

19



OFICINA ESPAÑOLA DE  
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 715 314**

51 Int. Cl.:

**H04N 21/44** (2011.01)  
**H04N 21/845** (2011.01)  
**H04N 21/236** (2011.01)  
**H04N 19/70** (2014.01)  
**H04N 19/196** (2014.01)  
**H04N 19/40** (2014.01)

12

## TRADUCCIÓN DE PATENTE EUROPEA

T3

86 Fecha de presentación y número de la solicitud internacional: **26.04.2013** **PCT/US2013/038450**

87 Fecha y número de publicación internacional: **31.10.2013** **WO13163569**

96 Fecha de presentación y número de la solicitud europea: **26.04.2013** **E 13721237 (9)**

97 Fecha y número de publicación de la concesión europea: **12.12.2018** **EP 2842344**

54 Título: **Acceso aleatorio completo desde imágenes de acceso aleatorio limpio en codificación de vídeo**

30 Prioridad:

**27.04.2012 US 201261639823 P**  
**25.04.2013 US 201313870846**

45 Fecha de publicación y mención en BOPI de la traducción de la patente:  
**03.06.2019**

73 Titular/es:

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

72 Inventor/es:

**WANG, YE-KUI**

74 Agente/Representante:

**FORTEA LAGUNA, Juan José**

ES 2 715 314 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

Acceso aleatorio completo desde imágenes de acceso aleatorio limpio en codificación de vídeo

### 5 CAMPO TÉCNICO

[0001] Esta divulgación se refiere, en general, al procesamiento de datos de vídeo.

### ANTECEDENTES

- 10 [0002] Las capacidades de vídeo digital se pueden incorporar a una amplia gama de dispositivos, incluyendo televisores digitales, sistemas de radiodifusión directa digital, sistemas de radiodifusión inalámbrica, asistentes digitales personales (PDA), ordenadores portátiles o de escritorio, cámaras digitales, dispositivos de grabación digitales, reproductores de medios digitales, dispositivos de videojuegos, consolas de videojuegos, teléfonos celulares o de radio por satélite, dispositivos de videoconferencia y similares. Entre las normas de codificación de vídeo se incluyen ITU-T H.261, MPEG-1 Visual de la ISO/IEC, ITU-T H.262 o MPEG-2 Visual de la ISO/IEC, ITU-T H.263, MPEG-4 Visual de la ISO/IEC e ITU-T H.264 (también conocida como AVC del MPEG-4 de la ISO/IEC), incluyendo sus ampliaciones de codificación de vídeo escalable (SVC) y de codificación de vídeo de múltiples visualizaciones (MVC). Además, la codificación de vídeo de alta eficiencia (HEVC) es una norma de codificación de vídeo que está en proceso de elaboración por el Equipo de Colaboración Conjunta en Codificación de Vídeo (JCT-VC) del Grupo de Expertos en Codificación de Vídeo (VCEG) de ITU-T y el Grupo de Expertos en Imágenes en Movimiento (MPEG) de ISO/IEC. Los dispositivos de vídeo pueden transmitir, recibir, codificar, descodificar y/o almacenar información de vídeo digital más eficazmente, implementando tales técnicas de codificación de vídeo.
- 20 [0003] Las técnicas de codificación de vídeo incluyen la predicción espacial (intra-imagen) y/o la predicción temporal (entre imágenes) para reducir o eliminar la redundancia intrínseca en las secuencias de vídeo. Para la codificación de vídeo predictiva basada en bloques, un fragmento de vídeo (por ejemplo, una trama de vídeo o una parte de una trama de vídeo) puede dividirse en bloques de vídeo, que también pueden denominarse macrobloques, bloques arbolados, unidades de codificación (CU) y/o nodos de codificación. Los bloques de vídeo en un fragmento intra-codificado (I) de una imagen se codifican utilizando predicción espacial con respecto a muestras de referencia en bloques contiguos en la misma imagen. Los bloques de vídeo en un fragmento inter-codificado (P o B) de una imagen pueden usar la predicción espacial con respecto a muestras de referencia en bloques contiguos en la misma imagen o la predicción temporal con respecto a muestras de referencia en otras imágenes de referencia. Las imágenes pueden denominarse tramas, y las imágenes de referencia pueden denominarse tramas de referencia.
- 30 [0004] La predicción espacial o temporal da como resultado un bloque predictivo para un bloque que se va a codificar. Los datos residuales representan diferencias de píxeles entre el bloque original a codificar y el bloque predictivo. Los píxeles también pueden denominarse elementos de imagen, imágenes o muestras. Un bloque inter-codificado se codifica de acuerdo con un vector de movimiento que apunta a un bloque de muestras de referencia que forman el bloque predictivo, y los datos residuales que indican la diferencia entre el bloque codificado y el bloque predictivo. Un bloque intra-codificado se codifica de acuerdo con un modo de intra-codificación y a los datos residuales. Para una mayor compresión, los datos residuales pueden transformarse desde el dominio de píxeles a un dominio de transformada, dando como resultado coeficientes de transformada residuales, que a continuación se pueden cuantificar. Los coeficientes de transformada cuantificados, dispuestos inicialmente en una formación bidimensional, pueden examinarse con el fin de generar un vector unidimensional de coeficientes de transformada, y puede aplicarse codificación por entropía para lograr aún más compresión.
- 40 [0005] En "WD6: High Efficiency Video Coding (HEVC) text specification draft 6 (version JCTVC-H1 003\_dA) [WD6: borrador de memoria descriptiva textual de codificación de vídeo de alta eficiencia (HEVC) 6 (versión JCTVC-H1 003\_dA)]", de Benjamin Bross et al, Equipo de colaboración conjunta sobre codificación de vídeo 7.<sup>a</sup> reunión: Ginebra, Suiza, 21-30 de noviembre de 2011, páginas 1-239, se publicó un borrador de memoria descriptiva para la HEVC que data del 13 de febrero de 2012.
- 50 [0006] El "Text of ISO/IEC 14496-15/FDIS [Texto de ISO/IEC 14496-15/FDIS]" de Singer et al. (65.<sup>a</sup> reunión MPEG; Trondheim; documento n.º N5780, 13 de agosto de 2003) describe cómo se puede almacenar el contenido de AVC en un formato compatible con formato de archivo de medios de base ISO. El documento hace referencia a conjuntos de parámetros y puntos de recuperación de acceso aleatorio.
- 55 [0007] El "Study Text of ISO/IEC 14496-1:2001/FPDAM7 [Texto de estudio de ISO/IEC 14496-1: 2001/FPDAM7]" de Lim (64.<sup>a</sup> reunión MPEG; Pattaya; documento n.º N5609, 5 de mayo de 2003) indica que un indicador en una cabecera de paquete de capa de sincronización MPEG-4 puede configurarse en "1" si está encapsulando una AU que contiene una imagen de actualización instantánea del descodificador o un mensaje de punto de recuperación SEI y los conjuntos de parámetros necesarios para su descodificación ya están disponibles o se entregan junto con la misma AU.

### 65 SUMARIO

**[0008]** En general, esta divulgación describe técnicas para proporcionar una indicación de si realizar acceso aleatorio a partir de una unidad de acceso particular en un flujo de bits requiere obtener conjuntos de parámetros a partir de unidades de acceso anteriores. Convencionalmente, una imagen de punto de acceso aleatorio (RAP) es una imagen de actualización de decodificación instantánea (IDR), que indica el inicio de una secuencia de vídeo codificada y siempre limpia una memoria intermedia de imágenes decodificadas (DPB) de un decodificador de vídeo. La norma de codificación de vídeo de alta eficiencia (HEVC) introduce una imagen de acceso aleatorio limpio (CRA) que se puede colocar en cualquier punto dentro de una secuencia de vídeo codificada y no limpia la DPB, lo cual proporciona una eficiencia de codificación mejorada. Las imágenes CRA, por lo tanto, proporcionan más opciones para el acceso aleatorio al flujo de bits sin una menor eficiencia de codificación.

**[0009]** En el caso de acceso aleatorio de una imagen CRA, no se decodificarán todas las imágenes anteriores a la imagen CRA en orden de salida o decodificación y no están disponibles para su uso como imágenes de referencia. En este caso, para decodificar la imagen CRA o cualquiera de las siguientes imágenes en orden de salida o decodificación, un decodificador de vídeo puede necesitar obtener uno o más conjuntos de parámetros incluidos en una unidad de acceso para una de las imágenes anteriores no disponibles, lo cual es una operación pesada. Las técnicas de esta divulgación proporcionan una indicación, para cada imagen CRA, que indica si los conjuntos de parámetros incluidos en las unidades de acceso anteriores son necesarios para realizar un acceso aleatorio desde la imagen. Cuando no se necesitan conjuntos de parámetros de unidades de acceso anteriores para el acceso aleatorio desde una imagen de CRA particular, un decodificador de vídeo puede determinar realizar un acceso aleatorio desde esa imagen sin una operación de búsqueda. A la inversa, cuando se necesitan conjuntos de parámetros de unidades de acceso anteriores para el acceso aleatorio desde una imagen CRA particular, el decodificador de vídeo puede determinar que no se realice un acceso aleatorio desde esa imagen para evitar la operación de búsqueda.

**[0010]** La invención se define mediante las reivindicaciones.

**[0011]** Los detalles de uno o más ejemplos se exponen en los dibujos adjuntos y en la descripción siguiente. Otras características, objetos y ventajas resultarán evidentes a partir de la descripción y de los dibujos, y a partir de las reivindicaciones. La divulgación habilitante para la invención se encuentra en el modo de realización de la Figura 7. Los modos de realización restantes deben entenderse como ejemplos que no describen partes de la presente invención.

## BREVE DESCRIPCIÓN DE LOS DIBUJOS

### **[0012]**

La FIG. 1 es un diagrama de bloques que ilustra un sistema de codificación y decodificación de vídeo de ejemplo que puede utilizar las técnicas descritas en esta divulgación.

La FIG. 2 es un diagrama de bloques que ilustra un codificador de vídeo de ejemplo que puede implementar las técnicas descritas en esta divulgación.

La FIG. 3 es un diagrama de bloques que ilustra un decodificador de vídeo de ejemplo que puede implementar las técnicas descritas en esta divulgación.

La FIG. 4 es un diagrama conceptual que ilustra una secuencia de vídeo codificada de acuerdo con técnicas de codificación de vídeo predictiva.

La FIG. 5 es un diagrama conceptual que ilustra un ejemplo de una secuencia de vídeo codificada.

La FIG. 6 es un diagrama de flujo que ilustra una operación para determinar si una actualización del conjunto de parámetros puede ocurrir en una parte de un flujo de bits.

La FIG. 7 es un diagrama de flujo que ilustra una operación para determinar si se puede realizar el acceso aleatorio desde una unidad de acceso CRA sin tener que obtener conjuntos de parámetros de unidades de acceso no disponibles anteriores.

La FIG. 8 es un diagrama de bloques que ilustra un conjunto a modo de ejemplo de dispositivos que forman parte de una red.

## DESCRIPCIÓN DETALLADA

**[0013]** Las técnicas de esta divulgación proporcionan una indicación de si la realización de acceso aleatorio a partir de una unidad de acceso particular en un flujo de bits requiere obtener conjuntos de parámetros a partir de unidades de acceso anteriores. Convencionalmente, una imagen de punto de acceso aleatorio (RAP) es una imagen de actualización de decodificación instantánea (IDR), que indica el inicio de una secuencia de vídeo codificada y siempre limpia una memoria intermedia de imágenes decodificadas (DPB) de un decodificador de vídeo. La norma de

codificación de vídeo de alta eficiencia (HEVC) introduce una imagen de acceso aleatorio limpio (CRA) que se puede colocar en cualquier punto dentro de una secuencia de vídeo codificada y no limpia la DPB, lo cual proporciona una eficiencia de codificación mejorada. Las imágenes CRA, por lo tanto, proporcionan más opciones para el acceso aleatorio al flujo de bits sin una menor eficiencia de codificación.

**[0014]** En el caso de acceso aleatorio de una imagen CRA, no se descodificarán todas las imágenes anteriores a la imagen CRA en orden de salida o descodificación y no están disponibles para su uso como imágenes de referencia. En este caso, para descodificar la imagen CRA o cualquiera de las siguientes imágenes en orden de salida o descodificación, un descodificador de vídeo puede necesitar obtener uno o más conjuntos de parámetros incluidos en una unidad de acceso para una de las imágenes anteriores no disponibles, lo cual es una operación pesada. Las técnicas de esta divulgación proporcionan una indicación, para cada imagen CRA, que indica si los conjuntos de parámetros incluidos en las unidades de acceso anteriores son necesarios para realizar un acceso aleatorio desde la imagen. Cuando no se necesitan conjuntos de parámetros de unidades de acceso anteriores para el acceso aleatorio desde una imagen de CRA particular, un descodificador de vídeo puede determinar realizar un acceso aleatorio desde esa imagen sin una operación de búsqueda. A la inversa, cuando se necesitan conjuntos de parámetros de unidades de acceso anteriores para el acceso aleatorio desde una imagen CRA particular, el descodificador de vídeo puede determinar que no se realice un acceso aleatorio desde esa imagen para evitar la operación de búsqueda.

**[0015]** En algunos casos, la indicación puede comprender un tipo particular de unidad de capa de abstracción de red (NAL) incluida en una unidad de acceso para cada imagen CRA en un flujo de bits. En otros casos, la indicación puede comprender un elemento sintáctico, como un indicador, incluido en una carga útil de unidad NAL incluida en una unidad de acceso para cada imagen CRA en el flujo de bits. En un ejemplo, la indicación puede comprender un elemento sintáctico, como un indicador, incluido en una unidad NAL de información de mejora suplementaria (SEI) en la unidad de acceso para la imagen CRA. Aunque se describen principalmente con respecto a las imágenes de CRA, las técnicas de esta divulgación se pueden usar de manera similar para otros tipos de imágenes RAP, incluidas las imágenes IDR y las imágenes de actualización de descodificación gradual (GDR).

**[0016]** La FIG. 1 es un diagrama de bloques que ilustra un sistema de codificación y descodificación de vídeo 10 de ejemplo que puede utilizar las técnicas descritas en esta divulgación. Como se muestra en la FIG. 1, el sistema 10 incluye un dispositivo de origen 12 que proporciona datos de vídeo codificados, a ser descodificados en un momento posterior por un dispositivo de destino 14. En particular, el dispositivo de origen 12 proporciona los datos de vídeo al dispositivo de destino 14 mediante un medio legible por ordenador 16. El dispositivo 12 de origen y el dispositivo 14 de destino pueden comprender cualquiera de entre una amplia gama de dispositivos, incluidos ordenadores de escritorio, ordenadores plegables (es decir, portátiles), ordenadores de tableta, descodificadores, teléfonos tales como los denominados teléfonos "inteligentes", los denominados paneles "inteligentes", televisores, cámaras, dispositivos de visualización, reproductores de medios digitales, consolas de videojuegos, un dispositivo de transmisión de vídeo continua o similares. En algunos casos, el dispositivo 12 de origen y el dispositivo 14 de destino pueden estar equipados para la comunicación inalámbrica.

**[0017]** El dispositivo de destino 14 puede recibir los datos de vídeo codificados a descodificar mediante el medio legible por ordenador 16. El medio legible por ordenador 16 puede comprender cualquier tipo de medio o dispositivo capaz de desplazar los datos de vídeo codificados desde el dispositivo de origen 12 al dispositivo de destino 14. En un ejemplo, el medio legible por ordenador 16 puede comprender un medio de comunicación que habilita el dispositivo de origen 12 para transmitir datos de vídeo codificados directamente al dispositivo de destino 14 en tiempo real. Los datos de vídeo codificados se pueden modular de acuerdo con una norma de comunicación, tal como un protocolo de comunicación inalámbrica, y transmitir al dispositivo 14 de destino. El medio de comunicación puede comprender cualquier medio de comunicación, inalámbrica o cableada, tal como un espectro de radiofrecuencia (RF) o una o más líneas de transmisión física. El medio de comunicación puede formar parte de una red basada en paquetes, tal como una red de área local, una red de área amplia o una red global tal como Internet. El medio de comunicación puede incluir routers, conmutadores, estaciones base o cualquier otro equipo que pueda ser útil para facilitar la comunicación desde el dispositivo 12 de origen hasta el dispositivo 14 de destino.

**[0018]** En algunos ejemplos, pueden emitirse datos codificados desde la interfaz de salida 22 hasta un dispositivo de almacenamiento. De forma similar, se puede acceder a los datos codificados del dispositivo de almacenamiento mediante una interfaz de entrada. El dispositivo de almacenamiento puede incluir cualquiera entre una diversidad de medios de almacenamiento de datos, de acceso distribuido o local, tales como una unidad de disco duro, discos Blu-ray, discos DVD, discos CD-ROM, memoria flash, memoria volátil o no volátil u otros medios adecuados cualesquiera de almacenamiento digital para almacenar datos de vídeo codificados. En un ejemplo adicional, el dispositivo de almacenamiento puede corresponder a un servidor de ficheros o a otro dispositivo de almacenamiento intermedio que pueda almacenar los datos de vídeo codificados, generados por el dispositivo de origen 12. El dispositivo de destino 14 puede acceder a datos de vídeo almacenados desde el dispositivo de almacenamiento mediante transmisión continua o descarga. El servidor de archivos puede ser cualquier tipo de servidor capaz de almacenar datos de vídeo codificados y transmitir esos datos de vídeo codificados al dispositivo 14 de destino. Los servidores de archivos de ejemplo incluyen un servidor web (por ejemplo, para un sitio web), un servidor FTP, dispositivos de almacenamiento conectados en red (NAS) o una unidad de disco local. El dispositivo de destino 14 puede acceder a los datos de vídeo codificados a través de cualquier conexión de datos estándar, incluyendo una conexión a Internet. Esto puede incluir

un canal inalámbrico (por ejemplo, una conexión wifi), una conexión cableada (por ejemplo, DSL, módem de cable, etc.) o una combinación de ambos que sea adecuada para acceder a datos de vídeo codificados almacenados en un servidor de archivos. La transmisión de datos de vídeo codificados desde el dispositivo de almacenamiento puede ser una transmisión continua, una transmisión de descarga o una combinación de ambas.

**[0019]** Las técnicas de esta divulgación no están limitadas necesariamente a aplicaciones o configuraciones inalámbricas. Las técnicas pueden aplicarse a la codificación de vídeo, para soportar cualquiera de una diversidad de aplicaciones multimedia, tales como radiodifusiones de televisión por el aire, transmisiones de televisión por cable, transmisiones de televisión por satélite, transmisiones de vídeo continuas, por ejemplo, a través de Internet, como transmisión continua adaptable dinámica por HTTP (DASH), codificación de vídeo digital para su almacenamiento en un medio de almacenamiento de datos, descodificación de vídeo digital almacenado en un medio de almacenamiento de datos, u otras aplicaciones. En algunos ejemplos, el sistema 10 se puede configurar para admitir la transmisión de vídeo unidireccional o bidireccional, para admitir aplicaciones tales como la transmisión de vídeo continua, la reproducción de vídeo, la radiodifusión de vídeo y/o la videotelefonía.

**[0020]** En el ejemplo de la FIG. 1, el dispositivo de origen 12 incluye una fuente de vídeo 18, un codificador de vídeo 20 y una interfaz de salida 22. El dispositivo de destino 14 incluye una interfaz de entrada 28, un descodificador de vídeo 30 y un dispositivo de visualización 32. En otros ejemplos, un dispositivo de origen y un dispositivo de destino pueden incluir otros componentes o disposiciones. Por ejemplo, el dispositivo de origen 12 puede recibir datos de vídeo desde un origen de vídeo externo 18, tal como una cámara externa. Del mismo modo, el dispositivo de destino 14 puede interactuar con un dispositivo de visualización externo, en lugar de incluir un dispositivo de visualización integrado.

**[0021]** El sistema ilustrado 10 de la FIG. 1 es solo un ejemplo. Las técnicas de esta divulgación pueden realizarse mediante cualquier dispositivo de codificación y/o descodificación de vídeo digital. Aunque en general las técnicas son realizadas por un dispositivo de codificación de vídeo, las técnicas también pueden ser realizadas por un codificador/descodificador de vídeo, típicamente denominado un "CODEC". Además, las técnicas de esta divulgación también pueden realizarse mediante un preprocesador de vídeo. El dispositivo de origen 12 y el dispositivo de destino 14 son simplemente ejemplos de dichos dispositivos de codificación en los que el dispositivo de origen 12 genera datos de vídeo codificados para su transmisión al dispositivo de destino 14. En algunos ejemplos, los dispositivos 12, 14 pueden funcionar de manera esencialmente simétrica, de modo que cada uno de los dispositivos 12, 14 incluya componentes de codificación y de descodificación de vídeo. Por lo tanto, el sistema 10 puede prestar soporte a la transmisión de vídeo unidireccional o bidireccional entre los dispositivos de vídeo 12, 14, por ejemplo, para la transmisión continua de vídeo, la reproducción de vídeo, la radiodifusión de vídeo o la videotelefonía.

**[0022]** La fuente de vídeo 18 del dispositivo de origen 12 puede incluir un dispositivo de captura de vídeo, tal como una cámara de vídeo, un archivo de vídeo que contiene vídeo grabado previamente y/o una interfaz de alimentación de vídeo para recibir vídeo desde un proveedor de contenidos de vídeo. Como otra alternativa, la fuente de vídeo 18 puede generar datos basados en gráficos por ordenador como el vídeo de origen, o una combinación de vídeo en directo, vídeo archivado y vídeo generado por ordenador. En algunos casos, si la fuente de vídeo 18 es una videocámara, el dispositivo de origen 12 y el dispositivo de destino 14 pueden formar los llamados teléfonos con cámara o videoteléfonos. Sin embargo, como se ha mencionado anteriormente, las técnicas descritas en esta divulgación pueden ser aplicables a la codificación de vídeo en general, y pueden aplicarse a aplicaciones inalámbricas y/o alámbricas. En cada caso, el codificador de vídeo 20 puede codificar el vídeo captado, precaptado o generado por ordenador. La información de vídeo codificada puede entonces ser emitida por la interfaz de salida 22 hacia un medio legible por ordenador 16.

**[0023]** El medio legible por ordenador 16 puede incluir medios transitorios, tales como una radiodifusión inalámbrica o una transmisión de red alámbrica, o medios de almacenamiento (es decir, medios de almacenamiento no transitorio), tales como un disco duro, una unidad de memoria flash, un disco compacto, un disco de vídeo digital, un disco Blu-ray u otros medios legibles por ordenador. En algunos ejemplos, un servidor de red (no mostrado) puede recibir datos de vídeo codificados desde el dispositivo de origen 12 y proporcionar los datos de vídeo codificados al dispositivo de destino 14, por ejemplo, mediante transmisión por red. De manera similar, un dispositivo informático de una utilidad de producción de un medio, tal como una utilidad de grabación de discos, puede recibir datos de vídeo codificados desde el dispositivo de origen 12 y producir un disco que contiene los datos de vídeo codificados. Por lo tanto, puede entenderse que el medio legible por ordenador 16 incluye uno o más medios legibles por ordenador de diversas formas, en varios ejemplos.

**[0024]** La interfaz de entrada 28 del dispositivo de destino 14 recibe información desde el medio legible por ordenador 16. La información del medio legible por ordenador 16 puede incluir información sintáctica definida por el codificador de vídeo 20, que también usa el descodificador de vídeo 30, que incluye elementos sintácticos que describen características y/o procesamiento de bloques y otras unidades codificadas, por ejemplo, GOP. El dispositivo de visualización 32 muestra los datos de vídeo descodificados a un usuario, y puede comprender cualquiera entre una diversidad de dispositivos de visualización, tales como un tubo de rayos catódicos (CRT), una pantalla de cristal líquido (LCD), una pantalla de plasma, una pantalla de diodos orgánicos emisores de luz (OLED) u otro tipo de dispositivo de visualización.

**[0025]** El codificador de vídeo 20 y el descodificador de vídeo 30 pueden funcionar de acuerdo con una norma de codificación de vídeo, tal como la norma de codificación de vídeo de alta eficiencia (HEVC), actualmente en fase de desarrollo por el Equipo de colaboración conjunta en codificación de vídeo (JCT-VC) del Grupo de expertos de codificación de vídeo (VCEG) de ITU-T y el Grupo de expertos en imágenes en movimiento (MPEG) de ISO/IEC, como producto de una asociación colectiva conocida como Equipo de Vídeo Conjunto (JVT). La norma HEVC puede cumplir con el Modelo de prueba HEVC (HM). Un borrador reciente de HEVC HM, denominado "HEVC Working Draft 6 [Borrador 6 de trabajo de la HEVC]" o "WD6", se describe en el documento JCTVC-H1003, de Bross et al., titulado "High efficiency video coding (HEVC) text specification draft 6 [Memoria descriptiva textual de Codificación de Vídeo de Alta Eficiencia (HEVC), Borrador 6]", Equipo de Colaboración Conjunta en Codificación de Vídeo (JCT-VC) de ITU-T SG16 WP3 e ISO/IEC JTC1/SC29/WG11, 8.<sup>a</sup> reunión: San José, California, EE.UU., febrero de 2012.

**[0026]** Sin embargo, las técnicas de esta divulgación no están limitadas a ninguna norma de codificación particular. Aunque no se muestra en la FIG. 1, en algunos aspectos, tanto el codificador 20 de vídeo como el descodificador 30 de vídeo pueden estar integrados en un codificador y descodificador de audio, y pueden incluir unidades MUX-DEMUX adecuadas, u otro tipo de hardware y software, para ocuparse de la codificación tanto de audio como de vídeo en un flujo de datos común o en flujos de datos separados. Si procede, las unidades MUX-DEMUX pueden ajustarse al protocolo de multiplexador H.223 de la ITU o a otros protocolos tales como el protocolo de datagramas de usuario (UDP).

