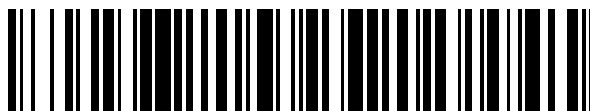


19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 898 452**

51 Int. Cl.:

H04N 19/70

(2014.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

86 Fecha de presentación y número de la solicitud internacional: **23.12.2013 PCT/US2013/077587**

87 Fecha y número de publicación internacional: **10.07.2014 WO14107377**

96 Fecha de presentación y número de la solicitud europea: **23.12.2013 E 13826913 (9)**

97 Fecha y número de publicación de la concesión europea: **20.10.2021 EP 2941889**

54 Título: **Señalización de la resolución espacial de las vistas de profundidad en el formato de archivo de codificación de múltiples vistas**

30 Prioridad:

04.01.2013 US 201361749089 P
20.12.2013 US 201314137358

45 Fecha de publicación y mención en BOPI de la traducción de la patente:
07.03.2022

73 Titular/es:

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

72 Inventor/es:

WANG, YE-KUI y
CHEN, YING

74 Agente/Representante:

ISERN JARA, Jorge

ES 2 898 452 T3

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

DESCRIPCIÓN

Señalización de la resolución espacial de las vistas de profundidad en el formato de archivo de codificación de múltiples vistas

REIVINDICACIÓN DE PRIORIDAD

Esta solicitud reivindica la prioridad de la solicitud provisional estadounidense n.º 61/749.089, presentada el 4 de enero de 2013.

CAMPO TÉCNICO

La presente divulgación se refiere a la codificación de vídeo y al almacenamiento de contenido de vídeo y, más en particular, a las técnicas para almacenar flujos de vídeo que contienen múltiples vistas codificadas.

ANTECEDENTES

Las capacidades de vídeo digital se pueden incorporar en una amplia gama de dispositivos, incluidos televisores digitales, sistemas de transmisión directa digital, sistemas de transmisión inalámbrica, asistentes digitales personales (PDA), ordenadores portátiles o de sobremesa, tabletas, lectores de libros electrónicos, cámaras digitales, dispositivos de grabación digital, reproductores de medios digitales, dispositivos de videojuegos, consolas de videojuegos, radiotelefonos móviles o por satélite, los denominados "teléfonos inteligentes", dispositivos de vídeo teleconferencia, dispositivos de transmisión de vídeo, transcodificadores, enrutadores u otros dispositivos de red, y similares. Los dispositivos de vídeo digital implementan técnicas de comprensión de vídeo, como las descritas en los estándares definidos por MPEG-2, MPEG-4, ITU-T H.263, ITU-T H.264/MPEG-4, Parte 10, Codificación de vídeo avanzada (AVC), el estándar de Codificación de vídeo de alta eficiencia (HEVC) actualmente en desarrollo, estándares de propiedad, formatos de compresión de vídeo abiertos como VP8, y extensiones de dichas normas, técnicas o formatos. Los dispositivos de vídeo pueden transmitir, recibir, codificar, decodificar y/o almacenar información de vídeo digital de manera más eficiente implementando tales técnicas de compresión de vídeo.

Las técnicas de compresión de vídeo realizan predicciones espaciales (intraimagen) y/o predicciones temporales (interimagen) para reducir o eliminar la redundancia inherente a las secuencias de vídeo. Para la codificación de vídeo basada en bloques, una rebanada de vídeo (es decir, un fotograma de vídeo o una parte de un fotograma de vídeo) se puede dividir en bloques de vídeo, que también pueden denominarse macrobloques, bloques de árbol, unidades de codificación (CU) y/o nodos de codificación. Los bloques de vídeo pueden subdividirse, por ejemplo, en bloques de vídeo más pequeños o en particiones, y las particiones pueden subdividirse aún más. Por ejemplo, los macrobloques pueden subdividirse en bloques de vídeo más pequeños, con frecuencia denominados particiones, y algunas de las particiones pueden subdividirse en bloques de vídeo aún más pequeños.

Los bloques de vídeo codificados en un modo intra (I) pueden codificarse al usar la predicción espacial con respecto a las muestras de referencia en bloques vecinos en la misma imagen. Los bloques de vídeo codificados en modo inter [por ejemplo, en modo predictivo (P) o bipredictivo (B)] pueden utilizar la predicción espacial con respecto a las muestras de referencia en los bloques vecinos de la misma imagen o la predicción temporal con respecto a las muestras de referencia en otras imágenes de referencia. La predicción espacial o temporal da como resultado un bloque predictivo para la codificación de un bloque. Los datos residuales representan las diferencias de píxeles entre el bloque original a codificar y el bloque predictivo. Un bloque intercodificado se codifica según un vector de movimiento que apunta a un bloque de muestras de referencia que forman el bloque predictivo, y los datos residuales indican la diferencia entre el bloque codificado y el bloque predictivo. Un bloque intracodificado se codifica según un modo de intracodificación y los datos residuales. Para una compresión adicional, los datos residuales pueden transformarse del dominio de píxeles a un dominio de transformación, lo que resulta en coeficientes de transformada residuales, que luego pueden cuantificarse. Los coeficientes de transformación cuantificados, dispuestos inicialmente en una matriz bidimensional, se pueden escanear para producir un vector unidimensional de coeficientes de transformación, y se puede aplicar codificación entrópica para lograr una compresión aún mayor.

SÍNTESIS

La presente divulgación se refiere a diversas técnicas de almacenamiento de datos de vídeo. En particular, esta divulgación describe la técnica para almacenar flujos de vídeo que contienen múltiples vistas codificadas.

La invención se define en las reivindicaciones independientes adjuntas. Las características opcionales se definen en las reivindicaciones dependientes.

Los detalles de uno o más ejemplos se exponen en los dibujos adjuntos y en la siguiente descripción. Otras características, objetos y ventajas serán evidentes a partir de la descripción, dibujos y reivindicaciones.

Los ejemplos ilustrados en las Fig. 12 y 13 están dentro del alcance de la invención reivindicada. Otros ejemplos, cuando no están comprendidos en el ámbito de las reivindicaciones, no forman parte de la invención reivindicada, pero son útiles para comprenderla.

5 BREVE DESCRIPCIÓN DE LOS DIBUJOS

La FIG. 1 es un diagrama conceptual que ilustra la estructura de un ejemplo de archivo multimedia MVC+D que puede generarse según las técnicas descritas en esta divulgación.

10 La FIG. 2 ilustra un ejemplo de una unidad de acceso MVC+D que puede incluirse en un archivo multimedia generado según las técnicas descritas en esta divulgación.

La FIG. 3 es un diagrama esquemático que ilustra un ejemplo de sistema de codificación y decodificación de vídeo que puede utilizar las técnicas descritas en esta divulgación.

15 La FIG. 4 es un diagrama esquemático que ilustra un ejemplo de codificador de vídeo que puede implementar las técnicas descritas en esta divulgación.

20 La FIG. 5 es un diagrama esquemático que ilustra un ejemplo de decodificador de vídeo y módulo de desencapsulación que puede implementar las técnicas descritas en esta divulgación.

La FIG. 6 es un diagrama esquemático que ilustra los componentes de un módulo de encapsulación de ejemplo que puede implementar las técnicas descritas en esta divulgación.

25 La FIG. 7 es un diagrama esquemático que ilustra un ejemplo de conjunto de dispositivos que forman parte de una red que puede implementar las técnicas descritas en esta divulgación.

La FIG. 8 es un diagrama de flujo que ilustra un método de ejemplo para analizar un archivo multimedia que incluye contenido de vídeo de múltiples vistas que indica la presencia de vistas de textura y profundidad según las técnicas descritas en esta divulgación.

30 La FIG. 9 es un diagrama de flujo que ilustra un método de ejemplo para crear un archivo multimedia con contenido de vídeo de múltiples vistas para indicar la presencia de vistas de textura y profundidad según las técnicas descritas en esta divulgación.

35 La FIG. 10 es un diagrama de flujo que ilustra un ejemplo de método de análisis de un archivo multimedia que contiene vídeo de múltiples vistas que indica la dependencia de las vistas según las técnicas descritas en esta divulgación.

40 La FIG. 11 es un diagrama de flujo que ilustra un método de ejemplo para crear un archivo multimedia con contenido de vídeo de múltiples vistas para indicar la dependencia de las vistas según las técnicas descritas en esta divulgación.

La FIG. 12 es un diagrama de flujo que ilustra un método de ejemplo para analizar un archivo multimedia que contiene vídeo de múltiples vistas para determinar una resolución espacial de una vista de profundidad según las técnicas descritas en esta divulgación.

45 La FIG. 13 es un diagrama de flujo que ilustra un método de ejemplo para crear un archivo multimedia con contenido de vídeo de múltiples vistas para indicar una resolución espacial de una vista de profundidad según las técnicas descritas en esta divulgación.

50 La FIG. 14 es un diagrama de flujo que ilustra un método de ejemplo para analizar un archivo multimedia que contiene vídeo de múltiples vistas en el que una textura y una vista de profundidad de una vista particular se almacenan en pistas separadas según las técnicas descritas en esta divulgación.

55 La FIG. 15 es un diagrama de flujo que ilustra un método de ejemplo para crear un archivo multimedia con contenido de vídeo de múltiples vistas en el que una textura y una vista de profundidad de una vista particular se almacenan en pistas separadas según las técnicas descritas en esta divulgación.

DESCRIPCIÓN DETALLADA

60 En general, la presente divulgación se refiere a la encapsulación y desencapsulación de contenido de vídeo y al almacenamiento del contenido de vídeo codificado en un archivo. La presente divulgación describe diversas técnicas para almacenar contenido de vídeo, que incluyen, por ejemplo, el almacenamiento de contenido de vídeo codificado según el estándar de vídeo de HEVC en un archivo que se basa en un formato de archivo multimedia base de la Organización Internacional de Estandarización (ISO) (ISO/BMFF). En particular, esta divulgación describe técnicas para encapsular flujos de vídeo que contienen múltiples vistas codificadas en un archivo multimedia. En esta divulgación, el término "archivo multimedia" puede usarse indistintamente con "archivo de vídeo". En el presente

documento se describen diversos métodos para crear y analizar flujos de vídeo que contengan múltiples vistas codificadas, en las que cada vista puede contener una vista de textura solamente, una vista de profundidad solamente, o ambas vistas de textura y de profundidad, en un archivo basado en ISO/BMFF.

Ejemplos de estándares de codificación de vídeo incluyen ITU-T H.261, ISO/IEC MPEG-1 Visual, ITU-T H.262 o ISO/IEC MPEG-2 Visual, ITU-T H.263, ISO/IEC MPEG-4 Visual e ITU-T H.264 (también conocido como ISO/IEC MPEG-4 AVC), incluidas sus extensiones Codificación de vídeo escalable (SVC) y Codificación de vídeo de múltiples vistas (MVC). Además, existe un nuevo estándar de codificación de vídeo, a saber, codificación de vídeo de alta eficiencia (HEVC), desarrollado por el equipo de colaboración conjunta sobre codificación de vídeo (JCT-VC) del grupo de expertos en codificación de vídeo ITU-T (VCEG) y el grupo de expertos en imágenes en movimiento ISO/IEC (MPEG). Un reciente borrador de trabajo (WD) de HEVC se titula "Borrador 9 de la especificación de texto de la codificación de vídeo de alta eficiencia (HEVC)", (denominado en la presente "Borrador de trabajo 9 de HEVC" o "WD9 de HEVC"), se describe en el documento JCTVC-K1003_v11, Bross et al., Equipo de colaboración conjunta sobre codificación de vídeo (JCT-VC) de ITU-T SG16 WP3 e ISO/IEC JTC1/SC29/WG11, 11ª reunión, Shanghai, China, 10 al 19 de octubre de 2012, y está disponible en http://phenix.int-evry.fr/jct/doc_end_user/documents/11_Shanghai/wg11/JCTVC-K1003-v11.zip.

En JCT-3V y JCT-VC también se están desarrollando diversas extensiones de AVC y HEVC. En JCT-3V se están desarrollando dos extensiones de HEVC, la extensión para múltiples vistas (MV-HEVC) y la extensión para vídeo 3D (3D-HEVC). Además, se están desarrollando dos extensiones de AVC, la MVC+D y la 3D-AVC.

Las versiones recientes de los estándares vigentes son las siguientes. Documento JCT3V-B1001, titulado "Texto de estudio de la extensión de MVC ISO/IEC 14496-10:2012/DAM2 para la inclusión de mapas de profundidad", elaborado por M. Hannuksela (Nokia), Y. Chen (Qualcomm), G. Sullivan (Microsoft), T. Suzuki, S. Hattori (Sony), con fecha de octubre de 2012 y disponible en http://phenix.it-sudparis.eu/jct2/doc_end_user/documents/2_Shanghai/wg11/JCT3V-B1001-v1.zip. También se denomina MVC+D o codificación de vídeo tridimensional (3VC), y el vídeo codificado puede denominarse vídeo tridimensional (3DV). La codificación de vídeo tridimensional también puede denominarse codificación de vídeo tridimensional. Documento JCT3V-B1002, titulado "Texto borrador 4 de 3D-AVC", de M. M. Hannuksela, Y. Chen, T. Suzuki, J.-R. Ohm, G. J. Sullivan, y disponible en http://phenix.it-sudparis.eu/jct2/doc_end_user/documents/2_Shanghai/wg11/JCT3V-B1002-v1.zip. Documento "JCT3V-B1004", titulado "Texto borrador MV-HEVC 2", de G. Tech, K. Wegner, Y. Chen, M. Hannuksela, y disponible en http://phenix.it-sudparis.eu/jct2/doc_end_user/documents/2_Shanghai/wg11/JCT3V-B1004-v1.zip. Documento JCT3V-B1005, titulado "Modelo de prueba 3DHEVC 2", de G. Tech, K. Wegner, Y. Chen, S. Yea, y disponible en http://phenix.it-sudparis.eu/jct2/doc_end_user/documents/2_Shanghai/wg11/JCT3V-B1005-v1.zip.

La presente divulgación se refiere al almacenamiento de contenidos de vídeo en un archivo de datos electrónico. Para almacenar contenido de vídeo codificado según un estándar de vídeo particular, puede ser útil una especificación de formato de archivo correspondiente al estándar de vídeo particular. En particular, este documento divulga varios métodos para el almacenamiento de flujos de vídeo que contienen múltiples vistas codificadas, en las que cada vista puede contener una vista de textura solamente, una vista de profundidad solamente, o ambas, en un archivo basado en el formato de archivo multimedia base ISO (ISO/BMFF). En varios de los estándares de extensiones de codificación, pueden estar presentes múltiples vistas o capas. Además, las diferentes capas, vistas de textura o profundidad pueden tener diferentes resoluciones espaciales. Se describen técnicas para encapsular el vídeo codificado en un archivo y también se describen técnicas para desencapsular el archivo para recuperar el contenido de vídeo codificado.

