

19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **3 020 557**

51 Int. Cl.:

G10L 19/008 (2013.01)

G10L 19/22 (2013.01)

G10L 19/24 (2013.01)

G10L 19/02 (2013.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

86 Fecha de presentación y número de la solicitud internacional: **01.02.2021** **PCT/CA2021/050114**

87 Fecha y número de publicación internacional: **12.08.2021** **WO21155460**

96 Fecha de presentación y número de la solicitud europea: **01.02.2021** **E 21751043 (7)**

97 Fecha y número de publicación de la concesión europea: **19.02.2025** **EP 4100948**

54 Título: **Conmutación entre modos de codificación estéreo en un códec de sonido multicanal**

30 Prioridad:

03.02.2020 US 202062969203 P

45 Fecha de publicación y mención en BOPI de la traducción de la patente:
23.05.2025

73 Titular/es:

VOICEAGE CORPORATION (100.00%)
750 Lucerne Road, Suite 250
Town of Mount Royal, Quebec H3R 2H6, CA

72 Inventor/es:

EKSLER, VACLAV

74 Agente/Representante:

SÁEZ MAESO, Ana

ES 3 020 557 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

Cambio entre modos de codificación estéreo en un códec de sonido multicanal

Ámbito técnico

La presente divulgación está relacionada con la codificación de sonido estéreo, en particular, pero no exclusivamente, con el cambio entre «modos de codificación estéreo» (en adelante, también «modos estéreo») en un códec de sonido multicanal capaz, entre otras cosas, de producir una buena calidad estéreo, por ejemplo, en una escena de audio compleja a baja velocidad de bits y con bajo retardo.

5 En la presente divulgación y las reivindicaciones adjuntas:

- El término «sonido» puede estar relacionado con la voz, el audio y cualquier otro sonido;
- El término «estéreo» es una abreviatura de «estereofónico»; y
- El término «mono» es una abreviatura de «monofónico».

Contexto

10 Históricamente, la telefonía conversacional se ha implementado con terminales con un único transductor para emitir sonido solo a uno de los oídos del usuario. En la última década, los usuarios han empezado a utilizar el móvil junto con unos auriculares para recibir el sonido en ambos oídos, principalmente para escuchar música, aunque a veces también para escuchar voz. Sin embargo, cuando se utiliza un teléfono móvil para transmitir y recibir el lenguaje hablado, el contenido sigue siendo mono, pero se presenta a los dos oídos del usuario cuando se utilizan auriculares.

15 Con el último estándar de codificación de la voz 3GPP, tal y como se describe en la Referencia [1], la calidad del sonido codificado, por ejemplo, la voz o el audio que se transmite y recibe a través de un dispositivo móvil, ha mejorado significativamente. El siguiente paso natural es transmitir información estéreo de manera que el receptor se acerque lo máximo posible a una escena de audio real captada en el otro extremo del enlace de comunicación.

20 En los códecs de audio, por ejemplo, tal y como se describe en Referencia [2], se suele utilizar la transmisión de información estéreo.

25 Para los códecs de lenguaje hablado, la señal mono es la habitual. Cuando se transmite una señal estéreo, a menudo es necesario duplicar la velocidad de bits, ya que tanto el canal izquierdo como el derecho de la señal estéreo están codificados mediante un códec mono. Esto funciona bien en la mayoría de los escenarios, pero presenta el inconveniente de duplicar la velocidad de bits y no aprovechar la redundancia potencial entre los dos canales (izquierdo y derecho de la señal estéreo). Además, para mantener la velocidad de bits general a un nivel razonable, se utiliza una velocidad de bits muy baja para cada canal, lo que afecta a la calidad general del sonido. Para reducir la velocidad de bits, se han desarrollado y utilizado técnicas de codificación estéreo eficientes. Como ejemplos no limitativos, en los párrafos siguientes se analiza el uso de tres técnicas de codificación estéreo que pueden utilizarse eficazmente a bajas velocidades de bits.

35 La primera técnica de codificación estéreo se llama estéreo paramétrico. La codificación estéreo paramétrica codifica dos canales, izquierdo y derecho, como una señal mono utilizando un códec mono común, además de cierta cantidad de información del lado estéreo (correspondiente a parámetros estéreo) que representa una imagen estéreo. Los dos canales de entrada, izquierdo y derecho, se mezclan en una señal mono, y los parámetros estéreo se calculan normalmente en el dominio de la transformada, por ejemplo, en el dominio de la Transformada Discreta de Fourier (DFT), y están relacionados con las llamadas pistas binaurales o intercanal. Las pistas binaurales (Referencia [3]) comprenden la diferencia de nivel interaural (ILD), la diferencia de tiempo interaural (ITD) y la correlación interaural (IC). Dependiendo de las características de la señal, la configuración de la escena estéreo, etc., algunas o todas las pistas binaurales se codifican y transmiten al decodificador. La información sobre qué pistas binaurales se codifican y transmiten se envía como información de señalización, que suele formar parte de la información del lado estéreo. Una pista binaural concreta también se puede cuantificar mediante diferentes técnicas de codificación, lo que da lugar a que se utilice un número variable de bits. Además de las pistas binaurales cuantificadas, la información del lado estéreo puede contener, normalmente a velocidades de bits medias y más altas, una señal residual cuantificada resultante de la mezcla de canales. La señal residual se puede codificar mediante una técnica de codificación de entropía, p. ej., un codificador aritmético. La codificación estéreo paramétrica con parámetros estéreo calculados en un dominio de transformación se denominará «estéreo DFT» en la presente divulgación.

Otra técnica de codificación estéreo es la que funciona en el dominio temporal (TD). Esta técnica de codificación estéreo mezcla los dos canales de entrada, izquierdo y derecho, en los llamados canales primario y secundario.

Por ejemplo, siguiendo el método descrito en la Referencia [4], la mezcla en el dominio temporal puede basarse en una relación de mezcla, que determina las contribuciones respectivas de los dos canales de entrada, izquierdo y derecho, tras la producción del canal primario y del canal secundario. La relación de mezcla se obtiene a partir de varias métricas, p. ej., correlaciones normalizadas de los canales de entrada izquierdo y derecho con respecto a una versión de señal mono o una diferencia de correlación a largo plazo entre los dos canales de entrada izquierdo y derecho. El canal primario se puede codificar mediante un códec mono común, mientras que el canal secundario puede hacerlo mediante un códec de velocidad de bits inferior. La codificación del canal secundario puede explotar la coherencia entre los canales primario y secundario y podría reutilizar algunos parámetros del canal primario. En la presente declaración se hará referencia a la codificación estéreo del dominio temporal como codificación «estéreo TD». En general, la codificación estéreo TD es más eficiente a velocidades de bits bajas y medias para codificar señales de voz.

Una tercera técnica de codificación estéreo es la que opera en la Transformada Discreta del Coseno Modificada (MDCT). Se basa en la codificación conjunta de los canales izquierdo y derecho mientras se calcula el procesamiento global de ILD y Mid/Side (M/S) en el dominio espectral blanqueado. Esta tercera técnica de codificación estéreo utiliza varias herramientas adaptadas de la codificación TCX (*Transform Coded eXcitation*) en códecos MPEG (*Moving Picture Experts Group*), tal y como se describe, por ejemplo, en las Referencias [6] y [7]. Estas herramientas pueden incluir codificación del núcleo TCX, análisis TCX LTP (*Long-Term Prediction*), relleno de ruido TCX, modelado de ruido en el dominio de la frecuencia (FDNS), relleno inteligente de huecos estéreo (IGF) o asignación de bits adaptativa entre canales. En general, esta tercera técnica de codificación estéreo es eficiente para codificar todo tipo de contenido de audio a velocidades de bits medias y altas. En la presente declaración, la técnica de codificación estéreo del dominio MDCT se denominará «codificación estéreo MDCT». En general, la codificación estéreo MDCT es más eficiente a velocidades de bits medias y altas para codificar señales de audio generales.

En los últimos años, la codificación estéreo se ha ampliado aún más a la codificación multicanal. Existen varias técnicas para proporcionar codificación multicanal, pero el núcleo fundamental de todas ellas se suele basar en una o varias instancias de técnicas de codificación mono o estéreo. Así, la presente divulgación presenta el cambio entre modos de codificación estéreo que pueden ser parte de técnicas de codificación multicanal tales como audio espacial asistido por metadatos (MASA), como se describe, por ejemplo, en la referencia [8].

En el enfoque MASA, los metadatos de MASA (por ejemplo, dirección, relación de energía, coherencia de dispersión, distancia, coherencia envolvente, todo en varias franjas de frecuencia temporal) se generan en un analizador de MASA, cuantificados, codificados y transmitidos al flujo de bits, mientras que los canales de audio MASA se tratan como señales de transporte (multi)mono o (multi)estéreo codificadas por los codificadores del núcleo. En el decodificador de MASA, los metadatos de MASA guían el proceso de decodificación y representación para recrear un sonido espacial de salida.

US 2017/0365263 A1 describe un codificador de audio multicanal configurado para cambiar entre un codificador de dominio de predicción lineal y un codificador de dominio de frecuencia, así como un decodificador correspondiente.

Resumen

La presente divulgación proporciona dispositivos y métodos de codificación y decodificación de señales de sonido estéreo, tal y como se definen en las reivindicaciones adjuntas.

Todo lo anterior y otros objetos, ventajas y características de los dispositivos y métodos de codificación y decodificación estéreo se harán más evidentes al leer la siguiente descripción no restrictiva de modalidades ilustrativas de los mismos, que se ofrece a modo de ejemplo solo con referencia a los esquemas adjuntos.

Descripción breve de los esquemas

En los esquemas adjuntos:

La Figura 1 es un diagrama esquemático de bloques de un sistema de procesamiento y comunicación de sonido que representa un posible contexto de implementación de los dispositivos y métodos de codificación y decodificación estéreo;

La Figura 2 es un diagrama de bloques de alto nivel que ilustra simultáneamente un dispositivo de codificación estéreo de servicios de voz y audio inmersivos (IVAS) y el método de codificación estéreo correspondiente, en el que el dispositivo de codificación estéreo IVAS incluye un codificador estéreo de dominio de frecuencia (FD), un codificador estéreo de dominio temporal (TD) y un codificador estéreo de la Transformada Discreta del Coseno Modificada (MDCT), donde la implementación del codificador estéreo FD se basa en la Transformada Discreta de Fourier (DFT) (en adelante, «codificador estéreo DFT») en esta modalidad ilustrativa y los esquemas adjuntos;

La Figura 3 es un diagrama de bloques que ilustra simultáneamente el codificador estéreo DFT de la Figura 2 y el método de codificación estéreo DFT correspondiente;

La Figura 4 es un diagrama de bloques que ilustra simultáneamente el codificador estéreo TD de la Figura 2 y el método de codificación estéreo TD correspondiente;

5 La Figura 5 es un diagrama de bloques que ilustra simultáneamente el codificador estéreo MDCT de la Figura 2 y el método de codificación estéreo MDCT correspondiente;

La Figura 6 es un diagrama de flujo que ilustra las operaciones de procesamiento en el dispositivo de codificación estéreo IVAS y el método al cambiar de un modo estéreo TD a un modo estéreo DFT;

10 La Figura 7a es un diagrama de flujo que ilustra las operaciones de procesamiento en el dispositivo de codificación estéreo IVAS y el método al cambiar del modo estéreo DFT al modo estéreo TD;

La Figura 7b es un diagrama de flujo que ilustra las operaciones de procesamiento relacionadas con las señales estéreo TD anteriores al cambiar del modo estéreo DFT al modo estéreo TD;

15 La Figura 8 es un diagrama de bloques de alto nivel que ilustra simultáneamente un dispositivo de decodificación estéreo IVAS y el método de decodificación correspondiente, en el que el dispositivo de decodificación estéreo IVAS comprende un decodificador estéreo DFT, un decodificador estéreo TD y un decodificador estéreo MDCT;

La Figura 9 es un diagrama de flujo que ilustra las operaciones de procesamiento en el dispositivo de decodificación estéreo IVAS y el método al cambiar del modo estéreo TD al modo estéreo DFT;

20 La Figura 10 es un diagrama de flujo que ilustra un caso B) de la Figura 9, que comprende la actualización de memorias de síntesis estéreo DFT en una trama estéreo TD en el lado del decodificador;

La Figura 11 es un diagrama de flujo que ilustra un caso C) de la Figura 9, que comprende el suavizado de una síntesis estéreo de salida en la primera trama estéreo DFT después de cambiar del modo estéreo TD al modo estéreo DFT, en el lado del decodificador;

25 La Figura 12 es un diagrama de flujo que ilustra las operaciones de procesamiento en el dispositivo de decodificación estéreo IVAS y el método al cambiar del modo estéreo DFT al modo estéreo TD;

La Figura 13 es un diagrama de flujo que ilustra un caso A) de la Figura 12, que comprende la actualización de una memoria de sincronización estéreo TD en una primera trama estéreo TD después de cambiar del modo estéreo DFT al modo estéreo TD, en el lado del decodificador; y

30 La Figura 14 es un diagrama de bloques simplificado de un ejemplo de configuración de componentes de hardware que implementan cada uno de los dispositivos y métodos de codificación estéreo IVAS y los dispositivos y métodos de decodificación estéreo IVAS.

Descripción detallada

35 Como se mencionó anteriormente, la presente divulgación se relaciona con la codificación de sonido estéreo, en particular, pero no exclusivamente, con el cambio entre modos de codificación estéreo en un sonido, que incluye voz o audio, códec capaz, en particular, pero no exclusivamente, de producir una buena calidad estéreo, por ejemplo, en una escena de audio compleja a baja velocidad de bits y con bajo retardo. En la presente divulgación, una escena de audio compleja incluye situaciones, por ejemplo, pero no exclusivamente, en las que (a) la correlación entre las señales de sonido que se graban por los micrófonos es baja, (b) hay una fluctuación importante del ruido de fondo o (c) hay un hablante que interfiere. Algunos ejemplos no limitativos de escenas de audio complejas incluyen una sala de conferencias anecoica grande con una configuración de micrófonos A/B, una sala ecoica pequeña con micrófonos binaurales y una sala ecoica pequeña con una configuración de micrófonos mono/laterales. Todas estas configuraciones de sala podrían incluir ruido de fondo fluctuante o hablantes que interfieran.

45 La Figura 1 es un diagrama esquemático de bloques de un sistema de procesamiento y comunicación de sonido estéreo 100 que representa un posible contexto de implementación del dispositivo y método de codificación estéreo IVAS y del dispositivo y método de decodificación estéreo IVAS.

El sistema de procesamiento y comunicación de sonido estéreo 100 de la Figura 1 admite la transmisión de una señal de sonido estéreo a través de un enlace de comunicación 101. El enlace de comunicación 101 puede incluir, por ejemplo, un cable o un enlace de fibra óptica. Por otra parte, el enlace de comunicación 101 puede comprender, al menos en parte, un enlace de radiofrecuencia. El enlace de radiofrecuencia suele admitir

múltiples comunicaciones simultáneas que requieren recursos de ancho de banda compartidos, como los que se pueden encontrar con la telefonía móvil. Aunque no se muestra, el enlace de comunicación 101 puede sustituirse por un dispositivo de almacenamiento en una implementación de dispositivo único del sistema 100 que grabe y almacene la señal de sonido estéreo codificada para su reproducción posterior.

- 5 Siguiendo con 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. Como se indica en la descripción anterior, la señal de sonido puede incluir, en particular, pero no exclusivamente, voz o audio.

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

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

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

Los canales sintetizados izquierdo 114 y derecho 134 de la señal de sonido estéreo analógica se reproducen
30 respectivamente en un par de unidades de altavoces o auriculares binaurales, 116 y 136. Como alternativa, los canales izquierdo 113 y derecho 133 de la señal de sonido estéreo digital del decodificador de sonido estéreo 110 también pueden suministrarse y grabarse en un dispositivo de almacenamiento (no se muestra).

Por ejemplo, (a) el canal izquierdo de la Figura 1 puede implementarse mediante el canal izquierdo de las Figuras 2-13, (b) el canal derecho de la Figura 1 puede implementarse mediante el canal derecho de las Figuras 2-13, (c) el codificador de sonido estéreo 106 de la Figura 1 puede implementarse mediante el dispositivo de
35 codificación estéreo IVAS de las Figuras 2-7 y (d) el decodificador de sonido estéreo 110 de la Figura 1 puede implementarse mediante el dispositivo de decodificación estéreo IVAS de las Figuras 8-13.

1. Cambio entre los modos estéreo en el codificador estéreo IVAS 200 y el método 250

La Figura 2 es un diagrama de bloques de alto nivel que ilustra simultáneamente el codificador estéreo IVAS 200 y el método de codificación estéreo IVAS 250 correspondiente; la Figura 3 es un diagrama de bloques que
40 ilustra simultáneamente el codificador estéreo FD 300 del dispositivo de codificación estéreo IVAS 200 de la Figura 2 y el método de codificación estéreo FD 350 correspondiente; la Figura 4 es un diagrama de bloques que ilustra simultáneamente el codificador estéreo TD 400 del dispositivo de codificación estéreo IVAS 200 de la Figura 2 y el método de codificación estéreo TD 450 correspondiente y la Figura 5 es un diagrama de bloques que ilustra simultáneamente el codificador estéreo MDCT 500 del dispositivo de codificación estéreo IVAS 200
45 de la Figura 2 y el método de codificación estéreo MDCT 550 correspondiente.

En la implementación ilustrativa y no limitativa de las Figuras 2-5, el marco del dispositivo codificador estéreo IVAS 200 (y correspondientemente el dispositivo decodificador estéreo IVAS 800 de la Figura 8) se basa en una versión modificada del códec de servicios de voz mejorados (EVS) (véase la Referencia [1]). En concreto,
50 el códec EVS se amplía para codificar (y decodificar) estéreo y multicanal y para abordar los servicios de voz y audio inmersivos (IVAS). Por esa razón, el dispositivo de codificación 200 y el método 250 se denominan dispositivo y método de codificación estéreo IVAS en la presente divulgación. En la implementación ilustrativa descrita, el dispositivo codificador estéreo IVAS 200 y el método 250 usan, como ejemplo no limitativo, tres modos de codificación estéreo: un modo estéreo de dominio de frecuencia (FD) basado en DFT (Transformada Discreta de Fourier), denominado en esta divulgación como «modo estéreo DFT»; un modo estéreo de dominio temporal (TD), denominado en esta divulgación como «modo estéreo TD», y un modo de codificación estéreo
55 conjunto basado en el modo estéreo de la Transformada Discreta del Coseno Modificada (MDCT), denominado en la presente divulgación como «modo estéreo MDCT». Debe tenerse en cuenta que se pueden usar otras

estructuras de códec como base para el marco del dispositivo codificador estéreo IVAS 200 (y correspondientemente el dispositivo decodificador estéreo IVAS 800).

En la implementación no limitativa descrita, el cambio de modo estéreo en el códec IVAS (dispositivo de codificación estéreo IVAS 200 y dispositivo de decodificación estéreo IVAS 800) se refiere al cambio entre los modos estéreo DFT, TD y MDCT.

1.1 Diferencias entre los diferentes codificadores estéreo y métodos de codificación

En la presente divulgación y en las imágenes adjuntas se usa la siguiente nomenclatura: las letras pequeñas indican señales del dominio temporal, las letras mayúsculas indican señales del dominio de transformación; l/L representa el canal izquierdo, r/R representa el canal derecho; m/M significa canal medio, s/S significa canal lateral, PCh significa canal primario y SCh significa canal secundario. Además, en las imágenes, los números sin unidad corresponden a un número de muestras a una frecuencia de muestreo de 16 kHz.

Existen diferencias entre (a) el codificador estéreo DFT 300 y el método de codificación 350, (b) el codificador estéreo TD 400 y el método de codificación 450 y (c) el codificador estéreo MDCT 500 y el método de codificación 550. Varias de estas diferencias se resumen en los siguientes párrafos y al menos algunas de ellas se explicarán mejor en la siguiente descripción.

El dispositivo codificador estéreo IVAS 200 y el método de codificación 250 realizan operaciones tales como almacenar en memoria intermedia una trama de 20 ms (como es bien sabido en la técnica, la señal de sonido estéreo se procesa en tramas sucesivas de una duración dada que contienen un número determinado de muestras de señal de sonido) de señal de entrada estéreo (canales izquierdo y derecho), algunos pasos de clasificación, mezcla de canales, preprocesamiento y codificación real. Hay disponible una anticipación de 8,75 ms que se usa principalmente para operaciones de análisis, clasificación y OverLap-Add (OLA) utilizadas en el dominio de transformación, como en un núcleo de Excitación Codificada por Transformación (TCX), un núcleo de alta calidad (HQ) y una extensión de ancho de banda del dominio de frecuencia (FD-BWE). Estas operaciones se describen en la Referencia [1], Cláusulas 5,3 y 5.2.6.2.

La anticipación es más corta en el dispositivo codificador estéreo IVAS 200 y el método de codificación 250 en comparación con el codificador EVS no modificado en 0,9375 ms (correspondiente a un retraso de remuestreo del filtro de respuesta de impulso finito (FIR) (véase la Referencia [1], Cláusula 5.1.3.1). Esto afecta al procedimiento de remuestreo de la señal procesada (señal mezclada para los modos estéreo TD y DFT) en cada trama:

- Codificador estéreo DFT 300 y método de codificación 350: El remuestreo se realiza en el dominio DFT y, por lo tanto, no introduce ningún retraso adicional;

- Codificador estéreo TD 400 y método de codificación 450: El remuestreo FIR (decimación) se realiza utilizando el retardo de 0,9375 ms. Dado que este retardo de remuestreo no está disponible en el dispositivo codificador estéreo IVAS 200, el retardo de remuestreo se compensa al agregar ceros al final de la señal mezclada. Por consiguiente, la parte compensada de 0,9375 ms de longitud de la señal mezclada debe recalcularse (volver a muestrearse) en la siguiente trama.

- Codificador estéreo MDCT 500 y método de codificación 550: igual que en el codificador estéreo TD 400 y método de codificación 450.

El remuestreo en el codificador estéreo DFT 300, el codificador estéreo TD 400 y el codificador estéreo MDCT 500 se realiza a partir de la frecuencia de muestreo de entrada (normalmente 16, 32, o 48 kHz) a la(s) velocidad(es) de muestreo interna(s) (normalmente 12,8, 16, 25,6 o 32 kHz). A continuación, la(s) señal(es) remuestreada(s) se utiliza(n) en el preprocesamiento y la codificación del núcleo.

Además, la anticipación contiene una parte de la señal procesada (señal mezclada para los modos estéreo TD y DFT) que no es precisa, sino más bien extrapolada o estimada, lo que también repercute en el proceso de remuestreo. La imprecisión de la señal procesada de la anticipación (señal mezclada para los modos estéreo TD y DFT) depende del modo de codificación estéreo actual:

- Codificador estéreo DFT 300 y método de codificación 350: La longitud de 8,75 ms de la anticipación corresponde a una parte de solapamiento en ventana de la señal mezclada relacionada con una parte de OLA de la ventana de análisis DFT o con una parte de OLA de la ventana de síntesis DFT. Para realizar el preprocesamiento en una señal lo más significativa posible, esta parte anticipada de la señal mezclada se rectifica (o se desvincula, es decir, la ventana inversa se aplica a la parte anticipada). Como consecuencia, la señal mezclada rectificada de 8,75 ms de longitud en la anticipación no se reconstruye con precisión en la trama actual;

- Codificador estéreo TD 400 y método de codificación 450: Antes de la mezcla del dominio temporal (TD), se realiza una alineación intercanal (ICA) utilizando una sincronización del retardo de tiempo intercanal (ITD) entre los dos canales de entrada l y r en el dominio temporal. Esto se logra retrasando uno de los canales de entrada (l o r) y extrapolando una parte faltante de la señal mezclada correspondiente a la duración del retardo de ITD; el valor máximo del retardo de ITD es de 7,5 ms. Por consiguiente, la señal mezclada extrapolada de hasta 7,5 ms de duración en la anticipación no se reconstruye con precisión en la trama actual.

- Codificador estéreo MDCT 500 y método de codificación 550: Por lo general, no se realiza ninguna mezcla ni cambio de tiempo, por lo que la parte anticipada de la señal de audio de entrada suele ser precisa.

La parte de señal rectificada/extrapolada en la anticipación no está sujeta a una codificación real, sino que se utiliza para el análisis y la clasificación. En consecuencia, la parte de señal rectificada/extrapolada en la anticipación se vuelve a calcular en la siguiente trama, y la señal procesada resultante (señal mezclada para los modos estéreo TD y DFT) se utiliza para la codificación real. La longitud de la señal recalculada depende del modo estéreo y del procesamiento de la codificación:

- Codificador estéreo DFT 300 y método de codificación 350: La señal de 8,75 ms de longitud se vuelve a calcular, tanto a la frecuencia de muestreo de la señal estéreo de entrada como a la frecuencia de muestreo interna;

- Codificador estéreo TD 400 y método de codificación 450: La señal de 7,5 ms de longitud se vuelve a calcular a la frecuencia de muestreo de la señal estéreo de entrada, mientras que la señal de $7,5 + 0,9375 = 8,4375$ ms de longitud se vuelve a calcular a la frecuencia de muestreo interna.

- Codificador estéreo MDCT 500 y método de codificación 550: El recálculo no suele ser necesario a la frecuencia de muestreo de la señal estéreo de entrada, mientras que la señal de 0,9375 ms de longitud se vuelve a calcular a la frecuencia de muestreo interna.

Cabe señalar que las longitudes de la parte de señal rectificada y respectivamente extrapolada en la anticipación se mencionan aquí a título ilustrativo, mientras que cualquier otra longitud puede implementarse en general.

Hay más información sobre el codificador estéreo DFT 300 y el método de codificación 350 en las Referencias [2] y [3]. Hay más información sobre el codificador estéreo TD 400 y el método de codificación 450 en la Referencia [4]. También hay más información sobre el codificador estéreo MDCT 500 y el método de codificación 550 en las Referencias [6] y [7].

1.2 Estructura del dispositivo de codificación estéreo IVAS 200 y procesamiento en el método de codificación estéreo IVAS 250

En la siguiente Tabla I se enumeran en orden secuencial las operaciones de procesamiento para cada trama dependiendo del modo de codificación estéreo actual (véanse también las Figuras 2-5).

Tabla I: operaciones de procesamiento en el dispositivo codificador estéreo IVAS 200.

