

## **ALMACENAMIENTO Y SUMINISTRO DE DATOS DE VIDEO PARA LA CODIFICACIÓN DE VIDEO**

**[0001]** Esta solicitud reivindica la prioridad de la solicitud estadounidense N.º 17/143.611, presentada el 7 de enero de 2021, que reivindica el beneficio de la solicitud de patente provisional estadounidense 62/958.561, presentada el 8 de enero de 2020, y el contenido completo se incorpora en la presente por referencia.

### **CAMPO TÉCNICO**

**[0002]** Esta divulgación se refiere al almacenamiento y suministro de datos de video

### **ANTECEDENTES**

**[0003]** Se pueden incorporar capacidades de video digitales en un amplio rango de dispositivos, incluso televisores digitales, sistemas de transmisión directa digital, sistemas de transmisión inalámbrica, asistentes digitales personales (PDA), computadoras portátiles o de escritorio, computadora tipo *tablet*, lectores de libros electrónicos, cámaras digitales, dispositivos de registro digital, reproductores de medios digitales, dispositivos de juegos de video, consolas de juegos de video, teléfonos celulares o de radio satélite, los denominados "teléfonos inteligentes", dispositivos de teleconferencia de video, dispositivos de *streaming* de video y similares. Los dispositivos de video digital implementan técnicas de codificación de video, tales como aquellas descritas en los estándares definidos por MPEG-2, MPEG-4, ITU-T H.263, ITU-T H.264/MPEG-4, parte 10, codificación de video avanzada (AVC), ITU-T H.265/codificación de video de alta eficiencia (HEVC) y las extensiones de estos estándares. Los dispositivos de video pueden transmitir, recibir, codificar, decodificar y/o almacenar información de video digital de manera más eficaz implementando las técnicas de codificación de video.

**[0004]** Las técnicas de codificación de video incluyen la predicción espacial (intraimagen) y/o la predicción temporal (interimagen) para reducir o eliminar la redundancia inherente en secuencias de video. Para la codificación de video basada en bloques, un plano de video (por ejemplo, una imagen de video o una porción de una imagen de video) se pueden dividir en bloques de video, que también se pueden denominar unidades de árboles de codificación (CTU), unidades de codificación (CU) y/o nodos de codificación. Los bloques de video en una porción intracodificada (I) de una imagen se codifican mediante el uso de predicción espacial con respecto a las muestras de referencia en bloques vecinos de la misma imagen. Los bloques de video

en una porción intercodificada (P o B) de la imagen pueden usar predicción espacial con respecto a las muestras de referencia en bloques vecinos de la misma imagen o predicción temporal con respecto a las muestras de referencia en otras imágenes de referencia. Las imágenes pueden denominarse fotogramas, y las imágenes de referencia pueden denominarse fotogramas de referencia.

### SÍNTESIS

**[0005]** En general, esta divulgación describe técnicas para el almacenamiento y el suministro de un flujo de bits conforme a un estándar de codificación de video basado en un formato de archivo. Esta divulgación también describe ejemplos de un perfil de formato de medios para el estándar de codificación de video, así como el suministro mediante el uso de una técnica de *streaming*. Como un ejemplo, el estándar de codificación de video es el estándar de codificación de video esencial (EVC) y el formato de archivo es el formato de archivo de medios base de la Organización Internacional de Normalización. El formato de archivo de medios base de la Organización Internacional de Normalización se denomina ISOBMFF. Un ejemplo del perfil de formato de medios es el perfil de formato de aplicación de medios comunes (CMAF), y un ejemplo de la técnica de *streaming* es el *streaming* dinámico adaptativo sobre HTTP (DASH). Esta divulgación también puede describir los requisitos de cifrado y los parámetros de códecs que se utilizarán con el tipo de extensiones de correo de Internet multipropósito (MIME). Aunque se describen con respecto a EVC, ISOBMFF, CMAF y DASH, las técnicas de ejemplo no deben considerarse tan limitadas.

**[0006]** De acuerdo con un ejemplo de esta divulgación, un método para procesar un flujo de video o un archivo de video incluye recibir un registro de configuración para decodificar un flujo de bits de los datos de video, en donde el registro de configuración para el flujo de bits incluye un elemento de sintaxis de indicación de conjunto de herramientas que incluye información que identifica, a partir de un conjunto de herramientas de decodificación de video, herramientas que se requieren para decodificar el flujo de bits asociado con el registro de configuración; en función del elemento de sintaxis de indicación del conjunto de herramientas, determinar si recuperar el flujo de bits asociado con el registro de configuración; y en función de la determinación de recuperar el flujo de bits asociado con el registro de configuración, recuperar el flujo de bits y enviar el flujo de bits a un decodificador de video para su decodificación.

**[0007]** De acuerdo con otro ejemplo de esta divulgación, un dispositivo para procesar un flujo de video incluye una memoria configurada para almacenar el flujo de video y

uno o más procesadores implementados en la circuitería, acoplados a la memoria y configurados para recibir un registro de configuración para decodificar un flujo de bits de los datos de video, en donde el registro de configuración para el flujo de bits incluye un elemento de sintaxis de indicación de conjunto de herramientas que incluye información que identifica, a partir de un conjunto de herramientas de decodificación de video, las herramientas que se requieren para decodificar el flujo de bits asociado con el registro de configuración; en función del elemento de sintaxis de indicación del conjunto de herramientas, determinar si recuperar el flujo de bits asociado con el registro de configuración; y en función de la determinación de recuperar el flujo de bits asociado con el registro de configuración, recuperar el flujo de bits y enviar el flujo de bits a un decodificador de video para su decodificación.

**[0008]** De acuerdo con otro ejemplo de esta divulgación, un medio de almacenamiento legible por computadora almacena instrucciones que cuando se ejecutan mediante uno o más procesadores hacen que uno o más procesadores reciban un registro de configuración para decodificar un flujo de bits de los datos de video, en donde el registro de configuración para el flujo de bits incluye un elemento de sintaxis de indicación de conjunto de herramientas que incluye información que identifica, a partir de un conjunto de herramientas de decodificación de video, las herramientas que se requieren para decodificar el flujo de bits asociado con el registro de configuración; en función del elemento de sintaxis de indicación del conjunto de herramientas, determinar si recuperar el flujo de bits asociado con el registro de configuración; y en función de la determinación de recuperar el flujo de bits asociado con el registro de configuración, recuperar el flujo de bits y enviar el flujo de bits a un decodificador de video para su decodificación.

**[0009]** De acuerdo con otro ejemplo de esta divulgación, un aparato para procesar un flujo de video o un archivo de video incluye medios para recibir un registro de configuración para decodificar un flujo de bits de los datos de video, en donde el registro de configuración para el flujo de bits incluye un elemento de sintaxis de indicación de conjunto de herramientas que incluye información que identifica, a partir de un conjunto de herramientas de decodificación de video, herramientas que se requieren para decodificar el flujo de bits asociado con el registro de configuración; medios para determinar si recuperar el flujo de bits asociado con el registro de configuración en función del elemento de sintaxis de indicación de conjunto de herramientas; medios para recuperar el flujo de bits asociado con el registro de configuración; y medios para enviar el flujo de bits a un decodificador de video para su

decodificación.

**[0010]** Los detalles de uno o más ejemplos se exponen en las figuras adjuntas y en la descripción siguiente. Otras características, objetivos y ventajas serán evidentes a partir de la descripción, figuras y reivindicaciones.

### **BREVE DESCRIPCIÓN DE LOS DIBUJOS**

**[0011]** La Figura 1 es un diagrama de bloques que ilustra un ejemplo de sistema de codificación y decodificación de video que puede realizar técnicas de la presente divulgación.

**[0012]** Las Figuras 2A y 2B son diagramas conceptuales que ilustran un ejemplo de estructura de árbol binario de árbol cuaternario (QTBT) y una unidad de árbol de codificación (CTU) correspondiente.

**[0013]** La Figura 3 es un diagrama de bloques que ilustra un codificador de video de ejemplo que puede realizar las técnicas de la presente divulgación.

**[0014]** La Figura 4 es un diagrama de bloques que ilustra un decodificador de video de ejemplo que puede realizar las técnicas de la presente divulgación.

**[0015]** La Figura 5 es un diagrama conceptual que ilustra una estructura de ejemplo de un archivo, de acuerdo con una o más técnicas de esta divulgación.

**[0016]** La Figura 6 es un diagrama conceptual que ilustra elementos de contenido de medios de ejemplo.

**[0017]** La Figura 7 es un diagrama de bloques que ilustra elementos de un archivo de video de ejemplo.

**[0018]** La Figura 8 es un diagrama de flujo que ilustra un método de ejemplo para procesar datos de video de acuerdo con las técnicas de esta divulgación.

### **DESCRIPCIÓN DETALLADA**

**[0019]** El estándar de codificación de video esencial (EVC), desarrollado por ISO/IEC JTC 1/SC 29/WG 11 (MPEG), ofrece una solución de codificación de video eficiente y de baja complejidad. Los flujos elementales de EVC están estructurados como unidades de capa de abstracción de red (NAL). El almacenamiento de las unidades de NAL en el formato de archivo de medios base ISO (ISOBMFF) sigue principios similares a los de otros formatos de video estructurados de NAL (por ejemplo, estándares de codificación de video de alta eficiencia (HEVC) y/o codificación de video versátil (VVC)).

**[0020]** El almacenamiento de flujos elementales de EVC se puede subdividir en dos partes: información estática que se usa globalmente en el flujo elemental e información dinámica que puede variar por muestra. El conjunto de parámetros de secuencia

(SPS) y el conjunto de parámetros de imagen (PPS) pueden ser parte de la información que rara vez cambia y puede considerarse estática. Se puede usar un conjunto de indicadores para indicar si se espera que los conjuntos de parámetros cambien en el flujo. En tal caso, se define una agrupación de muestras que indica las muestras en las que cambian los conjuntos de parámetros.

**[0021]** El conjunto de parámetros de adaptación (APS) puede ser información dinámica que puede cambiar por muestra. El APS se utiliza para transportar la información del filtro de bucle adaptativo (ALF). La presencia de ALF se señala a través de un indicador y las muestras que llevan información de APS pueden pertenecer al mismo grupo de muestras.

**[0022]** Un cuadro puede referirse a una estructura de sintaxis elemental en el ISO BMFF, que incluye un tipo de cuadro codificado de cuatro caracteres, el recuento de bytes del cuadro y la carga útil. Un archivo ISO BMFF incluye una secuencia de cuadros, y los cuadros pueden contener otros cuadros. Un cuadro de película ("moov") contiene los metadatos de los flujos de medios continuos presentes en el archivo, cada uno representado en el archivo como una pista.

**[0023]** Los metadatos de una pista pueden agregarse en un cuadro de pista ("trak"), mientras que el contenido de medios de una pista puede agregarse en un cuadro de datos de medios ("mdat") o directamente en un archivo separado. El contenido de medios de las pistas incluye una secuencia de muestras, tales como unidades de acceso de audio o video. Una unidad de acceso es, generalmente, una unidad de datos que incluye datos de medios codificados (por ejemplo, imágenes) para una instancia de tiempo común. Una muestra es una unidad de acceso definida por una especificación particular, tal como especificaciones de codificación de video descritas en la presente. Una entrada de muestra puede proporcionar una descripción de una muestra correspondiente.

**[0024]** ISO BMFF especifica los siguientes tipos de pistas: una pista de medios, que contiene un flujo de medios elemental, una pista de sugerencias, que incluye instrucciones de transmisión de medios o representa un flujo de paquetes recibido, y una pista de metadatos temporizados, que comprende metadatos sincronizados en el tiempo.

**[0025]** Esta divulgación describe un registro de configuración para un flujo de bits de datos de video que incluye un elemento de sintaxis de indicación de conjunto de herramientas que incluye información que identifica las herramientas que se requieren para decodificar el flujo de bits asociado con el registro de configuración. Esta

divulgación también describe técnicas para incluir un parámetro de tipo de extensiones de correo de Internet multipropósito (MIME), por ejemplo, en una transmisión HTTP, que incluye un par de valores clave. La clave puede indicar que el tipo de MIME identifica las herramientas de decodificación de video y el valor puede identificar las herramientas que se requieren para decodificar el flujo de bits. Por lo general, se puede considerar que se requiere una herramienta si la herramienta está habilitada en al menos un conjunto de parámetros para el flujo de medios. Mediante el uso de dicho elemento de sintaxis de indicación de conjunto de herramientas y tipo de MIME, las técnicas de esta divulgación pueden permitir ventajosamente que un dispositivo de procesamiento de video determine si un decodificador de video del dispositivo de procesamiento de video puede decodificar un archivo sin tener que analizar la mayor parte del archivo.

**[0026]** La Figura 1 es un diagrama de bloques que ilustra un sistema de codificación y decodificación de video 100 de ejemplo que puede realizar las técnicas de la presente divulgación. Las técnicas de la presente divulgación se dirigen generalmente a encriptar (codificar y/o decodificar) datos de video. En general, los datos de video incluyen cualquier dato para procesar un video. Por lo tanto, los datos de video pueden incluir video sin procesar, no codificado, video codificado, video decodificado (por ejemplo, reconstruido) y metadatos de video, tales como datos de señalización.

**[0027]** Como se muestra en la Figura 1, el sistema 100 incluye un dispositivo fuente 102 que proporciona datos de video codificados que se desean decodificar y mostrar posteriormente mediante un dispositivo de destino 116, en este ejemplo. En particular, el dispositivo fuente 102 proporciona los datos de video al dispositivo de destino 116 mediante un medio legible por computadora 110. El dispositivo fuente 102 o dispositivo de destino 116 puede comprender cualquiera de un amplio rango de dispositivos, lo que incluye computadoras de escritorio, computadoras portátiles (es decir, laptop), computadoras tipo *tablet*, decodificadores, aparatos telefónicos, tales como teléfonos inteligentes, televisores, cámaras, dispositivos de visualización, reproductores de medios digitales, consolas de juegos de video, dispositivos de *streaming* de video, o similares. En algunos casos, el dispositivo fuente 102 y el dispositivo de destino 116 pueden estar equipados para comunicación inalámbrica, y por lo tanto pueden denominarse dispositivos de comunicación inalámbrica.

**[0028]** En el ejemplo de la Figura 1, el dispositivo fuente 102 incluye la fuente de video 104, la memoria 106, el codificador de video 200 y la interfaz de salida 108. El dispositivo de destino 116 incluye la interfaz de entrada 122, el decodificador de video

300, la memoria 120 y el dispositivo de visualización 118. De acuerdo con esta divulgación, el dispositivo fuente 102 y el dispositivo de destino 116 pueden configurarse para aplicar las técnicas de almacenamiento y suministro de flujos de bits conforme a un estándar de codificación de video basado en un formato de archivo (por ejemplo, flujo de bits elemental de codificación de video esencial (EVC) basado en el formato de archivo de medios base ISO (ISOBMFF)). La divulgación también describe un perfil de medios (por ejemplo, formato de aplicación de medios común (perfil de medios CMAF para EVC), describe el suministro mediante el uso de *streaming* adaptativo dinámico sobre HTTP (DASH), describe los requisitos de cifrado y describe los parámetros de códec que se usarán con el tipo de extensiones de correo de Internet multipropósito (MIME) que puede ser utilizado por el codificador de video 200 y el decodificador de video 300. Por lo tanto, el dispositivo fuente 102 representa un ejemplo de un dispositivo cliente configurado para llevar a cabo la codificación y transmisión de video, mientras que el dispositivo de destino 116 representa un ejemplo de un dispositivo de destino configurado para recibir y decodificar video codificado. En otros ejemplos, un dispositivo fuente y un dispositivo de destino pueden incluir otros componentes o disposiciones. Por ejemplo, el dispositivo fuente 102 puede recibir datos de video de una fuente de video externa, tal como una cámara externa. Asimismo, el dispositivo de destino 116 puede interactuar con un dispositivo de visualización externo, en lugar de incluir un dispositivo de visualización integrado.

**[0029]** El sistema 100 como se muestra en la Figura 1 es solo un ejemplo. En general, cualquier dispositivo de codificación y/o decodificación de video digital puede llevar a cabo técnicas de almacenamiento y suministro de flujos de bits conforme a un estándar de codificación de video basado en un formato de archivo (por ejemplo, flujo de bits elemental de EVC basado en ISOBMFF). La divulgación también describe un perfil de medios (por ejemplo, perfil de medios CMAF para EVC), describe el suministro mediante el uso de DASH, describe los requisitos de cifrado y describe los parámetros de códec que se usarán con el tipo de MIME que puede ser utilizado por el codificador de video 200 y el decodificador de video 300. El dispositivo fuente 102 y el dispositivo de destino 116 son solo ejemplos de los dispositivos de codificación, en donde el dispositivo fuente 102 genera datos de video codificados para la transmisión al dispositivo de destino 116. Esta divulgación se refiere a un dispositivo de "encriptación" como un dispositivo que realiza la encriptación (codificación y/o decodificación) de datos. Por lo tanto, el codificador de video 200 y el decodificador de video 300 representan ejemplos de dispositivos de codificación, en particular, un

codificador de video y un decodificador de video, respectivamente. En algunos ejemplos, el dispositivo fuente 102 y de destino 116 pueden operar de manera sustancialmente simétrica, de modo que cada uno del dispositivo fuente 102 y el dispositivo de destino de destino 116 incluya componentes de codificación y decodificación de video. Por lo tanto, el sistema 100 puede permitir la transmisión de video unidireccional o bidireccional entre el dispositivo fuente 102 y el dispositivo de destino 116, por ejemplo, para streaming de video, repetición de video, transmisión de video o videotelefonía.

**[0030]** En general, la fuente de video 104 representa una fuente de datos de video (es decir, datos de video sin procesar, no codificados) y proporciona una serie secuencial de imágenes (también denominadas "cuadros") de los datos de video al codificador de video 200, que codifica los datos para las imágenes. La fuente de video 104 del dispositivo fuente 102 puede incluir un dispositivo de captura de video, tal como una cámara de video, un archivo de video que contiene video sin procesar capturado anteriormente y/o una interfaz de transmisión de video para recibir videos de un proveedor de contenidos de video. Como alternativa adicional, la fuente de video 104 puede generar datos basados en gráficos informáticos como el video fuente, o una combinación de video en vivo, video archivado y video generado por computadora. En cada caso, el codificador de video 200 codifica los datos de video capturado, precapturado o generado por computadora. El codificador de video 200 puede reorganizar las imágenes del orden recibido (a veces denominado "orden de visualización") en un orden de codificación para la codificación. El codificador de video 200 puede generar un flujo de bits que incluye datos de video codificados. El dispositivo fuente 102 puede emitir entonces los datos de video codificados a través de la interfaz de salida 108 en un medio legible por computadora 110 para su recepción y/o recuperación mediante, por ejemplo, la interfaz de entrada 122 del dispositivo de destino 116.

**[0031]** La memoria 106 del dispositivo fuente 102 y la memoria 120 del dispositivo de destino 116 representan memorias de propósito general. En algunos ejemplos, las memorias 106, 120 pueden almacenar datos de video sin procesar, por ejemplo, video sin procesar de la fuente de video 104 y datos de video sin procesar, decodificados del decodificador de video 300. Adicional o alternativamente, las memorias 106, 120 pueden almacenar instrucciones de software ejecutables por, por ejemplo, el codificador de video 200 y el decodificador de video 300, respectivamente. Aunque la memoria 106 y la memoria 120 se muestran por separado del codificador de video 200

y el decodificador de video 300 en este ejemplo, debe entenderse que el codificador de video 200 y el decodificador de video 300 también pueden incluir memorias internas para propósitos funcionalmente similares o equivalentes. Además, las memorias 106, 120 pueden almacenar datos de video codificados, por ejemplo, la salida del codificador de video 200 y la entrada al decodificador de video 300. En algunos ejemplos, las porciones de memorias 106, 120 se pueden ubicar como uno o más búfer de video, por ejemplo, para almacenar datos de video sin procesar, decodificados o codificados.

**[0032]** El medio legible por computadora 110 puede representar cualquier tipo de medio o dispositivo capaz de transportar los datos de video codificados desde el dispositivo fuente 102 al dispositivo de destino 116. En un ejemplo, el medio legible por computadora 110 representa un medio de comunicación para permitir que el dispositivo fuente 102 transmita datos de video codificados directamente al dispositivo de destino 116 en tiempo real, por ejemplo, a través de una red de radiofrecuencia o una red basada en computadora. La interfaz de salida 108 puede modular una señal de transmisión que incluye los datos de video codificados, y la interfaz de entrada 122 puede demodular la señal de transmisión recibida, de acuerdo con un estándar de comunicación, tal como un protocolo de comunicación inalámbrica. El medio de comunicación puede comprender cualquier medio de comunicación inalámbrica o por cable, tales como un espectro de frecuencia de radio (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 del dispositivo fuente 102 al dispositivo de destino 116.