**[0027]** El codificador de vídeo 20 y el descodificador de vídeo 30 se pueden implementar, cada uno, como cualquiera entre una variedad de circuitos codificadores adecuados, tales como uno o más microprocesadores, procesadores de señales digitales (DSP), circuitos integrados específicos de la aplicación (ASIC), formaciones de puertas programables *in situ* (FPGA), lógica discreta, software, hardware, firmware o cualquier combinación de los mismos. Cuando las técnicas se implementan parcialmente en software, un dispositivo puede almacenar instrucciones para el software en un medio adecuado no transitorio legible por ordenador, y ejecutar las instrucciones en hardware utilizando uno o más procesadores para realizar las técnicas de esta divulgación. Cada uno entre el codificador de vídeo 20 y el descodificador de vídeo 30 se puede incluir en uno o más codificadores o descodificadores, cualquiera de los cuales se puede integrar como parte de un codificador/descodificador (CÓDEC) combinado en un dispositivo respectivo.

**[0028]** El equipo JCT-VC está trabajando en el desarrollo de la norma HEVC. La iniciativa de normalización HEVC se basa en un modelo en evolución de un dispositivo de codificación de vídeo denominado modelo de prueba HEVC (HM). El HM supone varias capacidades adicionales relativas a las normas existentes. Por ejemplo, mientras que la norma H.264 proporciona nueve modos de codificación mediante predicción intra, el HM puede proporcionar hasta treinta y tres modos de codificación mediante predicción intra.

**[0029]** En general, el HM describe que una trama o imagen de vídeo puede dividirse en una secuencia de bloques arbolados o unidades de codificación de máximo tamaño (LCU), que incluyen muestras tanto de luma como de croma. Los datos sintácticos dentro de un flujo de bits pueden definir un tamaño para la LCU, que es la máxima unidad de codificación en lo que respecta al número de píxeles. Un fragmento incluye un número de bloques de árbol consecutivos en orden de codificación. Una trama o imagen de vídeo puede dividirse en uno o más fragmentos. Cada bloque de árbol puede dividirse en unidades de codificación (CU) de acuerdo con un árbol cuaternario. En general, una estructura de datos de árbol cuaternario incluye un nodo por CU, con un nodo raíz correspondiente al bloque arbolado. Si una CU se divide en cuatro sub-CU, el nodo correspondiente a la CU incluye cuatro nodos hoja, cada uno de los cuales corresponde a una de las sub-CU.

**[0030]** Cada nodo de la estructura de datos de árbol cuaternario puede proporcionar datos sintácticos para la CU correspondiente. Por ejemplo, un nodo en el árbol cuaternario puede incluir un indicador de división, que indica si la CU correspondiente al nodo está dividida en varias sub-CU. Los elementos sintácticos para una CU pueden definirse de manera recursiva y pueden depender de si la CU está dividida en varias sub-CU. Si una CU no está dividida adicionalmente, se denomina CU hoja. En esta divulgación, cuatro sub-CU de una CU hoja también se denominarán CU hojas incluso aunque no haya ninguna división explícita de la CU hoja original. Por ejemplo, si una CU con un tamaño de 16x16 no se divide adicionalmente, las cuatro sub-CU de tamaño 8x8 también se denominarán CU hojas aunque la CU de tamaño 16x16 no se haya dividido nunca.

**[0031]** Una CU tiene un propósito similar a un macrobloque de la norma H.264, excepto que una CU no tiene una distinción de tamaño. Por ejemplo, un bloque arbolado puede dividirse en cuatro nodos secundarios (también denominados sub-CU) y cada nodo secundario puede a su vez ser un nodo principal y dividirse en otros cuatro nodos secundarios. Un nodo secundario final, no dividido, denominado nodo hoja del árbol cuaternario, comprende un nodo de codificación, también denominado CU hoja. Los datos sintácticos asociados a un flujo de bits codificado pueden definir un número máximo de veces en que puede dividirse un bloque arbolado, denominado profundidad de CU máxima, y también pueden definir un tamaño mínimo de los nodos de codificación. Por consiguiente, un flujo de bits también puede definir una unidad de codificación de tamaño mínimo (SCU). Esta divulgación utiliza el término "bloque" para referirse a cualquiera entre una CU, PU o TU, en el contexto de la HEVC, o a estructuras de datos similares en el contexto de otras normas (por ejemplo, macrobloques y sub-bloques de los mismos en la norma H.264/AVC).

**[0032]** Una CU incluye un nodo de codificación y unidades de predicción (PU) y unidades de transformada (TU) asociadas al nodo de codificación. Un tamaño de la CU corresponde a un tamaño del nodo de codificación y debe ser de forma cuadrada. El tamaño de la CU puede variar desde 8x8 píxeles hasta el tamaño del bloque de árbol, con un máximo de 64x64 píxeles o más. Cada CU puede contener una o más PU y una o más TU. Los datos sintácticos asociados a una CU pueden describir, por ejemplo, la división de la CU en una o más PU. Los modos de división pueden diferir dependiendo de si la CU está codificada en modo de salto o directo, codificada en modo de predicción intra o codificada en modo de predicción inter. Las PU pueden dividirse para tener forma no cuadrada. Los datos sintácticos asociados a una CU también pueden describir, por ejemplo, la división de la CU en una o más TU de acuerdo con un árbol cuaternario. Una TU puede tener forma cuadrada o no cuadrada (por ejemplo, rectangular).

**[0033]** La norma HEVC soporta transformaciones de acuerdo con las TU, que pueden ser diferentes para diferentes CU. El tamaño de las TU se basa típicamente en el tamaño de las PU dentro de una CU definida dada para una LCU particionada, aunque puede que no sea siempre así. Las TU son típicamente del mismo tamaño o de un tamaño más pequeño que las PU. En algunos ejemplos, las muestras residuales correspondientes a una CU pueden subdividirse en unidades más pequeñas mediante una estructura de árbol cuaternario conocida como árbol cuaternario residual (RQT). Los nodos hoja del RQT pueden denominarse unidades de transformada (TU). Los valores de diferencia de píxel asociados a las TU pueden transformarse para generar coeficientes de transformación, que pueden cuantificarse.

**[0034]** Una CU hoja puede incluir una o más unidades de predicción (PU). En general, una PU representa una zona espacial correspondiente a la totalidad, o a una parte, de la CU correspondiente, y puede incluir datos para recuperar una muestra de referencia para la PU. Además, una PU incluye datos relacionados con la predicción. Por ejemplo, cuando la PU está codificada en modo intra, pueden incluirse datos para la PU en un árbol cuaternario residual (RQT), que pueden incluir datos que describen un modo de predicción intra para una TU correspondiente a la PU. Como otro ejemplo, cuando la PU está codificada en modo inter, la PU puede incluir datos que definen uno o más vectores de movimiento para la PU. Los datos que definen el vector de movimiento para una PU pueden describir, por ejemplo, un componente horizontal del vector de movimiento, un componente vertical del vector de movimiento, una resolución para el vector de movimiento (por ejemplo, precisión de píxeles de un cuarto o precisión de píxeles de un octavo), una imagen de referencia a la que apunta el vector de movimiento y/o una lista de imágenes de referencia (por ejemplo, la Lista 0 o la Lista 1) para el vector de movimiento.

**[0035]** Una CU hoja que tiene una o más PU también puede incluir una o más unidades de transformada (TU). Las unidades de transformada pueden especificarse usando un RQT (también denominado estructura de árbol cuaternario de TU), como se ha expuesto anteriormente. Por ejemplo, un indicador de división puede indicar si una CU hoja está dividida en cuatro unidades de transformada. A continuación, cada unidad de transformada puede dividirse adicionalmente en sub-TU adicionales. Cuando una TU no se divide adicionalmente, puede denominarse una TU hoja. En general, en lo que respecta a la intra-codificación, todas las TU hojas que pertenecen a una CU hoja comparten el mismo modo de predicción intra. Es decir, el mismo modo de predicción intra se aplica en general para calcular valores predichos para todas las TU de una CU hoja. En lo que respecta a la intra-codificación, un codificador de vídeo puede calcular un valor residual para cada TU hoja usando el modo de predicción intra, como una diferencia entre la parte de la CU correspondiente a la TU y el bloque original. Una TU no está necesariamente limitada al tamaño de una PU. De este modo, las TU pueden ser mayores o menores que una PU. En lo que respecta a la intra-codificación, una PU puede estar co-situada con una TU hoja correspondiente para la misma CU. En algunos ejemplos, el tamaño máximo de una TU hoja puede corresponder al tamaño de la CU hoja correspondiente.

**[0036]** Además, las TU de las CU hojas también pueden asociarse a las respectivas estructuras de datos de árbol cuaternario, denominadas árboles cuaternarios residuales (RQT). Es decir, una CU hoja puede incluir un árbol cuaternario que indica cómo la CU hoja está dividida en las TU. El nodo raíz de un árbol cuaternario de TU corresponde en general a una CU hoja, mientras que el nodo raíz de un árbol cuaternario de CU corresponde en general a un bloque arbolado (o LCU). Las TU del RQT que no están divididas se denominan TU hojas. En general, esta divulgación usa los términos CU y TU para hacer referencia a una CU hoja y a una TU hoja, respectivamente, a no ser que se indique lo contrario.

**[0037]** Una secuencia de vídeo incluye típicamente una serie de tramas o imágenes de vídeo. Un grupo de imágenes (GOP) comprende, en general, una serie de una o más de las imágenes de vídeo. Un GOP puede incluir datos sintácticos en una cabecera de GOP, en una cabecera de una o más de las imágenes, o en otras ubicaciones, que describen un número de imágenes incluidas en el GOP. Cada fragmento de una imagen puede incluir datos sintácticos de fragmento que describen un modo de codificación para el fragmento respectivo. El codificador 20 de vídeo actúa típicamente sobre bloques de vídeo dentro de fragmentos de vídeo individuales con el fin de codificar los datos de vídeo. Un bloque de vídeo puede corresponder a un nodo de codificación de una CU. Los bloques de vídeo pueden tener tamaños fijos o variables y pueden diferir en tamaño de acuerdo con una norma de codificación especificada.

**[0038]** Como ejemplo, el HM presta soporte a la predicción en diversos tamaños de PU. Si se supone que el tamaño de una CU particular es 2Nx2N, el HM admite predicción intra en tamaños de PU de 2Nx2N o NxN y predicción inter en tamaños de PU simétricas de 2Nx2N, 2NxN, Nx2N o NxN. El HM también admite la división asimétrica para predicción inter en tamaños de PU de 2NxN<sub>U</sub>, 2NxN<sub>D</sub>, nLx2N y nRx2N. En la división asimétrica, una dirección de una CU no está dividida, mientras que la otra dirección está dividida en un 25 % y un 75 %. La parte de la CU

correspondiente a la división del 25 % está indicada por una "n" seguida por una indicación de "arriba", "abajo", "izquierda" o "derecha". Así pues, por ejemplo, "2Nx2N" se refiere a una CU 2Nx2N que está dividida horizontalmente y tiene una PU 2Nx0,5N encima y una PU 2Nx1,5N debajo.

**[0039]** En esta divulgación, "NxN" y "N por N" pueden usarse indistintamente para hacer referencia a las dimensiones de píxeles de un bloque de vídeo en términos de dimensiones verticales y horizontales, por ejemplo, 16x16 píxeles o 16 por 16 píxeles. En general, un bloque de 16x16 tendrá 16 píxeles en una dirección vertical ( $y = 16$ ) y 16 píxeles en una dirección horizontal ( $x = 16$ ). Asimismo, un bloque de NxN tiene, en general, N píxeles en una dirección vertical y N píxeles en una dirección horizontal, donde N representa un valor entero no negativo. Los píxeles de un bloque se pueden disponer en filas y columnas. Además, no es necesario que los bloques tengan necesariamente el mismo número de píxeles en la dirección horizontal y en la dirección vertical. Por ejemplo, los bloques pueden comprender NxM píxeles, donde M no es necesariamente igual a N.

**[0040]** Tras la codificación de predicción intra o predicción inter mediante las PU de una CU, el codificador de vídeo 20 puede calcular los datos residuales para las TU de la CU. Las PU pueden comprender datos sintácticos que describen un procedimiento o modo de generación de datos de píxeles predictivos en el dominio espacial (también denominado el dominio de píxeles) y las TU pueden comprender coeficientes en el dominio de transformada, tras la aplicación de una transformada, por ejemplo, una transformada de coseno discreta (DCT), una transformada entera, una transformada de ondículas o una transformada conceptualmente similar, a los datos de vídeo residuales. Los datos residuales pueden corresponder a diferencias de píxeles entre píxeles de la imagen no codificada y los valores de predicción correspondientes a las PU. El codificador de vídeo 20 puede formar las TU incluyendo los datos residuales para la CU y, a continuación, transformar las TU para generar coeficientes de transformada para la CU.

**[0041]** Tras cualquier transformada para generar coeficientes de transformada, el codificador de vídeo 20 puede realizar la cuantificación de los coeficientes de transformada. La cuantificación se refiere, en general, a un proceso en el que los coeficientes de transformación se cuantifican para reducir posiblemente la cantidad de datos usados para representar los coeficientes, proporcionando compresión adicional. El proceso de cuantificación puede reducir la profundidad de bits asociada a algunos, o a la totalidad, de los coeficientes. Por ejemplo, un valor de  $n$  bits puede redondearse a la baja hasta un valor de  $m$  bits durante la cuantificación, donde  $n$  es mayor que  $m$ .

**[0042]** Después de la cuantificación, el codificador de vídeo puede recorrer los coeficientes de transformada, produciendo un vector unidimensional a partir de la matriz bidimensional que incluye los coeficientes de transformada cuantificados. El recorrido puede estar diseñado para colocar los coeficientes de energía más alta (y por lo tanto de menor frecuencia) en la parte frontal de la matriz y para colocar los coeficientes de energía más baja (y por lo tanto de mayor frecuencia) en la parte posterior de la matriz. En algunos ejemplos, el codificador 20 de vídeo puede utilizar un orden de escaneo predefinido para examinar los coeficientes de transformación cuantificados para producir un vector en serie que se pueda someter a una codificación por entropía. En otros ejemplos, el codificador 20 de vídeo puede realizar un escaneo adaptativo. Después de recorrer los coeficientes de transformada cuantificados para formar un vector unidimensional, el codificador de vídeo 20 puede realizar la codificación por entropía del vector unidimensional, por ejemplo, de acuerdo con la codificación de longitud variable adaptable al contexto (CAVLC), la codificación aritmética binaria adaptable al contexto (CABAC), la codificación aritmética binaria adaptable al contexto basada en la sintaxis (SBAC), la codificación por entropía por división de intervalos de probabilidad (PIPE) o con otra metodología de codificación por entropía. El codificador 20 de vídeo también puede realizar la codificación por entropía de los elementos sintácticos asociados a los datos de vídeo codificados, para su uso por el decodificador de vídeo 30 en la descodificación de los datos de vídeo.

**[0043]** Para realizar la CABAC, el codificador de vídeo 20 puede asignar un contexto dentro de un modelo contextual a un símbolo que se va a transmitir. El contexto se puede referir, por ejemplo, a si los valores contiguos del símbolo son distintos de cero o no. Para realizar la CAVLC, el codificador 20 de vídeo puede seleccionar un código de longitud variable para un símbolo que se vaya a transmitir. Las palabras de código en la VLC se pueden construir de modo que los códigos relativamente más cortos correspondan a símbolos más probables, mientras que los códigos más largos correspondan a símbolos menos probables. De esta forma, el uso de la VLC puede lograr un ahorro en bits con respecto, por ejemplo, al uso de palabras de código de igual longitud para cada símbolo que se vaya a transmitir. La determinación de la probabilidad se puede basar en un contexto asignado al símbolo.

**[0044]** Además, el codificador de vídeo 20 puede enviar datos sintácticos, tales como datos sintácticos basándose en bloques, datos sintácticos basándose en tramas y datos sintácticos basándose en GOP, al decodificador de vídeo 30, por ejemplo, en una cabecera de trama, una cabecera de bloque, una cabecera de fragmento o una cabecera de GOP. Los datos sintácticos de GOP pueden describir un cierto número de tramas en el respectivo GOP, y los datos sintácticos de trama pueden indicar un modo de codificación/predicción utilizada para codificar la trama correspondiente.

**[0045]** El codificador de vídeo 20 y el decodificador de vídeo 30 pueden implementarse, cada uno, como cualquiera entre una amplia variedad de circuitos de codificadores o decodificadores adecuados, según corresponda, tales como uno o más microprocesadores, procesadores de señales digitales (DSP), circuitos integrados específicos de la aplicación (ASIC), formaciones de puertas programables en el terreno (FPGA), circuitos de lógica discreta, software,



hardware, firmware o cualquier combinación de los mismos. Tanto el codificador de vídeo 20 como el descodificador de vídeo 30 pueden estar incluidos en uno o más codificadores o descodificadores, cada uno de los cuales puede estar integrado como parte de un codificador/descodificador (códec) de vídeo combinado. Un dispositivo que incluye el codificador de vídeo 20 y/o el descodificador de vídeo 30 puede comprender un circuito integrado, un microprocesador y/o un dispositivo de comunicación inalámbrica, tal como un teléfono celular.

**[0046]** El codificador de vídeo 20 y el descodificador de vídeo 30 pueden configurarse para implementar una o más de las técnicas descritas en esta divulgación. Como ejemplo, el codificador de vídeo 20 puede configurarse para codificar una indicación de si una actualización de conjunto de parámetros puede ocurrir en una parte de un flujo de bits. El descodificador de vídeo 30 puede configurarse para descodificar la indicación para determinar si se puede producir una actualización de un conjunto de parámetros almacenados sin realizar una comparación de contenido entre el conjunto de parámetros almacenados y un nuevo conjunto de parámetros del mismo tipo con el mismo valor de identificación. Cuando se produce una actualización del conjunto de parámetros, el descodificador de vídeo 30 almacena un conjunto de parámetros actual con un valor de identificación dado para reemplazar un conjunto de parámetros anterior del mismo tipo y que tiene el mismo valor de identificación. La parte del flujo de bits a la que se aplica la indicación puede ser una secuencia de vídeo codificada del flujo de bits o un GOP en la secuencia de vídeo codificada del flujo de bits.

**[0047]** En un caso, la indicación puede comprender un elemento sintáctico, tal como un indicador, en el flujo de bits que indica si cualquier conjunto de parámetros de cualquier tipo se puede actualizar en una parte de un flujo de bits. En otro caso, la indicación puede comprender varios indicadores diferentes en el flujo de bits, cada uno de los cuales indica si un conjunto de parámetros de un tipo particular puede actualizarse en la parte del flujo de bits. En cualquier caso, cuando un indicador indica que no se puede realizar una actualización del conjunto de parámetros, el descodificador de vídeo 30 puede almacenar y activar un único conjunto de parámetros de un tipo dado para la parte completa del flujo de bits. Al recibir un conjunto de parámetros posterior del mismo tipo y tener el mismo valor de identificación que el conjunto de parámetros inicial, el descodificador de vídeo 30 puede ignorar el conjunto de parámetros posterior. A la inversa, cuando un indicador indica que puede producirse una actualización del conjunto de parámetros, el descodificador de vídeo 30 puede activar automáticamente un nuevo conjunto de parámetros y desactivar un conjunto de parámetros almacenado de un tipo determinado, o puede determinar si actualizar el parámetro almacenado, por ejemplo, realizando una comparación de contenido.

**[0048]** Como otro ejemplo, el codificador de vídeo 20 puede estar configurado para codificar una indicación de si realizar acceso aleatorio a partir de una unidad de acceso CRA particular en un flujo de bits requiere obtener conjuntos de parámetros a partir de unidades de acceso anteriores. El descodificador de vídeo 30 puede configurarse para descodificar la indicación para determinar si los conjuntos de parámetros incluidos en las unidades de acceso anteriores son necesarios para realizar un acceso aleatorio desde una imagen CRA particular. En algunos casos, la indicación puede comprender un elemento sintáctico, como un indicador, incluido en una carga útil de unidad NAL incluida en una unidad de acceso para cada imagen CRA en el flujo de bits.

**[0049]** La unidad de acceso CRA se puede producir en cualquier punto dentro de una secuencia de vídeo codificada y no limpiar la DPB, la cual proporciona más opciones para el acceso aleatorio al flujo de bits sin la menor eficiencia de codificación. En el caso de acceso aleatorio desde una imagen CRA, todas las imágenes que preceden a la imagen CRA en orden de salida o descodificación no se descodificarán y no estarán disponibles para su uso como imágenes de referencia. Para descodificar la imagen CRA o cualquiera de las siguientes imágenes en orden de salida o descodificación, un descodificador de vídeo puede necesitar obtener uno o más conjuntos de parámetros incluidos en una unidad de acceso para una de las imágenes anteriores no disponibles, lo cual es una operación pesada. De acuerdo con algunas de las técnicas de esta divulgación, cuando el indicador indica que no se necesitan conjuntos de parámetros de unidades de acceso previas para el acceso aleatorio desde una imagen CRA particular, el descodificador de vídeo 30 puede determinar realizar un acceso aleatorio desde esa imagen porque no se requiere ninguna operación de búsqueda. A la inversa, cuando el indicador indica que los conjuntos de parámetros de las unidades de acceso anteriores son necesarios para el acceso aleatorio desde una imagen CRA particular, el descodificador de vídeo 30 puede determinar que no se realice un acceso aleatorio desde esa imagen para evitar la operación de búsqueda. Aunque se describe principalmente con respecto a las imágenes CRA, las técnicas se pueden usar de manera similar para otros tipos de imágenes RAP, incluidas las imágenes IDR y las imágenes de actualización de descodificación gradual (GDR).

**[0050]** La FIG. 2 es un diagrama de bloques que ilustra un ejemplo de codificador de vídeo 20 que puede implementar las técnicas descritas en esta divulgación. El codificador de vídeo 20 puede realizar intra-codificación e inter-codificación de bloques de vídeo dentro de fragmentos de vídeo. La intra-codificación se basa en la predicción espacial para reducir o eliminar la redundancia espacial en el vídeo dentro de una trama o imagen de vídeo dada. La inter-codificación se basa en la predicción temporal para reducir o eliminar la redundancia temporal en el vídeo dentro de tramas o imágenes adyacentes de una secuencia de vídeo. El modo intra (modo I) puede referirse a cualquiera de varios modos de codificación de base espacial. Los modos inter, tales como la predicción unidireccional (modo P) o la bipredicción (modo B), pueden referirse a cualquiera de varios modos de codificación de base temporal.

**[0051]** Como se muestra en la FIG. 2, el codificador de vídeo 20 recibe un bloque de vídeo actual dentro de una trama de vídeo para ser codificada. En el ejemplo de la FIG. 2, el codificador de vídeo 20 incluye una unidad de selección de modo 40, una memoria de imágenes de referencia 64, un sumador 50, una unidad de procesamiento de transformada 52, una unidad de cuantificación 54 y una unidad de codificación por entropía 56. A su vez, la unidad de selección de modo 40 incluye una unidad de compensación de movimiento 44, una unidad de estimación de movimiento 42, una unidad de procesamiento de predicción intra 46 y una unidad de división 48. Para la reconstrucción de bloques de vídeo, el codificador de vídeo 20 incluye también una unidad de cuantificación inversa 58, una unidad de procesamiento de transformada inversa 60 y un sumador 62.

**[0052]** Durante el proceso de codificación, el codificador de vídeo 20 recibe una trama o un fragmento de vídeo que va a codificarse. La trama o el fragmento pueden dividirse en múltiples bloques de vídeo. La unidad 42 de estimación de movimiento y la unidad 44 de compensación de movimiento llevan a cabo la codificación de predicción intra del bloque de vídeo recibido con respecto a uno o más bloques en una o más tramas de referencia, para proporcionar la predicción temporal. La unidad de procesamiento de predicción intra 46, de forma alternativa, puede llevar a cabo la codificación de predicción intra del bloque de vídeo recibido, con respecto a uno o más bloques contiguos de la misma trama o fragmento que el bloque que va a codificarse, para proporcionar predicción espacial. El codificador de vídeo 20 puede llevar a cabo múltiples pasadas de codificación, por ejemplo, para seleccionar un modo de codificación adecuada para cada bloque de datos de vídeo.

**[0053]** Además, la unidad de división 48 puede dividir bloques de datos de vídeo en sub-bloques, basándose en la evaluación de los anteriores esquemas de división en las pasadas de codificación anteriores. Por ejemplo, la unidad 48 de división puede dividir inicialmente una trama o un fragmento en varias LCU, y dividir cada una de las LCU en varias sub-CU, basándose en un análisis de velocidad-distorsión (por ejemplo, una optimización de velocidad-distorsión). La unidad de selección de modo 40 puede producir además una estructura de datos de árbol cuaternario, indicativa de la división de una LCU en las sub-CU. Las CU de nodos hoja del árbol cuaternario pueden incluir una o más PU y una o más TU.

**[0054]** La unidad de selección de modo 40 puede seleccionar uno de los modos de codificación, intra o inter, por ejemplo, basándose en los resultados de errores, y proporciona el bloque intra-codificado o inter-codificado resultante al sumador 50 para generar datos de bloque residuales y al sumador 62 para reconstruir el bloque codificado para su uso como una trama de referencia. La unidad de selección de modo 40 proporciona además elementos sintácticos, tales como vectores de movimiento, indicadores de modo intra, información de división y otra información sintáctica de este tipo, a la unidad de codificación por entropía 56.