Modo estéreo DFT	Modo estéreo TD	Modo estéreo MDCT
Clasificación estéreo y selección del modo estéreo		
Asignación/designación de memoria		
	Establecer el modo estéreo TD	
Actualizaciones del cambio de modo estéreo		
	Codificador ICA: alineación y escalado del tiempo	
Detectores de transitorios TD		
Configuración del codificador estéreo		

Análisis DFT	Análisis TD	
Procesamiento estéreo y mezcla en el dominio DFT	Mezcla ponderada en el dominio TD	
Síntesis de DFT		
Preprocesamiento frontal		
Configuración del codificador del núcleo		
	Configuración estéreo TD	
Codificación residual estéreo DFT		
Preprocesamiento posterior		
Codificación del núcleo		Codificación estéreo conjunta
Actualizaciones comunes de estéreo		

El método de codificación estéreo IVAS 250 comprende una operación (no se muestra) para controlar el cambio entre los modos estéreo DFT, TD y MDCT. Para realizar la operación de control del cambio, el dispositivo codificador estéreo IVAS 200 comprende un controlador (no se muestra) de cambio entre los modos estéreo DFT, TD y MDCT. El cambio entre los modos estéreo DFT y TD en el dispositivo codificador estéreo IVAS 200 y el método de codificación 250 implica el uso del controlador del cambio de modo estéreo (no se muestra) para mantener la continuidad de las siguientes señales de entrada 1) a 5) y permitir el procesamiento adecuado de las mismas en el dispositivo codificador estéreo IVAS 200 y el método 250:

1) La señal estéreo de entrada que incluye los canales izquierdo l/L y derecho r/R, usados, por ejemplo, para la detección de transitorios en el dominio temporal o BWE intercanal (IC-BWE);

2) La señal estéreo procesada (señal mezclada para los modos estéreo TD y DFT) a la frecuencia de muestreo de la señal estéreo de entrada:

- Codificador estéreo DFT 300 y método de codificación 350: canal medio m/M;

- Codificador estéreo TD 400 y método de codificación 450: Canal primario (PCh) y canal secundario (SCh);

- Codificador estéreo MDCT 500 y método de codificación 550: canales izquierdo y derecho originales (sin mezclar) l y r;

3) Señal procesada (señal mezclada para los modos estéreo TD y DFT) a una frecuencia de muestreo de 12,8 kHz, utilizada en el preprocesamiento;

4) Señal procesada (señal mezclada para los modos estéreo TD y DFT) a una frecuencia de muestreo interna, utilizada en la codificación del núcleo;

5) Señal de entrada de banda alta (HB); utilizada en la extensión de ancho de banda (BWE).

Aunque es sencillo mantener la continuidad para la señal 1) anterior, es difícil para las señales 2) - 5) debido a varios aspectos, por ejemplo, a una mezcla diferente, a una longitud diferente de la parte recalculada de la anticipación, al uso de la alineación intercanal (ICA) solo en el modo estéreo TD, etc.

1.2.1 Clasificación estéreo y selección del modo estéreo

La operación (no se muestra) de controlar el cambio entre los modos estéreo DFT, TD y MDCT comprende una operación 255 de clasificación estéreo y selección de modo estéreo, por ejemplo, como se describe en la Referencia 9.

Para realizar la operación 255, el controlador (no se muestra) para cambiar entre los modos estéreo DFT, TD y MDCT comprende un clasificador estéreo y un selector de modo estéreo 205.

5 El cambio entre el modo estéreo TD, el modo estéreo DFT y el modo estéreo MDCT responde a la selección del modo estéreo. La clasificación estéreo (Referencia [9]) se lleva a cabo en respuesta a los canales izquierdo l y derecho r de la señal estéreo de entrada o a la velocidad de bits codificada solicitada. La selección del modo estéreo (Referencia [9]) consiste en elegir uno de los modos estéreo DFT, TD y MDCT basándose en la clasificación estéreo.

10 El clasificador estéreo y el selector de modo estéreo 205 produce la señalización del modo estéreo 270 para identificar el modo de codificación estéreo seleccionado.

1.2.2 Asignación/desasignación de memoria

15 La operación (no se muestra) de control del cambio entre los modos estéreo DFT, TD y MDCT comprende una operación de asignación de memoria (no se muestra). Para realizar la operación de asignación de memoria, el controlador de cambio entre los modos estéreo DFT, TD y MDCT (no se muestra) asigna/desasigna dinámicamente estructuras de datos de memoria estática a/desde los modos estéreo DFT, TD y MDCT dependiendo del modo estéreo actual. Dicha asignación de memoria mantiene el impacto de la memoria estática del dispositivo codificador estéreo IVAS 200 lo más bajo posible al dejar solo aquellas estructuras de datos que se emplean en la trama actual.

20 Por ejemplo, en una primera trama estéreo DFT después de una trama estéreo TD, las estructuras de datos relacionadas con el modo estéreo TD (por ejemplo, manejo de datos estéreo TD, segunda estructura de datos del codificador del núcleo) se liberan (desasignan) y las estructuras de datos relacionadas con el modo estéreo DFT (por ejemplo, estructura de datos estéreo DFT) se asignan e inicializan en su lugar. Cabe señalar que la desasignación de las estructuras de datos no utilizadas se realiza primero, seguida de la asignación de las estructuras de datos recién utilizadas.

25 Este orden de operaciones es importante para no aumentar el impacto de la memoria estática en ningún punto de la codificación.

En la Tabla II se muestra un resumen de las principales estructuras de datos de la memoria estática utilizadas en los distintos modos estéreo.

Tabla II: asignación de estructuras de datos en diferentes modos estéreo.

30 «X» significa asignado -- «XX» significa asignado dos veces --
«-» significa desasignado y «--» significa desasignado dos veces.

Estructuras de datos	Modo estéreo DFT	Modo estéreo TD normal	Modo estéreo LRTD	Modo estéreo MDCT
Estructura principal de IVAS	X	X	X	X
Clasificador estéreo	X	X	X	X
DFT estéreo	X	-	-	-
TD estéreo	-	X	X	-
MDCT estéreo	-	-	-	X
Codificador del núcleo	X	XX	XX	XX
Núcleo ACELP	X	XX	XX	--
Núcleo TCX + IGF	X	X-	X-	XX
TD-BWE	X	X	XX	--

FD-BWE	X	X	XX	--
IC-BWE	X	X	-	-
ICA	X	X	X	-

A continuación, se muestra un ejemplo de implementación del módulo codificador de asignación/desasignación de memoria en el código fuente C.

```
void stereo_memory_enc(
    CPE_ENC_HANDLE hCPE,      /* i : CPE encoder structure */
    const int32_t input_Fs,    /* i : input sampling rate   */
    const int16_t max_bwidth,  /* i : maximum audio bandwidth */
    float *tdm_last_ratio      /* o : TD stereo last ratio   */
)
```

```

{
    Encoder_State *st;

    /*-----*
     * save parameters from structures that will be freed
     *-----*/

    if ( hCPE->last_element_mode == IVAS_CPE_TD )
    {
        *tdm_last_ratio = hCPE->hStereoTD->tdm_last_ratio; /* note: this
        must be set to local variable before data structures are
        allocated/deallocated */
    }

    if ( hCPE->hStereoTCA != NULL && hCPE->last_element_mode ==
    IVAS_CPE_DFT )
    {
        set_s( hCPE->hStereoTCA->prevCorrLagStats, (int16_t) hCPE-
        >hStereoDft->itd[1], 3 );
        hCPE->hStereoTCA->prevRefChanIndx = ( hCPE->hStereoDft->itd[1] >=
        0 ) ? ( L_CH_INDX ) : ( R_CH_INDX );
    }

    /*-----*
     * allocate/deallocate data structures
     *-----*/

    if ( hCPE->element_mode != hCPE->last_element_mode )
    {
        /*-----*
         * switching CPE mode to DFT stereo
         *-----*/

        if ( hCPE->element_mode == IVAS_CPE_DFT )
        {
            /* deallocate data structure of the previous CPE mode */
            if ( hCPE->hStereoTD != NULL )
            {
                count_free( hCPE->hStereoTD );
                hCPE->hStereoTD = NULL;
            }
        }
    }
}

```

```

if ( hCPE->hStereoMdct != NULL )
{
    count_free( hCPE->hStereoMdct );
    hCPE->hStereoMdct = NULL;
}

/* deallocate CoreCoder secondary channel */
deallocate_CoreCoder_enc( hCPE->hCoreCoder[1] );

/* allocate DFT stereo data structure */

stereo_dft_enc_create( &(amp; hCPE->hStereoDft ), input_Fs,
max_bwidth );

/* allocate ICBWE structure */
if ( hCPE->hStereoICBWE == NULL )
{
    hCPE->hStereoICBWE = (STEREO_ICBWE_ENC_HANDLE)
count_malloc( sizeof( STEREO_ICBWE_ENC_DATA ) );

    stereo_icbwe_init_enc( hCPE->hStereoICBWE );
}

/* allocate HQ core in M channel */
st = hCPE->hCoreCoder[0];
if ( st->hHQ_core == NULL )
{
    st->hHQ_core = (HQ_ENC_HANDLE) count_malloc( sizeof(
HQ_ENC_DATA ) );

    HQ_core_enc_init( st->hHQ_core );
}

/*-----*
* switching CPE mode to TD stereo
*-----*/

if ( hCPE->element_mode == IVAS_CPE_TD )
{
    /* deallocate data structure of the previous CPE mode */
    if ( hCPE->hStereoDft != NULL )
    {
        stereo_dft_enc_destroy( amp( hCPE->hStereoDft ) );
        hCPE->hStereoDft = NULL;
    }

    if ( hCPE->hStereoMdct != NULL )
    {
        count_free( hCPE->hStereoMdct );
        hCPE->hStereoMdct = NULL;
    }

    /* deallocated TCX/ISF structures for second channel */
    deallocate_CoreCoder_TCX_enc( hCPE->hCoreCoder[1] );

    /* allocate TD stereo data structure */

```

```

        hCPE->hStereoTD = (STEREO_TD_ENC_DATA_HANDLE) count_malloc(
sizeof( STEREO_TD_ENC_DATA ) );

        stereo_td_init_enc( hCPE->hStereoTD, hCPE->element_brate,
hCPE->last_element_mode );

/* allocate secondary channel */
allocate_CoreCoder_enc( hCPE->hCoreCoder[1] );
}

/*-----*
* allocate DFT/TD stereo structures after MDCT stereo frame
*-----*/

if ( hCPE->last_element_mode == IVAS_CPE_MDCT && ( hCPE->
element_mode == IVAS_CPE_DFT || hCPE->element_mode == IVAS_CPE_TD ) )
{
    /* allocate TCA data structure */
    hCPE->hStereoTCA = (STEREO_TCA_ENC_HANDLE) count_malloc(
sizeof( STEREO_TCA_ENC_DATA ) );

    stereo_tca_init_enc( hCPE->hStereoTCA, input_Fs );

    st = hCPE->hCoreCoder[0];

    /* allocate primary channel substructures */
    allocate_CoreCoder_enc( st );

    /* allocate CLDFB for primary channel */
    if ( st->cldfbAnaEnc == NULL )
    {
        openCldfb( &st->cldfbAnaEnc, CLDFB_ANALYSIS, input_Fs,
CLDFB_PROTOTYPE_1_25MS );
    }

    /* allocate BWEs for primary channel */
    if ( st->hBWE_TD == NULL )
    {
        st->hBWE_TD = (TD_BWE_ENC_HANDLE) count_malloc( sizeof(
TD_BWE_ENC_DATA ) );

        if ( st->cldfbSynTd == NULL )
        {
            openCldfb( &st->cldfbSynTd, CLDFB_SYNTHESIS, 16000,
CLDFB_PROTOTYPE_1_25MS );
        }

        InitSWBencBuffer( st->hBWE_TD );
        ResetSHBbuffer_Enc( st->hBWE_TD );

        st->hBWE_FD = (FD_BWE_ENC_HANDLE) count_malloc( sizeof(
FD_BWE_ENC_DATA ) );

```

```

        fd_bwe_enc_init( st->hBWE_FD );
    }
}

/*-----*
 * switching CPE mode to MDCT stereo
 *-----*/

if ( hCPE->element_mode == IVAS_CPE_MDCT )
{
    int16_t i;

    /* deallocate data structure of the previous CPE mode */
    if ( hCPE->hStereoDft != NULL )
    {
        stereo_dft_enc_destroy( &(amp; hCPE->hStereoDft) );
        hCPE->hStereoDft = NULL;
    }

    if ( hCPE->hStereoTD != NULL )
    {
        count_free( hCPE->hStereoTD );
        hCPE->hStereoTD = NULL;
    }

    if ( hCPE->hStereoTCA != NULL )
    {
        count_free( hCPE->hStereoTCA );
        hCPE->hStereoTCA = NULL;
    }

    if ( hCPE->hStereoICBWE != NULL )
    {
        count_free( hCPE->hStereoICBWE );
        hCPE->hStereoICBWE = NULL;
    }

    for ( i = 0; i < CPE_CHANNELS; i++ )
    {
        st = hCPE->hCoreCoder[i];

        /* deallocate core channel substructures */
        deallocate_CoreCoder_enc( hCPE->hCoreCoder[i] );
    }

    if ( hCPE->last_element_mode == IVAS_CPE_DFT )
    {
        /* allocate secondary channel */
        allocate_CoreCoder_enc( hCPE->hCoreCoder[1] );
    }

    /* allocate TCX/IGF structures for second channel */
    st = hCPE->hCoreCoder[1];

    st->hToxEnc = (TCX_ENC_HANDLE) count_malloc( sizeof(
TCX_ENC_DATA ) );
    st->hToxEnc->spectrum[0] = st->hToxEnc->spectrum_long;

```

```

        st->hTcxEnc->spectrum[1] = st->hTcxEnc->spectrum_long +
N__TCX10__MAX;

        set_f( st->hTcxEnc->old_out, 0, L__FRAME32k );

        set_f( st->hTcxEnc->spectrum_long, 0, N__MAX );

        if ( hCPE->last_element_mode == IVAS__CPE__DFT )
        {
            st->last_core = ACELP__CORE; /* needed to set-up TCX core
in SetTCXModeInfo() */
        }

        st->hTcxCfg = (TCX__CONFIG_HANDLE) count_malloc( sizeof(
TCX__config ) );

        st->hIGFEnc = (IGF__ENC__INSTANCE_HANDLE) count_malloc( sizeof(
IGF__ENC__INSTANCE ) );
        st->igf = getIgfpresent( st->element_mode, st->total_brte,
st->bwidth, st->rf_mode );

        /* allocate and initialize MDCT stereo structure */
        hCPE->hStereoMdct = (STEREO_MDCT_ENC_DATA_HANDLE)
count_malloc( sizeof( STEREO_MDCT_ENC_DATA ) );

        initMdctStereoEncData( hCPE->hStereoMdct, hCPE-
>element_brte, hCPE->hCoreCoder[0]->max_bwidth, SMDCT__MS_DECISION, 0,
NULL );
    }
}

return;
}

```

5 1.2.3 Ajuste del modo estéreo TD

El modo estéreo TD puede constar de dos submodos. Uno es el llamado submodo estéreo TD normal, para el que la relación de mezcla estéreo TD es superior a 0 e inferior a 1. El otro es un denominado submodo estéreo LRTD, para el cual la relación de mezcla estéreo TD es 0 o 1; por lo tanto, LRTD es un caso extremo del modo estéreo TD en el que la mezcla TD no mezcla realmente el contenido de los canales izquierdo l y derecho r del dominio de tiempo r para formar los canales primarios PCh y secundarios SCh, sino que los obtiene directamente de los canales l y r.

Cuando los dos submodos (normal y LRTD) del modo estéreo TD están disponibles, la operación de cambio de modo estéreo (no se muestra) comprende un ajuste del modo estéreo TD (no se muestra). Para realizar el ajuste del modo estéreo TD, que forma parte de la asignación de memoria, el controlador de cambio de modo estéreo (no se muestra) del dispositivo codificador estéreo IVAS 200 asigna/desasigna ciertas estructuras de datos de la memoria estática al cambiar entre el modo estéreo TD normal y el modo estéreo LRTD. Por ejemplo, una estructura de datos IC-BWE se asigna solo en tramas que utilizan el modo estéreo TD normal (véase la tabla II), mientras que varias estructuras de datos (BWE y Complex Low Delay Filter Bank (CLDFB) para el canal secundario SCh) se asignan solo en tramas que utilizan el modo estéreo LRTD (véase la tabla II). A continuación, se muestra un ejemplo de implementación del módulo codificador de asignación/desasignación de memoria en el código fuente C:

```

/* normal TD / LRTD switching */
if ( hCPE->hStereoTD->tdm_LRTD_flag == 0 )
{
    Encoder_State *st;
    st = hCPE->hCoreCoder[1];

    /* deallocate CLDFB ana for secondary channel */
    if ( st->cldfbAnaEnc != NULL )
    {
        deleteCldfb( &st->cldfbAnaEnc );
    }

    /* deallocate BWEs for secondary channel */
    if ( st->hBWE_TD != NULL )
    {
        if ( st->hBWE_TD != NULL )
        {
            count__free( st->hBWE_TD );
            st->hBWE_TD = NULL;
        }

        deleteCldfb( &st->cldfbSynTd );

        if ( st->hBWE_FD != NULL )
        {
            count__free( st->hBWE_FD );
            st->hBWE_FD = NULL;
        }
    }

    /* allocate ICBWE structure */
    if ( hCPE->hStereoICBWE == NULL )
    {

```

```

        ( hCPE->hStereoICBWE = (STEREO_ICBWE_ENC_HANDLE)
count_malloc( sizeof( STEREO_ICBWE_ENC_DATA ) );

        stereo_icBWE_init_enc( hCPE->hStereoICBWE );
    }
}
else /* tdm_LRTD_flag == 1 */
{
    Encoder_State *st;
    st = hCPE->hCoreCoder[1];

    /* deallocate ICBWE structure */
    if ( hCPE->hStereoICBWE != NULL )
    {
        /* copy past input signal to be used in BWE */
        mvr2r( hCPE->hStereoICBWE->dataChan[1], hCPE->
>hCoreCoder[1]->old_input_signal, st->input_Fs / 50 );

        count_free( hCPE->hStereoICBWE );
        hCPE->hStereoICBWE = NULL;
    }

    /* allocate CLDFB ana for secondary channel */
    if ( st->cldfbAnaEnc == NULL )
    {
        openCldfb( &st->cldfbAnaEnc, CLDFB_ANALYSIS, st->
>input_Fs, CLDFB_PROTOTYPE_1_25MS );
    }

    /* allocate BWEs for secondary channel */
    if ( st->hBWE_TD == NULL )
    {
        st->hBWE_TD = (TD_BWE_ENC_HANDLE) count_malloc( sizeof(
TD_BWE_ENC_DATA ) );

        openCldfb( &st->cldfbSynTd, CLDFB_SYNTHESIS, 16000,
CLDFB_PROTOTYPE_1_25MS );

        InitSWEencBuffer( st->hBWE_TD );
        ResetSHBbuffer_Enc( st->hBWE_TD );

        st->hBWE_FD = (FD_BWE_ENC_HANDLE) count_malloc( sizeof(
FD_BWE_ENC_DATA ) );

        fd_bwe_enc_init( st->hBWE_FD );
    }
}
}

```

Por lo general, en la presente divulgación solo se describirá en detalle el modo estéreo TD normal (que, para simplificar, se denominará únicamente modo estéreo TD). El modo estéreo LRTD se menciona como una posible implementación.

1.2.4 Actualizaciones del cambio de modo estéreo

- 5 La operación de control del cambio de modo estéreo (no se muestra) comprende una operación de actualizaciones del cambio de modo estéreo (no se muestra). Para realizar esta operación de actualización del cambio estéreo, el controlador de cambio de modo estéreo (no se muestra) actualiza los parámetros a largo plazo y actualiza o restablece los búferes de memoria anteriores.
- 10 Al cambiar del modo estéreo DFT al modo estéreo TD, el controlador de cambio de modo estéreo (no se muestra) restablece las estructuras de datos de la memoria estática TD estéreo e ICA. Estas estructuras de datos almacenan los parámetros y memorias del análisis estéreo TD y la mezcla ponderada (401 en la Figura 4), respectivamente, del algoritmo ICA (201 en la Figura 2). A continuación, el controlador de cambio de modo estéreo (no se muestra) establece un índice de relación de mezcla de tramas TD estéreo anteriores de acuerdo con el modo estéreo TD normal o el modo estéreo LRTD. Ejemplo ilustrativo y no limitativo:
 - 15 - El índice de relación de mezcla de tramas anteriores se establece en 15, lo que indica que el canal medio mezclado m/M se codifica como el canal primario PCh, donde la relación de mezcla es de 0,5, en el modo estéreo TD normal; o
 - El índice de relación de mezcla de tramas anteriores se establece en 31, lo que indica que el canal izquierdo l está codificado como el canal primario PCh, en el modo estéreo LRTD.
- 20 Al cambiar del modo estéreo TD al modo estéreo DFT, el controlador de cambio de modo estéreo (no se muestra) restablece la estructura de datos estéreo DFT. Esta estructura de datos estéreo DFT almacena parámetros y memorias relacionadas con el módulo de procesamiento y mezcla DFT estéreo (303 en la Figura 3).
- 25 Además, el controlador de cambio de modo estéreo (no se muestra) transfiere algunos parámetros relacionados con el estéreo entre estructuras de datos. Como ejemplo, los parámetros relacionados con el cambio de tiempo y la energía entre los canales l y r, a saber, una ganancia lateral (o parámetro ILD) y el parámetro ITD del modo estéreo DFT, se usan para actualizar una ganancia objetivo y retardos de correlación (parámetros ICA 202) del modo estéreo TD y viceversa. Estas ganancias objetivo y retrasos de correlación se describen con más detalle en la sección 1.2.5 de la presente divulgación.
- 30 Las actualizaciones/restablecimientos relacionados con los codificadores del núcleo (véanse las Figuras 3 y 4) se describen más adelante en la sección 1.4 de la presente divulgación. A continuación, se muestra un ejemplo de implementación del manejo de algunas memorias en el codificador.

```

void stereo_switching_end(
    CPE_ENC_HANDLE hCPE, /* i : CPE encoder structure
*/
    float old_input_signal_pri[], /* i : old input signal of primary
channel */
    const int16_t input_frame /* i : input frame length
*/
)
{
    int16_t i, n, dft_ovl, offset;
    float tmpF;
    Encoder_State **st;

    st = hCPE->hCoreCoder;
    dft_ovl = STEREO_DFT_OVL_MAX * input_frame / L_FRAME48k;

    /* update DFT analysis overlap memory */
    if ( hCPE->element_mode > IVAS_CPE_DFT && hCPE->input_mem[0] != NULL
)
    {
        for ( n = 0; n < CPE_CHANNELS; n++ )
        {
            mvr2r( st[n]->input + input_frame - dft_ovl, hCPE-
>input_mem[n], dft_ovl );
        }
    }

    /* TD/MDCT -> DFT stereo switching */
    if ( hCPE->element_mode == IVAS_CPE_DFT && hCPE->last_element_mode !=
IVAS_CPE_DFT )
    {
        /* window DFT synthesis overlap memory @input_fs, primary channel
*/
        for ( i = 0; i < dft_ovl; i++ )
        {

```

```

        hCPE->hStereoDft->output_mem_dmx[i] =
old_input_signal_pri[input_frame - dft_ovl + i] * hCPE->hStereoDft-
>win[dft_ovl - 1 - i];
    }

    /* reset 48kHz BWE overlap memory */
    set_f( hCPE->hStereoDft->output_mem_dmx_32k, 0,
STEREO_DFT_OVL_32k );

    stereo_dft_enc_reset( hCPE->hStereoDft );

    /* update ITD parameters */
    if ( hCPE->element_mode == IVAS_CPE_DFT && hCPE-
>last_element_mode == IVAS_CPE_PD )
    {
        set_f( hCPE->hStereoDft->itd, hCPE->hStereoTCA-
>prevCorrLagStats[2], STEREO_DFT_ENC_DFT_NE );
    }

    /* Update the side_gain[] parameters */
    if ( hCPE->hStereoTCA != NULL && hCPE->last_element_mode !=
IVAS_CPE_MDCT )
    {
        tmpF = usdequant( hCPE->hStereoTCA->indx_ica_gD,
STEREO_TCA_GDMIN, STEREO_TCA_GDSTEP );
        for ( i = 0; i < STEREO_DFT_BAND_MAX; i++ )
        {
            hCPE->hStereoDft->side_gain[STEREO_DFT_BAND_MAX + i] =
tmpF;
        }
    }

    /* do not allow differential coding of DFT side parameters */
    hCPE->hStereoDft->ipd_counter = STEREO_DFT_FEC_THRESHOLD;
    hCPE->hStereoDft->res_pred_counter = STEREO_DFT_FEC_THRESHOLD;

    /* update DFT synthesis overlap memory @12.8kHz */
    for ( i = 0; i < STEREO_DFT_OVL_12k8; i++ )
    {
        hCPE->hStereoDft->output_mem_dmx_12k8[i] = st[0]-
>buf_speech_enc[L_FRAME32k + L_FRAME - STEREO_DFT_OVL_12k8 + i] * hCPE-
>hStereoDft->win_12k8[STEREO_DFT_OVL_12k8 - 1 - i];
    }

    /* update DFT synthesis overlap memory @16kHz, primary channel
only */
    lerp( hCPE->hStereoDft->output_mem_dmx, hCPE->hStereoDft-
>output_mem_dmx_16k, STEREO_DFT_OVL_16k, dft_ovl );

    /* reset DFT synthesis overlap memory @8kHz, secondary channel */
    set_f( hCPE->hStereoDft->output_mem_res_8k, 0, STEREO_DFT_OVL_8k
);

```

```

        hCPE->vad_flag[1] = 0;
    }

    /* DFT/MDCT -> TD stereo switching */
    if ( hCPE->element_mode == IVAS_CPE_TD && hCPE->last_element_mode !=
IVAS_CPE_TD )

    {
        hCPE->hStereoTD->tdm_last_ratio_idx = LRTD_STEREO_MID_IS_PRIM;
        hCPE->hStereoTD->tdm_last_ratio_idx_SM = LRTD_STEREO_MID_IS_PRIM;
        hCPE->hStereoTD->tdm_last_SM_flag = 0;
        hCPE->hStereoTD->tdm_last_inst_ratio_idx
LRTD_STEREO_MID_IS_PRIM;

        /* First frame after DFT frame AND the content is uncorrelated or
xtalk -> the primary channel is forced to left */
        if ( hCPE->hStereoClassif->lrtdd_mode == 1 )
        {
            hCPE->hStereoTD->tdm_last_ratio
ratio_tab1[LRTD_STEREO_LEFT_IS_PRIM];
            hCPE->hStereoTD->tdm_last_ratio_idx
LRTD_STEREO_LEFT_IS_PRIM;

            if ( hCPE->hStereoTCA->instTargetGain < 0.05f && ( hCPE-
>vad_flag[0] || hCPE->vad_flag[1] ) ) /* but if there is no content in
the L channel -> the primary channel is forced to right */
            {
                hCPE->hStereoTD->tdm_last_ratio
ratio_tab1[LRTD_STEREO_RIGHT_IS_PRIM];
                hCPE->hStereoTD->tdm_last_ratio_idx
LRTD_STEREO_RIGHT_IS_PRIM;
            }
        }
    }

    /* DFT -> TD stereo switching */
    if ( hCPE->element_mode == IVAS_CPE_TD && hCPE->last_element_mode ==
IVAS_CPE_DFT )
    {
        offset = st[0]->cldfbAnaEnc->p_filter_length - st[0]-
>cldfbAnaEnc->no_channels;

        mvr2r( old_input_signal_pri + input_frame - offset - NS2SA(
input_frame * 50, L_MEM_RECALC_TBE_NS ), st[0]->cldfbAnaEnc->cldfb_state,
offset );

        cldfb_reset_memory( st[0]->cldfbSynTd );
        st[0]->currEnergyLookAhead = 6.1e-5f;

        if ( hCPE->hStereoICBWE == NULL )
        {
            offset = st[1]->cldfbAnaEnc->p_filter_length - st[1]-
>cldfbAnaEnc->no_channels;