**[0033]** En algunos ejemplos, el dispositivo fuente 102 puede enviar datos codificados desde la interfaz de salida 108 al dispositivo de almacenamiento 112. De manera similar, el dispositivo de destino 116 puede acceder a datos codificados desde el dispositivo de almacenamiento 112 a través de la interfaz de entrada 122. El dispositivo de almacenamiento 112 puede incluir cualquiera de una variedad de medios de almacenamiento de datos distribuidos o a los que se accede localmente como un disco duro, discos Blu-ray, DVD, CD-ROM, memoria flash, memoria volátil o no volátil, o cualquier otro medio de almacenamiento digital adecuado para el almacenamiento de datos de video codificado.

**[0034]** En algunos ejemplos, el dispositivo fuente 102 puede enviar datos de video

codificados al servidor de archivos 114 u otro dispositivo de almacenamiento intermedio que puede almacenar el video codificado generado por el dispositivo fuente 102. El dispositivo de destino 116 puede acceder a los datos de video almacenados desde el servidor de archivos 114 mediante *streaming* o descarga. El servidor de archivos 114 puede ser cualquier tipo de servidor capaz de almacenar datos de video y de transmitir los datos de video codificados al dispositivo de destino 116. El servidor de archivos 114 puede representar un servidor web (por ejemplo, para un sitio web), un servidor de protocolo de transferencia de archivos (FTP), un dispositivo de red de entrega de contenido o un dispositivo de almacenamiento conectado a la red (NAS). El dispositivo de destino 116 puede acceder a los datos de video codificados del servidor de archivos 114 a través de cualquier conexión de datos estándar, incluso una conexión a Internet. Esto puede incluir un canal inalámbrico (por ejemplo, una conexión Wi-Fi), una conexión por cable (por ejemplo, línea de abonado digital (DSL), cable módem, etc.), o una combinación de ambos que es adecuada para acceder a los datos de video codificados y almacenados en el servidor de archivos 114. El servidor de archivos 114 y la interfaz de entrada 122 pueden configurarse para operar de acuerdo con un protocolo de transmisión de *streaming*, un protocolo de transmisión de descarga o una combinación de estos.

**[0035]** La interfaz de salida 108 y la interfaz de entrada 122 pueden representar transmisores/receptores inalámbricos, módems, componentes de red cableados (por ejemplo, tarjetas Ethernet), componentes de comunicación inalámbrica que operan de acuerdo con cualquiera de una variedad de estándares IEEE 802.11 u otros componentes físicos. En los ejemplos en los que la interfaz de salida 108 y la interfaz de entrada 122 comprenden componentes inalámbricos, la interfaz de salida 108 y la interfaz de entrada 122 pueden configurarse para transferir datos, tales como datos de video codificados, de acuerdo con un estándar de comunicación celular, tal como 4G, 4G-LTE (evolución a largo plazo), LTE avanzada, 5G o similares. En algunos ejemplos en los que la interfaz de salida 108 comprende un transmisor inalámbrico, la interfaz de salida 108 y la interfaz de entrada 122 pueden configurarse para transferir datos, tales como datos de video codificados, de acuerdo con otros estándares inalámbricos, tales como una especificación IEEE 802.11, una especificación IEEE 802.15 (por ejemplo, ZigBee™), un estándar Bluetooth™ o similares. En algunos ejemplos, el dispositivo fuente 102 y/o el dispositivo de destino 116 pueden incluir respectivos dispositivos de sistema en un chip (SoC). Por ejemplo, el dispositivo fuente 102 puede incluir un dispositivo de SoC para realizar la funcionalidad atribuida al codificador de

video 200 y/o la interfaz de salida 108, y el dispositivo de destino 116 puede incluir un dispositivo de SoC para realizar la funcionalidad atribuida al decodificador de video 300 y/o la interfaz de entrada 122.

**[0036]** Las técnicas de la presente divulgación se pueden aplicar a la codificación de video para diversas aplicaciones multimedia, tales como transmisiones por televisión inalámbrica, transmisiones por televisión por cable, transmisiones por televisión satelital, transmisiones de video por *streaming* en Internet, tales como *streaming* adaptativo dinámico en HTTP (DASH), video digital que se codifica en un medio de almacenamiento de datos, decodificación de video digital almacenado en un medio de almacenamiento de datos, u otras aplicaciones.

**[0037]** La interfaz de entrada 122 del dispositivo de destino 116 recibe un flujo de bits de video codificado desde un medio legible por computadora 110 (por ejemplo, un medio de comunicación, el dispositivo de almacenamiento 112, el servidor de archivos 114 o similares). El flujo de bits de video codificado puede incluir información de señalización definida por el codificador de video 200, que también es utilizada por el decodificador de video 300, tal como elementos de sintaxis que tienen valores que describen características y/o procesamiento de bloques de video u otras unidades codificadas (por ejemplo, porciones, imágenes, grupos de imágenes, secuencias o similares). El dispositivo de visualización 118 muestra imágenes decodificadas de los datos de video decodificados a un usuario. El dispositivo de visualización 118 puede representar cualquiera de los diversos dispositivos de visualización, tales como un tubo de rayos catódicos (CRT), una pantalla de cristal líquido (LCD), una pantalla de plasma, una pantalla de diodo orgánico de emisión de luz (OLED), u otro tipo de dispositivo de visualización.

**[0038]** Aunque no se muestra en la Figura 1, en algunos ejemplos, el codificador de video 200 y el decodificador de video 300 pueden cada uno estar integrados con un codificador de audio y/o un decodificador de audio, y pueden incluir unidades MUX-DEMUX apropiadas, u otro *hardware* y/o *software*, para manejar flujos múltiples, incluso la codificación de audio y video en un flujo común de datos. Si correspondiera, las unidades MUX-DEMUX pueden adecuarse al protocolo de multiplexor ITU H.223, o a otros protocolos, tales como el protocolo de datagrama de usuario (UDP).

**[0039]** El codificador de video 200 y el decodificador de video 300 pueden implementarse como cualquiera de una variedad de circuitería codificadora y/o decodificadora adecuada, tal como uno o más microprocesadores, procesadores de señal digital (DSP), circuitos integrados de aplicación específica (ASIC), matrices de

puertas programables en campo (FPGA), circuitos discretos lógicos, *software*, *hardware*, *firmware* o cualquier combinación de estos. Cuando las técnicas se implementan parcialmente en *software*, un dispositivo puede almacenar las instrucciones para el *software* en un medio legible por computadora no transitorio adecuado y ejecutar las instrucciones en *hardware* mediante el uso de uno o más procesadores para realizar las técnicas de la presente divulgación. Cada uno del codificador de video 200 y el decodificador de video 300 pueden incluirse en uno o más codificadores o decodificadores, cualquiera de los cuales pueden estar integrados como parte de un codificador/decodificador combinado (CODEC) en un dispositivo respectivo. Un dispositivo que incluye un codificador de video 200 y/o un decodificador de video 300 puede comprender un circuito integrado, un microprocesador y/o un dispositivo de comunicación inalámbrica, tal como un teléfono celular.

**[0040]** El codificador de video 200 y el decodificador de video 300 pueden operar de acuerdo con otros estándares de propiedad o de la industria, como el modelo de prueba de exploración conjunta (JEM) o ITU-T H.266, también denominado codificación de video versátil (VVC). Un borrador reciente del estándar VVC se describe en Bross, *et al.* "Versatile Video Coding (Draft 7)", Joint Video Experts Team (JVET) de ITU-T SG 16 WP 3 e ISO/IEC JTC 1/SC 29/WG 11, 16.<sup>a</sup> reunión: Ginebra, CH, 1–11 de octubre de 2019, JVET-P2001-v14 (en adelante "VVC Draft 7"). El codificador de video 200 y el decodificador de video 300 pueden operar de acuerdo con el estándar EVC desarrollado por ISO/IEC JTC 1/SC 29/WG 11 (MPEG). Las técnicas de esta divulgación, sin embargo, no se limitan a ningún estándar de codificación particular.

**[0041]** En general, el codificador de video 200 y el decodificador de video 300 pueden realizar una codificación de imágenes basada en bloques. El término "bloque" generalmente se refiere a una estructura que incluye datos para ser procesados (por ejemplo, codificados, decodificados o usados de otro modo en el proceso de codificación y/o decodificación). Por ejemplo, un bloque puede incluir una matriz bidimensional de muestras de datos de luminancia y/o crominancia. En general, el codificador de video 200 y el decodificador de video 300 pueden codificar datos de video representados en un formato YUV (por ejemplo, Y, Cb, Cr). Es decir, en lugar de codificar datos de rojo, verde y azul (RGB) para muestras de una imagen, el codificador de video 200 y el decodificador de video 300 pueden codificar componentes de luminancia y crominancia, en donde los componentes de crominancia pueden incluir componentes de crominancia de tono rojo y azul. En algunos ejemplos,

el codificador de video 200 convierte los datos con formato RGB recibidos en una representación YUV antes de la codificación, y el decodificador de video 300 convierte la representación YUV en el formato RGB. Alternativamente, las unidades de procesamiento previo y posterior (no mostradas) pueden realizar estas conversiones.

**[0042]** Esta divulgación puede referirse generalmente a la codificación (por ejemplo, codificación y decodificación) de imágenes para incluir el proceso de codificación o decodificación de datos de la imagen. De manera similar, esta divulgación puede referirse a la codificación de bloques de una imagen para incluir el proceso de codificación o decodificación de datos para los bloques, por ejemplo, predicción y/o codificación residual. Un flujo de bits de video codificado generalmente incluye una serie de valores para elementos de sintaxis representativos de decisiones de codificación (por ejemplo, modos de codificación) y partición de imágenes en bloques. Por tanto, las referencias a la codificación de una imagen o un bloque deben entenderse generalmente como valores de codificación para los elementos de sintaxis que forman la imagen o el bloque.

**[0043]** Diversos estándares de codificación de video definen diversos bloques, a veces denominados unidades de codificación (CU), unidades de predicción (PU) y unidades de transformada (TU). Un codificador de video (tal como el codificador de video 200) divide una unidad de árbol de codificación (CTU) en CU de acuerdo con una estructura de árbol cuaternario. Es decir, el codificador de video divide las CTU y CU en cuatro cuadrados iguales que no se superponen, y cada nodo del árbol cuaternario tiene cero o cuatro nodos secundarios. Los nodos sin nodos secundarios pueden denominarse "nodos hoja", y las CU de dichos nodos hoja pueden incluir una o más PU y/o una o más TU. El codificador de video puede dividir además las PU y las TU. Por ejemplo, un árbol cuaternario residual (RQT) representa la partición de las TU, y las PU representan los datos de interpredicción, mientras que las TU representan los datos residuales. Las CU que son intrapredichas incluyen información de intrapredicción, tal como una indicación intramodo.

**[0044]** Como otro ejemplo, el codificador de video 200 y el decodificador de video 300 pueden configurarse para operar de acuerdo con JEM, VVC, EVC o cualquier otro estándar. Un codificador de video (tal como el codificador de video 200) divide una imagen en una pluralidad de unidades de árbol de codificación (CTU). El codificador de video 200 puede dividir una CTU de acuerdo con una estructura de árbol, tal como una estructura de árbol binario de árbol cuaternario (QTBT) o una estructura de árbol de tipos múltiples (MTT). La estructura QTBT elimina los conceptos de múltiples tipos

de partición, tales como la separación entre CU, PU y TU. Una estructura QTBT incluye dos niveles: un primer nivel dividido de acuerdo con la partición de árbol cuaternario y un segundo nivel dividido de acuerdo con la partición de árbol binario. Un nodo raíz de la estructura QTBT corresponde a una CTU. Los nodos hoja de los árboles binarios corresponden a unidades de codificación (CU).

**[0045]** En una estructura de partición MTT, los bloques se pueden particionar mediante el uso de una partición de árbol cuaternario (QT), una partición de árbol binario (BT) y uno o más tipos de particiones de árbol triple (TT) (también denominado árbol ternario). Una partición de árbol triple o ternario es una partición en la que un bloque se divide en tres subbloques. En algunos ejemplos, una partición de árbol triple o ternario divide un bloque en tres subbloques sin dividir el bloque original por el centro. Los tipos de partición en MTT (por ejemplo, QT, BT y TT) pueden ser simétricos o asimétricos.

**[0046]** En algunos ejemplos, el codificador de video 200 y el decodificador de video 300 pueden usar una sola estructura QTBT o MTT para representar cada uno de los componentes de luminancia y crominancia, mientras que en otros ejemplos, el codificador de video 200 y el decodificador de video 300 pueden usar dos o más estructuras QTBT o MTT, como una estructura QTBT/MTT para el componente de luminancia y otra estructura QTBT/MTT para ambos componentes de crominancia (o dos estructuras QTBT/MTT para los respectivos componentes de crominancia).

**[0047]** El codificador de video 200 y el decodificador de video 300 pueden configurarse para usar partición de árbol cuaternario, partición QTBT, partición MTT u otras estructuras de partición. Para fines de explicación, la descripción de las técnicas de esta divulgación se presenta con respecto a la partición QTBT. Sin embargo, debe entenderse que las técnicas de esta divulgación también se pueden aplicar a encodificadores de video configurados para usar partición de árbol cuaternario, así como también otros tipos de partición.

**[0048]** Los bloques (por ejemplo, CTU o CU) pueden agruparse en diversas formas en una imagen. A modo de ejemplo, un ladrillo puede referirse a una región rectangular de filas de CTU dentro de un mosaico particular en una imagen. Un mosaico puede ser una región rectangular de CTB dentro de una columna de mosaicos en particular y una fila de mosaicos en particular en una imagen. Una columna de mosaico se refiere a una región rectangular de CTU que tiene una altura igual a la altura de la imagen y un ancho determinado por los elementos de sintaxis (por ejemplo, tal como en un conjunto de parámetros de imagen). Una fila de mosaico se refiere a una región

rectangular de CTU que tiene una altura determinada por los elementos de sintaxis (por ejemplo, tal como en un conjunto de parámetros de imagen) y un ancho igual al ancho de la imagen.

**[0049]** En algunos ejemplos, un mosaico puede particionarse en múltiples ladrillos, cada uno de los cuales puede incluir una o más filas de CTU dentro del mosaico. Un mosaico que no está particionado en múltiples ladrillos también se puede denominar "ladrillo". Sin embargo, un ladrillo que es un subconjunto real de un mosaico no se puede denominar mosaico.

**[0050]** Los ladrillos en una imagen se pueden disponer en una porción. Una porción puede ser un número entero de ladrillos de una imagen que pueden contenerse exclusivamente en una única unidad de capa de abstracción de red (NAL). En algunos ejemplos, una porción incluye una cantidad de mosaicos completos o solo una secuencia consecutiva de ladrillos completos de un mosaico.

**[0051]** Esta divulgación puede usar "NxN" y "N por N" indistintamente para referirse a las dimensiones de muestra de un bloque (tal como una CU u otro bloque de video) en términos de dimensiones verticales y horizontales, por ejemplo, muestras de 16x16 o muestras de 16 por 16. En general, una CU de 16x16 tendrá 16 muestras en dirección vertical ( $y = 16$ ) y 16 muestras en dirección horizontal ( $x = 16$ ). Asimismo, una CU NxN generalmente tiene N muestras en una dirección vertical y N muestras en una dirección horizontal, donde N representa un valor entero no negativo. Las muestras en una CU pueden disponerse en filas y columnas. Además, las CU no necesitan tener necesariamente la misma cantidad de muestras en la dirección horizontal que en la dirección vertical. Por ejemplo, las CU pueden comprender muestras NxM, en donde M no es necesariamente igual a N.

**[0052]** El codificador de video 200 codifica datos de video para CU que representan información de predicción y/o residual y otra información. La información de predicción indica cómo se va a predecir la CU para formar un bloque de predicción para la CU. La información residual generalmente representa diferencias muestra por muestra entre muestras de la CU antes de la codificación y el bloque de predicción.

**[0053]** Para predecir una CU, el codificador de video 200 puede formar generalmente un bloque de predicción para la CU mediante interpredicción o intrapredicción. La interpredicción generalmente se refiere a predecir la CU a partir de datos de una imagen previamente codificada, mientras que la intrapredicción generalmente se refiere a predecir la CU a partir de datos previamente codificados de la misma imagen. Para realizar la interpredicción, el codificador de video 200 puede generar el bloque de

predicción mediante el uso de uno o más vectores de movimiento. El codificador de video 200 generalmente puede realizar una búsqueda de movimiento para identificar un bloque de referencia que se asemeje mucho a la CU, por ejemplo, en términos de diferencias entre la CU y el bloque de referencia. El codificador de video 200 puede calcular una métrica de diferencia mediante el uso de una suma de diferencia absoluta (SAD), suma de diferencias cuadradas (SSD), diferencia absoluta media (MAD), diferencias cuadráticas medias (MSD) u otros cálculos de diferencias similares para determinar si una referencia de bloque se asemeja mucho a la CU actual. En algunos ejemplos, el codificador de video 200 puede predecir la CU actual mediante el uso de predicción unidireccional o predicción bidireccional.

**[0054]** Algunos estándares de codificación de video también proporcionan un modo de compensación de movimiento afín, que puede considerarse un modo de interpredicción. En el modo de compensación de movimiento afín, el codificador de video 200 puede determinar dos o más vectores de movimiento que representan un movimiento no traslacional, tal como acercar o alejar, rotación, movimiento en perspectiva u otros tipos de movimiento irregular.

**[0055]** Para realizar la intrapredicción, el codificador de video 200 puede seleccionar un modo de intrapredicción para generar el bloque de predicción. Por ejemplo, el codificador de video 200 puede utilizar sesenta y siete o algún otro número de modos de intrapredicción, que incluyen diversos modos direccionales, así como el modo planar y el modo DC. En general, el codificador de video 200 selecciona un modo de intrapredicción que describe muestras cercanas a un bloque actual (por ejemplo, un bloque de una CU) desde el cual predecir muestras del bloque actual. Por lo general, dichas muestras pueden estar arriba, arriba y a la izquierda, o a la izquierda del bloque actual en la misma imagen que el bloque actual, asumiendo que el codificador de video 200 codifica CTU y CU en orden de exploración de trama (de izquierda a derecha, de arriba a abajo).

**[0056]** El codificador de video 200 codifica datos que representan el modo de predicción para un bloque actual. Por ejemplo, para los modos de interpredicción, el codificador de video 200 puede codificar datos que representen cuál de los diversos modos de interpredicción disponibles se usa, así como información de movimiento para el modo correspondiente. Para la interpredicción unidireccional o bidireccional, por ejemplo, el codificador de video 200 puede codificar vectores de movimiento mediante el uso de la predicción avanzada de vectores de movimiento (AMVP) o el modo de fusión. El codificador de video 200 puede usar modos similares para codificar

vectores de movimiento para el modo de compensación de movimiento afín.

**[0057]** Después de la predicción, tal como intrapredicción o interpredicción de un bloque, el codificador de video 200 puede calcular datos residuales para el bloque. Los datos residuales, tales como un bloque residual, representan las diferencias muestra por muestra entre el bloque y un bloque de predicción para el bloque, formado mediante el uso del modo de predicción correspondiente. El codificador de video 200 puede aplicar una o más transformadas al bloque residual, para producir datos transformados en un dominio de transformada en lugar del dominio de muestra. Por ejemplo, el codificador de video 200 puede aplicar una transformada de coseno discreta (DCT), una transformada de número entero, una transformada de ondículas o una transformada conceptualmente similar a los datos de video residuales. Además, el codificador de video 200 puede aplicar una transformada secundaria después de la primera transformada, tal como una transformada secundaria no separable dependiente del modo (MDNSST), una transformada dependiente de la señal, una transformada de Karhunen-Loeve (KLT) o similares. El codificador de video 200 produce coeficientes de transformada tras la aplicación de la una o más transformadas.