Los estándares de formato de archivo incluyen el formato de archivo multimedia base ISO (ISO/BMFF, ISO/IEC 14496-12) (véase, por ejemplo, "Texto de ISO/IEC 14496-12 4ª edición" 100. MPEG MEETING; 30-4-2012 - 4-5-2012; GINEBRA; (GRUPO DE EXPERTOS EN IMÁGENES EN MOVIMIENTO O ISO/IEC JTC1/SC29/WG11), N.º N12640, 7 de junio de 2012) y otros formatos derivados del ISO/BMFF, incluidos el formato de archivo MPEG-4 (ISO/IEC 14496-14), el formato de archivo del Proyecto de Asociación de Tercera Generación (3GPP) (3GPP TS 26.244) y el formato de archivo AVC (ISO/IEC 14496-15). El ISO/BMFF puede utilizarse como base para muchos formatos de encapsulación de codificadores-decodificadores (códecs), como el formato de archivo AVC, así como para muchos formatos de contenedores multimedia, como el formato de archivo MPEG-4, el formato de archivo 3GPP (3GP) y el formato de archivo DVB.

El formato de archivo HEVC utiliza el ISO/BMFF como base para su formato de encapsulación. El último borrador del formato de archivo HEVC (denominado en el presente documento "HEVCFF") se define de acuerdo con la norma ISO/IEC 14496-15:2010/Amd.1:2011(E) e incluye la siguiente referencia normativa: ISO/IEC 14496-15, Tecnología de la información - Codificación de objetos audiovisuales - Parte 15: Transporte de vídeo estructurado en unidades NAL en el formato de archivo multimedia base ISO (es decir, ISO/BMFF).

El ISO/BMFF se utiliza como base para muchos formatos de encapsulación de normas de codificación de vídeo, así como para muchos formatos de contenedores multimedia. Algunos ejemplos de otros formatos de archivo basados en el ISO/BMFF son, por ejemplo, el formato de archivo MPEG4 (ISO/IEC 14496-14), el formato de archivo 3GPP (3GPP TS 26.244) y el formato de archivo AVC (ISO/IEC 14496-15) (es decir, AVCFF). Además, cabe señalar que, si bien las técnicas divulgadas en el presente documento se describen en algunos ejemplos con respecto a HEVC y pueden

hacer referencia a formatos de archivo concretos, pueden ser igualmente aplicables a otros estándares de codificación de vídeo y formatos de archivo.

Además de los medios continuos (como, por ejemplo, el audio y el vídeo), los medios estáticos (como, por ejemplo, las imágenes), así como los metadatos, pueden almacenarse en un archivo conforme al ISOBMFF. Los archivos estructurados de acuerdo con el ISOBMFF pueden utilizarse para muchos propósitos, entre ellos, la reproducción de archivos multimedia locales, la descarga progresiva de un archivo remoto, los segmentos para la Transmisión Dinámica Adaptativa sobre el Protocolo de Transferencia de Hipertexto (HTTP) (DASH), los contenedores para el contenido que se va a transmitir y sus instrucciones de paquetización, y la grabación de los flujos multimedia recibidos en tiempo real. En general, el ISOBMFF permite almacenar datos multimedia continuos, como datos de audio y vídeo, independientemente de los metadatos. En un ejemplo, los datos multimedia pueden incluir un grupo de imágenes codificadas según HEVC. Los metadatos suelen incluir información sobre la presentación y la temporización, de modo que las partes de los datos multimedia pueden decodificarse de forma independiente.

ISOBMFF utiliza un esquema orientado a objetos en el que los elementos de sintaxis elementales se utilizan como bloques de construcción para formar un archivo multimedia. En ISOBMFF, estos elementos de sintaxis elemental se denominan "cuadros". Sin embargo, los formatos de archivo específicos basados en ISOBMFF pueden referirse a los cuadros al utilizar una terminología diferente, como "átomos". Además, cabe señalar que, al describir formatos de archivo específicos basados en el ISOBMFF, la terminología específica puede utilizarse a menudo de forma intercambiable con la terminología del HEVCFF de manera no limitativa. Por ejemplo, al describir el HEVCFF, el término "unidad de acceso" puede corresponder al término "muestra" en el ISOBMFF y una unidad de acceso puede describirse con cualquiera de los dos términos.

En ISOBMFF, un cuadro, el elemento de sintaxis elemental, incluye un tipo de cuatro caracteres, el recuento de bytes del cuadro y la carga útil. Un tipo de cuadro define la relación lógica entre un cuadro particular y otros tipos de cuadros. El tipo de cuadro también puede describir qué datos son obligatorios para el tipo de cuadro en particular y qué datos son opcionales para el tipo de cuadro en particular. Un cuadro puede formar parte de una secuencia o grupo de cuadros y puede contener otros cuadros, que pueden denominarse subcuadros.

En ISOBMFF, un cuadro de película ("moov") contiene metadatos para los flujos de medios continuos presentes en el archivo. Los metadatos de los flujos de medios continuos presentes en el archivo pueden estar encerrados en cuadros adicionales del cuadro de película, como un cuadro de pista ("trak"). Los flujos de medios continuos pueden representarse en un archivo ISOBMFF como una pista, donde una pista es una estructura lógica que hace referencia a los metadatos que pueden estar encerrados en el cuadro de película y donde el flujo de medios consiste en una secuencia de muestras. En ISOBMFF, un flujo de medios puede estar encerrado dentro de cuadros, como un cuadro de datos multimedia ("mdat") y subcuadros de este. Además, en algunos ejemplos, un flujo de medios puede almacenarse directamente en un archivo ISOBMFF. El contenido multimedia de una pista consiste en una secuencia de muestras, como unidades de acceso de audio o vídeo. Una secuencia de muestras puede denominarse flujo. El ISOBMFF especifica una pista de medios que contiene un flujo de medios elemental. Los metadatos de cada pista incluyen una lista de entradas de descripción de muestras, cada una de las cuales proporciona el formato de codificación o encapsulación utilizado en la pista y los datos de inicialización para procesar ese formato. Cada muestra está asociada a una de las entradas de descripción de muestra de la pista. Algunos tipos de archivos basados en ISOBMFF también incluyen pistas de sugerencias. Las pistas de sugerencia incluyen información que facilita la reconstrucción del contenido multimedia, de manera que la calidad y la fiabilidad se mejoran cuando el contenido multimedia se transmite a través de una red.

El ISOBMFF especifica los siguientes tipos de pistas: una pista de medios, que contiene un flujo de medios elemental, una pista de pistas, que incluye instrucciones de transmisión de medios o representa un flujo de paquetes recibidos, y una pista de metadatos temporizada, que comprende metadatos sincronizados en el tiempo.

Si bien se diseñó originalmente para el almacenamiento, el ISOBMFF demostró ser muy valioso para la transmisión, por ejemplo, para la descarga progresiva o DASH. Para la transmisión, se pueden utilizar los fragmentos de película definidos en ISOBMFF.

Los metadatos de cada pista incluyen una lista de entradas de descripción de muestras, cada una de las cuales proporciona el formato de codificación o encapsulación utilizado en la pista y los datos de inicialización para procesar ese formato. Cada muestra está asociada a una de las entradas de descripción de muestra de la pista.

El ISOBMFF permite especificar metadatos específicos de la muestra con diversos mecanismos. Se han estandarizado cuadros específicos dentro del cuadro de la tabla de muestras ("stbl") para responder a necesidades comunes. Por ejemplo, un cuadro de muestra de sincronización ("stss") se utiliza para enumerar las muestras de acceso aleatorio de la pista. Una muestra de sincronización identifica la presencia de una unidad de acceso de refresco de decodificación instantánea (IDR) del flujo de bits de MVC o 3VC para cualquier entrada de muestra que incluya un registro de configuración de MVC o 3VC, respectivamente. El mecanismo de agrupación de muestras permite la asignación de muestras según un tipo de agrupación de cuatro caracteres en grupos de muestras que comparten la

misma propiedad especificada como una entrada de descripción de grupo de muestras en el archivo. En el ISOBMFF se han especificado varios tipos de agrupación.

Al igual que un archivo almacenado según ISOBMFF, un archivo HEVC puede incluir una serie de flujos elementales que hacen referencia a los metadatos. En el HEVCFF, una muestra puede denominarse unidad de acceso, tal como se define en la norma ISO/IEC 23008-2. En el HEVCFF, los metadatos también pueden incluirse en las entradas de descripción de muestras. El HEVCFF puede permitir el uso de pistas, pero debe tenerse en cuenta que lo que las pistas de información denominan "fotogramas B" en algunos formatos de archivo pueden ser en realidad imágenes "desechables" o imágenes que no son de referencia, tal como se definen en la norma ISO/IEC 23008-2. Asimismo, un formato de archivo utilizado para el almacenamiento de contenido de vídeo de HEVC puede utilizar las capacidades existentes del ISOBMFF, pero también puede definir extensiones para admitir características específicas del contenido de vídeo codificado de HEVC, como: conjuntos de parámetros, escalabilidad temporal y punto de recuperación de acceso aleatorio.

En la mayoría de los casos, para almacenar el contenido de vídeo de un códec de vídeo concreto basado en el ISOBMFF, se utiliza una especificación de formato de archivo para ese códec de vídeo. En algunos casos, para el almacenamiento de flujos de vídeo que contienen múltiples vistas codificadas con profundidad que se codifican como una extensión de un estándar, se puede utilizar un formato de archivo compatible con ese estándar para las múltiples vistas codificadas con profundidad. Por ejemplo, el formato de archivo para MVC (que forma parte del formato de archivo AVC) puede reutilizarse para almacenar flujos de vídeo que contengan múltiples vistas codificadas más profundidad (MVC+D). El uso del formato de archivo para MVC para almacenar MVC+D es posible porque el diseño de MVC+D se basa en MVC. Sin embargo, la aplicación directa del formato de archivo de MVC para el almacenamiento de contenidos de MVC+D presenta al menos las siguientes deficiencias. En primer lugar, no habría forma de indicar si una vista incluida en una pista contiene sólo la vista de textura, sólo la vista de profundidad, o ambos componentes de la vista de textura y de profundidad. Además, no habría forma de indicar qué parte (la vista de textura, la vista de profundidad o ambas) de una primera vista de la que depende una segunda vista es necesaria para decodificar la segunda vista. Además, no habría forma de señalar la resolución espacial de las vistas de profundidad sin analizar los conjuntos de parámetros de secuencia (SPS) a los que se refieren los componentes de la vista de profundidad.

Las técnicas de esta divulgación pueden resolver uno o más de los problemas anteriores, así como proporcionar otras mejoras, para permitir el almacenamiento eficiente y flexible de flujos de vídeo que contienen múltiples vistas codificadas con profundidad. Aunque esta descripción es específica para el almacenamiento de flujos de vídeo de MVC+D basados en el formato de archivo MVC, las técnicas aquí descritas pueden aplicarse a cualquier formato de archivo para el almacenamiento de dicho contenido de vídeo codificado por cualquier códec de vídeo. La cláusula 7 de la norma ISO/IEC 14496-15 especifica el formato de archivo MVC. Las técnicas de esta divulgación pueden ampliar la cláusula 7 de la norma ISO/IEC 14496-15 al almacenamiento de datos de codificación de vídeo 3D (3VC) que contengan múltiples vistas con profundidad. En esta divulgación se definen varios cuadros nuevos y tipos de entrada de muestra.

Un único archivo multimedia puede contener múltiples pistas (por ejemplo, pistas de película y de audio, como flujos de datos de vídeo 146 y flujos de datos de audio 148). Una entrada de muestra proporciona el almacenamiento de información de encabezamiento de nivel superior para una muestra o una pista (que, en algunos ejemplos, una pista puede consistir en una o más muestras). Una muestra puede ser, por ejemplo, una unidad de acceso, como la unidad de acceso 150-1, o parte de una unidad de acceso (por ejemplo, en una pista de MVC o 3VC). Una entrada de muestra proporciona un códec para una pista con información de codificación relevante para la pista o la muestra, tal como información de perfil, nivel y capa, así como otra información de alto nivel. En algunos ejemplos, una o más entradas de muestra están presentes en el archivo multimedia de múltiples vistas con profundidad 142.

De acuerdo con las técnicas descritas en el presente documento, se cambia un cuadro del Identificador de vista definido en el formato de archivo MVC. En un ejemplo, el cuadro del Identificador de vista puede cambiarse para indicar, para cada una de las vistas incluidas en una pista asociada con el cuadro del identificador de vista, la presencia de vistas de textura y/o vistas de profundidad en la pista y en el flujo. En otro ejemplo, el cuadro del Identificador de la Vista puede cambiarse para indicar cómo se requiere una vista de textura y una vista de profundidad de una vista de referencia para decodificar una vista específica. Por ejemplo, sólo la vista de textura, sólo la vista de profundidad, o ambas vistas de la vista de referencia pueden ser necesarias para decodificar la vista específica. El cuadro del identificador de vista puede incluirse en una entrada de muestra o en una entrada de grupo de vistas múltiples.

Las técnicas de esta divulgación también admiten la señalización de las resoluciones espaciales de las vistas de profundidad. La resolución espacial de una vista de profundidad puede indicarse en un cuadro de resolución de profundidad de 3VC recién definido. El cuadro de resolución de profundidad de 3VC puede incluirse en un cuadro de configuración de 3VC recién definido. El cuadro de configuración de 3VC incluye un registro de configuración del decodificador 3VC, RegistroConfiguraciónDecodificador3VC, y está incluido en al menos una de las entradas de muestra de tipo "3vc1" o "3vc2". El cuadro de configuración de 3VC puede estar incluido en otra entrada de muestra. El registro de configuración del decodificador 3VC se define para tener la misma sintaxis que RegistroConfiguraciónDecodificadorMVC, pero con una semántica diferente para indicar características relacionadas con un flujo de vídeo 3VC.

En algunos ejemplos, se define un cuadro de mensaje de información de mejora complementaria (SEI) de la información de escalabilidad de 3VC (denominado en el presente documento "cuadro de mensaje de SEI"). El cuadro de mensaje de SEI contiene el mensaje de SEI de información de escalabilidad de 3VC definido en la especificación del códec MVC+D. El cuadro de mensaje de SEI puede incluirse en la entrada de muestra.

Además, las técnicas de esta divulgación proporcionan el almacenamiento de una vista de textura y una vista de profundidad de una vista particular en pistas separadas. En algunos ejemplos, se define un nuevo tipo de referencia de pista con valor de tipo 'deps', cuya presencia indica que la pista actual contiene la vista de profundidad que está asociada con una vista de textura en la pista de referencia.