```

```

        if ( hCPE->hStereoTD->tdm_last_ratio_idx ==
LRTD_STEREO_LEFT_IS_PRIM )
        {
            v_multc( hCPE->hCoreCoder[1]->old_input_signal +
input_frame - offset - NS2SA( input_frame * 50, L_MEM_RECALC_TBE_NS ), -
1.0f, st[1]->cldfbAnaEnc->cldfb_state, offset );
        }
        else
        {
            mvr2r( hCPE->hCoreCoder[1]->old_input_signal +
input_frame - offset - NS2SA( input_frame * 50, L_MEM_RECALC_TBE_NS ),
st[1]->cldfbAnaEnc->cldfb_state, offset );
        }
        cldfb_reset_memory( st[1]->cldfbSynTd );
        st[1]->currEnergyLookAhead = 6.1e-5f;
    }

    st[1]->last_extl = -1;

    /* no secondary channel in the previous frame -> memory resets */
    set_zero( st[1]->old_inp_12k8, L_INP_MEM );
    /*set_zero( st[1]->old_inp_16k, L_INP_MEM );*/
    set_zero( st[1]->mem_decim, 2 * L_FILT_MAX );
    /*set_zero( st[1]->mem_decim16k, 2*L_FILT_MAX );*/
    st[1]->mem_preemph = 0;
    /*st[1]->mem_preemph16k = 0;*/

    set_zero( st[1]->buf_speech_enc, L_PAST_MAX_32k + L_FRAME32k +
L_NEXT_MAX_32k );
    set_zero( st[1]->buf_speech_enc_pe, L_PAST_MAX_32k + L_FRAME32k +
L_NEXT_MAX_32k );

    if ( st[1]->hTcxEnc != NULL )
    {
        set_zero( st[1]->hTcxEnc->buf_speech_ltp, L_PAST_MAX_32k +
L_FRAME32k + L_NEXT_MAX_32k );
    }
    set_zero( st[1]->buf_wspeech_enc, L_FRAME16k + L_SUBFR +
L_FRAME16k + L_NEXT_MAX_16k );
    set_zero( st[1]->buf_synth, OLD_SYNTH_SIZE_ENC + L_FRAME32k );
    st[1]->mem_wsp = 0.0f;
    st[1]->mem_wsp_enc = 0.0f;
    init_gp_clip( st[1]->clip_var );

    set_f( st[1]->Bin_E, 0, L_FFT );
    set_f( st[1]->Bin_E_old, 0, L_FFT / 2 );

    /* st[1]->hLPDmem reset already done in allocation of handles */

    st[1]->last_L_frame = st[0]->last_L_frame;

    pitch_ol_init( &st[1]->old_thres, &st[1]->old_pitch, &st[1]-
>delta_plt, &st[1]->old_corr );

```

```

set_zero( st[1]->old_wsp, L_WSP_MEM );
set_zero( st[1]->old_wsp2, ( L_WSP_MEM - L_INTERPOL ) / OPL_DECIM
);

set_zero( st[1]->mem_decim2, 3 );

st[1]->Nb_ACELP_frames = 0;

/* populate PCh memories into the Sch */
mvr2r(      st[0]->hLPDmem->old_exc,      st[1]->hLPDmem->old_exc,
L_EXC_MEM );
mvr2r( st[0]->lsp_old, st[1]->lsp_old, M );
mvr2r( st[0]->lsp_old, st[1]->lsp_old, M );
mvr2r( st[0]->lsp_old1, st[1]->lsp_old1, M );
mvr2r( st[0]->lsp_old1, st[1]->lsp_old1, M );

st[1]->GSC_noisy_speech = 0;
}
else if ( hCPE->element_mode == IVAS_CPE_TD && hCPE-
>last_element_mode == IVAS_CPE_MDCT )
{
set_f( st[0]->hLPDmem->old_exc, 0.0f, L_EXC_MEM );
set_f( st[1]->hLPDmem->old_exc, 0.0f, L_EXC_MEM );
}

```

1.2.5 Codificador ICA

En tramas estéreo TD, la operación de control de cambio de modo estéreo (no se muestra) comprende una operación temporal de alineación intercanal (ICA) 251. Para realizar la operación 251, el controlador de cambio de modo estéreo (no se muestra) comprende un codificador ICA 201 para alinear temporalmente los canales l y r de la señal estéreo de entrada y luego escalar el canal r.

Como se indica en la descripción anterior, antes de la mezcla de TD, la ICA se realiza utilizando la sincronización de ITD entre los dos canales de entrada l y r en el dominio de tiempo. Esto se logra retrasando uno de los canales de entrada (l o r) y extrapolando una parte faltante de la señal mezclada correspondiente a la duración del retardo de ITD; el valor máximo del retardo de ITD es de 7,5 ms. La alineación del tiempo, es decir, el cambio de tiempo ICA, se aplica primero y altera la mayor parte de la trama estéreo TD actual. La parte extrapolada de la señal mezclada de la anticipación se vuelve a calcular y, por lo tanto, se ajusta temporalmente en la siguiente trama en función de la ITD estimada en esa trama siguiente.

Cuando no se prevé ningún cambio de modo estéreo, la señal extrapolada de 7,5 ms se vuelve a calcular en el codificador ICA 201. Sin embargo, cuando puede ocurrir un cambio de modo estéreo, es decir, el cambio de modo estéreo DFT al modo estéreo TD, se vuelve a calcular una señal más larga. En este caso, la longitud corresponde a la longitud de la señal estéreo rectificada DFT más el retardo de remuestreo FIR, es decir, 8,75 ms + 0,9375 ms = 9,6875 ms. En la sección 1.4 se explican estas características con más detalle.

Otro propósito del codificador ICA 201 es el escalado del canal de entrada r. La ganancia de escalado, es decir, la ganancia objetivo mencionada anteriormente, se estima como una relación logarítmica de las energías de los canales l y r suavizadas con la ganancia objetivo de la trama anterior en cada trama, independientemente del modo estéreo DFT o TD que se utilice. La ganancia objetivo estimada en la trama actual (20 ms) se aplica a los últimos 15 ms del canal de entrada actual r, mientras que los primeros 5 ms del canal actual r se escalan mediante una combinación de las ganancias objetivo de la trama anterior y actual en un fundido de entrada y salida.

El codificador ICA 201 produce parámetros ICA 202 tales como el retardo ITD, la ganancia objetivo y un índice de canal objetivo.

1.2.6 Detectores de transitorios del dominio temporal

La operación de control de cambio de modo estéreo (no se muestra) comprende una operación 253 de detección de transitorios en el dominio de tiempo en el canal l del codificador ICA 201. Para realizar la operación 253, el controlador de cambio de modo estéreo (no se muestra) comprende un detector 203 para detectar transitorios en el dominio de tiempo en el canal l.

De la misma manera, la operación de control de cambio de modo estéreo (no se muestra) comprende una operación 254 de detectar transitorios en el dominio de tiempo en el canal r del codificador ICA 201. Para

realizar la operación 254, el controlador de cambio de modo estéreo (no se muestra) comprende un detector 204 para detectar transitorios en el dominio de tiempo en el canal r.

- 5 La detección de transitorios en el dominio temporal en los canales de dominio temporal l y r es un paso de preprocesamiento que permite detectar y, por lo tanto, procesar y codificar adecuadamente dichos transitorios en los módulos de codificación del núcleo del dominio de transformación (TCX principal, HQ principal, FD-BWE).

Se puede encontrar más información sobre los detectores de transitorios del dominio temporal 203 y 204 y las operaciones de detección de transitorios en el dominio temporal 253 y 254, por ejemplo, en la Referencia [1], Cláusula 5.1.8.

1.2.7 Configuraciones del codificador estéreo

- 10 Para realizar configuraciones de codificadores estéreo, el dispositivo codificador estéreo IVAS 200 establece parámetros de los codificadores estéreo 300, 400 y 500. Por ejemplo, se fija una velocidad de bits nominal para los codificadores del núcleo.

1.2.8 Análisis DFT, procesamiento estéreo y mezcla en el dominio DFT y síntesis IDFT

- 15 Con referencia a la Figura 3, el método de codificación estéreo DFT 350 comprende una operación 351 para aplicar una transformada DFT al canal l desde el detector de transitorios del dominio temporal 203 de la Figura 2. Para realizar la operación 351, el codificador estéreo DFT 300 comprende una calculadora 301 de la transformada DFT del canal l (análisis DFT) para producir un canal L en el dominio DFT.

- 20 El método de codificación estéreo DFT 350 también incluye una operación 352 para aplicar una transformada DFT al canal r desde el detector de transitorios del dominio temporal 204 de la Figura 2. Para realizar la operación 352, el codificador estéreo DFT 300 comprende una calculadora 302 de la transformada DFT del canal r (análisis DFT) para producir un canal R en el dominio DFT.

- 25 El método de codificación estéreo DFT 350 también incluye una operación 353 de procesamiento estéreo y mezcla en el dominio DFT. Para realizar la operación 353, el codificador estéreo DFT 300 comprende un procesador estéreo y un mezclador 303 para producir información lateral en un canal lateral S. La mezcla de los canales L y R también produce una señal residual en el canal lateral S. La información lateral y la señal residual del canal lateral S se codifican, por ejemplo, mediante una operación de codificación 354 y el codificador 304 correspondiente; después se multiplexan en un flujo de bits de salida 310 del codificador estéreo DFT 300. El procesador estéreo y el mezclador 303 también mezclan los canales izquierdo L y derecho R de las calculadoras DFT 301 y 302 para producir el canal medio M en el dominio DFT. Se puede encontrar más información sobre la operación 353 de procesamiento estéreo y mezcla, el procesador estéreo y mezclador 303, el canal medio M y la información lateral y la señal residual del canal lateral S, por ejemplo, en la Referencia [3].

- 35 En una operación de síntesis DFT inversa (IDFT) 355 del método de codificación estéreo de DFT 350, una calculadora 305 del codificador estéreo DFT 300 calcula la transformada IDFT m del canal medio M a la frecuencia de muestreo de la señal estéreo de entrada, por ejemplo, 12,8 kHz. De la misma manera, en una operación de síntesis DFT inversa (IDFT) 356 del método de codificación estéreo DFT 350, una calculadora 306 del codificador estéreo DFT 300 calcula la transformada IDFT m del canal M a la frecuencia de muestreo interna.

1.2.9 Análisis TD y mezcla en el dominio de TD

- 40 Con referencia a la Figura 4, el método de codificación estéreo TD 450 comprende una operación 451 de análisis del dominio temporal y mezcla ponderada en el dominio TD. Para realizar la operación 451, el codificador estéreo TD 400 comprende un analizador del dominio de tiempo y un mezclador 401 para calcular parámetros del lado estéreo 402 tales como un indicador de submodo, un índice de relación de mezcla o un indicador de reutilización de predicción lineal, que se multiplexan en un flujo de bits de salida 410 del codificador estéreo TD 400. El analizador del dominio de tiempo y el mezclador 401 también realizan la mezcla ponderada de los canales l y r de los detectores 203 y 204 (Figura 2) para producir el canal primario PCh y el canal secundario SCh utilizando una relación de mezcla estimada, en alineación con la escala ICA. Se puede encontrar más información sobre el analizador del dominio de tiempo y el mezclador 401 y la operación 451, por ejemplo, en la Referencia [4].

- 50 La mezcla mediante la relación de mezcla de trama actual se realiza, por ejemplo, en los últimos 15 ms de la trama actual de los canales de entrada l y r, mientras que los primeros 5 ms de la trama actual se mezclan utilizando una combinación de las relaciones de mezcla de la trama anterior y actual en un fundido de entrada y salida para suavizar la transición de un canal al otro. Los dos canales (canal primario PCh y canal secundario SCh) muestreados a la frecuencia de muestreo del canal de entrada estéreo, por ejemplo, 32 kHz, se vuelven

a muestrear usando filtros de decimación FIR a sus representaciones a 12,8 kHz y a la frecuencia de muestreo interna.

En el modo estéreo TD, no solo se mezcla la señal de entrada estéreo de la trama actual. También se vuelven a mezclar las señales mezcladas almacenadas que correspondan a la trama anterior. La longitud de la señal anterior sujeta a este recálculo se corresponde con la longitud de la señal desplazada temporalmente y recalculada en el módulo ICA, es decir, $8,75 \text{ ms} + 0,9375 \text{ ms} = 9,6875 \text{ ms}$.

1.2:10 Preprocesamiento frontal

En el códec IVAS (dispositivo de codificación estéreo IVAS 200 y dispositivo de decodificación estéreo IVAS 800), se produce una reestructuración del preprocesamiento tradicional, de forma que algunas decisiones de clasificación se toman en función de la velocidad de bits global del códec, mientras que otras decisiones se toman en función de la velocidad de bits de la codificación del núcleo. En consecuencia, el preprocesamiento tradicional, utilizado por ejemplo en el códec EVS (Referencia [1]), se divide en dos partes para garantizar que en cada trama procesada se utilice la mejor configuración de códec posible. Por lo tanto, la configuración del códec puede cambiar de trama a trama, mientras que ciertos cambios de configuración se pueden realizar lo más rápido posible, por ejemplo, aquellos basados en la actividad de la señal o en la clase de señal. Por otro lado, algunos cambios en la configuración del códec no deben producirse con demasiada frecuencia, por ejemplo, la selección del ancho de banda del audio codificado, la selección de la frecuencia de muestreo interna o la distribución del presupuesto de bits entre la codificación de banda baja y banda alta; los cambios demasiado frecuentes en dicha configuración del códec pueden dar lugar a una calidad inestable en la señal codificada o incluso a artefactos audibles.

La primera parte del preprocesamiento, el preprocesamiento frontal, puede incluir módulos de preprocesamiento y clasificación, como el muestreo a la velocidad de muestreo de preprocesamiento, el análisis espectral, la detección de ancho de banda (BWD), la detección de actividad sonora (SAD), el análisis de predicción lineal (LP), la búsqueda de altura de tono en bucle abierto, la clasificación de señales o la clasificación de voz/música. Cabe señalar que las decisiones en el preprocesamiento frontal dependen exclusivamente de la velocidad de bits general del códec. Se puede encontrar más información sobre las operaciones realizadas durante el preprocesamiento descrito anteriormente, por ejemplo, en la Referencia [1].

En el modo estéreo DFT (codificador estéreo DFT 300 de la Figura 3), el preprocesamiento frontal se realiza con un preprocesador frontal 307 y la operación de preprocesamiento frontal correspondiente 357 en el dominio de tiempo del canal medio m a la frecuencia de muestreo interna de la calculadora IDFT 306.

En el modo estéreo TD, el preprocesamiento frontal se realiza con (a) un preprocesador frontal 403 y la operación de preprocesamiento frontal correspondiente 453 en el canal primario PCh desde el analizador del dominio de tiempo y el mezclador 401, y (b) un preprocesador frontal 404 y la operación de preprocesamiento frontal correspondiente 454 en el canal secundario SCh desde el analizador del dominio de tiempo y el mezclador 401.

En el modo estéreo MDCT, el preprocesamiento frontal se realiza con un preprocesador frontal 503 y la operación de preprocesamiento frontal correspondiente 553 en el canal izquierdo de entrada l desde el detector de transitorios en el dominio de tiempo 203 (Figura 2), y (b) un preprocesador frontal 504 y la operación de preprocesamiento frontal correspondiente 554 en el canal derecho de entrada r desde el detector de transitorios en el dominio de tiempo 204 (Figura 2).

1.2:11 Configuración del codificador del núcleo

La configuración de los codificadores del núcleo se realiza en función de la velocidad de bits general del códec y del preprocesamiento frontal.

En concreto, en el codificador estéreo DFT 300 y el método de codificación estéreo DFT correspondiente 350 (Figura 3), un configurador del codificador del núcleo 308 y la operación de configuración del codificador del núcleo correspondiente 358 responden al canal medio m en el dominio temporal de la calculadora IDFT 305 y a la salida desde el preprocesador frontal 307 para configurar el codificador del núcleo 311 y la operación de codificación del núcleo correspondiente 361. El configurador del codificador del núcleo 308 es responsable, por ejemplo, de establecer la frecuencia de muestreo interna o de modificar la clasificación de tipo del codificador del núcleo. Se puede encontrar más información sobre la configuración del codificador del núcleo en el dominio DFT, por ejemplo, en las Referencias [1] y [2].

En el codificador estéreo TD 400 y el método de codificación estéreo TD correspondiente 450 (Figura 4), un configurador de codificadores del núcleo 405 y la operación de configuración de codificadores del núcleo correspondiente 455 responden al canal primario PCh y al canal secundario SCh preprocesados frontalmente desde los preprocesadores frontales 403 y 404, respectivamente, para realizar la configuración del codificador del núcleo 406 y la operación de codificación del núcleo correspondiente 456 del canal primario PCh y el

codificador del núcleo 407 y la operación de codificación del núcleo correspondiente 457 del canal secundario SCh. El configurador del codificador del núcleo 405 es responsable, por ejemplo, de establecer la frecuencia de muestreo interna o de modificar la clasificación de tipo del codificador del núcleo. Se puede encontrar más información sobre la configuración de codificadores del núcleo en el dominio TD, por ejemplo, en las Referencias [1] y [4].

1.2:12 Preprocesamiento posterior

El método de codificación DFT 350 comprende una operación 362 de preprocesamiento posterior. Para realizar la operación 362, un denominado preprocesador posterior 312 del codificador estéreo DFT 300 lleva a cabo una segunda parte del preprocesamiento, que puede incluir la clasificación, la selección de núcleo, el preprocesamiento a la frecuencia de muestreo interna de codificación, etc. Las decisiones en el preprocesador frontal 307 dependen de la velocidad de bits de la codificación del núcleo que, por lo general, fluctúa durante una sesión. Se puede encontrar más información sobre las operaciones realizadas durante dicho preprocesamiento posterior en el dominio DFT, por ejemplo, en la Referencia [1].

El método de codificación TD 450 comprende una operación 458 de preprocesamiento posterior. Para realizar la operación 458, un denominado preprocesador posterior 408 del codificador estéreo TD 400 realiza, antes de la codificación del núcleo del canal primario PCh, una segunda parte del preprocesamiento que puede incluir la clasificación, la selección del núcleo, el preprocesamiento a la frecuencia de muestreo interna de codificación, etc. Las decisiones en el preprocesador posterior 408 dependen de la velocidad de bits de la codificación del núcleo que, por lo general, fluctúa durante una sesión.

Además, el método de codificación TD 450 incluye una operación 459 de preprocesamiento posterior. Para realizar la operación 459, el codificador estéreo TD 400 comprende un denominado preprocesador posterior 409 para, antes de la codificación del núcleo del canal secundario SCh, realizar una segunda parte del preprocesamiento, que puede incluir clasificación, la selección del núcleo, el preprocesamiento a la frecuencia de muestreo interna de codificación, etc. Las decisiones en el preprocesador posterior 409 dependen de la velocidad de bits de la codificación del núcleo que, por lo general, fluctúa durante una sesión.

Se puede encontrar más información sobre dicho preprocesamiento posterior en el dominio TD, por ejemplo, en la Referencia [1].

El método de codificación MDCT 550 incluye una operación 555 de preprocesamiento posterior del canal izquierdo l. Para realizar la operación 555, un denominado preprocesador adicional 505 del codificador estéreo MDCT 500 lleva a cabo una segunda parte del preprocesamiento del canal izquierdo l, que puede incluir la clasificación, la selección de núcleo, el preprocesamiento a la frecuencia de muestreo interna de codificación, etc., antes de una operación 556 de codificación del núcleo conjunta del canal izquierdo l y el canal derecho r realizada por el codificador del núcleo conjunto 506 del codificador estéreo MDCT 500.

El método de codificación MDCT 550 incluye una operación 557 de preprocesamiento posterior del canal derecho r. Para realizar la operación 557, un denominado preprocesador posterior 507 del codificador estéreo MDCT 500 lleva a cabo una segunda parte del preprocesamiento del canal izquierdo l, que puede incluir la clasificación, la selección de núcleo, el preprocesamiento a la frecuencia de muestreo interno de codificación, etc., antes de la operación 556 de codificación del núcleo conjunta del canal izquierdo l y el canal derecho r realizada por el codificador del núcleo conjunto 506 del codificador estéreo MDCT 500.

Se puede encontrar más información sobre dicho preprocesamiento posterior en el dominio MDCT, por ejemplo, en la Referencia [1].

1.2:13 Codificación del núcleo

En general, el codificador del núcleo 311 en el codificador estéreo DFT 300 (que realiza la operación de codificación del núcleo 361) y los codificadores del núcleo 406 (que realiza la operación de codificación de núcleo 456) y 407 (que realiza la operación de codificación del núcleo 457) en el codificador estéreo TD 400 pueden ser cualquier códec mono con velocidad de bits variable. En la implementación ilustrativa de la presente divulgación, se usa el códec EVS (véase la Referencia [1]), con una capacidad de velocidad de bits fluctuante (véase la Referencia [5]). Por supuesto, se pueden considerar e implementar otros códecs adecuados. En el codificador estéreo MDCT 500, se emplea el codificador del núcleo conjunto 506, que puede ser, en general, un módulo de codificación estéreo con herramientas estereofónicas que procesa y cuantifica los canales l y r de manera conjunta.

1.2:14 Actualizaciones comunes de estéreo

Por último, se realizan actualizaciones estéreo comunes. Se puede encontrar más información sobre actualizaciones comunes de estéreo, por ejemplo, en la Referencia [1].

1.2:15 Flujos de bits

Con referencia a las Figuras 2 y 3, la señalización del modo estéreo 270 desde el clasificador estéreo y el selector de modo estéreo 205, un flujo de bits 313 desde la información lateral, el codificador de señal residual 304 y un flujo de bits 314 desde el codificador del núcleo 311 se multiplexan para formar el flujo de bits del codificador estéreo DFT 310 (y luego formar un flujo de bits de salida 206 del dispositivo codificador estéreo IVAS 200 (Figura 2)).

Con referencia a las Figuras 2 y 4, la señalización de modo estéreo 270 desde el clasificador estéreo y el selector de modo estéreo 205, los parámetros laterales 402 desde el analizador del dominio temporal y el mezclador 401, los parámetros ICA 202 desde el codificador ICA 201, un flujo de bits 411 desde el codificador del núcleo 406 y un flujo de bits 412 desde el codificador del núcleo 407 se multiplexan para formar el flujo de bits del codificador estéreo TD 410 (y luego formar el flujo de bits de salida 206 del dispositivo codificador estéreo IVAS 200 (Figura 2)).

Con referencia a las Figuras 2 y 5, la señalización de modo estéreo 270 desde el clasificador estéreo y el selector de modo estéreo 205, y un flujo de bits 509 desde el codificador del núcleo conjunto 506 se multiplexan para formar el flujo de bits del codificador estéreo MDCT 508 (y luego formar el flujo de bits de salida 206 del dispositivo codificador estéreo IVAS 200 (Figura 2)).

1.3 Cambio del modo estéreo TD al modo estéreo DFT en el codificador estéreo IVAS 200

El cambio de modo estéreo TD (codificador estéreo TD 400) al modo estéreo DFT (codificador estéreo DFT 300) es relativamente sencillo, como se muestra en la Figura 6.

En concreto, la Figura 6 es un diagrama de flujo que ilustra las operaciones de procesamiento en el dispositivo codificador estéreo IVAS 200 y el método 250 al cambiar del modo estéreo TD al modo estéreo DFT. Como se puede ver, la Figura 5 muestra dos tramas de señal de entrada estéreo, es decir, una trama estéreo TD 601 seguida de una trama estéreo DFT 602, con diferentes operaciones de procesamiento e instancias de tiempo relacionadas al cambiar del modo estéreo TD al modo estéreo DFT.

Hay disponible una anticipación lo suficientemente larga, el remuestreo se realiza en el dominio DFT (por lo tanto, no hay manejo de memoria de filtro de decimación FIR) y hay una transición desde dos codificadores del núcleo 406 y 407 en la última trama estéreo TD 501 hasta un codificador del núcleo 311 en la primera trama estéreo DFT 502.

Las siguientes operaciones realizadas al cambiar del modo estéreo TD (codificador estéreo TD 400) al modo estéreo DFT (codificador estéreo DFT 300) se realizan mediante el controlador de cambio de modo estéreo mencionado anteriormente (no se muestra) en respuesta a la selección del modo estéreo.

El caso A) de la Figura 6 se refiere a una actualización de la memoria de análisis DFT, específicamente a la memoria de análisis DFT estéreo OLA como parte de la estructura de datos estéreo DFT que está sujeta a la división en ventanas previa a las operaciones de cálculo DFT 351 y 352. Esta actualización la realiza el controlador de cambio de modo estéreo (no se muestra) antes de la alineación intercanal (ICA) (véase 251 en la Figura 2) y comprende el almacenamiento de muestras relacionadas con los últimos 8,75 ms de la trama estéreo TD actual 601 de los canales l y r de la señal estéreo de entrada. Esta actualización se realiza en cada trama estéreo TD en ambos canales l y r. Se puede encontrar más información sobre la memoria de análisis DFT, por ejemplo, en las Referencias [1] y [2].

El caso B) de la Figura 6 se refiere a una actualización de la memoria de síntesis DFT, específicamente a la memoria de síntesis OLA como parte de la estructura de datos estéreo DFT que resulta de la división en ventanas después de las operaciones de cálculo IDFT 355 y 356, al cambiar del modo estéreo TD al modo estéreo DFT. El controlador de cambio de modo estéreo (no se muestra) realiza esta actualización en la primera trama estéreo DFT 602 después de la trama estéreo TD 601 y, para esta actualización, usa las memorias estéreo TD como parte de la estructura de datos estéreo TD que se emplea para el procesamiento estéreo TD correspondiente al canal primario PCh mezclado. Se puede encontrar más información sobre la memoria de síntesis DFT, por ejemplo, en las Referencias [1] y [2], y sobre las memorias estéreo TD, por ejemplo, en la Referencia [4].

Comenzando con la primera trama estéreo DFT 602, ciertas estructuras de datos relacionadas con el estéreo TD, por ejemplo, la estructura de datos estéreo TD (utilizada en el codificador estéreo TD 400) y una estructura de datos del codificador del núcleo 407 relacionada con el canal secundario SCh, ya no son necesarias y, por lo tanto, se desasignan, es decir, se liberan mediante el controlador de cambio de modo estéreo (no se muestra).

En la trama estéreo DFT 602 que sigue a la trama estéreo TD 601, el controlador de cambio de modo estéreo (no se muestra) continúa la operación de codificación del núcleo 361 en el codificador del núcleo 311 del codificador estéreo DFT 300 con memorias del codificador del núcleo del canal PCh primario 406 (por ejemplo,

memoria de síntesis, memoria de preénfasis, señales y parámetros anteriores, etc.) en la trama estéreo TD 601 anterior mientras controla las diferencias de instancia de tiempo entre los modos estéreo TD y DFT para asegurar la continuidad de varios búferes del codificador del núcleo, por ejemplo, búferes de señal de entrada de preénfasis, búferes de entrada HB, etc., que se utilizarán posteriormente en el codificador de banda baja o en el codificador de banda alta FD-BWE. Se puede encontrar más información sobre la operación de codificación del núcleo 361, las memorias del codificador del núcleo del canal PCh 406, los búferes de señal de entrada de preénfasis, los búferes de entrada HB, etc., por ejemplo, en la Referencia [1].