**[0058]** Como se indicó anteriormente, tras cualquier transformada para producir coeficientes de transformada, el codificador de video 200 puede llevar a cabo la cuantificación de los coeficientes de transformada. Cuantificación se refiere generalmente a un proceso donde los coeficientes de transformada se cuantifican para reducir posiblemente la cantidad de datos utilizados para representar los coeficientes de transformada, lo que proporciona una compresión adicional. Al realizar el proceso de cuantificación, el codificador de video 200 puede reducir la profundidad de bits asociada con algunos o todos los coeficientes de transformada. Por ejemplo, el codificador de video 200 puede redondear un valor de  $n$  bits a un valor de  $m$  bits durante la cuantificación, donde  $n$  es mayor que  $m$ . En algunos ejemplos, para realizar la cuantificación, el codificador de video 200 puede realizar un desplazamiento a la derecha bit a bit del valor que se va a cuantificar.

**[0059]** Después de la cuantificación, el codificador de video 200 puede explorar los coeficientes de transformada, lo que produce un vector unidimensional a partir de la matriz bidimensional que incluye los coeficientes de transformada cuantificados. La exploración puede diseñarse para colocar coeficientes de transformada de energía más alta (y por lo tanto de frecuencia más baja) en el frente del vector y para colocar coeficientes de transformada de energía más baja (y por lo tanto de frecuencia más

alta) en la parte posterior del vector. En algunos ejemplos, el codificador de video 200 puede utilizar un orden de exploración predefinido para explorar los coeficientes de transformada cuantificados para producir un vector serializado, y luego codificar por entropía los coeficientes de transformada cuantificados del vector. En otros ejemplos, el codificador de video 200 puede realizar una exploración adaptativa. Después de explorar los coeficientes de transformada cuantificados para formar el vector unidimensional, el codificador de video 200 puede codificar por entropía el vector unidimensional, por ejemplo, de acuerdo con la codificación aritmética binaria adaptativa al contexto (CABAC). El codificador de video 200 también puede codificar por entropía valores para elementos de sintaxis que describen metadatos asociados con los datos de video codificados para su uso por el decodificador de video 300 en la decodificación de los datos de video.

**[0060]** Para llevar a cabo CABAC, el codificador de video 200 puede asignar un contexto dentro de un modelo de contexto a un símbolo que se va a transmitir. El contexto puede referirse, por ejemplo, a si los valores cercanos del símbolo tienen valor cero o no. La determinación de la probabilidad puede ser en función de un contexto asignado al símbolo.

**[0061]** El codificador de video 200 puede generar además datos de sintaxis, tales como datos de sintaxis basados en bloques, datos de sintaxis basados en imágenes y datos de sintaxis basados en secuencias, al decodificador de video 300, por ejemplo, en un encabezado de imagen, un encabezado de bloque, un encabezado de porción, u otros datos de sintaxis, tales como un conjunto de parámetros de secuencia (SPS), un conjunto de parámetros de imagen (PPS) o un conjunto de parámetros de video (VPS). El decodificador de video 300 puede decodificar igualmente tales datos de sintaxis para determinar cómo decodificar los datos de video correspondientes.

**[0062]** De esta manera, el codificador de video 200 puede generar un flujo de bits que incluye datos de video codificados, por ejemplo, elementos de sintaxis que describen la partición de una imagen en bloques (por ejemplo, CU) e información de predicción y/o residual para los bloques. Finalmente, el decodificador de video 300 puede recibir el flujo de bits y decodificar los datos de video codificados.

**[0063]** En general, el decodificador de video 300 lleva a cabo un proceso recíproco al llevado a cabo por el codificador de video 200 para decodificar los datos de video codificados del flujo de bits. Por ejemplo, el decodificador de video 300 puede decodificar valores para elementos de sintaxis del flujo de bits mediante el uso de CABAC de una manera sustancialmente similar, aunque recíproca al proceso de

codificación CABAC del codificador de video 200. Los elementos de sintaxis pueden definir la información de partición de una imagen en CTU, y la partición de cada CTU de acuerdo con una estructura de partición correspondiente, tal como una estructura QTBT, para definir las CU de la CTU. Los elementos de sintaxis pueden definir además la información de predicción y residual para los bloques (por ejemplo, CU) de datos de video.

**[0064]** La información residual puede representarse, por ejemplo, mediante coeficientes de transformada cuantificados. El decodificador de video 300 puede cuantificar de forma inversa y transformar de forma inversa los coeficientes de transformada cuantificados de un bloque para reproducir un bloque residual para el bloque. El decodificador de video 300 utiliza un modo de predicción señalado (intrapredicción o interpredicción) e información de predicción relacionada (por ejemplo, información de movimiento para la interpredicción) para formar un bloque de predicción para el bloque. El decodificador de video 300 puede combinar entonces el bloque de predicción y el bloque residual (muestra por muestra) para reproducir el bloque original. El decodificador de video 300 puede llevar a cabo un procesamiento adicional, tal como llevar a cabo un proceso de desbloqueo para reducir los artefactos visuales a lo largo de los límites del bloque.

**[0065]** Esta divulgación en general puede referirse a la "señalización" de cierta información, tal como elementos de sintaxis. El término "señalización" puede referirse generalmente a la comunicación de valores para elementos de sintaxis y/u otros datos usados para decodificar datos de video codificados. Es decir, el codificador de video 200 puede señalar valores para elementos de sintaxis en el flujo de bits. En general, la señalización se refiere a generar un valor en el flujo de bits. Como se señaló anteriormente, el dispositivo fuente 102 puede transportar el flujo de bits al dispositivo de destino 116 sustancialmente en tiempo real, o no en tiempo real, como podría ocurrir cuando se almacenan elementos de sintaxis en el dispositivo de almacenamiento 112 para su posterior recuperación por el dispositivo de destino 116.

**[0066]** A continuación se describen técnicas de ejemplo de EVC. En algunos ejemplos, las técnicas de EVC pueden ser similares a las de VVC y/o HEVC (por ejemplo, técnicas de codificación basadas en bloques).

**[0067]** Una muestra de EVC contiene una unidad de acceso como se define en la cláusula 3.1 de ISO/IEC 23094-1: Tecnología de la información - Codificación general de video - Parte 1: Codificación de video esencial. En EVC, puede haber un orden canónico para que un flujo de bits se ajuste a EVC y puede haber ciertas restricciones.

El formato de flujo canónico es un flujo elemental de EVC que cumple las siguientes condiciones además de las condiciones generales de ISO/IEC 14496-15: Tecnología de la información – Codificación de objetos audiovisuales – Parte 15: Formato de archivo de codificación de video avanzada (AVC), sección 4.3.2:

- a. **Unidades de NAL del delimitador de unidad de acceso:** Las restricciones que obedecen las unidades de NAL del delimitador de unidad de acceso se definen en ISO/IEC 23094-1.
- b. **SPS y PPS:** Un SPS o PPS que se utilizará en una imagen puede enviarse (y en algunos ejemplos debe enviarse) antes de la muestra que contiene esa imagen o en la muestra para esa imagen. Al menos el SPS y el PPS con id igual a 0 pueden almacenarse en la entrada de muestra de la pista que contiene el flujo elemental de EVC.
- c. **APS:** Un APS que se utilizará en un segmento puede (y en algunos ejemplos debe) enviarse antes de la unidad de NAL de VCL que contiene esa porción. Los APS pueden almacenarse en la entrada de muestra y/o en las muestras.
- d. **Mensajes SEI:** Los mensajes SEI de naturaleza declarativa pueden almacenarse en la entrada de muestra; no existe ninguna prescripción sobre la eliminación de tales mensajes SEI de las muestras.
- e. **Datos de relleno.** Los datos de video se representan naturalmente como una tasa de bits variable en el formato de archivo y deben completarse para la transmisión si es necesario.

**[0068]** La eliminación o adición de unidades de NAL de datos de relleno, códigos de inicio, mensajes SEI o mensajes SEI de datos de relleno pueden cambiar las características del flujo de bits con respecto a la conformidad con el decodificador de referencia hipotético (HRD) cuando se opera el HRD en modo de tasa de bits constante (CBR) como se especifica en ISO/IEC 23094-1, Anexo C.

**[0069]** El siguiente es el registro y la descripción de la configuración del decodificador EVC. El decodificador de video 300 puede configurarse de acuerdo con la configuración del decodificador EVC, en algunos ejemplos. Como parte de la descripción de la configuración del decodificador EVC, el siguiente especifica la información de configuración del decodificador para el contenido de video ISO/IEC 23094-1.

**[0070]** Este registro contiene un campo de versión. Los analizadores de formato de archivo (por ejemplo, lectores) pueden (por ejemplo, deben) no intentar decodificar

este registro o los flujos a los que se aplica si no se reconoce el número de versión.

**[0071]** Las extensiones compatibles con este registro pueden ampliarlo y no cambiar el código de versión de configuración. Los lectores pueden ignorar los datos no reconocidos más allá de la definición de los datos para los que los lectores están configurados para analizar.

**[0072]** Los valores para `profile_idc`, `level_idc`, `toolset_idc`, `chroma_format_idc`, `pic_width_in_luma_samples`, `pic_height_in_luma_samples`, `bit_depth_luma_minus8`, y `bit_depth_chroma_minus8` pueden ser (y en algunos ejemplos deben ser) válidos para todos los conjuntos de parámetros que se activan cuando se decodifica el flujo (denominados "todos los conjuntos de parámetros"). En concreto, pueden aplicarse las siguientes restricciones:

- a. La indicación de `profile_idc` puede (por ejemplo, debe) indicar un perfil al que se ajusta el flujo asociado con este registro de configuración. Si los SPS están marcados con diferentes perfiles, entonces el flujo podría necesitar un examen para determinar con qué perfil, si lo hubiere, se ajusta todo el flujo. Si no se examina todo el flujo, o si el examen revela que no hay un perfil al que se ajuste todo el flujo, entonces todo el flujo puede dividirse (por ejemplo, debe dividirse) en dos o más subflujos con registros de configuración separados en los que estas reglas se pueden cumplir.
- b. La indicación de nivel `level_idc` puede (por ejemplo, debe) indicar un nivel de capacidad igual o mayor que el nivel más alto indicado en todos los conjuntos de parámetros de este registro de configuración.
- c. Tanto `pic_width_in_luma_samples` como `pic_height_in_luma_samples` pueden (por ejemplo, deben) contener los valores más altos de todos los conjuntos de parámetros de este registro de configuración.
- d. La indicación del conjunto de herramientas `toolset_idc` puede (por ejemplo, debe) señalar todas las herramientas que se requieren para decodificar el flujo asociado con este registro de configuración. Los indicadores de las herramientas pueden (por ejemplo, deben) cumplir con los requisitos de conformidad que se proporcionan en la Tabla A.6 (reproducida a continuación) de ISO/IEC 23094-1 y pueden (por ejemplo, deben) ser idénticos al campo `toolset_idc` que se indica en el SPS.
- e. El valor de `chroma_format_idc` en todos los conjuntos de parámetros

puede (por ejemplo, debe) ser idéntico.

- f. El valor de `bit_depth_luma_minus8` en todos los conjuntos de parámetros puede (por ejemplo, debe) ser idéntico.
- g. El valor de `bit_depth_chroma_minus8` en todos los conjuntos de parámetros puede (por ejemplo, debe) ser idéntico.

**[0073]** Se proporciona una indicación explícita en el registro de configuración del decodificador de EVC sobre el formato de croma y la profundidad de bits utilizados por el flujo elemental de video de EVC. Cada tipo de dicha información puede (por ejemplo, debe) ser idéntico en todos los conjuntos de parámetros, si están presentes, en un único registro de configuración de EVC. Si dos secuencias difieren en cualquier tipo de información, pueden (por ejemplo, deben) usarse dos entradas de muestra de EVC diferentes.

**[0074]** Hay un conjunto de matrices para transportar unidades de NAL de inicialización. Los tipos de unidades de NAL pueden estar restringidos para indicar unidades SPS, PPS, APS y SEI NAL únicamente. Se pueden definir los tipos de unidades de NAL que están reservados en ISO/IEC 23094-1 y en esta especificación, y se pueden ignorar las unidades de NAL con tipos de unidades de NAL no reconocidos. Este comportamiento "tolerante" (por ejemplo, ignorar los tipos de unidades de NAL no reconocidas) puede diseñarse para que no se generen errores, lo que permite la posibilidad de extensiones compatibles con versiones anteriores de estas matrices en especificaciones futuras.

**[0075]** El campo de longitud se puede utilizar en cada muestra para indicar la longitud de las unidades de NAL que contiene, así como los conjuntos de parámetros, si están almacenados en la entrada de la muestra. En algunos ejemplos, las matrices están en el orden de SPS, PPS, APS, SEI.

**[0076]** La Tabla A.6 de ISO/IEC 23094-1 es la siguiente:

| <b>tools_idcbinIdx</b> | <b>Indicador de herramienta SPS</b> | <b>Requisito de conformidad</b>                                                                                                               |
|------------------------|-------------------------------------|-----------------------------------------------------------------------------------------------------------------------------------------------|
| 0                      | <code>sps_btt_flag</code>           | <code>sps_btt_flag &lt;= toolset_idc_h &amp; 0x1 &amp;&amp; sps_btt_flag &gt;= toolset_idc_l &amp; 0x1</code>                                 |
| 1                      | <code>sps_suco_flag</code>          | <code>sps_suco_flag &lt;= ( toolset_idc_h &amp; 0x2 ) &gt;&gt; 1 &amp;&amp; sps_suco_flag &gt;= ( toolset_idc_l &amp; 0x2 ) &gt;&gt; 1</code> |

|    |                 |                                                                                                              |
|----|-----------------|--------------------------------------------------------------------------------------------------------------|
| 2  | sps_amvr_flag   | sps_amvr_flag <= ( toolset_idc_h & 0x4 )<br>>> 2 && sps_amvr_flag >= ( toolset_idc_l & 0x4 ) >> 2            |
| 3  | sps_mmvd_flag   | sps_mmvd_flag <= ( toolset_idc_h & 0x8 )<br>>> 3 && sps_mmvd_flag >= ( toolset_idc_l & 0x8 ) >> 3            |
| 4  | sps_affine_flag | sps_affine_flag <= ( toolset_idc_h & 0x10 )<br>>> 4 && sps_affine_flag >= ( toolset_idc_l & 0x10 ) >> 4      |
| 5  | sps_dmvr_flag   | sps_dmvr_flag <= ( toolset_idc_h & 0x20 )<br>>> 5 && sps_dmvr_flag >= ( toolset_idc_l & 0x20 ) >> 5          |
| 6  | sps_alf_flag    | sps_alf_flag <= ( toolset_idc_h & 0x40 ) >><br>6 && sps_alf_flag >= ( toolset_idc_l & 0x40 )<br>>> 6         |
| 7  | sps_admvp_flag  | sps_admvp_flag <= ( toolset_idc_h & 0x80<br>) >> 7 && sps_admvp_flag >= ( toolset_idc_l<br>& 0x80 ) >> 7     |
| 8  | sps_eipd_flag   | sps_eipd_flag <= ( toolset_idc_h & 0x100 )<br>>> 8 && sps_eipd_flag >= ( toolset_idc_l &<br>0x100 ) >> 8     |
| 9  | sps_adcc_flag   | sps_adcc_flag <= ( toolset_idc_h & 0x200 )<br>>> 9 && sps_adcc_flag >= ( toolset_idc_l &<br>0x200 ) >> 9     |
| 10 | sps_ibc_flag    | sps_ibc_flag <= ( toolset_idc_h & 0x400 )<br>>> 10 && sps_ibc_flag >= ( toolset_idc_l &<br>0x400 ) >> 10     |
| 11 | sps_iqt_flag    | sps_iqt_flag <= ( toolset_idc_h & 0x800 )<br>>> 11 && sps_iqt_flag >= ( toolset_idc_l &<br>0x800 ) >> 11     |
| 12 | sps_htdf_flag   | sps_htdf_flag <= ( toolset_idc_h & 0x1000 )<br>>> 12 && sps_htdf_flag >= ( toolset_idc_l &<br>0x1000 ) >> 12 |

|        |                                        |                                                                                                              |
|--------|----------------------------------------|--------------------------------------------------------------------------------------------------------------|
| 13     | sps_addb_flag                          | sps_addb_flag <= ( toolset_idc_h & 0x2000 ) >>13 && sps_addb_flag >= ( toolset_idc_l & 0x2000 ) >>13         |
| 14     | sps_cm_init_flag                       | sps_cm_init_flag <= ( toolset_idc_h & 0x4000 ) >>14 && sps_cm_init_flag >= ( toolset_idc_l & 0x4000 ) >>14   |
| 15     | sps_ats_flag                           | sps_ats_flag <= ( toolset_idc_h & 0x8000 ) >> 15 && sps_ats_flag >= ( toolset_idc_l & 0x8000 ) >> 15         |
| 16     | sps_rpl_flag                           | sps_rpl_flag <= ( toolset_idc_h & 0x10000 ) >> 16 && sps_rpl_flag >= ( toolset_idc_l & 0x10000 ) >> 16       |
| 17     | sps_pocs_flag                          | sps_pocs_flag <= ( toolset_idc_h & 0x20000 ) >> 17 && sps_pocs_flag >= ( toolset_idc_l & 0x20000 ) >> 17     |
| 18     | sps_dquant_flag                        | sps_dquant_flag <= ( toolset_idc_h & 0x40000 ) >> 18 && sps_dquant_flag >= ( toolset_idc_l & 0x40000 ) >> 18 |
| 19     | sps_dra_flag                           | sps_dra_flag <= ( toolset_idc_h & 0x80000 ) >> 19 && sps_dra_flag >= ( toolset_idc_l & 0x80000 ) >> 19       |
| 20     | sps_hmvp_flag                          | sps_hmvp_flag <= ( toolset_idc_h & 0x100000 ) >> 20 && sps_hmvp_flag >= ( toolset_idc_l & 0x100000 ) >> 20   |
| 21..31 | Reservado para uso futuro por ISO/IEC. | Reservado para uso futuro por ISO/IEC.                                                                       |

**[0077]** A continuación se describe la sintaxis utilizada por el codificador de video 200 y el decodificador de video 300.

```
aligned(8) class EVCDDecoderConfigurationRecord {
    unsigned int(8) configurationVersion = 1;
    unsigned int(8) profile_idc;  unsigned int(8) level_idc;
    unsigned int(32) toolset_idc;
    unsigned int(2) chroma_format_idc; unsigned int(3) bit_depth_luma_minus8;
```

```

    unsigned int(3) bit_depth_chroma_minus8;
    unsigned int(32) pic_width_in_luma_samples;
    unsigned int(32) pic_height_in_luma_samples;
    unsigned int(5) reserved = '00000'b;
    unsigned int(1) sps_in_stream;
    unsigned int(1) pps_in_stream;
    unsigned int(1) aps_in_stream;
    unsigned int(8) numOfArrays;
    for (j=0; j < numOfArrays; j++) {
        bit(2) reserved = '00'b;
        unsigned int(6) NAL_unit_type;
        unsigned int(16) numNalus;
        for (i=0; i < numNalus; i++) {
            unsigned int(16) nalUnitLength;
            bit(8*nalUnitLength) nalUnit;
        }
    }
}

```

**[0078]** La siguiente descripción ilustra la semántica para definir los términos de la sintaxis anterior. Los elementos de sintaxis `profile_idc`, `level_idc`, `toolset_idc`, `chroma_format_idc`, `toolset_idc`, `bit_depth_luma_minus8` y `bit_depth_chroma_minus8` contienen los valores correspondientes a los campos del PPS para todos los conjuntos de parámetros del registro de configuración. La notación "(32)" indica que el elemento de sintaxis `toolset_idc` es de 32 bits. Estos 32 bits pueden incluir indicadores de un bit que corresponden cada uno a una herramienta específica, o en algunos casos, bits múltiples de los 32 bits pueden, por ejemplo, corresponder a combinaciones de herramientas o la selección de herramientas de un conjunto de herramientas.

**[0079]** Los elementos de sintaxis `pic_width_in_luma_samples` y `pic_height_in_luma_samples` contienen los valores más grandes para los campos de todo el SPS de este registro de configuración cuando el valor del campo `sps_in_stream` es "0". Estos elementos de sintaxis pueden contener los valores más grandes para los campos de todo el SPS de este registro de configuración y todo el SPS del flujo cuando el valor del campo `sps_in_stream` es "1". El valor "0" se puede (por ejemplo, se debe) usar si el valor más grande de estos campos del SPS para todos los conjuntos de parámetros en este registro no se indica a través de este

campo cuando el valor del campo `sps_in_stream` es "0" o el valor de estos campos del SPS en el flujo tiene el valor más grande que el valor más grande del campo en este registro cuando el valor de `sps_in_stream` es "1".