**[0055]** La unidad de estimación de movimiento 42 y la unidad de compensación de movimiento 44 pueden estar altamente integradas, pero se ilustran por separado con fines conceptuales. La estimación del movimiento, realizada por la unidad 42 de estimación del movimiento, es el proceso de generación de vectores de movimiento, que estiman el movimiento de los bloques de vídeo. Un vector de movimiento, por ejemplo, puede indicar el desplazamiento de una PU de un bloque de vídeo dentro de una trama o imagen de vídeo actual, con respecto a un bloque predictivo dentro de una trama de referencia (u otra unidad codificada), con respecto al bloque actual que se está codificando dentro de la trama actual (u otra unidad codificada). Un bloque predictivo es un bloque que se revela como estrechamente coincidente con el bloque que se va a codificar, en lo que respecta a la diferencia de píxeles, lo cual puede determinarse mediante la suma de diferencias absolutas (SAD), suma de las diferencias al cuadrado (SSD) u otras métricas de diferencia. En algunos ejemplos, el codificador de vídeo 20 puede calcular valores para posiciones de píxel de subentero de imágenes de referencia almacenadas en la memoria de imágenes de referencia 64, que también puede denominarse memoria intermedia de imágenes descodificadas (DPB). Por ejemplo, el codificador de vídeo 20 puede interpolar valores de posiciones de un cuarto de píxel, posiciones de un octavo de píxel u otras posiciones fraccionarias de píxel de la imagen de referencia. Por lo tanto, la unidad 42 de estimación de movimiento puede realizar una búsqueda de movimiento relativa a las posiciones de píxeles completas y las posiciones de píxeles fraccionarias, y generar un vector de movimiento con una precisión de píxel fraccionaria.

**[0056]** La unidad de estimación de movimiento 42 calcula un vector de movimiento para una PU de un bloque de vídeo en un fragmento inter-codificado, comparando la posición de la PU con la posición de un bloque predictivo de una imagen de referencia. La imagen de referencia puede seleccionarse de una primera lista (lista 0) de imágenes de referencia o una segunda lista (lista 1) de imágenes de referencia, cada una de las cuales identifica una o más imágenes de referencia almacenadas en una memoria 64 de imágenes de referencia. La unidad 42 de estimación del movimiento envía el vector de movimiento calculado a la unidad 56 de codificación por entropía y a la unidad 44 de compensación de movimiento.

**[0057]** La compensación de movimiento, llevada a cabo por la unidad de compensación de movimiento 44, puede implicar capturar o generar el bloque predictivo basándose en el vector de movimiento determinado por la unidad de estimación de movimiento 42. De nuevo, la unidad 42 de estimación de movimiento y la unidad 44 de compensación de movimiento pueden integrarse funcionalmente, en algunos ejemplos. Tras recibir el vector de movimiento para la PU del bloque de vídeo actual, la unidad 44 de compensación de movimiento puede localizar el bloque predictivo al que apunta el vector de movimiento en una de las listas de imágenes de referencia. El sumador 50 forma un bloque de vídeo residual restando los valores de píxel del bloque predictivo a los valores de píxel del bloque de vídeo actual.

que se está codificando, generando valores de diferencias de píxel, como se analiza posteriormente. En general, la unidad de estimación de movimiento 42 lleva a cabo la estimación de movimiento con respecto a los componentes de luma, y la unidad de compensación de movimiento 44 utiliza los vectores de movimiento calculados basándose en los componentes de luma, tanto para los componentes de croma como para los componentes de luma. La unidad de selección de modo 40 también puede generar elementos sintácticos asociados a los bloques de vídeo y el fragmento de vídeo para su uso por el descodificador de vídeo 30 al descodificar los bloques de vídeo del fragmento de vídeo.

**[0058]** La unidad de procesamiento de predicción intra 46 puede realizar predicción intra de un bloque actual, de forma alternativa a la predicción inter llevada a cabo por la unidad de estimación de movimiento 42 y la unidad de compensación de movimiento 44, como se ha descrito anteriormente. En particular, la unidad de procesamiento de predicción intra 46 puede determinar un modo de predicción intra que se vaya a usar para codificar un bloque actual. En algunos ejemplos, la unidad de procesamiento de predicción intra 46 puede codificar un bloque actual usando diversos modos de predicción intra, por ejemplo, durante diferentes pases de codificación, y la unidad de procesamiento de predicción intra 46 (o la unidad de selección de modo 40, en algunos ejemplos) puede seleccionar un modo adecuado de predicción intra para su uso a partir de los modos probados.

**[0059]** Por ejemplo, la unidad de procesamiento de predicción intra 46 puede calcular valores de distorsión de velocidad usando un análisis de distorsión de velocidad para las diversos modos de predicción intra probados, y seleccionar el modo de predicción intra que tenga las mejores características de distorsión de velocidad entre los modos probados. El análisis de distorsión de velocidad determina en general una magnitud de distorsión (o de errores) entre un bloque codificado y un bloque original, no codificado, que se codificó para generar el bloque codificado, así como una velocidad binaria (es decir, un número de bits) usada para generar el bloque codificado. La unidad de procesamiento de predicción intra 46 puede calcular razones a partir de las distorsiones y velocidades para los diversos bloques codificados, para determinar qué modo de predicción intra presenta el mejor valor de distorsión de velocidad para el bloque.

**[0060]** Después de seleccionar un modo de predicción intra para un bloque, la unidad de procesamiento de predicción intra 46 puede proporcionar información, indicativa del modo de predicción intra seleccionado para el bloque, a la unidad de codificación por entropía 56. La unidad 56 de codificación por entropía puede codificar la información que indica el modo de predicción intra seleccionada. El codificador de vídeo 20 puede incluir datos de configuración en el flujo de bits transmitido, que pueden incluir una pluralidad de tablas de índices de modos de predicción intra y una pluralidad de tablas de índices de modos de predicción intra modificadas (también denominadas tablas de correlación de palabras de código), definiciones de contextos de codificación para diversos bloques e indicaciones de un modo de predicción intra más probable, una tabla de índices de modos de predicción intra y una tabla de índices de modos de predicción intra modificadas para su uso para cada uno de los contextos.

**[0061]** El codificador de vídeo 20 forma un bloque de vídeo residual restando los datos de predicción, de la unidad de selección de modo 40, del bloque de vídeo original que está codificado. El sumador 50 representa el componente o los componentes que realizan esta operación de resta. La unidad 52 de procesamiento de transformación aplica una transformada, tal como una transformada discreta de coseno (DCT) o una transformada conceptualmente similar, al bloque residual, generando un bloque de vídeo que comprende valores residuales de coeficientes de transformación. La unidad 52 de procesamiento de transformadas puede llevar a cabo otras transformadas que son conceptualmente similares a la DCT. También podrían usarse transformadas de ondículas, transformadas de números enteros, transformadas de subbandas u otros tipos de transformadas. En cualquier caso, la unidad 52 de procesamiento de transformación aplica la transformada al bloque residual, generando un bloque de coeficientes de transformación residuales. La transformada puede convertir la información residual, desde un dominio de valor de píxel a un dominio de transformada, tal como un dominio de frecuencia. La unidad 52 de procesamiento de transformación puede enviar los coeficientes de transformación resultantes a la unidad 54 de cuantificación. La unidad 54 de cuantificación cuantifica los coeficientes de transformación para reducir todavía más la velocidad de transmisión de bits. El proceso de cuantificación puede reducir la profundidad de bits asociada a algunos, o a la totalidad, de los coeficientes. El grado de cuantificación se puede modificar ajustando un parámetro de cuantificación. En algunos ejemplos, la unidad 54 de cuantificación puede realizar, a continuación, una exploración de la matriz que incluye los coeficientes de transformación cuantificados. De forma alternativa, la unidad 56 de codificación por entropía puede realizar la exploración.

**[0062]** Tras la cuantificación, la unidad de codificación por entropía 56 codifica por entropía los coeficientes de transformada cuantificados. Por ejemplo, la unidad 56 de codificación por entropía puede llevar a cabo la codificación de longitud variable adaptativa según el contexto (CAVLC), la codificación aritmética binaria adaptativa según el contexto (CABAC), la codificación aritmética binaria adaptativa según el contexto y basada en sintaxis (SBAC), la codificación por entropía mediante la división en intervalos de probabilidades (PIPE) u otra técnica de codificación por entropía. En el caso de la codificación por entropía basada en el contexto, el contexto puede basarse en bloques contiguos. Tras la codificación por entropía realizada por la unidad 56 de codificación por entropía, el flujo de bits codificado puede transmitirse a otro dispositivo (por ejemplo, el descodificador de vídeo 30) o archivar para su posterior transmisión o recuperación.

**[0063]** La unidad de cuantificación inversa 58 y la unidad de procesamiento de transformada inversa 60 aplican la cuantificación inversa y la transformación inversa, respectivamente, para reconstruir el bloque residual en el dominio del píxel, por ejemplo, para su posterior uso como bloque de referencia. La unidad 44 de compensación de movimiento puede calcular un bloque de referencia añadiendo el bloque residual a un bloque predictivo de una de las tramas de la memoria 64 de imágenes de referencia. La unidad 44 de compensación del movimiento también puede aplicar uno o más filtros de interpolación al bloque residual reconstruido para calcular valores fraccionarios de píxel para su uso en la estimación de movimiento. El sumador 62 añade el bloque residual reconstruido al bloque de predicción compensado por movimiento, generado por la unidad 44 de compensación de movimiento, para generar un bloque de vídeo reconstruido para su almacenamiento en la memoria 64 de imágenes de referencia.

**[0064]** El codificador de vídeo 20 de la FIG. 2 representa un ejemplo de un codificador de vídeo configurado para implementar una o más de las técnicas descritas en esta divulgación. En un ejemplo de las técnicas, el codificador de vídeo 20 puede configurarse para codificar una indicación de si una actualización de conjunto de parámetros puede ocurrir en una parte de un flujo de bits. El codificador de vídeo 20 puede generar la indicación para notificar a un descodificador de vídeo si se puede producir una actualización de un conjunto de parámetros almacenados. De esta manera, basándose en la indicación, el descodificador de vídeo primero determina si se permite una actualización del conjunto de parámetros en la parte del flujo de bits antes de determinar si se debe realizar una actualización, por ejemplo, realizando una comparación de contenido entre un conjunto de parámetros activos almacenados y un nuevo conjunto de parámetros del mismo tipo con el mismo valor de identificación.

**[0065]** En algunos casos, la indicación puede comprender una única indicación de si cualquiera de los conjuntos de parámetros para una parte de un flujo de bits, por ejemplo, una secuencia de vídeo codificada o un GOP, se puede actualizar. En este caso, se puede utilizar la misma indicación para todos los tipos de conjuntos de parámetros, incluidos los conjuntos de parámetros de vídeo (VPS), los conjuntos de parámetros de secuencia (SPS), los conjuntos de parámetros de imagen (PPS) y/o los conjuntos de parámetros adaptativos (APS). En otros casos, la indicación puede comprender varias indicaciones diferentes, cada una de las cuales indica si un conjunto de parámetros de un tipo particular puede actualizarse en la parte del flujo de bits. La indicación puede ser una indicación de nivel de flujo de bits, como un elemento sintáctico incluido en la parte del flujo de bits. Por ejemplo, la indicación puede ser un elemento sintáctico, como un indicador, incluido en las unidades de la capa de abstracción de red (NAL) de información de mejora suplementaria (SEI) en el flujo de bits. Como otro ejemplo, la indicación puede ser un elemento sintáctico, como un indicador, incluido en las unidades NAL de conjunto de parámetros de secuencia (SPS) en el flujo de bits.

**[0066]** En otros ejemplos, la indicación puede ser un requisito en una definición de perfil. En ejemplos adicionales, la indicación puede definirse como un parámetro de tipo de medio e incluirse en el protocolo de descripción de sesión (SDP) cuando el flujo de bits de vídeo se utiliza en un entorno de transmisión basado en el protocolo de transporte en tiempo real (RTP), como la transmisión continua o multidifusión, usando RTP para el transporte de datos multimedia. En este caso, la indicación puede usarse en una fase de negación de sesión, también conocida como intercambio de capacidad. En aún otros ejemplos, la indicación puede definirse como una propiedad de transmisión continua incluida en una descripción de presentación de medios (MPD) en un entorno de transmisión continua de protocolo de transferencia de hipertexto (HTTP), como DASH.

**[0067]** En otro ejemplo de las técnicas, el codificador de vídeo 20 puede estar configurado para codificar una indicación para cada imagen CRA de si realizar el acceso aleatorio a un flujo de bits de la unidad de acceso CRA requiere obtener conjuntos de parámetros a partir de unidades de acceso anteriores. El codificador de vídeo 20 puede generar la indicación para notificar a un descodificador de vídeo si los conjuntos de parámetros incluidos en las unidades de acceso anteriores son necesarios para realizar un acceso aleatorio desde una unidad de acceso CRA particular. De esta manera, basándose en la indicación, el descodificador de vídeo primero determina si se requiere una operación de búsqueda para la descodificación de acceso aleatorio desde la unidad de acceso CRA particular, y puede evitar realizar una operación de búsqueda seleccionando una imagen RAP diferente desde la cual realizar el acceso aleatorio. La indicación es especialmente útil para la unidad de acceso CRA introducida en la norma HEVC porque las imágenes CRA se pueden colocar en cualquier punto dentro de una secuencia de vídeo codificada y no limpiar la DPB, lo cual brinda más opciones para el acceso aleatorio al flujo de bits sin una menor eficiencia de codificación.

**[0068]** En algunos casos, la indicación puede comprender un tipo particular de unidad NAL, por ejemplo, un tipo particular de SPS, SEI o unidad NAL delimitadora de unidad de acceso, incluido en una unidad de acceso para cada imagen CRA en un flujo de bits. En otros casos, la indicación puede comprender un elemento sintáctico, como un indicador, incluido en una carga útil de unidad NAL, por ejemplo, una carga útil de unidad NAL delimitadora de unidad de acceso, SPS, SEI, o una cabecera de unidad NAL incluida en una unidad de acceso para cada imagen CRA en el flujo de bits. En un ejemplo, la indicación puede comprender un indicador incluido en una unidad SEI NAL en la unidad de acceso para la imagen CRA. Aunque se describe principalmente con respecto a las imágenes CRA, el codificador de vídeo 20 puede codificar indicaciones similares para otros tipos de imágenes RAP, incluidas las imágenes IDR y las imágenes de actualización de descodificación gradual (GDR).

**[0069]** La FIG. 3 es un diagrama de bloques que ilustra un ejemplo de descodificador de vídeo 30, que descodifica una secuencia de vídeo codificada. En el ejemplo de la FIG. 3, el descodificador de vídeo 30 incluye una unidad de descodificación por entropía 70, una unidad de compensación de movimiento 72, una unidad de procesamiento de

predicción intra 74, una unidad de cuantificación inversa 76, una unidad de procesamiento de transformada inversa 78, un sumador 80 y una memoria de imágenes de referencia 82. En algunos ejemplos, el descodificador de vídeo 30 puede realizar una pasada de descodificación, en general, recíproca a la pasada de codificación descrita con respecto al codificador de vídeo 20 (FIG. 2).

**[0070]** Durante el proceso de descodificación, el descodificador de vídeo 30 recibe un flujo de bits de vídeo codificado, que representa los bloques de vídeo de un fragmento de vídeo codificado y elementos sintácticos asociados, desde el codificador de vídeo 20. El descodificador de vídeo 30 puede recibir el flujo de bits de vídeo codificado desde una entidad 29 de red. La entidad de red 29 puede ser, por ejemplo, ser un servidor de transmisión continua, un elemento de red compatible con medios (MANE), un editor/empalmador de vídeo, un elemento de red intermedio u otro dispositivo configurado para implementar una o más de las técnicas descritas anteriormente. La entidad de red 29 puede incluir un medio externo configurado para realizar algunas de las técnicas de esta divulgación. Algunas de las técnicas descritas en esta divulgación pueden ser implementadas por la entidad de red 29 antes de que la entidad de red 29 transmita el flujo de bits de vídeo codificado al descodificador de vídeo 30. En algunos sistemas de descodificación de vídeo, la entidad 29 de red y el descodificador de vídeo 30 pueden ser partes de dispositivos separados, mientras que en otros casos la función descrita con respecto a la entidad 29 de red puede ser realizada por el mismo dispositivo que comprende el descodificador de vídeo 30.

**[0071]** La unidad de descodificación por entropía 70 del descodificador de vídeo 30 descodifica por entropía el flujo de bits para generar coeficientes cuantificados, vectores de movimiento o indicadores de modo de predicción intra y otros elementos sintácticos. La unidad de descodificación por entropía 70 remite los vectores de movimiento y otros elementos sintácticos a la unidad de compensación de movimiento 72. El descodificador de vídeo 30 puede recibir los elementos sintácticos en el nivel de fragmento de vídeo y/o el nivel de bloque de vídeo.

**[0072]** Cuando el fragmento de vídeo se codifica como un fragmento intra-codificado (I), la unidad de procesamiento de predicción intra 74 puede generar datos de predicción para un bloque de vídeo del fragmento de vídeo actual, basándose en un modo de predicción intra señalado, y datos de bloques previamente descodificados de la trama o imagen actual. Cuando la trama de vídeo se codifica como un fragmento inter-codificado (es decir, B o P), la unidad de compensación de movimiento 72 genera bloques predictivos para un bloque de vídeo del fragmento de vídeo actual basándose en los vectores de movimiento y en otros elementos sintácticos recibidos desde la unidad de descodificación por entropía 70. Los bloques predictivos se pueden producir a partir de una de las imágenes de referencia dentro de una de las listas de imágenes de referencia. El descodificador de vídeo 30 puede construir las listas de tramas de referencia, Lista 0 y Lista 1, basándose en las imágenes de referencia almacenadas en la memoria de tramas de referencia 82, que puede denominarse una memoria intermedia de imágenes descodificadas (DPB).

**[0073]** La unidad de compensación de movimiento 72 determina la información de predicción para un bloque de vídeo del fragmento de vídeo actual, analizando los vectores de movimiento y otros elementos sintácticos, y usa la información de predicción para producir los bloques predictivos del bloque de vídeo actual que se está descodificando. Por ejemplo, la unidad de compensación de movimiento 72 usa algunos de los elementos sintácticos recibidos para determinar un modo de predicción (por ejemplo, intra o predicción inter), usado para codificar los bloques de vídeo del fragmento de vídeo, un tipo de fragmento de predicción inter (por ejemplo, fragmento B o fragmento P), información de construcción para una o más de las listas de imágenes de referencia del fragmento, vectores de movimiento para cada bloque de vídeo inter-codificado del fragmento, el estado de predicción inter para cada bloque de vídeo inter-codificado del fragmento y otra información, para descodificar los bloques de vídeo en el fragmento de vídeo actual.

**[0074]** La unidad de compensación de movimiento 72 también puede realizar la interpolación basándose en filtros de interpolación. La unidad 72 de compensación de movimiento puede usar filtros de interpolación como los usados por el codificador de vídeo 20 durante la codificación de los bloques de vídeo, para calcular valores interpolados para píxeles fraccionarios de los bloques de referencia. En este caso, la unidad 72 de compensación de movimiento puede determinar los filtros de interpolación usados por el codificador de vídeo 20 a partir de los elementos sintácticos recibidos y usar los filtros de interpolación para generar bloques predictivos.

**[0075]** La unidad de cuantificación inversa 76 cuantifica de manera inversa, es decir, descuantifica, los coeficientes de transformada cuantificados, proporcionados en el flujo de bits y descodificados por la unidad de descodificación por entropía 70. El proceso de cuantificación inversa puede incluir el uso de un parámetro de cuantificación  $QP_V$ , calculado por el descodificador de vídeo 30 para cada bloque de vídeo en el fragmento de vídeo para determinar un grado de cuantificación y, asimismo, un grado de cuantificación inversa que debería aplicarse. La unidad de procesamiento de transformada inversa 78 aplica una transformada inversa, por ejemplo una DCT inversa, una transformada inversa de enteros o un proceso de transformada inversa conceptualmente similar, a los coeficientes de transformada, con el fin de generar bloques residuales en el dominio de píxeles.

**[0076]** Después de que la unidad de compensación de movimiento 72 genera el bloque predictivo para el bloque de vídeo actual, basándose en los vectores de movimiento y otros elementos sintácticos, el descodificador de vídeo 30 forma un bloque de vídeo descodificado sumando los bloques residuales procedentes de la unidad de transformación inversa 78 a los correspondientes bloques predictivos generados por la unidad de compensación de movimiento 72. El sumador 80 representa el componente o los componentes que realizan esta operación de suma. Si se desea,

también se puede aplicar un filtro de eliminación de bloques para filtrar los bloques descodificados con el fin de eliminar distorsiones de efecto pixelado. También se pueden usar otros filtros de bucle (en el bucle de codificación o bien después del bucle de codificación) para suavizar las transiciones de píxeles o mejorar de otro modo la calidad del vídeo. Los bloques de vídeo descodificados de una trama o imagen determinada se almacenan a continuación en la memoria 82 de imágenes de referencia, que almacena imágenes de referencia usadas para una compensación de movimiento posterior. La memoria de tramas de referencia 82 almacena también vídeo descodificado para su presentación posterior en un dispositivo de visualización, tal como el dispositivo de visualización 32 de la FIG. 1.

**[0077]** El descodificador de vídeo 30 de la FIG. 3 representa un ejemplo de un descodificador de vídeo configurado para implementar una o más de las técnicas descritas en esta divulgación. En un ejemplo de las técnicas, el descodificador de vídeo 30 puede configurarse para descodificar una indicación de si una actualización del conjunto de parámetros puede ocurrir en una parte de un flujo de bits. Cuando la indicación indica que una actualización del conjunto de parámetros no puede ocurrir en una parte de un flujo de bits, el descodificador de vídeo 30 activa un conjunto de parámetros inicial de un tipo dado para la parte completa del flujo de bits. En este caso, al recibir un conjunto de parámetros posterior del mismo tipo y tener el mismo valor de identificación que el conjunto de parámetros inicial, el descodificador de vídeo 30 no necesita determinar si actualizar el conjunto de parámetros inicial para la secuencia de vídeo codificada y puede ignorar el conjunto de parámetros posteriores.

**[0078]** El HEVC HM incluye el concepto de conjuntos de parámetros. Un conjunto de parámetros es una estructura sintáctica que incluye elementos sintácticos configurados para indicar información que permite a un descodificador de vídeo reconstruir vídeo codificado. Los elementos sintácticos pueden incluirse en un tipo particular de conjunto de parámetros basado en la información indicada por los elementos sintácticos y la frecuencia con la que se espera que los elementos sintácticos cambien en el flujo de bits. Por ejemplo, un conjunto de parámetros de vídeo (VPS) incluye elementos sintácticos que describen las características generales de las secuencias de vídeo codificadas, incluidas las dependencias entre subcapas temporales. Un conjunto de parámetros de secuencia (SPS) incluye elementos sintácticos que se espera que permanezcan sin cambios para las imágenes dentro de una secuencia de imágenes (por ejemplo, orden de imágenes, número de imágenes de referencia y tamaño de imagen). Un conjunto de parámetros de imagen (PPS) incluye elementos sintácticos que pueden cambiar de imagen a imagen dentro de una secuencia (por ejemplo, modo de codificación por entropía, parámetros de cuantificación y profundidad de bits). Un conjunto de parámetros adaptativos (APS) incluye elementos sintácticos que se espera que cambien dentro de las imágenes de una secuencia (por ejemplo, tamaño de bloque y filtrado de desbloqueo).

**[0079]** El concepto de conjunto de parámetros desacopla la transmisión de información que cambia poco frecuentemente de la transmisión de datos de vídeo codificados. En algunos ejemplos, los conjuntos de parámetros pueden transmitirse "fuera de banda" (es decir, no transportarse junto con las unidades que contienen datos de vídeo codificados). Puede ser deseable transmitir los conjuntos de parámetros fuera de banda, ya que la transmisión fuera de banda es típicamente más fiable que la transmisión "dentro de banda" dentro del canal de datos. En el HEVC HM, cada VPS incluye un valor de identificación (ID) del VPS, cada SPS incluye un ID de SPS y hace referencia a un VPS que utiliza un ID de VPS, y cada PPS incluye un ID de PPS y hace referencia a un SPS utilizando un ID de SPS. Además, cada cabecera de fragmento hace referencia a un PPS que utiliza un ID de PPS.

**[0080]** En un flujo de bits, cuando un conjunto de parámetros de un tipo particular con un valor de ID de conjunto de parámetros en particular tiene un contenido diferente que un conjunto de parámetros anterior (en orden de flujo de bits o en orden de descodificación) del mismo tipo con el mismo valor de ID, el tipo particular de conjunto de parámetros con ese valor de ID particular puede actualizarse. Cuando se produce una actualización del conjunto de parámetros, el descodificador de vídeo 30 almacena el conjunto de parámetros actual con el valor de ID dado para reemplazar el conjunto de parámetros anterior del mismo tipo y que tiene el mismo valor de ID.

**[0081]** La actualización del conjunto de parámetros puede ocurrir cuando se necesita un nuevo conjunto de parámetros de un tipo particular, pero todos los posibles valores de ID se han utilizado. La actualización del conjunto de parámetros también puede ocurrir en flujos de bits empalmados incluso cuando los espacios de valor de ID del conjunto de parámetros no se utilizan completamente. El empalme del flujo de bits se refiere a la combinación de dos o más flujos de bits o partes de flujos de bits. Un primer flujo de bits puede ser anexado por un segundo flujo de bits, posiblemente con algunas modificaciones a uno o ambos flujos de bits para generar un flujo de bits empalmado. La primera imagen codificada en el segundo flujo de bits también se denomina punto de empalme. La imagen en el punto de empalme en el flujo de bits empalmado se originó a partir del segundo flujo de bits, mientras que la imagen inmediatamente anterior al punto de empalme en el flujo de bits empalmado se originó a partir del primer flujo de bits.

**[0082]** Los flujos de bits típicamente utilizan unos pocos conjuntos de parámetros de cada tipo, por ejemplo, unos pocos SPS y unos pocos PPS, y en ocasiones utilizan solo un conjunto de parámetros de cada tipo. Además, cada uno de los flujos de bits típicamente comienza con el valor de ID de conjunto de parámetros más pequeño 0 para cada uno de los diferentes tipos de conjuntos de parámetros. De esta manera, si se empalman dos flujos de bits, es probable que tanto el punto de empalme como la imagen que precede inmediatamente a la imagen del punto de empalme se refieran al mismo valor de ID para un tipo particular de conjunto de parámetros, por ejemplo, SPS con valor de ID 0, pero que los dos conjuntos de parámetros incluyan contenido diferente. Por lo tanto, el conjunto de parámetros al que hace referencia la imagen que precede al punto de empalme debe actualizarse con el conjunto de parámetros del

mismo tipo con el mismo valor de ID al que se hace referencia en la imagen del punto de empalme. En este caso, los conjuntos de parámetros del flujo de bits empalmado no pueden transmitirse al comienzo del flujo de bits empalmado y, por lo tanto, no pueden transmitirse fuera de banda desde la secuencia de vídeo codificada del flujo de bits.