1.4 Cambio del modo estéreo DFT al modo estéreo TD en el codificador estéreo IVAS 200

El cambio de modo estéreo DFT al modo estéreo TD es más complicado que el cambio de modo estéreo TD al modo estéreo DFT, debido a la estructura más compleja del codificador estéreo TD 400. Las siguientes operaciones realizadas al cambiar del modo estéreo DFT (codificador estéreo DFT 300) al modo estéreo TD (codificador estéreo TD 400) las realiza el controlador de cambio de modo estéreo (no se muestra) en respuesta a la selección del modo estéreo.

La Figura 7a es un diagrama de flujo que ilustra las operaciones de procesamiento en el dispositivo codificador estéreo IVAS 200 y el método 250 al cambiar del modo estéreo DFT al modo estéreo TD. En particular, la Figura 7a muestra dos tramas de la señal de entrada estéreo, es decir, una trama estéreo DFT 701 seguida de una trama estéreo TD 702, en diferentes operaciones de procesamiento con instancias de tiempo relacionadas al cambiar del modo estéreo DFT al modo estéreo TD.

El caso A) de la Figura 7a se refiere a la actualización de la memoria del filtro de remuestreo FIR (utilizada en el remuestreo FIR desde la frecuencia de muestreo de la señal estéreo de entrada hasta la frecuencia de muestreo de 12,8 kHz y la frecuencia de muestreo del codificador del núcleo interno) usada en el canal primario PCh del modo de codificación estéreo TD. El controlador de cambio de modo estéreo (no se muestra) realiza esta actualización en cada trama estéreo DFT usando el canal medio mezclado m, y se corresponde con un segmento 703 de $2 \times 0,9375$ ms de longitud antes del último segmento de 7,5 ms de longitud en la trama estéreo DFT 701 (véase 704), lo que garantiza la continuidad de la memoria de remuestreo FIR para el canal primario PCh.

Dado que el canal lateral s (Figura 3) del método de codificación estéreo DFT 350 no está disponible aunque se use a, por ejemplo, la frecuencia de muestreo de 12,8 kHz, a la frecuencia de muestreo de la señal estéreo de entrada y a la frecuencia de muestreo interna, el controlador de cambio de modo estéreo (no se muestra) llena de manera diferente la memoria del filtro de remuestreo FIR del canal secundario mezclado SCh. Con el fin de reconstruir la longitud completa de la señal mezclada a la frecuencia de muestreo interna para el codificador del núcleo 407, se recalcula un segmento de 8,75 ms (véase 705) de la señal mezclada de la trama anterior en la trama estéreo TD 702. Así, la actualización de la memoria del filtro de remuestreo del canal secundario mezclado SCh FIR corresponde a un segmento 708 de $2 \times 0,9375$ ms de longitud del canal medio mezclado m antes del último segmento de 8,75 ms de longitud (véase 705); esto se hace en la primera trama estéreo TD 702 después de cambiar desde la trama estéreo DFT 701 anterior. La actualización de la memoria del filtro de remuestreo FIR del canal secundario SCh se denomina caso C) en la Figura 7a. Como se puede observar, el controlador de cambio de modo estéreo (no se muestra) recalcula en la trama estéreo TD una longitud (véase 706) de la señal mezclada que es más larga en el canal secundario SCh que la longitud recalculada de la señal mezclada en el canal primario PCh (véase 707).

El caso B) de la figura 7a se refiere a la actualización (recálculo) de los canales primario PCh y secundario SCh en la primera trama estéreo TD 702 que sigue a la trama estéreo DFT 701. Las operaciones del caso B) realizadas por el controlador de cambio de modo estéreo (no se muestra) se ilustran con más detalle en la Figura 7b. Como se menciona en la descripción anterior, la Figura 7b es un diagrama de flujo que ilustra las operaciones de procesamiento al cambiar del modo estéreo DFT al modo estéreo TD.

Con referencia a la Figura 7b, en una operación 710, el controlador de cambio de modo estéreo (no se muestra) recalcula la memoria ICA empleada en el análisis y cálculo ICA (véase la operación 251 en la Figura 2) y, más tarde, como señal de entrada para los codificadores del núcleo y de preprocesamiento (véanse las operaciones 453-454 y 456-459) de 9,6875 ms de longitud (tal y como se señala en las secciones 1.2.7-1.2.9 de la presente divulgación) de los canales l y r correspondientes a la trama estéreo DFT 701 anterior.

Así, en las operaciones 712 y 713, el controlador de cambio de modo estéreo (no se muestra) recalcula los canales primarios PCh y secundarios SCh de la trama estéreo DFT 701 al mezclar los canales l y r procesados por ICA utilizando la relación de mezcla estéreo de esa trama 701.

Para el canal secundario SCh, la longitud (véase 714) del segmento pasado que debe recalcular el controlador de cambio de modo estéreo (no se muestra) en la operación 712 es 9,6875 ms, aunque se recalcula un segmento de longitud de solo 7,5 ms (véase 715) cuando no hay un cambio de modo de codificación estéreo. Para el canal primario PCh (véase la operación 713), la longitud del segmento que debe recalcular el controlador de cambio de modo estéreo (no se muestra) usando la relación de mezcla estéreo TD de la trama

anterior 701 es siempre 7,5 ms (véase 715). Esto garantiza la continuidad de los canales primario PCh y secundario SCh.

Se emplea una señal mezclada continua al cambiar del canal medio m de la trama estéreo DFT 701 al canal primario PCh de la trama estéreo TD 702. Para ello, el controlador de cambio de modo estéreo (no se muestra) realiza un fundido cruzado (717) en el segmento de 7,5 ms de longitud (véase 715) del canal medio DFT m con el canal primario recalculado PCh (713) de la trama estéreo DFT 701 para suavizar la transición y ecualizar para diferentes energías de señal mezclada entre el modo estéreo DFT y el modo estéreo TD. La reconstrucción del canal secundario SCh en la operación 712 usa la relación de mezcla de la trama 701, y no se aplica más suavizado porque el canal secundario SCh de la trama estéreo DFT 701 no está disponible.

La codificación del núcleo en la primera trama estéreo TD 702 posterior a la trama estéreo DFT 701 continúa después con el remuestreo de las señales mezcladas usando los filtros FIR, el preénfasis de estas señales, el cálculo de las señales HB, etc. Se puede encontrar más información sobre estas operaciones, por ejemplo, en la Referencia [1].

Con respecto al filtro de preénfasis implementado como un filtro de paso alto de primer orden utilizado para enfatizar las frecuencias más altas de la señal de entrada (véase la Referencia [1], Cláusula 5.1.4), el controlador de cambio de modo estéreo (no se muestra) almacena dos valores de la memoria del filtro de preénfasis en cada marco estéreo DFT. Estos valores de memoria se corresponden con instancias de tiempo basadas en diferentes longitudes de recálculo de los modos estéreo DFT y TD. Este mecanismo garantiza un recálculo óptimo de la señal de preénfasis en el canal m o en el canal primario PCh con una longitud de señal mínima. Para el canal secundario SCh del modo estéreo TD, la memoria del filtro de preénfasis se establece en cero antes de procesar la primera trama estéreo TD.

Comenzando con la primera trama estéreo TD 702 después de la trama estéreo DFT 701, ciertas estructuras de datos estéreo DFT relacionadas (por ejemplo, la estructura de datos estéreo DFT mencionada anteriormente) no son necesarias, por lo que son desasignadas/liberadas por el controlador de cambio de modo estéreo (no se muestra). Por otro lado, se asigna e inicializa una segunda instancia de la estructura de datos del codificador del núcleo para la codificación del núcleo (operación 457) del canal secundario SCh. La mayoría de las estructuras de datos del codificador del núcleo del canal secundario SCh se restablecen, aunque algunas de ellas se estiman para que las transiciones de cambio sean más suaves. Por ejemplo, el búfer de excitación anterior (libro de código adaptativo del núcleo ACELP), los parámetros LSF anteriores y los parámetros LSP (véase la Referencia [1]) del canal secundario SCh se rellenan a partir de sus homólogos en el canal primario PCh. El restablecimiento o la estimación de los búferes anteriores del canal secundario SCh pueden originar una serie de artefactos. Aunque muchos de estos artefactos se suprimen significativamente en los procesos basados en suavizado en el decodificador, algunos de ellos pueden seguir siendo una fuente de artefactos subjetivos.

1.5 Cambio del modo estéreo TD al modo estéreo MDCT en el codificador estéreo IVAS 200

Cambiar del modo estéreo TD al modo estéreo MDCT es relativamente sencillo, pues ambos modos estéreo manejan dos canales de entrada y emplean dos instancias de codificador del núcleo. El principal obstáculo es mantener la fase correcta de los canales de entrada izquierdo y derecho.

Para mantener la fase correcta de los canales de entrada izquierdo y derecho de la señal de sonido estéreo, el controlador de cambio de modo estéreo (no se muestra) altera la mezcla estéreo TD. En la última trama estéreo TD antes de la primera trama estéreo MDCT, la relación de mezcla estéreo TD se establece en $\beta = 1,0$ y se implementa una mezcla de fase opuesta de los canales izquierdo y derecho de la señal de sonido estéreo utilizando, por ejemplo, la siguiente fórmula para la mezcla estéreo TD:

$$PCh(i) = r(i) \cdot (1 - \beta) + l(i) \cdot \beta$$

$$SCh(i) = l(i) \cdot (1 - \beta) + r(i) \cdot \beta$$

donde $PCh(i)$ es el canal primario TD, $SCh(i)$ es el canal secundario TD, $l(i)$ es el canal izquierdo, $r(i)$ es el canal derecho, β es la relación de mezcla estéreo TD, e i es el índice de tiempo discreto.

A su vez, esto significa que el canal primario estéreo TD $PCh(i)$ es idéntico al canal izquierdo anterior MDCT estéreo $l_{past}(i)$ y el canal secundario estéreo TD $SCh(i)$ es idéntico al canal derecho anterior MDCT estéreo $r_{past}(i)$, donde i es el índice de tiempo discreto. En aras de la exhaustividad, cabe señalar que el controlador de cambio de modo estéreo (no se muestra) puede usar en la última trama estéreo TD una mezcla estéreo TD predeterminada utilizando, por ejemplo, la siguiente fórmula:

$$PCh(i) = r(i) \cdot (1 - \beta) + l(i) \cdot \beta$$

$$SCh(i) = l(i) \cdot (1 - \beta) + r(i) \cdot \beta$$

A continuación, durante el procesamiento estéreo MDCT habitual (sin cambio de modo estéreo), el preprocesamiento frontal (preprocesadores frontales 503 y 504 y operaciones de preprocesamiento frontales 553 y 554) no recalcula la anticipación de los canales izquierdo l y derecho r de la señal de sonido estéreo, excepto para su último segmento de 0,9375 ms de longitud. Sin embargo, en la práctica, la anticipación de la longitud de 7,5 + 0,9375 ms está sujeta a un nuevo cálculo a la frecuencia de muestreo interna (12,8 kHz en esta implementación ilustrativa y no limitativa). Por lo tanto, no se necesita ninguna manipulación específica para mantener la continuidad de las señales de entrada a la frecuencia de muestreo de entrada.

A continuación, en el procesamiento estéreo MDCT habitual (sin cambio de modo estéreo), el preprocesamiento posterior (otros preprocesadores 505 y 507 y operaciones de preprocesamiento frontal 555 y 557) no recalcula la anticipación de los canales izquierdo l y derecho r de la señal de sonido estéreo, excepto su último segmento de 0,9375 ms de longitud. A diferencia del preprocesamiento frontal, en el preprocesamiento posterior se vuelven a calcular las señales de entrada (canales izquierdo l y derecho r de la señal de sonido estéreo) a la frecuencia de muestreo interna (12,8 kHz en esta implementación ilustrativa y no limitativa) de solo 0,9375 ms de longitud.

En otras palabras:

el codificador estéreo MDCT 500 comprende (a) los preprocesadores frontales 503 y 504 que, en el segundo modo estéreo MDCT, recalculan la anticipación de la primera duración de los canales izquierdo l y derecho r de la señal de sonido estéreo a la frecuencia de muestreo interna y (b) otros preprocesadores que, en el segundo modo estéreo MDCT, recalculan un último segmento de una duración determinada de la anticipación de los canales izquierdo l y derecho r de la señal de sonido estéreo a la frecuencia de muestreo interna, donde la primera y la segunda duración son diferentes.

La operación de codificación estéreo MDCT 550 comprende, en el segundo modo estéreo MDCT, (a) el recálculo de la anticipación de la primera duración de los canales izquierdo l y derecho r de la señal de sonido estéreo a la frecuencia de muestreo interna y (b) el recálculo de un último segmento de duración determinada de la vista anticipación de los canales izquierdo l y derecho r de la señal de sonido estéreo a la frecuencia de muestreo interna, donde la primera y la segunda duración son diferentes.

1.6 Cambio del modo estéreo MDCT al modo estéreo TD en el codificador estéreo IVAS 200

De forma similar al cambio de modo estéreo TD al modo estéreo MDCT, en este escenario siempre hay dos canales de entrada disponibles y siempre se emplean dos instancias de codificador del núcleo. El principal obstáculo es, de nuevo, mantener la fase correcta de los canales de entrada izquierdo y derecho. Por lo tanto, en la primera trama estéreo TD después de la última trama estéreo MDCT, el controlador de cambio de modo estéreo (no se muestra) establece la relación de mezcla estéreo TD a $\beta = 1,0$ y altera la mezcla estéreo TD usando el esquema de mezcla de la fase opuesta, de manera similar a lo que se describe en la Sección 1.5.

Otra característica específica del cambio de modo estéreo MDCT al modo estéreo TD es que el controlador de cambio de modo estéreo (no se muestra) reconstruye adecuadamente en la primera trama TD el segmento anterior de canales de entrada de la señal de sonido estéreo a la frecuencia de muestreo interna. Por lo tanto, se reconstruye una parte de la anticipación que corresponde a $8,75 - 7,5 = 1,25$ ms (remuestreada y con preénfasis) en la primera trama estéreo TD.

1.7 Cambio del modo estéreo DFT al modo estéreo MDCT en el codificador estéreo IVAS 200

En este escenario se usa un mecanismo similar al cambio de modo estéreo DFT al modo estéreo TD que se describe anteriormente, donde los canales primarios PCh y secundarios SCh del modo estéreo TD se sustituyen por los canales izquierdo l y derecho r del modo estéreo MDCT.

1.8 Cambio del modo estéreo MDCT al modo estéreo DFT en el codificador estéreo IVAS 200

En este escenario se usa un mecanismo similar al cambio de modo estéreo TD al modo estéreo DFT que se describe anteriormente, donde los canales primarios PCh y secundarios SCh del modo estéreo TD se sustituyen por los canales izquierdo l y derecho r del modo estéreo MDCT.

2. Cambio entre los modos estéreo en el decodificador estéreo IVAS 800 y el método 850

La Figura 8 es un diagrama de bloques de alto nivel que ilustra simultáneamente un dispositivo de decodificación estéreo IVAS 800 y el método de decodificación correspondiente 850, donde el dispositivo de decodificación estéreo IVAS 800 comprende un decodificador estéreo DFT 801 y el método de decodificación estéreo DFT correspondiente 851, un decodificador estéreo TD 802 y el método de decodificación estéreo TD correspondiente 852 y un decodificador estéreo MDCT 803 y el método de decodificación estéreo MDCT correspondiente 853. Para simplificar, solo se muestran y describen los modos estéreo DFT, TD y MDCT; sin embargo, está dentro del alcance de la presente divulgación usar e implementar otros tipos de modos estéreo.

El dispositivo de decodificación estéreo IVAS 800 y el método de decodificación correspondiente 850 reciben un flujo de bits 830 transmitido desde el dispositivo de codificación estéreo IVAS 200. En términos generales, el dispositivo de decodificación estéreo IVAS 800 y el método de decodificación correspondiente 850 decodifica, a partir del flujo de bits 830, tramas sucesivas de una señal estéreo codificada, por ejemplo, tramas de 20 ms de longitud como en el caso del códec EVS; realiza una mezcla de las tramas decodificadas, y finalmente, produce una señal de salida estéreo que incluye los canales l y r.

2.1 Diferencias entre los distintos decodificadores estéreo y métodos de decodificación

La decodificación del núcleo, realizada a la frecuencia de muestreo interna, es básicamente la misma independientemente del modo estéreo real; sin embargo, la decodificación del núcleo se realiza una vez (canal medio m) para una trama estéreo DFT y dos veces para una trama estéreo TD (canales primario PCh y secundario SCh) o para una trama estéreo MDCT (canales izquierdo l y derecho r). La cuestión es mantener (actualizar) las memorias del canal secundario SCh de una trama estéreo TD al cambiar de una trama estéreo DFT a una trama estéreo TD para mantener (actualizar) memorias del canal r de una trama estéreo MDCT al cambiar de una trama estéreo DFT a una trama estéreo MDCT.

Además, otras operaciones de decodificación posteriores a la decodificación del núcleo dependen en gran medida del modo estéreo real, lo que complica el cambio entre los modos estéreo. Las diferencias más fundamentales son las siguientes:

Decodificador estéreo DFT 801 y método de decodificación 851:

- El remuestreo de la síntesis del núcleo descodificado desde la frecuencia de muestreo interna a la frecuencia de muestreo de la señal estéreo de salida se realiza en el dominio DFT con un análisis DFT y una longitud de la ventana de superposición de síntesis de 3,125 ms.

- El ajuste del filtrado posterior de graves de banda baja (LB) (en tramas ACELP) se realiza en el dominio DFT.

- El cambio de núcleo (núcleo ACELP <-> núcleo TCX/HQ) se realiza en el dominio DFT con un retardo disponible de 3,125 ms.

- La sincronización entre la síntesis de LB y la síntesis de HB (en tramas ACELP) no requiere ningún retardo adicional.

- La mezcla estéreo se realiza en el dominio DFT con un retardo disponible de 3,125 ms.

- La sincronización de tiempo para coincidir con un retardo general del decodificador (que es de 3,25 ms) se aplica con una longitud de 0,125 ms.

Decodificador estéreo TD 802 y método de decodificación 852: (Se puede encontrar más información sobre el decodificador estéreo TD, por ejemplo, en la Referencia [4])

- El remuestreo de la síntesis del núcleo descodificado de la frecuencia de muestreo interna a la frecuencia de muestreo de la señal estéreo de salida se realiza utilizando los filtros CLDFB con un retardo de 1,25 ms.

- El ajuste del filtrado posterior de graves LB (en tramas ACELP) se realiza en el dominio CLDFB.

- El cambio de núcleo (núcleo ACELP <-> núcleo TCX/HQ) se realiza en el dominio de tiempo con un retardo disponible de 1,25 ms.

- La sincronización entre la síntesis de LB y la síntesis de HB (en marcos ACELP) introduce un retardo adicional.

- La mezcla estéreo se realiza en el dominio TD con un retardo cero.

- La sincronización de tiempo para coincidir con un retardo general del decodificador se aplica con una duración de 2,0 ms.

Decodificador estéreo MDCT 803 y método de decodificación 853:

- Solo se utiliza un decodificador del núcleo basado en TCX, por lo que solo se utiliza un ajuste de retardo de 1,25 ms para sincronizar las señales de síntesis del núcleo entre diferentes núcleos.
 - Se omite el filtrado posterior de graves LB (en tramas ACELP).
 - El cambio de núcleo (núcleo ACELP <-> núcleo TCX/HQ) se realiza en el dominio de tiempo solo en la primera trama estéreo MDCT después de la trama estéreo TD o DFT con un retardo disponible de 1,25 ms.
 - La sincronización entre la síntesis de LB y la síntesis de HB no es relevante.
 - Se omite la mezcla estéreo.
 - La sincronización de tiempo para coincidir con un retardo general del decodificador se aplica con una duración de 2,0 ms.
- 10 Las diferentes operaciones durante la decodificación, principalmente el procesamiento del dominio DFT «vs» TD, y los diferentes esquemas de retardo entre el modo estéreo DFT y el modo estéreo TD se consideran cuidadosamente en el procedimiento descrito a continuación para cambiar entre los modos estéreo DFT y TD.
- 2.2 Procesamiento en el dispositivo de decodificación estéreo IVAS 800 y método de decodificación 850
- 15 La siguiente Tabla III enumera en un orden secuencial las operaciones de procesamiento en el dispositivo de decodificación estéreo IVAS 800 para cada trama dependiendo del modo estéreo DFT, TD o MDCT actual (Véase también la Figura 8).

Tabla III - Pasos de procesamiento en el dispositivo de decodificación estéreo IVAS 800

Modo estéreo DFT	Modo estéreo TD	Modo estéreo MDCT
Lectura de información sobre el modo estéreo y el ancho de banda de audio		
Asignación de memoria		
Actualizaciones del cambio de modo estéreo		
Configuración del decodificador estéreo		
Configuración del codificador del núcleo		
	Configuración del decodificador estéreo TD	
Decodificación del núcleo		Decodificación estéreo conjunta
Cambio de núcleo en el dominio DFT	Cambio de núcleo en el dominio TD	
	Actualización de memorias de superposición del modo estéreo DFT Actualización del búfer de superposición MDCT estéreo TCX	Restablecimiento/actualización de memorias de solapamiento estéreo DFT
Análisis DFT		
Decodificación estéreo DFT incl. decodificación residual		
Mezcla en el dominio DFT	Mezcla en el dominio TD	
Síntesis de DFT		

Sincronización de síntesis	
IC-BWE, adición de síntesis de HB	
Decodificador ICA - ajuste temporal	
Actualizaciones comunes de estéreo	

El método de decodificación estéreo IVAS 850 comprende una operación (no se muestra) para controlar el cambio entre los modos estéreo DFT, TD y MDCT. Para realizar la operación de control de cambio, el dispositivo de decodificación estéreo IVAS 800 comprende un controlador (no se muestra) de cambio entre los

5

10

1) Señales y memorias mezcladas de filtros posteriores del núcleo a la frecuencia de muestreo interna, utilizadas en la decodificación del núcleo;

- Decodificador estéreo DFT 801: canal medio m;

- Decodificador estéreo TD 802: canal primario PCh y canal secundario SCh;

- Decodificador estéreo MDCT 803: canal izquierdo l y canal derecho r (sin mezclar).

15

2) Memorias de filtro posterior TCX-LTP (*Transform Coded eXcitation - Long Term Prediction*). El filtro posterior TCX-LTP se utiliza para interpolar entre muestras de síntesis anteriores utilizando filtros de interpolación FIR de polifase (véase la referencia [1], Cláusula 6.9.2);

20

3) Memorias de análisis OLA de DFT a la frecuencia de muestreo interna y a la frecuencia de muestreo de la señal estéreo de salida utilizada en la parte de OLA de la división en ventanas en las tramas anteriores y actuales antes de la operación de DFT 854;

4) Memorias de síntesis OLA de DFT utilizadas en la parte OLA de la división en ventanas en las tramas anteriores y actuales después de las operaciones IDFT 855 y 856 a la frecuencia de muestreo de la señal estéreo de salida;

5) Salida de la señal estéreo, incluidos los canales l y r; y

25

6) Memorias de la señal HB (véase la Referencia [1], Cláusula 6.1.5), canales l y r, utilizados en BWE e IC-BWE.

Aunque es relativamente sencillo mantener la continuidad para un canal (canal medio m en el modo estéreo DFT, canal primario PCh en el modo estéreo TD o canal l en el modo estéreo MDCT) en el punto 1) anterior, es difícil para el canal secundario SCh en el punto 1) anterior y también para las señales/memorias en los puntos 2) - 6) debido a varios aspectos, por ejemplo, la señal anterior y las memorias del canal secundario SCh que faltan por completo, una mezcla diferente, un retardo predeterminado diferente entre el modo estéreo DFT y el modo estéreo TD, etc. Además, un retardo del decodificador más corto (3,25 ms) en comparación con el retardo del codificador (8,75 ms) complica aún más el proceso de decodificación.

30

2.2.1 Lectura de información sobre el modo estéreo y el ancho de banda de audio

35

El método de decodificación estéreo IVAS 850 comienza con la lectura (no se muestra) de información sobre el modo estéreo y el ancho de banda de audio del flujo de bits 830. Basándose en el modo estéreo leído actualmente, las operaciones de decodificación relacionadas se realizan para cada modo estéreo particular (véase la Tabla III) mientras se mantienen las memorias y los búferes de los otros modos estéreo.

2.2.2 Asignación de memoria

40

De manera similar al dispositivo codificador estéreo IVAS 200, en una operación de asignación de memoria (no se muestra), el controlador de cambio de modo estéreo (no se muestra) asigna/designa dinámicamente estructuras de datos (memoria estática) dependiendo del modo estéreo actual. El controlador de cambio de modo estéreo (no se muestra) mantiene el impacto de la memoria estática del códec lo más bajo posible

manteniendo solo las partes de la memoria estática que se utilizan en la trama actual. En la Tabla II se resumen las estructuras de datos asignadas en un modo estéreo concreto.

Además, el controlador de cambio de modo estéreo (no se muestra) lee un indicador de submodo estéreo LRTD para distinguir entre el modo estéreo TD normal y el modo estéreo LRTD. Basándose en el indicador de submodo, el controlador de cambio de modo estéreo (no se muestra) asigna/desasigna las estructuras de datos relacionadas dentro del modo estéreo TD, tal y como se muestra en la Tabla II.

2.2.3 Actualizaciones del cambio de modo estéreo

De forma similar al dispositivo codificador estéreo IVAS 200, el controlador de cambio de modo estéreo (no se muestra) maneja memorias en caso de cambiar de uno de los modos estéreo DFT, TD, y MDCT a otro modo estéreo. Esto mantiene actualizados los parámetros a largo plazo y actualiza o restablece los búferes anteriores.

Al recibir una primera trama estéreo DFT después de una trama estéreo TD o MDCT, el controlador de cambio de modo estéreo (no se muestra) realiza una operación de restablecimiento de la estructura de datos estéreo DFT (ya definida en relación con el codificador estéreo DFT 300). Al recibir una primera trama estéreo TD después de una trama estéreo DFT o MDCT, el controlador de cambio de modo estéreo realiza una operación de restablecimiento de la estructura de datos estéreo TD (ya descrita en relación con el decodificador estéreo TD 400). Finalmente, al recibir una primera trama estéreo MDCT después de una trama estéreo DFT o TD, el controlador de cambio de modo estéreo (no se muestra) realiza una operación de restablecimiento de la estructura de datos estéreo MDCT. De nuevo, al cambiar de uno de los modos estéreo DFT y TD al otro modo estéreo, el controlador de cambio de modo estéreo (no se muestra) realiza una operación de transferencia de algunos parámetros relacionados con el estéreo entre las estructuras de datos, tal y como se describe en relación con el dispositivo codificador estéreo IVAS 200 (véase la Sección 1.2.4 anterior).