**[0080]** El elemento de sintaxis `sps_in_stream` indica que el flujo puede contener SPS adicional, que no se incluye en la matriz de unidades de NAL de este registro de configuración. El elemento de sintaxis `pps_in_stream` indica que el flujo puede contener PPS adicional, que no se incluye en la matriz de unidades de NAL de este registro de configuración. El elemento de sintaxis `aps_in_stream` indica que el flujo puede contener APS adicional, que no se incluye en la matriz de unidades de NAL de este registro de configuración.

**[0081]** El elemento de sintaxis `numArrays` indica el número de matrices de unidades de NAL de los tipos indicados. El elemento de sintaxis `NAL_unit_type` indica el tipo de unidades de NAL en la siguiente matriz (que puede (por ejemplo, debe) ser todo de ese tipo). `NAL_unit_type` toma un valor como se define en ISO/IEC 23094-1, y puede limitarse a tomar uno de los valores que indican una unidad de NAL SPS, PPS, APS o SEI.

**[0082]** El elemento de sintaxis `numNalus` indica el número de unidades de NAL del tipo indicado incluidas en el registro de configuración del flujo al que se aplica este registro de configuración. El elemento de sintaxis `nalUnitLength` indica la longitud en bytes de la unidad de NAL. El elemento de sintaxis `nalUnit` contiene una unidad de NAL SPS, PPS, APS o SEI, como se especifica en ISO/IEC 23094-1.

**[0083]** Los archivos de video de acuerdo con el formato de archivo de medios base ISO y extensiones de este almacenan datos en una serie de objetos, denominados "cuadros". A continuación se describe el formato de archivo de medios base ISO, que incluye la definición de flujo de video EVC y nombre y formato de una entrada de muestra.

- a. Entrada de muestra y Tipos de cuadro: 'evc1', 'evcC'
- b. Contenedor: Cuadro de tabla de muestra ('stbl')
- c. Obligatorio: Una entrada de muestra "evc1" es obligatoria.
- d. Cantidad: Puede haber una o más entradas de muestra

**[0084]** Una entrada de muestra visual EVC puede (por ejemplo, debe) contener un Cuadro de configuración EVC, como se define a continuación. Este incluye un `EVCDecoderConfigurationRecord`.

**[0085]** Puede haber un `BitRateBox` opcional en la entrada de muestra visual EVC para señalar la información de la tasa de bits del flujo de video EVC.

**[0086]** Se pueden usar múltiples entradas de muestra, según lo que permite la especificación del formato de archivo de medios base ISO, para indicar secciones de video que usan diferentes configuraciones o conjuntos de parámetros.

**[0087]** Cuando el nombre de la entrada de muestra es "evc1", el flujo al que se aplica esta entrada de muestra puede (por ejemplo, debe) ser un flujo EVC compatible tal como lo ve un decodificador EVC (por ejemplo, el decodificador de video 300) que funciona con la configuración (que incluye el perfil, nivel y conjunto de herramientas) proporcionada en el EVCConfigurationBox.

**[0088]** La entrada de muestra 'evc1' permite almacenar los conjuntos de parámetros en la entrada de muestra y en el flujo. Cuando se establecen en 0, sps\_in\_stream, pps\_in\_stream, y aps\_in\_stream indican que se completó la matriz de las unidades de NAL del tipo correspondiente.

**[0089]** A continuación se muestra un conjunto de ejemplo de sintaxis para un cuadro de configuración para el formato de archivo.

```
class EVCConfigurationBox extends Box('evcC') {
    EVCDecoderConfigurationRecord() EVCConfig;
}

class EVCSampleEntry() extends VisualSampleEntry ('evc1'){
    EVCConfigurationBox config;
    MPEG4ExtensionDescriptorsBox (); // optional}
```

**[0090]** A continuación se muestra un conjunto de ejemplo de semántica para la sintaxis descrita anteriormente.

**[0091]** Compressorname en la clase base VisualSampleEntry indica el nombre del compresor utilizado con el valor "\012EVC Coding" que se recomienda (\012 es 10, la longitud de la cadena en bytes). EVCDecoderConfigurationRecord se define en la sección 5.3.3 (por ejemplo, tal como posible, pero no se limita a ISO/IEC 23904-1).

**[0092]** A continuación se describen los conjuntos de parámetros. A modo de resumen, al menos los SPS y PPS iniciales con ids iguales a 0 pueden (por ejemplo, deben) incluirse en la entrada de la muestra. En caso de establecer sps\_in\_stream y/o pps\_in\_stream en "1", puede haber SPS y/o PPS adicionales en banda en el flujo.

**[0093]** Las muestras que incluyen un conjunto de parámetros pueden (por ejemplo, deben) pertenecer a un grupo de muestras que corresponde al tipo de ese conjunto de parámetros. En esta memoria descriptiva se definen tres ejemplos de group\_type "pss1". Otro parámetro de tipo de agrupación se utiliza para diferenciar entre SPS, PPS y APS, donde "sps1" identifica un grupo de muestras que incluye un SPS, "pps1"

identifica un grupo de muestras que incluye un PPS, y "aps1" identifica un grupo de muestras que incluye un APS.

**[0094]** A continuación se describe la entrada del grupo de muestra del conjunto de parámetros que incluye las definiciones.

- a. Tipos de grupo: "pss1"
- b. Contenedor: Cuadro de descripción de grupo de muestras ("sgpd")
- c. Obligatorio: No
- d. Cantidad: Cero o más

**[0095]** Un grupo de muestras de conjuntos de parámetros identifica muestras que contienen un conjunto de parámetros de tipo SPS, PPS o APS. El `grouping_type_parameter` identifica además el tipo del conjunto de parámetros y puede tomar un valor de "sps1", "pps1" o "aps1".

**[0096]** A continuación se describe parte de la sintaxis de la entrada del grupo de muestra del conjunto de parámetros.

```
class PSSSampleEntry() extends VisualSampleGroupEntry('pss1')
{
}
```

**[0097]** A continuación se describe la muestra de sincronización. Una muestra de sincronización en las pistas 'evc1' puede (por ejemplo, debe) contener unidades de NAL de VCL que indiquen que la imagen codificada con `nuh_temporal_id` igual a 0 en la muestra es una imagen de actualización de decodificación instantánea (IDR).

**[0098]** La Tabla 1 indica el mapeo entre los tipos de unidades de NAL de VCL EVC, el estado de muestra de sincronización de ISOBMFF y los tipos SAP como se documenta en ISOBMFF.

**[0099] Tabla 1 – Mapeo del estado de muestra de sincronización y tipos SAP al tipo de unidad de NAL**

| Tipo de unidad de NAL | Estado de muestra de sincronización de ISOBMFF | Tipo SAP DASH |
|-----------------------|------------------------------------------------|---------------|
| IDR_N_LP              | verdadero                                      | 1             |

**[0100]** A continuación se define una submuestra para EVC. Para el uso de `SubSampleInformationBox` (8.7.7 de ISO/IEC 14496-12) en un flujo EVC, se define

una submuestra sobre la base del valor del campo de indicadores del cuadro de información de submuestra como se especifica a continuación. La presencia de este cuadro es opcional; sin embargo, si está presente en una pista que contiene datos EVC, el campo "codec\_specific\_parameters" del cuadro puede (por ejemplo, debe) tener la semántica definida en la presente.

**[0101]** Indicadores especifican el tipo de información de submuestra proporcionada en el cuadro de la siguiente manera:

- a. 0: Submuestras basadas en unidades de NAL: Una submuestra contiene una o más unidades de NAL contiguas.
- b. 1: Submuestras basadas en mosaicos: Una submuestra contiene las unidades de NAL de VCL con todas las CTU de un mosaico junto con las unidades de NAL no VCL asociadas, si las hubiere.
- c. 2: Submuestras basadas en porciones; una submuestra contiene una porción (es decir, una unidad de NAL VCL) y las unidades de NAL no VCL asociadas, si las hubiere.
- d. Se reservan otros valores de los indicadores.

**[0102]** El campo subsample\_priority se puede (por ejemplo, debe) establecer a un valor de acuerdo con la especificación de este campo en ISO/IEC 14496-12.

**[0103]** El campo desecheable se puede (por ejemplo, debe) establecer en 1 solo si esta muestra sigue siendo decodificable si se descarta esta submuestra (por ejemplo, la submuestra consiste en una unidad de NAL SEI).

**[0104]** Cuando el primer byte de una unidad de NAL se incluye en una submuestra, el campo de longitud precedente puede (por ejemplo, debe) incluirse también en la misma submuestra.

**[0105]** El campo codec\_specific\_parameters del SubSampleInformationBox se define para EVC de la siguiente manera:

```
if (flags == 1) {
    unsigned int(16) tile_col_idx;
    unsigned int(16) tile_row_idx;
}
```

**[0106]** tile\_col\_idx para las submuestras basadas en mosaicos, este parámetro indica el índice basado en 0 de la columna de mosaicos que contiene el mosaico de esta submuestra.

**[0107]** tile\_row\_idx para las submuestras basadas en mosaicos, este parámetro indica el índice basado en 0 de la fila de mosaicos que contiene el mosaico de esta

submuestra.

**[0108]** A continuación se describe el perfil de medios CMAF. El formato de aplicación de medios comunes (CMAF) ISO/IEC 23000-19 define las restricciones estructurales de los archivos ISOBMFF adicionales a las de ISO/IEC 14496-12 con el fin de, por ejemplo, realizar *streaming* adaptativo o suministrar archivos encriptados. La conformidad con estas restricciones estructurales se indica mediante la presencia de una marca estructural definida por CMAF en el FileTypeBox.

**[0109]** En caso de que una pista ISO BMFF use la marca "cevc", se denomina una Pista CMAF EVC y se aplican las siguientes restricciones, lo que define el perfil de medios CMAF para EVC:

- a. puede (por ejemplo, debe) usar la entrada de muestra "evc1" como se define en la cláusula 6.3 de ISO/IEC 23094-1.
- b. la pista puede (por ejemplo, debe) adecuarse a las restricciones de pista CMAF generales en ISO/IEC 23000-19, cláusula 7.
- c. la pista puede (por ejemplo, debe) adecuarse a las restricciones de pista de video CMAF generales en ISO/IEC 23000-19, cláusula 9.

**[0110]** En caso de que se proporcione medios EVC en un conjunto de conmutación CMAF, entonces

- a. cada pista CMAF en el conjunto de conmutación CMAF puede (por ejemplo, debe) adecuarse a una Pista CMAF EVC
- b. el conjunto de conmutación CMAF puede (por ejemplo, debe) adecuarse a las restricciones de conjunto de conmutación CMAF generales en ISO/IEC 23000-19, cláusula 7, y
- c. las restricciones de conjunto de conmutación de pista de video CMAF generales definidas en ISO/IEC 23000-19, cláusula 9.

**[0111]** Un conjunto de conmutación CMAF de acuerdo con estos requisitos se define como el perfil de medios CMAF EVC "cevc". El cifrado de las pistas CMAF EVC y de los conjuntos de conmutación CMAF EVC puede (por ejemplo, debe) cumplir con ISO/IEC 23000-19 cláusula 8, mediante el uso del esquema "cenc" AES-CTR o el esquema de cifrado de patrones de submuestra "cbcs" AES-CBC, como se especifica en ISO/IEC 23001-7, cláusula 10.1 y 10.4, respectivamente.

**[0112]** Además, en caso de que el modo "cbcs" de Common Encryption use el cifrado de patrones definido en la cláusula 9.6 de ISO/IEC 23001-7, entonces se puede (por ejemplo, se debe) aplicar una longitud de bloque de patrones de 10 y un patrón de encrypt:skip de 1:9 (por ejemplo, como se describe en la cláusula 10.4 de ISO/IEC

23001-7).

**[0113]** A continuación se describe el mapeo al suministro de DASH. Si se proporcionan medios codificados con EVC en una presentación de medios DASH en un conjunto de adaptación, el conjunto de adaptación puede (por ejemplo, debe) adecuarse al perfil DASH para CMAF definido en ISO/IEC 23009-1. Los siguientes parámetros pueden (por ejemplo, deben) estar presentes en el nivel del conjunto de adaptación y establecer:

- a. @codecs se establece de acuerdo con el Anexo A
- b. @mimeType se establece para ser compatible con "video/mp4 profiles='cevc'"

**[0114]** A continuación se describen los subparámetros del parámetro "códigos" de tipo MIME. DASH y otras aplicaciones requieren valores definidos para el parámetro de códecs especificado en IETF RFC 6381 para las pistas de medios ISO BMFF. La cadena de parámetros de los códecs para el códec EVC es la siguiente: <entrada de ejemplo 4CC>.<clave1><valor1>.<clave2><valor2>.....<claveN><valorN>

**[0115]** Las claves se definen como 4CC. Se define un conjunto inicial de claves y los pares de valores asociados en la Tabla 2. Se pueden especificar claves adicionales como 4CC. En algunos ejemplos, las claves se adecuan a la ISO/IEC 23091-2.

**[0116]** Si no se proporciona una clave específica, entonces se aplica un valor predeterminado especificado o, si no se anuncia (n/a), se desconoce el valor de la clave.

**Tabla 2 Definición del conjunto inicial de claves y valores definidos para EVC**

| Clave  | Definición de clave                                                  | Valor                                                                                                                 | Valor predeterminado |
|--------|----------------------------------------------------------------------|-----------------------------------------------------------------------------------------------------------------------|----------------------|
| 'vprf' | Define el perfil de video                                            | profile_idc                                                                                                           | ?                    |
| 'vlev' | Define el nivel de video                                             | level_idc                                                                                                             | ?                    |
| 'vtoo' | Define el conjunto de herramientas                                   | toolset_idc in hex<br>decimal                                                                                         | 0000                 |
| 'vbit' | Define la profundidad de bits de video para luminancia y crominancia | El primer dígito es la profundidad de bits de luminancia -8 y el segundo es la profundidad de bits de crominancia - 8 | 00                   |

|        |                                                                                                                  |                                                                                                                                                                                                  |                                  |
|--------|------------------------------------------------------------------------------------------------------------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----------------------------------|
| 'vcss' | Define el submuestreo de crominancia                                                                             | El primer dígito es el primer dígito, el segundo y el tercero                                                                                                                                    | 420                              |
| 'vcpr' | Define los colores primarios (ColourPrimaries) como se define en la ISO/IEC 23091-2.                             | Define los valores de colores primarios (ColourPrimaries) como se define en la ISO/IEC 23091-2. El valor es un decimal de 2 dígitos con un posible 0 precedente.                                 | 01                               |
| 'vtrc' | Define las características de transferencia (TransferCharacteristics) como se define en la norma ISO/IEC23091-2. | Define los valores de las características de transferencia (TransferCharacteristics) como se define en la norma ISO/IEC23091-2. El valor es un decimal de 2 dígitos con un posible 0 precedente. | 01                               |
| 'vmac' | Define los coeficientes de la matriz (MatrixCoefficients) como se define en la norma ISO/IEC23091-2.             | Define los valores de los coeficientes de la matriz (MatrixCoefficients) como se define en la norma ISO/IEC23091-2. El valor es un decimal de 2 dígitos con un posible 0 precedente.             | 01                               |
| 'vfrf' | Define VideoFullRangeFlag como se define en la norma ISO/IEC23091-2.                                             | Define VideoFullRangeFlag como se define en la norma ISO/IEC23091-2.                                                                                                                             | 0                                |
| 'vfpq' | Define el tipo de compresión de los cuadros de vídeo                                                             | Define los valores del tipo de compresión de los cuadros de vídeo                                                                                                                                | En caso de que no esté presente, |

|        |                                                                                                                               |                                                                                                                                                                                                  |                                                                           |
|--------|-------------------------------------------------------------------------------------------------------------------------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|---------------------------------------------------------------------------|
|        | (VideoFramePackingType) como se define en ISO/IEC23091-2 junto con QuincunxSamplingFlag.                                      | (VideoFramePackingType) como se define en ISO/IEC23091-2. El valor es un decimal de 2 dígitos, en donde el primero es el QuincunxSamplingFlag y el segundo dígito el VideoFramePackingType.      | entonces no se usa la compresión de cuadros.                              |
| 'vpci' | Define el tipo de interpretación del contenido comprimido (PackedContentInterpretationType) como se define en ISO/IEC23091-2. | Tipo de interpretación del contenido comprimido (PackedContentInterpretationType) como se define en ISO/IEC23091-2. El valor es de 1 dígito.                                                     | En caso de que no esté presente, entonces no se usa contenido comprimido. |
| 'vsar' | Define el indicador de la relación de aspecto de la muestra (SampleAspectRatio), como se define en ISO/IEC23091-2.            | Define los valores del indicador de la relación de aspecto de la muestra (SampleAspectRatio), como se define en ISO/IEC23091-2. El valor es un decimal de 2 dígitos con un posible 0 precedente. | 01                                                                        |

**[0117]** Por ejemplo,

codecs="evc1.vprf3.vlev51.vtoo03FF.vbit20.vcss420.vcpr09.vtrc16.vmac09.vsar01"

representa el perfil principal de EVC, nivel 5.1, con submuestreo de crominancia 4:2:0 ubicado junto con (0, 0) muestra de luminancia, un conjunto de herramientas restringido, colores primarios ITU-R BT.2100, características de transferencia PQ ITU-R BT.2100, matriz de color YCbCr ITU-R BT.2100 y relación de aspecto de muestra 1:1. Todas las claves de la Tabla 2 pueden ser (por ejemplo, deben ser) reconocidas, si se reconoce la entrada de muestra evc1. Si no se reconoce una clave, se ignora el

par clave valor. En algunos ejemplos, se pueden definir otras claves, por ejemplo, solo 2CC.

**[0118]** El par clave-valor de 'vtoo' y toolset\_idc transmite la misma información que el elemento de sintaxis toolset\_idc en el registro de configuración del decodificador, pero representa una manera diferente de exponer esta información a un dispositivo de procesamiento de video.

**[0119]** De acuerdo con las técnicas descritas anteriormente, el dispositivo de destino 116 se puede configurar para recibir un registro de configuración para decodificar un flujo de bits de los datos de video, en donde el registro de configuración para el flujo de bits incluye un elemento de sintaxis de indicación de conjunto de herramientas que incluye información que identifica, a partir de un conjunto de herramientas de decodificación de video, herramientas que se requieren para decodificar el flujo de bits asociado con el registro de configuración; en función del elemento de sintaxis de indicación del conjunto de herramientas, determinar si recuperar el flujo de bits asociado con el registro de configuración; y en función de la determinación de recuperar el flujo de bits asociado con el registro de configuración, recuperar el flujo de bits y enviar el flujo de bits a un decodificador de video para su decodificación. El dispositivo de destino 116, de manera adicional o alternativa, puede recibir un parámetro de tipo MIME que incluye un par de valores clave, en donde la clave indica que el tipo de MIME identifica las herramientas de decodificación de video, y el valor identifica, a partir de las herramientas de decodificación de video, las herramientas que se requieren para decodificar el flujo de bits. El flujo de bits de datos de video puede incluir uno o más conjuntos de parámetros, y cada una de las herramientas identificadas en el registro de configuración como requeridas para decodificar el flujo de bits se pueden habilitar en al menos uno del uno o más conjuntos de parámetros.

**[0120]** El elemento de sintaxis de indicación de conjunto de herramientas se señala en un cuadro de configuración de información de formato de archivo. El elemento de sintaxis de indicación de conjunto de herramientas que incluye información que identifica las herramientas que se requieren para decodificar el flujo de bits asociado con el registro de configuración puede identificar todas las herramientas que se requieren para decodificar el flujo de bits asociado con el registro de configuración. El elemento de sintaxis de indicación de conjunto de herramientas puede ser un número entero de 32 bits sin signo, en donde cada bit del número entero de 32 bits sin signo corresponde a una herramienta única para decodificar el flujo de bits.

**[0121]** El registro de configuración se puede formatear de acuerdo con el estándar

EVC esencial. El registro de configuración para el flujo de bits puede incluir un elemento de sintaxis de perfil y/o un elemento de sintaxis de nivel antes del elemento de sintaxis de indicación de conjunto de herramientas. El registro de configuración para el flujo de bits puede incluir un elemento de sintaxis de formato de croma después del elemento de sintaxis de indicación de conjunto de herramientas.

