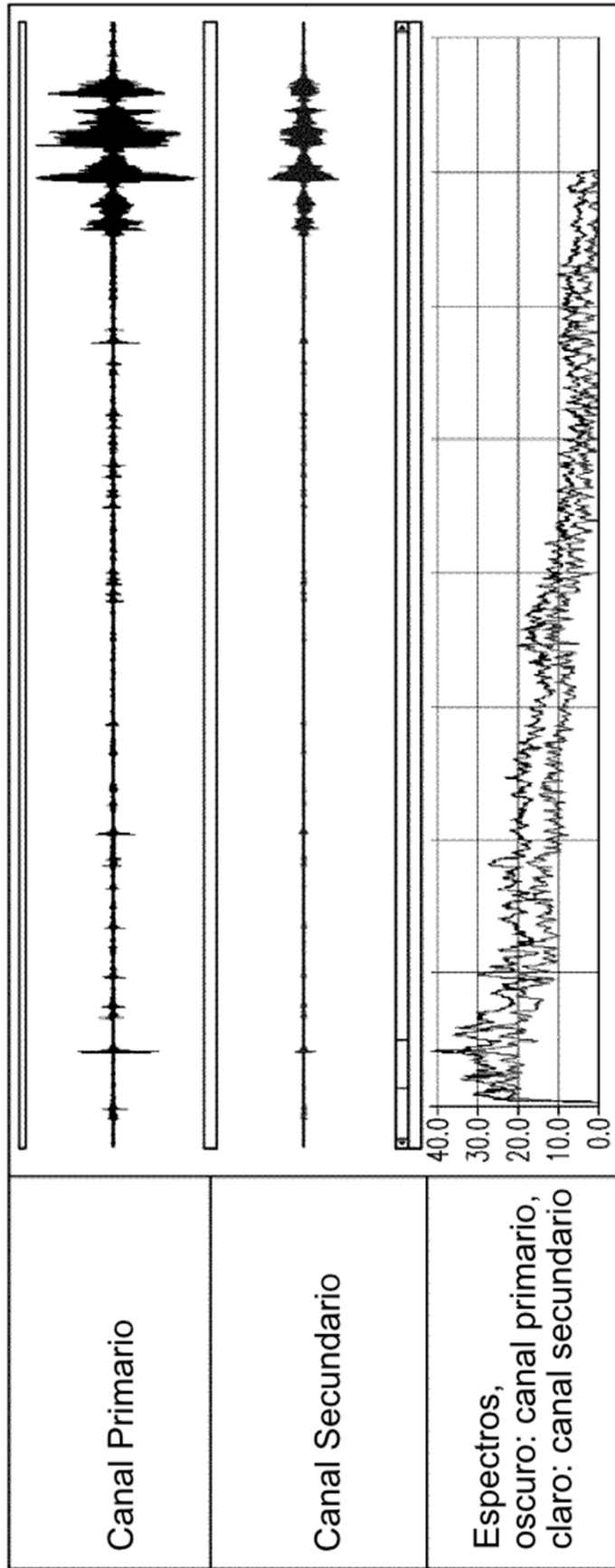


FIG.7

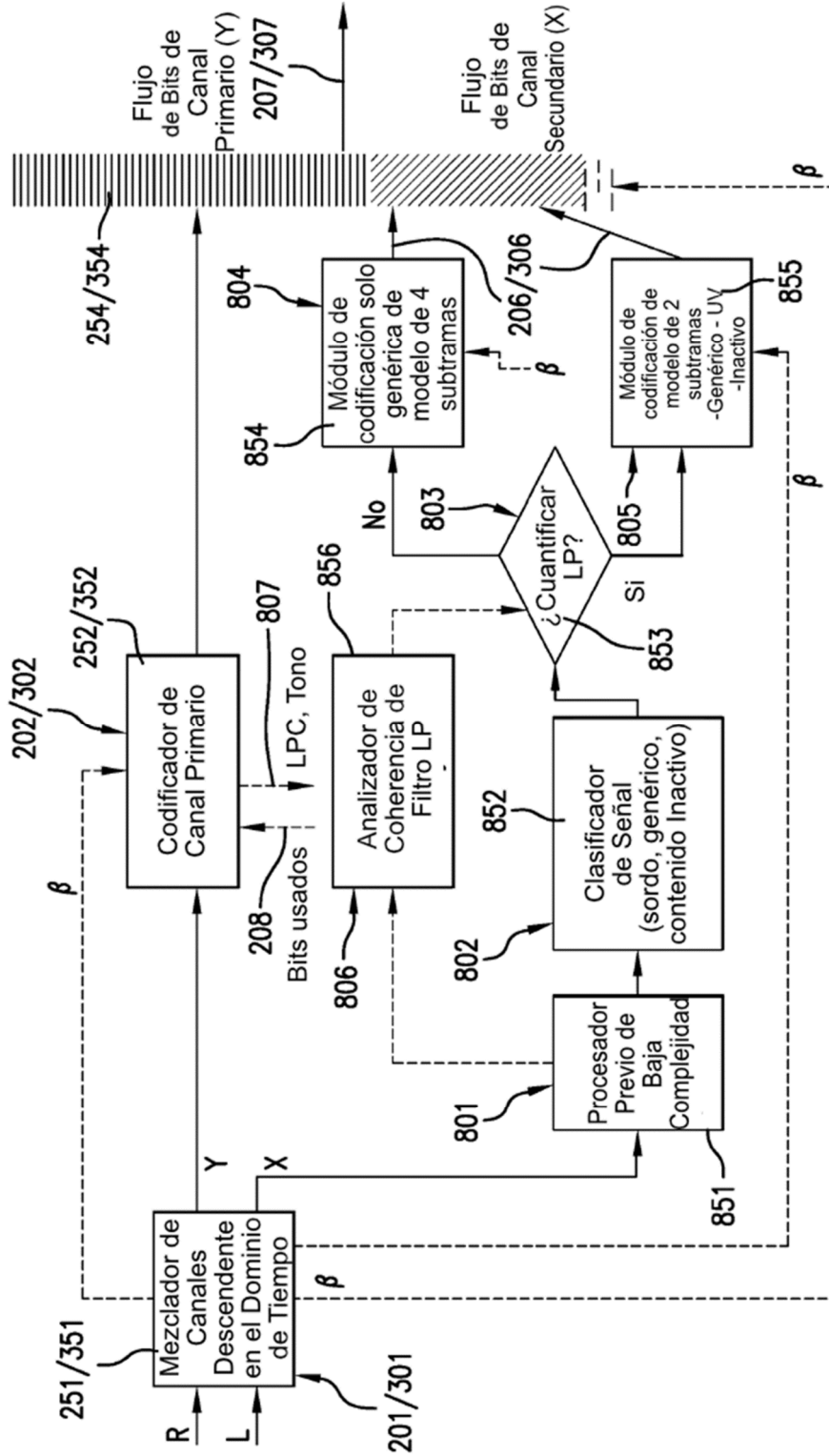