Las actualizaciones/restablecimientos relacionados con el canal secundario SCh de la decodificación del núcleo se describen en la Sección 2.4.

También se puede encontrar más información sobre las operaciones de configuración del decodificador estéreo, la configuración del decodificador del núcleo, la configuración del decodificador estéreo TD, la decodificación del núcleo, el cambio del núcleo en el dominio DFT, el cambio del núcleo en el dominio TD, por ejemplo, en la Tabla III, en las Referencias [1] y [2].

2.2.4 Actualización de las memorias de superposición del modo estéreo DFT

El controlador de cambio de modo estéreo (no se muestra) mantiene o actualiza las memorias OLA DFT en cada trama estéreo TD o MDCT (véase «Actualización de las memorias de superposición del modo estéreo DFT», «Actualización del búfer de superposición TCX del MDCT estéreo» y «Restablecimiento/actualización de las memorias de superposición DFT estéreo» de la Tabla III). De esta manera, las memorias OLA DFT actualizadas están disponibles para la siguiente trama estéreo DFT. El mecanismo de mantenimiento/actualización real y los búferes de memoria relacionados se describen más adelante, en la Sección 2.3 de la presente divulgación. A continuación se ofrece un ejemplo de implementación de actualización de las memorias DFT estéreo OLA realizadas en tramas estéreo TD o MDCT en el código fuente C.

```
if ( st[n]->element_mode != IVAS_CPE_DFT )
{
    ivas_post_proc( ... );

    /* update OLA buffers - needed for switching to DFT stereo */
    stereo_td2dft_update( hCPE, n, output[n], synth[n], hb_synth[n],
output_frame );

    /* update ovl buffer for possible switching from TD stereo SCh ACELP
frame to MDCT stereo TCX frame */
    if ( st[n]->element_mode == IVAS_CPE_TD && n == 1 && st[n]->hTorDec
== NULL )
    {
        mvr2r( output[n] + st[n]->L_frame / 2, hCPE->hStereoTD-
>TCX_old_syn_Overl, st[n]->L_frame / 2 );
    }
}
```

```

}

void stereo_td2dft_update(
    CPE_DEC_HANDLE hCPE,          /* i/o: CPE decoder structure */
    const int16_t n,              /* i : channel number */
    float output[],               /* i/o: synthesis @internal Fs */
    float synth[],                /* i/o: synthesis @output Fs */
    float hb_synth[],             /* i/o: hb synthesis */
    const int16_t output_frame /* i : frame length */
)
{
    int16_t ovl, ovl_TCX, dft32ms_ovl, hq_delay_comp;
    Decoder_State **st;

    /* initialization */
    st = hCPE->hCoreCoder;
    ovl = NS2SA( st[n]->L_frame * 50, STEREO_DFT32MS_OVL_NS );
    dft32ms_ovl = ( STEREO_DFT32MS_OVL_MAX * st[0]->output_Fs ) / 48000;
    hq_delay_comp = NS2SA( st[0]->output_Fs, DELAY_CLDFB_NS );

    if ( hCPE->element_mode >= IVAS_CPE_DFT && hCPE->element_mode !=
        IVAS_CPE_MDCT )
    {
        if ( st[n]->core == ACELP_CORE )
        {
            if ( n == 0 )
            {
                /* update DFT analysis overlap memory @internal_Fs: core
synthesis */
                mvr2r( output + st[n]->L_frame - ovl, hCPE-
>input_mem_LB[n], ovl );

                /* update DFT analysis overlap memory @internal_Fs: BPF
*/
                if ( st[n]->p_bpf_noise_buf )
                {
                    mvr2r( st[n]->p_bpf_noise_buf + st[n]->L_frame - ovl,
hCPE->input_mem_BPF[n], ovl );
                }

                /* update DFT analysis overlap memory @output_Fs: BWE */
                if ( st[n]->ext1 != -1 || ( st[n]->bws_cnt > 0 && st[n]-
>core == ACELP_CORE ) )
                {
                    mvr2r( hb_synth + output_frame - dft32ms_ovl, hCPE-
>input_mem[n], dft32ms_ovl );
                }
            }
            else
            {
                /* update DFT analysis overlap memory @internal_Fs: core
synthesis, secondary channel */

```

```

        mvr2r( output + st[n]->L_frame - ovl, hCPE->input_mem_LB[n], ovl );
    }
    else /* TCX core */
    {
        /* LB-TCX synthesis */
        mvr2r( output + st[n]->L_frame - ovl, hCPE->input_mem_LB[n],
        ovl );

        /* BPF */
        if ( n == 0 && st[n]->p_bpf_noise_buf )
        {
            mvr2r( st[n]->p_bpf_noise_buf + st[n]->L_frame - ovl,
            hCPE->input_mem_BPF[n], ovl );
        }

        /* TCX synthesis (it was already delayed in TD stereo in
        core_switching_post_dec()) */
        if ( st[n]->hTcxDec != NULL )
        {
            ovl_TCX = NS2SA( st[n]->hTcxDec->L_frameTCX * 50,
            STEREO_DFT32MS_OVL_NS );
            mvr2r( synth + st[n]->hTcxDec->L_frameTCX + hq_delay_comp
            - ovl_TCX, hCPE->input_mem[n], ovl_TCX - hq_delay_comp );
            mvr2r( st[n]->delay_buf_out, hCPE->input_mem[n] + ovl_TCX
            - hq_delay_comp, hq_delay_comp );
        }
    }
    else if ( hCPE->element_mode == IVAS_CPE_MDCT && hCPE->input_mem[0]
    != NULL )
    {
        /* reset DFT stereo OLA memories */
        set_zero( hCPE->input_mem[n], NS2SA( st[0]->output_Fs,
        STEREO_DFT32MS_OVL_NS ); );
        set_zero( hCPE->input_mem_LB[n], STEREO_DFT32MS_OVL_16k );
        if ( n == 0 )
        {
            set_zero( hCPE->input_mem_BPF[n], STEREO_DFT32MS_OVL_16k );
        }
    }

    return;
}

```

2.2.5 Decodificador estéreo DFT 801 y método de decodificación 851

- 5 El método de decodificación DFT 851 comprende una operación 857 de decodificación del núcleo del canal medio m. Para realizar la operación 857, un decodificador del núcleo 807 decodifica el canal medio m en el dominio de tiempo en respuesta al flujo de bits recibido 830. El decodificador del núcleo 807 (que realiza la operación de decodificación de núcleo 857) en el decodificador estéreo DFT 801 puede ser cualquier códec mono con velocidad de bits variable. En la implementación ilustrativa de la presente divulgación, se usa el códec EVS (véase la Referencia [1]), con una capacidad de velocidad de bits fluctuante (véase la Referencia [5]). Por supuesto, se pueden considerar e implementar otros códecs adecuados.

En una operación de cálculo DFT 854 del método de decodificación DFT 851 (análisis DFT de la Tabla III), una calculadora 804 calcula el DFT del canal medio m para recuperar el canal medio M en el dominio DFT.

- 15 El método de decodificación DFT 851 comprende, además, una operación 858 de decodificación de información del lado estéreo y de la señal residual S (decodificación residual de la Tabla III). Para realizar la operación 858, un decodificador 808 responde al flujo de bits 830 para recuperar la información del lado estéreo y la señal residual S.

En una operación de decodificación estéreo DFT (decodificación estéreo DFT de la Tabla III) y mezcla (mezcla en el dominio DFT de la Tabla III) 859, un decodificador estéreo DFT y un mezclador 809 produce los canales L y R en el dominio DFT en respuesta al canal medio M y a la información lateral y a la señal residual S. Generalmente, la operación de decodificación estéreo DFT y mezcla 859 es la inversa a la operación de procesamiento estéreo DFT y mezcla 353 de la Figura 3.

En la operación de cálculo de la IDFT 855 (síntesis de DFT de la Tabla III), una calculadora 805 calcula la IDFT del canal L para recuperar el canal l en el dominio de tiempo. De manera similar, en la operación de cálculo de la IDFT 856 (síntesis de DFT de la Tabla III), una calculadora 806 calcula la IDFT del canal R para recuperar el canal r en el dominio de tiempo.

2.2.6 Decodificador estéreo TD 802 y método de decodificación 852

El método de decodificación TD 852 comprende una operación 860 de decodificación del núcleo del canal primario PCh. Para realizar la operación 860, un decodificador del núcleo 810 decodifica el canal primario PCh en respuesta al flujo de bits recibido 830.

El método de decodificación TD 852 también comprende una operación 861 de decodificación del núcleo del canal secundario SCh. Para realizar la operación 861, un decodificador del núcleo 811 decodifica el canal secundario SCh en respuesta al flujo de bits recibido 830.

De nuevo, el decodificador del núcleo 810 (que realiza la operación de decodificación del núcleo 860 en el decodificador estéreo TD 802) y el decodificador del núcleo 811 (que realiza la operación de decodificación de núcleo 861 en el decodificador estéreo TD 802) pueden ser cualquier códec mono con velocidad de bits variable. En la implementación ilustrativa de la presente divulgación, se usa el códec EVS (véase la Referencia [1]), con una capacidad de velocidad de bits fluctuante (véase la Referencia [5]). Por supuesto, se pueden considerar e implementar otros códecs adecuados.

En una operación de mezcla en el dominio de tiempo (TD) 862 (mezcla en el dominio TD de la Tabla III), un mezclador 812 recibe y mezcla los canales primario PCh y secundario SCh para recuperar los canales de dominio de tiempo l y r de la señal estéreo basándose en el factor de mezcla estéreo de TD.

2.2.7 Decodificador estéreo MDCT 803 y método de decodificación 853

El método de decodificación MDCT 853 comprende una operación 863 de decodificación del núcleo conjunta (decodificación estéreo conjunta de la Tabla III) en el canal izquierdo l y el canal derecho r. Para realizar la operación 863, un decodificador del núcleo conjunto 813 decodifica el canal izquierdo l y el canal derecho r en respuesta al flujo de bits recibido 830. Cabe señalar que no se realiza ninguna operación de mezcla y que no se emplea ningún mezclador en el modo estéreo MDCT.

2.2.8 Sincronización de síntesis

Para realizar una sincronización de tiempo de síntesis estéreo (sincronización de síntesis de la Tabla III) y una operación de cambio de estéreo 864, el controlador de cambio de modo estéreo (no se muestra) comprende un sincronizador de tiempo y un conmutador estéreo 814 para recibir los canales l y r del decodificador estéreo DFT 801, el decodificador estéreo TD 802 o el decodificador estéreo MDCT 803 y para sincronizar los canales estéreo de salida mezclados l y r. El sincronizador de tiempo y el conmutador estéreo 814 retrasan los canales estéreo de salida mezclados l y r para que coincidan con el valor de retardo general del códec y maneja las transiciones entre los canales de salida estéreo DFT, los canales de salida estéreo TD y los canales de salida estéreo MDCT.

Por defecto, en el modo estéreo DFT, el sincronizador de tiempo y el conmutador estéreo 814 introducen un retardo de 3,125 ms en el decodificador estéreo DFT 801. Para adaptarse al retardo total del códec de 32 ms (longitud de la trama de 20 ms, retardo del codificador de 8,75 ms, retardo del decodificador de 3,25 ms), el sincronizador de tiempo y el conmutador estéreo 814 aplican una sincronización de retardo de 0,125 ms. En el caso del modo estéreo TD o MDCT, el sincronizador de tiempo y el conmutador estéreo 814 aplican un retardo que consiste en el retardo de remuestreo de 1:25 ms y el retardo de 2 ms usado para la sincronización entre la síntesis LB y HB y para coincidir con el retardo general del códec de 32 ms.

Después de realizar la sincronización de tiempo y el cambio de estéreo (véase la operación de sincronización de tiempo de síntesis y cambio de estéreo 864 y el sincronizador de tiempo y conmutador estéreo 814 de la Figura 8), la síntesis de HB (de BWE o IC-BWE) se añade a la síntesis del núcleo (IC-BWE, adición de síntesis de HB de la Tabla III; véase también en la Figura 8 la operación de cálculo BWE o IC-BWE 865 y la calculadora BWE o IC-BWE 815) y la decodificación ICA (decodificador ICA - ajuste temporal de la Tabla III que desincroniza dos canales de salida l y r) se realiza antes de que la síntesis estéreo final de los canales l y r se emita desde el dispositivo de decodificación estéreo IVAS 800 (véase la operación ICA temporal 866 y el decodificador ICA correspondiente 816). Estas operaciones 865 y 866 se omiten en el modo estéreo MDCT.

Finalmente, como se muestra en la Tabla III, se realizan actualizaciones estéreo comunes.

2.3 Cambio del modo estéreo TD al modo estéreo DFT en el decodificador estéreo IVAS

Se puede encontrar más información sobre los elementos, las operaciones y las señales mencionadas en las secciones 2.3 y 2.4, por ejemplo, en las Referencias [1] y [2].

5 El mecanismo de cambiar del modo estéreo TD al modo estéreo DFT en el dispositivo de decodificación estéreo IVAS 800 se complica por el hecho de que las etapas de decodificación entre estos dos modos estéreo son fundamentalmente diferentes (véase la Sección 2.1 anterior para más detalles), incluyendo una transición de dos decodificadores del núcleo 810 y 811 en la última trama estéreo TD a un decodificador del núcleo 807 en la primera trama estéreo DFT.

10 La Figura 9 es un diagrama de flujo que ilustra las operaciones de procesamiento en el dispositivo de decodificación estéreo IVAS 800 y el método 850 al cambiar del modo estéreo TD al modo estéreo DFT. En concreto, la Figura 9 muestra dos tramas de la señal estéreo decodificada en diferentes operaciones de procesamiento con instancias de tiempo relacionadas al cambiar de una trama estéreo TD 901 a una trama estéreo DFT 902.

15 Primero, los decodificadores del núcleo 810 y 811 del decodificador estéreo TD 802 se usan tanto para los canales primario PCh y secundario SCh, y cada uno emite la síntesis de núcleo decodificada correspondiente a la frecuencia de muestreo interna. En la trama estéreo TD 901, la síntesis de núcleo decodificada de los dos decodificadores del núcleo 810 y 811 se usa para actualizar los búferes de memoria OLA de DFT estéreo (un búfer de memoria por canal, es decir, dos búferes de memoria OLA en total; véanse los análisis OLA DFT y las memorias de síntesis descritas anteriormente). Estos búferes de memoria OLA se actualizan en cada trama

20 estéreo TD para estar actualizadas en caso de que la siguiente trama sea una trama estéreo DFT.

El caso A) de la Figura 9 se refiere, al recibir una primera trama estéreo DFT 902 después de una trama estéreo TD 901, a una operación (no se muestra) de actualización de memorias de análisis estéreo DFT (estas se usan en la parte OLA de la división en ventanas en la trama anterior y actual antes de la operación de cálculo DFT 854) a la frecuencia de muestreo interna, *input_mem_LB[]*, usando el controlador de cambio de modo estéreo (no se muestra). Para ello, el controlador de cambio de modo estéreo (no se muestra) utiliza un número L_{ovl} de las últimas muestras 903 de la síntesis estéreo TD a la frecuencia de muestreo interna del canal primario PCh y el canal secundario SCh en la trama estéreo TD 901 para actualizar las memorias de análisis estéreo DFT del canal medio estéreo DFT m y el canal lateral s. La longitud del segmento de superposición 903, L_{ovl} , corresponde a la parte de superposición de 3,125 ms de longitud de la ventana de análisis DFT 905, por ejemplo, $L_{ovl} = 40$ muestras a una frecuencia de muestreo interna de 12,8 kHz.

25

30

De forma similar, el controlador de cambio de modo estéreo (no se muestra) actualiza la memoria de análisis del filtro posterior de bajos (BPF) DFT estéreo (que se usa en la parte OLA de la división en ventanas en la trama anterior y actual antes de la operación de cálculo de DFT 854) del canal medio m a la frecuencia de muestreo interna, *input_mem_BPF[]*, usando las últimas muestras L_{ovl} de la señal de error BPF (véase la Referencia [1], Cláusula 6.1.4.2) del canal primario TD PCh. Además, la memoria de análisis de banda completa (FB) DFT estéreo (esta memoria se usa en la parte OLA de la división en ventanas en la trama anterior y actual antes de la operación de cálculo DFT 854) del canal medio m a la frecuencia de muestreo de la señal estéreo de salida, *input_mem[]*, se actualiza usando las últimas muestras de 3,125 ms de la síntesis de HB PCh TD estéreo (núcleo ACELP) o bien la síntesis TCX PCh. Las memorias de análisis estéreo BPF y FB de DFT

35

40

estéreo no se emplean para el canal de información lateral s, por lo que no se actualizan utilizando la síntesis del núcleo del canal secundario SCh.

A continuación, en la trama estéreo TD 901, la síntesis del núcleo ACELP decodificada (canales primario PCh y secundario SCh) a la frecuencia de muestreo interna se vuelve a muestrear usando el filtrado del dominio CLDFB, que introduce un retardo de 1:25 ms. En el caso de la trama del núcleo TCX/HQ, se utiliza un retardo de compensación de 1,25 ms para sincronizar la síntesis del núcleo entre diferentes núcleos. A continuación, se aplica el filtro posterior TCX-LTP a los dos canales del núcleo PCh y SCh.

45

En la siguiente operación, los canales primario PCh y secundario SCh de la síntesis estéreo TD a la frecuencia de muestreo de la señal estéreo de salida de la trama estéreo TD 901 están sujetos a la mezcla estéreo TD (combinación de los canales primario PCh y secundario SCh usando la relación de mezcla estéreo TD en el mezclador TD 812 (véase la Referencia [4]), lo que da lugar a los canales estéreo mezclados l y r en el dominio temporal. Dado que la operación de mezcla 862 se realiza en el dominio de tiempo, no introduce ningún retardo de mezcla.

50

A continuación, los canales izquierdo l y derecho r mezclados de la trama estéreo TD 901 desde el mezclador 812 del decodificador estéreo TD 802 se usan en una operación (no se muestra) de actualización de las memorias de síntesis estéreo DFT (estas se usan en la parte OLA de la división en ventanas de la trama anterior y actual después de la operación de cálculo IDFT 855). De nuevo, esta actualización la realiza en cada

55

trama estéreo TD el controlador de cambio de modo estéreo (no se muestra) en caso de que la siguiente trama sea una trama estéreo DFT. El caso B) de la Figura 9 ilustra que el número de las últimas muestras disponibles de la síntesis de canales TD estéreo izquierdo l y derecho r es insuficiente para usarse para una actualización directa de las memorias de síntesis estéreo DFT. Las memorias de síntesis estéreo DFT de 3,125 ms de longitud se reconstruyen así en dos segmentos mediante aproximaciones. El primer segmento corresponde a la señal de 3,125-1,25 ms de longitud que está disponible (es decir, la síntesis mezclada a la frecuencia de muestreo de la señal estéreo de salida), mientras que el segundo segmento corresponde a la señal restante de 1,25 ms de longitud, que no está disponible debido al retraso de muestreo del decodificador del núcleo.

En concreto, el controlador de cambio de modo estéreo (no se muestra) actualiza las memorias de síntesis estéreo DFT mediante las suboperaciones que se muestran en la Figura 10. La Figura 10 es un diagrama de flujo que ilustra el caso B) de la Figura 9, que comprende la actualización de memorias de síntesis estéreo DFT en una trama estéreo TD en el lado del decodificador:

(a) Los dos canales l y r de las memorias de análisis estéreo DFT a la frecuencia de muestreo interna, *input_mem_LB[]*, como se reconstruyeron anteriormente durante el método de decodificación 850 (son idénticos a la síntesis del núcleo a la frecuencia de muestreo interna), están sujetos a procesamiento adicional dependiendo del núcleo de decodificación real:

- Núcleo ACELP: las últimas muestras L_{ovl} 1001 de la síntesis del núcleo LB de los canales primario PCh y secundario SCh a la frecuencia de muestreo interna se vuelven a muestrear a la frecuencia de muestreo de la señal estéreo de salida usando una interpolación lineal simple con retardo cero (véase 1003).

- TCX/HQ core: las últimas muestras L_{ovl} 1001 de la síntesis del núcleo LB de los canales primario PCh y secundario SCh a la frecuencia de muestreo interna se vuelven a muestrear de manera similar a la frecuencia de muestreo de la señal estéreo de salida usando una interpolación lineal simple con retardo cero (véase 1003). Sin embargo, en este caso, la memoria de sincronización de TCX (el último segmento de 1,25 ms de la síntesis de TCX de la trama anterior) se utiliza para actualizar los últimos 1,25 ms de la síntesis del núcleo remuestreado.

(b) Las señales LB remuestreadas linealmente correspondientes a la parte de 3,125 ms de longitud de los canales primario PCh y secundario SCh de la trama estéreo TD 901 se mezclan (véase 1003) para formar los canales izquierdo l y derecho r usando la rutina común de mezcla estéreo TD mientras se usa la relación de mezcla estéreo TD de la trama actual (véase la operación de mezcla TD 862). La señal resultante se denomina, además, «*síntesis reconstruida*» 1002.

(c) La reconstrucción de la primera parte de (3,125 - 1,25 ms) de las memorias de síntesis estéreo DFT depende del núcleo de decodificación real:

- Núcleo ACELP: Se realiza un fundido cruzado 1004 entre la síntesis 1005 remuestreada basada en CLDFB y mezclada con TD a la frecuencia de muestreo de la señal estéreo de salida y la síntesis 1002 reconstruida (de la suboperación anterior (b)) para ambos canales l y r durante la primera parte de (3,125 - 1,25) ms de los canales de la trama estéreo TD 901.

- Núcleo TCX/HQ: La primera parte de (3,125 - 1,25) ms de las memorias de síntesis estéreo DFT se actualiza usando la síntesis mixta 1005.

(d) La última parte de 1,25 ms de longitud de las memorias de síntesis estéreo DFT se llena con la última porción de la síntesis reconstruida 1002.

(e) La ventana de síntesis DFT (904 en la Figura 9) se aplica a las memorias de síntesis OLA DFT (definidas anteriormente en la presente divulgación) solo en la primera trama estéreo DFT 902 (si se produce el cambio de modo estéreo TD al DFT). Cabe señalar que la última parte de 1,25 ms de las memorias de síntesis OLA DFT tiene una importancia limitada, ya que la forma de ventana de síntesis de DFT 904 converge a cero y, por lo tanto, enmascara las muestras aproximadas de la síntesis reconstruida 1002 resultantes del remuestreo basado en una interpolación lineal simple.

Finalmente, la síntesis reconstruida mezclada 1002 de la trama estéreo TD 901 se alinea, es decir, se retrasa 2 ms en el sincronizador de tiempo y conmutador estéreo 814 para coincidir con el retardo general del códec. En concreto, esto significa lo siguiente:

- En caso de que se produzca un cambio de una trama estéreo TD a una trama estéreo DFT, el controlador de cambio de modo estéreo (no se muestra) restablece otras memorias estéreo DFT (que no sean memorias superpuestas), es decir, los parámetros y búferes del decodificador estéreo DFT anteriores del marco.

- A continuación tiene lugar la decodificación estéreo DFT (véase 859), la mezcla (véase 859) y la síntesis DFT (véase 855 y 856), y la síntesis de salida estéreo (canales l y r) se alinea, es decir, se retrasa 0,125 ms en el sincronizador de tiempo y conmutador estéreo 814 para coincidir con el retardo general del códec.

La Figura 11 es un diagrama de flujo que ilustra un caso C) de la Figura 9, que comprende el suavizado de la síntesis estéreo de salida en la primera trama estéreo DFT 902 después del cambio de modo estéreo, en el lado del decodificador.

Con referencia a la Figura 11, una vez que la síntesis estéreo DFT está alineada y sincronizada con el retardo general del códec en la primera trama estéreo DFT 902, el controlador de cambio de modo estéreo (no se muestra) realiza una operación de fundido cruzado 1151 entre la síntesis estéreo TD alineada y sincronizada 1101 (desde la operación 864) y la síntesis estéreo DFT alineada y sincronizada 1102 (desde la operación 864) para suavizar la transición del cambio. El fundido cruzado se realiza en un segmento 1103 de 1,875 ms de longitud comenzando después de un retardo 1104 de 0,125 ms al principio de ambos canales de salida l y r (todas las señales están a la frecuencia de muestreo de la señal estéreo de salida). Este caso se corresponde con el caso C) de la Figura 9.

A continuación, la decodificación continúa independientemente del modo estéreo actual con la calculadora IC-BWE 815, el decodificador ICA 816 y las actualizaciones comunes del decodificador estéreo.

2.4 Cambio del modo estéreo DFT al modo estéreo TD en el decodificador estéreo IVAS

Las operaciones de decodificación fundamentalmente diferentes entre el modo estéreo DFT y el modo estéreo TD y la presencia de dos decodificadores del núcleo 810 y 811 en el decodificador estéreo TD 802 hacen difícil el cambio de modo estéreo DFT al modo estéreo TD en el dispositivo de decodificación estéreo IVAS 800. La Figura 12 es un diagrama de flujo que ilustra las operaciones de procesamiento en el dispositivo de decodificación estéreo IVAS 800 y el método 850 al cambiar del modo estéreo DFT al modo estéreo TD. En concreto, la Figura 12 muestra dos tramas de señal estéreo decodificada en diferentes operaciones de procesamiento con instancias de tiempo relacionadas al cambiar de una trama estéreo DFT 1201 a una trama estéreo TD 1202.

La decodificación del núcleo puede utilizar el mismo procesamiento independientemente del modo estéreo real, con dos excepciones.

Primera excepción: En las tramas estéreo DFT, el remuestreo de la frecuencia de muestreo interna a la frecuencia de muestreo de la señal estéreo de salida se realiza en el dominio DFT, pero el remuestreo CLDFB se ejecuta en paralelo para mantener/actualizar el análisis CLDFB y las memorias de síntesis en caso de que la siguiente trama sea una trama estéreo TD.

Segunda excepción: A continuación, se aplica el BPF (*Bass Post-Filter*, un procedimiento de mejora del tono de baja frecuencia; véase la Referencia [1], Cláusula 6.1.4.2) en el dominio DFT en tramas estéreo DFT, mientras que el análisis BPF y el cálculo de la señal de error se realizan en el dominio de tiempo independientemente del modo estéreo.

Por lo demás, todos los estados internos y las memorias del decodificador del núcleo son simplemente continuos y se mantienen bien al cambiar del canal medio DFT m al canal primario TD PCh.

En la trama estéreo DFT 1201, la decodificación continúa después con la decodificación del núcleo (857) del canal medio m, el cálculo (854) de la transformada DFT del canal medio m en el dominio de tiempo para obtener el canal medio M en el dominio DFT y la decodificación estéreo y la mezcla (859) de los canales M y S en los canales L y R en el dominio DFT, que incluye la decodificación (858) de la señal residual. El análisis y la síntesis del dominio DFT introducen un retraso de OLA de 3,125 ms. Las transiciones de síntesis se manejan después en el sincronizador de tiempo y conmutador estéreo 814.