**[0122]** Las Figuras 2A y 2B son diagramas conceptuales que ilustran una estructura de árbol binario de árbol cuaternario (QTBT) 130 de ejemplo y una unidad de árbol de codificación (CTU) 132 correspondiente. Las líneas continuas representan la división de árbol cuaternario y las líneas de puntos indican la división de árbol binario. En cada nodo dividido (es decir, no hoja) del árbol binario, se señala un indicador para indicar qué tipo de división (es decir, horizontal o vertical) se utiliza, donde 0 indica división horizontal y 1 indica división vertical en este ejemplo. Para la división de árbol cuaternario, no es necesario indicar el tipo de división, porque los nodos de árbol cuaternario dividen un bloque horizontal y verticalmente en 4 subbloques con el mismo tamaño. En consecuencia, el codificador de video 200 puede codificar, y el decodificador de video 300 puede decodificar, elementos de sintaxis (tal como dividir información) para un nivel de árbol de región de la estructura QTBT 130 (es decir, las líneas continuas) y elementos de sintaxis (tal como dividir información) para un nivel de árbol de predicción de la estructura QTBT 130 (es decir, las líneas discontinuas). El codificador de video 200 puede codificar, y el decodificador de video 300 puede decodificar, datos de video, tales como datos de predicción y transformación, para CU representadas por nodos hoja terminal de la estructura QTBT 130.

**[0123]** En general, la CTU 132 de la Figura 2B puede estar asociada con parámetros que definen tamaños de bloques correspondientes a nodos de la estructura QTBT 130 en el primer y segundo nivel. Estos parámetros pueden incluir un tamaño de CTU (que representa un tamaño de CTU 132 en muestras), un tamaño mínimo de árbol cuaternario (MinQTSIZE, que representa un tamaño mínimo de nodo hoja de árbol cuaternario permitido), un tamaño máximo de árbol binario (MaxBTSIZE, que representa un tamaño de nodo raíz de árbol binario máximo permitido), una profundidad máxima de árbol binario (MaxBTDepth, que representa una profundidad máxima permitida de árbol binario) y un tamaño mínimo de árbol binario (MinBTSIZE, que representa el tamaño mínimo permitido de nodo hoja de árbol binario).

**[0124]** El nodo raíz de una estructura QTBT correspondiente a una CTU puede tener cuatro nodos secundarios en el primer nivel de la estructura QTBT, cada uno de los cuales puede particionarse de acuerdo con la partición de árbol cuaternario. Es decir,

los nodos del primer nivel son nodos hoja (que no tienen nodos secundarios) o tienen cuatro nodos secundarios. El ejemplo de la estructura QTBT 130 representa tales nodos que incluyen el nodo principal y los nodos secundarios que tienen líneas continuas para las ramas. Si los nodos del primer nivel no son más grandes que el tamaño máximo permitido del nodo raíz del árbol binario (MaxBTSIZE), se pueden particionar aún más mediante los respectivos árboles binarios. La división del árbol binario de un nodo se puede iterar hasta que los nodos resultantes de la división alcancen el tamaño mínimo permitido del nodo hoja del árbol binario (MinBTSIZE) o la profundidad máxima permitida del árbol binario (MaxBTDepth). El ejemplo de la estructura QTBT 130 representa tales nodos que tienen líneas discontinuas para las ramas. El nodo hoja del árbol binario se denomina unidad de codificación (CU), que se utiliza para la predicción (por ejemplo, predicción intraimagen o interimagen) y transformada, sin partición adicional. Como se analizó anteriormente, las CU también se pueden denominar "bloques de video" o "bloques".

**[0125]** En un ejemplo de la estructura de partición QTBT, el tamaño de CTU se establece en 128x128 (muestras de luminancia y dos muestras de crominancia de 64x64 correspondientes), MinQTSIZE se establece en 16x16, MaxBTSIZE se establece en 64x64, MinBTSIZE (para ancho y altura) se establece en 4 y MaxBTDepth se establece en 4. La partición del árbol cuaternario se aplica primero a la CTU para generar nodos hoja de árbol cuaternario. Los nodos hoja de árbol cuaternario pueden tener un tamaño de 16x16 (es decir, MinQTSIZE) a 128x128 (es decir, el tamaño de CTU). Si el nodo hoja de árbol cuaternario es 128x128, el árbol binario no dividirá más el nodo hoja de árbol cuaternario, porque el tamaño excede el MaxBTSIZE (es decir, 64x64, en este ejemplo). De lo contrario, el árbol binario dividirá aún más el nodo hoja de árbol cuaternario. Por lo tanto, el nodo hoja de árbol cuaternario es también el nodo raíz del árbol binario y tiene la profundidad del árbol binario como 0. Cuando la profundidad del árbol binario alcanza MaxBTDepth (4, en este ejemplo), no se permite más división. Cuando el nodo de árbol binario tiene un ancho igual a MinBTSIZE (4, en este ejemplo), implica que no se permite más división horizontal. De manera similar, un nodo de árbol binario que tiene una altura igual a MinBTSIZE implica que no se permite más división vertical para ese nodo de árbol binario. Como se señaló anteriormente, los nodos hoja de árbol binario se denominan CU y se procesan adicionalmente de acuerdo con la predicción y la transformada sin más particiones.

**[0126]** La Figura 3 es un diagrama de bloques que ilustra un codificador de video 200 que puede llevar a cabo técnicas de esta divulgación. La Figura 3 se proporciona para

los propósitos de explicación y no se debe considerar limitativo de las técnicas de acuerdo con lo ejemplificado y descrito de manera más amplia en esta divulgación. Para fines de explicación, esta divulgación describe el codificador de video 200 de acuerdo con las técnicas de JEM, EVC, VVC (ITU-T H.266, en desarrollo) y HEVC (ITU-T H.265). Sin embargo, las técnicas de esta divulgación pueden llevarse a cabo mediante dispositivos de codificación de video que están configurados para otros estándares de codificación de video.

**[0127]** En el ejemplo de la Figura 3, el codificador de video 200 incluye la memoria de datos de video 230, la unidad de selección de modo 202, la unidad de generación residual 204, la unidad de procesamiento de transformada 206, la unidad de cuantificación 208, la unidad de cuantificación inversa 210, la unidad de procesamiento de transformada inversa 212, la unidad de reconstrucción 214, la unidad de filtro 216, el búfer de imagen decodificada (DPB) 218 y la unidad de codificación por entropía 220. Cualquiera o todos de la memoria de datos de video 230, la unidad de selección de modo 202, la unidad de generación residual 204, la unidad de procesamiento de transformada 206, la unidad de cuantificación 208, la unidad de cuantificación inversa 210, la unidad de procesamiento de transformada inversa 212, la unidad de reconstrucción 214, la unidad de filtro 216, el DPB 218 y la unidad de codificación por entropía 220 pueden implementarse en uno o más procesadores o en circuitería de procesamiento. Por ejemplo, las unidades del codificador de video 200 pueden implementarse como uno o más circuitos o elementos lógicos como parte de una circuitería de *hardware*, o como parte de un procesador, ASIC, de FPGA. Además, el codificador de video 200 puede incluir procesadores o circuitería de procesamiento adicionales o alternativos para llevar a cabo estas y otras funciones.

**[0128]** La memoria de datos de video 230 puede almacenar los datos de video que van a codificarse por los componentes del codificador de video 200. El codificador de video 200 puede recibir los datos de video almacenados en la memoria de datos de video 230 desde, por ejemplo, la fuente de video 104 (Figura 1). El DPB 218 puede actuar como una memoria de imágenes de referencia que almacena datos de video de referencia para su uso en la predicción de datos de video posteriores por el codificador de video 200. La memoria de datos de video 230 y el DPB 218 se pueden formar mediante cualquiera de una variedad de dispositivos de memoria, tales como memoria dinámica de acceso aleatorio (DRAM), incluso DRAM síncrona (SDRAM), RAM magnetorresistiva (MRAM), RAM resistiva (RRAM) u otros tipos de dispositivos de memoria. La memoria de datos de video 230 y el DPB 218 pueden ser proporcionados

por el mismo dispositivo de memoria o por dispositivos de memoria separados. En diversos ejemplos, la memoria de datos de video 230 puede estar en un chip con otros componentes del codificador de video 200, como se ilustra, o fuera del chip con respecto a esos componentes.

**[0129]** En esta divulgación, la referencia a la memoria de datos de video 230 no debe interpretarse como limitada a la memoria interna del codificador de video 200, a menos que se describa específicamente como tal, o la memoria externa al codificador de video 200, a menos que se describa específicamente como tal. Más bien, la referencia a la memoria de datos de video 230 debe entenderse como una memoria de referencia que almacena datos de video que el codificador de video 200 recibe para codificar (por ejemplo, datos de video para un bloque actual que se va a codificar). La memoria 106 de la Figura 1 también puede proporcionar almacenamiento temporal de salidas de las diversas unidades del codificador de video 200.

**[0130]** Las diversas unidades de la Figura 3 se ilustran para ayudar a comprender las operaciones llevadas a cabo por el codificador de video 200. Las unidades pueden implementarse como circuitos de función fija, circuitos programables o una combinación de estos. Los circuitos de función fija se refieren a los circuitos que proporcionan una funcionalidad particular y están preestablecidos en las operaciones que se pueden llevar a cabo. Los circuitos programables se refieren a los circuitos que se pueden programar para realizar diversas tareas y proporcionan una funcionalidad flexible en las operaciones que se pueden realizar. Por ejemplo, los circuitos programables pueden ejecutar *software* o *firmware* que hacen que los circuitos programables funcionen de la manera definida por las instrucciones del *software* o *firmware*. Los circuitos de función fija pueden ejecutar instrucciones de *software* (por ejemplo, para recibir parámetros o emitir parámetros), pero los tipos de operaciones que llevan a cabo los circuitos de función fija son generalmente inmutables. En algunos ejemplos, una o más de las unidades pueden ser bloques de circuitos distintos (de función fija o programables) y, en algunos ejemplos, una o más de una o más unidades pueden ser circuitos integrados.

**[0131]** El codificador de video 200 puede incluir unidades lógicas aritméticas (ALU), unidades de función elemental (EFU), circuitos digitales, circuitos analógicos y/o núcleos programables, formados a partir de circuitos programables. En los ejemplos donde las operaciones del codificador de video 200 se llevan a cabo mediante el uso de *software* ejecutado por los circuitos programables, la memoria 106 (Figura 1) puede almacenar instrucciones (por ejemplo, el código objeto) del *software* que el codificador

de video 200 recibe y ejecuta, u otra memoria dentro del codificador de video 200 (no mostrado) puede almacenar tales instrucciones.

**[0132]** La memoria de datos de video 230 está configurada para almacenar datos de video recibidos. El codificador de video 200 puede recuperar una imagen de los datos de video de la memoria de datos de video 230 y proporcionar los datos de video a la unidad de generación residual 204 y a la unidad de selección de modo 202. Los datos de video en la memoria de datos de video 230 pueden ser datos de video sin procesar que se van a codificar.

**[0133]** La unidad de selección de modo 202 incluye una unidad de estimación de movimiento 222, una unidad de compensación de movimiento 224 y una unidad de intrapredicción 226. La unidad de selección de modo 202 puede incluir unidades funcionales adicionales para llevar a cabo la predicción de video de acuerdo con otros modos de predicción. Como ejemplos, la unidad de selección de modo 202 puede incluir una unidad de paleta, una unidad de copia intrabloque (que puede ser parte de la unidad de estimación de movimiento 222 y/o la unidad de compensación de movimiento 224), una unidad afín, una unidad de modelo lineal (LM) o similares.

**[0134]** La unidad de selección de modo 202 generalmente coordina múltiples pasadas de codificación para probar combinaciones de parámetros de codificación y valores de distorsión de velocidad resultantes para tales combinaciones. Los parámetros de codificación pueden incluir la partición de CTU en CU, modos de predicción para las CU, tipos de transformada para datos residuales de las CU, parámetros de cuantificación para datos residuales de las CU, etc. La unidad de selección de modo 202 puede finalmente seleccionar la combinación de parámetros de codificación que tienen valores de distorsión de velocidad que son mejores que las otras combinaciones probadas.

**[0135]** El codificador de video 200 puede dividir una imagen recuperada de la memoria de datos de video 230 en una serie de CTU e incluir una o más CTU dentro de una porción. La unidad de selección de modo 202 puede dividir una CTU de la imagen de acuerdo con una estructura de árbol, tal como la estructura QTBT o la estructura de árbol cuaternario descrita anteriormente. Como se describió anteriormente, el codificador de video 200 puede formar una o más CU a partir de la partición de una CTU de acuerdo con la estructura de árbol. Tal CU también puede denominarse generalmente "bloque de video" o "bloque".

**[0136]** En general, la unidad de selección de modo 202 también controla los componentes de esta (por ejemplo, unidad de estimación de movimiento 222, unidad

de compensación de movimiento 224 y unidad de intrapredicción 226) para generar un bloque de predicción para un bloque actual (por ejemplo, una CU actual, o la porción superpuesta de una PU y una TU). Para la interpredicción de un bloque actual, la unidad de estimación de movimiento 222 puede llevar a cabo una búsqueda de movimiento para identificar uno o más bloques de referencia que coincidan estrechamente en una o más imágenes de referencia (por ejemplo, una o más imágenes codificadas previamente almacenadas en el DPB 218). En particular, la unidad de estimación de movimiento 222 puede calcular un valor representativo de qué tan similar es un bloque de referencia potencial al bloque actual, por ejemplo, de acuerdo con la suma de diferencia absoluta (SAD), la suma de diferencias cuadradas (SSD), la diferencia absoluta media (MAD), las diferencias cuadráticas medias (MSD) o similares. La unidad de estimación de movimiento 222 generalmente puede llevar a cabo estos cálculos mediante el uso de las diferencias muestra por muestra entre el bloque actual y el bloque de referencia que se está considerando. La unidad de estimación de movimiento 222 puede identificar un bloque de referencia que tiene un valor más bajo resultante de estos cálculos, lo que indica un bloque de referencia que se asemeja más al bloque actual.

**[0137]** La unidad de estimación de movimiento 222 puede formar uno o más vectores de movimiento (MV) que definen las posiciones de los bloques de referencia en las imágenes de referencia con respecto a la posición del bloque actual en una imagen actual. La unidad de estimación de movimiento 222 puede entonces proporcionar los vectores de movimiento a la unidad de compensación de movimiento 224. Por ejemplo, para la interpredicción unidireccional, la unidad de estimación de movimiento 222 puede proporcionar un único vector de movimiento, mientras que para la interpredicción bidireccional, la unidad de estimación de movimiento 222 puede proporcionar dos vectores de movimiento. La unidad de compensación de movimiento 224 puede generar entonces un bloque de predicción mediante los vectores de movimiento. Por ejemplo, la unidad de compensación de movimiento 224 puede recuperar datos del bloque de referencia mediante el vector de movimiento. Como otro ejemplo, si el vector de movimiento tiene una precisión de muestra fraccionada, la unidad de compensación de movimiento 224 puede interpolar valores para el bloque de predicción de acuerdo con uno o más filtros de interpolación. Además, para la interpredicción bidireccional, la unidad de compensación de movimiento 224 puede recuperar datos para dos bloques de referencia identificados por vectores de movimiento respectivos y combinar los datos recuperados, por ejemplo, mediante

promediado muestra por muestra o promediado ponderado.

**[0138]** Como otro ejemplo, para la intrapredicción o codificación de intrapredicción, la unidad de intrapredicción 226 puede generar el bloque de predicción a partir de muestras cercanas al bloque actual. Por ejemplo, para los modos direccionales, la unidad de intrapredicción 226 generalmente puede combinar matemáticamente valores de muestras cercanas y completar estos valores calculados en la dirección definida a través del bloque actual para producir el bloque de predicción. Como otro ejemplo, para el modo DC, la unidad de intrapredicción 226 puede calcular un promedio de las muestras cercanas al bloque actual y generar el bloque de predicción para incluir este promedio resultante para cada muestra del bloque de predicción.

**[0139]** La unidad de selección de modo 202 proporciona el bloque de predicción a la unidad de generación residual 204. La unidad de generación residual 204 recibe una versión en bruto, no codificada del bloque actual desde la memoria de datos de video 230 y el bloque de predicción desde la unidad de selección de modo 202. La unidad de generación residual 204 calcula las diferencias muestra por muestra entre el bloque actual y el bloque de predicción. Las diferencias muestra por muestra resultantes definen un bloque residual para el bloque actual. En algunos ejemplos, la unidad de generación residual 204 también puede determinar diferencias entre valores de muestra en el bloque residual para generar un bloque residual mediante el uso de modulación por impulsos codificados diferencial residual (RDPCM). En algunos ejemplos, la unidad de generación residual 204 puede formarse mediante el uso de uno o más circuitos sustractores que llevan a cabo una resta binaria.

**[0140]** En los ejemplos donde la unidad de selección de modo 202 divide las CU en PU, cada PU puede estar asociada con una unidad de predicción de luminancia y las correspondientes unidades de predicción de crominancia. El codificador de video 200 y el decodificador de video 300 pueden soportar PU que tienen diversos tamaños. De acuerdo con lo indicado con anterioridad, el tamaño de una CU se puede referir al tamaño del bloque de codificación de luminancia de la CU y el tamaño de una PU se puede referir al tamaño de una unidad de predicción de luminancia de la PU. Suponiendo que el tamaño de una CU particular es  $2N \times 2N$ , el codificador de video 200 puede soportar tamaños de PU de  $2N \times 2N$  o  $N \times N$  para la intrapredicción y tamaños simétricos de PU de  $2N \times 2N$ ,  $2N \times N$ ,  $N \times 2N$ ,  $N \times N$  o similares para la interpredicción. El codificador de video 200 y el decodificador de video 300 también pueden soportar la partición asimétrica para los tamaños de PU de  $2N \times nU$ ,  $2N \times nD$ ,  $nL \times 2N$  y  $nR \times 2N$  para la interpredicción.

**[0141]** En los ejemplos en donde la unidad de selección de modo 202 no divide adicionalmente una CU en PU, cada CU puede estar asociada con un bloque de codificación de luminancia y los correspondientes bloques de codificación de crominancia. De acuerdo con lo anterior, el tamaño de una CU se puede referir al tamaño del bloque de codificación de luminancia de la CU. El codificador de video 200 y el decodificador de video 300 pueden soportar tamaños de CU de  $2N \times 2N$ ,  $2N \times N$  o  $N \times 2N$ .

**[0142]** Para otras técnicas de codificación de video, tal como una codificación de modo de copia intrabloque, una codificación de modo afín y una codificación de modo de modelo lineal (LM), como algunos ejemplos, la unidad de selección de modo 202, a través de las unidades respectivas asociadas con las técnicas de codificación, genera un bloque de predicción para el bloque actual que se está codificando. En algunos ejemplos, como la codificación del modo de paleta, es posible que la unidad de selección de modo 202 no genere un bloque de predicción y, en su lugar, genere elementos de sintaxis que indican la manera de reconstruir el bloque en función de una paleta seleccionada. En tales modos, la unidad de selección de modo 202 puede proporcionar estos elementos de sintaxis a la unidad de codificación por entropía 220 a codificar.

**[0143]** Como se describió anteriormente, la unidad de generación residual 204 recibe los datos de video para el bloque actual y el bloque de predicción correspondiente. La unidad de generación residual 204 genera entonces un bloque residual para el bloque actual. Para generar el bloque residual, la unidad de generación residual 204 calcula las diferencias muestra por muestra entre el bloque de predicción y el bloque actual.

**[0144]** La unidad de procesamiento de transformada 206 aplica una o más transformadas al bloque residual para generar un bloque de coeficientes de transformada (denominado en la presente "bloque de coeficientes de transformada"). La unidad de procesamiento de transformada 206 puede aplicar diversas transformadas a un bloque residual para formar el bloque de coeficientes de transformada. Por ejemplo, la unidad de procesamiento de transformada 206 puede aplicar una transformada de coseno discreta (DCT), una transformada direccional, una transformada de Karhunen-Loeve (KLT) o una transformada conceptualmente similar a un bloque residual. En algunos ejemplos, la unidad de procesamiento de transformada 206 puede llevar a cabo múltiples transformadas en un bloque residual, por ejemplo, una transformada primaria y una transformada secundaria, tal como una transformada rotacional. En algunos ejemplos, la unidad de procesamiento de transformada 206 no

aplica transformadas a un bloque residual.