En algunos ejemplos, se recomienda un nuevo valor de Nombrecompresor para un archivo que contiene flujos de vídeo 3VC. El nuevo nombre es "\Codificación 3VC 012."

Las técnicas aquí descritas introducen dos nuevos tipos de entrada de muestras, '3vc1' y '3vc2' para pistas de 3VC con unidades de profundidad NAL solamente. Los extractores no están permitidos en '3vc1' y sí en '3vc2'. Algunos tipos de entrada de muestra existentes se definen con más detalle. El tipo de entrada de muestra 'avc1' se define para contener configuraciones de AVC, MVC y 3VC, para pistas de 3VC con unidades AVC, MVC y NAL de profundidad. El tipo de entrada de muestra "mvc1" se define además para contener configuraciones de MVC y 3VC, para pistas de 3VC sin unidades NAL de AVC, pero con unidades NAL de MVC y de profundidad en las que no se permiten los extractores. El tipo de entrada de muestra "mvc2" se define además para contener configuraciones de MVC y 3VC, para pistas de 3VC sin unidades NAL de AVC, pero con unidades NAL de MVC y de profundidad en las que se permiten los extractores. Cabe señalar que pueden utilizarse otros nombres de tipos de entrada de muestra con una semántica similar a la descrita anteriormente.

La FIG. 1 es un diagrama conceptual que ilustra la estructura de un ejemplo de archivo multimedia de múltiples vistas con profundidad 142 que puede generarse según las técnicas descritas en esta divulgación. El archivo multimedia de múltiples vistas con profundidad 142 se basa en el ISOBMFF y en el formato de archivo AVC actual. Cabe señalar que el ejemplo de archivo multimedia de múltiples vistas con profundidad 142 de la FIG. 1 pretende ilustrar la relación lógica entre los flujos de datos y los metadatos. En otros ejemplos, el archivo multimedia 142 puede basarse en otros formatos de archivo multimedia. Técnicas similares a las descritas en el presente documento para el archivo multimedia de múltiples vistas con profundidad 142 pueden aplicarse a otros formatos de archivo para adaptar el vídeo tridimensional. A efectos de brevedad, no se ha ilustrado la estructura de encapsulación completa del archivo multimedia de múltiples vistas con profundidad 142. Sin embargo, el archivo multimedia de múltiples vistas con profundidad 142 puede utilizar cuadros y estructuras de encapsulación definidos según el ISOBMFF y el formato de archivo de AVC.

El almacenamiento de flujos de MVC y 3VC puede ser admitido por una serie de estructuras, que incluyen la información en la entrada de la muestra, el cuadro de información multimedia y los grupos de muestra. La Tabla 1 proporciona una visión general de las estructuras proporcionadas, sus nombres y una breve descripción de sus funciones. Más específicamente, la Tabla 1 proporciona un ejemplo de cuadro, entrada de muestra y estructuras de grupo para flujos de MVC y 3VC. Obsérvese que cada grupo de filas que comienza con una entrada en la columna de la izquierda (por ejemplo, 'minf y '?vc?') documenta una estructura de contención dentro de ese contenedor; sin embargo, no se muestra la contención de nivel superior

Tabla 1

				Nombre del cuadro	Breve descripción
minf				Cuadro de información multimedia	
	mvci			Cuadro de información de múltiples vistas	
		mvcg		Cuadro de grupo de múltiples vistas	Especifica un grupo de múltiples vistas para las vistas del flujo de vídeo de múltiples vistas que se emiten
			uff	Cuadro de información de almacenamiento en búfer	Contiene la información de almacenamiento en búfer del subconjunto de flujos de bits especificado por el grupo de múltiples vistas
			vra	Cuadro de atributos de la relación de múltiples vistas	Indica la relación de las pistas o niveles del respectivo grupo de múltiples vistas entre sí (cuando está contenido en un cuadro de grupo de múltiples vistas)
(continuación)					
			ibr	Cuadro de velocidad de bits por capas	Proporciona información sobre los valores de la tasa de bits del subconjunto del flujo de bits especificado por el grupo de múltiples vistas

			iri	Cuadro de información de capas	Proporciona información sobre el perfil, el nivel, el tamaño de los fotogramas, el descarte y la velocidad de los fotogramas del subconjunto del flujo de bits especificado por el grupo de múltiples vistas
			wdi	Cuadro de información de escena de múltiples vistas	Indica la máxima disparidad en una escena con múltiples vistas
		wtc		Cuadro de relación de grupo de múltiples vistas	Especifica un conjunto de grupos de múltiples vistas de los cuales un grupo de múltiples vistas es decodificado y reproducido en cualquier momento
			vra	Cuadro de atributos de la relación de múltiples vistas	Indica la relación de los grupos de múltiples vistas entre sí (cuando están contenidos en un cuadro de relación de grupos de múltiples vistas)
?vc?				Entrada de muestra	(Nota: se utilizan varios códigos para las entradas de muestra)
	vsib			Cuadro de mensaje de SEI de información de escalabilidad de la vista	Contiene una unidad NAL SEI que sólo contiene un mensaje de SEI de información de escalabilidad de la vista, tal como se especifica en el Anexo H de la norma ISO/IEC 14496-10
	ecam			Cuadro de parámetros de la cámara extrínseca	Contiene los parámetros de la cámara que definen la ubicación y la orientación de los fotogramas de referencia de la cámara con respecto a un fotograma de referencia mundial conocido
	icam			Cuadro de parámetros de la cámara intrínseca	Contiene parámetros de la cámara que vinculan las coordenadas de los píxeles de un punto de la imagen con las coordenadas correspondientes en el fotograma de referencia de la cámara
	vwid			Cuadro del identificador de vista	Indica las vistas incluidas en la pista (cuando se incluye en una entrada de muestra)
	vcP			Cuadro de asignación de prioridades de la vista de MVC	Proporciona un URI que contiene un nombre único del método utilizado para asignar valores de _id_prioridad_contenido para la agrupación de muestras de vista de prioridad
	mvcC			Cuadro de configuración de MVC	
	3vcC			Cuadro de configuración de 3VC	Contiene el registro de configuración del decodificador 3VC y el cuadro de resolución de profundidad de 3VC (sólo para flujos de 3VC)
		3dpr		Cuadro de resolución de profundidad de 3VC	Proporciona la resolución de las vistas de profundidad (sólo para flujos de 3VC)
	3sib			Cuadro de mensaje de SEI de información de escalabilidad de 3VC	Contiene una unidad de SEI NAL que sólo contiene un mensaje de SEI de información de escalabilidad 3VC, tal y como se especifica en el Anexo I de ISO/IEC 14496-10
gpd				Cuadro de descripción del grupo de muestra	
	vif			Entrada del grupo de múltiples vistas	Contiene los siguientes cuadros
		uff		Cuadro de información de almacenamiento en búfer	Contiene la información del búfer de la capa
		dep		Cuadro de dependencia de capas	Identifica las capas de las que depende la capa actual
		vip		Cuadro de conjuntos de parámetros iniciales	Contiene los conjuntos de parámetros necesarios para decodificar este nivel y todos los niveles de los que depende
		vpr		Cuadro de rango de prioridad	Informa de la prioridad mínima y máxima _id de las unidades NAL asignadas a esta capa
(continuación)					
		ibr		Cuadro de velocidad de bits por capas	Proporciona información sobre los valores de la tasa de bits de una capa

		iri		Cuadro de información de capas	Proporciona información sobre el perfil, el nivel, el tamaño de los fotogramas, el descarte y la velocidad de los fotogramas de una capa
		ipr		Cuadro del identificador de vista	Etiqueta las vistas con prioridades según el contenido
		vwid		Cuadro del identificador de vista	Indica las vistas incluidas en la capa (cuando se incluyen en una entrada de grupo de múltiples vistas.)
	trt			Decodificación de la entrada del grupo de reajuste de tiempo	Proporciona tiempos de decodificación ajustados cuando se descartan las capas temporales altas
	cnm			Muestra de entrada de un grupo de mapas	Proporciona la asignación de unidades NAL a grupos de múltiples vistas para todas las muestras de la pista

Las estructuras dentro de una entrada de muestra proporcionan información para la decodificación o el uso de las muestras (por ejemplo, información de vídeo) que están asociadas a esa entrada de muestra. Los grupos de muestras proporcionan información variable en el tiempo sobre la pista en su conjunto, lo que ayuda, por ejemplo, a la extracción de subconjuntos de medios dentro de una pista. La información del cuadro de información de múltiples vistas (que aparece en el cuadro de información multimedia) puede abarcar varias pistas y describe colecciones de pistas, aunque el cuadro de información de múltiples vistas resida en la pista que contiene la vista base del flujo.

En algunos ejemplos, los flujos elementales de vídeo de MVC y 3VC contienen todas las unidades NAL relacionadas con la codificación de vídeo (es decir, aquellas unidades NAL que contienen datos de vídeo o estructura de vídeo de señalización) y pueden contener unidades NAL no relacionadas con la codificación de vídeo, como mensajes de SEI y unidades NAL delimitadoras de unidades de acceso. Además, puede haber agregadores o extractores en los flujos de vídeo de MVC y 3VC. Es posible que haya otras unidades NAL que no estén expresamente prohibidas, y si no se reconocen, pueden ignorarse (por ejemplo, no colocarse en el búfer de salida al acceder al archivo). Los flujos de MVC y 3VC también pueden almacenarse al utilizar flujos de conjuntos de parámetros asociados, cuando sea necesario. En un flujo elemental de vídeo de MVC puede haber unidades NAL de AVC, unidades NAL de MVC y unidades NAL no-VCL. En un flujo elemental de vídeo 3VC puede haber unidades NAL VCL de AVC, unidades NAL VCL de MVC, unidades NAL VCL de profundidad y unidades NAL no VCL. Además, puede haber unidades NAL agregadoras o extractoras presentes en un flujo elemental de vídeo de MVC o 3VC.

El archivo multimedia de múltiples vistas con profundidad 142 puede almacenar medios continuos (por ejemplo, audio y vídeo), medios estáticos (por ejemplo, imágenes) y metadatos. La estructura sintáctica elemental en el ISOBMFF es un cuadro, que incluye un tipo de cuadro codificado de cuatro caracteres, el recuento de bytes del cuadro y la carga útil. El archivo multimedia de múltiples vistas con profundidad 142 consiste en una secuencia de cuadros. En el ejemplo ilustrado en la FIG. En el ejemplo ilustrado en la FIG. 1, el archivo multimedia de múltiples vistas con profundidad 142 incluye un cuadro de película 144, flujos de datos de vídeo 146-1 a 146-N (aquí denominados colectivamente "flujos de datos de vídeo 146"), y flujos de datos de audio 148-1 a 148-N (aquí denominados colectivamente "flujos de datos de audio 148"), en los que N puede ser cualquier número entero positivo, y puede ser el mismo o diferente número para las diferentes instancias de N en la FIG. 1.

Los flujos de datos de vídeo 146 pueden corresponder a las pistas de vídeo o visuales descritas en el ISOBMFF. Un flujo de vídeo puede estar representado por una pista de vídeo en múltiples vistas con el archivo multimedia de profundidad 142. Cada uno de los flujos de datos de vídeo 146-1 a 146-N puede incluir una o más unidades de acceso 150-1 a 150-N (aquí denominadas colectivamente "unidades de acceso 150"). Las unidades de acceso 150 en el archivo multimedia de múltiples vistas con profundidad 142 pueden definirse de acuerdo con el ISOBMFF, o más específicamente, MVC o HEVCFF. Las unidades de acceso 150 pueden corresponder a una muestra en ISOBMFF. Las unidades de acceso 150 pueden estar enmarcadas externamente y tener un tamaño suministrado por ese marco externo. En algunos casos, el cuadro de metadatos 154 puede incluir información que defina el tamaño de las unidades de acceso 150. Las unidades de acceso 150 pueden incluir un conjunto de unidades NAL (es decir, capa de abstracción de red) 152-1 a 152-N (aquí denominadas colectivamente "unidades NAL 152"). Las unidades NAL 152 pueden definirse de acuerdo con el HEVCFF.

Cada unidad NAL 152 puede incluir un campo de longitud que indica la longitud de la unidad NAL, por ejemplo, en bytes u otras unidades de longitud o tamaño asociadas con el almacenamiento de datos. En algunos ejemplos, el campo de longitud puede estar configurado para ser de uno, dos o cuatro bytes. Los campos de longitud permiten un fácil análisis de las unidades NAL dentro de una unidad de acceso 150. Las unidades NAL 152 pueden contener datos de unidades NAL como se especifica en la norma ISO/IEC 23008-2. Además de otros tipos de unidades NAL, el HEVCFF define los siguientes tipos de unidades NAL: Unidades NAL de datos de vídeo, que pueden denominarse unidades NAL VCL, unidades NAL de información de mejora complementaria (SEI) y unidades NAL delimitadoras de unidades de acceso (AU), así como otros tipos de unidades NAL. Cabe señalar que los tipos de unidad NAL reservados en la norma ISO/IEC 23008-2 pueden adquirir una definición en el futuro. Algunos lectores de archivos pueden estar configurados para ignorar las unidades NAL con valores reservados de tipo de unidad NAL cuando estén presentes

en estas matrices. Cabe señalar que este comportamiento "tolerante" está diseñado para que no se produzcan errores, lo que permite la posibilidad de realizar extensiones compatibles con el pasado para estas matrices en futuras especificaciones. Las unidades NAL de VCL pueden estar formateadas de tal manera que todas las unidades NAL de rebanada codificada para una sola imagen estén contenidas dentro de la unidad de acceso cuyo tiempo de decodificación y tiempo de composición sean los de la imagen. Además, las unidades de acceso 150 pueden estar obligadas a contener al menos una unidad NAL de datos de vídeo.

El cuadro de película 144 puede ser similar a un cuadro ISOBMFF 'moov'. Como se describe aquí, "moov" es un tipo de cuadro codificado de cuatro caracteres para el cuadro de película 144 definido en el ISOBMFF. Además, los cuadros pueden contener otros cuadros. El cuadro de película 144 ("moov") contiene metadatos para cualquier flujo de medios continuos 146 y 148 presentes en el archivo multimedia de múltiples vistas con profundidad 142.