Al cambiar de la trama estéreo DFT 1201 a la trama estéreo TD 1202, el hecho de que solo haya un decodificador del núcleo 807 en el decodificador estéreo DFT 801 complica la decodificación del núcleo del canal secundario TD SCh, ya que los estados internos y las memorias del segundo decodificador del núcleo 811 del decodificador estéreo TD 802 no se mantienen continuamente (al contrario, los estados y memorias internos del primer decodificador del núcleo 810 se mantienen continuamente usando los estados internos y las memorias del decodificador del núcleo 807 del decodificador estéreo DFT 801). Por lo tanto, el controlador de cambio de modo estéreo (no se muestra) suele restablecer las memorias del segundo decodificador del núcleo 811 en las actualizaciones de cambio de modo estéreo (véase la Tabla III). Sin embargo, hay algunas excepciones en las que la memoria del canal primario SCh se llena con la memoria de ciertos búferes PCh, por ejemplo, la excitación anterior, los parámetros LSF anteriores y los parámetros LSP anteriores. En cualquier caso, la síntesis al inicio de la primera trama del canal secundario SCh TD después de cambiar de la trama estéreo DFT 1201 a la trama estéreo TD 1202 sufre, en consecuencia, una reconstrucción imperfecta. Por consiguiente, mientras que la síntesis del primer decodificador del núcleo 810 se decodifica bien y suavemente

durante el cambio de modo estéreo, la síntesis de calidad limitada del segundo decodificador del núcleo 811 introduce discontinuidades durante la mezcla de estéreo y la síntesis final (862). Estas discontinuidades se suprimen usando las memorias DFT estéreo OLA durante la primera reconstrucción de síntesis de salida estéreo TD que se describe más adelante.

- 5 El controlador de cambio de modo estéreo (no se muestra) suprime posibles discontinuidades y diferencias entre los canales mezclados DFT estéreo y TD estéreo mediante una simple ecualización de la energía de la señal. Si la ganancia objetivo ICA, g_{ICA} , es inferior a 1,0, el canal l, $y_L(i)$, después de la mezcla (862) y antes de que la sincronización de tiempo (864), se altera en la primera trama estéreo TD 1202 después del cambio de modo estéreo usando la siguiente relación:

$$y'_L(i) = \alpha \cdot y_L(i) \quad \text{for } i = 0, \dots, L_{eq} - 1$$

10 donde L_{eq} es la longitud de las señales a igualar, que se corresponde en el dispositivo de decodificación estéreo IVAS 800 con un segmento de 8,75 ms de longitud (que corresponde, por ejemplo, a $L_{eq} = 140$ muestras a una frecuencia de muestreo de señal estéreo de salida de 16 kHz). A continuación, el valor del factor de ganancia α se obtiene utilizando la siguiente relación:

$$\alpha = g_{ICA} + i \cdot \frac{1 - g_{ICA}}{L_{eq}} \quad \text{for } i = 0, \dots, L_{eq} - 1.$$

15 Con referencia a la Figura 12, el caso A) se relaciona con una parte faltante 1203 de la síntesis sincronizada y mezclada de estéreo TD (de la operación 864) de la trama estéreo TD 1202 correspondiente a una memoria de síntesis sincronizada combinada de estéreo DFT anterior de la trama estéreo DFT 1201. Esta memoria de longitud de (3,25 - 1,25) ms no está disponible al cambiar de la trama estéreo DFT 1201 a la trama estéreo TD 1202, excepto para su primer segmento de 0,125 ms de longitud 1204.

La Figura 13 es un diagrama de flujo que ilustra el caso A) de la Figura 12, que comprende la actualización de la memoria de síntesis de sincronización estéreo mezclada TD en una primera trama estéreo TD después de cambiar del modo estéreo DFT al modo estéreo TD, en el lado del decodificador.

25 Con referencia a ambas Figuras 12 y 13, el controlador de cambio de modo estéreo (no se muestra) reconstruye los 3,25 ms 1205 de la síntesis sincronizada y mezclada TD estéreo usando las siguientes operaciones (a) a (e) para los canales izquierdo l y derecho r:

(a) Las memorias de síntesis de OLA DFT estéreo (definidas anteriormente en la presente divulgación) se rectifican (es decir, la ventana de síntesis inversa se aplica a las memorias de síntesis de OLA; véase 1301).

30 (b) La primera parte 1302 de 0,125 ms (véase 1204 en la Figura 12) de la síntesis sincronizada combinada estéreo TD 1303 es idéntica a la memoria de síntesis sincronizada combinada estéreo DFT 1304 anterior (último segmento de 0,125 ms de longitud de la memoria de síntesis sincronizada combinada estéreo DFT anterior) y, por lo tanto, se reutiliza para formar esta primera parte de la síntesis sincronizada y mezclada estéreo TD 1303.

35 (c) La segunda parte (véase 1203 en la Figura 12) de la síntesis sincronizada y mezclada de TD estéreo 1303, que tiene una longitud de (3,125 - 1,25) ms, se aproxima a las memorias de síntesis de DFT estéreo OLA rectificadas 1301.

(d) La parte de la síntesis sincronizada y mezclada estéreo TD 1303 con una longitud de 2 ms de las dos etapas anteriores (b) y (c) se rellena después con la síntesis estéreo de salida en la primera trama estéreo TD 1202.

40 (e) Al principio de la síntesis mixta estéreo sincronizada TD 1305, se realiza un suavizado de la transición entre la memoria de síntesis estéreo OLA DFT 1301 anterior y la síntesis mezclada y sincronizada TD 1305 de la operación 864 de la trama estéreo TD 1202 actual. El segmento de transición es de 1,25 ms de longitud (véase 1306) y se obtiene usando un fundido cruzado 1307 entre la memoria de síntesis estéreo OLA DFT rectificada 1301 y la síntesis mezclada y sincronizada TD estéreo 1305.

2.5 Cambio del modo estéreo TD al modo estéreo MDCT en el decodificador estéreo IVAS

45 Cambiar del modo estéreo TD al modo estéreo MDCT es relativamente sencillo porque ambos modos estéreo manejan dos canales de transporte y emplean dos instancias de codificador del núcleo.

Como se empleó un esquema de mezcla de fase opuesta en el codificador estéreo TD 400, el controlador de cambio de modo estéreo (no se muestra) altera de forma similar la mezcla del canal estéreo TD para mantener la fase correcta de los canales izquierdo y derecho de la señal de sonido estéreo en la última trama estéreo TD antes de la primera trama estéreo MDCT. En concreto, el controlador de cambio de modo estéreo (no se muestra) establece la relación de mezcla $\beta = 1,0$ e implementa una mezcla de fase opuesta (inversa a la mezcla de fase opuesta empleada en el codificador estéreo TD 400) del canal primario estéreo TD $PCh(i)$ y el canal secundario estéreo TD $SCh(i)$ para calcular el anterior canal izquierdo MDCT estéreo $l_{past(i)}$ y el anterior canal derecho MDCT estéreo $r_{past(i)}$. Por consiguiente, el canal primario TD estéreo $PCh(i)$ es idéntico al anterior canal izquierdo MDCT estéreo $l_{past(i)}$, y la señal del canal secundario TD estéreo $SCh(i)$ es idéntica a la del anterior canal derecho MDCT estéreo $r_{past(i)}$.

2.6 Cambio del modo estéreo MDCT al modo estéreo TD en el decodificador estéreo IVAS

De forma similar al cambio de modo estéreo TD al modo estéreo MDCT, hay dos canales de transporte disponibles y se emplean dos instancias de codificador del núcleo en este escenario. Para mantener la fase correcta de los canales izquierdo y derecho de la señal de sonido estéreo, la relación de mezcla estéreo TD se establece en 1,0, y el controlador de cambio de modo estéreo (no se muestra) utiliza de nuevo el esquema de mezcla de fase opuesta en la primera trama estéreo TD después de la última trama estéreo MDCT.

2.7 Cambio del modo estéreo DFT al modo estéreo MDCT en el decodificador estéreo IVAS

En este escenario se usa un mecanismo similar al cambio de lado del decodificador del modo estéreo DFT al modo estéreo TD, donde los canales primario PCh y secundario SCh del modo estéreo TD se sustituyen por los canales izquierdo l y derecho r del modo estéreo MDCT.

2.8 Cambio del modo estéreo MDCT al modo estéreo DFT en el decodificador estéreo IVAS

En este escenario se usa un mecanismo similar al cambio de lado del decodificador del modo estéreo TD al modo estéreo DFT, donde los canales primario PCh y secundario SCh del modo estéreo TD se sustituyen por los canales izquierdo l y derecho r del modo estéreo MDCT.

Finalmente, la decodificación continúa independientemente del modo estéreo actual con la decodificación IC-BWE 865 (omitida en el modo estéreo MDCT), la adición de la síntesis HB (omitida en el modo estéreo MDCT), la alineación ICA temporal 866 (omitida en el modo estéreo MDCT) y las actualizaciones comunes del decodificador estéreo.

2.9 Implementación del hardware

La Figura 14 es un diagrama de bloques simplificado de un ejemplo de configuración de componentes de hardware que forman cada uno de los dispositivos de codificación estéreo IVAS 200 y de decodificación estéreo IVAS 800 descritos anteriormente.

Cada uno de los dispositivos de codificación estéreo IVAS 200 y decodificación estéreo IVAS 800 puede implementarse como parte de un terminal móvil, como parte de un reproductor de medios portátil o en cualquier dispositivo similar. Cada uno de los dispositivos de codificación estéreo IVAS 200 y de decodificación estéreo IVAS 800 (identificado como 1400 en la Figura 14) comprende una entrada 1402, una salida 1404, un procesador 1406 y una memoria 1408.

La entrada 1402 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 dispositivo codificador estéreo IVAS 200, o el flujo de bits 803 en el caso del dispositivo decodificador estéreo IVAS 800. La salida 1404 está configurada para suministrar el flujo de bits multiplexado 206 en el caso del dispositivo codificador estéreo IVAS 200 o el canal izquierdo l y el canal derecho r descodificados en el caso del dispositivo codificador estéreo IVAS 800. La entrada 1402 y la salida 1404 pueden implementarse en un módulo común, por ejemplo, un dispositivo de entrada/salida en serie.

El procesador 1406 se conecta operativamente a la entrada 1402, a la salida 1404 y a la memoria 1408. El procesador 1406 consiste en uno o más procesadores para ejecutar instrucciones de código en apoyo de las funciones de los diversos elementos y operaciones del dispositivo codificador estéreo IVAS 200 descrito anteriormente, el método de codificación estéreo IVAS 250, el dispositivo de decodificación estéreo IVAS 800 y el método de decodificación estéreo IVAS 850, tal y como se muestra en las figuras adjuntas o como se describe en la presente divulgación.

La memoria 1408 puede comprender una memoria no transitoria para almacenar instrucciones de código ejecutables por el procesador 1406, específicamente, una memoria legible por procesador que almacena instrucciones no transitorias que, cuando se ejecutan, hacen que un procesador implemente los elementos y operaciones del dispositivo codificador estéreo IVAS 200, el método de codificación estéreo IVAS 250, el dispositivo de decodificación estéreo IVAS 800 y el método de decodificación estéreo IVAS 850. La memoria

1408 también puede incluir una memoria de acceso aleatorio o búfer(es) para almacenar datos de procesamiento intermedios de las diversas funciones realizadas por el procesador 1406.

Los expertos en la materia se darán cuenta de que las descripciones del dispositivo de codificación estéreo IVAS 200, el método de codificación estéreo IVAS 250, el dispositivo de decodificación estéreo IVAS 800 y el método de decodificación estéreo IVAS 850 tienen un carácter meramente ilustrativo y no pretenden ser limitativas en modo alguno. A aquellas personas con conocimientos ordinarios en la materia que se beneficien de la presente divulgación se les ocurrirán fácilmente otras modalidades. Además, el dispositivo codificador estéreo IVAS 200 descrito, el método de codificación estéreo IVAS 250, el dispositivo de decodificación estéreo IVAS 800 y el método de decodificación estéreo IVAS 850 pueden personalizarse para ofrecer soluciones valiosas a las necesidades y problemas existentes de codificación y decodificación de sonido estéreo.

Para mayor claridad, no se muestran ni describen todas las características rutinarias de las implementaciones del dispositivo de codificación estéreo IVAS 200, el método de codificación estéreo IVAS 250, el dispositivo de decodificación estéreo IVAS 800 y el método de decodificación estéreo IVAS 850. Por supuesto, se tendrá en cuenta que, al desarrollar cualquier implementación real del dispositivo de codificación estéreo IVAS 200, el método de codificación estéreo IVAS 250, dispositivo de decodificación estéreo IVAS 800 y el método de decodificación estéreo IVAS 850, es posible que se deban tomar numerosas decisiones específicas de implementación para lograr los objetivos específicos del desarrollador, como el cumplimiento de las normas 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.

También se tendrá en cuenta que un esfuerzo de desarrollo puede ser complejo y lento, pero, sin embargo, sería una tarea rutinaria de ingeniería para aquellos con conocimientos ordinarios en el campo del procesamiento de sonido que se beneficien de la presente divulgación.

De conformidad con la presente divulgación, los elementos, las operaciones de procesamiento o las estructuras de datos que aquí se describen pueden implementarse utilizando varios tipos de sistemas operativos, plataformas informáticas, dispositivos de red, programas informáticos o máquinas de uso general. Además, los expertos en la materia reconocerán que también pueden usarse dispositivos de naturaleza menos general, como dispositivos cableados, matrices de puertas lógicas programables en campo (FPGA), circuitos integrados de aplicación específica (ASIC) o similares. Cuando un procesador, ordenador o máquina implementa un método que comprende una serie de operaciones y suboperaciones y estas pueden guardarse como una serie de instrucciones de código no transitorio legibles por el procesador, el ordenador o la máquina, pueden almacenarse en un medio tangible o no transitorio.

Los elementos y las operaciones de procesamiento del dispositivo de codificación estéreo IVAS 200, el método de codificación estéreo IVAS 250, el dispositivo de decodificación estéreo IVAS 800 y el método de decodificación estéreo IVAS 850, tal y como se describen en la presente divulgación, pueden incluir software, firmware, hardware o cualquier combinación de software, firmware o hardware adecuada para los fines descritos en este documento.

En el método de codificación estéreo IVAS 250 y el método de decodificación estéreo IVAS 850, tal y como se describen en la presente divulgación, las diversas operaciones y suboperaciones de procesamiento pueden realizarse en varios órdenes, y algunas de las operaciones y suboperaciones de procesamiento pueden ser opcionales.

La presente divulgación menciona las siguientes referencias:

[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. Multus, N. Rettelbach, G. Fuchs, J. Robillard, J. Lecompte, S. Wilde y S. Bayer, S. Disch, C. Helmrich, R. Lefevre y P. Gournay, et al., «The ISO/MPEG Unified Speech and Audio Coding Standard - Consistent High Quality for All Content Types and at All Bit Rates», *J. Audio Eng. Soc.*, vol. 61, n.º 12, pp. 956-977, Dic. 2013.

[3] F. Baumgarte, C. Faller, «Binaural cue coding - Part I: Psychoacoustic fundamentals and design principles», *IEEE Trans. Speech Audio Processing*, vol. 11, pp. 509-519, Nov. 2003.

[4] T. Vaillancourt, «Method and system using a long-term correlation difference between left and right channels for time domain down mixing a stereo sound signal into primary and secondary channels», PCT Application WO2017/049397A1.

[5] V. Eksler, «Method and Device for Assigning a Bit-Budget between Sub-Frames in a CELP Codec», PCT Application WO2019/056107A1.

- [6] M. Neuendorf et. al., «MPEG Unified Speech and Audio Coding - The ISO/MPEG Standard for High-Efficiency Audio Coding of all Content Types», Journal of the Audio Engineering Society, vol. 61, n.º 12, pp. 956-977, diciembre de 2013.
- 5 [7] J. Herre et al., «MPEG-H Audio - The New Standard for Universal Spatial / 3D Audio Coding», en la 137.ª Convención Internacional AES, Paper 9095, Los Ángeles, 9-12 de octubre de 2014.
- [8] Contribución a 3GPP SA4 S4-180462, «On spatial metadata for IVAS spatial audio input format», reunión SA4 n.º 98, 9-13 de abril de 2018, https://www.3gpp.org/ftp/tsg_sa/WG4_CODEC/TSGS4_98/Docs/S4-180462.zip
- 10 [9] V. Malenovsky, T. Vaillancourt, «Method and Device for Classification of Uncorrelated Stereo Content, Cross-Talk Detection, and Stereo Mode Selection in a Sound Codec», solicitud de patente provisional de EE. UU. 63/075,984 presentada el 9 de septiembre de 2020.

REIVINDICACIONES

1. Un dispositivo para codificar una señal de sonido estéreo, que comprende:

un primer codificador estéreo de la señal de sonido estéreo que utiliza un primer modo estéreo que opera en el dominio de tiempo, TD, en el que el primer modo estéreo TD, en tramas TD de la señal de sonido estéreo, (a) produce una primera señal mezclada y (b) usa las primeras estructuras de datos y memorias;

un segundo codificador estéreo de la señal de sonido estéreo que utiliza un segundo modo estéreo que opera en el dominio de frecuencia, FD, en el que el segundo modo estéreo FD, en tramas FD de la señal de sonido estéreo, (a) produce una segunda señal mezclada y (b) usa las segundas estructuras de datos y memorias;

un controlador configurado para cambiar entre (i) el primer modo estéreo TD y el primer codificador estéreo, y (ii) el segundo modo estéreo FD y el segundo codificador estéreo para codificar la señal de sonido estéreo en el dominio de tiempo o el dominio de frecuencia;

caracterizado porque, al cambiar de uno de los primeros modos estéreo TD y de los segundos modos estéreo FD a otro de los primeros modos estéreo TD y segundos modos estéreo FD, el controlador de cambio de modo estéreo está configurado para recalcular al menos una longitud de señal mezclada en una trama actual de la señal de sonido estéreo, donde la longitud de la señal mezclada recalculada en el primer modo estéreo TD es diferente de la longitud de la señal mezclada recalculada en el segundo modo estéreo FD.
2. Un dispositivo codificador de la señal de sonido estéreo de conformidad con la reivindicación 1, en el que el segundo modo estéreo FD es un modo estéreo de la Transformada Discreta de Fourier, DFT.
3. Un dispositivo codificador de la señal de sonido estéreo de conformidad con la reivindicación 2, en el que, al cambiar del primer modo estéreo TD al segundo modo estéreo DFT, el segundo codificador estéreo está configurado para continuar una operación de codificación de núcleo en una trama estéreo DFT después de una trama estéreo TD con memorias de un codificador del núcleo del canal primario PCh.
4. Un dispositivo codificador de la señal de sonido estéreo de conformidad con la reivindicación 2 o 3, en el que el controlador de cambio de modo estéreo está configurado para usar y actualizar los parámetros relacionados con el estéreo de dicho modo estéreo al cambiar de ese modo estéreo al otro modo estéreo.
5. Un dispositivo codificador de la señal de sonido estéreo de conformidad con la reivindicación 4, en el que el controlador de cambio de modo estéreo está configurado para transferir los parámetros relacionados con el estéreo entre estructuras de datos.
6. Un dispositivo codificador de la señal de sonido estéreo de conformidad con la reivindicación 4 o 5, en el que los parámetros relacionados con el estéreo comprenden una ganancia lateral y un retardo de tiempo intercanal ITD, parámetros del segundo modo estéreo DFT y una ganancia objetivo y retrasos de correlación del primer modo estéreo TD.
7. Un dispositivo codificador de la señal de sonido estéreo de conformidad con cualquiera de las reivindicaciones de 2 a 6, en el que, al cambiar del segundo modo estéreo DFT al primer modo estéreo TD, el controlador de cambio de modo estéreo está configurado para recalcular en una trama TD actual una longitud de la señal mezclada que es más larga en un canal secundario SCh con respecto a una longitud recalculada de la señal mezclada en un canal primario PCh.
8. Un dispositivo codificador de señal de sonido estéreo de conformidad con cualquiera de las reivindicaciones de 2 a 7, en el que, al cambiar del segundo modo estéreo DFT al primer modo estéreo TD, el controlador de cambio de modo estéreo está configurado para hacer un fundido cruzado de un canal primario PCh recalculado y de un canal medio DFT m de un canal estéreo DFT para recalcular un canal primario mezclado PCh en una primera trama TD después de una trama DFT.
9. Un dispositivo codificador de la señal de sonido estéreo de conformidad con cualquiera de las reivindicaciones de 2 a 8, en el que, al cambiar del segundo modo estéreo DFT al primer modo estéreo TD, el controlador de cambio de modo estéreo está configurado para recalcular una memoria ICA de un canal izquierdo l y derecho r correspondiente a una trama DFT que precede a una trama TD.

10. Un dispositivo de codificación de la señal de sonido estéreo de conformidad con la reivindicación 9, en el que el controlador de cambio de modo estéreo está configurado para recalcular los canales primario PCh y secundario SCh de la trama DFT mezclando los canales procesados por ICA l y r y utilizando una relación de mezcla estéreo de la trama DFT.
- 5 11. Un dispositivo codificador de la señal de sonido estéreo de conformidad con la reivindicación 10, en el que el controlador de cambio de modo estéreo está configurado para recalcular una longitud más corta del canal secundario SCh cuando no hay cambio de modo estéreo.
12. Un dispositivo codificador de la señal de sonido estéreo de conformidad con la reivindicación 10 o 11, en el que el controlador de cambio de modo estéreo está configurado para recalcular, en la trama DFT que precede a la trama TD, una primera longitud del canal primario PCh y una segunda longitud del canal secundario SCh, y donde la primera longitud es más corta que la segunda.
- 10 13. Un dispositivo para decodificar una señal de sonido estéreo, que comprende:
- un primer decodificador estéreo de la señal de sonido estéreo usando un primer modo estéreo que opera en el dominio de tiempo, TD, en el que el primer decodificador estéreo, en tramas TD de la señal de sonido estéreo, (a) decodifica una señal mezclada y (b) usa las primeras estructuras de datos y memorias;
- 15 un segundo decodificador estéreo de la señal de sonido estéreo usando un segundo modo estéreo que opera en el dominio de frecuencia, FD, en el que el segundo decodificador estéreo, en tramas FD de la señal de sonido estéreo, (a) decodifica una segunda señal mezclada y (b) usa las segundas estructuras de datos y memorias;
- 20 un controlador para cambiar entre (i) el primer modo estéreo TD y el primer decodificador estéreo y (ii) el segundo modo estéreo FD y el segundo decodificador estéreo;
- caracterizado porque, al cambiar de uno de los primeros modos estéreo TD y de los segundos modos estéreo FD a otro de los primeros modos estéreo TD y segundos modos estéreo FD, el controlador de cambio de modo estéreo está configurado para recalcular al menos una longitud de señal mezclada en una trama actual de la señal de sonido estéreo, donde la longitud de la señal mezclada recalculada en el primer modo estéreo TD es diferente de la longitud de la señal mezclada recalculada en el segundo modo estéreo FD.
- 25 14. Un dispositivo de decodificación de la señal de sonido estéreo de conformidad con la reivindicación 13, en el que el segundo modo estéreo FD es un modo estéreo de la Transformada Discreta de Fourier, DFT.
- 30 15. Un dispositivo de decodificación de la señal de sonido estéreo de conformidad con la reivindicación 14, en el que el controlador de cambio de modo estéreo está configurado para asignar/desasignar estructuras de datos a/desde el primer modo estéreo TD y el segundo modo estéreo DFT, dependiendo del modo estéreo actual, para reducir un impacto de memoria estática al mantener solo aquellas estructuras de datos que se emplean en la trama actual.
- 35 16. Un dispositivo de decodificación de la señal de sonido estéreo de conformidad con la reivindicación 14 o 15, en el que, al recibir una primera trama DFT después de una trama TD, el controlador de cambio de modo estéreo está configurado para restablecer una estructura de datos estéreo DFT.
- 40 17. Un dispositivo de decodificación de la señal de sonido estéreo de conformidad con cualquiera de las reivindicaciones de 14 a 16, en el que, al recibir una primera trama TD después de una trama DFT, el controlador de cambio de modo estéreo está configurado para restablecer una estructura de datos estéreo TD.
- 45 18. Un dispositivo de decodificación de la señal de sonido estéreo de conformidad con cualquiera de las reivindicaciones de 14 a 17, en el que el controlador de cambio de modo estéreo está configurado para actualizar las memorias de síntesis estéreo DFT en cada trama estéreo TD.
- 50 19. Un dispositivo de decodificación de la señal de sonido estéreo de conformidad con la reivindicación 18, en el que, para actualizar las memorias de síntesis estéreo DFT y para un núcleo ACELP, el controlador de cambio de modo estéreo está configurado para reconstruir en cada trama TD una primera parte de las memorias de síntesis estéreo DFT haciendo un fundido cruzado de (a) una síntesis del

canal izquierdo y derecho remuestreada y mezclada en TD basada en CLDFB y (b) una síntesis del canal izquierdo y derecho reconstruida, remuestreada y mezclada.

5 20. Un dispositivo de decodificación de la señal de sonido estéreo de conformidad con cualquiera de las reivindicaciones de 14 a 19, en el que el controlador de cambio de modo estéreo está configurado para reconstruir una síntesis sincronizada mezclada de TD estéreo.

21. Un dispositivo de decodificación de la señal de sonido estéreo de conformidad con la reivindicación 20, en el que el controlador de cambio de modo estéreo está configurado para usar las siguientes operaciones de (a) a (e), tanto para un canal izquierdo como para un canal derecho, para reconstruir la síntesis sincronizada mezclada de TD estéreo:

10 (a) rectificar una memoria de síntesis de OLA DFT estéreo;

(b) reutilizar una memoria de síntesis de sincronización mezclada de DFT estéreo como una primera parte de la síntesis sincronizada mezclada de DT estéreo;

(c) aproximar una segunda parte de la síntesis sincronizada mezclada de TD estéreo usando la memoria de síntesis estéreo OLA DFT rectificada; y

15 (d) suavizar una transición entre la memoria de síntesis de sincronización mezclada de DFT estéreo y una síntesis mezclada sincronizada de TD estéreo al principio de la síntesis combinada sincronizada de TD estéreo haciendo un fundido cruzado de la memoria de síntesis combinada de OLA DFT estéreo rectificada con la síntesis mezclada sincronizada de TD estéreo.

20 22. Un método para codificar una señal de sonido estéreo, que incluye lo siguiente:

proporcionar un primer codificador estéreo de la señal de sonido estéreo usando un primer modo estéreo que opera en el dominio de tiempo, TD, donde el primer modo estéreo TD, en tramas TD de la señal de sonido estéreo, (a) produce una primera señal mezclada y (b) usa las primeras estructuras de datos y memorias;

25 proporcionar un segundo codificador estéreo de la señal de sonido estéreo usando un segundo modo estéreo que opera en el dominio de frecuencia, FD, donde el segundo modo estéreo FD, en tramas FD de la señal de sonido estéreo, (a) produce una segunda señal mezclada y (b) usa las segundas estructuras de datos y memorias;

30 controlar el cambio entre (i) el primer modo estéreo TD y el primer codificador estéreo, y (ii) el segundo modo estéreo FD y el segundo codificador estéreo para codificar la señal de sonido estéreo en el dominio de tiempo o el dominio de frecuencia;

35 caracterizado porque, al cambiar de uno de los primeros modos estéreo TD y de los segundos modos estéreo FD a otro de los primeros modos estéreo TD y de los segundos modos estéreo FD, el control del cambio de modo estéreo comprende recalcular al menos una longitud de la señal mezclada en una trama actual de la señal de sonido estéreo, donde la longitud de la señal mezclada recalculada en el primer modo estéreo TD es diferente de la longitud de señal mezclada recalculada en el segundo modo estéreo FD.