FIG.8

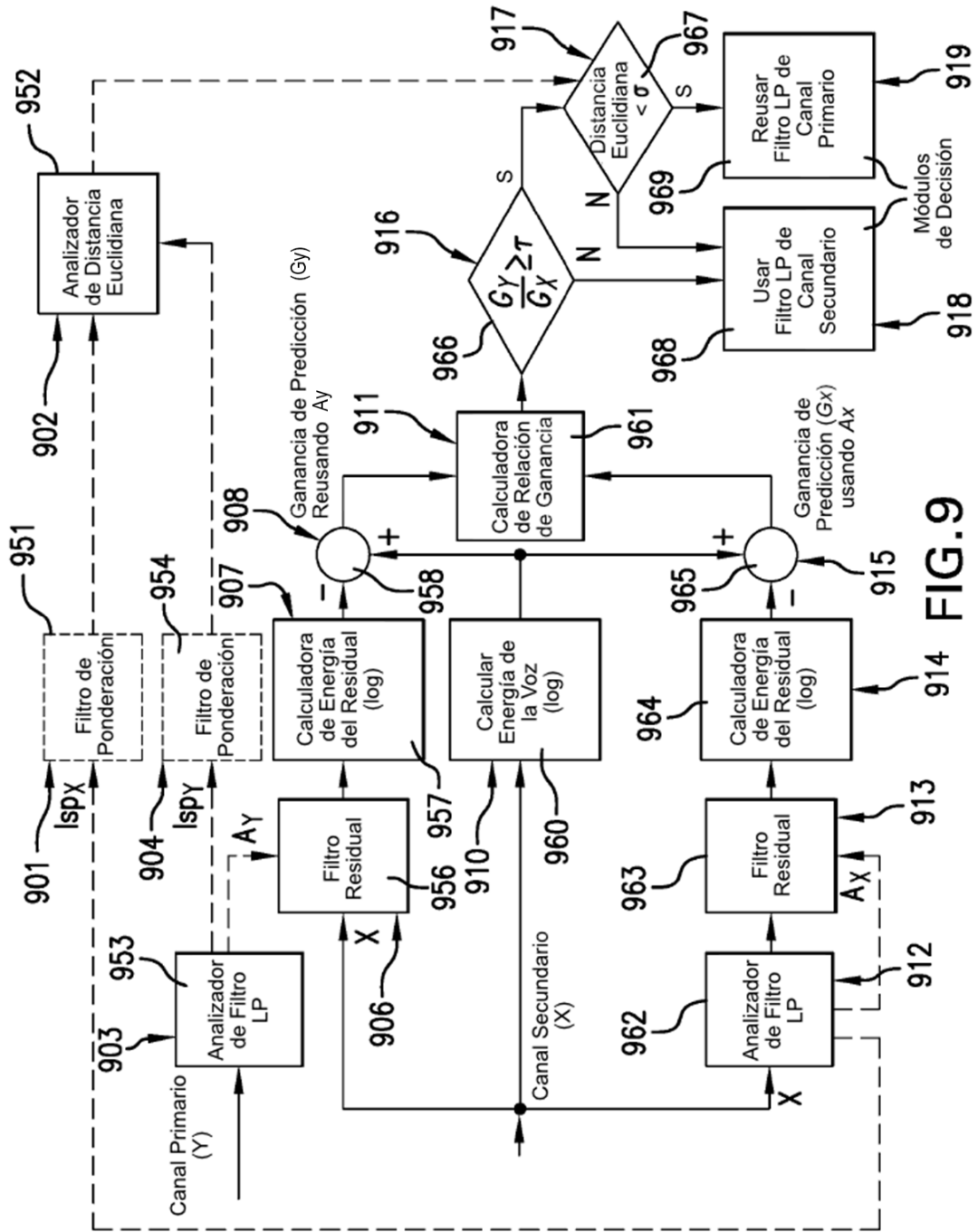


FIG. 9