**[0083]** Convencionalmente, con el fin de determinar si un conjunto de parámetros actual de cualquier tipo con un valor de ID en particular debe ser activado para actualizar un parámetro anterior establecido con el mismo valor de ID, un descodificador de vídeo compara contenido del conjunto de parámetros actual con el contenido del conjunto de parámetros anterior para determinar si el contenido es diferente. Cuando el contenido del conjunto de parámetros actual es diferente al conjunto de parámetros anterior, el descodificador de vídeo activa el conjunto de parámetros actual para el flujo de bits y desactiva el conjunto de parámetros anterior. El descodificador de vídeo realiza esta comparación de contenido cada vez que recibe un nuevo conjunto de parámetros de un tipo particular que tiene un valor de ID particular que es el mismo que un conjunto de parámetros recibidos anteriormente. Cuando los conjuntos de parámetros se transmiten en banda, es decir, con una secuencia de vídeo codificada del flujo de bits, el descodificador de vídeo puede recibir transmisiones repetidas de los conjuntos de parámetros para mejorar la resistencia al error. El descodificador de vídeo luego realiza la comparación de contenido para cada conjunto de parámetros repetidos. Esta comparación de contenido es innecesaria e impone una carga computacional al descodificador de vídeo.

**[0084]** Las técnicas descritas en esta divulgación proporcionan un indicador configurado para indicar si una actualización del conjunto de parámetros puede ocurrir en una parte de un flujo de bits (por ejemplo, una secuencia de vídeo codificada o un GOP en la secuencia de vídeo codificada). Cuando el indicador indica que la actualización del conjunto de parámetros no puede ocurrir en la parte del flujo de bits, el descodificador de vídeo 30 almacena y activa solo un primer conjunto de parámetros con un valor de ID particular para la parte completa del flujo de bits. Además, cuando el indicador indica que no se puede realizar la actualización del conjunto de parámetros, el descodificador de vídeo 30 no realizará ninguna actualización del conjunto de parámetros y, por lo tanto, no necesita comparar el contenido del primer conjunto de parámetros con el contenido de un segundo conjunto de parámetros del mismo tipo con el mismo valor de ID. De esta manera, el descodificador de vídeo 30 puede funcionar como si el contenido del primer conjunto de parámetros fuera idéntico al contenido incluido en cualquier otro conjunto de parámetros del mismo tipo que tenga el mismo valor de ID.

**[0085]** Cuando el indicador indica que una actualización de conjunto de parámetros puede ocurrir en la parte del flujo de bits, el descodificador de vídeo 30 almacena y activa un primer conjunto de parámetros con un valor de ID en particular, pero puede actualizar el primer conjunto de parámetros almacenado usando un segundo conjunto de parámetros del mismo tipo con el mismo valor de ID para reemplazar el primer conjunto de parámetros almacenado. En algunos casos, el descodificador de vídeo 30 puede activar automáticamente el segundo conjunto de parámetros para la parte del flujo de bits, y desactivar el primer conjunto de parámetros. En otros casos, el descodificador de vídeo 30 puede comparar el contenido del primer conjunto de parámetros con el contenido del segundo conjunto de parámetros. Cuando el contenido del segundo conjunto de parámetros es diferente al contenido del primer conjunto de parámetros, el descodificador de vídeo 30 activa el segundo conjunto de parámetros para la parte del flujo de bits, y desactiva el primer conjunto de parámetros.

**[0086]** La indicación puede comprender un elemento sintáctico de nivel de flujo de bits, tal como un indicador, incluido en el flujo de bits. Por ejemplo, la indicación puede comprender un indicador incluido en una o más unidades de capa de abstracción de red (NAL) de información de mejora suplementaria (SEI) asociadas con una secuencia de vídeo codificada en el flujo de bits. Como otro ejemplo, la indicación puede comprender un indicador incluido en una o más unidades NAL de conjunto de parámetros de secuencia (SPS) para la secuencia de vídeo codificada en el flujo de bits.

**[0087]** En un ejemplo, el indicador puede indicar si cualquier conjunto de parámetros puede ser actualizado en la parte del flujo de bits. Por lo tanto, se puede usar un único indicador para indicar el estado de actualización para diferentes tipos de conjuntos de parámetros. En este caso, se puede usar el mismo indicador para todos los tipos de conjuntos de parámetros, incluidos VPS, SPS, PPS y/o APS. Cuando el indicador es igual a 1, se puede actualizar un conjunto de parámetros. Cuando el indicador es igual a 0, no se actualizará ningún conjunto de parámetros y el descodificador de vídeo 30 no realiza comparaciones de contenido entre los conjuntos de parámetros del mismo tipo y con los mismos valores de ID de conjunto de parámetros. Cuando el indicador es igual a 0, las comparaciones de contenido no son necesarias porque el descodificador de vídeo 30 puede funcionar como si los conjuntos de parámetros del mismo tipo y con los mismos valores de ID tuvieran el mismo contenido.

**[0088]** En otro ejemplo, el indicador puede indicar si un conjunto de parámetros de un tipo particular puede ser actualizada en la parte del flujo de bits. En este caso, se puede usar un indicador diferente para cada uno de los diferentes tipos de conjuntos de parámetros, incluidos VPS, SPS, PPS y APS. Por lo tanto, cada tipo de conjunto de parámetros puede, en efecto, tener su propio indicador para indicar si se puede actualizar. Cuando el indicador es igual a 1, se puede actualizar un conjunto de parámetros del tipo particular. Cuando el indicador es igual a 0, no se actualizará ningún conjunto de parámetros del tipo particular y el descodificador de vídeo 30 no realiza comparaciones de contenido entre los conjuntos de parámetros del tipo particular y con los mismos valores de ID de conjunto de parámetros. Cuando el indicador es igual a 0, las comparaciones de contenido no son necesarias porque el

descodificador de vídeo 30 puede funcionar como si los conjuntos de parámetros del tipo particular y con los mismos valores de ID tuvieran el mismo contenido.

**[0089]** En un ejemplo adicional, el indicador puede indicar si dos o más conjuntos de parámetros de los diferentes tipos pueden ser actualizados en la parte del flujo de bits. En este caso, se puede usar un indicador para dos o más de los diferentes tipos de conjuntos de parámetros, por ejemplo, SPS y PPS, y otro indicador se puede usar para uno o más de los diferentes tipos de conjuntos de parámetros, por ejemplo, VPS. Por lo tanto, algunos tipos de conjuntos de parámetros pueden compartir un indicador para indicar el estado de actualización conjunta, mientras que otros tipos de conjuntos de parámetros pueden tener indicadores individuales para indicar el estado de actualización individual. Además, en algunos casos, la indicación puede comprender una máscara de bits o una palabra de código que es un elemento sintáctico único con varios bits, donde cada bit corresponde al estado de actualización de uno o más de los conjuntos de parámetros, o una palabra de código tiene un valor particular para Indica diferentes combinaciones de los estados de actualización para los diferentes conjuntos de parámetros.

**[0090]** En algunos casos, la indicación de si una actualización del conjunto de parámetros puede ocurrir en la parte del flujo de bits puede determinar cómo los conjuntos de parámetros se transmiten al descodificador de vídeo 30. Por ejemplo, cuando las actualizaciones de conjuntos de parámetros pueden ocurrir en la parte del flujo de bits, los conjuntos de parámetros pueden no ser transmitidos fuera de banda desde la secuencia de vídeo codificada del flujo de bits porque las actualizaciones pueden ocurrir para diferentes partes del flujo de bits, por ejemplo, en el empalme de imágenes de puntos de un flujo de bits empalmado.

**[0091]** De acuerdo con las técnicas, en un ejemplo, cuando el indicador indica que una actualización de conjunto de parámetros no puede ocurrir en la parte del flujo de bits, el descodificador de vídeo 30 puede recibir el primer conjunto de parámetros en una transmisión fuera de banda, es decir, una transmisión separada de una secuencia de vídeo codificada del flujo de bits. En este caso, dado que el primer conjunto de parámetros es el único conjunto de parámetros activado para todo el flujo de bits, un codificador de vídeo, como el codificador de vídeo 20, puede codificar el primer conjunto de parámetros fuera de banda, de manera que el primer conjunto de parámetros sea independiente de los datos de secuencia de vídeo codificada. En otro ejemplo, cuando el indicador indica que no se puede producir una actualización del conjunto de parámetros en la parte del flujo de bits, el descodificador de vídeo 30 puede recibir el primer conjunto de parámetros en una transmisión en banda, es decir, una transmisión con una secuencia de vídeo codificada del flujo de bits, sin comparar el contenido de cualquier transmisión repetida del primer conjunto de parámetros. En este caso, cuando un codificador de vídeo, como el codificador de vídeo 20, codifica transmisiones repetidas del primer conjunto de parámetros para mejorar la resistencia a los errores, el descodificador de vídeo 30 no realiza una comparación de contenido para los conjuntos de parámetros repetidos. Las comparaciones de contenido no son necesarias porque, de acuerdo con la indicación, el descodificador de vídeo 30 puede funcionar como si los conjuntos de parámetros del mismo tipo y con los mismos valores de ID tuvieran el mismo contenido.

**[0092]** En algunos ejemplos, la indicación puede ser un elemento sintáctico, tal como un indicador, incluido en un flujo de bits, por ejemplo, en un VPS, SPS, PPS o APS. Como ejemplo, la indicación puede ser un elemento sintáctico incluido en una secuencia de vídeo codificada. Como otro ejemplo, la indicación puede ser un requisito en una definición de perfil. En otros ejemplos, la indicación puede definirse como un parámetro de tipo de medio e incluirse en el protocolo de descripción de sesión (SDP) cuando el flujo de bits de vídeo se utiliza en un entorno de transmisión basado en el protocolo de transporte en tiempo real (RTP), como la transmisión continua o la multidifusión usando RTP para el transporte de datos multimedia. En este caso, la indicación puede usarse en una fase de negación de sesión, también conocida como intercambio de capacidades. En aún otros ejemplos, la indicación puede definirse como una propiedad de transmisión continua incluida en una descripción de presentación de medios (MPD) en un entorno de transmisión continua de protocolo de transferencia de hipertexto (HTTP), como DASH.

**[0093]** En otro ejemplo de las técnicas, el descodificador de vídeo 30 pueden estar configurado para descodificar una indicación para cada imagen CRA de si realizar el acceso aleatorio a un flujo de bits de la unidad de acceso CRA requiere obtener conjuntos de parámetros a partir de unidades de acceso anteriores. La indicación es especialmente útil para las unidades de acceso CRA introducidas en la norma HEVC porque las imágenes CRA pueden colocarse en cualquier punto dentro de una secuencia de vídeo codificada y no limpian la DPB, lo cual brinda más opciones para el acceso aleatorio al flujo de bits sin una menor eficiencia de codificación.

**[0094]** Cuando la indicación indica que no se necesitan conjuntos de parámetros a partir de unidades de acceso anteriores para el acceso aleatorio de una imagen CRA en particular, el descodificador de vídeo 30 puede determinar para realizar acceso aleatorio a partir de esa imagen CRA con ninguna operación obtener. Cuando la indicación indica que se necesitan conjuntos de parámetros de unidades de acceso anteriores para el acceso aleatorio desde una imagen CRA particular, el descodificador de vídeo 30 puede determinar que no se realice un acceso aleatorio desde esa imagen para evitar la operación de búsqueda. En algunos casos, el descodificador de vídeo 30 puede descodificar indicaciones para varias imágenes CRA diferentes dentro de una parte, por ejemplo, una secuencia de vídeo codificada del flujo de bits para identificar una unidad de acceso desde la cual se puede realizar un acceso aleatorio sin realizar una operación de búsqueda de conjuntos de parámetros. De esta manera, el descodificador de vídeo 30 puede evitar realizar una operación de búsqueda seleccionando una imagen CRA diferente desde la cual realizar el acceso aleatorio.



**[0095]** El acceso aleatorio se refiere a una descodificación de un flujo de bits de vídeo a partir de una imagen codificada que no es la primera imagen codificada en el flujo de bits. El acceso aleatorio a un flujo de bits se necesita en muchas aplicaciones de vídeo, como la radiodifusión y la transmisión continua, por ejemplo, para que los usuarios alternen entre canales diferentes, para pasar a partes específicas del vídeo, para empalmar diferentes flujos de bits o para cambiar a un flujo de bits diferente para adaptación del flujo de la velocidad de transmisión de bits, velocidad de tramas, resolución espacial, etc. Esta característica se habilita insertando imágenes de acceso aleatorio o puntos de acceso aleatorios en intervalos regulares, en el flujo de bits de vídeo. Se puede utilizar una imagen de actualización instantánea del descodificador (IDR) para el acceso aleatorio. Sin embargo, dado que una imagen IDR comienza una secuencia de vídeo codificada y siempre limpia la memoria intermedia de imágenes descodificadas (DPB), las imágenes que siguen a la imagen IDR en orden de descodificación no pueden usar imágenes descodificadas antes que la imagen IDR como referencia. Por consiguiente, los flujos de bits que dependen de imágenes IDR para el acceso aleatorio pueden tener una eficiencia de codificación significativamente menor.

**[0096]** Con el fin de mejorar la eficiencia de la codificación, el concepto de imágenes de acceso aleatorio limpio (CRA) se introdujo en el HEVC HM. A diferencia de las imágenes IDR, las imágenes CRA se pueden colocar en cualquier punto dentro de una secuencia de vídeo codificada y no limpian la DPB, lo cual brinda más opciones para el acceso aleatorio al flujo de bits sin una menor eficiencia de codificación. Debido a que una imagen CRA no limpia la DPB, las imágenes que siguen a la imagen CRA en orden de descodificación pero que la preceden en orden de salida pueden usar imágenes descodificadas antes de la imagen CRA como referencia. Las imágenes que siguen a una imagen CRA en orden de descodificación, pero que preceden a la imagen CRA en orden de salida, se denominan imágenes iniciales de la imagen CRA. Las imágenes principales de una imagen CRA se pueden descodificar correctamente si la descodificación comienza desde una imagen IDR o CRA que precede a la imagen CRA actual. Sin embargo, las imágenes iniciales de una imagen CRA no se pueden descodificar correctamente cuando se produce un acceso aleatorio desde la imagen CRA actual. Las imágenes principales, por lo tanto, típicamente se descartan de la DPB durante la descodificación de acceso aleatorio. Para evitar la propagación de errores por imágenes de referencia que pueden no estar disponibles dependiendo de dónde comienza la descodificación, todas las imágenes que siguen a una imagen CRA tanto en orden de descodificación como en orden de salida pueden estar restringidas a no usar ninguna imagen que preceda a la imagen CRA en orden de descodificación u orden de salida (lo cual incluye las imágenes iniciales) como referencia.

**[0097]** Las funciones de acceso aleatorio similares son soportadas en la norma H.264/AVC con un mensaje SEI de punto de recuperación. Una implementación de descodificador H.264/AVC puede o no soportar las funciones de mensaje SEI de punto de recuperación. En HEVC, un flujo de bits que comienza con una imagen CRA se considera un flujo de bits conforme. Cuando un flujo de bits comienza con una imagen CRA, las imágenes iniciales de la imagen CRA pueden referirse a imágenes de referencia no disponibles y, por lo tanto, no pueden descodificarse. La norma HEVC especifica que no se emiten las imágenes iniciales de la imagen CRA inicial. Sin embargo, para el establecimiento de la conformidad del flujo de bits, la norma HEVC especifica un proceso de descodificación para generar imágenes de referencia no disponibles para descodificar las imágenes principales sin salida. Las implementaciones de descodificador conformes no tienen que seguir ese proceso de descodificación siempre que pueda generarse una salida idéntica en comparación con cuando se realiza el proceso de descodificación desde el comienzo de la secuencia de vídeo codificada.

**[0098]** Además, un flujo de bits conforme en la norma HEVC puede no contener ninguna imagen IDR, y en consecuencia puede contener un subconjunto de una secuencia de vídeo codificada o una secuencia de vídeo codificada incompleta. En la norma HEVC, una secuencia de vídeo codificada se define como una secuencia de unidades de acceso que incluye, en orden de descodificación, una unidad de acceso IDR seguida de cero o más unidades de acceso no IDR incluidas todas las unidades de acceso posteriores hasta, pero sin incluir, cualquier unidad de acceso IDR posterior. Las unidades de acceso incluidas en la secuencia de vídeo codificada corresponden a imágenes de la secuencia de vídeo. Cada una de las unidades de acceso incluye una o más unidades de capa de abstracción de red (NAL). La norma HEVC define las unidades NAL de capa de codificación de vídeo (VCL) y las unidades de NAL no VCL. Una unidad VCL NAL contiene un fragmento de datos de vídeo para la imagen. Una unidad NAL no VCL contiene información distinta a un fragmento de datos de vídeo. Por ejemplo, una unidad NAL no VCL puede contener datos delimitadores o un conjunto de parámetros, como un VPS, SPS o PPS. Otro ejemplo de una unidad NAL no VCL es una unidad SEI NAL que puede incluir mensajes de temporización de imagen utilizados por un descodificador de vídeo al descodificar el flujo de bits.

**[0099]** Cuando descodificador de vídeo 30 realiza acceso aleatorio de una imagen CRA incluida en un flujo de bits, no se descodifican todas las imágenes anteriores a la imagen CRA en orden de salida o descodificación y no están disponibles para su uso como imágenes de referencia. En este caso, para descodificar la imagen CRA o cualquiera de las siguientes imágenes en orden de salida o descodificación, el descodificador de vídeo 30 puede necesitar obtener uno o más conjuntos de parámetros incluidos en una unidad de acceso para una de las imágenes anteriores no disponibles, lo cual es una operación pesada.

**[0100]** Como un ejemplo, un SPS se pueden incluir en una unidad de acceso para una primera imagen CRA del flujo de bits y ninguna SPS se puede incluir en una segunda imagen CRA posterior del flujo de bits. En este caso, si el

acceso aleatorio se realiza desde la segunda imagen CRA y todas las imágenes anteriores, incluida la primera imagen CRA, no están disponibles, el SPS tampoco está disponible para descodificar la segunda imagen CRA y cualquier imagen posterior en el flujo de bits. Para realizar un acceso aleatorio desde la segunda imagen de CRA, el descodificador de vídeo 30 necesita obtener el SPS de la primera imagen de CRA no disponible.

**[0101]** Convencionalmente, para evitar la operación de búsqueda, un descodificador de vídeo puede realizar una comprobación a través de un flujo de bits recibidos para determinar qué conjuntos de parámetros se utilizan para diferentes partes del flujo de bits y si el acceso aleatorio está disponible a partir de imágenes RAP en el flujo de bits. Tanto la operación de búsqueda como la operación de comprobación del flujo de bits pueden requerir un uso intensivo del procesador, ser operaciones pesadas para el descodificador de vídeo y sería deseable evitar estas operaciones.

**[0102]** Las técnicas de la presente divulgación proporcionan una indicación de si realizar acceso aleatorio a partir de una unidad de acceso CRA particular en un flujo de bits requiere obtener conjuntos de parámetros a partir de unidades de acceso anteriores. Cuando no se necesitan conjuntos de parámetros previos para el acceso aleatorio desde una unidad de acceso CRA particular, el descodificador de vídeo 30 puede realizar un acceso aleatorio desde esa unidad de acceso CRA particular sin una operación de búsqueda. A la inversa, cuando se necesitan conjuntos de parámetros anteriores para el acceso aleatorio desde la unidad de acceso CRA particular, el descodificador de vídeo 30 puede no realizar el acceso aleatorio desde esa unidad de acceso CRA particular para evitar la operación de búsqueda. En algunos ejemplos, cuando se necesitan conjuntos de parámetros anteriores para el acceso aleatorio desde la unidad de acceso CRA particular, el descodificador de vídeo 30 puede realizar un acceso aleatorio desde esa unidad de acceso CRA después de obtener los conjuntos de parámetros de las unidades de acceso anteriores.

**[0103]** De acuerdo con las técnicas descritas en esta divulgación, una indicación puede incluirse en el flujo de bits para cada unidad de acceso CRA para indicar si los conjuntos de parámetros incluidos en unidades de acceso anteriores en el orden de salida o descodificación son necesarios para descodificar la unidad de acceso CRA y todas las unidades de acceso posteriores en orden de salida o descodificación. Si no se necesitan conjuntos de parámetros anteriores para el acceso aleatorio desde una unidad de acceso CRA, el descodificador de vídeo 30 puede realizar un acceso aleatorio al flujo de bits desde esa unidad de acceso sin depender de ninguna unidad NAL de conjunto de parámetros incluida en las unidades de acceso anteriores. De esta manera, todas las unidades NAL anteriores pueden simplemente omitirse o ignorarse.

**[0104]** Como en el ejemplo expuesto anteriormente, una unidad SPS NAL puede ser incluido en una unidad de acceso para una primera imagen CRA de la secuencia de vídeo y ninguna unidad SPS NAL pueden incluirse en una segunda imagen CRA posterior de la secuencia de vídeo. De acuerdo con las técnicas descritas en esta divulgación, se puede incluir una indicación en la unidad de acceso para la primera imagen de CRA para indicar que no se necesita ningún conjunto de parámetros de una unidad de acceso anterior en orden de salida o descodificación para descodificar la primera imagen de CRA o cualquier otra imagen posterior en orden de salida o descodificación. Además, se puede incluir una indicación en la unidad de acceso para que la segunda imagen CRA indique que se necesita un conjunto de parámetros de una unidad de acceso anterior en orden de salida o descodificación (es decir, el SPS en la unidad de acceso para la primera imagen CRA) para descodificar la segunda imagen CRA o cualquier imagen posterior en orden de salida o descodificación. En este ejemplo, basándose en las indicaciones, el descodificador de vídeo 30 puede determinar realizar un acceso aleatorio desde la primera imagen CRA, pero no realizar un acceso aleatorio desde la segunda imagen CRA para evitar la operación de búsqueda.

**[0105]** Aunque se describen principalmente con respecto a las imágenes CRA, las técnicas de esta divulgación se pueden usar de manera similar para otros tipos de imágenes RAP, incluidas las imágenes IDR y las imágenes de actualización de descodificación gradual (GDR). En algunos casos, la indicación puede comprender un tipo particular de unidad de capa de abstracción de red (NAL) incluida en una unidad de acceso para cada imagen IDR o CRA en una secuencia de vídeo codificada. En otros casos, la indicación puede comprender un elemento sintáctico, como un indicador, incluido en una carga útil de unidad NAL incluida en una unidad de acceso para cada imagen IDR o CRA en la secuencia de vídeo codificada. En un ejemplo, la indicación puede comprender un indicador incluido en una unidad NAL de información de mejora suplementaria (SEI) en la unidad de acceso para la imagen CRA.

**[0106]** Además, por ejemplo, la indicación de una imagen dada puede ser señalizada utilizando un tipo particular de SPS con un tipo de unidad NAL distinto. En este ejemplo, puede requerirse que el tipo particular de SPS se incluya en la primera unidad NAL en la unidad de acceso para la imagen, o la primera unidad NAL después de una unidad NAL delimitadora de la unidad de acceso (si está presente). En otro ejemplo, la indicación de una imagen dada puede ser señalizada utilizando un tipo particular de delimitador de unidad de acceso con un tipo de unidad NAL distinto o un indicador incluido en una carga útil de una unidad NAL delimitadora de unidad de acceso en la unidad de acceso para la imagen. En un ejemplo adicional, la indicación de una imagen dada puede ser señalizada utilizando un indicador incluido en la cabecera de una unidad NAL en la unidad de acceso para la imagen. Además, la indicación de una imagen dada puede ser señalizada utilizando un tipo particular de mensaje SEI o un indicador en un mensaje SEI. En el HEVC HM, el mensaje SEI puede incluirse en una unidad SEI NAL en la unidad de acceso para la imagen.

**[0107]** La FIG. 4 es un diagrama conceptual que ilustra una secuencia de vídeo 100 codificada de acuerdo con técnicas de codificación de vídeo predictiva. Como se ilustra en la FIG. 4, la secuencia de vídeo 100 incluye las imágenes Pic<sub>1</sub>-

Pic<sub>10</sub>. Las imágenes Pic<sub>1</sub>-Pic<sub>10</sub> están ordenadas y numeradas secuencialmente de acuerdo con el orden en el que se muestran. Como se describe en más detalle a continuación, el orden de salida o visualización no se corresponde necesariamente con el orden de descodificación. Como se ilustra en la FIG. 4, la secuencia de vídeo 100 incluye GOP<sub>1</sub> y GOP<sub>2</sub>, donde las imágenes Pic<sub>1</sub>-Pic<sub>5</sub> se incluyen en GOP<sub>1</sub> y las imágenes Pic<sub>6</sub>-Pic<sub>10</sub> se incluyen en GOP<sub>2</sub>. La FIG. 4 ilustra además la Pic<sub>5</sub> como dividida en slice<sub>1</sub> y slice<sub>2</sub>, donde cada uno de los fragmentos <sub>1</sub> y <sub>2</sub> incluye LCU consecutivas de acuerdo con un escaneo de ráster de izquierda a derecha de arriba a abajo. Además, la última LCU en Pic<sub>5</sub> se ilustra como una división adicional en varias CU de acuerdo con un árbol cuaternario. Aunque no se muestran, las otras imágenes ilustradas en la FIG. 4 se puede dividir en una o más partes de una manera similar.