**[0145]** La unidad de cuantificación 208 puede cuantificar los coeficientes de transformada en un bloque de coeficientes de transformada, para producir un bloque de coeficientes de transformada cuantificado. La unidad de cuantificación 208 puede cuantificar los coeficientes de transformada de un bloque de coeficientes de transformada de acuerdo con un valor de parámetro de cuantificación (QP) asociado con el bloque actual. El codificador de video 200 (por ejemplo, mediante la unidad de selección de modo 202) puede ajustar el grado de cuantificación aplicado a los bloques de coeficientes de transformada asociados con el bloque actual al ajustar el valor de QP asociado con la CU. La cuantificación puede introducir pérdida de información y, por lo tanto, los coeficientes de transformada cuantificados pueden tener una precisión menor que los coeficientes de transformada originales producidos por la unidad de procesamiento de transformada 206.

**[0146]** La unidad de cuantificación inversa 210 y la unidad de procesamiento de transformada inversa 212 pueden aplicar cuantificación inversa y transformadas inversas a un bloque de coeficientes de transformada cuantificado, respectivamente, para reconstruir un bloque residual a partir del bloque de coeficientes de transformada. La unidad de reconstrucción 214 puede producir un bloque reconstruido correspondiente al bloque actual (aunque potencialmente con algún grado de distorsión) en función del bloque residual reconstruido y un bloque de predicción generado por la unidad de selección de modo 202. Por ejemplo, la unidad de reconstrucción 214 puede agregar muestras del bloque residual reconstruido a las muestras correspondientes del bloque de predicción generado por la unidad de selección de modo 202 para producir el bloque reconstruido.

**[0147]** La unidad de filtro 216 puede llevar a cabo una o más operaciones de filtro en bloques reconstruidos. Por ejemplo, la unidad de filtro 216 puede llevar a cabo operaciones de desbloqueo para reducir los artefactos de bloqueo a lo largo de los bordes de las CU. Las operaciones de la unidad de filtro 216 pueden omitirse, en algunos ejemplos.

**[0148]** El codificador de video 200 almacena bloques reconstruidos en DPB 218. Por ejemplo, en los ejemplos en los que no se necesitan las operaciones de la unidad de filtro 216, la unidad de reconstrucción 214 puede almacenar bloques reconstruidos en el DPB 218. En los ejemplos en los que se necesitan las operaciones de la unidad de filtro 216, la unidad de filtro 216 puede almacenar los bloques reconstruidos filtrados en el DPB 218. La unidad de estimación de movimiento 222 y la unidad de

compensación de movimiento 224 pueden recuperar una imagen de referencia del DPB 218, formada a partir de los bloques reconstruidos (y potencialmente filtrados), para interpredecir bloques de imágenes codificadas posteriormente. Además, la unidad de intrapredicción 226 puede usar bloques reconstruidos en el DPB 218 de una imagen actual para intraprededir otros bloques en la imagen actual.

**[0149]** En general, la unidad de codificación por entropía 220 puede codificar por entropía elementos de sintaxis recibidos de otros componentes funcionales del codificador de video 200. Por ejemplo, la unidad de codificación por entropía 220 puede codificar por entropía bloques de coeficientes de transformada cuantificados a partir de la unidad de cuantificación 208. Como otro ejemplo, la unidad de codificación por entropía 220 puede codificar por entropía elementos de sintaxis de predicción (por ejemplo, información de movimiento para la interpredicción o información de intramodo para la intrapredicción) a partir de la unidad de selección de modo 202. La unidad de codificación por entropía 220 puede llevar a cabo una o más operaciones de codificación por entropía en los elementos de sintaxis, que son otro ejemplo de datos de video, para generar datos codificados por entropía. Por ejemplo, la unidad de codificación por entropía 220 puede llevar a cabo una operación de codificación adaptativa según el contexto de longitud variable (CAVLC), una operación de CABAC, una operación de codificación de longitud de variable a variable (V2V), una operación de codificación aritmética binaria adaptativa según el contexto con base en la sintaxis (SBAC), una operación de codificación por Entropía de Partición de Intervalo de Probabilidad (PIPE), una operación de exponencial de codificación Golomb, u otro tipo de operación de codificación por entropía sobre los datos. En algunos ejemplos, la unidad de codificación por entropía 220 puede funcionar en modo de derivación donde los elementos de sintaxis no están codificados por entropía.

**[0150]** El codificador de video 200 puede generar un flujo de bits que incluye los elementos de sintaxis codificados por entropía necesarios para reconstruir bloques de una porción o imagen. En particular, la unidad de codificación por entropía 220 puede generar el flujo de bits.

**[0151]** Las operaciones descritas anteriormente se describen con respecto a un bloque. Tal descripción debe entenderse como operaciones para un bloque de codificación de luminancia y/o bloques de codificación de crominancia. Como se describió anteriormente, en algunos ejemplos, el bloque de codificación de luminancia y los bloques de codificación de crominancia son componentes de luminancia y crominancia de una CU. En algunos ejemplos, el bloque de codificación de luminancia

y los bloques de codificación de crominancia son componentes de luminancia y crominancia de una PU.

**[0152]** En algunos ejemplos, las operaciones llevadas a cabo con respecto a un bloque de codificación de luminancia no necesitan repetirse para los bloques de codificación de crominancia. Como ejemplo, las operaciones para identificar un vector de movimiento (MV) y una imagen de referencia para un bloque de codificación de luminancia no necesitan repetirse para identificar un MV y una imagen de referencia para los bloques de crominancia. Más bien, el MV para el bloque de codificación de luminancia se puede escalar para determinar el MV para los bloques de crominancia, y la imagen de referencia puede ser la misma. Como otro ejemplo, el proceso de intrapredicción puede ser el mismo para el bloque de codificación de luminancia y los bloques de codificación de crominancia.

**[0153]** El codificador de video 200 representa un ejemplo de un dispositivo configurado para codificar datos de video que incluye una memoria configurada para almacenar datos de video, y una o más unidades de procesamiento implementadas en circuitería y configuradas para llevar a cabo una o más de las técnicas de ejemplo descritas en esta divulgación.

**[0154]** En algunos ejemplos, un dispositivo fuente puede configurarse para recibir la salida del codificador de video 200 y determinar (por ejemplo, generar) un registro de configuración para decodificar un flujo de bits que incluye datos de video codificados mediante el codificador de video 200 de acuerdo con el estándar de codificación de video esencial (EVC) que incluye información para decodificar el flujo de bits. El registro de configuración puede incluir un elemento de sintaxis de indicación de conjunto de herramientas (por ejemplo, `toolset_idc`). Se puede proporcionar el registro de configuración en un nivel de formato de archivo, por ejemplo, como datos de nivel de formato de archivo, o en un nivel que incluye datos de medios codificados a nivel de capa de codificación de video (VCL). Por ejemplo, como se analizó en la presente, el registro de configuración se puede incluir en un cuadro de configuración en un archivo de medios, donde un cuadro de configuración es independiente de un cuadro de fragmento de película u otro cuadro que incluye datos de medios codificados reales. En algunos ejemplos, el elemento de sintaxis de indicación de conjunto de herramientas puede incluir información de todas las herramientas que se requieren para decodificar el flujo de bits asociado con el registro de configuración. El elemento de sintaxis de indicación de conjunto de herramientas puede indicar herramientas que coinciden con otro campo `toolset_idc` que se señala en un conjunto de parámetros de

secuencia (SPS).

**[0155]** En algunos ejemplos, se puede considerar que el elemento de sintaxis de indicación de conjunto de herramientas incluye herramientas que se requieren para reproducir el archivo en la entrada de muestra que forma parte de los parámetros codificados. Para habilitar el elemento de sintaxis de indicación de conjunto de herramientas, se puede utilizar un esquema de codificación para los parámetros codificados en el tipo de MIME.

**[0156]** La Figura 4 es un diagrama de bloques que ilustra un decodificador de video 300 de ejemplo que puede realizar las técnicas de la presente divulgación. La Figura 4 se proporciona con fines explicativos y no se limita a las técnicas que se ejemplifican y describen ampliamente en esta divulgación. Para fines de explicación, esta divulgación describe el decodificador de video 300 de acuerdo con las técnicas de JEM, EVC, VVC (ITU-T H.266, en desarrollo), y HEVC (ITU-T H.265). Sin embargo, las técnicas de esta divulgación pueden llevarse a cabo mediante dispositivos de codificación de video que están configurados para otros estándares de codificación de video.

**[0157]** En el ejemplo de la Figura 4, el decodificador de video 300 incluye la memoria de búfer de imagen codificada (CPB) 320, la unidad de decodificación por entropía 302, la unidad de procesamiento de predicción 304, la unidad de cuantificación inversa 306, la unidad de procesamiento de transformada inversa 308, la unidad de reconstrucción 310, la unidad de filtro 312 y el búfer de imagen decodificada (DPB) 314. Cualquiera o todos de la memoria de CPB 320, la unidad de decodificación por entropía 302, la unidad de procesamiento de predicción 304, la unidad de cuantificación inversa 306, la unidad de procesamiento de transformada inversa 308, la unidad de reconstrucción 310, la unidad de filtro 312 y el DPB 314 pueden implementarse en uno o más procesadores o en circuitería de procesamiento. Por ejemplo, las unidades del decodificador de video 300 pueden implementarse como uno o más circuitos o elementos lógicos como parte de una circuitería de hardware, o como parte de un procesador, ASIC, de FPGA. Además, el decodificador de video 300 puede incluir procesadores o circuitería de procesamiento adicionales o alternativos para llevar a cabo estas y otras funciones.

**[0158]** La unidad de procesamiento de predicción 304 incluye la unidad de compensación de movimiento 316 y la unidad de intrapredicción 318. La unidad de procesamiento de predicción 304 puede incluir unidades adicionales para llevar a cabo la predicción de acuerdo con otros modos de predicción. Como ejemplos, la unidad de procesamiento de predicción 304 puede incluir una unidad de paleta, una unidad de

copia intrabloque (que puede formar parte de la unidad de compensación de movimiento 316), una unidad afín, una unidad de modelo lineal (LM) o similar. En otros ejemplos, el decodificador de video 300 puede incluir más, menos o diferentes componentes funcionales.

**[0159]** La memoria de CPB 320 puede almacenar datos de video, tales como un flujo de bits de video codificado, para ser decodificado por los componentes del decodificador de video 300. Los datos de video almacenados en la memoria de CPB 320 pueden obtenerse, por ejemplo, del medio legible por computadora 110 (Figura 1). La memoria de CPB 320 puede incluir un CPB que almacena datos de video codificados (por ejemplo, elementos de sintaxis) de un flujo de bits de video codificado. Además, la memoria de CPB 320 puede almacenar datos de video distintos de los elementos de sintaxis de una imagen codificada, tales como datos temporales que representan salidas de las diversas unidades del decodificador de video 300. El DPB 314 generalmente almacena imágenes decodificadas, cuyo decodificador de video 300 puede emitir y/o usar como datos de video de referencia al decodificar datos o imágenes subsiguientes del flujo de bits de video codificado. La memoria de CPB 320 y el DPB 314 se pueden formar mediante cualquiera de una variedad de dispositivos de memoria, tal como DRAM, que incluye SDRAM, MRAM, RRAM u otros tipos de dispositivos de memoria. La memoria de CPB 320 y el DPB 314 pueden ser proporcionados por el mismo dispositivo de memoria o por dispositivos de memoria separados. En diversos ejemplos, la memoria de CPB 320 puede estar integrada con otros componentes del decodificador de video 300, o puede ser externa con respecto a esos componentes.

**[0160]** Adicional o alternativamente, en algunos ejemplos, el decodificador de video 300 puede recuperar datos de video codificados de la memoria 120 (Figura 1). Es decir, la memoria 120 puede almacenar datos como se analizó anteriormente con la memoria de CPB 320. Del mismo modo, la memoria 120 puede almacenar instrucciones que el decodificador de video 300 ejecutará, cuando algunas o todas las funciones del decodificador de video 300 se implementan en *software* para que la circuitería de procesamiento del decodificador de video 300 las ejecute.

**[0161]** Las diversas unidades que se muestran en la Figura 4 se ilustran para ayudar a comprender las operaciones llevadas a cabo por el decodificador de video 300. Las unidades pueden implementarse como circuitos de función fija, circuitos programables o una combinación de estos. De manera similar a la Figura 3, los circuitos de función fija se refieren a los circuitos que proporcionan una funcionalidad particular y están

preestablecidos en las operaciones que se pueden llevar a cabo. Los circuitos programables se refieren a los circuitos que se pueden programar para realizar diversas tareas y proporcionan una funcionalidad flexible en las operaciones que se pueden realizar. Por ejemplo, los circuitos programables pueden ejecutar *software* o *firmware* que hacen que los circuitos programables funcionen de la manera definida por las instrucciones del *software* o *firmware*. Los circuitos de función fija pueden ejecutar instrucciones de *software* (por ejemplo, para recibir parámetros o emitir parámetros), pero los tipos de operaciones que llevan a cabo los circuitos de función fija son generalmente inmutables. En algunos ejemplos, una o más de las unidades pueden ser bloques de circuitos distintos (de función fija o programables) y, en algunos ejemplos, una o más de una o más unidades pueden ser circuitos integrados.

**[0162]** El decodificador de video 300 puede incluir ALU, EFU, circuitos digitales, circuitos analógicos y/o núcleos programables formados a partir de circuitos programables. En los ejemplos en los que las operaciones del decodificador de video 300 se llevan a cabo mediante *software* que se ejecuta en los circuitos programables, la memoria integrada o externa puede almacenar instrucciones (por ejemplo, código objeto) del *software* que el decodificador de video 300 recibe y ejecuta.

**[0163]** La unidad de decodificación por entropía 302 puede recibir datos de video codificados desde el CPB y decodificar por entropía los datos de video para reproducir elementos de sintaxis. La unidad de procesamiento de predicción 304, la unidad de cuantificación inversa 306, la unidad de procesamiento de transformada inversa 308, la unidad de reconstrucción 310, y la unidad de filtro 312 pueden generar datos de video decodificados en función de los elementos de sintaxis extraídos del flujo de bits.

**[0164]** En general, el decodificador de video 300 reconstruye una imagen bloque por bloque. El decodificador de video 300 puede llevar a cabo una operación de reconstrucción en cada bloque individualmente (donde el bloque que se está reconstruyendo actualmente, es decir, decodificando, puede denominarse "bloque actual").

**[0165]** La unidad de decodificación por entropía 302 puede decodificar por entropía elementos de sintaxis que definen coeficientes de transformada cuantificados de un bloque de coeficientes de transformada cuantificados, así como información de transformada, tal como un parámetro de cuantificación (QP) y/o indicaciones de modo de transformada. La unidad de cuantificación inversa 306 puede usar el QP asociado con el bloque de coeficientes de transformada cuantificado para determinar un grado de cuantificación e, igualmente, un grado de cuantificación inversa para que se aplique

la unidad de cuantificación inversa 306. La unidad de cuantificación inversa 306 puede, por ejemplo, llevar a cabo una operación de desplazamiento a la izquierda bit a bit para cuantificar de forma inversa los coeficientes de transformada cuantificados. La unidad de cuantificación inversa 306 puede formar de ese modo un bloque de coeficientes de transformada que incluye coeficientes de transformada.

**[0166]** Después de que la unidad de cuantificación inversa 306 forme el bloque de coeficientes de transformada, la unidad de procesamiento de transformada inversa 308 puede aplicar una o más transformadas inversas al bloque de coeficiente de transformada para generar un bloque residual asociado con el bloque actual. Por ejemplo, la unidad de procesamiento de transformada inversa 308 puede aplicar una DCT inversa, una transformada inversa de números enteros, una transformada inversa de Karhunen-Loeve (KLT), una transformada inversa de rotación, una transformada direccional inversa u otra transformada inversa al bloque de coeficientes de transformada.

**[0167]** Además, la unidad de procesamiento de predicción 304 genera un bloque de predicción de acuerdo con los elementos de sintaxis de información de predicción que fueron decodificados por entropía por la unidad de decodificación por entropía 302. Por ejemplo, si los elementos de sintaxis de información de predicción indican que el bloque actual está interpredicho, la unidad de compensación de movimiento 316 puede generar el bloque de predicción. En este caso, los elementos de sintaxis de información de predicción pueden indicar una imagen de referencia en el DPB 314 de la cual recuperar un bloque de referencia, así como un vector de movimiento que identifica una ubicación del bloque de referencia en la imagen de referencia en relación con la ubicación del bloque actual en la imagen actual. La unidad de compensación de movimiento 316 generalmente puede llevar a cabo el proceso de interpredicción de una manera que es sustancialmente similar a la descrita con respecto a la unidad de compensación de movimiento 224 (Figura 3).

**[0168]** Como otro ejemplo, si los elementos de sintaxis de información de predicción indican que el bloque actual está intrapredicho, la unidad de intrapredicción 318 puede generar el bloque de predicción de acuerdo con un modo de intrapredicción indicado por los elementos de sintaxis de información de predicción. De nuevo, la unidad de intrapredicción 318 generalmente puede llevar a cabo el proceso de intrapredicción de una manera que es sustancialmente similar a la descrita con respecto a la unidad de intrapredicción 226 (Figura 3). La unidad de intrapredicción 318 puede recuperar datos de muestras cercanas al bloque actual del DPB 314.

**[0169]** La unidad de reconstrucción 310 puede reconstruir el bloque actual mediante el uso del bloque de predicción y el bloque residual. Por ejemplo, la unidad de reconstrucción 310 puede agregar muestras del bloque residual a las muestras correspondientes del bloque de predicción para reconstruir el bloque actual.

**[0170]** La unidad de filtro 312 puede llevar a cabo una o más operaciones de filtro en bloques reconstruidos. Por ejemplo, la unidad de filtro 312 puede llevar a cabo operaciones de desbloqueo para reducir los artefactos de bloqueo a lo largo de los bloques reconstruidos. Las operaciones de la unidad de filtro 312 no se llevan a cabo necesariamente en todos los ejemplos.

**[0171]** El decodificador de video 300 puede almacenar los bloques reconstruidos en el DPB 314. Por ejemplo, en los ejemplos en los que no se llevan a cabo las operaciones de la unidad de filtro 312, la unidad de reconstrucción 310 puede almacenar bloques reconstruidos en el DPB 314. En los ejemplos en los que se llevan a cabo las operaciones de la unidad de filtro 312, la unidad de filtro 312 puede almacenar los bloques reconstruidos filtrados en el DPB 314. Como se analizó anteriormente, el DPB 314 puede proporcionar información de referencia, tal como muestras de una imagen actual para intrapredicción e imágenes previamente decodificadas para una compensación de movimiento posterior, a la unidad de procesamiento de predicción 304. Además, el decodificador de video 300 puede emitir imágenes decodificadas (por ejemplo, video decodificado) desde el DPB 314 para su presentación posterior en un dispositivo de visualización, tal como el dispositivo de visualización 118 de la Figura 1.

**[0172]** De esta manera, el decodificador de video 300 representa un ejemplo de un dispositivo de decodificación de video que incluye una memoria configurada para almacenar datos de video, y una o más unidades de procesamiento implementadas en circuitería y configuradas para llevar a cabo técnicas de ejemplo descritas en esta divulgación.

**[0173]** En algunos ejemplos, un dispositivo cliente puede configurarse para recibir un registro de configuración para decodificar un flujo de bits de datos de video de acuerdo con el estándar de codificación de video esencial (EVC). El registro de configuración puede incluir un elemento de sintaxis de indicación de conjunto de herramientas (por ejemplo, toolset\_idc). En algunos ejemplos, el elemento de sintaxis de indicación de conjunto de herramientas puede incluir información de todas las herramientas que se requieren para decodificar el flujo de bits asociado con el registro de configuración. El elemento de sintaxis de indicación de conjunto de herramientas del registro de configuración puede coincidir con un elemento de sintaxis de indicación de conjunto de

herramientas señalado en un conjunto de parámetros de secuencia (SPS). En función del elemento de sintaxis de indicación de conjunto de herramientas, el dispositivo cliente puede determinar si recuperar el flujo de bits asociado con el registro de configuración. En función de la determinación de recuperar el flujo de bits asociado con el registro de configuración, el dispositivo cliente puede recuperar el flujo de bits y enviar el flujo de bits al decodificador de video 300 para decodificar.