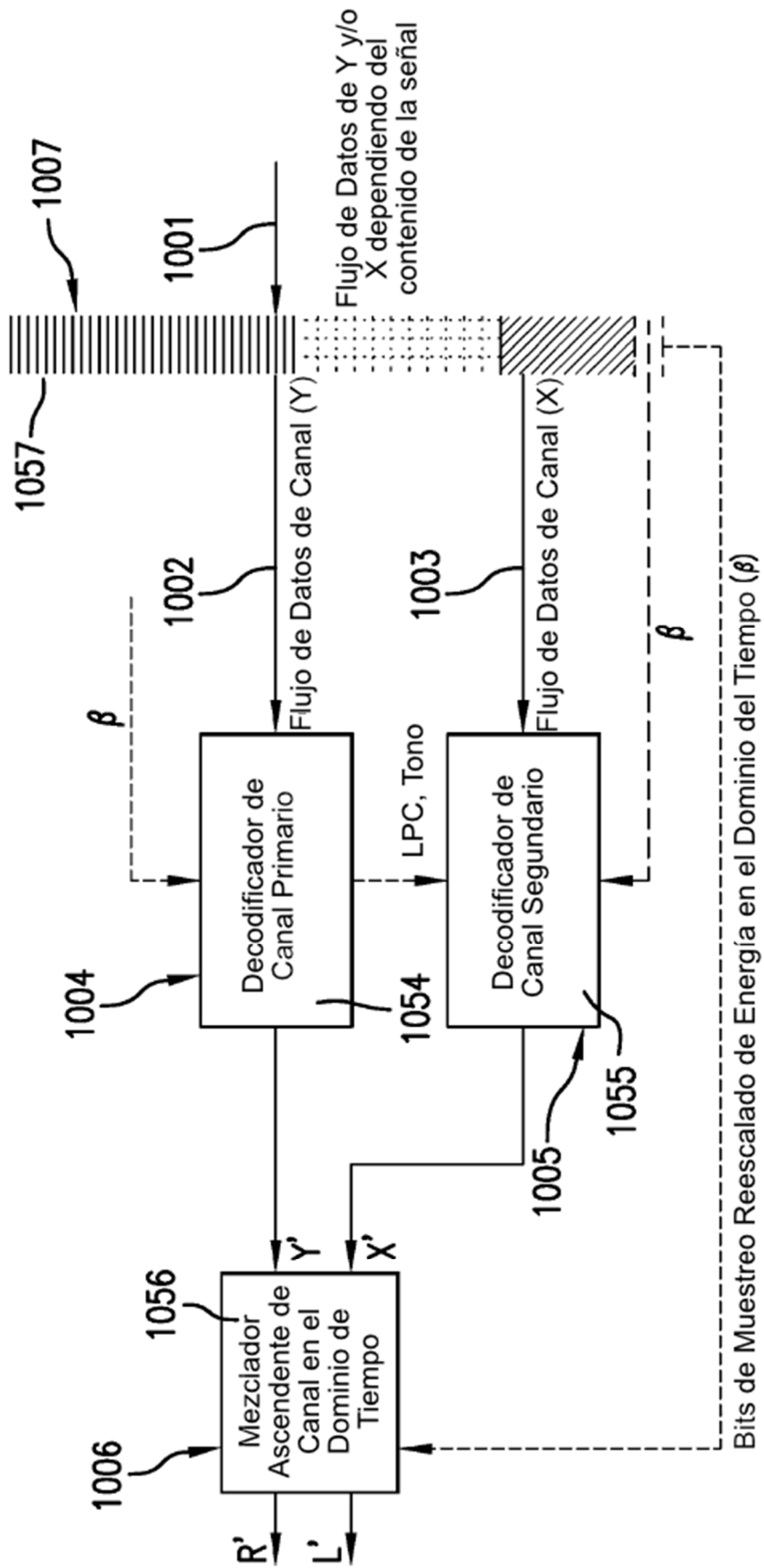


FIG.10

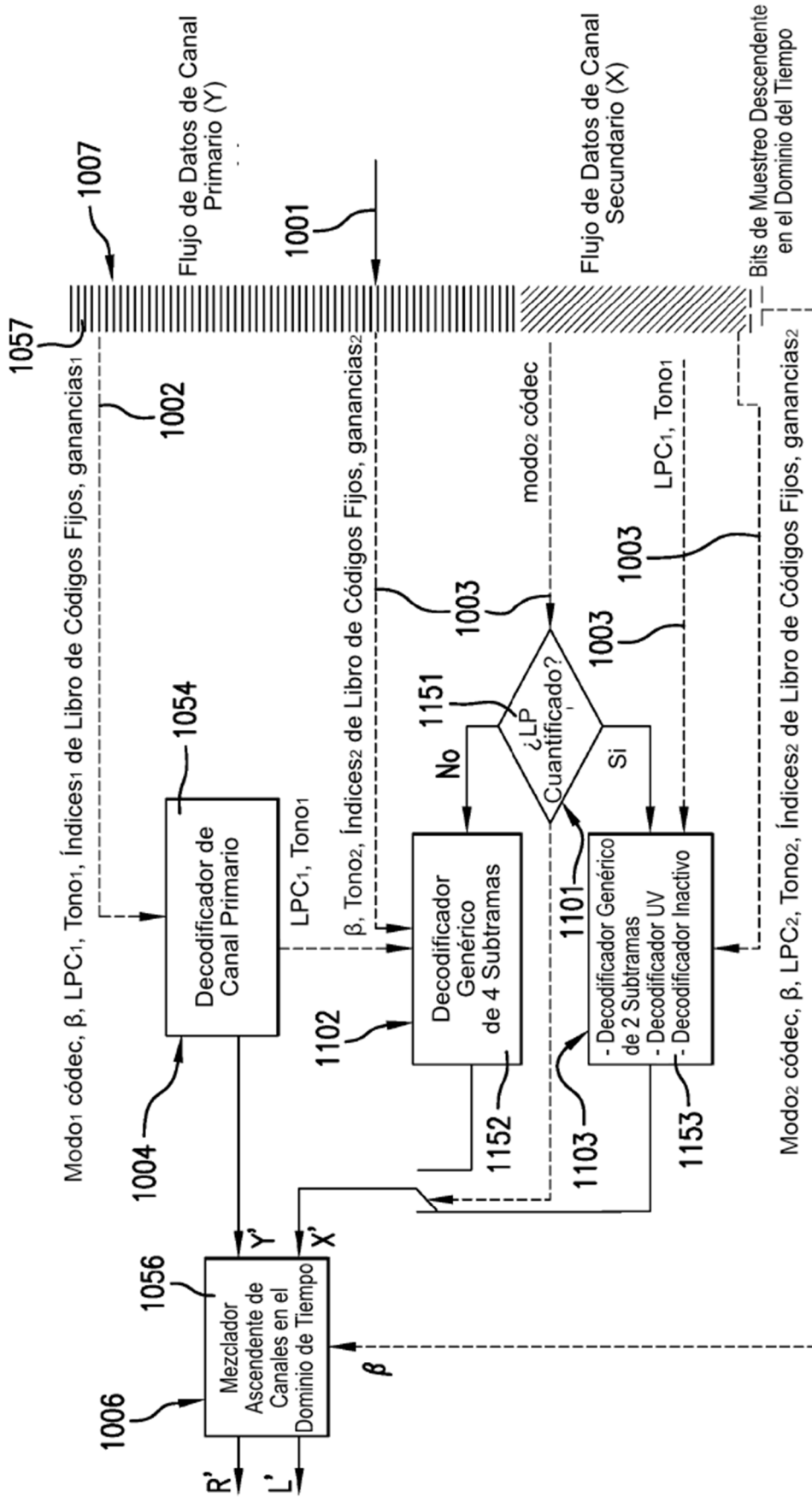