El cuadro de película 144 puede contener uno o más cuadros de metadatos 154. El cuadro de metadatos 154 puede ser similar y puede contener metadatos para cada uno de los flujos de datos de vídeo 146 y/o flujos de datos de audio 148. En un ejemplo, los metadatos pueden estar incluidos en una o más tablas de metadatos 156 incluidas en uno o más cuadros de metadatos 154. Las tablas de metadatos 156 pueden incluir entradas de muestra que corresponden a los flujos de datos de vídeo 146. Un ejemplo de tabla de metadatos 156 es un cuadro de tabla de muestra. Cada uno de los flujos de datos de vídeo 146 puede hacer referencia a una o más tablas de metadatos 156.

El ISOBMFF permite especificar metadatos específicos de la muestra con diversos mecanismos. El cuadro de metadatos 154 puede incluir además cuadros específicos dentro de un cuadro de tabla de muestras 162 para contener otra información de metadatos, como cuadros de descripción de muestras que contengan además entradas de muestras. Un tipo de cuadro codificado con cuatro caracteres para el cuadro de Tabla de Muestras 162 puede ser "stbl". Por ejemplo, un cuadro 164 de muestras de sincronización puede utilizarse para enumerar las muestras de acceso aleatorio de la pista. Un tipo de cuadro codificado de cuatro caracteres para el cuadro de muestra de sincronización 164 puede ser "stss". El mecanismo de agrupación de muestras permite el mapeo de muestras de acuerdo con un tipo de agrupación de cuatro caracteres en grupos de muestras que comparten la misma propiedad especificada como una entrada de descripción de grupo de muestras en el archivo multimedia de múltiples vistas con profundidad 142. En el ISOBMFF se han especificado varios tipos de agrupación.

En algunos ejemplos, un cuadro de sincronización de sombra no se utiliza para los datos de vídeo descritos por cualquier entrada de muestra de MVC o 3VC.

En algunos ejemplos, se puede utilizar un cuadro de muestras independiente y desechable. Si se utiliza un cuadro de muestras independientes y desechables en una pista que es compatible con AVC y MVC, entonces la información indicada por el cuadro de muestras independientes y desechables es verdadera sin importar qué subconjunto válido de los datos de MVC (y posiblemente sólo los datos de AVC) se utiliza. Los valores "desconocidos" (el valor 0 de los campos muestra-depende-de, muestra-es-dependiente-de y muestra-tiene-redundancia) pueden utilizarse si la información varía. Si se utiliza un cuadro de muestras independiente y desechable en una pista que es compatible con todas las AVC, MVC y 3VC, se debe tener cuidado de que las declaraciones sean verdaderas sin importar qué subconjunto válido de los datos 3VC (posiblemente sólo los datos de AVC o sólo los datos de MVC) se utilice. Los valores "desconocidos" (el valor 0 de los campos muestra-depende-de, muestra-es-dependiente-de y muestra-tiene-redundancia) pueden utilizarse si la información varía.

El archivo multimedia de múltiples vistas con profundidad 142 puede incluir un grupo de muestra de recuperación de acceso aleatorio y un grupo de muestra de punto de acceso aleatorio. Para los datos de vídeo descritos por una entrada de muestra de tipo "avc1" o "avc2", el grupo de muestra de recuperación de acceso aleatorio y el grupo de muestra de punto de acceso aleatorio identifican puntos de recuperación de acceso aleatorio y puntos de acceso aleatorio, respectivamente, para todos los decodificadores AVC, MVC y 3VC que operan en todo el flujo de bits. Para los datos de vídeo descritos por un tipo de entrada de muestra de MVC, el grupo de muestra de recuperación de acceso aleatorio identifica la recuperación de acceso aleatorio en todo el flujo de bits de MVC y el grupo de muestra de punto de acceso aleatorio identifica los puntos de acceso aleatorio en todo el flujo de bits de MVC. Para los datos de vídeo descritos por un tipo de entrada de muestra de 3VC, el grupo de muestra de recuperación de acceso aleatorio puede identificar la recuperación de acceso aleatorio en todo el flujo de bits de 3VC y el grupo de muestra de punto de acceso aleatorio identifica los puntos de acceso aleatorio en todo el flujo de bits de 3VC.

El cuadro de película 144 puede contener además uno o más cuadros de pista 158 y uno o más cuadros de datos multimedia 160. Los metadatos de una pista están encerrados en al menos un cuadro de pista 158. El tipo de cuadro codificado de cuatro caracteres para el cuadro de pista 158 puede ser "trak". El contenido multimedia de una pista está encerrado en un cuadro de datos multimedia 160 o directamente en un archivo separado. El tipo de cuadro codificado de cuatro caracteres para el cuadro de Datos multimedia 160 puede ser "mdat". El contenido multimedia de pistas consiste en una secuencia de muestras, como unidades de acceso de audio o vídeo.

Los flujos de medios continuos presentes en el archivo multimedia de múltiples vistas con profundidad 142 pueden representarse en una pista. El ISOBMFF especifica los siguientes tipos de pistas: una pista de medios, que contiene un flujo de medios elemental, una pista de pistas, que incluye instrucciones de transmisión de medios o representa un

flujo de paquetes recibidos, y una pista de metadatos temporizada, que comprende metadatos sincronizados en el tiempo.

Un flujo de MVC o 3VC puede colocarse en pistas de varias maneras, entre las que se encuentran las siguientes:

- 5 todas las vistas en una pista, etiquetadas con grupos de muestra; cada vista, incluidas las vistas de textura y las vistas de profundidad cuando ambas están presentes, en su propia pista, etiquetada en las entradas de muestra; un híbrido,
- 10 una pista que contenga todas las vistas, y una o más pistas de una sola vista, cada una de las cuales contenga una vista que pueda codificarse de forma independiente; los puntos de operación esperados cada uno en una pista (por ejemplo, la base de AVC, el par estéreo o la escena de 3VC); y, sólo para 3VC, cada vista de textura o de profundidad en su propia pista, etiquetada en las entradas de muestra.

El formato de archivo de MVC y 3VC permite almacenar una o más vistas en una pista. El almacenamiento de varias vistas por pista puede utilizarse, por ejemplo, cuando un proveedor de contenidos quiere proporcionar un flujo de bits de múltiples vistas que no está pensado para el subconjunto o cuando el flujo de bits se ha creado para unos pocos

- 15 conjuntos predefinidos de vistas de salida (como 1, 2, 5 o 9 vistas) en los que se pueden crear pistas en consecuencia. Si se almacena más de una vista en una pista y hay varias pistas (más de una) que representan el flujo de bits de MVC o SVC, se puede recomendar el uso del mecanismo de agrupación de muestras. El mecanismo de agrupación de muestras puede utilizarse para definir capas que identifiquen las vistas presentes en la vía y para extraer convenientemente las unidades NAL necesarias para determinados puntos de operación. El mecanismo de agrupación
- 20 de muestras suele utilizarse con unidades NAL agregadoras para formar patrones regulares de unidades NAL dentro de las muestras. Así, la agrupación de muestras, los agregadores y las definiciones de vistas para los grupos de muestras se especifican para MVC y 3VC.

En esta divulgación se incluyen varias técnicas para indicar si una vista incluida en una pista contiene sólo vista de

- 25 textura, sólo vista de profundidad, o ambas vistas de textura y de profundidad, indicar qué parte (la vista de textura, la vista de profundidad, o ambas) de una primera vista de la que depende una segunda vista es necesaria para decodificar la segunda vista, señalar la resolución espacial de las vistas de profundidad de manera que no sea necesario analizar los conjuntos de parámetros de secuencia (SPS) a los que se refieren los componentes de la vista de profundidad, y mejorar la eficiencia y la flexibilidad del almacenamiento de flujos de vídeo que contienen múltiples
- 30 vistas codificadas con profundidad.

Además, en algunos ejemplos se admite el almacenamiento de la vista de textura y la vista de profundidad de una vista concreta en pistas separadas. Se define un nuevo tipo de referencia de pista con valor de tipo 'deps'. La presencia del nuevo tipo de referencia de pista 'deps' puede indicar que una pista actual contiene la vista de profundidad que

- 35 está asociada a una vista de textura en una pista de referencia. En un ejemplo alternativo, se define un nuevo tipo de referencia de pista con valor de tipo 'tref', cuya presencia en una pista indica que la pista actual depende de la vista de textura almacenada en la pista de referencia. En otro ejemplo alternativo, se define otro nuevo tipo de referencia de pista con valor de tipo 'dref', cuya presencia en una pista indica que la pista actual depende de la vista de profundidad almacenada en la pista de referencia.

El archivo multimedia de múltiples vistas con profundidad 142 puede contener además uno o más cuadros del

- 40 identificador de vista 172, que pueden incluirse en una entrada de muestra o en una entrada de grupo de múltiples vistas. Por ejemplo, el cuadro del identificador de vista 172 puede indicar, para cada vista incluida en una pista, la presencia de vistas de textura y/o profundidad en la pista y en el flujo de vídeo. En otro ejemplo, el cuadro del
- 45 identificador de vista 172 puede indicar cómo se requiere la vista de textura y la vista de profundidad de una vista de referencia para decodificar una vista específica. Es decir, el cuadro del identificador de vista 172 puede contener información sintáctica que indique si sólo se requiere la vista de la textura, sólo se requiere la vista de la profundidad, o se requieren ambas vistas, la de la textura y la de la profundidad, para decodificar una vista específica. Tal como se utiliza en este documento, la codificación puede referirse a la composición o el almacenamiento de información según
- 50 un formato de archivo específico y la decodificación puede referirse al análisis de la información según el formato de archivo. En general, la codificación puede referirse a la composición o al análisis.

El archivo multimedia de múltiples vistas con profundidad 142 puede contener además uno o más cuadros de configuración de 3VC 168. El cuadro de configuración de 3VC 168 puede incluir un registro de configuración del

- 55 decodificador 3VC. El registro de configuración del decodificador 3VC puede denominarse RegistroConfiguraciónDecodificador3VC. El registro de configuración del decodificador 3VC puede incluir una sintaxis similar a la de RegistroConfiguraciónDecodificadorMVC, pero tener una semántica diferente para indicar características relacionadas con un flujo de vídeo 3DV. El cuadro de configuración de 3VC puede incluirse en una entrada de muestra, en algunos ejemplos, debe incluirse en una entrada de muestra de tipo '3vc1' o '3vc2'.

Las técnicas de la divulgación también admiten la señalización de la resolución espacial de las vistas de profundidad. El cuadro de configuración de 3VC 168 puede incluir además un cuadro de resolución de profundidad de 3VC 170. Las resoluciones espaciales de las vistas de profundidad pueden señalizarse en el cuadro de resolución de

- 60 profundidad de 3VC 170. En otros ejemplos, que quedan fuera de la invención reivindicada, el cuadro de resolución de profundidad de 3VC 170 puede añadirse directamente en una entrada de muestra. En otro ejemplo alternativo, la
- 65

anchura y la altura de las vistas de profundidad pueden señalarse directamente en el registro de configuración del decodificador 3VC.

El archivo multimedia de múltiples vistas con profundidad 142 puede contener además uno o más cuadros de mensaje de SEI de información de escalabilidad de 3VC 166. El cuadro de mensaje de SEI de información de escalabilidad de 3VC 166 puede contener un mensaje de información de mejora complementaria (SEI) de información de escalabilidad de 3VC definido en la especificación del códec MVC+D. En algunos ejemplos, el cuadro de mensaje de SEI de información de escalabilidad de 3VC 166 puede incluirse en una entrada de muestra.

El archivo multimedia de múltiples vistas con profundidad 142 puede contener además un cuadro de grupo de múltiples vistas 174. El cuadro de agrupación de múltiples vistas 174 puede indicar los puntos de operación utilizados en la codificación de flujos de datos de vídeo 146, incluidos los puntos de operación 3DV. En un ejemplo, la agrupación de múltiples vistas es un mecanismo de estructura y agrupación que indica la asociación de unidades NAL con diferentes tipos y niveles jerárquicos de escalabilidad. Un tipo de cuadro codificado de cuatro caracteres para el cuadro de grupo de múltiples vistas 174 puede ser "mvcg". En algunos ejemplos, el cuadro de grupo de múltiples vistas 174 puede estar contenido en un cuadro de información de múltiples vistas ("mvti"). Puede haber cero o más cuadros de grupo de múltiples vistas 174 presentes en el archivo multimedia múltiples vistas con profundidad 142.

El cuadro de grupo de múltiples vistas 174 puede especificar un grupo de múltiples vistas para las vistas del flujo de MVC o 3VC que se emiten. Las vistas de salida objetivo pueden indicarse sobre la base de `id_pista`, `id_capa` o `id_vista`. Cuando las vistas incluidas en una pista coinciden con un punto operativo, se puede utilizar `id_pista` (es decir, `tipo_entrada` igual a 0) dentro del cuadro de grupo de múltiples vistas 174. Cuando se utiliza la agrupación de muestras de múltiples vistas y las capas cubren más de una vista o algunas capas contienen un subconjunto temporal del flujo de bits, se puede utilizar la indicación de capa (es decir, `tipo_entrada` igual a 1) dentro del cuadro de grupo de múltiples vistas 174. De lo contrario, se puede utilizar una de las indicaciones basadas en `id_vista` (es decir, `tipo_entrada` igual a 2 o 3).

Cuando `tipo_entrada` es igual a 0 o 1, puede aplicarse lo siguiente. Cada vista de una pista o capa incluida en el cuadro de grupo de múltiples vistas 174 es una vista de salida de destino, y si una pista o capa incluida en este cuadro contiene múltiples vistas, todas las vistas contenidas son vistas de salida de destino.

La semántica del cuadro de grupo de múltiples vistas 174 incluye un elemento sintáctico `núm_entradas`, que puede indicar un número de pistas (tipo de entrada 0), capas (tipo de entrada 1), vistas de salida de destino (tipo de entrada 2), o conjuntos continuos de vistas de salida de destino (tipo de entrada 3) incluidos en el grupo de múltiples vistas.

Otra técnica descrita en el presente documento incluye la definición de un atributo común y/o un atributo diferenciador para indicar si están presentes los puntos de operación de 3DV indicados en el cuadro de grupo múltiples vistas 174. En otro ejemplo, puede incluirse información en el cuadro de grupo de múltiples vistas 174 para indicar si un punto de operación indicado en el cuadro de grupo de múltiples vistas 174 es un punto de operación de 3DC o un punto de operación de MVC. Un punto de operación puede ser un subconjunto decodificable independientemente de un flujo de bits escalable. Un subconjunto de flujo de bits de un punto de operación de MVC o 3VC puede representar un conjunto particular de vistas de salida de destino en una resolución temporal particular, y consiste en datos que pueden utilizarse para decodificar el subconjunto de flujo de bits particular. En 3VC, cada vista de salida objetivo en el subconjunto de flujo de bits de un punto operativo de 3VC puede contener una vista de textura, una vista de profundidad o ambas. Además, en algunos ejemplos, se introducen elementos sintácticos, que pueden incluirse en el cuadro de grupo de múltiples vistas 174, para indicar si las vistas de salida indicadas en el cuadro de grupo de múltiples vistas 174 contienen sólo vistas de textura, sólo vistas de profundidad, o tanto vistas de textura como de profundidad.