**[0174]** En algunos ejemplos, se puede considerar que el elemento de sintaxis de indicación de conjunto de herramientas incluye herramientas que se requieren para reproducir el archivo en la entrada de muestra que forma parte de los parámetros codificados. Para habilitar el elemento de sintaxis de indicación de conjunto de herramientas, se puede utilizar un esquema de codificación para los parámetros codificados en el tipo de MIME.

**[0175]** La Figura 5 es un diagrama conceptual que ilustra una estructura de ejemplo de un archivo 500, de acuerdo con una o más técnicas de esta divulgación. En el ejemplo de la Figura 5, archivo 500 incluye un cuadro de película 502 y una pluralidad de cuadros de datos de medios 504. Si bien se ilustra en el ejemplo de la Figura 5 como si estuvieran en el mismo archivo, en otros ejemplos, el cuadro de película 502 y los cuadros de datos de medios 504 pueden estar en archivos separados. Como se indicó anteriormente, un cuadro puede ser un bloque de construcción orientado a objetos definido por un identificador de tipo único y una longitud. Por ejemplo, un cuadro puede ser la estructura de sintaxis elemental en el ISOBMFF, que incluye un tipo de cuadro codificado de cuatro caracteres, un recuento de bytes del cuadro y una carga útil.

**[0176]** El cuadro de película 502 puede contener metadatos para las pistas del archivo 500. Cada pista del archivo 500 puede comprender un flujo continuo de datos de medios. Cada uno de los cuadros de datos de medios 504 puede incluir una o más muestras 505. Cada una de las muestras 505 puede comprender una unidad de acceso de audio o video. Cada unidad de acceso puede comprender múltiples imágenes codificadas para codificación de vistas múltiples o codificación de video escalable. Por ejemplo, una unidad de acceso puede incluir una o más imágenes codificadas para cada capa.

**[0177]** Además, en el ejemplo de la Figura 5, el cuadro de película 502 incluye un cuadro de pista 506. El cuadro de pista 506 puede incluir metadatos para una pista del archivo 500. En otros ejemplos, el cuadro de película 502 puede incluir múltiples cuadros de pista para diferentes pistas del archivo 500. Cuadro de pista 506 incluye un

cuadro de referencia de pista 508 y un cuadro de medios 510. El cuadro de referencia de pista 508 puede incluir un cuadro de tipo de referencia de pista 509. El cuadro de tipo de referencia de pista 509 puede estar asociado con un tipo (por ejemplo, 'tbas') y un identificador de pista que identifica otra pista. De acuerdo con la sexta técnica de esta divulgación, cuadros de referencia de pista de pistas de mosaico pueden o no incluir cuadros de tipo de referencia de pista asociadas con el identificador de tipo 'tbas'.

**[0178]** El cuadro de medios 510 puede contener todos los objetos que declaran información sobre los datos de medios dentro de la pista. El cuadro de medios 510 incluye un cuadro de información de medios 512. El cuadro de información de medios 512 puede contener todos los objetos que declaran información característica de los medios de la pista. El cuadro de información de medios 512 incluye un cuadro de tabla de muestra 514. El cuadro de tabla de muestra 514 puede especificar metadatos específicos de la muestra.

**[0179]** En el ejemplo de la Figura 5, el cuadro de tabla de muestra 514 incluye un cuadro SampleDescription 516, y un cuadro SampleDescription 516 incluye SampleEntry 518. La entrada de muestra 518 incluye la configuración del decodificador EVC 520, que como se describió anteriormente, puede incluir un elemento de sintaxis de indicación de conjunto de herramientas. En otros ejemplos, el cuadro de tabla de muestra 514 puede incluir otros cuadros además del cuadro SampleDescription 516 y SampleEntry 518. Aunque no se muestra en la Figura 5, la entrada de muestra 518 puede incluir un registro de configuración de decodificador. Como se describió con más detalle anteriormente, esta divulgación describe técnicas para incluir nuevos tipos de información importante en una entrada de muestra tal como la entrada de muestra 518.

**[0180]** La Figura 6 es un diagrama conceptual que ilustra elementos de contenido de medios de ejemplo 620. En el ejemplo de la Figura 6, el contenido multimedia 620 incluye MPD 622 y una pluralidad de representaciones 624A-624N (representaciones 624). Representación 624A incluye datos de encabezado 626 opcionales y los segmentos 628A-628N (segmentos 628), mientras que la representación 624N incluye datos de encabezado 630 opcionales y los segmentos 632A-632N (segmentos 632). La letra N se usa para designar el último fragmento de película en cada una de las representaciones 624 por conveniencia. En algunos ejemplos, puede haber diferentes números de fragmentos de película entre las representaciones 624.

**[0181]** MPD 622 puede comprender una estructura de datos separada de las

representaciones 624. En general, MPD 422 puede incluir datos que generalmente describen las características de las representaciones 424, tal como las características de codificación y representación, conjuntos de adaptación, un perfil al que corresponde MPD 422, información de tipo de texto, información del ángulo de la cámara, información de calificación, información del control de movimiento (por ejemplo, información indicativa de representaciones que incluyen subsecuencias temporales), y/o información para recuperar períodos remotos (por ejemplo, para la inserción de anuncios dirigidos en contenido de medios durante la reproducción).

**[0182]** Los datos de encabezado 626, cuando están presentes, pueden describir las características de los segmentos 628, por ejemplo, las ubicaciones temporales de los puntos de acceso aleatorio (RAP, también denominados puntos de acceso de flujo (SAP)), de los que los segmentos 628 incluyen puntos de acceso aleatorio, desplazamientos de bytes a puntos de acceso aleatorio dentro de los segmentos 628, localizadores uniformes de recursos (URL) de los segmentos 628, u otros aspectos de los segmentos 628. Los datos de encabezado 630, cuando están presentes, pueden describir características similares para los segmentos 632. Adicional o alternativamente, tales características pueden incluirse completamente dentro de MPD 622.

**[0183]** Los segmentos 628, 632 incluyen una o más muestras de video codificado, cada una de las cuales puede incluir fotogramas o porciones de datos de video. Cada una de las muestras de video codificadas de los segmentos 628 puede tener características similares, por ejemplo, requisitos de altura, ancho y ancho de banda. Tales características pueden describirse mediante los datos de MPD 622, aunque tales datos no se ilustran en el ejemplo de la Figura 6. MPD 622 puede incluir características como se describe mediante la Especificación 3GPP, con la adición de cualquiera o toda la información señalada descrita en esta divulgación.

**[0184]** Cada uno de los segmentos 628, 632 puede estar asociado con un único localizador uniforme de recursos (URL). Por lo tanto, cada uno de los segmentos 628, 632 puede recuperarse de manera independiente mediante el uso de un protocolo de red de transmisión, tal como DASH. De esta manera, un dispositivo de destino puede usar una solicitud HTTP GET para recuperar los segmentos 628 o 632. En algunos ejemplos, el dispositivo de destino puede usar solicitudes GET parciales de HTTP para recuperar rangos de bytes específicos de los segmentos 628 o 632.

**[0185]** La Figura 7 es un diagrama de bloques que ilustra elementos de un archivo de video de ejemplo 750. Puede decirse que el archivo de video 750 incluye un

segmento. Como se describió anteriormente, los archivos de video de acuerdo con el formato de archivo de medios base ISO y extensiones de este almacenan datos en una serie de objetos, denominados "cuadros". En el ejemplo de la Figura 7, el archivo de video 750 incluye un cuadro de tipo de archivo (FTYP) 752, un cuadro de película (MOOV) 754, cuadros de índice de segmento (sidx) 762, cuadros de fragmento de película (MOOF) 764, y un cuadro de acceso aleatorio a un fragmento de película (MFRA) 766. Si bien la Figura 7 representa un ejemplo de un archivo de video, debe entenderse que otros archivos de medios puede incluir otros tipos de datos de medios (por ejemplo, datos de audio, datos de texto temporizados, o similares) que están estructurados de manera similar a los datos del archivo de video 750, de acuerdo con el formato de archivo de medios base ISO y sus extensiones.

**[0186]** El cuadro FTYP 752 generalmente describe un tipo de archivo para el archivo de video 750. El cuadro de tipo de archivo 752 puede incluir datos que identifican una especificación que describen un mejor uso para el archivo de video 750. El cuadro de tipo de archivo 752 puede colocarse alternativamente antes del cuadro MOOV 754, cuadros de fragmento de película 764, y/o cuadro MFRA 766.

**[0187]** En algunos ejemplos, un segmento, tal como el archivo de video 750, puede incluir un cuadro de actualización de MPD (no se muestra) antes del cuadro FTYP 752. El cuadro de actualización de MPD puede incluir información que indica que se va a actualizar una MPD correspondiente a una representación que incluye el archivo de video 750, junto con información para actualizar la MPD. Por ejemplo, el cuadro de actualización de MPD puede proporcionar un URI o URL para que un recurso se use para actualizar la MPD. Como otro ejemplo, el cuadro de actualización de MPD puede incluir datos para actualizar la MPD. En algunos ejemplos, el cuadro de actualización de MPD puede seguir inmediatamente a un cuadro de tipo de segmento (STYP) (no se muestra) del archivo de video 750, donde el cuadro STYP puede definir un tipo de segmento para el archivo de video 750.

**[0188]** El cuadro MOOV 754, en el ejemplo de la Figura 7, incluye un cuadro de encabezado de película (MVHD) 756, un cuadro de pista (TRAK) 758, y uno o más cuadros de extensión de película (MVEX) 760. En general, el cuadro MVHD 756 puede describir características generales del archivo de video 750. Por ejemplo, el cuadro MVHD 756 puede incluir datos que describen cuándo se creó originalmente el archivo de video 750, cuándo se modificó por última vez el archivo de video 750, una escala de tiempo para el archivo de video 750, la duración de la reproducción para el archivo de video 750, u otros datos que generalmente describen el archivo de video 750.

**[0189]** El cuadro TRAK 758 puede incluir datos de una pista del archivo de video 750. El cuadro TRAK 758 puede incluir un cuadro de encabezado de pista (TKHD) que describe características de la pista correspondiente al cuadro TRAK 758. En algunos ejemplos, el cuadro TRAK 758 puede incluir imágenes de video codificadas, mientras que en otros ejemplos, las imágenes de video codificadas de la pista pueden incluirse en fragmentos de película 764, a los que se puede hacer referencia mediante datos del cuadro TRAK 758 y/o cuadros sidx 762.

**[0190]** En algunos ejemplos, el archivo de video 750 puede incluir más de una pista. En consecuencia, el cuadro MOOV 754 puede incluir una cantidad de cuadros TRAK igual a la cantidad de pistas en el archivo de video 750. El cuadro TRAK 758 puede describir características de una pista correspondiente del archivo de video 750. Por ejemplo, el cuadro TRAK 758 puede describir información temporal y/o espacial de la pista correspondiente. Un cuadro TRAK similar al cuadro TRAK 758 del cuadro MOOV 754 puede describir características de una pista de conjunto de parámetros, cuando una pista de conjunto de parámetros se incluye en un archivo de video, tal como el archivo de video 750. La presencia de mensajes SEI de nivel de secuencia puede señalarse en la pista de conjunto de parámetros dentro del cuadro TRAK que describe la pista de conjunto de parámetros.

**[0191]** Los cuadros MVEX 760 pueden describir características de correspondientes fragmentos de película 764, por ejemplo, para señalar que el archivo de video 750 incluye fragmentos de película 764, además de datos de video incluidos dentro del cuadro MOOV 754, si los hubiere. En el contexto de transmisión de datos de video, las imágenes de video codificadas pueden incluirse en fragmentos de película 764 en lugar de en el cuadro MOOV 754. En consecuencia, todas las muestras de video codificado pueden incluirse en fragmentos de película 764, en lugar de en el cuadro MOOV 754.

**[0192]** El cuadro MOOV 754 puede incluir una cantidad de cuadros MVEX 760 igual a la cantidad de fragmentos de película 764 en el archivo de video 750. Cada uno de los cuadros MVEX 760 puede describir características de uno correspondiente de uno de los fragmentos de película 764. Por ejemplo, cada cuadro MVEX puede incluir una película que extiende el cuadro de encabezado (MEHD) cuadro que describe una duración temporal para el correspondiente uno de los fragmentos de película 764.

**[0193]** Un conjunto de datos de secuencia puede almacenarse en una muestra de video que no incluye datos de video codificados reales. Una muestra de video generalmente puede corresponder a una unidad de acceso, que es una representación

de una imagen codificada en una instancia de tiempo específica. En el contexto de AVC, la imagen codificada incluye una o más unidades de NAL de VCL que contienen la información para construir todos los píxeles de la unidad de acceso y otras unidades de NAL no VCL asociadas, tales como mensajes SEI. En consecuencia, un conjunto de datos de secuencia, que puede incluir mensajes SEI de nivel de secuencia, puede incluirse en uno de los fragmentos de película 764. La presencia de un conjunto de datos de secuencia y/o mensajes SEI de nivel de secuencia pueden señalizarse como presentes en uno de los fragmentos de película 764 dentro del uno de los cuadros MVEX 760 correspondiente al uno de los fragmentos de película 764.

**[0194]** Los cuadros SIDX 762 son elementos opcionales del archivo de video 750. Es decir, los archivos de video conforme al formato de archivo 3GPP, u otros formatos de archivo similares, no incluyen necesariamente los cuadros sidx 762. De acuerdo con el ejemplo del formato de archivo 3GPP, se puede usar un cuadro SIDX para identificar un subsegmento de un segmento (por ejemplo, un segmento contenido dentro del archivo de video 750). El formato de archivo 3GPP define un subsegmento como “un conjunto autónomo de uno o más cuadros de fragmentos de película consecutivos con los correspondientes cuadros de datos de medios y un cuadro de datos de medios que contiene datos a los que hace referencia un cuadro de fragmento de película debe seguir a ese cuadro de fragmento de película y preceder al siguiente cuadro de fragmento de película que contiene información sobre la misma pista”. El formato de archivo 3GPP también indica que un cuadro sidx “contiene una secuencia de referencias a subsegmentos del (sub)segmento documentado por el cuadro. Los subsegmentos a los que se hace referencia son contiguos en el tiempo de presentación. De manera similar, los bytes a los que hace referencia un cuadro de índice de segmento siempre son contiguos dentro del segmento. El tamaño al que se hace referencia da la cuenta de la cantidad de bytes en el material al que se hace referencia”.

**[0195]** Los cuadros SIDX 762 generalmente proporcionan información representativa de uno o más subsegmentos de un segmento incluido en el archivo de video 750. Por ejemplo, tal información puede incluir tiempos de reproducción en los que comienzan y/o finalizan los subsegmentos, desplazamientos de bytes para los subsegmentos, si los subsegmentos incluyen (por ejemplo, comienzan con) un punto de acceso de flujo (SAP), un tipo para el SAP (por ejemplo, si el SAP es una imagen de actualización instantánea del decodificador, una imagen de acceso aleatorio limpio (CRA), una imagen de acceso de enlace roto (BLA), o similar), una posición del SAP (en términos

de tiempo de reproducción y/o desplazamiento de bytes) en el subsegmento, y similares.

**[0196]** Los fragmentos de película 764 pueden incluir una o más imágenes de video codificadas. En algunos ejemplos, los fragmentos de película 764 pueden incluir uno o más grupos de imágenes (GOP), cada uno de los cuales puede incluir varias imágenes de video codificadas, por ejemplo, fotogramas o imágenes. Además, como se describió anteriormente, los fragmentos de película 764 pueden incluir conjuntos de datos de secuencia en algunos ejemplos. Cada uno de fragmentos de película 764 puede incluir un cuadro de encabezado de fragmento de película (MFHD, no se muestra en la Figura 7). El cuadro MFHD puede describir características del correspondiente fragmento de película, tal como un número de secuencia para el fragmento de película. Los fragmentos de película 764 pueden incluirse en el orden del número de secuencia en el archivo de video 750.

**[0197]** El cuadro MFRA 766 puede describir puntos de acceso aleatorio dentro de fragmentos de película 764 del archivo de video 750. Esto puede ayudar a llevar a cabo controles de movimiento, tal como llevar a cabo búsquedas en ubicaciones temporales particulares (es decir, tiempos de reproducción) dentro de un segmento comprendido por el archivo de video 750. El cuadro MFRA 766 generalmente es opcional y no es necesario incluirlo en los archivos de video, en algunos ejemplos. Asimismo, no es necesario que un dispositivo cliente haga referencia al cuadro MFRA 766 para decodificar y mostrar correctamente los datos de video del archivo de video 750. El cuadro MFRA 766 puede incluir una cantidad de cuadros de acceso aleatorio de fragmento de pista (TFRA) (no se muestra) igual a la cantidad de pistas del archivo de video 750, o en algunos ejemplos, igual a la cantidad de pistas de medios (por ejemplo, pistas no de sugerencias) del archivo de video 750.

**[0198]** En algunos ejemplos, los fragmentos de película 764 pueden incluir uno o más puntos de acceso de flujo (SAP), tal como imágenes IDR. Asimismo, el cuadro MFRA 766 puede proporcionar indicaciones de ubicaciones dentro del archivo de video 750 de los SAP. En consecuencia, una subsecuencia temporal del archivo de video 750 puede formarse a partir de los SAP del archivo de video 750. La subsecuencia temporal también puede incluir otras imágenes, como fotogramas P y/o fotogramas B que dependen de los SAP. Los fotogramas y/o porciones de la subsecuencia temporal pueden disponerse dentro los segmentos de manera tal que fotogramas/porciones de la subsecuencia temporal que dependen de otros fotogramas/porciones de la subsecuencia puedan decodificarse adecuadamente. Por ejemplo, en la disposición

jerárquica de los datos, los datos utilizados para la predicción de otros datos también pueden incluirse en la subsecuencia temporal.

**[0199]** El archivo de video 750 también contiene un cuadro de descripción de muestra 768, en este ejemplo. En particular, el cuadro de descripción de muestra 768 está incluido dentro del cuadro TRAK 758, en este ejemplo. En el ejemplo de la Figura 7, el cuadro de descripción de muestra 768 no incluye datos de video codificados de acuerdo con el códec.

**[0200]** Un cuadro de descripción de muestra 768 de ejemplo puede definirse de la siguiente manera:

Entrada de muestra y tipos de cuadro: 'hvc2', 'hev2', 'lhv1', 'lhe1', 'lhvC'

- Contenedor: Cuadro de descripción de muestra ('std')
- Obligatorio: Una entrada de muestra 'hvc1', 'hev1', 'hvc2', 'hev2', 'lhv1' o 'lhe1' es obligatoria
- Cantidad: Puede haber una o más entradas de muestra

**[0201]** En esta definición de ejemplo para el cuadro de descripción de muestra 768, cuando el nombre de la entrada de muestra es 'lhv1', el valor predeterminado y obligatorio de array\_completeness es 4 para matrices de todos los tipos de conjuntos de parámetros y 0 para todas las otras matrices. Cuando el nombre de la entrada de muestra es 'lhe1', el valor predeterminado de array\_completeness es 0 para todas las matrices.

**[0202]** La Figura 8 es un diagrama de flujo que ilustra un método de ejemplo para procesar datos de video de acuerdo con las técnicas de esta divulgación. Si bien se describe con respecto al dispositivo de destino 116 de la Figura 1, debe entenderse que se pueden configurar otros dispositivos para llevar a cabo un método similar al de la Figura 8.

**[0203]** El dispositivo de destino 116 puede configurarse para recibir un registro de configuración para decodificar un flujo de bits de los datos de video (800). El registro de configuración para el flujo de bits puede incluir un elemento de sintaxis de indicación de conjunto de herramientas que incluye información que identifica, a partir de un conjunto de herramientas de decodificación de video, herramientas que se requieren para decodificar el flujo de bits asociado con el registro de configuración. En función del elemento de sintaxis de indicación de conjunto de herramientas en el registro de configuración, el dispositivo de destino 116 puede determinar si recuperar el flujo de bits asociado con el registro de configuración (802). Luego, el dispositivo de destino 116 puede recuperar el flujo de bits (804). El dispositivo de destino 116 puede

enviar el flujo de bits a un decodificador de video para su decodificación (806). El dispositivo de destino 116, de manera adicional o alternativa, puede recibir un parámetro de tipo MIME que incluye un par de valores clave, en donde la clave indica que el tipo de MIME identifica las herramientas de decodificación de video, y el valor identifica, a partir de las herramientas de decodificación de video, las herramientas que se requieren para decodificar el flujo de bits. El flujo de bits de datos de video puede incluir uno o más conjuntos de parámetros, y cada una de las herramientas identificadas en el registro de configuración como requeridas para decodificar el flujo de bits se pueden habilitar en al menos uno del uno o más conjuntos de parámetros.