**[0108]** Como se describe en más detalle a continuación, el HEVC HM incluye dos tipos de esquemas de fragmento de división, a saber fragmentos regulares y fragmentos de entropía. Adicionalmente, además de esquemas de división de fragmentos, el HM también incluye los esquemas de división de procesamiento en paralelo de frente de onda (WPP) y en elementos. Se puede seleccionar y aplicar un esquema de división basándose en una coincidencia de tamaño de unidad de transmisión máxima (MTU) y/o procesamiento en paralelo. Slice<sub>1</sub> y slice<sub>2</sub> de Pic<sub>5</sub> o cualquiera de las otras imágenes, pueden ser fragmentos regulares o fragmentos de entropía. Además, en otros ejemplos, cada uno de Pic<sub>1</sub>-Pic<sub>10</sub> se puede dividir usando elementos o WPP.

**[0109]** Además, la FIG. 4 ilustra el concepto de fragmentos I, fragmentos P y fragmentos B con respecto al GOP<sub>2</sub> de la secuencia de vídeo 100. Las flechas asociadas con cada una de las Pic<sub>6</sub> - Pic<sub>10</sub> en GOP<sub>2</sub> indican si una imagen incluye fragmentos I, fragmentos P o fragmentos B basados en una imagen referenciada indicada por las flechas. En la secuencia de vídeo 100, Pic<sub>6</sub> y Pic<sub>9</sub> representan imágenes que incluyen fragmentos I (es decir, fragmentos predichos basándose en las referencias dentro de la imagen en sí), las imágenes Pic<sub>7</sub> y Pic<sub>10</sub> representan imágenes que incluyen fragmentos P (es decir, fragmentos con predicción inter basados en referencias en una imagen anterior) y Pic<sub>8</sub> representa una imagen que incluye fragmentos B (es decir, fragmentos con predicción inter basados en referencias tanto en una imagen anterior como en una imagen posterior).

**[0110]** La FIG. 5 es un diagrama conceptual que ilustra un ejemplo de una secuencia de vídeo codificada 120 que corresponde al GOP<sub>2</sub> ilustrado en la FIG. 4. Un HEVC HM define una secuencia de vídeo codificada como una secuencia de unidades de acceso que consta, en orden de descodificación, de una unidad de acceso de actualización de descodificación instantánea (IDR) seguida de cero o más unidades de acceso no IDR incluidas todas las unidades de acceso posteriores hasta, pero sin incluir, cualquier unidad de acceso IDR posterior. Una imagen codificada es una representación codificada de una imagen que contiene todos los bloques de árboles de la imagen. Una unidad de acceso es un conjunto de unidades de capa de abstracción de red (NAL) que son consecutivas en orden de descodificación y representan una imagen codificada. Una unidad de acceso puede incluir unidades NAL de capa de codificación de vídeo (VCL) que representan fragmentos codificados de la imagen codificada, y unidades NAL no VCL, incluidos datos de delimitadores, conjuntos de parámetros y mensajes de información de mejora suplementaria (SEI).

**[0111]** Como se ilustra en la FIG. 5, la secuencia de vídeo codificada 120 incluye una unidad de acceso correspondiente a cada una de las Pic<sub>6</sub>-Pic<sub>10</sub>. Las unidades de acceso del flujo de datos de vídeo 120 están dispuestas secuencialmente de acuerdo con el orden de descodificación. Cabe señalar que la unidad de acceso correspondiente a Pic<sub>9</sub> está ubicada antes de la unidad de acceso correspondiente a Pic<sub>8</sub>. Por lo tanto, el orden de descodificación no corresponde al orden de salida ilustrado en la FIG. 4. En este ejemplo, la diferencia entre el orden de salida y el orden de descodificación se debe a que Pic<sub>8</sub> hace referencia a Pic<sub>9</sub>. Por lo tanto, Pic<sub>9</sub> debe descodificarse antes de que Pic<sub>8</sub> pueda descodificarse.

**[0112]** Como se ilustra en la FIG. 5, la unidad de acceso correspondiente a Pic<sub>9</sub> incluye una unidad NAL 122 delimitadora de unidad de acceso (AU), una unidad NAL 124 del conjunto de parámetros de imagen (PPS), una unidad NAL 126 de slice<sub>1</sub> y una unidad NAL 128 de slice<sub>2</sub>. Cada una de la unidad NAL 126 de slice<sub>1</sub> y la unidad NAL 128 de slice<sub>2</sub> contiene un fragmento de datos de vídeo y son ejemplos de unidades VCL NAL. Como se describió anteriormente con respecto a la FIG. 4, Pic<sub>9</sub> representa una imagen que incluye fragmentos I. En el ejemplo de la FIG. 5, por lo tanto, cada una de la unidad NAL 126 de slice<sub>1</sub>, y la unidad NAL 128 de slice<sub>2</sub> de la unidad de acceso correspondiente a Pic<sub>9</sub> puede representar fragmentos I.

**[0113]** Una unidad no VCL NAL incluye información distinta de un fragmento de datos de vídeo. Por ejemplo, un no VCL puede contener datos delimitadores, un conjunto de parámetros o mensajes SEI. En el ejemplo ilustrado en la FIG. 5, la unidad NAL delimitadora de AU 122 incluye información para delimitar la unidad de acceso correspondiente a Pic<sub>9</sub> de la unidad de acceso correspondiente a Pic<sub>7</sub>. Además, la unidad PPS NAL 124 incluye un conjunto de parámetros de imagen. La unidad NAL 122 delimitadora de AU y la unidad 124 de PPS NAL son ejemplos de unidades NAL no VCL. Otro ejemplo de una unidad NAL no VCL es una unidad SEI NAL. Las unidades SEI NAL pueden incluir mensajes SEI de temporización de imagen o mensajes SEI de período de memoria intermedia que son utilizados por un descodificador de vídeo cuando se descodifica un flujo de bits. Los mensajes SEI de sincronización de imagen pueden incluir información que indica cuándo un descodificador de vídeo debe comenzar a descodificar una unidad VCL NAL. Los mensajes SEI del período de almacenamiento en memoria intermedia pueden incluir información de retardo de eliminación de la memoria intermedia de imágenes codificadas (CPB) que indica cuándo un descodificador de vídeo debe eliminar las imágenes de la CPB para evitar el desbordamiento.

**[0114]** En el HEVC HM, cada uno de una secuencia de vídeo, un GOP, una imagen, un fragmento, una CU, una PU y una TU pueden estar asociados con elementos sintácticos definidos para indicar propiedades de codificación de vídeo para esa estructura de datos de vídeo. El HM emplea además un concepto de conjunto de parámetros. Un conjunto de parámetros es una estructura sintáctica que incluye elementos sintácticos definidos para indicar las propiedades de codificación de vídeo aplicadas a cualquier estructura de datos de vídeo que se refiera al conjunto de parámetros. El HM emplea un mecanismo de conjunto de parámetros jerárquico donde los elementos sintácticos se incluyen en un tipo de conjunto de parámetros basándose en la frecuencia con la que se espera que cambien las propiedades de codificación de vídeo. El mecanismo de conjunto de parámetros, por consiguiente, desacopla la transmisión de información que cambia poco frecuentemente de la transmisión de datos de bloque codificados. En algunas aplicaciones, los conjuntos de parámetros pueden ser transmitidos "fuera de banda", es decir, no transportados junto con datos de vídeo codificados, porque la transmisión fuera de banda es típicamente más fiable que la transmisión dentro de banda.

**[0115]** El HM define varios tipos diferentes de conjuntos de parámetros, incluyendo uno o más de un conjunto de parámetros de vídeo (VPS), un conjunto de parámetros de secuencia (SPS), un conjunto de parámetros de imagen (PPS), y un conjunto de parámetros de adaptación (APS). Un flujo de bits puede incluir varios conjuntos de parámetros diferentes de cada tipo. Un conjunto de parámetros particular se identifica mediante un identificador de conjunto de parámetros (ID). Cada VPS incluye un ID de VPS, cada SPS incluye un ID de SPS y puede referirse a un ID de VPS, cada PPS incluye un ID de PPS y se refiere a un ID de SPS, y cada cabecera de fragmento se refiere a un ID de PPS, y posiblemente a un ID de APS. Además, cada mensaje de información de mejora suplementaria (SEI) del período de almacenamiento en memoria intermedia también hace referencia a un ID de SPS y, posiblemente, a un ID de VPS. En cada uno de estos ejemplos, los ID de conjunto de parámetros se pueden codificar como elementos sintácticos utilizando un procedimiento de codificación exp-Golomb ( $ue(v)$ ) de entero sin signo de longitud variable.

**[0116]** Un VPS es una estructura sintáctica que incluye elementos sintácticos que se aplican a cero o más secuencias de vídeo codificadas completas. Más específicamente, un VPS incluye elementos sintácticos que describen las características generales de las secuencias de vídeo codificadas, incluida la información de perfiles, grados y niveles, y las dependencias entre subcapas temporales. Un VPS particular se identifica utilizando un ID de VPS particular codificado como un elemento sintáctico en un mensaje SPS o SEI. Un SPS es una estructura sintáctica que incluye elementos sintácticos que se aplican a cero o más secuencias de vídeo codificadas completas. Más específicamente, un SPS incluye elementos sintácticos definidos para indicar las propiedades de codificación de vídeo que se espera que permanezcan sin cambios para una secuencia de tramas (por ejemplo, orden de imágenes, número de tramas de referencia y tamaño de imagen). Un SPS particular se identifica utilizando un ID de SPS particular codificado como un elemento sintáctico en un PPS o un mensaje SEI.

**[0117]** Un PPS es una estructura sintáctica que incluye elementos sintácticos que se aplican a cero o más imágenes codificadas completas. Más específicamente, un PPS incluye elementos sintácticos definidos para indicar propiedades de codificación de vídeo que pueden cambiar de imagen a imagen dentro de una secuencia (por ejemplo, modo de codificación por entropía, parámetros de cuantificación y profundidad de bits). Un PPS particular se identifica utilizando un ID de PPS particular codificado como un elemento sintáctico en una cabecera de fragmento. Un APS es una estructura sintáctica que incluye elementos sintácticos que se aplican a cero o más imágenes codificadas completas. Un APS, si está codificado, puede incluir elementos sintácticos definidos para indicar las propiedades de codificación del vídeo, con el cambio con mayor frecuencia que otras propiedades indicadas en un PPS. Un APS particular se identifica mediante un ID de APS codificado como un elemento sintáctico en una cabecera de fragmento.

**[0118]** El acoplamiento de flujo de bits se refiere a la concatenación de dos o más flujos de bits o partes de los flujos de bits. Por ejemplo, un primer flujo de bits puede ser anexado por un segundo flujo de bits, posiblemente con algunas modificaciones a uno o ambos flujos de bits para generar un flujo de bits empalmado. La primera imagen codificada en el segundo flujo de bits también se denomina punto de empalme. Por lo tanto, las imágenes después del punto de empalme en el flujo de bits empalmado se originaron a partir del segundo flujo de bits, mientras que las imágenes anteriores al punto de empalme en el flujo de bits empalmado se originaron a partir del primer flujo de bits. En algunos casos, un empalmador de flujo de bits puede realizar el empalme de flujo de bits. Los empalmadores de flujos de bits pueden ser ligeros y menos inteligentes que los codificadores de vídeo. Por ejemplo, un empalmador de flujo de bits puede no estar equipado con capacidades de codificación por entropía y descodificación por entropía. El empalme del flujo de bits puede ser útil para las aplicaciones de escalabilidad (por ejemplo, para la velocidad de transmisión de bits, la velocidad de tramas o la escalabilidad de resolución espacial).

**[0119]** Como se describió anteriormente, un flujo de bits puede incluir uno o más conjuntos de parámetros diferentes de cada tipo, por ejemplo, SPS y PPS. Un conjunto de parámetros de un tipo particular se identifica mediante un identificador de conjunto de parámetros (ID). En algunos casos, un descodificador de vídeo puede recibir un conjunto de parámetros de un tipo particular con un ID de conjunto de parámetros particular que tiene un contenido diferente al de un conjunto de parámetros anterior (en orden de flujo de bits o de descodificación) del mismo tipo con el mismo ID de conjunto de parámetros. En este caso, el conjunto de parámetros del tipo particular con el ID de conjunto de parámetros particular (por ejemplo, SPS 0) se actualiza. La actualización del conjunto de parámetros incluye el almacenamiento del contenido del conjunto de parámetros recibido con el ID de conjunto de parámetros en particular

para reemplazar el conjunto de parámetros anterior del mismo tipo y que tenga el mismo valor de ID de conjunto de parámetros.

**[0120]** Una actualización de conjunto de parámetros se puede producir cuando se necesite un nuevo conjunto de parámetros de un tipo particular, pero se hayan utilizado todos los valores posibles del ID de conjunto de parámetros. Una actualización del conjunto de parámetros también puede ocurrir en flujos de bits empalmados incluso cuando los espacios de valor de ID de conjunto de parámetros no se utilizan completamente. Los flujos de bits a menudo usan solo unos pocos SPS y unos pocos PPS, o incluso solo un SPS y un PPS, y usan el valor de ID de SPS más pequeño y el valor de ID de PPS más pequeño (es decir, 0). Por lo tanto, cuando se empalman dos flujos de bits o partes de flujos de bits, es probable que el punto de empalme, originalmente del segundo flujo de bits, y la imagen que precede inmediatamente a la imagen del punto de empalme, originalmente del primer flujo de bits, hagan referencia al mismo ID de SPS o ID de PPS. En este caso, el contenido del SPS o PPS al que hace referencia la imagen del punto de empalme y el contenido de la imagen del SPS o PPS al que hace referencia la imagen inmediatamente anterior a la imagen del punto de empalme, es probablemente diferente.

**[0121]** En el caso particular de los conjuntos de parámetros SPS, cuando se hace referencia a la misma ID de SPS en la imagen del punto de empalme y en la imagen inmediatamente anterior al punto de empalme, y en realidad se utilizan dos SPS diferentes, entonces, en el flujo de bits empalmado, el SPS con el ID de SPS particular al que hace referencia la imagen que precede inmediatamente al punto de empalme se actualiza de manera efectiva mediante el SPS al que se hace referencia en la imagen del punto de empalme. En este caso, los SPS del flujo de bits empalmado no se pueden poner al principio del flujo de bits empalmado. Aunque se describió anteriormente con respecto a los SPS, lo mismo se aplica a los VPS, PPS y APS. La actualización del empalme del flujo de bits y el conjunto de parámetros asociados se ha descrito basándose en el supuesto de que solo hay un punto de empalme en un flujo de bits empalmado. Sin embargo, un flujo de bits puede incluir múltiples puntos de empalme, y las técnicas descritas en el presente documento pueden aplicarse a cada uno de los puntos de empalme individualmente.

**[0122]** El acceso aleatorio se refiere a una descodificación de un flujo de bits de vídeo a partir de una imagen codificada que no es la primera imagen codificada en el flujo de bits. El acceso aleatorio a un flujo de bits es útil en muchas aplicaciones de vídeo, como la radiodifusión y la transmisión continua. Por ejemplo, el acceso aleatorio es útil para que los usuarios cambien entre diferentes canales, para saltar a partes específicas del vídeo, o para cambiar a un flujo de bits diferente para la adaptación del flujo (por ejemplo, para la velocidad de transmisión de bits, la velocidad de tramas o la escalabilidad de resolución espacial). El acceso aleatorio se habilita insertando imágenes de punto de acceso aleatorio (RAP) o unidades de acceso muchas veces a intervalos regulares en un flujo de bits de vídeo.

**[0123]** Se puede usar una imagen de Actualización de descodificación Instantánea (IDR) para el acceso aleatorio a un flujo de bits. Una imagen IDR comienza una secuencia de vídeo codificada e incluye solo fragmentos con predicción intra (es decir, fragmentos I). Además, una imagen IDR siempre limpia la memoria intermedia de imágenes descodificadas (DPB). Por lo tanto, las imágenes que siguen a la IDR en orden de descodificación no pueden usar imágenes descodificadas antes de la imagen IDR como referencia. Por consiguiente, los flujos de bits que dependen de imágenes IDR para el acceso aleatorio pueden tener una eficiencia de codificación significativamente menor.

**[0124]** Para mejorar la eficiencia de la codificación, el concepto de imágenes de acceso aleatorio limpio (CRA) se introdujo en el HEVC HM. Una imagen CRA, al igual que una imagen IDR, incluye solo fragmentos de predicción intra, es decir, fragmentos I. Las imágenes CRA se diferencian de las imágenes IDR en que las imágenes CRA no limpian la DPB y pueden ubicarse en cualquier posición dentro de una secuencia de vídeo codificada. Por consiguiente, se permite que las imágenes que siguen a una imagen CRA en orden de descodificación, pero que preceden a las imágenes CRA en orden de salida, utilicen imágenes descodificadas antes de la imagen CRA como referencia. Las imágenes que siguen a una imagen CRA en orden de descodificación, pero preceden a la imagen CRA en orden de salida, se denominan imágenes principales asociadas con la imagen CRA (o imágenes principales de la imagen CRA).

**[0125]** Las imágenes principales de una imagen CRA se pueden descodificar correctamente si la descodificación comienza desde una imagen IDR o CRA antes de la imagen CRA actual. Sin embargo, las imágenes principales de una imagen CRA pueden no ser descodificables correctamente cuando se inicia un acceso aleatorio desde la imagen CRA actual. Por lo tanto, las imágenes principales se descartan típicamente durante la descodificación de acceso aleatorio de la imagen CRA actual. Refiriéndose al ejemplo ilustrado en las FIGs. 4 y 5,  $pic_9$  puede ser una imagen CRA y  $pic_8$  puede ser una imagen principal de  $pic_9$ .  $Pic_8$  se puede descodificar correctamente si se accede al  $GOP_2$  en la  $pic_6$ , pero es posible que no se pueda descodificar correctamente si se accede al  $GOP_2$  en  $pic_9$ . Esto se debe al hecho de que la  $pic_7$  puede no estar disponible si se accede al  $GOP_2$  como  $pic_9$ . Para evitar la propagación de errores por imágenes de referencia que pueden no estar disponibles dependiendo de dónde comienza la descodificación, todas las imágenes que siguen a una imagen CRA tanto en orden de descodificación como en orden de salida pueden estar restringidas a no usar ninguna imagen que preceda a la imagen CRA en orden de descodificación u orden de salida (lo cual incluye las imágenes iniciales) como referencia.

**[0126]** En el HEVC HM, un flujo de bits que comienza con una imagen CRA se considera un flujo de bits conforme. Cuando un flujo de bits comienza con una imagen CRA, las imágenes iniciales de la imagen CRA pueden referirse a imágenes de referencia no disponibles y, por lo tanto, pueden no descodificarse correctamente. Sin embargo, el HM

especifica que las imágenes iniciales de la imagen CRA inicial no se emiten, de ahí el nombre "acceso aleatorio limpio". En el ejemplo ilustrado en las FIGs. 4 y 5, si un flujo de bits comienza en  $P_{i-3}$ ,  $P_8$  no se emite. Para el establecimiento de requisitos de conformidad con el flujo de bits, el HM especifica un proceso de descodificación para generar imágenes de referencia no disponibles para la descodificación de las imágenes iniciales no emitidas. Sin embargo, las implementaciones de descodificador conformes no tienen que seguir ese proceso de descodificación, siempre que el descodificador conforme pueda generar un resultado idéntico al del proceso de descodificación que se realiza desde el comienzo de la secuencia de vídeo codificada.

**[0127]** Hay que señalar que similares funciones de acceso aleatorio a las del HEVC HM se soportan en H.264/AVC con el mensaje SEI de punto de recuperación. Una implementación de descodificador H.264/AVC puede o no soportar la función de acceso aleatorio. Además, se debe tener en cuenta que, en el HEVC HM, un flujo de bits conforme puede no contener imágenes IDR. El HM define una secuencia de vídeo codificada es una secuencia de unidades de acceso que incluye, en orden de descodificación, una unidad de acceso IDR seguida de cero o más unidades de acceso no IDR incluidas todas las unidades de acceso posteriores hasta, pero sin incluir, cualquier unidad de acceso IDR posterior. Por consiguiente, un flujo de bits conforme que no contiene ninguna imagen IDR puede contener un subconjunto de una secuencia de vídeo codificada o una secuencia de vídeo codificada incompleta.

**[0128]** Ahora se describirán los problemas y las técnicas propuestas relacionadas con las actualizaciones de conjuntos de parámetros en un flujo de bits. Convencionalmente, para determinar si realizar una actualización del conjunto de parámetros, un descodificador compara el contenido del conjunto de parámetros cada vez que se recibe un nuevo conjunto de parámetros de un tipo particular con un ID de conjunto de parámetros particular que tiene el mismo ID de conjunto de parámetros que un conjunto de parámetros recibido anteriormente del mismo tipo. La comparación de conjuntos de parámetros se realiza para verificar si los conjuntos de parámetros incluyen el mismo contenido o uno diferente. Basándose en si los conjuntos de parámetros tienen el mismo contenido, un descodificador determina si activar el conjunto de parámetros recibido actualmente y desactivar el conjunto de parámetros previamente activo. En el caso de que los conjuntos de parámetros se transmitan en banda con datos de vídeo codificados, los conjuntos de parámetros pueden transmitirse repetidamente para mejorar la resistencia a los errores. En este caso, un descodificador puede hacer una comparación del contenido del conjunto de parámetros para cada conjunto de parámetros repetidos. Hacer comparaciones de conjuntos de parámetros repetidos puede imponer una carga innecesaria en el descodificador.

**[0129]** Esta divulgación describe técnicas que pueden proporcionar un uso más eficiente de los recursos de procesamiento de descodificación, particularmente en el caso de actualizaciones de conjuntos de parámetros. En un ejemplo, esta divulgación describe un dispositivo de procesamiento de vídeo configurado para usar un indicador de nivel de flujo de bits para determinar si un conjunto de parámetros se puede actualizar en una parte de un flujo de bits. En algunos casos, se puede usar una indicación de nivel de flujo de bits para indicar si se puede actualizar un conjunto de parámetros. La indicación puede ser un indicador incluido en el flujo de bits, por ejemplo, en el SPS, y el indicador puede ser idéntico en todos los SPS. Cuando el indicador es igual a 1, se puede actualizar un conjunto de parámetros. Cuando el indicador es igual a 0, no se puede actualizar ningún conjunto de parámetros. En este caso, un descodificador de vídeo no necesita comparar el contenido de dos conjuntos de parámetros del mismo tipo y con los mismos valores de ID de conjunto de parámetros determinar si los conjuntos de parámetros tienen el mismo contenido, ya que, cuando el indicador es igual a 0, no se puede realizar la actualización de un conjunto de parámetros y el descodificador de vídeo funciona como si los dos conjuntos de parámetros tuvieran el mismo contenido.

**[0130]** En otros casos, una indicación de nivel de flujo de bits puede usarse para indicar si un conjunto de parámetros de un tipo particular (por ejemplo, un VPS, SPS, PPS, o APS) se puede actualizar. La indicación puede ser un indicador incluido en el flujo de bits, por ejemplo, en el SPS, y el indicador puede ser idéntico en todos los SPS. Cuando el indicador es igual a 1, se puede actualizar un conjunto de parámetros del tipo particular. Cuando el indicador es igual a 0, no se puede actualizar ningún conjunto de parámetros del tipo en particular. En este caso, un descodificador de vídeo no necesita comparar el contenido de dos conjuntos de parámetros del tipo en particular y con los mismos valores de ID de conjunto de parámetros para determinar si los conjuntos de parámetros tienen el mismo contenido, porque, cuando el indicador es igual a 0, no se puede realizar la actualización de un conjunto de parámetros y el descodificador de vídeo funciona como si los dos conjuntos de parámetros tuvieran el mismo contenido. Cuando hay tres tipos de conjuntos de parámetros, por ejemplo, SPS, PPS y APS, se utilizan tres de estas indicaciones, una para cada tipo de conjunto de parámetros.

**[0131]** En los ejemplos descritos anteriormente, la indicación se incluye en el nivel de flujo de bits en un conjunto de parámetros. En otro ejemplo, la indicación puede incluirse en un mensaje SEI. En otro ejemplo más, la indicación puede incluirse como un requisito en una definición de un perfil. En algunos ejemplos, la indicación puede incluirse en el nivel de secuencia de vídeo codificada en lugar de en el nivel de flujo de bits. En otros ejemplos, la indicación puede no estar incluida en el flujo de bits, pero en su lugar, la indicación puede definirse como un parámetro de tipo de medios e incluirse en el protocolo de descripción de sesión (SDP) cuando el flujo de bits de vídeo se utiliza en un entorno de transmisión basado en protocolo de transporte en tiempo real (RTP), por ejemplo, transmisión continua o multidifusión utilizando RTP para el transporte de datos de medios, durante la negociación de la sesión o la fase de intercambio de capacidades. Como ejemplo adicional, en un entorno de transmisión continua HTTP, por ejemplo, transmisión continua

dinámica adaptativa sobre HTTP (DASH), la indicación puede definirse como una propiedad de transmisión continua incluida en la descripción de presentación de medios (MPD).

**[0132]** Ahora se describirán los problemas y las técnicas propuestas relacionados con la activación de nuevos conjuntos de parámetros en imágenes IDR. Como se describió anteriormente, en muchos casos las imágenes de puntos de empalme en el flujo de bits son imágenes IDR. Las imágenes de puntos de empalme típicamente tienen características de vídeo diferentes a las imágenes inmediatamente anteriores al punto de empalme en un flujo de bits empalmado. Por lo tanto, habitualmente se necesita un conjunto de parámetros diferente de cada tipo en el punto de empalme, y sería una pérdida de recursos de procesamiento de descodificación comparar si el contenido de los conjuntos de parámetros asociados con la imagen del punto de empalme es idéntico a los conjuntos de parámetros asociados con una imagen anterior. En este caso, puede ser más eficiente activar un nuevo conjunto de parámetros de cada tipo asociado con la imagen del punto de empalme sin comparar el contenido del conjunto de parámetros. Además, cuando una imagen IDR es una imagen de punto de empalme, para evitar un desbordamiento de DPB, puede ser necesario no generar ninguna imagen antes de la imagen IDR, incluso si las imágenes aún no se han emitido al descodificar la imagen IDR.