En técnicas adicionales de esta divulgación, se definen nuevos tipos de entrada de muestras. Por ejemplo, '3vcl' y '3vc2' se definen para pistas de 3VC con unidades NAL de profundidad solamente, donde los extractores no están permitidos y están permitidos, respectivamente. Un extractor es una estructura dentro del flujo (por ejemplo, una estructura que reside dentro de los datos de muestra) para permitir la extracción eficiente de unidades NAL de pistas distintas de la que contiene los datos multimedia relevantes. Por ejemplo, un extractor puede contener instrucciones sobre cómo extraer datos de otras pistas.

El ejemplo de la FIG. 1 ilustra sólo una forma posible de organizar el archivo multimedia de múltiples vistas con profundidad 142. En algunos ejemplos, uno o más de los cuadros pueden ser subcuadros de otros cuadros. Es decir, en diferentes ejemplos, los diversos cuadros pueden organizarse y configurarse de maneras diferentes a las mostradas en la FIG. 1. Además, una o más de las técnicas descritas en el presente documento pueden utilizarse en cualquier combinación con una o más de las otras técnicas descritas en el presente documento.

La FIG. 2 ilustra un ejemplo de una unidad de acceso MVC+D 150 que puede incluirse en un archivo multimedia generado según las técnicas descritas en esta divulgación. La FIG. 2 ilustra un ejemplo de estructura de una muestra, que en este ejemplo es una unidad de acceso completa 150. Las muestras están enmarcadas externamente y tienen un tamaño suministrado por ese enmarcado externo. La sintaxis de una muestra se configura a través de la configuración específica del decodificador para el flujo elemental. Un ejemplo de la estructura de una muestra de vídeo

se representa en la siguiente figura. Una muestra de 3VC puede incluir uno o más componentes de vista tal y como se define en el Anexo I de la norma ISO/IEC 14496-10 y las unidades NAL no VCL asociadas, donde cada componente de vista puede contener un componente de vista de textura, un componente de vista de profundidad, o ambos.

Como se ilustra en la FIG. 2, la unidad de acceso 150 incluye un conjunto de unidades NAL, entre ellas una unidad NAL delimitadora AU, una unidad NAL SEI, y dos unidades NAL de rebanada. Cada unidad NAL se representa con una longitud, que indica una longitud en bytes de la unidad NAL que le sigue en el orden del flujo de bits. El campo de longitud puede estar configurado para ser de 1, 2 o 4 bytes, por ejemplo. Una unidad NAL contiene los datos de la unidad NAL como se especifica en el estándar de codificación de vídeo aplicable.

En algunos ejemplos, se proporciona una extensión de rebanada codificada para las unidades NAL de rebanada codificada en profundidad. En algunos ejemplos, todas las unidades NAL de rebanada codificada en profundidad para un único instante se incluirán en una muestra cuyo tiempo de composición es el de una imagen representada por la unidad de acceso. Una muestra de 3VC deberá contener al menos una unidad NAL de AVC, MVC o VCL de profundidad.

En algunos ejemplos, puede reconstruirse una unidad de acceso, como la unidad de acceso 150-1. Para reconstruir la unidad de acceso 150-1 a partir de muestras de una o más pistas de MVC o 3VC, puede ser necesario determinar las vistas de salida de destino, al examinar el cuadro de grupo de múltiples vistas 174 y un cuadro de relación de grupo de múltiples vistas. Un indicador pista_au_explicita igual a 1 indica que esta pista es un punto en funcionamiento completo; no obstante, la pista puede examinarse para determinar qué vistas enviadas por esta pista son las vistas de salida.

Si las vistas de salida de destino no están representadas exactamente por ninguna pista marcada con pista_au_explicita igual a 1 en un registro de configuración del decodificador MVC o en un registro de configuración del decodificador 3VC, la unidad de acceso 150-1 puede reconstruirse de la siguiente forma. Las vistas necesarias para la decodificación de las vistas de salida de destino determinadas pueden concluirse a partir de los identificadores de vistas de referencia incluidos en el cuadro del identificador de vistas, las referencias de la pista "sello" o los cuadros de dependencia de capas.

Si varias pistas contienen datos para la unidad de acceso 150-1, la alineación de las respectivas muestras en las pistas se realiza en tiempo de decodificación, por ejemplo, al utilizar sólo la tabla de tiempo a muestra sin considerar las listas de edición. La unidad de acceso 150-1 se reconstruye a partir de las muestras respectivas en las pistas y capas que se requieren, al disponer sus unidades NAL en un orden conforme a la norma ISO/IEC 14496-10. El siguiente orden proporciona un esquema del procedimiento para formar una unidad de acceso conforme: todas las unidades NAL del conjunto de parámetros (de las pistas del conjunto de parámetros asociadas y de las pistas del flujo elemental asociadas); todas las unidades NAL de SEI (de las pistas del conjunto de parámetros asociadas y de las pistas del flujo elemental asociadas); los componentes de la vista en orden ascendente del valor del índice de orden de la vista; y dentro de un componente de la vista, si están presentes tanto la textura como la profundidad, el componente de la vista de la textura precede al componente de la vista de la profundidad. Las unidades NAL dentro de un componente de vista de textura o de profundidad están en su orden de aparición dentro de la muestra.

La FIG. 3 es un diagrama esquemático que ilustra un ejemplo de sistema de codificación y decodificación de vídeo 10 que puede utilizar las técnicas descritas en esta divulgación. En un ejemplo, el sistema 10 puede estar configurado para componer o analizar un archivo multimedia que contenga múltiples vistas codificadas más profundidad. En este documento, los términos componer y encapsular pueden utilizarse indistintamente. Asimismo, tal y como se utilizan aquí, los términos analizar y desencapsular pueden usarse indistintamente. Además, en algunos ejemplos, el término codificar se utiliza de forma similar al término componer y el término decodificar se utiliza de forma similar al término analizar.

Según lo mostrado en la FIG. 3, el sistema 10 incluye un dispositivo de origen 12 que genera datos de vídeo codificados para ser decodificados en un momento posterior por un dispositivo de destino 14. El dispositivo de origen 12 y el dispositivo de destino 14 pueden comprender cualquiera de una amplia gama de dispositivos, que incluyen ordenadores de escritorio, ordenadores portátiles (es decir, ordenadores portátiles), tabletas, decodificadores, aparato telefónico con microteléfono como los denominados teléfonos "inteligentes", los denominados teclados "inteligentes", televisores, cámaras, dispositivos de visualización, reproductores de medios digitales, consolas de videojuegos, dispositivos de transmisión de vídeo o similares. En algunos casos, el dispositivo de origen 12 y el dispositivo de destino 14 pueden estar equipados para comunicación inalámbrica.

El dispositivo de destino 14 puede recibir los datos de vídeo codificados para decodificarlos a través de un enlace 16. El enlace 16 puede comprender cualquier tipo de medio o dispositivo capaz de mover los datos de vídeo codificados desde el dispositivo de origen 12 al dispositivo de destino 14. En un ejemplo, el enlace 16 puede comprender un medio de comunicación para permitir que el dispositivo de origen 12 transmita datos de vídeo codificados directamente al dispositivo de destino 14 en tiempo real. Los datos de vídeo codificados pueden modularse de acuerdo con un estándar de comunicación, tal como un protocolo de comunicación inalámbrica, y transmitirse al dispositivo 14 de destino. El medio de comunicación puede comprender cualquier medio de comunicación inalámbrica o cableada, como un

espectro de radiofrecuencia (RF) o una o más líneas de transmisión físicas. El medio de comunicación puede formar parte de una red basada en paquetes, como una red de área local, una red de área amplia o una red global como Internet. El medio de comunicación puede incluir enrutadores, conmutadores, estaciones base o cualquier otro equipo que pueda ser útil para facilitar la comunicación desde el dispositivo de origen 12 al dispositivo de destino 14.

Alternativamente, los datos codificados, como un archivo multimedia MVC+D compuesto, pueden salir de la interfaz de salida 22 a un dispositivo de almacenamiento 32. De manera similar, se puede acceder a los datos codificados desde el dispositivo de almacenamiento 32 mediante la interfaz de entrada 28. El dispositivo de almacenamiento 32 puede incluir cualquiera de una variedad de medios de almacenamiento de datos distribuidos o de acceso local, como un disco duro, discos Bluray, DVD, CD-ROM, memoria flash, memoria volátil o no volátil o cualquier otro medio de almacenamiento digital adecuado para almacenar datos de video codificados. En un ejemplo adicional, el dispositivo de almacenamiento 32 puede corresponder a un servidor de archivos u otro dispositivo de almacenamiento intermedio que puede contener el video codificado generado por el dispositivo de origen 12. En algunos ejemplos, el dispositivo de almacenamiento 32 almacena contenido de video codificado que un módulo de encapsulación 21 recupera posteriormente para su encapsulación. El dispositivo de destino 14 puede acceder a los datos de video almacenados desde el dispositivo de almacenamiento 32 mediante transmisión o descarga. El servidor de archivos puede ser cualquier tipo de servidor capaz de almacenar datos de video codificados y transmitir esos datos de video codificados al dispositivo de destino 14. Ejemplos de los servidores de archivos incluyen un servidor web (por ejemplo, para un sitio web), un servidor FTP, dispositivos de almacenamiento conectados a la red (NAS) o una unidad de disco local. El dispositivo de destino 14 puede acceder a los datos de video codificados a través de cualquier conexión de datos estándar, incluida una conexión a Internet. Puede incluir un canal inalámbrico (por ejemplo, una conexión Wi-Fi), una conexión por cable (por ejemplo, DSL, módem por cable, etc.) o una combinación de ambos que sea adecuada para acceder a datos de video codificados almacenados en un servidor de archivos. La transmisión de datos de video codificados desde el dispositivo de almacenamiento 32 puede ser una transmisión en directo, una transmisión de descarga o una combinación de ambas.

Las técnicas de esta divulgación no se limitan necesariamente a aplicaciones o configuraciones inalámbricas. Las técnicas se pueden aplicar a la codificación de video para respaldar cualquiera de una variedad de aplicaciones multimedia, como transmisiones de televisión por aire, transmisiones de televisión por cable, transmisiones de televisión por satélite, transmisiones de video en directo, por ejemplo, a través de Internet, codificación de video digital para almacenamiento en un medio de almacenamiento de datos, decodificación de video digital almacenado en un medio de almacenamiento de datos u otras aplicaciones. En algunos ejemplos, el sistema 10 puede configurarse para admitir transmisión de video unidireccional o bidireccional para admitir aplicaciones tales como transmisión de video, reproducción de video, difusión de video y/o telefonía de video.

En el ejemplo de la FIG. 1, el dispositivo fuente 12 incluye una fuente de video 18, un codificador de video 20, un módulo de encapsulación 21 y una interfaz de salida 22. En algunos casos, la interfaz de salida 22 puede incluir un modulador/demodulador (módem) y/o un transmisor. En el dispositivo de origen 12, la fuente de video 18 puede incluir una fuente tal como un dispositivo de captura de video, por ejemplo, una cámara de video, un archivo de video que contiene video capturado previamente, una interfaz de alimentación de video para recibir videos de un proveedor de contenido de videos y/o un sistema de gráficos por ordenador para generar datos de gráficos por ordenador como el video de origen, o una combinación de tales fuentes. Como ejemplo, si la fuente de video 18 es una cámara de video, el dispositivo de origen 12 y el dispositivo de destino 14 pueden formar los denominados teléfonos con cámara o videoteléfonos. Sin embargo, las técnicas descritas en esta divulgación pueden ser aplicables a la codificación de video en general, y pueden aplicarse a aplicaciones inalámbricas y/o cableadas.

El video capturado, precapturado o generado por ordenador puede ser codificado por el codificador de video 20. Los datos de video codificados pueden transmitirse directamente al dispositivo de destino 14 a través de la interfaz de salida 22 del dispositivo de origen 12. Los datos de video codificados también pueden (o alternativamente) almacenarse en el dispositivo de almacenamiento 32 para un acceso posterior por el dispositivo de destino 14 u otros dispositivos, para la decodificación y/o reproducción.

El módulo de encapsulación 21 puede recibir contenido de video codificado del codificador de video 20 y generar un archivo de video de acuerdo con las técnicas aquí descritas. En un ejemplo, el módulo de encapsulación 21 puede recibir contenido de video codificado de MVC o 3VC y generar un archivo de video al utilizar un formato de archivo basado en el ISOBMFF y la versión actual del AVCFF. La FIG. 3 muestra el dispositivo fuente 12 que incluye un codificador de video 20. Sin embargo, en algunos ejemplos, el dispositivo fuente 12 puede no incluir ningún codificador de video, pero sí incluye el módulo de encapsulación 21 que encapsula los datos de video codificados. En algunos ejemplos, los datos de video codificados pueden almacenarse en una ubicación de almacenamiento remota, como el dispositivo de almacenamiento 32, por ejemplo. Como se describe en el presente documento, el módulo de encapsulación 21 puede ser un dispositivo de codificación de video.

El dispositivo de destino 14 incluye una interfaz de entrada 28, un módulo de desencapsulación 29, un decodificador de video 30 y un dispositivo de visualización 31. En algunos casos, la interfaz de entrada 28 puede incluir un receptor y/o un módem. La interfaz de entrada 28 del dispositivo de destino 14 recibe los datos de video codificados a través del enlace 16. Los datos de video codificados comunicados a través del enlace 16, o proporcionados en el dispositivo

de almacenamiento 32, pueden incluir una variedad de elementos de sintaxis generados por el codificador de vídeo 20 para su uso por un decodificador de vídeo, como el decodificador de vídeo 30, al decodificar los datos de vídeo. Tales elementos de sintaxis pueden incluirse con los datos de vídeo codificados transmitidos en un medio de comunicación, almacenados en un medio de almacenamiento o almacenados en un servidor de archivos.

El módulo de desencapsulación 29 puede estar configurado para recibir un archivo de vídeo y analizarlo y alimentar el flujo de bits desencapsular al decodificador de vídeo 30, de manera que el decodificador de vídeo 30 pueda emitir datos de vídeo decodificados. En algunos ejemplos, el módulo de desencapsulación 29 puede realizar procesos recíprocos del módulo de encapsulación 21. El decodificador de vídeo 30 puede, en algunos ejemplos, realizar un pase de decodificación generalmente recíproco al pase de codificación descrito con respecto al codificador de vídeo 20. En algunos ejemplos, el módulo de desencapsulación 29 está incluido en el decodificador de vídeo 30. En el resto de este documento, el decodificador de vídeo 30 que analiza un archivo puede referirse a un módulo de desencapsulación, como el módulo de desencapsulación 29 incluido en el decodificador de vídeo 30, que analiza un archivo. Como se describe en el presente documento, el módulo de encapsulación 21 puede ser un dispositivo de codificación de vídeo.