FIG.11

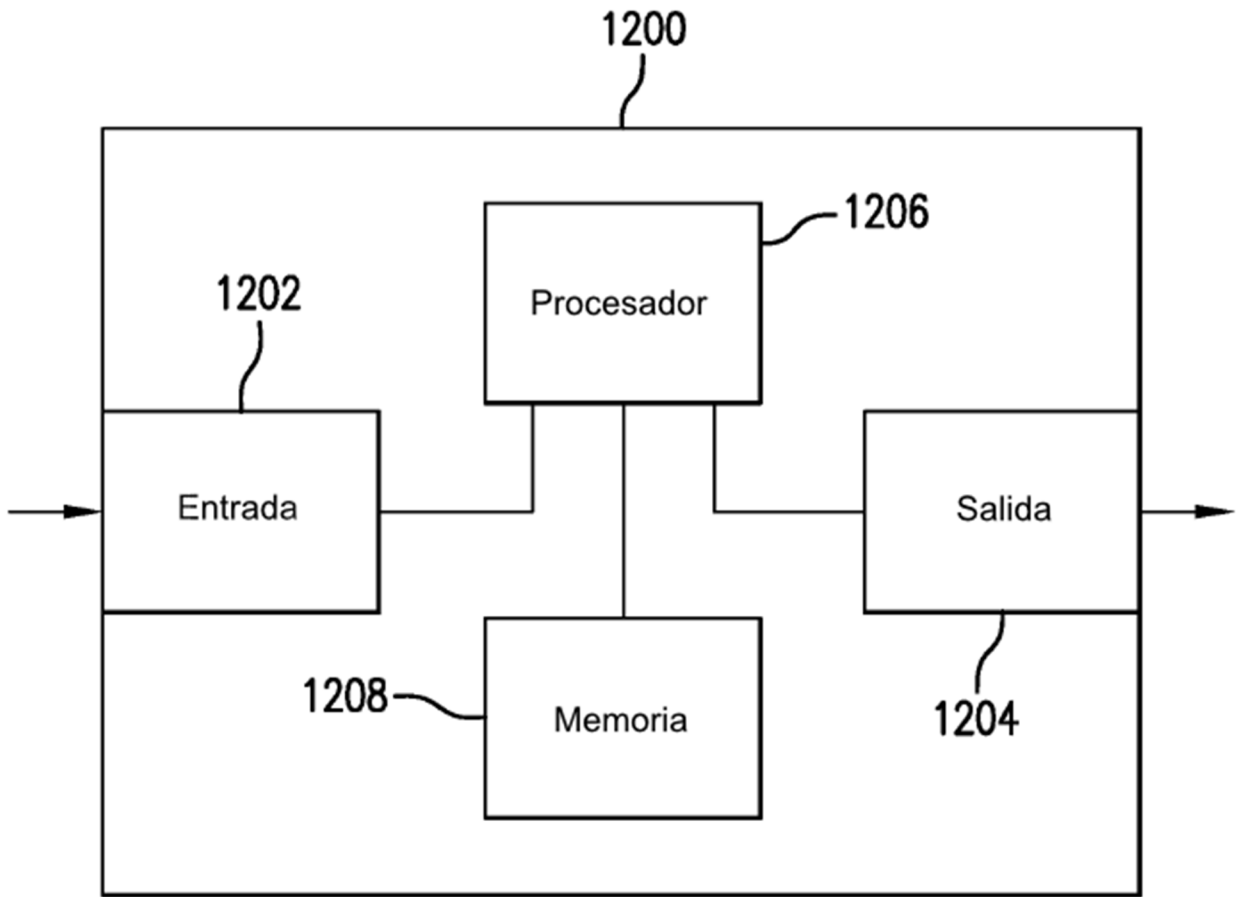


FIG. 12