**[0133]** Esta divulgación describe técnicas que pueden proporcionar un uso más eficiente de los recursos de procesamiento de descodificación, particularmente en el caso de actualizaciones de conjuntos de parámetros en puntos de empalme de flujo de bits. En un ejemplo, esta divulgación describe un dispositivo de procesamiento de vídeo configurado para activar automáticamente un nuevo conjunto de parámetros de cada tipo asociado con una imagen de punto de empalme de un flujo de bits empalmado. En particular, en muchos casos, las imágenes de puntos de empalme en el flujo de bits son imágenes IDR. De acuerdo con las técnicas, en cada imagen IDR, un descodificador de vídeo puede activar cada tipo de conjunto de parámetros (y, por lo tanto, desactivar el conjunto de parámetros previamente activo) aun cuando el conjunto de parámetros recién activado pueda ser exactamente el mismo que el conjunto de parámetros previamente activo, incluyendo el contenido y valor de ID del conjunto de parámetros. Un posible inconveniente potencial de activar nuevos conjuntos de parámetros en cada imagen IDR es que se puede realizar alguna activación de conjuntos de parámetros innecesarios.

**[0134]** En otro ejemplo, una indicación puede incluirse en el flujo de bits para cada imagen IDR para indicar si la imagen IDR es una imagen IDR de empalme (SIDR). En este ejemplo, cuando se indica una imagen SIDR, un descodificador de vídeo puede activar cada tipo de conjunto de parámetros (por lo tanto, desactivar el conjunto de parámetros activado previamente), incluso aunque el conjunto de parámetros recién activado sea exactamente igual al conjunto de parámetros previamente activo, incluyendo el contenido y valor de ID del conjunto de parámetros. Al activar solo nuevos conjuntos de parámetros en las imágenes SIDR, en lugar de en cada imagen IDR, se puede reducir la activación innecesaria del conjunto de parámetros. Además, cuando se indica una imagen SIDR, el comportamiento de salida de la imagen puede ser el mismo que si estuviera presente `no_output_of_prior_pics_flag` y fuera igual a 1, lo cual indica que no se emitirán imágenes antes de la imagen SIDR. El `no_output_of_prior_pics_flag` puede o no estar presente, por ejemplo, en la cabecera de fragmento. En algunos casos, la indicación puede ser un tipo de unidad NAL distinto especial. En otros casos, la indicación puede ser un indicador de nivel de imagen en la cabecera de fragmento, el PPS o el APS.

**[0135]** A continuación se describirán problemas y técnicas propuestas relacionadas con el acceso aleatorio completo a un flujo de bits desde una unidad de acceso RAP. Cuando los conjuntos de parámetros se transportan en banda con datos de vídeo codificados, es decir, en el flujo de bits de vídeo codificado, el acceso aleatorio al flujo de bits desde una unidad de acceso para una imagen IDR o una imagen CRA puede requerir la búsqueda de conjuntos de parámetros desde las unidades de acceso anteriores al IDR o unidad de acceso CRA en orden de descodificación. Cuando no haya necesidad de obtener conjuntos de parámetros de unidades de acceso anteriores a la unidad de acceso IDR o CRA en orden de descodificación, sería conveniente notificar a un descodificador de vídeo para evitar operaciones de búsqueda innecesarias.

**[0136]** Esta divulgación describen las técnicas que se pueden utilizar con el fin de proporcionar un uso más eficiente de los recursos de procesamiento de descodificación en particular en el caso de acceso aleatorio. En un ejemplo, esta divulgación describe un dispositivo de procesamiento de vídeo configurado para usar una indicación en un flujo de bits para determinar si se necesita una unidad NAL de conjunto de parámetros en una unidad de acceso que precede a una unidad de acceso IDR o CRA para descodificar la unidad de acceso IDR o CRA. En particular, un flujo de bits puede incluir una indicación para cada unidad de acceso IDR o CRA, para indicar si se necesita una unidad NAL de conjunto de parámetros que precede a la unidad de acceso IDR o CRA (en orden de flujo de bits o descodificación) para descodificar la unidad de acceso IDR o CRA y cualquiera de las unidades de acceso posteriores (en flujo de bits o en orden de descodificación).

**[0137]** Si la indicación es verdadera para una unidad de acceso IDR o CRA, entonces el descodificador de vídeo puede acceder aleatoriamente al flujo de bits desde la unidad de acceso IDR o CRA sin depender de ninguna unidad NAL de conjunto de parámetros que preceda a la unidad de acceso IDR o CRA. En este caso, todas las unidades NAL anteriores pueden simplemente omitirse o ignorarse. En un ejemplo, la indicación para una imagen IDR o una imagen CRA puede ser señalizada usando un tipo particular de SPS (por ejemplo, con un tipo de unidad NAL distinto) donde SPS puede ser la primera unidad NAL de la unidad de acceso o la primera NAL Unidad después de la unidad NAL

delimitadora de la unidad de acceso, si está presente en la unidad de acceso. En otro ejemplo, la indicación puede ser señalizada utilizando un tipo particular de delimitador de la unidad de acceso, ya sea con un tipo de unidad NAL delimitadora de la unidad de acceso distinto o con un indicador dentro de la carga útil de la unidad NAL de una unidad NAL delimitadora de la unidad de acceso. En otro ejemplo más, la indicación puede ser señalizada utilizando un indicador en la cabecera de la unidad NAL en la unidad de acceso. Como ejemplo adicional, la indicación puede ser señalizada utilizando un tipo particular de mensaje SEI en la unidad de acceso o un indicador en un mensaje SEI en la unidad de acceso.

**[0138]** Ahora se describirán los problemas y las técnicas propuestas relacionadas con las imágenes principales de imágenes CRA en un flujo de bits. Cuando el flujo de bits se inicia desde una imagen CRA, y las imágenes principales asociadas con la imagen CRA no están presentes en un decodificador de vídeo, la memoria intermedia de imágenes codificadas (CPB) puede desbordarse. En consecuencia, puede ser necesario aplicar un conjunto diferente de parámetros del decodificador de referencia hipotético (HRD) para evitar los desbordamientos de memoria intermedia. S. Deshpande, et al., "Signaling of CRA Pictures [Señalización de imágenes CRA]", 9.<sup>a</sup> reunión, Ginebra, Suiza, 27 de abril - 7 de mayo de 2012, Doc., JCTVC-I0278 (en adelante, "Deshpande"), describe una indicación (por ejemplo, un nuevo tipo de unidad NAL o un indicador en una cabecera de fragmento) para una imagen CRA para indicar si las imágenes principales asociadas están presentes, de manera que el decodificador de vídeo puede determinar qué conjunto de parámetros HRD se deben aplicar. Sin embargo, la indicación descrita por Deshpande requiere que una entidad de red externa (por ejemplo, un servidor de transmisión continua o un elemento de red consciente de medios (MANE)) realice cambios en las cabeceras de unidad NAL o cabeceras de fragmento de la imagen CRA, que pueden ser pesados o imposibles para la entidad de red externa.

**[0139]** Esta divulgación se describen las técnicas que se pueden utilizarse con el fin de evitar posibles desbordamientos de memoria intermedia. En un ejemplo, esta divulgación describe un dispositivo de procesamiento de vídeo configurado para usar una indicación para determinar si están presentes las imágenes principales asociadas con una imagen CRA que comienza un flujo de bits. En particular, en un ejemplo, cuando las imágenes iniciales de una imagen CRA que comienza un flujo de bits no están presentes, una entidad de red externa puede cambiar el tipo de unidad NAL para indicar la imagen CRA como una imagen IDR. En otro ejemplo, la entidad de red externa puede notificar al decodificador de vídeo la presencia o ausencia de imágenes principales asociadas con una imagen CRA que comienza el flujo de bits.

**[0140]** La entidad de red externa puede generar un mensaje de protocolo de transmisión continua en tiempo real (RTSP) para sistemas de transmisión continua basados en RTSP, o un parámetro de protocolo de descripción de sesión (SDP) enviado durante las negociaciones de la sesión tanto para transmisión continua basada en RTSP como para radiodifusión/multidifusión basada en (SAP) de protocolo de anuncio de sesión como una configuración estática. En la transmisión continua de HTTP, por ejemplo, la transmisión continua dinámica adaptativa a través de HTTP (DASH), la indicación puede realizarse asignando diferentes localizadores de recursos uniformes (URL) o asignando una URL con diferentes desviaciones de bytes para la imagen CRA y las imágenes principales. En este caso, el decodificador de vídeo puede determinar por sí mismo si están presentes las imágenes iniciales de la imagen CRA, ya que es el decodificador de vídeo el que determina si se deben solicitar las imágenes principales.

**[0141]** Ahora se describirán los problemas y las técnicas propuestas relacionadas con la predicción del vector de movimiento temporal. Refiriéndose nuevamente a la FIG. 4, cada una de las Pic<sub>6</sub>, Pic<sub>7</sub> y Pic<sub>9</sub> se utilizan como imágenes de referencia. Como se describió anteriormente, los elementos sintácticos de información de movimiento pueden identificar la ubicación de un bloque de vídeo predictivo en una imagen de referencia para la predicción inter de un bloque de vídeo en una imagen actual. La predicción de movimiento puede referirse al proceso de definir información de movimiento (por ejemplo, vectores de movimiento e índices de imágenes de referencia) para un bloque de vídeo dado en relación con la información de movimiento para un bloque de vídeo codificado previamente. Por ejemplo, la predicción de movimiento puede implicar el uso de información de movimiento de una CU codificada previamente para generar información de movimiento para una CU actual que se está codificando. La predicción de movimiento se puede aplicar para reducir la cantidad de datos necesarios para comunicar el vector de movimiento de la CU actual.

**[0142]** Un ejemplo de predicción de movimiento es la predicción de movimiento de modo de fusión. En la predicción de movimiento en modo de fusión, la información de movimiento para una CU actual se hereda de un bloque de vídeo contiguo codificado previamente. Las técnicas de predicción de movimiento pueden usar un valor de índice para identificar un bloque de vídeo contiguo del cual el bloque de vídeo actual obtiene su información de movimiento. Los posibles bloques de vídeo contiguos pueden incluir, por ejemplo, bloques de vídeo en la imagen actual que son adyacentes a la CU actual (por ejemplo, arriba, abajo, a la izquierda o a la derecha). Además, los posibles bloques de vídeo contiguos también pueden incluir bloques de vídeo que se ubican en imágenes diferentes a la imagen actual (por ejemplo, un bloque de vídeo co-ubicado desde una imagen temporalmente adyacente). Cuando un posible bloque de vídeo contiguo se encuentra en una imagen diferente a la imagen actual, se le conoce como predicción de vector de movimiento temporal (TMVP). Por ejemplo, en el ejemplo ilustrado en la FIG. 4, si Pic<sub>8</sub> hereda un vector de movimiento de Pic<sub>7</sub>, este es un TMVP. En algunos casos, puede ser ventajoso permitir que algunas imágenes, fragmentos y CU utilicen TMVP, mientras que no permiten que otros utilicen TMVP.



**[0143]** El HEVC HM permite que TMVP se habilite o deshabilite imagen a imagen, y puede señalar la habilitación de TMVP usando el indicador `enable_temporal_mv_flag` en el PPS. Como se describió anteriormente, cada cabecera de fragmento hace referencia a un PPS particular utilizando un ID de PPS particular y determina si TMVP está habilitado para la imagen basándose en el valor del indicador en el PPS particular. Por ejemplo, cuando un fragmento se refiere a un PPS con un valor `enable_temporal_mv_flag` igual a 0 y tiene una identificación temporal igual a 0, se invoca un proceso de marcado de imágenes de referencia en la DPB. Es decir, todas las imágenes en la memoria intermedia de imágenes descodificadas (DPB) están marcadas como "no utilizadas para la predicción del vector de movimiento temporal" y TMVP de imágenes de referencia anteriores no está permitido.

**[0144]** Como una alternativa para incluir la habilitación temporal de `mv_flag` en el PPS, se ha descrito, en Lim, Chong Soon et al. "High-level Syntax: Proposed fix on signaling of TMVP disabling flag [Sintaxis de alto nivel: arreglo propuesto para la señalización del indicador de deshabilitación de TMVP]" 9.ª reunión, Ginebra, Suiza, 27 de abril - 7 de mayo de 2012, Doc., JCTVC-10420, (en adelante "Lim"), para indicar explícitamente el `enable_temporal_mv_flag` en cada cabecera de fragmento P y B. Si bien la técnica descrita en Lim elimina la necesidad de un proceso de marcado y determina las imágenes TMVP de forma independiente en lugar de confiar en la actualización de imagen a imagen, la técnica descrita en Lim no es eficiente cuando muchas imágenes o todas las imágenes en una secuencia de vídeo codificada no utilice TMVP. Bajo la técnica descrita en Lim, las cabeceras de fragmento de todas las fragmentos P y B dentro de una imagen o secuencia incluirían el `enable_temporal_mv_flag` incluso si ninguno de los fragmentos usa TMVP.

**[0145]** Esta divulgación describe técnicas que se pueden usar para reducir la señalización de `enable_temporal_mvflag` para cada cabecera de fragmento P y B. En un ejemplo, esta divulgación describe un dispositivo de procesamiento de vídeo configurado para identificar un indicador en el SPS para determinar si la predicción del vector de movimiento temporal está habilitada. En particular, en algunos ejemplos, se puede agregar un indicador (por ejemplo, `enable_temporal_mv_ps_flag`) en el SPS o PPS. Cuando este indicador es igual a 0 en un SPS o PPS en particular, `enable_temporal_mv_flag` no está presente en ninguna cabecera de fragmento que se refiera al SPS o PPS en particular. Cuando este indicador es igual a 1 en un SPS o PPS en particular, `enable_temporal_mvflag` está presente en cada cabecera de fragmento que se refiere al SPS o PPS en particular. La semántica de `enable_temporal_mv_flag` en las cabeceras de fragmento puede ser la misma que en la propuesta de Lim.

**[0146]** Ahora se describirán los problemas y técnicas propuestas relacionados con la señalización de conjunto de imágenes de referencia para imágenes codificadas en un flujo de bits. Como se describió anteriormente, además de los vectores de movimiento, los elementos sintácticos de información de movimiento también pueden incluir índices de imágenes de referencia. El HEVC HM utiliza un mecanismo basado en el Conjunto de imágenes de referencia (RPS) para gestionar las imágenes de referencia. Un RPS hace referencia a un conjunto de imágenes de referencia asociadas con una imagen, que consta de todas las imágenes de referencia que son anteriores a la imagen codificada en orden de descodificación, que puede usarse para la predicción inter de la imagen asociada o cualquier imagen que siga a la imagen asociada en orden de descodificación. En el HM, se señala directamente un RPS para cada imagen codificada. Los elementos sintácticos para señalar el RPS se incluyen tanto en el SPS como en la cabecera de fragmento. Para una imagen codificada en particular, el RPS puede ser uno de los posibles conjuntos de imágenes incluidos en el SPS, como lo indica un indicador en la cabecera de fragmento, o está directamente señalado en la cabecera de fragmento.

**[0147]** El RPS para cada imagen comprende cinco listas diferentes de imágenes de referencia, también denominados los cinco subconjuntos RPS. Los cinco subconjuntos RPS incluyen: `RefPicSetStCurrBefore`, `RefPicSetStCurrAfter`, `RefPicSetStFoll`, `RefPicSetLtCurr` y `RefPicSetLtFoll`. `RefPicSetStCurrBefore` incluye todas las imágenes de referencia a corto plazo (STRP) que son anteriores a la imagen actual tanto en orden de descodificación como en orden de salida, y que pueden usarse en la predicción inter de la imagen actual. `RefPicSetStCurrAfter` consta de todas las imágenes de referencia a corto plazo que son anteriores a la imagen actual en orden de descodificación, que vienen a continuación de la imagen actual en orden de salida, y que pueden usarse en predicción inter de la imagen actual. `RefPicSetStFoll` consta de todas las imágenes de referencia a corto plazo que se pueden usar en la predicción inter de una o más de las imágenes que siguen a la imagen actual en orden de descodificación, y que no se usan en la predicción inter de la imagen actual. `RefPicSetLtCurr` consta de todas las imágenes de referencia a largo plazo (LTRP) que se pueden utilizar en la predicción inter de la imagen actual. `RefPicSetLtFoll` consta de todas las imágenes de referencia a largo plazo que se pueden usar en la predicción inter de una o más de las imágenes que siguen a la imagen actual en orden de descodificación, y que no se usan en la predicción inter de la imagen actual.

**[0148]** Se ha descrito, en K. Suehring et al. "Long-term Picture Signaling for error-free environments [Señalización de imagen a largo plazo para entornos sin errores]" 9.ª Reunión, Ginebra, Suiza, 27 de abril - 7 de mayo de 2012, Doc., JCTVC-I0112 (en adelante, "Suehring"), para identificar un LTRP mediante un índice a la lista ordenada de recuento de orden de imágenes (POC) de imágenes de referencia en la memoria intermedia de imágenes descodificadas (DPB) que están: 1) marcadas como "usadas como referencia" y 2) no incluidas en los conjuntos de imágenes de referencia a corto plazo. Suehring afirma que, en un entorno sin errores, tal señalización de las LTRP es eficiente. Sin embargo, en un entorno sin errores, las STRP también pueden identificarse mediante índices a la lista ordenada de POC de imágenes de referencia en la DPB.

**[0149]** Esta divulgación describe técnicas que pueden usarse para identificar las STRP. En un ejemplo, esta divulgación describe un dispositivo de procesamiento de vídeo configurado para identificar una STRP mediante un índice a la lista ordenada de POC de imágenes de referencia en la DPB. En particular, en algunos ejemplos, cada STRP puede identificarse mediante un índice a la lista ordenada de POC de imágenes de referencia en la DPB. En algunos casos, el orden puede estar en orden creciente de valores POC. En otros casos, el orden puede estar en orden decreciente de valores POC. Para la primera STRP indicada, el índice se codifica directamente, por ejemplo, usando uno de la codificación de enteros sin signo exp-Golomb (ue (v)), la codificación de enteros truncados exp-Golomb (te (v)), o la codificación de enteros sin signo (u (v)). Si la codificación es u(v), entonces la longitud, en bits, del índice es  $\text{Ceil}(\log_2(N))$ , donde N es el número de imágenes de referencia en la DPB. Para otra STRP indicada, se señala la diferencia entre el índice de la STRP actual y el índice de la STRP anterior. De forma alternativa, las imágenes de referencia en la DPB se pueden dividir en dos listas. En este ejemplo, una de las listas incluye una imagen de referencia con un POC menor que el POC de la imagen actual y puede ordenarse en orden decreciente de valores de POC. La otra de las listas incluye imágenes de referencia con un valor POC mayor que el valor POC de la imagen actual y puede ordenarse en orden creciente de valores POC. A continuación, cada STRP que se incluirá en el conjunto de imágenes de referencia a corto plazo puede ser identificado por la lista (es decir, una de las dos listas) y el índice a esa lista.

**[0150]** Ahora se describirán los problemas y las técnicas propuestas relacionadas con procesamiento por lotes de datos con prefijo de categoría para los esquemas de división de imagen de procesamiento paralelo de frente de onda y elementos (WPP). Como se describió anteriormente, el HEVC HM incluye cuatro esquemas diferentes de división de imágenes, a saber, fragmentos regulares, fragmentos de entropía, elementos y WPP. Los fragmentos regulares son similares a los de H.264/AVC. Cada fragmento regular se encapsula en su propia unidad NAL, y la predicción dentro de la imagen (por ejemplo, la predicción dentro de la muestra, la predicción de la información de movimiento, la predicción del modo de codificación) y la dependencia de la codificación por entropía a través de los límites de los fragmentos se deshabilitan. Por lo tanto, una trama regular puede reconstruirse independientemente de otras tramas regulares dentro de la misma imagen (aunque todavía puede haber interdependencias debido a las operaciones de filtrado de bucle). Dado que la unidad básica en HEVC, es decir, un bloque de árbol, puede ser de un tamaño relativamente grande (por ejemplo, un bloque de árbol puede ser de 64x64), se incluye un concepto de "fragmentos de granularidad fina" en el HM para permitir que el tamaño de MTU coincida en los límites de fragmento dentro de un bloque de árbol, como una forma especial de fragmentos regulares. La granularidad del fragmento se señala en el conjunto de parámetros de imagen, mientras que la dirección de un fragmento de granularidad fina aún se señala en la cabecera de fragmento.

**[0151]** Los fragmentos de entropía, al igual que los fragmentos regulares, rompen las dependencias de descodificación por entropía, pero permiten que la predicción (y el filtrado) en la imagen crucen los límites de los fragmentos de entropía. Por lo tanto, los fragmentos de entropía se pueden usar como un mecanismo ligero para paralelizar la descodificación por entropía sin tener un impacto en otros pasos de descodificación. Si bien cada fragmento de entropía está encapsulado en su propia unidad NAL, tiene una cabecera de fragmento mucho más corta en comparación con un fragmento regular, ya que la mayoría de los elementos sintácticos de cabecera de fragmento no están presentes y deben heredarse de la cabecera de fragmento completa anterior. Debido a que se permite la predicción en imagen entre los fragmentos de entropía contiguos dentro de una imagen, la comunicación requerida entre el procesadores/entre núcleos para habilitar la predicción en imagen puede ser sustancial. No se recomienda que los fragmentos de entropía se utilicen para la coincidencia de tamaño de MTU en entornos propensos a errores, ya que debido a la predicción en imagen, una pérdida de un fragmento de entropía también provoca el fallo de la descodificación del siguiente fragmento de entropía en orden de descodificación. Los fragmentos de entropía parecen tener un uso limitado para las arquitecturas de sistemas que ejecutan el proceso de descodificación por entropía en una arquitectura de múltiples núcleos o de múltiples CPU, pero ejecutan la función de descodificación restante en un hardware de procesamiento de señales dedicado.

**[0152]** En WPP, la imagen se divide en filas de bloques de árbol. La descodificación y predicción por entropía pueden usar datos de bloques de árbol en otras divisiones. El procesamiento paralelo es posible a través de la descodificación paralela de filas de bloques de árbol, donde el inicio de la descodificación de una fila se retarda dos bloques de árbol, para asegurar que los datos relacionados con un bloque de árbol arriba y a la derecha del bloque de árbol del asunto estén disponibles antes que el bloque de árbol está siendo descodificado. Usando este inicio escalonado (que aparece como un frente de onda cuando se representa gráficamente), la paralelización es posible con tantos procesadores/núcleos como la imagen contenga filas de bloques de árbol. El HEVC HM incluye un mecanismo para organizar los bits codificados de diferentes filas de bloques de árbol para ser beneficiosos para un número particular de procesadores/núcleos paralelos. Por ejemplo, es posible que todos los bits codificados de números pares de filas de bloques de árbol (filas de bloques de árbol 0, 2, 4,...) aparezcan antes de los bits codificados de números impares de filas de bloques de árbol (filas de bloques de árbol 1, 3, 5,...), de modo que el flujo de bits sea descodificable por dos procesadores/núcleos paralelos, aunque la descodificación de una fila de bloque de árbol que viene más temprano (por ejemplo, la fila 2 del bloque de árbol) se refiere a una fila de bloque de árbol que viene más adelante (por ejemplo, la fila 1 del bloque de árbol). De manera similar a los fragmentos de entropía, debido a la tolerancia de la predicción en imagen entre las filas de bloques de árbol contiguas dentro de una imagen, la comunicación necesaria entre procesadores/entre núcleos para permitir la predicción en imagen puede ser sustancial. La división WPP no da como

resultado más unidades NAL en comparación con cuando no se aplica; por lo tanto, WPP no se puede utilizar para la coincidencia de tamaño MTU.

**[0153]** Los elementos definen límites horizontales y verticales que dividen una imagen en columnas y filas de elementos. El orden de escaneo de los bloques de árbol se cambia para que sea local dentro de un elemento (en el orden de un escaneo de ráster de bloque de árbol de un elemento), antes de descodificar el bloque de árbol superior izquierdo del siguiente elemento en el orden de escaneo de ráster de elementos de una imagen. Al igual que los fragmentos regulares, los elementos rompen las dependencias de predicción en imagen, así como las dependencias de descodificación por entropía. Sin embargo, no es necesario que los elementos se incluyan en unidades NAL individuales (al igual que WPP en este aspecto) y, por lo tanto, no se pueden usar para la coincidencia de tamaño de MTU. Cada elemento puede procesarse mediante un procesador/núcleo, y la comunicación entre procesadores/entre núcleos requerida para la predicción en imagen entre las unidades de procesamiento que descodifican los elementos contiguos se limita a transmitir la cabecera de fragmento compartida, en los casos en que un fragmento se extiende más de un elemento, y el filtrado en bucle está relacionado con el intercambio de muestras reconstruidas y metadatos. Por lo tanto, los elementos son menos exigentes en términos de ancho de banda de memoria en comparación con WPP debido a la independencia en imagen entre dos divisiones contiguas.

**[0154]** En el HEVC HM, el fragmento regular es la única herramienta que se puede utilizar para la paralelización que también está disponible, de forma prácticamente idéntica, en H.264/AVC. La paralelización basada en fragmentos regulares no requiere mucha comunicación entre procesadores ni entre núcleos (excepto para el intercambio de datos entre procesadores o entre núcleos para la compensación de movimiento cuando se descodifica una imagen codificada de manera predecible, que es típicamente mucho más pesado que intercambio de datos entre procesadores o entre núcleos debido a la predicción en imagen). Sin embargo, por la misma razón, los fragmentos regulares pueden requerir cierta sobrecarga de codificación. Además, los fragmentos regulares (en contraste con algunos de los otros esquemas de división mencionados anteriormente) también sirven como mecanismo clave para que la división del flujo de bits coincida con los requisitos de tamaño de MTU, debido a la independencia en imagen de los fragmentos regulares y que cada fragmento regular está encapsulado su propia unidad NAL. En muchos casos, el objetivo de la paralelización y el objetivo de la coincidencia de tamaño de MTU pueden imponer demandas contradictorias al diseño de fragmentos en una imagen.