El dispositivo de visualización 31 puede estar integrado o ser externo al dispositivo de destino 14. En algunos ejemplos, el dispositivo de destino 14 puede incluir un dispositivo de visualización integrado y también estar configurado para interactuar con un dispositivo de visualización externo. En otros ejemplos, el dispositivo de destino puede ser un dispositivo de visualización 14. En general, el dispositivo de visualización 31 muestra los datos de vídeo decodificados a un usuario y puede comprender cualquiera de una variedad de dispositivos de visualización tales como una pantalla de cristal líquido (LCD), una pantalla de plasma, una pantalla de diodo emisor de luz orgánica (OLED), u otro tipo de dispositivo de visualización.

El codificador de vídeo 20 y el decodificador de vídeo 30 pueden funcionar según un estándar de compresión de vídeo, como el estándar de codificación de vídeo avanzada o el estándar de codificación de vídeo de alta eficiencia (HEVC) y las extensiones actualmente en desarrollo, y pueden ajustarse al modelo de prueba de HEVC (HM). En particular, el codificador de vídeo 20 y el decodificador de vídeo 30 pueden funcionar de acuerdo con la norma de codificación AVC basada en la codificación de múltiples vistas más profundidad (MVC+D) y la norma de codificación de vídeo 3D basada en AVC (3D-AVC) actualmente en desarrollo por JCT-3V, como se ha indicado anteriormente. De manera alternativa o adicional, el codificador de vídeo 20 y el decodificador de vídeo 30 pueden funcionar según otros estándares de propiedad o de la industria, como el estándar ITU-T H.264, también conocido como MPEG-4, Parte 10 o extensiones de tales estándares. Las técnicas de esta divulgación, sin embargo, no se limitan a ningún estándar de codificación particular. Otros ejemplos de normas de compresión de vídeo son MPEG-2 y UIT-T H.263, así como formatos abiertos como VP8.

En algunos aspectos, el codificador de vídeo 20 y el decodificador de vídeo 30 pueden estar integrados cada uno con un codificador y decodificador de audio, y pueden incluir unidades MUX-DEMUX adecuadas, u otro hardware y software, para manejar la codificación de audio y vídeo en un flujo de datos común o flujos de datos separados. Si corresponde, en algunos ejemplos, las unidades MUX-DEMUX pueden ajustarse al protocolo multiplexor ITU H.223 u otros protocolos como el protocolo de datagramas de usuario (UDP).

El codificador de vídeo 20 y el decodificador de vídeo 30 pueden implementarse cada uno como cualquiera de una variedad de circuitos de codificador adecuados, como uno o más microprocesadores, procesadores de señales digitales (DSP), circuitos integrados de aplicación específica (ASIC), matrices de puertas programables en campo (FPGA), lógica discreta, software, hardware, firmware o cualquier combinación de estos. Cuando las técnicas se implementan parcialmente en software, un dispositivo puede almacenar instrucciones para el software en un medio legible por ordenador adecuado y no transitorio y ejecutar las instrucciones en hardware usando uno o más procesadores para realizar las técnicas de esta divulgación. Cada codificador de vídeo 20 y decodificador de vídeo 30 se puede incluir en uno o más codificadores o decodificadores, cualquiera de los cuales puede integrarse como parte de un codificador/decodificador combinado (CODEC) en un dispositivo respectivo.

El JCT-VC está trabajando en el desarrollo de extensiones del estándar HEVC. Los esfuerzos de estandarización de HEVC se basan en un modelo en evolución de un dispositivo de codificación de vídeo denominado modelo de prueba de HEVC (HM). El HM presupone varias capacidades adicionales de los dispositivos de codificación de vídeo en relación con los dispositivos existentes según, por ejemplo, ITU-T H.264/AVC.

Algunos detalles de la codificación de vídeo se describen aquí en el contexto de HEVC. Sin embargo, las técnicas pueden aplicarse a UIT-T H.264/AVC, y más concretamente a la codificación de la vista de textura y/o de la vista de profundidad consistente con un formato de archivo multimedia base ISO (ISOBMFF).

El modelo de trabajo del HM describe que un fotograma de vídeo o una imagen se puede dividir en una secuencia de macrobloques, bloques de árbol o unidades de codificación más grandes (LCU) que incluyen tanto muestras de luminancia como de crominancia. Un bloque de árbol puede servir para algunos propósitos similares a los de un macrobloque del estándar H.264, aunque un bloque de árbol tiene muchas diferencias con respecto a un macrobloque.

La FIG. 4 es un diagrama esquemático que ilustra un ejemplo de codificador de vídeo 20 que puede implementar las técnicas descritas en esta divulgación. El codificador de vídeo 20 puede componer un archivo multimedia que incluya datos de vídeo de múltiples vistas basado en el formato de archivo multimedia base ISO, como el archivo multimedia de múltiples vistas con profundidad 142 de la FIG. 1. El codificador de vídeo 20 puede realizar intra e intercodificación de bloques de vídeo dentro de rebanadas de vídeo. La intracodificación se basa en la predicción espacial para reducir o eliminar la redundancia espacial en el vídeo dentro de un fotograma o imagen de vídeo determinado. La intercodificación se basa en la predicción temporal para reducir o eliminar la redundancia temporal en el vídeo dentro de fotogramas o imágenes adyacentes de una secuencia de vídeo. El intramodo (modo I) puede referirse a cualquiera de varios modos de compresión basados en el espacio. Los intermodos, como la predicción unidireccional (modo P) o la predicción bidireccional (modo B), pueden referirse a cualquiera de varios modos de compresión basados en el tiempo.

En el ejemplo de la FIG. 4, el codificador de vídeo 20 incluye una unidad de partición 35, una unidad de procesamiento de predicción 41, una unidad de filtro 63, una memoria de imagen de referencia 64, un sumador 50, una unidad de procesamiento de transformación 52, una unidad de cuantificación 54 y una unidad de codificación de entropía 56. La unidad de procesamiento de predicción 41 incluye la unidad de estimación de movimiento 42, la unidad de compensación de movimiento 44 y la unidad de intrapredicción 46. Para la reconstrucción de bloques de vídeo, el codificador de vídeo 20 también incluye la unidad de cuantificación inversa 58, la unidad de transformada inversa 60 y el sumador 62. La unidad de filtro 63 puede representar generalmente uno o más filtros de bucle como un filtro de desbloqueo, un filtro de bucle adaptativo (ALF) y un filtro de compensación adaptativo de muestra (SAO). Si bien la unidad de filtro 63 se muestra en la FIG. 4 como un filtro de bucle, en otras configuraciones, la unidad de filtro 63 puede implementarse como un filtro de bucle posterior, en cuyo caso los datos no filtrados pueden utilizarse en el bucle de codificación.

Según lo mostrado en la FIG. 4, el codificador de vídeo 20 recibe datos de vídeo, y la unidad de partición 35 divide los datos en bloques de vídeo. La partición también puede incluir la partición en rebanadas, mosaicos u otras unidades más grandes, así como la partición de bloques de vídeo, por ejemplo, de acuerdo con macrobloques y particiones de macrobloques, o posiblemente según una estructura de cuatro árboles de LCU y CU. El codificador de vídeo 20 en general ilustra los componentes que codifican bloques de vídeo dentro de una rebanada de vídeo a codificar. La rebanada se puede dividir en múltiples bloques de vídeo (y posiblemente en conjuntos de bloques de vídeo denominados mosaicos). La unidad de procesamiento de predicción 41 puede seleccionar uno de una pluralidad de modos de codificación posibles, como uno de una pluralidad de modos de intracodificación o uno de una pluralidad de modos de intercodificación, para el bloque de vídeo actual basándose en resultados de error (por ejemplo, tasa de codificación y nivel de distorsión). La unidad de procesamiento de predicción 41 puede proporcionar el bloque intra o intercodificado resultante al sumador 50 para generar datos de bloque residuales y al sumador 62 para reconstruir el bloque codificado para su uso como imagen de referencia.

La unidad de predicción interna 46 dentro de la unidad de procesamiento de predicción 41 puede realizar una codificación intrapredictiva del bloque de vídeo actual en relación con uno o más bloques vecinos en el mismo fotograma o rebanada que el bloque actual a codificar para proporcionar compresión espacial. La unidad de estimación de movimiento 42 y la unidad de compensación de movimiento 44 dentro de la unidad de procesamiento de predicción 41 realizan una codificación interpredictiva del bloque de vídeo actual con respecto a uno o más bloques predictivos en una o más imágenes de referencia para proporcionar compresión temporal.

La unidad de estimación de movimiento 42 puede configurarse para determinar el modo de interpredicción para una rebanada de vídeo de acuerdo con un patrón predeterminado para una secuencia de vídeo. El patrón predeterminado puede designar rebanadas de vídeo en la secuencia como rebanadas P, rebanadas B o rebanadas GPB. La unidad de estimación de movimiento 42 y la unidad de compensación de movimiento 44 pueden estar muy integradas, pero se ilustran por separado con fines conceptuales. La estimación de movimiento, realizada por la unidad 42 de estimación de movimiento, es el proceso de generar vectores de movimiento, que estiman el movimiento para bloques de vídeo. Un vector de movimiento, por ejemplo, puede indicar el desplazamiento de una PU (o una partición de macrobloque o subbloque) de un bloque de vídeo dentro de una imagen o fotograma de vídeo actual con respecto a un bloque predictivo dentro de una imagen de referencia.

Un bloque predictivo es un bloque que coincide estrechamente con la PU (o una partición de macrobloque o subbloque) del bloque de vídeo que se codificará en términos de diferencia de píxeles, que puede determinarse mediante la suma de la diferencia absoluta (SAD), la suma de la diferencia cuadrada (SSD) u otras métricas de diferencia. En algunos ejemplos, el codificador de vídeo 20 puede calcular valores para posiciones de píxeles subenteros de imágenes de referencia almacenadas en la memoria de la imagen de referencia 64. Por ejemplo, el codificador de vídeo 20 puede interpolar valores de posiciones de un cuarto de píxel, posiciones de un octavo de píxel u otras posiciones de píxeles fraccionarios de la imagen de referencia. Por lo tanto, la unidad de estimación de movimiento 42 puede realizar una búsqueda de movimiento en relación con las posiciones de píxeles completos y las posiciones de píxeles fraccionarios y generar un vector de movimiento con precisión de píxeles fraccionarios.

La unidad de estimación de movimiento 42 calcula un vector de movimiento para una PU (o una partición de macrobloque o subbloque) de un bloque de vídeo en una rebanada intercodificado mediante la comparación de la

posición de la PU (o una partición de macrobloque o subbloque) con la posición de un bloque predictivo de una imagen de referencia. La imagen de referencia puede seleccionarse de una primera lista de imágenes de referencia (Lista 0) o una segunda lista de imágenes de referencia (Lista 1), cada una de las cuales identifica una o más imágenes de referencia almacenadas en la memoria de la imagen de referencia 64. La unidad de estimación de movimiento 42 envía el vector de movimiento calculado a la unidad de codificación por entropía 56 y a la unidad de compensación de movimiento 44.

La compensación de movimiento, realizada por la unidad de compensación de movimiento 44, puede implicar buscar o generar el bloque predictivo en base al vector de movimiento determinado por la estimación de movimiento, posiblemente realizando interpolaciones con precisión de subpíxeles. Al recibir el vector de movimiento para la PU (o una partición de macrobloque o subbloque) del bloque de vídeo actual, la unidad 44 de compensación de movimiento puede localizar el bloque predictivo al que apunta el vector de movimiento en una de las listas de imágenes de referencia. El codificador de vídeo 20 forma un bloque de vídeo residual restando los valores de píxeles del bloque predictivo de los valores de píxeles del bloque de vídeo actual que se está codificando, formando valores de diferencia de píxeles. Los valores de diferencia de píxeles forman datos residuales para el bloque y pueden incluir componentes de diferencia de luminancia y crominancia. El sumador 50 representa el componente o componentes que realizan esta operación de resta. La unidad de compensación de movimiento 44 también puede generar elementos de sintaxis asociados con los bloques de vídeo y la rebanada de vídeo para su uso por el decodificador de vídeo 30 al decodificar los bloques de vídeo de la rebanada de vídeo.

La unidad de intrapredicción 46 puede intrapredecir un bloque actual, como una alternativa a la interpredicción realizada por la unidad de estimación de movimiento 42 y la unidad de compensación de movimiento 44, como se describió anteriormente. En particular, la unidad de intrapredicción 46 puede determinar un modo de intrapredicción para su uso para codificar un bloque actual. En algunos ejemplos, la unidad de intrapredicción 46 puede codificar un bloque actual usando varios modos de intrapredicción, por ejemplo, durante pases de codificación separados, y la unidad de intrapredicción 46 (o la unidad de selección de modo 40, en algunos ejemplos) puede seleccionar un modo de intrapredicción adecuado para utilizar los modos evaluados. Por ejemplo, la unidad de intrapredicción 46 puede calcular valores de distorsión de velocidad usando un análisis de distorsión de velocidad para los diversos modos de intrapredicción evaluados, y seleccionar el modo de intrapredicción que tiene las mejores características de distorsión de velocidad entre los modos evaluados. El análisis de distorsión de velocidad generalmente determina una cantidad de distorsión (o error) entre un bloque codificado y un bloque no codificado original que fue codificado para producir el bloque codificado, así como una tasa de bits (es decir, una cantidad de bits) que se usa para producir el bloque codificado. La unidad de intrapredicción 46 puede calcular las relaciones a partir de las distorsiones y velocidades para los diversos bloques codificados para determinar qué modo de intrapredicción exhibe el mejor valor de distorsión de velocidad para el bloque.

En cualquier caso, después de seleccionar un modo de intrapredicción para un bloque, la unidad de intrapredicción 46 puede proporcionar información indicativa del modo de intrapredicción seleccionado al bloque a la unidad de codificación de entropía 56. La unidad de codificación de entropía 56 puede codificar la información que indica el modo de intrapredicción seleccionado de acuerdo con las técnicas de esta divulgación. El codificador de vídeo 20 puede incluir en los datos de configuración del flujo de bits transmitidos, que pueden incluir una pluralidad de tablas de índice de modo de intrapredicción y una pluralidad de tablas de índice de modo de intrapredicción modificadas (también denominadas tablas de mapeo de palabras de código), definiciones de contextos de codificación para varios bloques, e indicaciones de un modo de intrapredicción más probable, una tabla de índice de modo de intrapredicción y una tabla de índice de modo de intrapredicción modificada para usar para cada uno de los contextos.