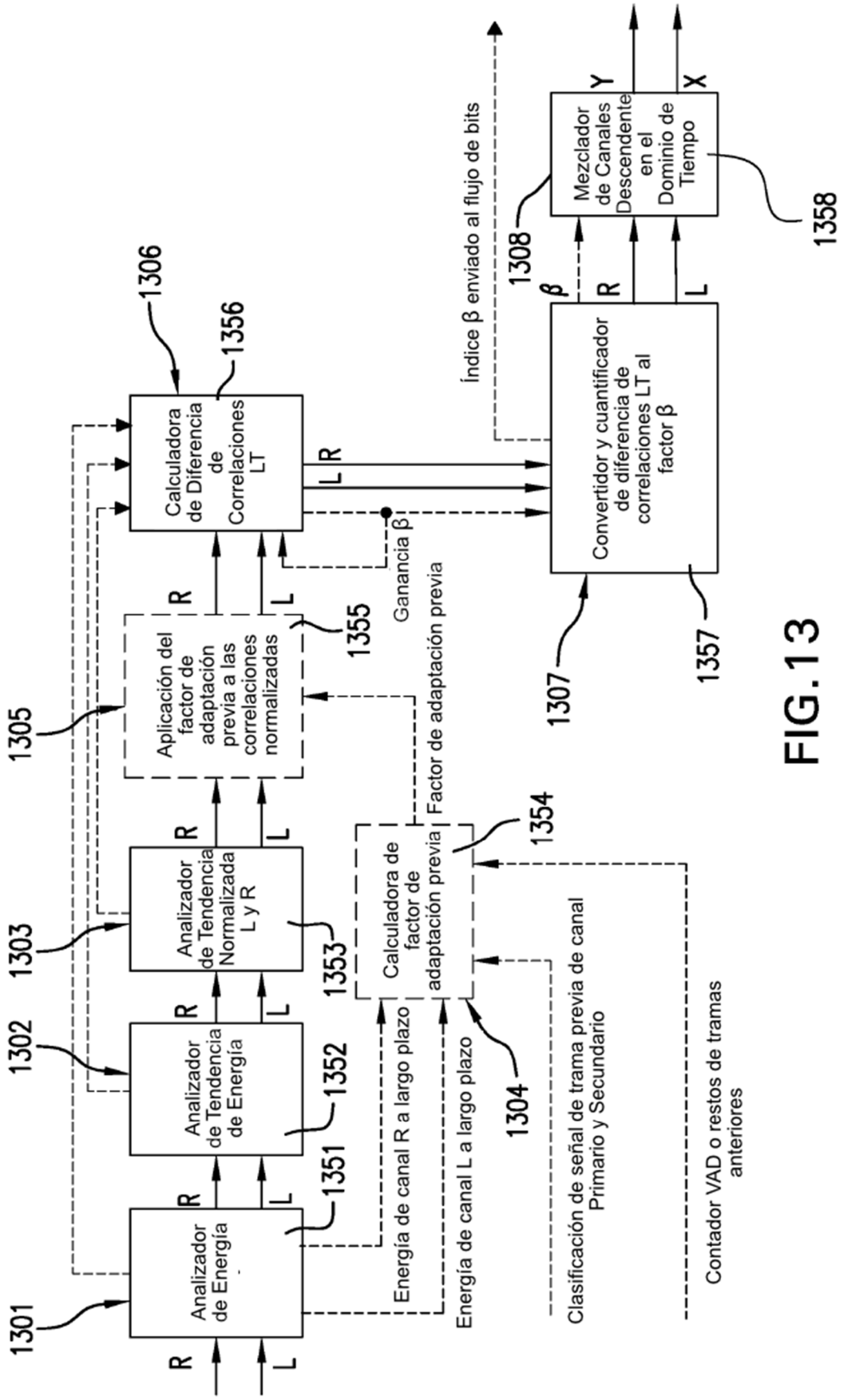


FIG. 13

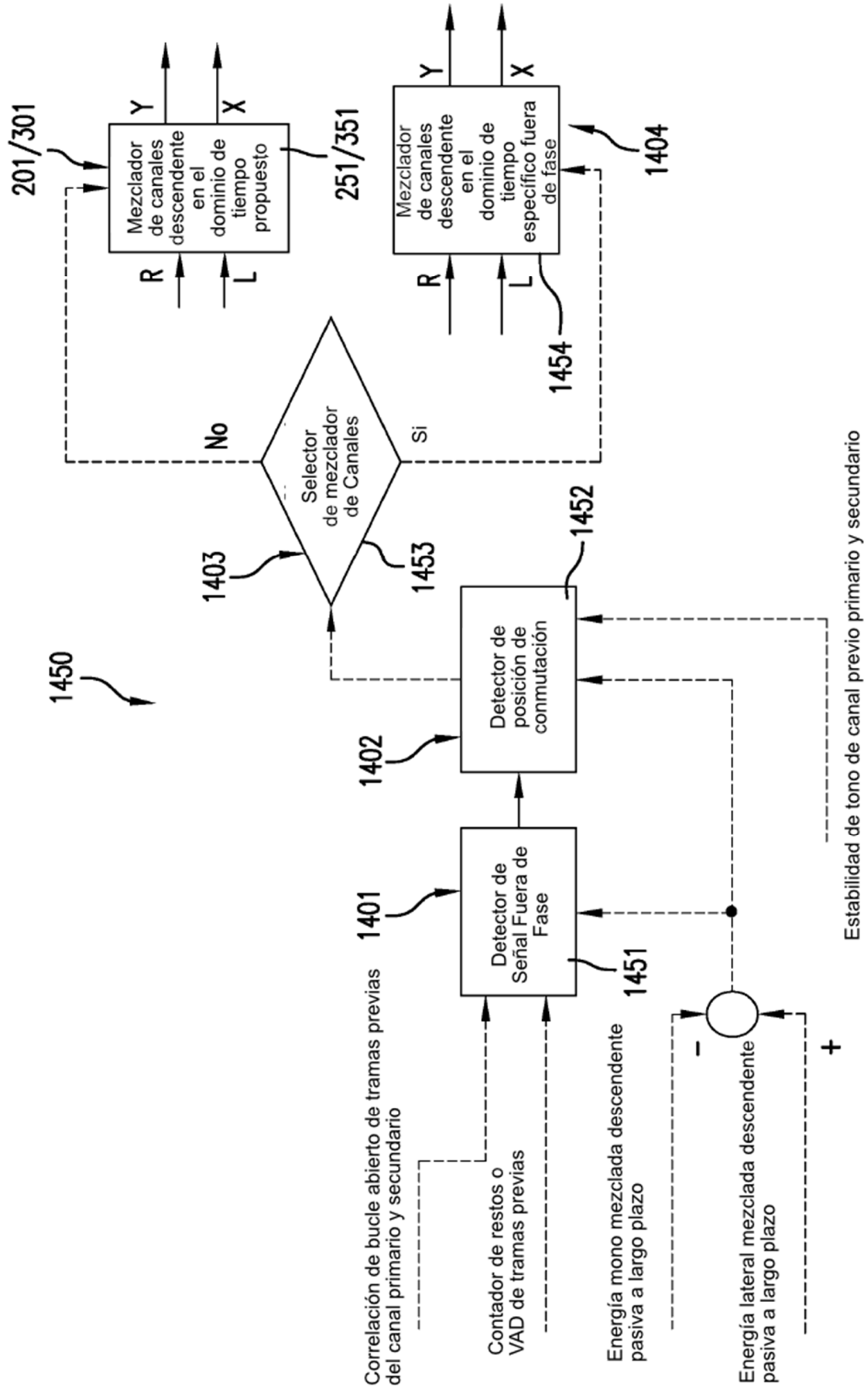