**[0155]** Por simplicidad, las restricciones de los cuatro esquemas diferentes de división de imágenes se especifican en el HM. Por ejemplo, en el HM, no se puede aplicar ninguna combinación de dos de los fragmentos, elementos y WPP de entropía al mismo tiempo. Además, en el HM, para cada fragmentos y elemento, se deben cumplir una o las dos condiciones siguientes: 1) todos los bloques de árbol codificados en un fragmento pertenecen al mismo elemento, y 2) todos los bloques de árbol codificados en un elemento pertenecen al mismo fragmento.

**[0156]** Un orden de datos intercalados para elemento y WPP, que también hace que los dos procedimientos de señalización de punto de entrada sean menos útiles, se ha descrito en S. Kanumuri, et al. "Category-prefixed data batching for tiles and wavefronts [Procesamiento por lotes de datos con prefijo de categoría para elementos y frentes de onda]" 9.ª Reunión, Ginebra, Suiza 27 de abril - 7 de mayo de 2012, Doc., JCTVC-I0427 (en adelante "Kanumuri"). Un problema asociado con el enfoque propuesto en Kanumuri es que el número máximo de categorías admitidas es 32. Sin embargo, es posible que haya más de 32 sub-flujos WPP (por ejemplo, cuando cada fila de bloque de árbol es un sub-flujo y hay más de 32 filas de bloque de árbol) o más de 32 elementos, y se requiere que se conozca cada punto de entrada de elemento o sub-flujo WPP, por ejemplo, para procesar bloques de árbol (tanto de análisis como de descodificación) en el escaneo de ráster de bloques de árbol de una imagen mediante un descodificador de un solo núcleo. Además, el enfoque propuesto en Kanumuri impone una carga en la descodificación de un solo núcleo, ya que el proceso de fusión para fusionar datos intercalados o saltar hacia atrás y hacia adelante en el flujo de bits de una unidad NAL es necesario, dependiendo del orden de procesamiento.

**[0157]** Esta divulgación describe técnicas que pueden usarse para indicar si se aplican el procesamiento por lotes de datos con prefijo de categoría. En un ejemplo, esta divulgación describe un dispositivo de procesamiento de vídeo configurado para identificar con un indicador si se aplica el procesamiento por lotes de datos con prefijo de categoría. En particular, en algunos ejemplos, se puede agregar una indicación en el SPS o PPS para indicar si se aplica el procesamiento por lotes de datos con prefijo de categoría. Si se indica que no se aplica el procesamiento por lotes de datos con prefijo de categoría, entonces se puede aplicar la forma convencional de manejar los elementos. En otros ejemplos, el número máximo de categorías puede variar, y el valor del número máximo de categorías puede señalarse en el SPS o PPS. En ejemplos adicionales, la asignación de elementos o sub-flujos de WPP a categorías (es decir, qué elementos o sub-flujos de WPP pertenecen a cada categoría) puede señalarse, por ejemplo, en el SPS o PPS.

**[0158]** La FIG. 6 es un diagrama de flujo que ilustra una operación para determinar si una actualización del conjunto de parámetros puede ocurrir en una parte de un flujo de bits. La operación ilustrada se describe como realizada por el descodificador de vídeo 30 de la FIG. 3 basándose en una indicación generada por el codificador de vídeo 20 de la FIG. 2.

**[0159]** El descodificador de vídeo 30 puede recibir un flujo de bits del codificador de vídeo 20 con una o más secuencias de vídeo codificadas que incluyen múltiples unidades de acceso que representan fragmentos de vídeo codificados y

conjuntos de parámetros codificados. El descodificador de vídeo 30 descodifica un indicador definido para indicar si una actualización del conjunto de parámetros puede ocurrir en una parte del flujo de bits, como una de las secuencias de vídeo codificadas (130). En un ejemplo, el indicador puede incluirse en una unidad SEI NAL incluida en al menos una unidad de acceso de la secuencia de vídeo codificada. En otro ejemplo, el indicador puede incluirse en una unidad SPS NAL incluida en al menos una unidad de acceso de la secuencia de vídeo codificada.

**[0160]** En algunos casos, el indicador puede indicar si los conjuntos de parámetros de cualquier tipo pueden actualizarse en la secuencia de vídeo codificada, de tal manera que solo un indicador puede ser utilizado para todos los tipos de conjuntos de parámetros. En otros casos, el indicador puede indicar si un conjunto de parámetros de un tipo particular puede actualizarse en la secuencia de vídeo codificada, de manera que se pueda usar un indicador diferente para cada tipo de conjunto de parámetros. En otros casos, el indicador puede indicar si los conjuntos de parámetros de dos o más tipos pueden actualizarse en la secuencia de vídeo codificada, de manera que un indicador se puede utilizar para indicar el estado de actualización para los dos o más tipos de conjuntos de parámetros y otro indicador puede ser utilizado para indicar el estado de actualización para tipos adicionales de conjuntos de parámetros.

**[0161]** El descodificador de vídeo 30 luego descodifica un primer conjunto de parámetros de un tipo particular con un valor de ID particular de una unidad NAL de conjunto de parámetros incluida en una de las unidades de acceso de la secuencia de vídeo codificada (132). Cuando el indicador indica que no puede ocurrir una actualización de conjunto de parámetros en la secuencia de vídeo codificada para el tipo particular de conjunto de parámetros (rama NO de 134), el descodificador de vídeo 30 activa el primer conjunto de parámetros para la parte completa del flujo de bits, es decir, la secuencia de vídeo codificada en este ejemplo (136). En este caso, al descodificar un conjunto de parámetros posterior del mismo tipo y tener un mismo valor de identificación que el primer conjunto de parámetros, el descodificador de vídeo 30 no necesita determinar si actualizar el primer conjunto de parámetros para la secuencia de vídeo codificada y puede ignorar el conjunto de parámetros posteriores. Por lo tanto, el descodificador de vídeo 30 puede funcionar como si el contenido del primer conjunto de parámetros fuera idéntico al contenido incluido en cualquier otro conjunto de parámetros del mismo tipo que tenga el mismo valor de ID.

**[0162]** Cuando el indicador indica que se puede producir una actualización del conjunto de parámetros en la secuencia de vídeo codificada para el tipo particular de conjunto de parámetros (rama SÍ de 134), el descodificador de vídeo 30 activa el primer conjunto de parámetros (138). El descodificador de vídeo 30 luego descodifica un segundo conjunto de parámetros del mismo tipo y con el mismo valor de ID que el primer conjunto de parámetros de una unidad NAL de conjunto de parámetros incluida en una de las unidades de acceso de la secuencia de vídeo codificada (140). En este caso, debido a que pueden ocurrir actualizaciones de conjuntos de parámetros, el descodificador de vídeo 30 determina si almacenar el segundo conjunto de parámetros para reemplazar el primer conjunto de parámetros (142).

**[0163]** En algunos casos, el descodificador de vídeo 30 puede activar automáticamente el segundo conjunto de parámetros para la secuencia de vídeo codificada, y desactivar el primer conjunto de parámetros. En otros casos, el descodificador de vídeo 30 puede comparar el contenido del primer conjunto de parámetros con el contenido del segundo conjunto de parámetros. Cuando el contenido del segundo conjunto de parámetros es diferente al contenido del primer conjunto de parámetros, el descodificador de vídeo 30 activa el segundo conjunto de parámetros para la secuencia de vídeo codificada, y desactiva el primer conjunto de parámetros.

**[0164]** La FIG. 7 es un diagrama de flujo que ilustra una operación para determinar si se puede realizar el acceso aleatorio desde una unidad de acceso CRA sin tener que obtener conjuntos de parámetros de unidades de acceso no disponibles anteriores. La operación ilustrada se describe como realizada por el descodificador de vídeo 30 de la FIG. 3 basándose en una indicación generada por el codificador de vídeo 20 de la FIG. 2.

**[0165]** El HEVC HM introdujo imágenes CRA para proporcionar más opciones para el acceso aleatorio al flujo de bits sin la menor eficiencia de codificación, como ocurre con imágenes IDR. Las unidades de acceso CRA se pueden colocar en cualquier punto dentro de una secuencia de vídeo codificada y no limpian la memoria intermedia de imágenes descodificadas (DPB), lo cual permite a las imágenes iniciales de una unidad de acceso CRA determinada (es decir, imágenes que siguen la unidad de acceso CRA en orden de descodificación pero preceden a la unidad de acceso CRA en orden de salida) usar imágenes descodificadas antes de la unidad de acceso CRA como imágenes de referencia. Sin embargo, las imágenes iniciales solo pueden descodificarse correctamente cuando el acceso aleatorio comienza desde una imagen RAP descodificada antes de la unidad de acceso CRA dada. En el caso de que el acceso aleatorio se inicie desde la imagen CRA particular, todas las imágenes que preceden a la imagen CRA en orden de salida o descodificación no se descodificarán y no estarán disponibles para su uso como imágenes de referencia. Para descodificar la imagen CRA o cualquier unidad de acceso posterior en orden de salida o descodificación, el descodificador de vídeo 30 puede necesitar obtener uno o más conjuntos de parámetros incluidos en una unidad de acceso para una de las imágenes anteriores no disponibles, lo cual es una operación pesada.

**[0166]** El descodificador de vídeo 30 puede recibir un flujo de bits del codificador de vídeo 20 con una o más secuencias de vídeo codificadas que incluyen múltiples unidades de acceso que representan fragmentos de vídeo codificados y conjuntos de parámetros codificados. El descodificador de vídeo 30 descodifica una o más unidades de acceso CRA a partir de una secuencia de vídeo codificada del flujo de bits (150). El descodificador de vídeo 30 también descodifica

un indicador definido para indicar si el acceso aleatorio al flujo de bits desde una unidad de acceso CRA particular requiere conjuntos de parámetros de unidades de acceso anteriores, no disponibles (152). En un ejemplo, el indicador puede incluirse en una unidad SEI NAL incluida en la unidad de acceso CRA particular en la secuencia de vídeo codificada del flujo de bits.

**[0167]** Cuando el indicador indica que no es necesario obtener conjuntos de parámetros de unidades de acceso anteriores (es decir, no se requiere búsqueda) para el acceso aleatorio desde la unidad de acceso CRA particular (rama NO de 154), el descodificador de vídeo 30 realiza un acceso aleatorio al flujo de bits desde la unidad de acceso CRA particular sin realizar una operación de búsqueda (156). A la inversa, cuando el indicador indica que los conjuntos de parámetros de las unidades de acceso anteriores deben ser obtenidos (es decir, se requiere búsqueda) para el acceso aleatorio desde la imagen CRA particular (rama SÍ de 154), el descodificador de vídeo 30 determina si realizar el acceso aleatorio a la flujo de bits desde la unidad de acceso CRA particular en orden (158). En algunos casos, el descodificador de vídeo 30 puede determinar que el acceso aleatorio al flujo de bits está disponible desde otra imagen RAP sin realizar una operación de búsqueda, y además determinar que no se realice el acceso aleatorio al flujo de bits desde la unidad de acceso CRA particular para evitar la búsqueda de conjuntos de parámetros desde las unidades de acceso anteriores. En otros casos, el descodificador de vídeo 30 puede determinar que el acceso aleatorio al flujo de bits desde otra imagen RAP no está disponible, y determinar además realizar un acceso aleatorio al flujo de bits desde la unidad de acceso CRA particular después de obtener los conjuntos de parámetros de las unidades de acceso anteriores.

**[0168]** De esta manera, basándose en el indicador, el descodificador de vídeo 30 primero determina si se requiere una operación de búsqueda para la descodificación de acceso aleatorio desde la unidad de acceso CRA particular, y puede evitar realizar una operación de búsqueda seleccionando una imagen RAP diferente desde la cual realizar acceso aleatorio. La indicación es especialmente útil para las unidades de acceso CRA, ya que las unidades de acceso CRA pueden ubicarse en cualquier punto dentro de una secuencia de vídeo codificada y no limpiar la DPB, lo cual brinda más opciones para el acceso aleatorio al flujo de bits sin una menor eficiencia de codificación.

**[0169]** La FIG. 8 es un diagrama de bloques que ilustra un conjunto de dispositivos de ejemplo que forman parte de la red 300. En este ejemplo, la red 300 incluye los dispositivos de encaminamiento 304A, 304B (colectivamente denominados "dispositivos de encaminamiento 304") y el dispositivo de transcodificación 306. Los dispositivos 304 de encaminamiento y el dispositivo 306 de transcodificación están concebidos para representar un pequeño número de dispositivos que pueden formar parte de la red 300. Otros dispositivos de red, tales como conmutadores, concentradores, pasarelas, cortafuegos, puentes y otros dispositivos de ese tipo también pueden estar incluidos en la red 300. Además, pueden proporcionarse dispositivos de red adicionales a lo largo de un trayecto de red entre el dispositivo servidor 302 y el dispositivo cliente 308. En algunos ejemplos, el dispositivo servidor 302 puede corresponder al dispositivo de origen 12 de la FIG. 1, mientras que el dispositivo cliente 308 puede corresponder al dispositivo de destino 14 de la FIG. 1.

**[0170]** En general, los dispositivos de encaminamiento 304 implementan uno o más protocolos de encaminamiento para intercambiar datos de red a través de la red 300. En algunos ejemplos, los dispositivos 304 de encaminamiento pueden estar configurados para realizar operaciones de *proxy* o de memoria caché. Por lo tanto, en algunos ejemplos, los dispositivos de encaminamiento 304 pueden denominarse también dispositivos *proxy*. En general, los dispositivos 304 de encaminamiento ejecutan protocolos de encaminamiento para descubrir rutas a través de la red 300. Al ejecutar dichos protocolos de encaminamiento, el dispositivo 304B de encaminamiento puede descubrir una ruta de red desde sí mismo hasta el dispositivo servidor 302, a través del dispositivo 304A de encaminamiento.

**[0171]** Las técnicas de esta divulgación pueden ser implementadas por dispositivos de red tales como los dispositivos de encaminamiento 304 y el dispositivo de transcodificación 306, pero también pueden ser implementadas por el dispositivo cliente 308. De esta manera, los dispositivos de encaminamiento 304, el dispositivo de transcodificación 306 y el dispositivo cliente 308 representan ejemplos de dispositivos configurados para implementar una o más de las técnicas descritas en esta divulgación.

**[0172]** En uno o más ejemplos, las funciones descritas se pueden implementar en hardware, software, firmware o cualquier combinación de los mismos. Si se implementan en software, las funciones pueden almacenarse en, y transmitirse por, un medio legible por ordenador, como una o más instrucciones o código, y ejecutarse mediante una unidad de procesamiento basada en hardware. Los medios legibles por ordenador pueden incluir medios de almacenamiento legibles por ordenador, que correspondan a un medio tangible tal como medios de almacenamiento de datos o medios de comunicación que incluyan cualquier medio que facilite la transferencia de un programa informático desde un lugar a otro, por ejemplo, de acuerdo con un protocolo de comunicación. De esta manera, los medios legibles por ordenador pueden corresponder en general a (1) medios de almacenamiento tangibles legibles por ordenador que sean no transitorios o (2) un medio de comunicación tal como una señal o una onda portadora. Los medios de almacenamiento de datos pueden ser medios disponibles cualesquiera a los que se pueda acceder desde uno o más ordenadores o uno o más procesadores para recuperar instrucciones, código y/o estructuras de datos para la implementación de las técnicas descritas en esta divulgación. Un producto de programa informático puede incluir un medio legible por ordenador.

[0173] A modo de ejemplo, y no de limitación, dichos medios de almacenamiento legibles por ordenador pueden comprender RAM, ROM, EEPROM, CD-ROM u otro almacenamiento de disco óptico, almacenamiento de disco magnético u otros dispositivos de almacenamiento magnético, memoria flash o cualquier otro medio que se pueda usar para almacenar un código de programa deseado en forma de instrucciones o estructuras de datos y al que se pueda acceder mediante un ordenador. Además, cualquier conexión recibe debidamente la denominación de medio legible por ordenador. Por ejemplo, si las instrucciones se transmiten desde una sede de la Red, un servidor u otro origen remoto usando un cable coaxial, un cable de fibra óptica, un par trenzado, una línea de abonado digital (DSL) o tecnologías inalámbricas tales como infrarrojos, radio y microondas, entonces el cable coaxial, el cable de fibra óptica, el par trenzado, la DSL o las tecnologías inalámbricas tales como infrarrojos, radio y microondas se incluyen en la definición de medio. Sin embargo, debería entenderse que los medios de almacenamiento legibles por ordenador y los medios de almacenamiento de datos no incluyen conexiones, ondas portadoras, señales u otros medios transitorios, sino que, en cambio, están orientados a medios de almacenamiento tangibles no transitorios. Los discos, como se usan en el presente documento, incluyen un disco compacto (CD), un disco láser, un disco óptico, un disco versátil digital (DVD), un disco flexible y un disco Blu-ray, donde algunos discos reproducen usualmente los datos magnéticamente, mientras que otros discos reproducen los datos ópticamente con láseres. Las combinaciones de lo anterior también deberían incluirse dentro del alcance de los medios legibles por ordenador.

[0174] Las instrucciones se pueden ejecutar por uno o más procesadores, tales como uno o más procesadores de señales digitales (DSP), microprocesadores de propósito general, circuitos integrados específicos de la aplicación (ASIC), matrices lógicas programables *in situ* (FPGA) u otros circuitos lógicos integrados o discretos equivalentes. En consecuencia, el término "procesador", como se usa en el presente documento, se puede referir a cualquiera de las estructuras anteriores o a cualquier otra estructura adecuada para la implementación de las técnicas descritas en el presente documento. Además, en algunos aspectos, la función descrita en el presente documento se puede proporcionar dentro de módulos de hardware y/o software dedicados, configurados para la codificación y la decodificación, o incorporados en un códec combinado. Además, las técnicas se podrían implementar totalmente en uno o más circuitos o elementos lógicos.

[0175] Las técnicas de esta divulgación se pueden implementar en una amplia variedad de dispositivos o aparatos, incluyendo un teléfono inalámbrico, un circuito integrado (IC) o un conjunto de IC (por ejemplo, un conjunto de chips). Diversos componentes, módulos o unidades se describen en esta divulgación para enfatizar aspectos funcionales de dispositivos configurados para realizar las técnicas divulgadas, pero no requieren necesariamente su realización mediante diferentes unidades de hardware. En cambio, como se ha descrito anteriormente, diversas unidades se pueden combinar en una unidad de hardware de códec, o ser proporcionadas por un grupo de unidades de hardware interoperativas, incluyendo uno o más procesadores, como se ha descrito anteriormente, juntamente con software y/o firmware adecuados.

## REIVINDICACIONES

1. Un procedimiento para descodificar datos de vídeo de codificación de vídeo de alta eficiencia (HEVC), el procedimiento que comprende:

descodificar (152), en un mensaje de información de mejora complementaria (SEI), un elemento sintáctico que indica si el acceso aleatorio al flujo de bits desde una unidad de acceso de acceso aleatorio limpio (CRA) en particular en una secuencia de vídeo codificada (120) del flujo de bits requiere uno o más conjuntos de parámetros de al menos una unidad de acceso anterior en el flujo de bits para descodificar la unidad de acceso CRA particular o las unidades de acceso posteriores, en el que la unidad de acceso CRA particular se coloca en cualquier punto dentro de la secuencia de vídeo codificada del flujo de bits y no limpia una memoria intermedia de imágenes descodificadas (DPB); y

basándose en el elemento sintáctico que indica que no se necesitan conjuntos de parámetros de una unidad de acceso anterior, realizar (156) acceso aleatorio al flujo de bits desde la unidad de acceso CRA particular sin obtener los conjuntos de parámetros de la al menos una unidad de acceso anterior en el flujo de bits.

2. Un procedimiento para codificar datos de vídeo de codificación de vídeo de alta eficiencia (HEVC), el procedimiento que comprende:

codificar una o más unidades de acceso de acceso aleatorio limpio (CRA) en una secuencia de vídeo codificada de un flujo de bits de datos de vídeo, en el que las unidades de acceso CRA están ubicadas en cualquier punto dentro de la secuencia de vídeo codificada del flujo de bits y no limpian una memoria intermedia de imágenes descodificadas (DPB); **caracterizado por**

codificar, en un mensaje de información de mejora suplementaria (SEI), un elemento sintáctico que indica si el acceso aleatorio al flujo de bits desde una unidad de acceso CRA particular en la secuencia de vídeo codificada del flujo de bits requiere uno o más conjuntos de parámetros de al menos una unidad de acceso anterior en el flujo de bits para descodificar la unidad de acceso CRA particular o las unidades de acceso posteriores, en el que, basándose en el elemento sintáctico que indica que no se necesitan conjuntos de parámetros de una unidad de acceso anterior, se puede realizar un acceso aleatorio al flujo de bits desde la unidad de acceso CRA particular sin obtener los conjuntos de parámetros de la al menos una unidad de acceso anterior en el flujo de bits.

3. Un dispositivo para descodificar datos de vídeo de codificación de vídeo de alta eficiencia (HEVC), el dispositivo que comprende:

medios para descodificar, en un mensaje de información de mejora suplementaria (SEI), un elemento sintáctico que indica si el acceso aleatorio al flujo de bits desde una unidad de acceso de acceso aleatorio limpio (CRA) en particular en una secuencia de vídeo codificada del flujo de bits requiere uno o más conjuntos de parámetros de la al menos una unidad de acceso anterior en el flujo de bits para descodificar la unidad de acceso CRA particular o las unidades de acceso posteriores, en el que la unidad de acceso CRA particular está situada en cualquier punto dentro de la secuencia de vídeo codificada del flujo de bits y no limpia una memoria intermedia de imágenes descodificadas (DPB); y

medios para, basándose en el elemento sintáctico que indica que no se necesitan conjuntos de parámetros de una unidad de acceso anterior, realizar un acceso aleatorio al flujo de bits desde la unidad de acceso CRA en particular sin obtener los conjuntos de parámetros de la al menos una unidad de acceso anterior en el flujo de bits.

4. Un dispositivo para codificar datos de vídeo de codificación de vídeo de alta eficiencia (HEVC), el dispositivo que comprende:

medios para codificar una o más unidades de acceso de acceso aleatorio limpio (CRA) en una secuencia de vídeo codificada de un flujo de bits de datos de vídeo, en el que las unidades de acceso CRA están ubicadas en cualquier punto dentro de la secuencia de vídeo codificada del flujo de bits y no limpian una memoria intermedia de imágenes descodificadas (DPB); **caracterizado por**

medios para codificar, en un mensaje de información de mejora suplementaria (SEI), un elemento sintáctico que indica si el acceso aleatorio al flujo de bits desde una unidad de acceso CRA particular en la secuencia de vídeo codificada del flujo de bits requiere uno o más conjuntos de parámetros de al menos una unidad de acceso en el flujo de bits para descodificar la unidad de acceso CRA particular o las unidades de acceso posteriores, en el que, basándose en el elemento sintáctico que indica que no se necesitan conjuntos de parámetros de una unidad de acceso anterior, el acceso aleatorio al flujo de bits desde la unidad de acceso CRA particular puede realizarse sin obtener los conjuntos de parámetros de la al menos una unidad de acceso anterior en el flujo de bits.

5. Un medio legible por ordenador que comprende instrucciones que, cuando se ejecutan, hacen que uno o más procesadores lleven a cabo un procedimiento de acuerdo con la reivindicación 1 o la reivindicación 2.