Después de que la unidad de procesamiento de predicción 41 genera el bloque predictivo para el bloque de vídeo actual, ya sea mediante interpredicción o intrapredicción, el codificador de vídeo 20 forma un bloque de vídeo residual restando el bloque predictivo del bloque de vídeo actual. Los datos de vídeo residuales en el bloque residual pueden aplicarse a la unidad de procesamiento de transformación 52. La unidad de procesamiento de transformación 52 transforma los datos de vídeo residuales en coeficientes de transformación residuales utilizando una transformada, como una transformada de coseno discreta (DCT) o una transformada conceptualmente similar. La unidad de procesamiento de transformación 52 puede convertir los datos de vídeo residuales de un dominio de píxeles a un dominio de transformación, tal como un dominio de frecuencia.

La unidad de procesamiento de transformación 52 puede enviar los coeficientes de transformación resultantes a la unidad de cuantificación 54. La unidad de cuantificación 54 cuantifica los coeficientes de transformada para reducir aún más la tasa de bits. El proceso de cuantificación puede reducir la profundidad de bits asociada con algunos o todos los coeficientes. El grado de cuantificación se puede modificar ajustando un parámetro de cuantificación. En algunos ejemplos, la unidad de cuantificación 54 puede luego realizar un análisis de la matriz que incluye los coeficientes de transformación cuantificados. Alternativamente, la unidad de codificación por entropía 56 puede realizar el análisis.

Después de la cuantificación, la entropía de la unidad de codificación de entropía 56 codifica los coeficientes de transformada cuantificados. Por ejemplo, la unidad de codificación de entropía 56 puede realizar codificación de longitud variable adaptativa al contexto (CAVLC), codificación aritmética binaria adaptativa al contexto (CABAC),

codificación aritmética binaria adaptativa al contexto basada en sintaxis (SBAC), codificación de entropía de partición de intervalo de probabilidad (PIPE) u otra metodología o técnica de codificación de entropía. Después de la codificación de entropía mediante la unidad de codificación de entropía 56, el flujo de bits codificado puede transmitirse al decodificador de vídeo 30 o archivar para su posterior transmisión o recuperación mediante el decodificador de vídeo 30. La unidad de codificación de entropía 56 también puede codificar la entropía en los vectores de movimiento y los otros elementos de sintaxis para la rebanada de vídeo actual que se codifica.

La unidad de cuantificación inversa 58 y la unidad de transformación inversa 60 aplican cuantificación inversa y transformación inversa, respectivamente, para reconstruir el bloque residual en el dominio de píxeles para su uso posterior como un bloque de referencia de una imagen de referencia. La unidad de compensación de movimiento 44 puede calcular un bloque de referencia añadiendo el bloque residual a un bloque predictivo de una de las imágenes de referencia dentro de una de las listas de imágenes de referencia. La unidad de compensación de movimiento 44 también puede aplicar uno o más filtros de interpolación al bloque residual reconstruido para calcular valores de píxeles subenteros para su uso en la estimación de movimiento. El sumador 62 añade el bloque residual reconstruido al bloque de predicción compensado por movimiento producido por la unidad de compensación de movimiento 44 para producir un bloque de referencia para el almacenamiento en la memoria de la imagen de referencia 64. El bloque de referencia puede ser utilizado por la unidad de estimación de movimiento 42 y la unidad de compensación de movimiento 44 como un bloque de referencia para interpredecir un bloque en un fotograma de vídeo o imagen posterior.

El codificador de vídeo 20 de la FIG. 4 representa un ejemplo de un codificador de vídeo configurado para implementar una o más de las técnicas descritas en el presente documento. Por ejemplo, el módulo de encapsulación 21 del codificador de vídeo 20 puede componer un formato de archivo multimedia a partir de flujos de vídeo que contienen múltiples vistas codificadas. Por ejemplo, el módulo de encapsulación 21 puede componer un archivo multimedia que incluye un cuadro del identificador de vista, como el cuadro del identificador de vista 172 de la FIG. 1, que indica, para cada una de las vistas incluidas en la pista, la presencia de textura y/o profundidad en la pista y en el flujo. El módulo de encapsulación 21 puede incluir el cuadro del identificador de vista 172 en una entrada de muestra o en una entrada de grupo de múltiples vistas. En otro ejemplo, el módulo de encapsulación 21 puede componer un archivo multimedia de tal manera que el cuadro del identificador de vista 172 indique cómo se requiere una vista de textura y una vista de profundidad de una vista de referencia para decodificar una vista específica: si sólo se requiere la vista de textura, sólo se requiere la vista de profundidad, o se requieren ambas. Según lo mostrado en la FIG. 4, el módulo de encapsulación 21 forma parte del codificador de vídeo 20. Sin embargo, en otros ejemplos, el módulo de encapsulación 21 está separado del codificador de vídeo 20.

Además, el codificador de vídeo 20 puede soportar el almacenamiento de una vista de textura y una vista de profundidad de una vista particular en pistas separadas. En algunos ejemplos, el codificador de vídeo 20 está configurado para utilizar e interpretar uno o más tipos nuevos de referencias de pista. Por ejemplo, el codificador de vídeo 20 puede utilizar un nuevo tipo de referencia de pista con valor de tipo 'deps', cuya presencia puede indicar que la pista actual contiene la vista de profundidad que está asociada con una vista de textura en la pista de referencia. En un ejemplo alternativo, el codificador de vídeo 20 puede utilizar otro nuevo tipo de referencia de pista con valor de tipo "tref", cuya presencia en una pista indica que la pista actual depende de la vista de textura almacenada en la pista referenciada. En otro ejemplo alternativo, se define un nuevo tipo de referencia de pista, por ejemplo, con un valor de tipo 'dref', cuya presencia en una pista indica que la pista actual depende de la vista de profundidad almacenada en la pista de referencia.

El codificador de vídeo 20 puede admitir la señalización de la resolución espacial de las vistas de profundidad. Por ejemplo, el codificador de vídeo 20 puede señalar la resolución espacial de las vistas de profundidad en un cuadro de resolución de profundidad de 3VC recién definido, que puede incluirse en un cuadro de configuración de 3VC. En algunos ejemplos, el cuadro de configuración de 3VC incluye el registro de configuración del decodificador 3VC y puede incluirse en una entrada de muestra (y debe incluirse en una entrada de muestra de tipo "3vc1" o "3vc2"). En algunos ejemplos, el CuadroResoluciónProfundidad3VC se añade directamente en la entrada de muestra. En otro ejemplo alternativo, el codificador de vídeo 20 señala directamente la anchura y la altura de las vistas de profundidad en el registro de configuración del decodificador 3VC. Por ejemplo, el codificador de vídeo 20 puede componer un archivo multimedia, como el archivo multimedia de múltiples vistas con profundidad 142 de la FIG. 1, y añadir un cuadro de resolución de profundidad de 3VC, tal como el cuadro de resolución de profundidad de 3VC 170, en una entrada de muestra. En otro ejemplo alternativo, el codificador de vídeo 20 puede señalar una anchura y una altura de las vistas de profundidad directamente en un registro de configuración del decodificador 3VC, que puede incluirse en un cuadro de configuración de 3VC 168.

El codificador de vídeo 20 también puede utilizar tipos de entrada de muestra recientemente definidos, '3vc1' y '3vc2', para pistas de 3VC con unidades NAL de profundidad solamente, donde no se permiten extractores y se permiten, respectivamente. El codificador de vídeo 20 puede incluir, además, en un archivo multimedia codificado, un cuadro de mensaje de SEI de información de escalabilidad de 3VC para contener el mensaje de SEI de información de escalabilidad de 3VC definido en la especificación del códec MVC+D.

En otro ejemplo alternativo, el codificador de vídeo 20 incluye un atributo común y/o un atributo diferenciador en un archivo multimedia compuesto que indica si los puntos de operación de 3VC indicados en un cuadro de grupo de

múltiples vistas (como el cuadro de grupo de múltiples vistas 174) están presentes, si un punto de operación indicado en un cuadro de grupo de múltiples vistas es 3VC o MVC, o si las vistas de salida indicadas en un cuadro de grupo de múltiples vistas contienen sólo textura, sólo profundidad, o ambas.

El codificador de vídeo 20 puede estar configurado para emplear una o más de las técnicas descritas en el presente documento para indicar si una vista incluida en una pista consiste en vistas de textura, vistas de profundidad, o ambas vistas de textura y profundidad, indicar la dependencia de la vista, señalar la resolución espacial de las vistas de profundidad sin requerir el análisis de un SPS, y mejorar la eficiencia y la flexibilidad del almacenamiento de flujos de vídeo que contienen múltiples vistas codificadas con profundidad.

En referencia nuevamente a la FIG. 3, el módulo de encapsulación 21 puede recibir contenido de vídeo codificado del codificador de vídeo 20 y generar un archivo de vídeo de acuerdo con las técnicas aquí descritas. En un ejemplo, el módulo de encapsulación 21 recibe contenido de vídeo codificado AVC o HEVC y genera un archivo de vídeo utilizando un formato de archivo basado en el ISOBMFF y la versión actual del AVCFF o HEVCFF, respectivamente.

La FIG. 5 es un diagrama de bloques que ilustra un ejemplo de decodificador de vídeo 30 y un módulo de desencapsulación 29 que puede implementar las técnicas de esta divulgación. El módulo de desencapsulación 29 puede analizar un archivo multimedia que incluya datos de vídeo de múltiples vistas basado en el formato de archivo multimedia base ISO, como el archivo multimedia de múltiples vistas con profundidad 142 de la FIG. 1. El módulo de desencapsulación 29 puede proporcionar el archivo multimedia analizado al decodificador de vídeo 30.

En el ejemplo de la FIG. 5, el decodificador de vídeo 30 incluye una unidad de codificación de entropía 80, una unidad de procesamiento de predicción 81, una unidad de cuantificación inversa 86, una unidad de transformación inversa 88, un sumador 90, una unidad de filtro 91 y una memoria de imagen de referencia 92. La unidad de procesamiento de predicción 81 incluye la unidad de compensación de movimiento 82 y la unidad de intrapredicción 84. El decodificador de vídeo 30 puede, en algunos ejemplos, realizar un pase de decodificación generalmente recíproco al pase de codificación descrito con respecto al codificador de vídeo 20 de la FIG. 3.

Durante el proceso de decodificación, el decodificador de vídeo 30 recibe un flujo de bits de vídeo codificado que representa bloques de vídeo de una rebanada de vídeo codificado y elementos de sintaxis asociados del codificador de vídeo 20. El decodificador de vídeo 30 puede recibir el flujo de bits de vídeo codificado de una entidad de red 29. La entidad de red 29 puede, por ejemplo, ser un servidor, un elemento de red consciente de los medios (MANE), un editor/empalmador de vídeo, u otro dispositivo similar configurado para implementar una o más de las técnicas descritas anteriormente. Como se ha descrito anteriormente, algunas de las técnicas descritas en esta divulgación pueden ser implementadas por la entidad de red 29 antes de que la red 29 transmita el flujo de bits de vídeo codificado al decodificador de vídeo 30. En algunos sistemas de decodificación de vídeo, la entidad de red 29 y el decodificador de vídeo 30 pueden ser partes de dispositivos separados, mientras que, en otros casos, la funcionalidad descrita con respecto a la entidad de red 29 puede ser realizada por el mismo dispositivo que comprende el decodificador de vídeo 30.

La unidad de codificación de entropía 80 del decodificador de vídeo 30 decodifica el flujo de bits para generar coeficientes cuantificados, vectores de movimiento y otros elementos de sintaxis. La unidad de codificación de entropía 80 envía los vectores de movimiento y otros elementos de sintaxis a la unidad de procesamiento de predicción 81. El decodificador de vídeo 30 puede recibir los elementos de sintaxis en el nivel de rebanada de vídeo y/o en el nivel de bloque de vídeo.

Cuando la rebanada de vídeo se codifica como una rebanada intracodificada (I), la unidad de intrapredicción 84 de la unidad de procesamiento de predicción 81 puede generar datos de predicción para un bloque de vídeo de la rebanada de vídeo actual en base a un modo de intrapredicción señalado y datos de bloques previamente decodificados del fotograma o imagen actual. Cuando el fotograma de vídeo se codifica como una rebanada intercodificada (es decir, B, P o GPB), la unidad de compensación de movimiento 82 de la unidad de procesamiento de predicción 81 produce bloques predictivos para un bloque de vídeo de la rebanada de vídeo actual basado en los vectores de movimiento y otros elementos de sintaxis recibidos de la unidad de codificación de entropía 80. Los bloques predictivos pueden producirse a partir de una de las imágenes de referencia dentro de una de las listas de imágenes de referencia. El decodificador de vídeo 30 puede construir las listas de tramas de referencia, Lista 0 y Lista 1, usando técnicas de construcción predeterminadas basadas en imágenes de referencia almacenadas en la memoria de la imagen de referencia 92.

La unidad de compensación de movimiento 82 determina información de predicción para un bloque de vídeo de la rebanada de vídeo actual analizando los vectores de movimiento y otros elementos de sintaxis, y usa la información de predicción para producir los bloques predictivos para el bloque de vídeo actual que se está decodificando. Por ejemplo, la unidad de compensación de movimiento 82 usa algunos de los elementos de sintaxis recibidos para determinar un modo de predicción (por ejemplo, intra o interpredicción) usado para codificar los bloques de vídeo de la rebanada de vídeo, un tipo de rebanada de interpredicción (por ejemplo, la rebanada B o la rebanada P o la rebanada GPB), información de construcción para una o más de las listas de imágenes de referencia para la rebanada, vectores de movimiento para cada bloque de vídeo intercodificado de la rebanada, estado de interpredicción para cada bloque

de video intercodificado de la rebanada, y otra información para decodificar los bloques de video en la rebanada de video actual.

La unidad de compensación de movimiento 82 también puede realizar una interpolación basada en filtros de interpolación. La unidad de compensación de movimiento 82 puede usar filtros de interpolación como los usa el codificador de video 20 durante la codificación de los bloques de video para calcular valores interpolados para píxeles subenteros de bloques de referencia. En este caso, la unidad de compensación de movimiento 82 puede determinar los filtros de interpolación usados por el codificador de video 20 a partir de los elementos de sintaxis recibidos y usar los filtros de interpolación para producir bloques predictivos.