FIG. 14

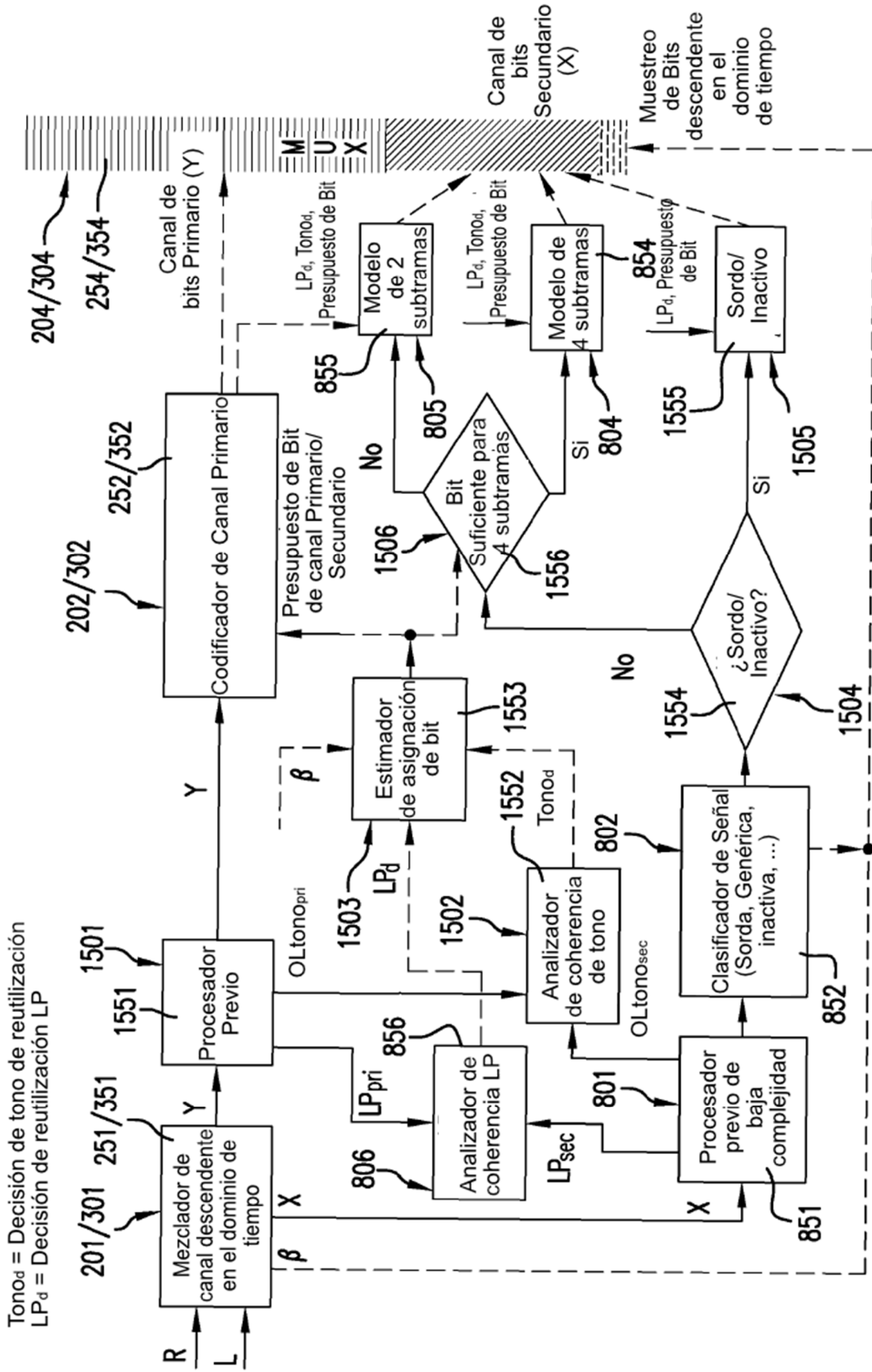


FIG.15

Tono_d = Decisión de