40 23. Un método de codificación de la señal de sonido estéreo de conformidad con la reivindicación 22, en el que el segundo modo estéreo FD es un modo estéreo de la Transformada Discreta de Fourier, DFT.

24. Un método de codificación de la señal de sonido estéreo de conformidad con la reivindicación 23, en el que, al cambiar de uno de los primeros modos estéreo TD y de los segundos modos estéreo DFT a otro de los primeros modos estéreo TD y de los segundos modo estéreo DFT, el control de cambio de modo estéreo comprende mantener la continuidad de al menos una de las siguientes señales:

- 45
- una señal estéreo de entrada que incluye los canales izquierdo y derecho;
 - un canal medio usado en el segundo modo estéreo DFT;
 - un canal primario y un canal secundario usados en el primer modo estéreo TD;
 - una señal mezclada usada en el preprocesamiento; y

- una señal mezclada utilizada en la codificación del núcleo.

25. Un método de codificación de la señal de sonido estéreo de conformidad con la reivindicación 23 o 24, en el que, al cambiar de uno de los primeros modos estéreo TD y de los segundos modos estéreo DFT a otro de los primeros modos estéreo TD y de los segundos modos estéreo DFT, el control de cambio de modo estéreo comprende asignar/desasignar estructuras de datos a/desde el primer modo estéreo TD y el segundo modo estéreo DFT, dependiendo del modo estéreo actual, para reducir el impacto en la memoria al mantener solo aquellas estructuras de datos que se emplean en la trama actual.
26. Un método de codificación de la señal de sonido estéreo de conformidad con la reivindicación 25, en el que, al cambiar del primer modo estéreo TD al segundo modo estéreo DFT, el control de cambio de modo estéreo comprende desasignar estructuras de datos relacionadas con el estéreo TD.
27. Un método de codificación de la señal de sonido estéreo de conformidad con la reivindicación 26, en el que las estructuras de datos relacionadas con el estéreo TD comprenden una estructura de datos estéreo TD o estructuras de datos de un codificador del núcleo del primer codificador estéreo.
28. Un método de codificación de la señal de sonido estéreo de conformidad con cualquiera de las reivindicaciones de 23 a 27, en el que el control de cambio de modo estéreo comprende actualizar una memoria de análisis DFT cada trama estéreo TD almacenando muestras relacionadas con un último período de tiempo de una trama estéreo TD actual.
29. Un método de codificación de la señal de sonido estéreo de conformidad con cualquiera de las reivindicaciones de 23 a 28, en el que el control de cambio de modo estéreo comprende mantener memorias relacionadas con DFT durante las tramas estéreo TD.
30. Un método de codificación de la señal de sonido estéreo de conformidad con cualquiera de las reivindicaciones de 23 a 29, en el que el control del cambio de modo estéreo comprende, al cambiar del primer modo estéreo TD al segundo modo estéreo DFT, la actualización en una trama DFT después de una trama TD de una memoria de síntesis DFT usando memorias estéreo TD correspondientes a un canal primario PCh de la trama TD.
31. Un método de codificación de la señal de sonido estéreo de conformidad con cualquiera de las reivindicaciones de 23 a 30, en el que el control de cambio de modo estéreo comprende mantener una respuesta de impulso finito, FIR, volver a muestrear la memoria de filtro durante las tramas DFT.
32. Un método de codificación de la señal de sonido estéreo de conformidad con la reivindicación 31, en el que el control de cambio de modo estéreo comprende actualizar en cada trama DFT la memoria de filtro de remuestreo FIR usada en un canal primario PCh en el primer codificador estéreo, usando un segmento de un canal medio m antes de un último segmento de la primera longitud del canal medio m en la trama DFT.
33. Un método de codificación de la señal de sonido estéreo de conformidad con la reivindicación 31 o 32, en el que el cambio de control comprende llenar una memoria de filtro de remuestreo FIR usada en un canal secundario SCh en el primer codificador estéreo, de manera diferente a la actualización de la memoria de filtro de remuestreo FIR usada en el canal primario PCh en el primer codificador estéreo.
34. Un método de codificación de la señal de sonido estéreo de conformidad con la reivindicación 33, en el que el control de cambio de modo estéreo comprende actualizar en una trama TD actual la memoria de filtro de remuestreo FIR usada en el canal secundario SCh en el primer codificador estéreo, llenando la memoria de filtro de remuestreo FIR mediante un segmento de un canal medio m en la trama DFT antes de un último segmento de la segunda longitud del canal medio m.
35. Un método de codificación de la señal de sonido estéreo de conformidad con cualquiera de las reivindicaciones de 23 a 34, en el que el control de cambio de modo estéreo comprende almacenar dos valores de una memoria de filtro de preénfasis en cada trama DFT.
36. Un método de codificación de la señal de sonido estéreo de conformidad con cualquiera de las reivindicaciones de 23 a 35, que incluye estructuras de datos del codificador del núcleo de canal secundario SCh, en el que, al cambiar del segundo modo estéreo DFT al primer modo estéreo TD, el control de cambio de modo estéreo comprende restablecer o estimar las estructuras de datos del

codificador del núcleo del canal secundario SCh basadas en estructuras de datos del codificador del núcleo del canal primario PCh.

37. Un método para decodificar una señal de sonido estéreo que comprende lo siguiente:

5 proporcionar un primer decodificador estéreo de la señal de sonido estéreo usando un primer modo estéreo que opera en el dominio de tiempo, TD, en el que el primer decodificador estéreo, en tramas TD de la señal de sonido estéreo, (a) decodifica una señal mezclada y (b) usa las primeras estructuras de datos y memorias;

10 proporcionar un segundo decodificador estéreo de la señal de sonido estéreo usando un segundo modo estéreo que opera en el dominio de frecuencia, FD, en el que el segundo decodificador estéreo, en tramas FD de la señal de sonido estéreo, (a) decodifica una segunda señal mezclada y (b) usa las segundas estructuras de datos y memorias;

controlar el cambio entre (i) el primer modo estéreo TD y el primer decodificador estéreo y (ii) el segundo modo estéreo FD y el segundo decodificador estéreo;

15 caracterizado porque, al cambiar de uno de los primeros modos estéreo TD y de los segundos modos estéreo FD a otro de los primeros modos estéreo TD y de los segundos modos estéreo FD, el control de cambio de modo estéreo comprende recalcular al menos una longitud de señal mezclada en una trama actual de la señal de sonido estéreo, en el que la longitud de la señal mezclada recalculada en el primer modo estéreo es diferente de la longitud de señal mezclada recalculada en el segundo modo estéreo.

20 38. Un método de decodificación de la señal de sonido estéreo de conformidad con la reivindicación 37, en el que el segundo modo estéreo FD es un modo estéreo de la Transformada Discreta de Fourier, DFT.

25 39. Un método de decodificación de la señal de sonido estéreo de conformidad con la reivindicación 38, en el que el primer modo estéreo usa primeros retardos de procesamiento, el segundo modo estéreo usa segundos retardos de procesamiento y el primer y segundo retardos de procesamiento son diferentes y comprenden retrasos de procesamiento de remuestreo y mezcla.

30 40. Un método de decodificación de la señal de sonido estéreo de conformidad con la reivindicación 38 o 39, en el que, al cambiar de uno de los primeros modos estéreo TD y de los segundos modo estéreo DFT a otro de los primeros modos estéreo FD y de los segundos modos estéreo DFT, el control de cambio de modo estéreo comprende mantener la continuidad de al menos una de las siguientes señales y memorias:

- un canal medio m usado en el segundo modo estéreo DFT;
- un canal primario PCh y un canal secundario SCh usados en el primer modo estéreo TD;
- memorias de filtro posterior TCX-LTP;

35 - memorias de análisis OLA DFT a una frecuencia de muestreo interna y a una frecuencia de muestreo de la señal estéreo de salida;

- memorias de síntesis OLA DFT a la frecuencia de muestreo de la señal estéreo de salida;

- una señal estéreo de salida, que incluye los canales l y r; y

- Memorias de señal HB y canales l y r utilizados en BWE e IC-BWE.

40

41. Un método de decodificación de la señal de sonido estéreo de conformidad con cualquiera de las reivindicaciones de 38 a 40, en el que el control de cambio de modo estéreo comprende actualizar los búferes de memoria estéreo OLA DFT en cada trama TD.

45 42. Un método de decodificación de la señal de sonido estéreo de conformidad con cualquiera de las reivindicaciones de 38 a 41, en el que el control de cambio de modo estéreo comprende actualizar memorias de análisis estéreo DFT.

43. Un método de decodificación de la señal de sonido estéreo de conformidad con la reivindicación 42, en el que, al recibir una primera trama DFT después de una trama TD, el control del cambio de modo

estéreo comprende el uso de un número de últimas muestras de un canal primario PCh y un canal secundario SCh de la trama TD para actualizar en la trama DFT las memorias de análisis estéreo DFT de un canal medio estéreo DFT m y un canal lateral s, respectivamente

5 44. Un método de decodificación de la señal de sonido estéreo de conformidad con cualquiera de las reivindicaciones de 38 a 43, en el que el control de cambio de modo estéreo comprende un fundido cruzado de una síntesis sincronizada y alineada en TD con una síntesis sincronizada y alineada en DFT estéreo para suaviza la transición al cambiar de una trama TD a una trama DFT.

10 45. Un método de decodificación de la señal de sonido estéreo de conformidad con cualquiera de las reivindicaciones de 38 a 44, en el que el control de cambio de modo estéreo comprende la actualización de memorias de síntesis estéreo TD durante tramas DFT en caso de que una siguiente trama sea una trama TD.

15 46. Un método de decodificación de la señal de sonido estéreo de conformidad con cualquiera de las reivindicaciones de 38 a 45, en el que, al cambiar de una trama DFT a una trama TD, el control de la conmutación comprende el restablecimiento de las memorias de un decodificador del núcleo de un canal secundario SCh en el primer decodificador estéreo.

47. Un método de decodificación de la señal de sonido estéreo de conformidad con cualquiera de las reivindicaciones de 38 a 46, en el que, al cambiar de una trama DFT a una trama TD, el control de cambio de modo estéreo comprende la supresión de discontinuidades y diferencias entre los canales mezclados estéreo DFT y TD usando la ecualización de energía de la señal.

20 48. Un método de decodificación de la señal de sonido estéreo de conformidad con la reivindicación 47, en el que, para suprimir las discontinuidades y diferencias entre los canales mezclados estéreo DFT y TD, el control de cambio de modo estéreo comprende, si una ganancia objetivo ICA, g_{ICA} , es inferior a 1,0, la alteración del canal izquierdo l, $y_L(i)$, después de la mezcla y antes de la sincronización de tiempo en la trama TD utilizando la siguiente relación:

25
$$y'_L(i) = \alpha \cdot y_L(i) \quad \text{for } i = 0, \dots, L_{eq} - 1$$

donde L_{eq} es una longitud de una señal para ecualizar, y α es un valor de un factor de ganancia obtenido utilizando la siguiente relación:

$$\alpha = g_{ICA} + i \cdot \frac{1 - g_{ICA}}{L_{eq}} \quad \text{for } i = 0, \dots, L_{eq} - 1.$$

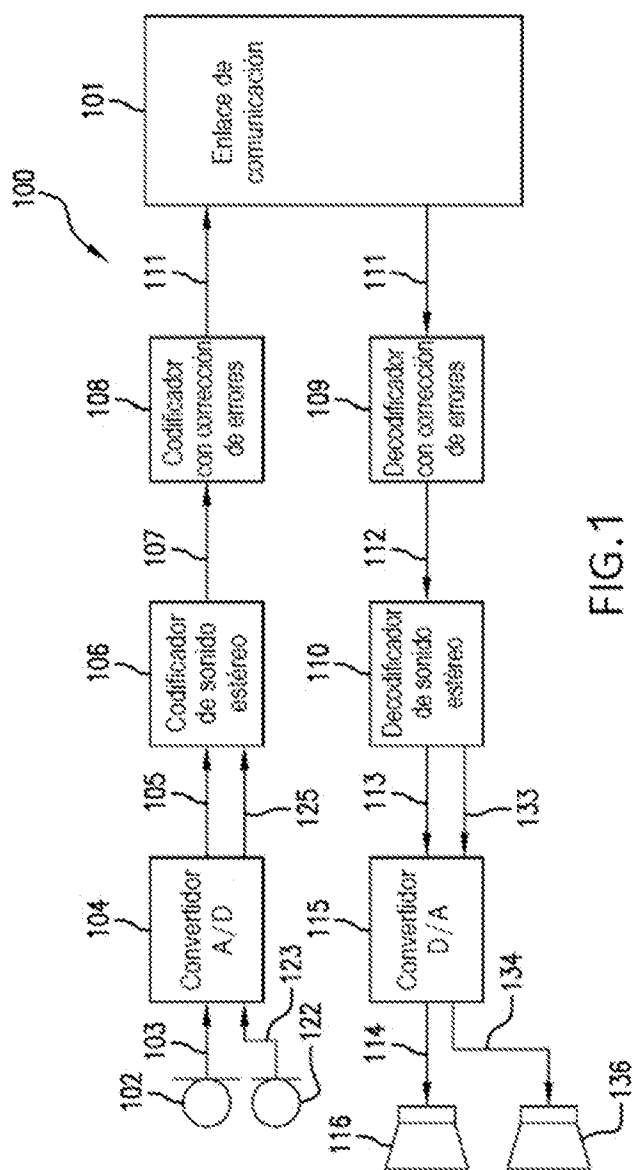


FIG.1

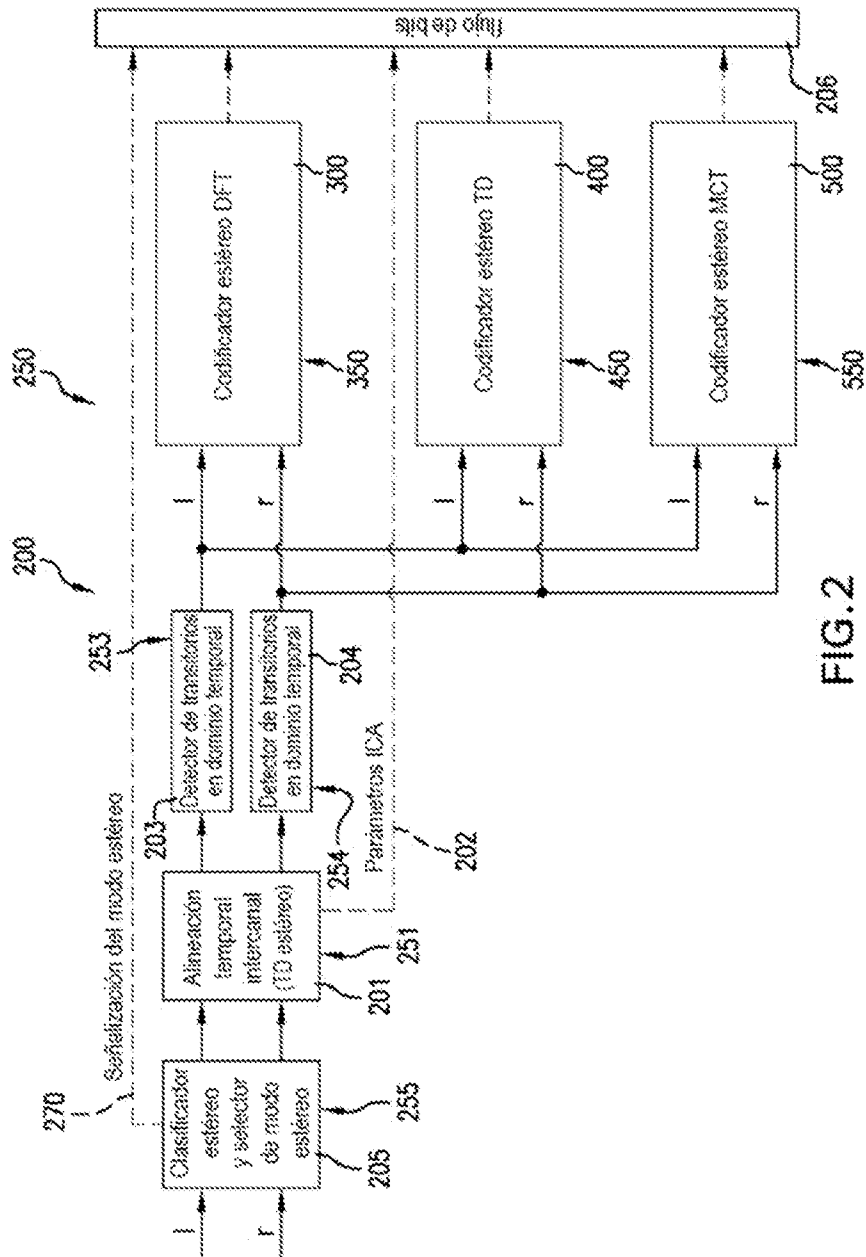


FIG. 2

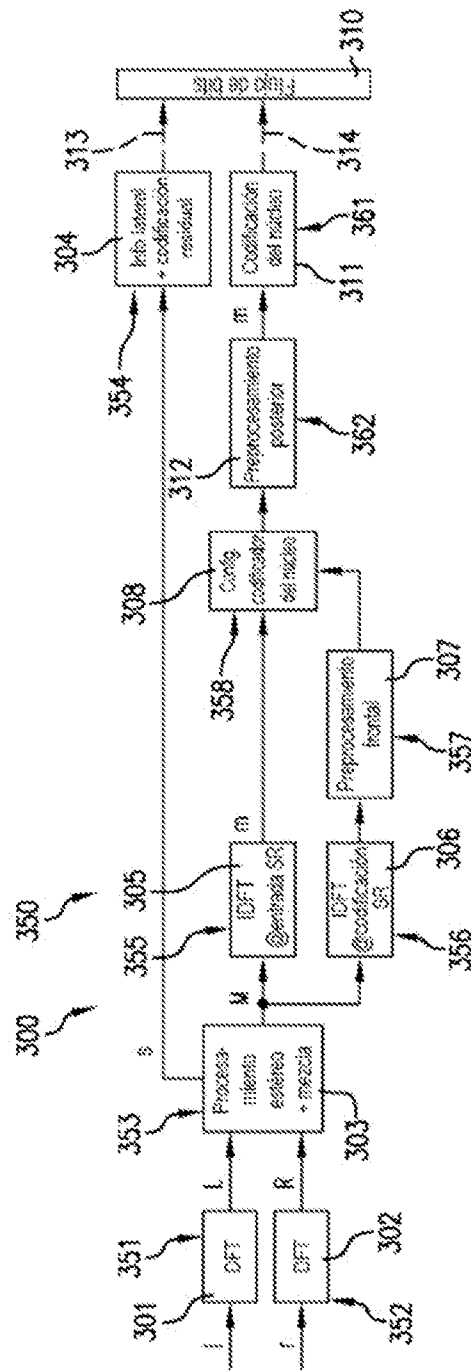


FIG. 3

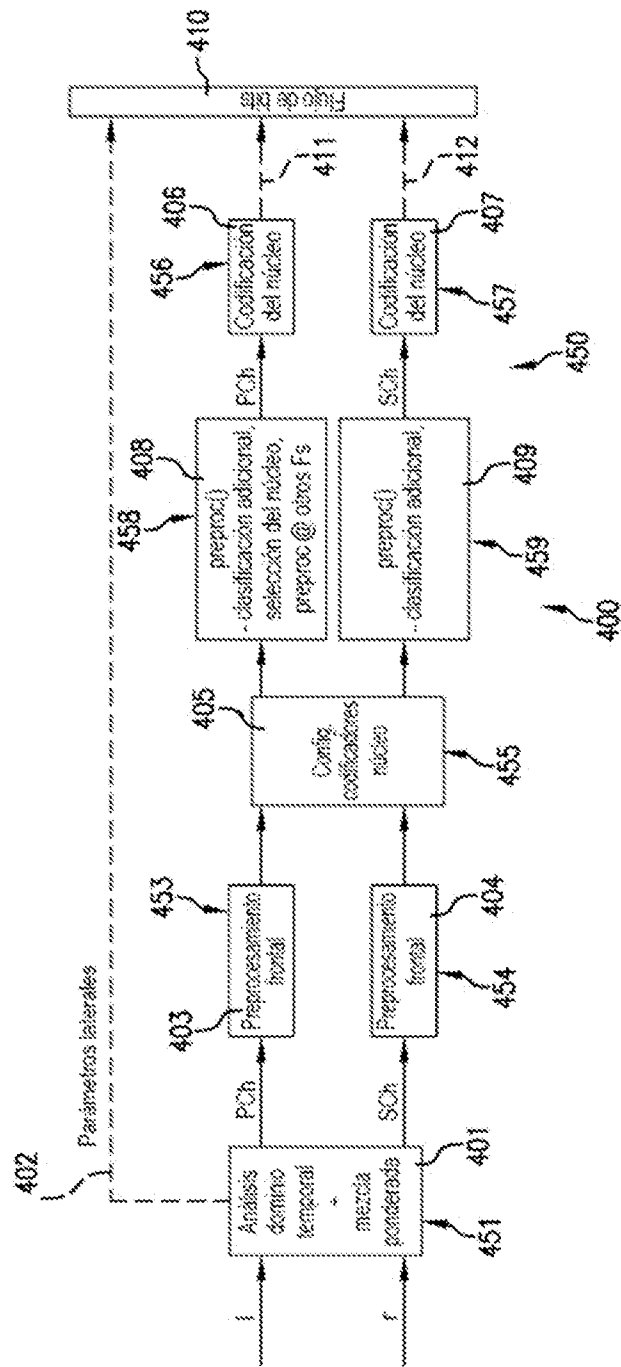


FIG.4

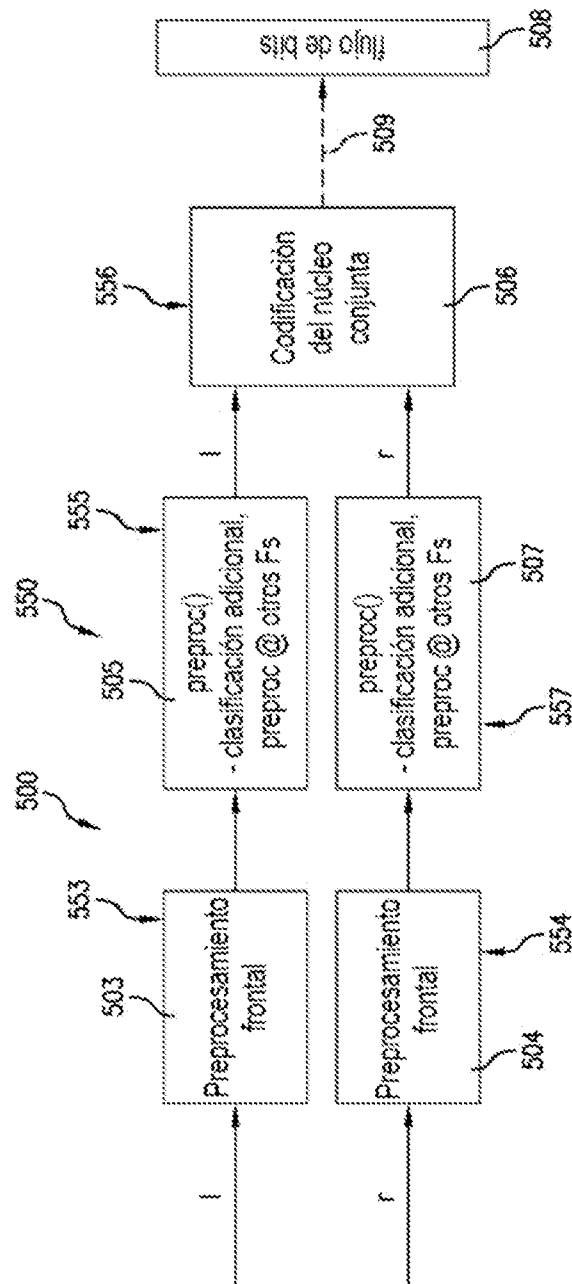
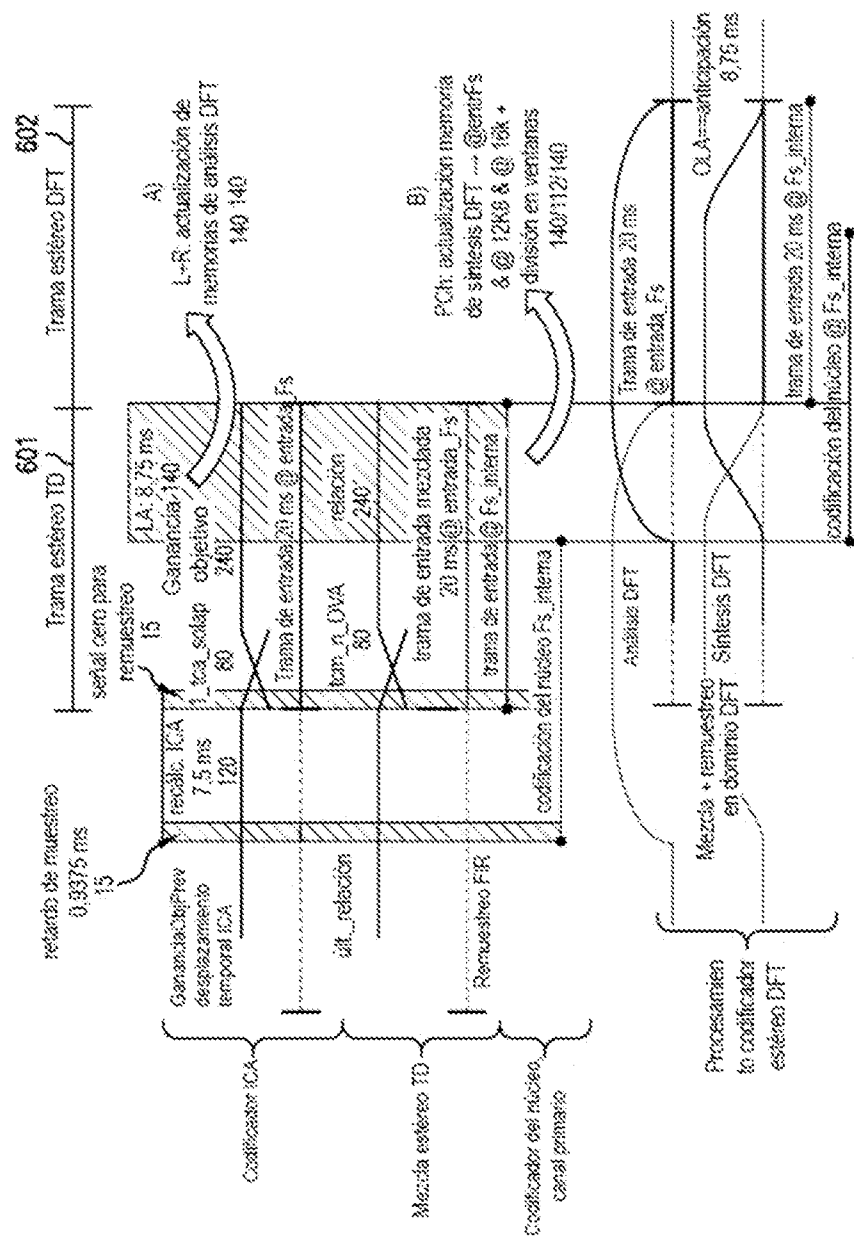