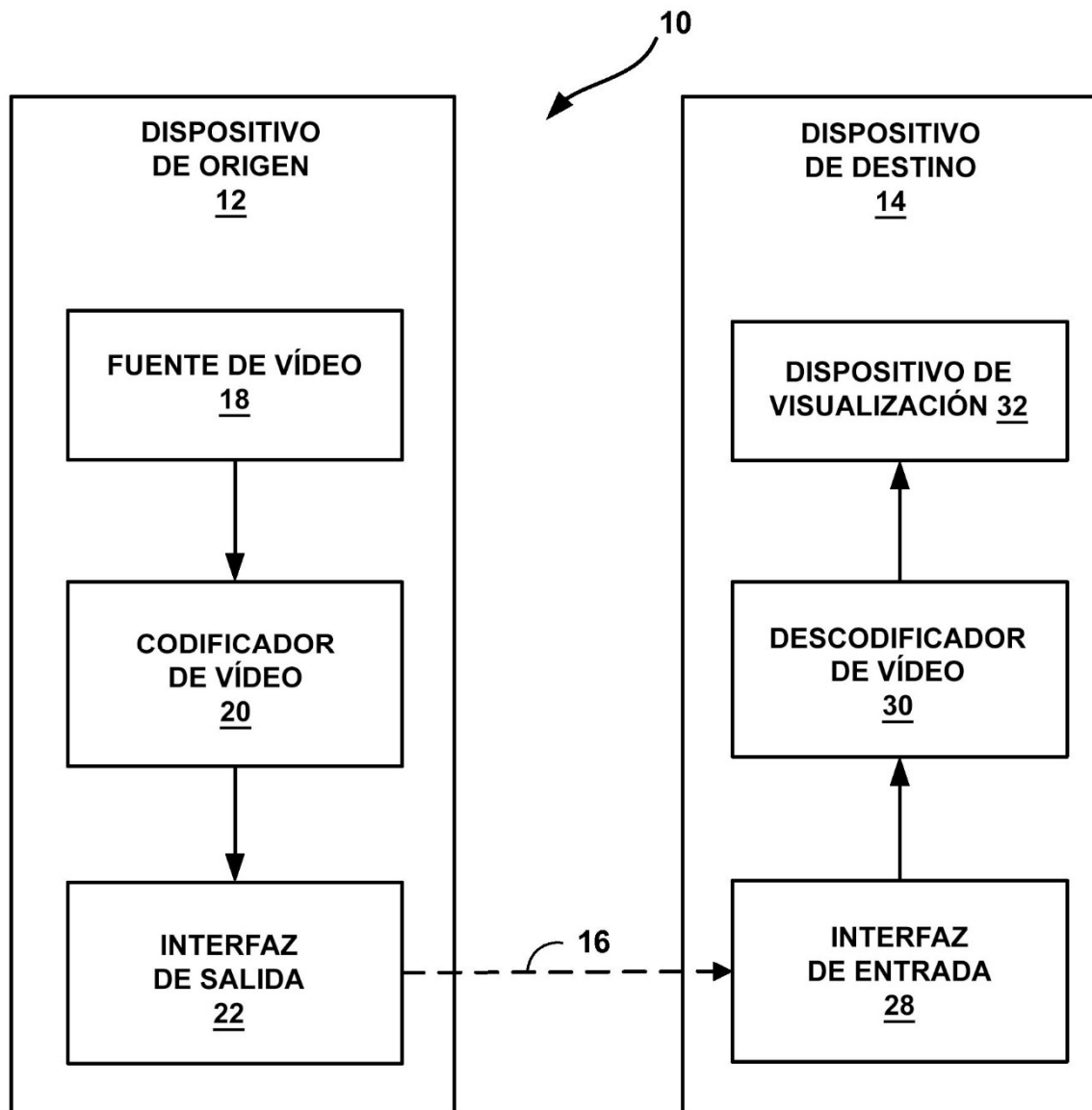
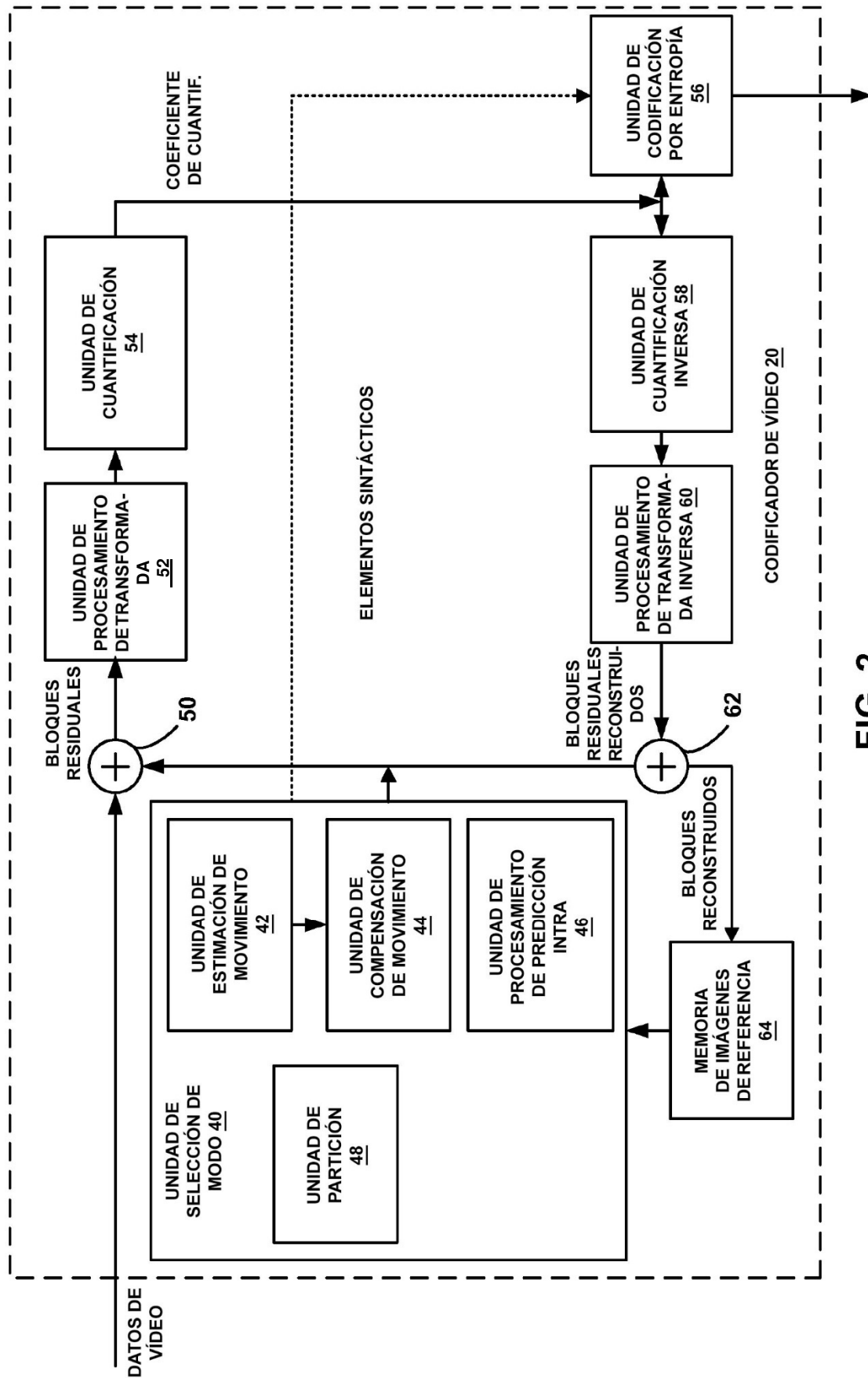


FIG. 1



**FIG. 2**

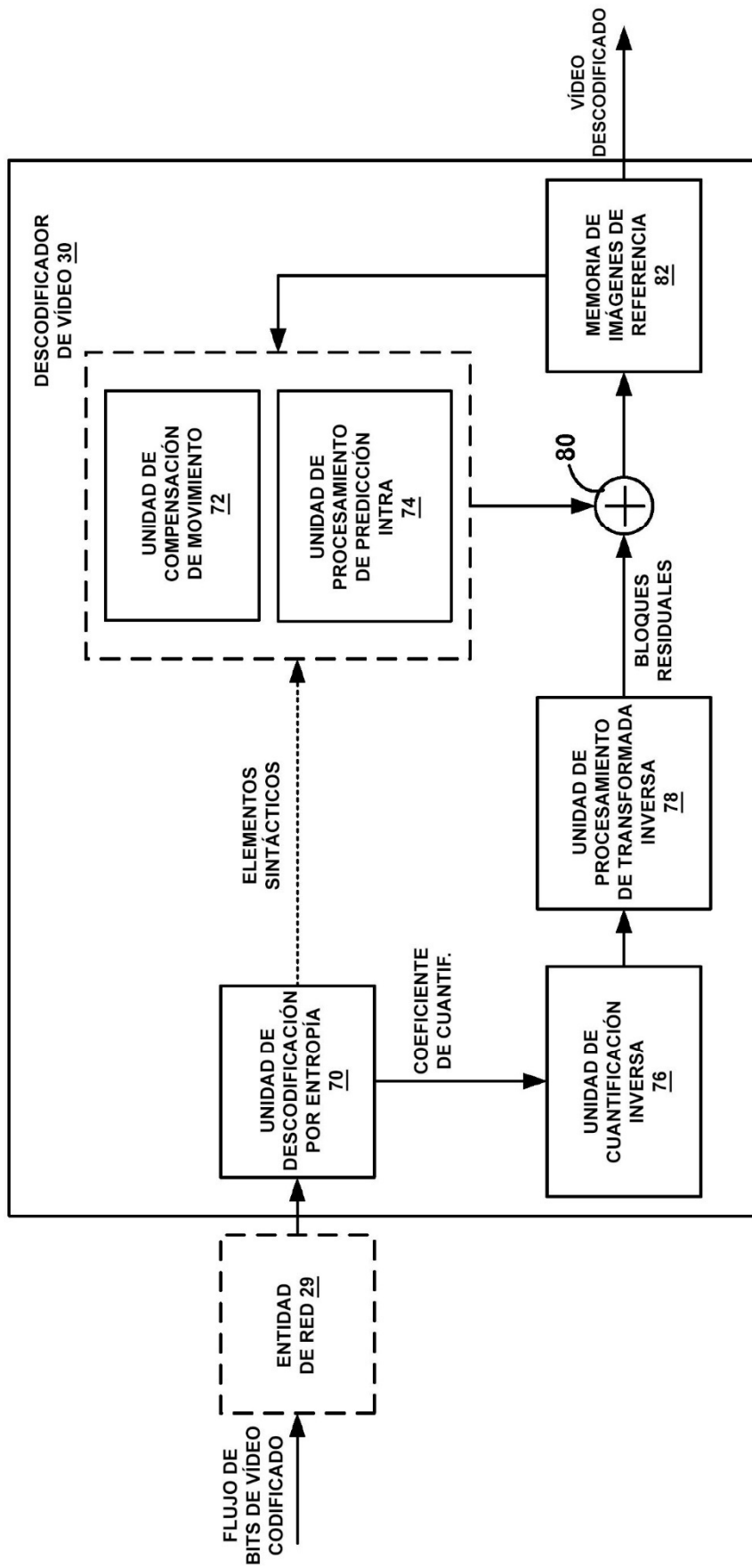


FIG. 3

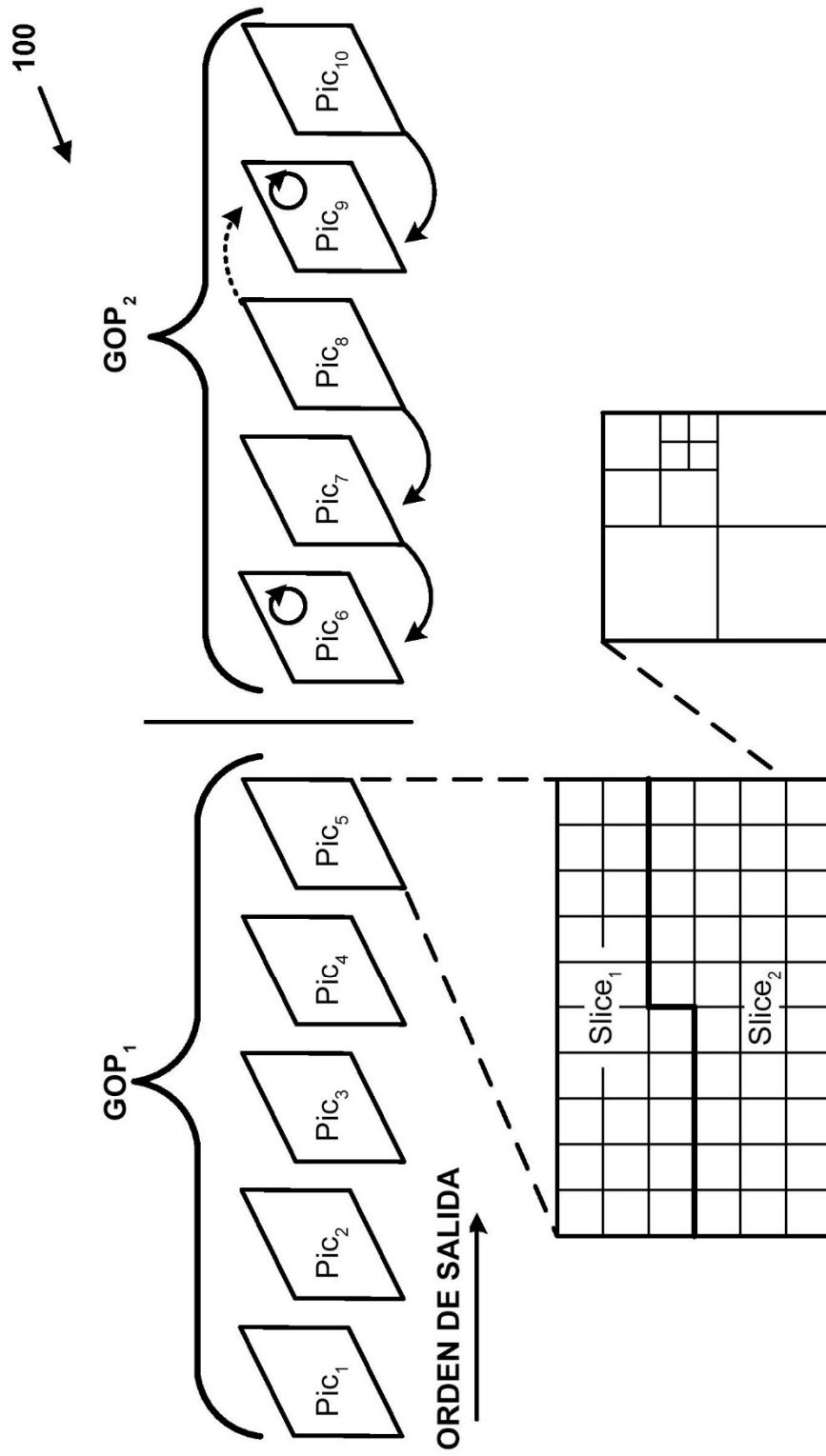


FIG. 4

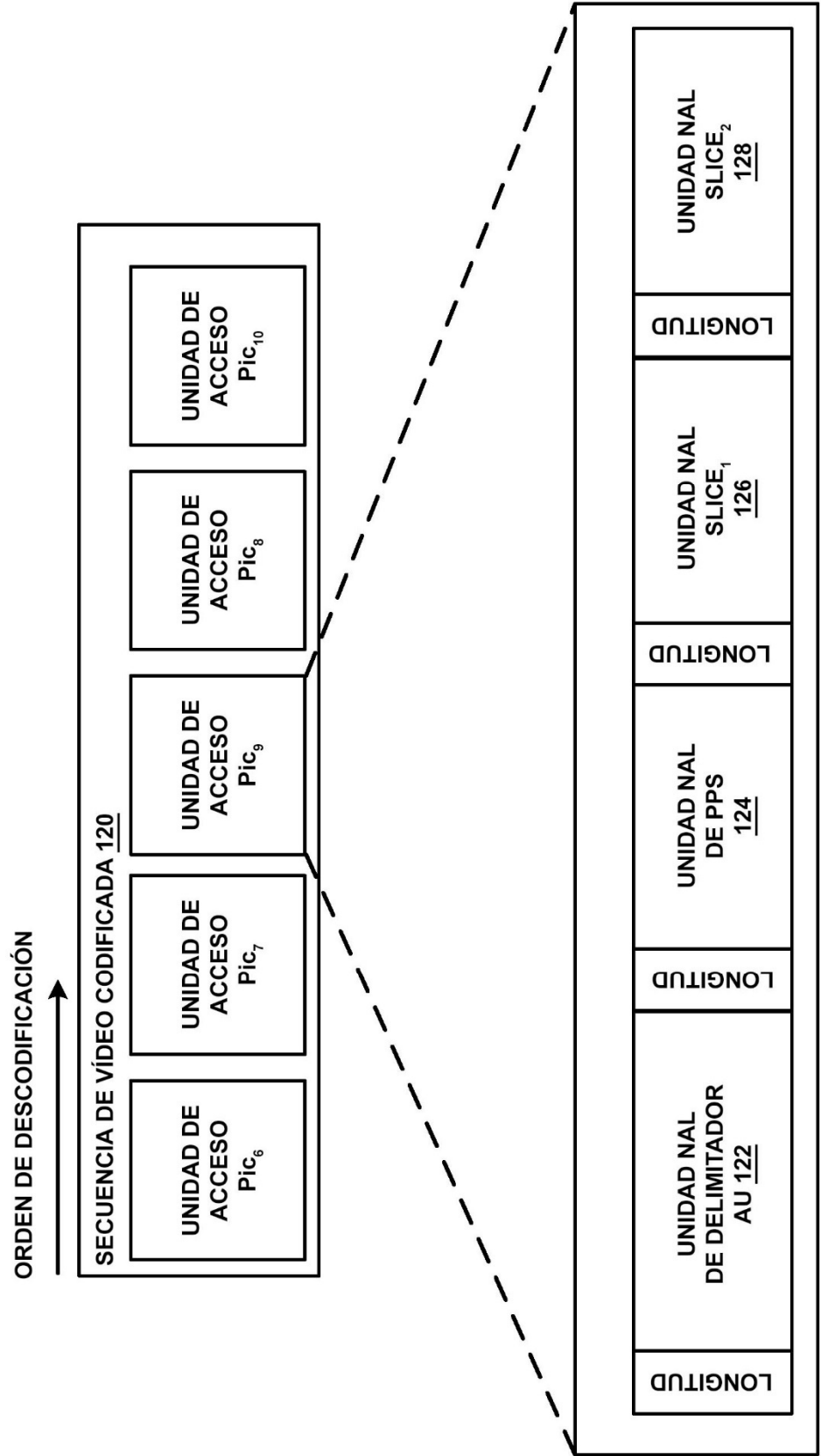
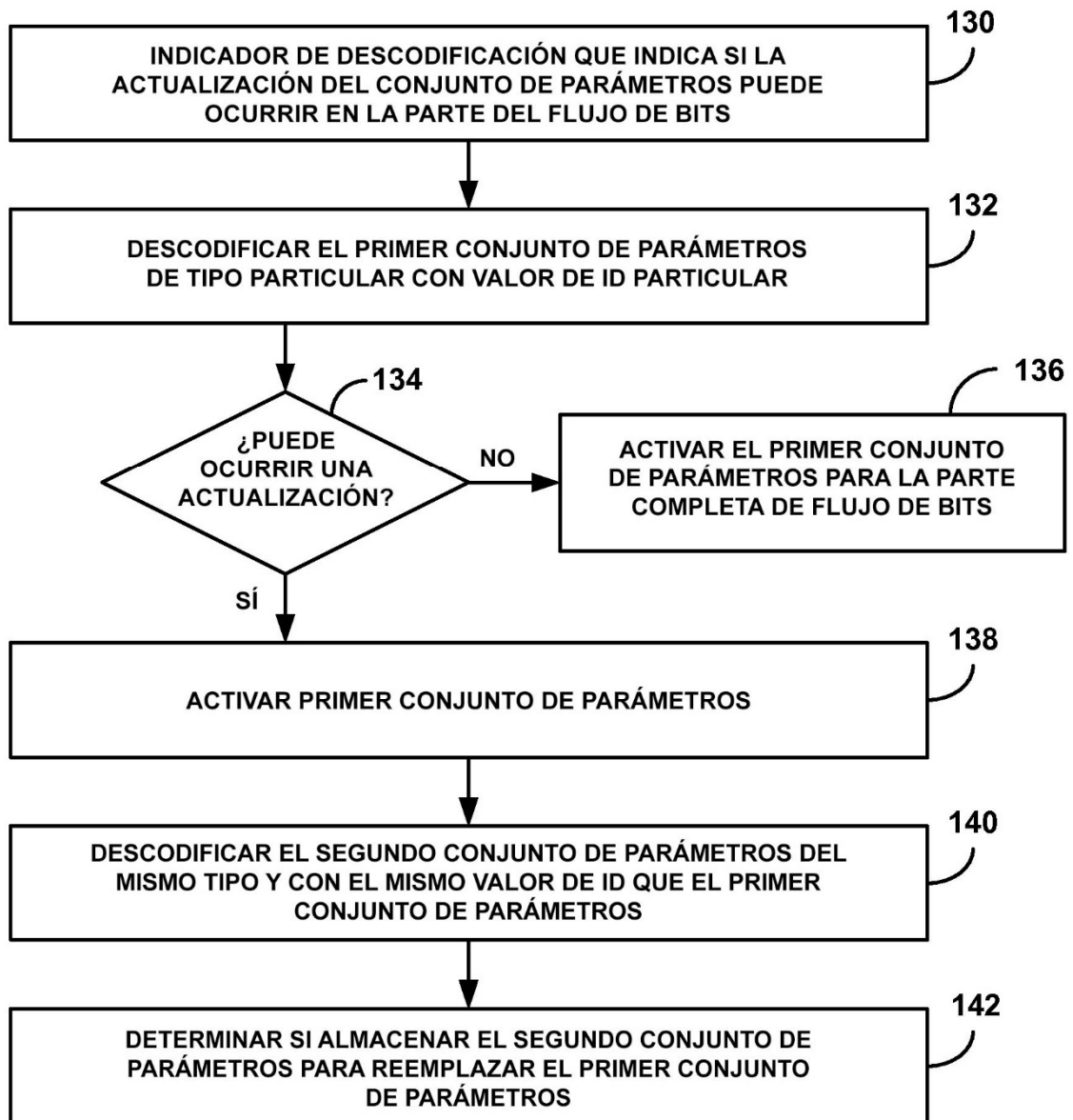
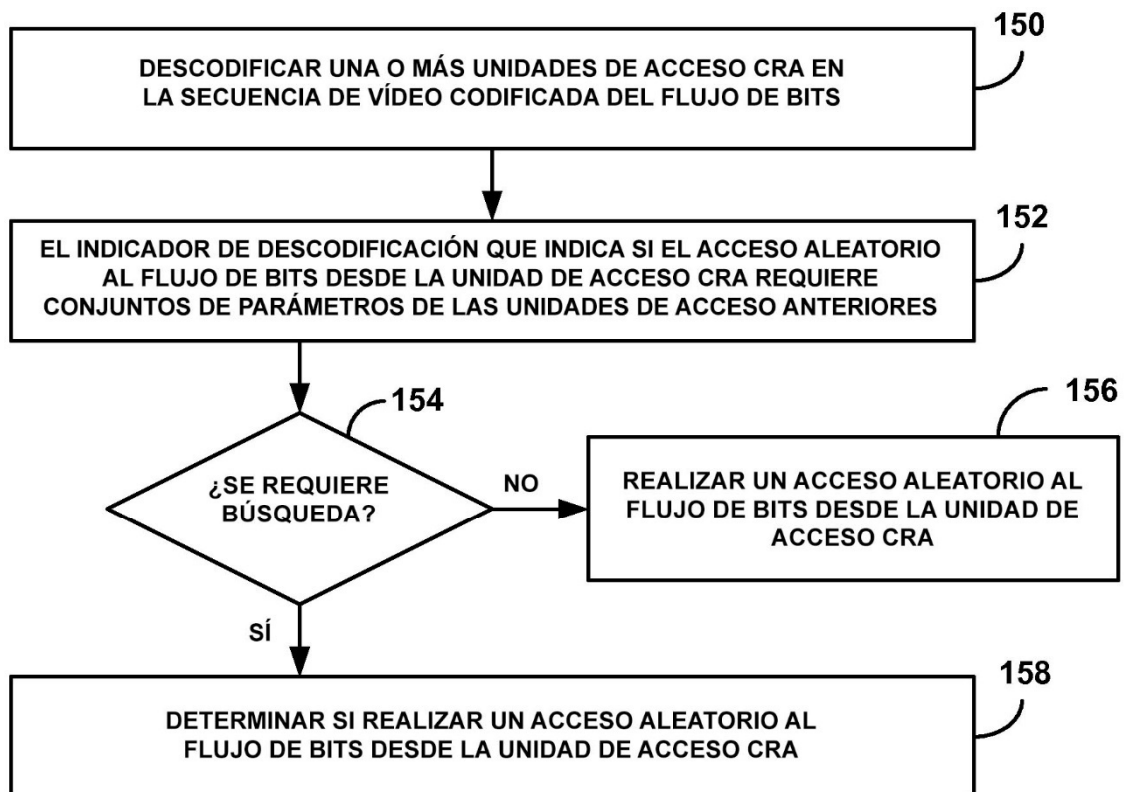


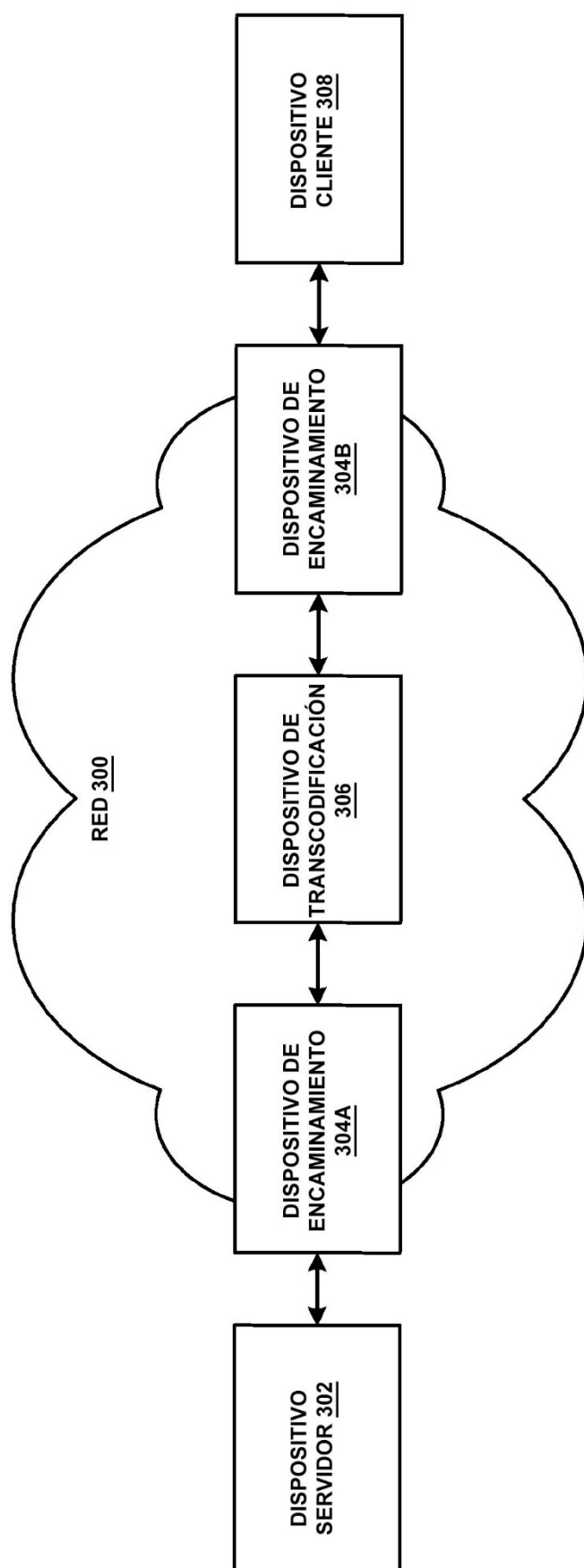
FIG. 5



**FIG. 6**



**FIG. 7**



**FIG. 8**