tono de reutilización
 LP_d = Decisión de reutilización LP

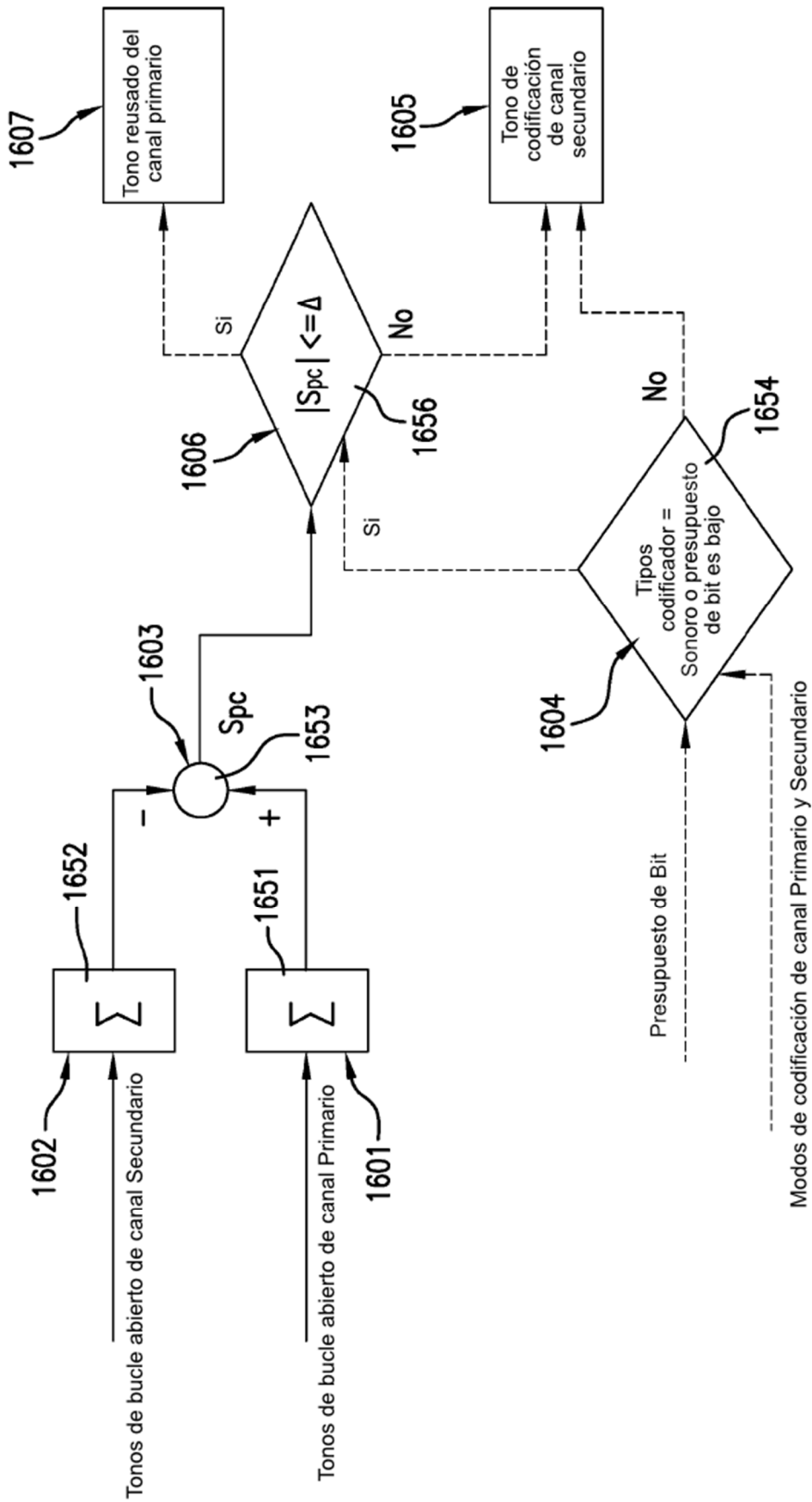