FIG.5



66
67
68
69

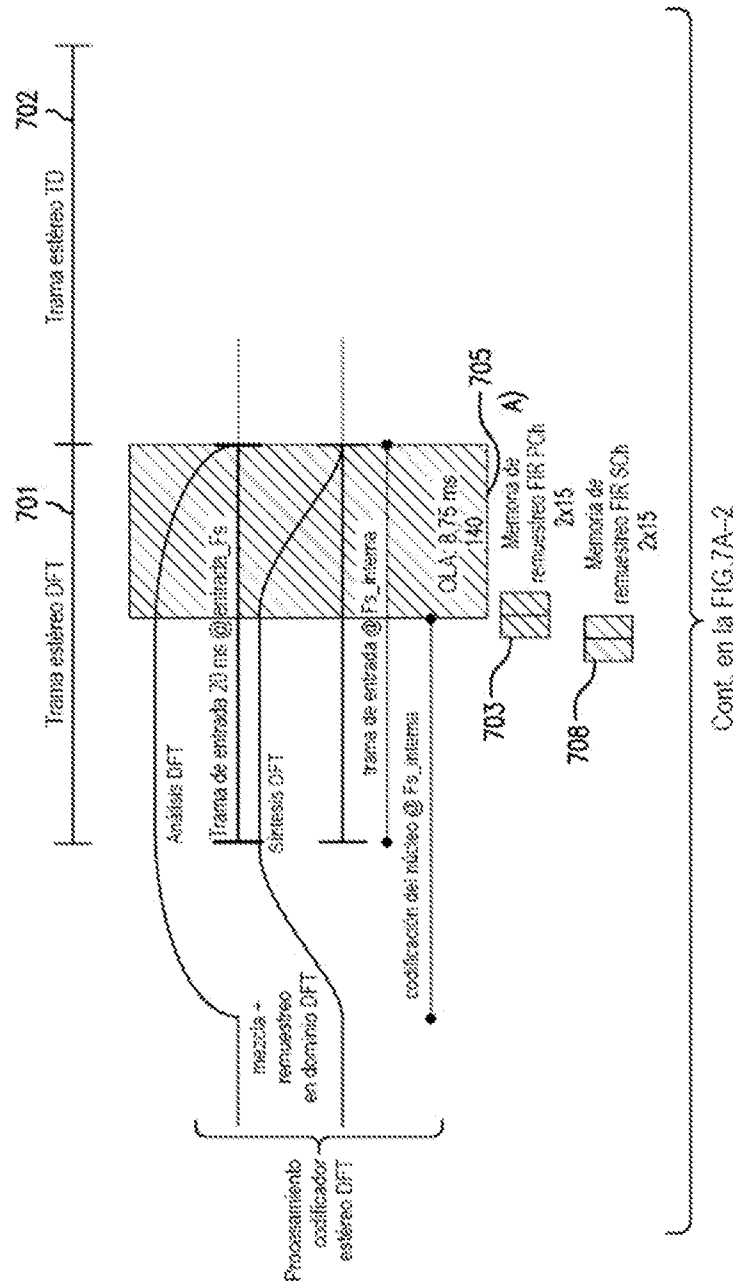
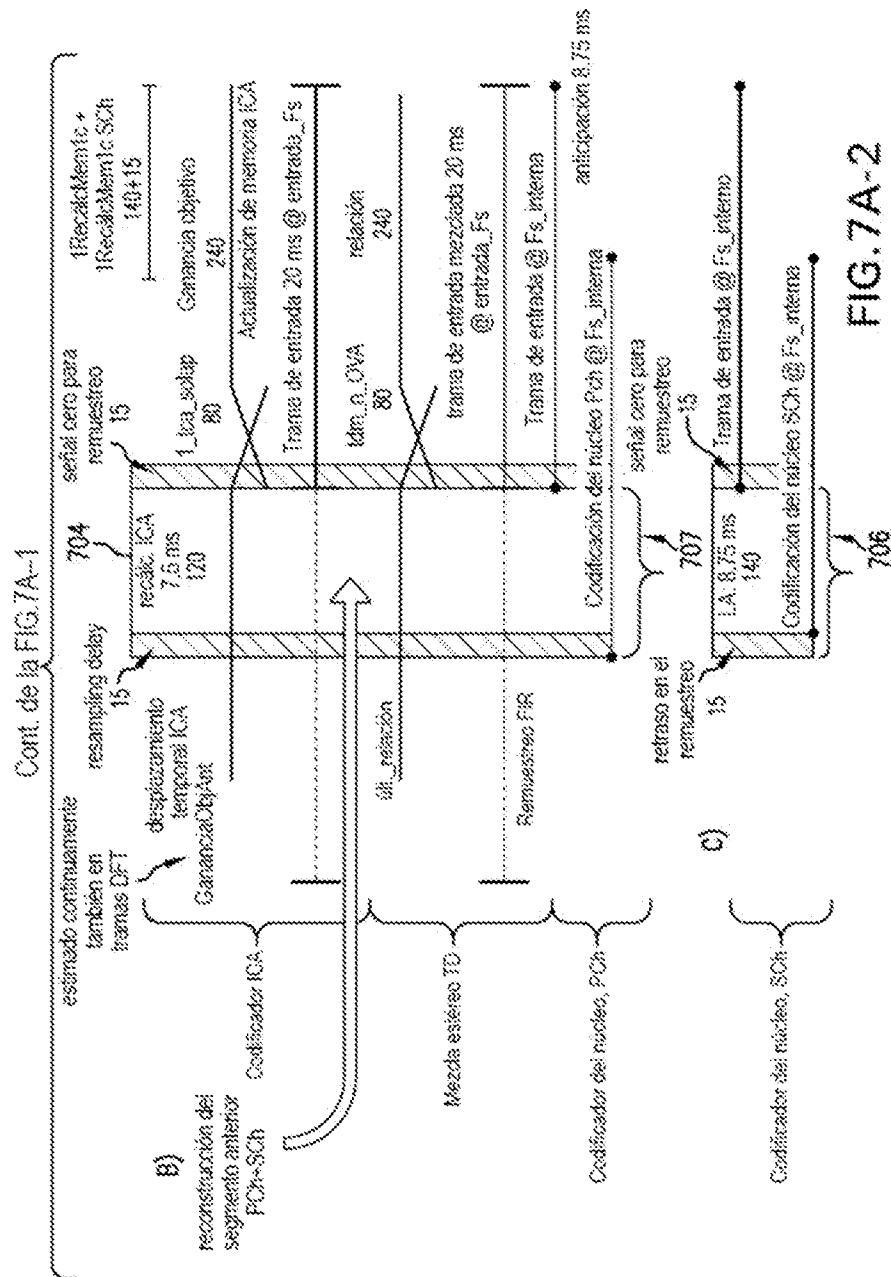


FIG. 7A-1



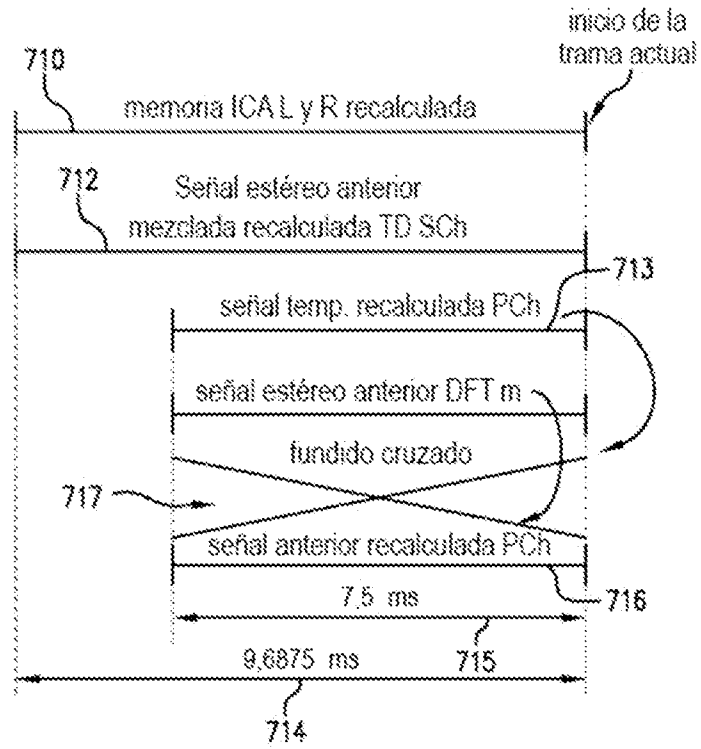


FIG.7B

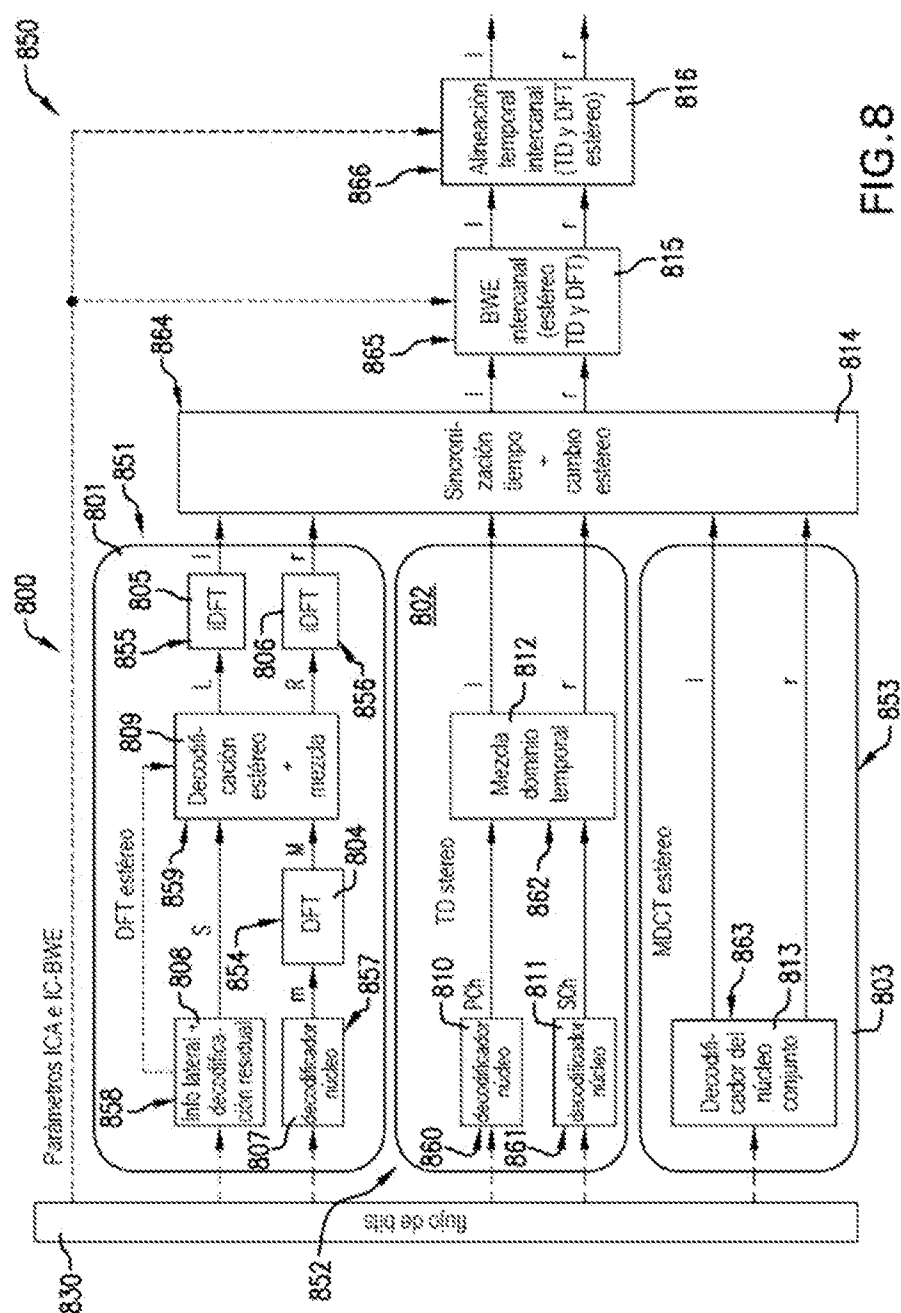


FIG. 8

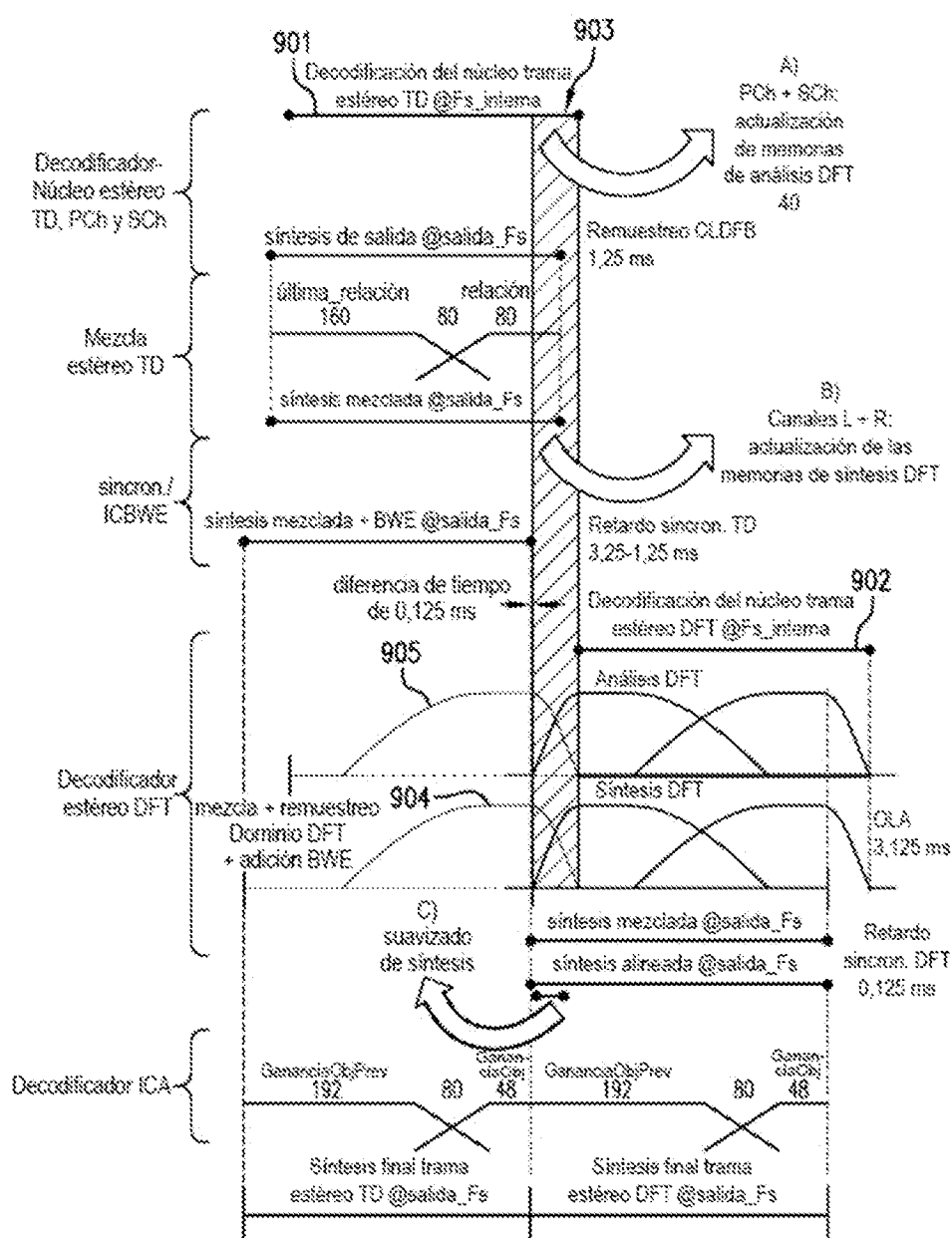


FIG. 9

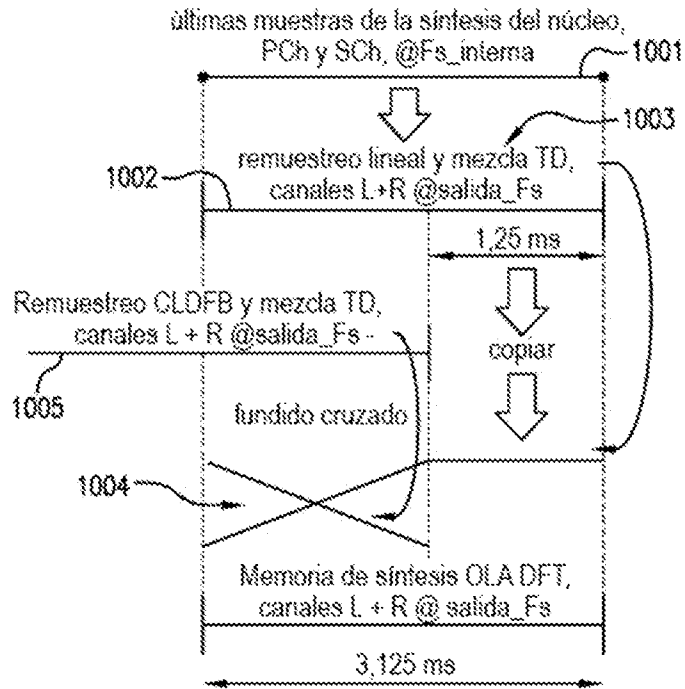


FIG.10

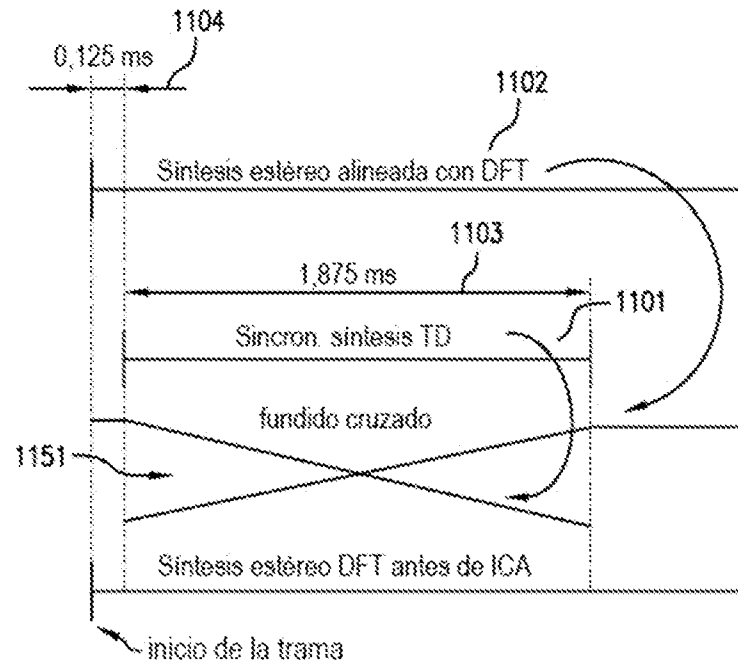


FIG. 11

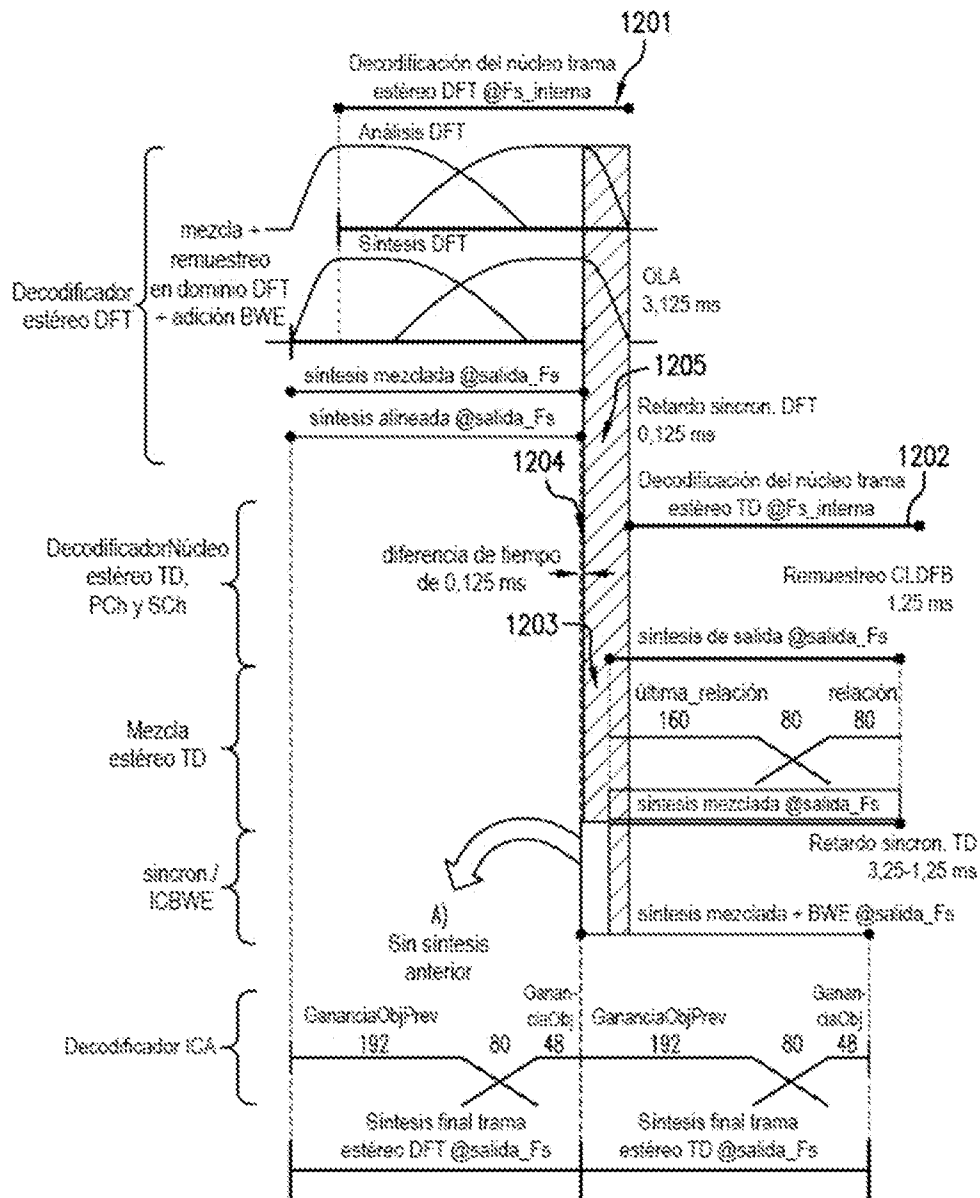


FIG. 12

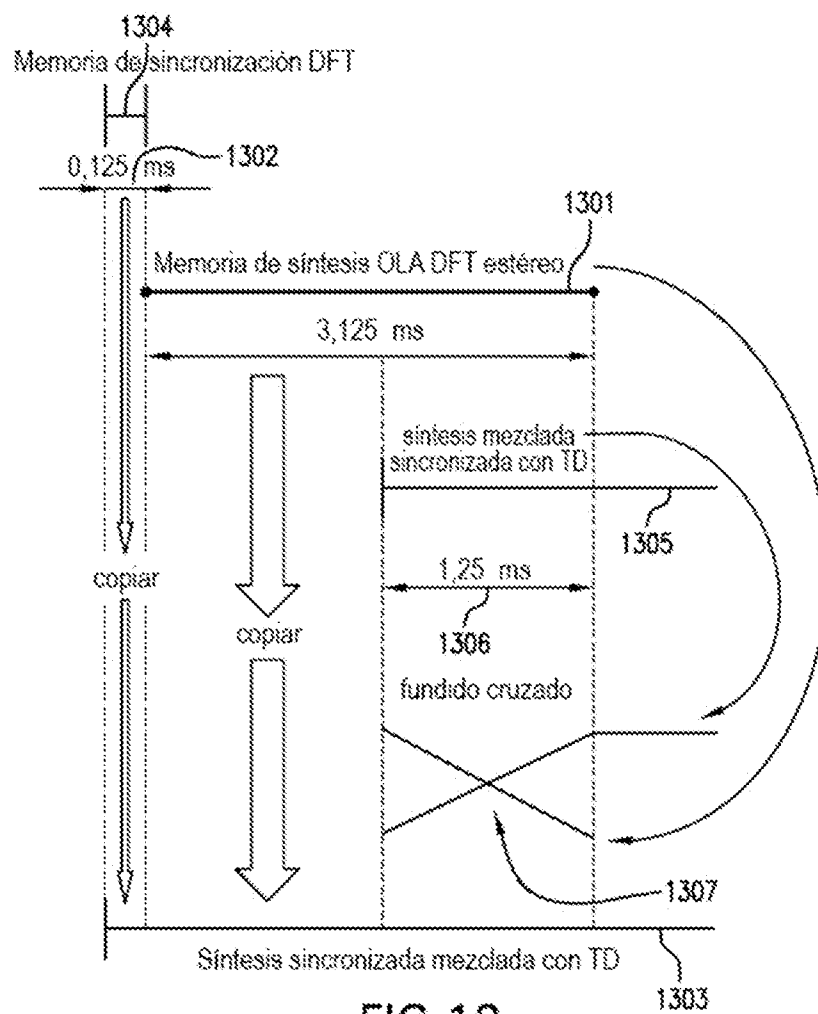


FIG.13

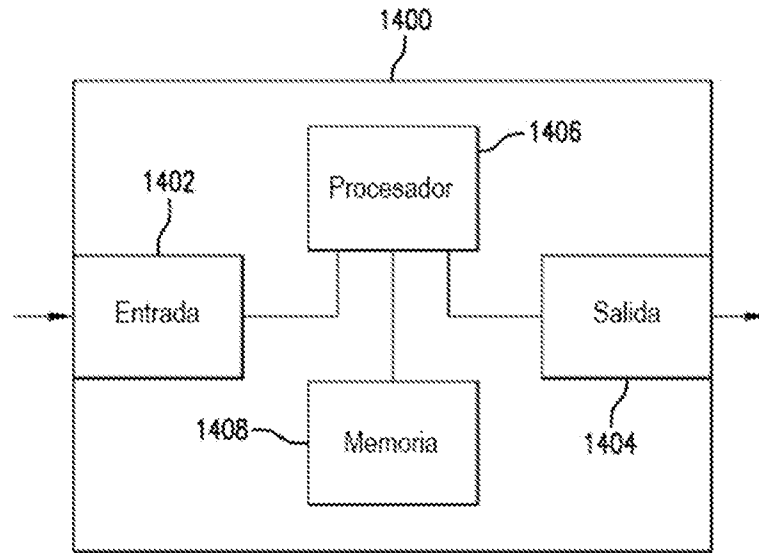


FIG.14