**[0204]** El elemento de sintaxis de indicación de conjunto de herramientas se señala en un cuadro de configuración de información de formato de archivo. El elemento de sintaxis de indicación de conjunto de herramientas que incluye información que identifica las herramientas que se requieren para decodificar el flujo de bits asociado con el registro de configuración puede identificar todas las herramientas que se requieren para decodificar el flujo de bits asociado con el registro de configuración. El elemento de sintaxis de indicación de conjunto de herramientas puede ser un número entero de 32 bits sin signo, en donde cada bit del número entero de 32 bits sin signo corresponde a una herramienta única para decodificar el flujo de bits.

**[0205]** El registro de configuración se puede formatear de acuerdo con el estándar EVC esencial. El registro de configuración para el flujo de bits puede incluir un elemento de sintaxis de perfil y/o un elemento de sintaxis de nivel antes del elemento de sintaxis de indicación de conjunto de herramientas. El registro de configuración para el flujo de bits puede incluir un elemento de sintaxis de formato de croma después del elemento de sintaxis de indicación de conjunto de herramientas.

**[0206]** Cabe aclarar que según el ejemplo, algunos actos o eventos de cualquiera de las técnicas descritas en la presente se pueden realizar en una secuencia diferente, se pueden agregar, fusionar o excluir completamente (por ejemplo, no todos los actos o eventos descritos son necesarios para la práctica de las técnicas). Además, en algunos ejemplos, los actos o eventos se pueden llevar a cabo concurrentemente, por ejemplo, a través de procesamiento multihilo, procesamiento interrumpido o procesadores múltiples, en lugar de secuencialmente.

**[0207]** Las siguientes cláusulas representan implementaciones de ejemplo de los sistemas y técnicas descritos anteriormente.

**[0208]** Cláusula 1. Un método para procesar un flujo de video o un archivo de video, en donde el método comprende: recibir un registro de configuración para decodificar

un flujo de bits de datos de video, en donde el registro de configuración para el flujo de bits incluye un elemento de sintaxis de indicación de conjunto de herramientas que incluye información que identifica, a partir de un conjunto de herramientas de decodificación de video, herramientas que se requieren para decodificar el flujo de bits asociado con el registro de configuración; en función del elemento de sintaxis de indicación de conjunto de herramientas, determinar si recuperar el flujo de bits asociado con el registro de configuración; y en función de la determinación de recuperar el flujo de bits asociado con el registro de configuración, recuperar el flujo de bits y enviar el flujo de bits a un decodificador de video para su decodificación.

**[0209]** Cláusula 2. El método de acuerdo con la cláusula 1, que comprende además: recibir un parámetro de tipo de extensiones de correo de Internet multipropósito (MIME) que incluye un par de valores clave, en donde la clave indica que el tipo de MIME identifica las herramientas de decodificación de video, y el valor identifica las herramientas, a partir de las herramientas de decodificación de video, que se requieren para decodificar el flujo de bits.

**[0210]** Cláusula 3. El método de acuerdo con la cláusula 1 o 2, en donde el flujo de bits de datos de video comprende uno o más conjuntos de parámetros, y en donde cada una de las herramientas identificadas en el registro de configuración como requeridas para decodificar el flujo de bits se habilitan en al menos uno del uno o más conjuntos de parámetros.

**[0211]** Cláusula 4. El método de acuerdo con cualquiera de las cláusulas 1-3, en donde el elemento de sintaxis de indicación de conjunto de herramientas que incluye información que identifica las herramientas que se requieren para decodificar el flujo de bits asociado con el registro de configuración identifica todas las herramientas que se requieren para decodificar el flujo de bits asociado con el registro de configuración.

**[0212]** Cláusula 5. El método de acuerdo con cualquiera de las cláusulas 1-4, en donde el registro de configuración se formatea de acuerdo con el estándar de codificación de video esencial (EVC).

**[0213]** Cláusula 6. El método de acuerdo con cualquiera de las cláusulas 1-5, en donde el elemento de sintaxis de indicación de conjunto de herramientas se señala en un cuadro de configuración de información de formato de archivo.

**[0214]** Cláusula 7. El método de acuerdo con cualquiera de las cláusulas 1-6, en donde el elemento de sintaxis de indicación de conjunto de herramientas comprende un número entero de 32 bits sin signo.

**[0215]** Cláusula 8. El método de acuerdo con la cláusula 7, en donde los bits del

número entero de 32 bits sin signo corresponden cada uno a una herramienta única para decodificar el flujo de bits.

**[0216]** Cláusula 9. El método de acuerdo con cualquiera de las cláusulas 1-8, en donde el registro de configuración para el flujo de bits incluye un elemento de sintaxis de perfil antes del elemento de sintaxis de indicación de conjunto de herramientas.

**[0217]** Cláusula 10. El método de acuerdo con cualquiera de las cláusulas 1-9, en donde el registro de configuración para el flujo de bits incluye un elemento de sintaxis de nivel antes del elemento de sintaxis de indicación de conjunto de herramientas.

**[0218]** Cláusula 11. El método de acuerdo con cualquiera de las cláusulas 1-10, en donde el registro de configuración para el flujo de bits incluye un elemento de sintaxis de formato de croma después del elemento de sintaxis de indicación de conjunto de herramientas.

**[0219]** Cláusula 12. El método de acuerdo con cualquiera de las cláusulas 1-11, en donde el registro de configuración comprende datos de nivel de formato de archivo que incluyen datos de medios codificados a nivel de capa de codificación de video (VCL) en el flujo de bits de los datos de video, en donde los datos de medios codificados a nivel de VCL incluyen uno o más conjuntos de parámetros, y en donde cada una de las herramientas identificadas en el registro de configuración como requeridas para decodificar el flujo de bits se habilitan en al menos uno del uno o más conjuntos de parámetros.

**[0220]** Cláusula 13. Un dispositivo para procesar datos de video, en donde el dispositivo comprende: una memoria; y uno o más procesadores implementados en la circuitería, acoplados a la memoria y configurados para: recibir un registro de configuración para decodificar un flujo de bits de datos de video, en donde el registro de configuración para el flujo de bits incluye un elemento de sintaxis de indicación de conjunto de herramientas que incluye información que identifica, a partir de un conjunto de herramientas de decodificación de video, herramientas que se requieren para decodificar el flujo de bits asociado con el registro de configuración; en función del elemento de sintaxis de indicación de conjunto de herramientas, determinar si recuperar el flujo de bits asociado con el registro de configuración; y en función de la determinación de recuperar el flujo de bits asociado con el registro de configuración, recuperar el flujo de bits y enviar el flujo de bits a un decodificador de video para su decodificación.

**[0221]** Cláusula 14. El dispositivo de acuerdo con la cláusula 13, en donde el uno o más procesadores además están configurados para: recibir un parámetro de tipo de

extensiones de correo de Internet multipropósito (MIME) que incluye un par de valores clave, en donde la clave indica que el tipo de MIME identifica las herramientas de decodificación de video, y el valor identifica las herramientas, a partir de las herramientas de decodificación de video, que se requieren para decodificar el flujo de bits.

**[0222]** Cláusula 15. El dispositivo de acuerdo con la cláusula 13 o 14, en donde el flujo de bits de datos de video comprende uno o más conjuntos de parámetros, y en donde cada una de las herramientas identificadas en el registro de configuración como requeridas para decodificar el flujo de bits se habilitan en al menos uno del uno o más conjuntos de parámetros.

**[0223]** Cláusula 16. El dispositivo de acuerdo con cualquiera de las cláusulas 13-15, en donde el elemento de sintaxis de indicación de conjunto de herramientas que incluye información que identifica las herramientas que se requieren para decodificar el flujo de bits asociado con el registro de configuración identifica todas las herramientas que se requieren para decodificar el flujo de bits asociado con el registro de configuración.

**[0224]** Cláusula 17. El dispositivo de acuerdo con cualquiera de las cláusulas 13-16, en donde el registro de configuración se formatea de acuerdo con el estándar de codificación de video esencial (EVC).

**[0225]** Cláusula 18. El dispositivo de acuerdo con cualquiera de las cláusulas 13-117, en donde el elemento de sintaxis de indicación de conjunto de herramientas se señala en un cuadro de configuración de información de formato de archivo.

**[0226]** Cláusula 19. El dispositivo de acuerdo con cualquiera de las cláusulas 13-15, en donde el elemento de sintaxis de indicación de conjunto de herramientas comprende un número entero de 32 bits sin signo.

**[0227]** Cláusula 20. El dispositivo de acuerdo con la cláusula 19, en donde los bits del número entero de 32 bits sin signo corresponden cada uno a una herramienta única para decodificar el flujo de bits.

**[0228]** Cláusula 21. El dispositivo de acuerdo con cualquiera de las cláusulas 13-20, en donde el registro de configuración para el flujo de bits incluye un elemento de sintaxis de perfil antes del elemento de sintaxis de indicación de conjunto de herramientas.

**[0229]** Cláusula 22. El dispositivo de acuerdo con cualquiera de las cláusulas 13-21, en donde el registro de configuración para el flujo de bits incluye un elemento de sintaxis de nivel antes del elemento de sintaxis de indicación de conjunto de

herramientas.

**[0230]** Cláusula 23. El dispositivo de acuerdo con cualquiera de las cláusulas 13-22, en donde el registro de configuración para el flujo de bits incluye un elemento de sintaxis de formato de croma después del elemento de sintaxis de indicación de conjunto de herramientas.

**[0231]** Cláusula 24. El dispositivo de acuerdo con cualquiera de las cláusulas 13-23, en donde el registro de configuración comprende datos de nivel de formato de archivo que incluyen datos de medios codificados a nivel de capa de codificación de video (VCL) en el flujo de bits de los datos de video, en donde los datos de medios codificados a nivel de VCL incluyen uno o más conjuntos de parámetros, y en donde cada una de las herramientas identificadas en el registro de configuración como requeridas para decodificar el flujo de bits se habilitan en al menos uno del uno o más conjuntos de parámetros.

**[0232]** Cláusula 25. Un medio de almacenamiento legible por computadora almacena instrucciones que cuando se ejecutan mediante uno o más procesadores hacen a los uno o más procesadores recibir un registro de configuración para decodificar un flujo de bits de los datos de video, en donde el registro de configuración para el flujo de bits incluye un elemento de sintaxis de indicación de conjunto de herramientas que incluye información que identifica, a partir de un conjunto de herramientas de decodificación de video, las herramientas que se requieren para decodificar el flujo de bits asociado con el registro de configuración; en función del elemento de sintaxis de indicación del conjunto de herramientas, determinar si recuperar el flujo de bits asociado con el registro de configuración; y en función de la determinación de recuperar el flujo de bits asociado con el registro de configuración, recuperar el flujo de bits y enviar el flujo de bits a un decodificador de video para su decodificación.

**[0233]** Cláusula 26. El medio de almacenamiento legible por computadora de acuerdo con la cláusula 25, en donde las instrucciones hacen al uno o más procesadores: recibir un parámetro de tipo de extensiones de correo de Internet multipropósito (MIME) que incluye un par de valores clave, en donde la clave indica que el tipo de MIME identifica las herramientas de decodificación de video, y el valor identifica las herramientas, a partir de las herramientas de decodificación de video, que se requieren para decodificar el flujo de bits.

**[0234]** Cláusula 27. El medio de almacenamiento legible por computadora de acuerdo con la cláusula 25 o 26, en donde el flujo de bits de datos de video comprende uno o más conjuntos de parámetros, y en donde cada una de las herramientas identificadas

en el registro de configuración como requeridas para decodificar el flujo de bits se habilitan en al menos uno del uno o más conjuntos de parámetros.

**[0235]** Cláusula 28. El medio de almacenamiento legible por computadora de acuerdo con cualquiera de las cláusulas 25-27, en donde el elemento de sintaxis de indicación de conjunto de herramientas que incluye información que identifica las herramientas que se requieren para decodificar el flujo de bits asociado con el registro de configuración identifica todas las herramientas que se requieren para decodificar el flujo de bits asociado con el registro de configuración.

**[0236]** Cláusula 29. El medio de almacenamiento legible por computadora de acuerdo con cualquiera de las cláusulas 25-28, en donde el registro de configuración comprende datos de nivel de formato de archivo que incluyen datos de medios codificados a nivel de capa de codificación de video (VCL) en el flujo de bits de los datos de video, en donde los datos de medios codificados a nivel de VCL incluyen uno o más conjuntos de parámetros, y en donde cada una de las herramientas identificadas en el registro de configuración como requeridas para decodificar el flujo de bits se habilitan en al menos uno del uno o más conjuntos de parámetros.

**[0237]** Cláusula 30. Un aparato para procesar un flujo de video o un archivo de video, en donde el aparato comprende: medios para recibir un registro de configuración para decodificar un flujo de bits de los datos de video, en donde el registro de configuración para el flujo de bits incluye un elemento de sintaxis de indicación de conjunto de herramientas que incluye información que identifica, a partir de un conjunto de herramientas de decodificación de video, herramientas que se requieren para decodificar el flujo de bits asociado con el registro de configuración; medios para determinar si recuperar el flujo de bits asociado con el registro de configuración en función del elemento de sintaxis de indicación de conjunto de herramientas; medios para recuperar el flujo de bits en función de la determinación de recuperar el flujo de bits asociado con el registro de configuración; y medios para enviar el flujo de bits a un decodificador de video para su decodificación.

**[0238]** Cláusula 31. Un método para procesar datos de video, en donde el método comprende: recibir un registro de configuración para decodificar un flujo de bits de datos de video de acuerdo con el estándar de codificación de video esencial (EVC), en donde el registro de configuración para el flujo de bits incluye un elemento de sintaxis de indicación de conjunto de herramientas que incluye información de todas las herramientas que se requieren para decodificar el flujo de bits asociado con el registro de configuración; en función del elemento de sintaxis de indicación de conjunto de

herramientas, determinar si recuperar el flujo de bits asociado con el registro de configuración; y en función de la determinación de recuperar el flujo de bits asociado con el registro de configuración, recuperar el flujo de bits y enviar el flujo de bits a un decodificador de video para su decodificación.

**[0239]** Cláusula 32. Un método para procesar datos de video, en donde el método comprende: determinar un registro de configuración para un flujo de bits de los datos de video generados por un codificador de video de acuerdo con el estándar de codificación de video esencial (EVC), en donde el registro de configuración para el flujo de bits incluye información para decodificar el flujo de bits, y en donde el registro de configuración incluye un elemento de sintaxis de indicación de conjunto de herramientas que incluye información de todas las herramientas que se requieren para decodificar el flujo de bits asociado con el registro de configuración; y señalizar el registro de configuración.

**[0240]** Cláusula 33. El método de acuerdo con la cláusula 31 o 32 en donde el elemento de sintaxis de indicación de conjunto de herramientas se señala en un cuadro de configuración de información de formato de archivo.

**[0241]** Cláusula 34. Un dispositivo para procesar datos de video, en donde el dispositivo comprende: circuitería de procesamiento configurada para llevar a cabo el método de acuerdo con cualquiera o una combinación de las cláusulas 1–3.

**[0242]** Cláusula 35. El dispositivo de acuerdo con la cláusula 34, que comprende además una memoria para almacenar los datos de video.

**[0243]** Cláusula 36. El dispositivo de acuerdo con cualquiera de las cláusulas 34 y 35, que comprende además una pantalla configurada para mostrar datos de video decodificados.

**[0244]** Cláusula 37. El dispositivo de acuerdo con cualquiera de las cláusulas 34–36, en donde el dispositivo comprende uno o más de una cámara, una computadora, un dispositivo móvil, un dispositivo receptor de transmisión o un decodificador.

**[0245]** Cláusula 38. Un medio de almacenamiento legible por computadora que tiene almacenadas instrucciones que, cuando se ejecutan, hacen que uno o más procesadores lleven a cabo el método de acuerdo con cualquiera de las cláusulas 31–33.

**[0246]** Cláusula 39. Un dispositivo para procesar datos de video, en donde el dispositivo comprende medios para llevar a cabo el método de acuerdo con cualquiera de las cláusulas 31–33.

**[0247]** En uno o más ejemplos, las funciones descritas pueden implementarse en

*hardware*, *software*, *firmware* o cualquier combinación de estos. Si se implementan en *software*, las funciones pueden almacenarse o transmitirse como una o más instrucciones o código en un medio legible por computadora y pueden ejecutarse mediante una unidad de procesamiento de *hardware*. Los medios legibles por computadora pueden incluir medios de almacenamiento legibles por computadora, que corresponden a un medio tangible tal como un medio de almacenamiento de datos, o medios de comunicación, incluso cualquier medio que facilite la transferencia de un programa informático de un lugar a otro, por ejemplo, de acuerdo con un protocolo de comunicación. De esta manera, los medios legibles por computadora generalmente pueden corresponder a (1) medios de almacenamiento legibles por computadora tangibles no transitorios o (2) un medio de comunicación tal como una onda de señal o portadora. Los medios de almacenamiento de datos pueden ser cualquier medio disponible al que se puede acceder mediante una o más computadoras o uno o más procesadores para obtener instrucciones, códigos 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 computadora.

**[0248]** A modo de ejemplo, y no de limitación, tal medio de almacenamiento legible por computadora puede comprender RAM, ROM, EEPROM, CD-ROM u otros dispositivos de almacenamiento en disco óptico, de almacenamiento en disco magnético u otro almacenamiento magnético, o cualquier otro medio que pueda utilizarse para almacenar el código de programación deseado en forma de instrucciones o estructuras de datos y al que pueda accederse mediante una computadora. Además, cualquier conexión se denomina de manera adecuada medio legible por computadora. Por ejemplo, si las instrucciones se transmiten desde un sitio web, servidor u otra fuente remota mediante un cable coaxial, cable de fibra óptica, par trenzado, línea de abonado digital (DSL) o tecnologías inalámbricas tales como infrarroja, radio y microondas, entonces el cable coaxial, cable de fibra óptica, par trenzado, DSL o tecnologías inalámbricas, tales como infrarroja, radio y microondas, están incluidos en la definición de medio. Sin embargo, cabe aclarar que los medios de almacenamiento legibles por computadora y los medios de almacenamiento de datos no incluyen conexiones, ondas portadoras, señales u otros medios transitorios, sino que, en su lugar, se refieren a medios de almacenamiento tangibles no transitorios. Como se usa en la presente, disco incluye disco compacto (CD), disco láser, disco óptico, disco versátil digital (DVD), disquete y disco Blu-ray, en donde los discos reproducen datos magnéticamente y datos ópticamente con láseres. También

deberían incluirse combinaciones de los anteriores dentro del alcance de medios legibles por computadora.

**[0249]** Las instrucciones pueden ejecutar mediante uno o más procesadores, tales como uno o más procesadores de señal digital (DSP), procesadores de propósito general, circuitos integrados de aplicación específica (ASIC), matrices de puertas programables por campo (FPGA) u otro tipo de circuitería lógica integrada o separada equivalente. En consecuencia, las expresiones "procesador" y "circuitería de procesamiento", como se usa en la presente, pueden referirse a cualquiera de las estructuras anteriores o cualquier otra estructura adecuada para la aplicación de las técnicas descritas en la presente. Además, en algunos aspectos, la funcionalidad descrita en la presente puede proporcionarse dentro de los módulos de *hardware* y/o *software* dedicados configurados para codificar y decodificar, o incorporarse en un códec combinado. Además, las técnicas podrían implementarse totalmente en uno o más circuitos o elementos lógicos.

**[0250]** Las técnicas de esta divulgación se pueden implementar en una amplia variedad de dispositivos o aparatos, que incluyen un microteléfono inalámbrico, un circuito integrado (IC) o un conjunto de IC (por ejemplo, un conjunto de chips). En esta divulgación, se describen distintos componentes, módulos o unidades para enfatizar los aspectos funcionales de los dispositivos configurados para realizar las técnicas divulgadas, pero no es necesario que sean realizadas por diferentes unidades de *hardware*. Más bien, como se describió anteriormente, las diversas unidades pueden combinarse en una unidad de *hardware* de códec o proporcionarse mediante una colección de unidades de *hardware* interoperativas, que incluye uno o más procesadores como se describió anteriormente, en conjunción con el *software* y/o *firmware* adecuado.

**[0251]** Se han descrito diversos ejemplos. Estos y otros ejemplos se encuentran dentro del alcance de las siguientes reivindicaciones.