FIG.16

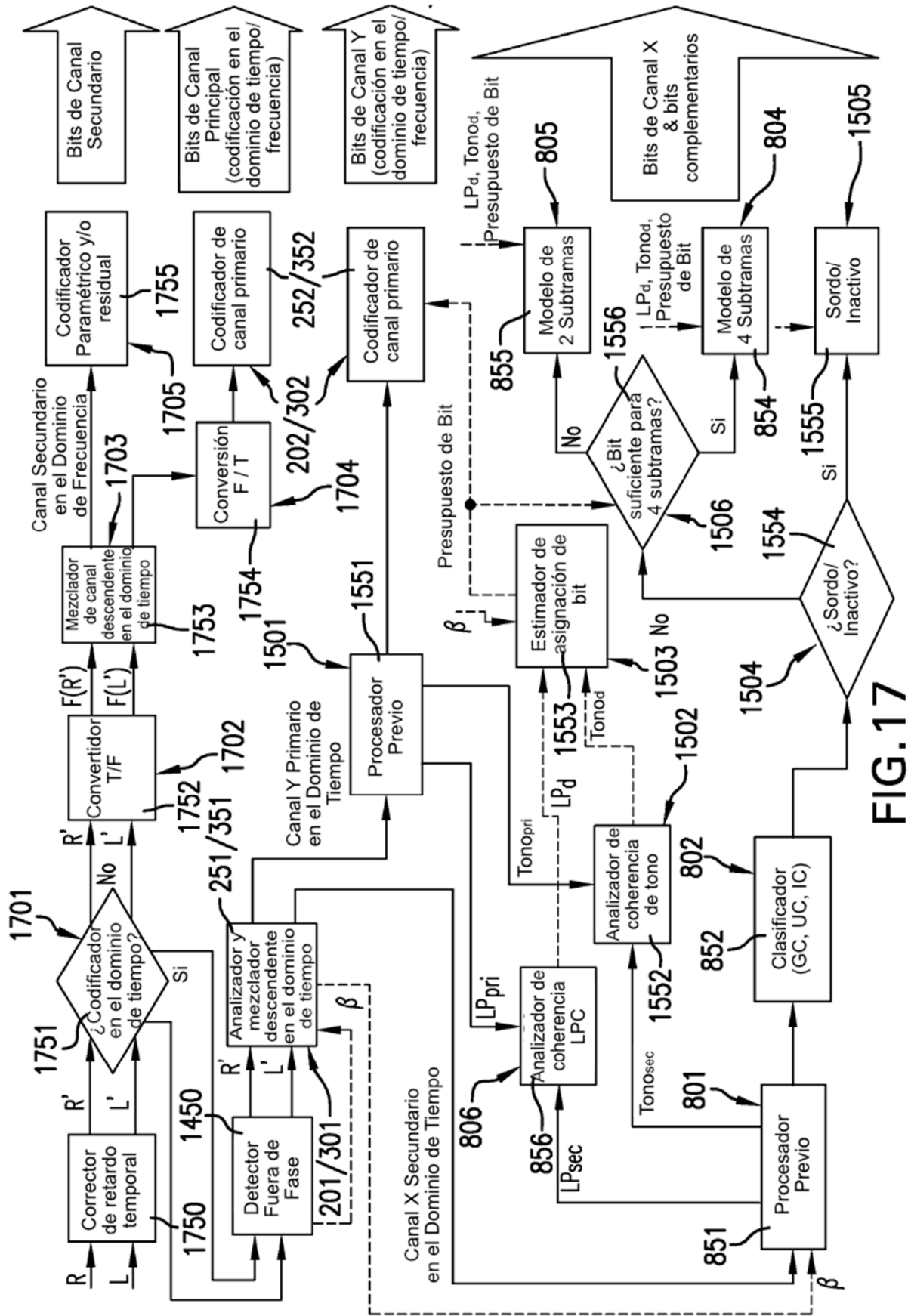


FIG.17