La unidad de cuantificación inversa 86 cuantifica inversamente, es decir, descuantifica, los coeficientes de transformación cuantificados proporcionados en el flujo de bits y decodificados por la unidad de codificación de entropía 80. El proceso de cuantificación inversa puede incluir el uso de un parámetro de cuantificación calculado por el codificador de video 20 para cada bloque de video en la rebanada de video para determinar un grado de cuantificación e, igualmente, un grado de cuantificación inversa que debería aplicarse. La unidad de transformada inversa 88 aplica una transformación inversa, por ejemplo, una DCT inversa, una transformada inversa de número entero o un proceso de transformada inversa conceptualmente similar, a los coeficientes de transformada para producir bloques residuales en el dominio de píxeles.

Después de que la unidad de compensación de movimiento 82 genera el bloque predictivo para el bloque de video actual basado en los vectores de movimiento y otros elementos de sintaxis, el decodificador de video 30 forma un bloque de video decodificado sumando los bloques residuales de la unidad de transformada inversa 88 con los correspondientes bloques predictivos generados por la unidad de compensación de movimiento 82. El sumador 90 representa el componente o componentes que realizan esta operación de suma. Si se desea, también se pueden utilizar otros filtros de bucle (ya sea en el bucle de codificación o después del bucle de codificación) para suavizar las transiciones de píxeles o mejorar de otro modo la calidad del video. La unidad de filtro 91 puede representar uno o más filtros de bucle como un filtro de desbloqueo, un filtro de bucle adaptativo (ALF) y un filtro de compensación adaptativo de muestra (SAO). Si bien la unidad de filtro 91 se muestra en la FIG. 5 como un filtro de bucle, en otras configuraciones, la unidad de filtro 91 puede implementarse como un filtro de bucle posterior. Los bloques de video decodificados en una trama o imagen determinadas se almacenan luego en la memoria de la imagen de referencia 92, que almacena imágenes de referencia utilizadas para la compensación de movimiento posterior. La memoria de imagen de referencia 92 también almacena video decodificado para una presentación posterior en un dispositivo de visualización, tal como el dispositivo de visualización 31 de la FIG. 1.

El decodificador de video 30 de la FIG. 5 representa un ejemplo de decodificador de video que puede utilizarse con un módulo de desencapsulación 29 configurado para implementar una o más de las técnicas descritas en el presente documento. Por ejemplo, el módulo de desencapsulación 29 puede analizar un archivo multimedia que incluye flujos de video almacenados que contienen múltiples vistas codificadas y proporcionarlo al decodificador de video 30. Por ejemplo, el módulo de desencapsulación 29 puede analizar un archivo multimedia que incluye un cuadro del identificador de vista, como el cuadro del identificador de vista 172 de la FIG. 1, que indica, para cada una de las vistas incluidas en la pista, la presencia de textura y/o profundidad en la pista y en el flujo. El módulo de desencapsulación 29 puede decodificar el cuadro del identificador de vista 172 de una entrada de muestra o en una entrada de grupo de múltiples vistas. En otro ejemplo, el decodificador de video 30 analiza el cuadro del identificador de vista 172 en un archivo multimedia para determinar si se requiere una vista de textura y una vista de profundidad de una vista de referencia para decodificar una vista específica. Es decir, el módulo de desencapsulación 29 puede determinar si sólo se requiere la vista de textura, sólo se requiere la vista de profundidad, o se requieren ambas.

Además, el módulo de desencapsulación 29 puede analizar un archivo multimedia que contenga una vista de textura y una vista de profundidad de una vista particular en pistas separadas. En algunos ejemplos, el módulo de desencapsulación 29 está configurado para analizar e interpretar uno o más tipos nuevos de referencias de pista. Por ejemplo, el módulo de desencapsulación 29 analiza un nuevo tipo de referencia de pista con valor de tipo "deps", cuya presencia indica que la pista actual contiene la vista de profundidad que está asociada con una vista de textura en una pista de referencia. En un ejemplo alternativo, el módulo de desencapsulación 29 analiza otro nuevo tipo de referencia de pista con valor de tipo "tref", cuya presencia en una pista indica que la pista actual depende de la vista de textura almacenada en la pista de referencia. En otro ejemplo alternativo, el módulo de desencapsulación 29 analiza un nuevo tipo de referencia de pista con valor de tipo "dref", cuya presencia en una pista indica que la pista actual depende de la vista de profundidad almacenada en la pista de referencia.

El módulo de desencapsulación 29 puede determinar las resoluciones espaciales de las vistas de profundidad a partir del análisis de un archivo multimedia compuesto según las técnicas descritas en el presente documento. Por ejemplo, el módulo de desencapsulación 29 puede determinar las resoluciones espaciales de las vistas de profundidad a partir del cuadro de resolución de profundidad de 3VC recién definido, que puede incluirse en un cuadro de configuración de 3VC. En otro ejemplo alternativo, el módulo de desencapsulación 29 determina directamente la anchura y la altura de las vistas de profundidad a partir del registro de configuración del decodificador 3VC. Por ejemplo, el módulo de desencapsulación 29 puede analizar un registro de configuración del decodificador 3VC de un archivo multimedia,

como el archivo multimedia de múltiples vistas con profundidad 142 de la FIG. 1 para determinar la anchura y la altura de las vistas de profundidad asociadas con el archivo multimedia.

El módulo de desencapsulación 29 también puede analizar los tipos de entrada de muestra recientemente definidos, "3vc1" y "3vc2", para las pistas de 3VC con unidades NAL de profundidad solamente. El módulo de desencapsulación 29 puede decodificar, además, a partir de un archivo multimedia codificado, un cuadro de mensaje de SEI de información de escalabilidad de 3VC para acceder al mensaje del SEI de información de escalabilidad de 3VC definido en la especificación del códec MVC+D.

En otro ejemplo alternativo, el módulo de desencapsulación 29 analiza un atributo común y/o un atributo diferenciador del archivo multimedia para determinar si los puntos de operación de 3VC indicados en un cuadro de grupo de múltiples vistas (como el cuadro de grupo de múltiples vistas 174) están presentes, si un punto de operación indicado en el cuadro de grupo de múltiples vistas 174 es 3VC o MVC, o si las vistas de salida indicadas en el cuadro de grupo de múltiples vistas 174 contienen sólo textura, sólo profundidad, o ambas.

El módulo de desencapsulación 29 puede estar configurado para emplear una o más de las técnicas descritas en el presente documento para determinar si una vista incluida en una pista consiste en vistas de textura, vistas de profundidad, o ambas vistas de textura y profundidad, determinar la dependencia de la vista, determinar la resolución espacial de las vistas de profundidad sin requerir el análisis de un SPS, y para mejorar la eficiencia y la flexibilidad del almacenamiento de flujos de vídeo que contienen múltiples vistas codificadas con profundidad.

La FIG. 6 es un diagrama de bloques que ilustra los componentes de un ejemplo de módulo de encapsulación 21. En el ejemplo de la FIG. 6, el módulo de encapsulación 21 incluye una interfaz de entrada de vídeo 70, un módulo de creación de archivos de vídeo 72 y una interfaz de salida de archivos de vídeo 76. El módulo de creación de archivos de vídeo 70, en este ejemplo, incluye el constructor de unidades de capa de abstracción de red (NAL) 74.

La interfaz de entrada de vídeo 70 puede recibir contenido de vídeo codificado. La interfaz de entrada de vídeo 70 puede recibir contenido de vídeo codificado del codificador de vídeo 20 o puede recuperar contenido de vídeo codificado de un dispositivo de almacenamiento. El contenido de vídeo codificado puede codificarse según el estándar de AVC o HEVC y puede incluir muestras de contenido de vídeo, conjuntos de parámetros y mensajes de SEI. Al recibir el contenido de vídeo codificado, la interfaz de entrada de vídeo 70 puede enviar el contenido de vídeo codificado al módulo de creación de archivos de vídeo 72 para su ensamblaje en un archivo multimedia. En algunos casos, la interfaz de entrada de vídeo 70 puede facilitar el ensamblaje de un archivo multimedia al organizar o almacenar en el búfer el contenido de vídeo codificado antes de que se emita al módulo de creación de archivos de vídeo 72.

En general, el módulo de creación de archivos de vídeo 72 puede generar un archivo de vídeo que incluya el contenido de vídeo codificado recibido. En un ejemplo, el módulo de creación de archivos de vídeo 72 puede crear un archivo de vídeo, como el archivo multimedia de múltiples vistas con profundidad 142 descrito con respecto a la FIG. 1. El módulo de creación de archivos de vídeo 72 puede corresponder a una unidad de control que incluye hardware, software y/o firmware configurado para realizar las funciones y procedimientos que se le atribuyen. La unidad de control puede realizar además las funciones atribuidas al módulo de encapsulación 21 en general. Para los ejemplos en los que el módulo de creación de archivos de vídeo 72 está incorporado en software y/o firmware, el módulo de encapsulación 21 puede incluir un medio legible por ordenador que comprende instrucciones para el módulo de creación de archivos de vídeo 72 y una unidad de procesamiento para ejecutar las instrucciones. Los submódulos del módulo de creación de archivos de vídeo 72 (constructor de unidades NAL 74 en este ejemplo) pueden implementarse como módulos de hardware individuales y/o módulos de software, y pueden integrarse funcionalmente o separarse en submódulos adicionales. El módulo de creación de archivos de vídeo 72 puede corresponder a cualquier unidad de procesamiento o circuito de procesamiento adecuado, como, por ejemplo, uno o más microprocesadores, circuitos integrados de aplicación específica (ASIC), matrices de puertas programables en campo (FPGA), procesadores de señales digitales (DSP), o cualquiera de sus combinaciones. El módulo de creación de archivos de vídeo 72 puede incluir además un medio legible por ordenador no transitorio que almacena instrucciones para el constructor de unidades NAL 74, así como un procesador para ejecutar las instrucciones.

En un ejemplo, el módulo de creación de archivos de vídeo 72 puede estar configurado para emitir datos de vídeo codificados de acuerdo con un formato de archivo que puede almacenar flujos de vídeo que contienen múltiples vistas codificadas con profundidad, en el que las propiedades y relaciones de las vistas de profundidad pueden ser señaladas en el archivo.

El módulo de creación de archivos de vídeo 72 puede estar configurado para implementar las técnicas descritas en la presente divulgación. Por ejemplo, el módulo de creación de archivos de vídeo 72 puede componer un archivo multimedia que tenga un nuevo valor del elemento de sintaxis Nombrecompresor. Nombrecompresor, definido en una clase base de entrada de muestra visual, puede indicar el nombre del compresor utilizado con el valor "\Codificación 0123VC" recomendado para un archivo que contenga flujos de vídeo de 3DV (como MVC+D). El elemento sintáctico Nombrecompresor puede definirse además según AVCFF o HEVCFF.

Además de la sintaxis y la semántica de ejemplo descritas anteriormente, un formato de archivo utilizado por el módulo de creación de archivos de vídeo 72 puede incluir restricciones adicionales con respecto a un registro de configuración del decodificador 3VC. Según las técnicas descritas en el presente documento, el elemento de sintaxis RegistroConfiguraciónDecodificadorMVC puede ser estructuralmente idéntico al elemento de sintaxis RegistroConfiguraciónDecodificadorAVC. Sin embargo, se redefinen uno o más bits reservados que preceden y suceden al campo tamañoLongitudMenosUno. Un ejemplo de sintaxis es el siguiente:

```
RegistroConfiguraciónDecodificadorMVC de clase alineada(8) {
  versión de configuración ent(8) sin signo = 1;
  IndicaciónPerfilAVC ent(8) sin signo;
  Compatibilidad_perfil ent(8) sin signo;
  IndicaciónNivelAVC ent(8) sin signo;
  representación completa bit(1);
  pista_au_explicita bit(1);
  bit(4) reservado = '1111'b;
  tamañoLongitudMenosUno ent(2) sin signo;
  bit(1) reservado = '0'b;
  númDeConjuntosParámetrosSecuencia ent(7) sin signo;
  para (i=0; i< númDeConjuntosParámetrosSecuencia; i++) {
    LongitudConjuntoParámetroSecuencia ent(16) sin signo ;
    bit(8*LongitudConjuntoParámetroSecuencia) UnidadNALConjuntoParámetroSecuencia;
  }
  númDeConjuntosParámetrosSecuencia ent(8) sin signo;
  para (i=0; i< númDeConjuntosParámetrosImagen; i++) {
    númDeConjuntosParámetrosImagen ent(16) sin signo;
    bit(8*LongitudConjuntosParámetrosImagen) UnidadNALConjuntoParámetroImagen;
  }
}
```

En algunos ejemplos, la estructura sintáctica de RegistroConfiguraciónDecodificador3VC puede ser exactamente la misma que la de RegistroConfiguraciónDecodificadorMVC. Cuando se utiliza un registro de configuración del decodificador AVC (como se define en la cláusula 5.3.3.1 de la norma ISO/IEC 14496-15) para un flujo que puede interpretarse como un flujo de 3VC, el registro de configuración del decodificador AVC puede reflejar las propiedades de la vista base compatible con AVC. Por ejemplo, el registro de configuración del decodificador AVC puede contener sólo los conjuntos de parámetros utilizados para decodificar la vista base AVC.

Cuando el registro de configuración del decodificador MVC (definido en la norma ISO/IEC 14496-15) se utiliza para un flujo que puede interpretarse como un flujo de MVC o 3VC, el registro de configuración del decodificador MVC deberá reflejar las propiedades de un subconjunto de flujo de bits compatible con MVC. Es decir, por ejemplo, el registro de configuración del decodificador MVC puede contener sólo los conjuntos de parámetros utilizados para decodificar el subconjunto de flujos de bits compatible con MVC. Un subconjunto compatible con MVC puede ser, por ejemplo, un subconjunto de flujo de bits con sólo las vistas de textura.

Un flujo de conjuntos de parámetros puede utilizarse con flujos de 3VC, al igual que con flujos de AVC o MVC. En este ejemplo, los conjuntos de parámetros no se incluirán en el registro de configuración del decodificador. Los conjuntos de parámetros de secuencia (SPS), incluidos los conjuntos de parámetros de secuencia de subconjuntos, pueden numerarse en orden de almacenamiento de 1 a númDeConjuntosParámetrosSecuencia o númDeConjuntosParámetrosImagen, respectivamente. Los SPS y los conjuntos de parámetros de imagen (PPS) almacenados en este registro en un archivo pueden consultarse al usar este índice basado en 1 mediante el CuadroConjuntoParámetrosInicial.

La semántica de los campos IndicaciónPerfilAVC, compatibilidad_perfil e IndicaciónNivelAVC puede diferir de la del RegistroConfiguraciónDecodificadorMVC de la siguiente manera. IndicaciónPerfilAVC, compatibilidad_perfil e IndicaciónNivelAVC, si son distintos de cero, deben tener valores tales que un decodificador 3VC conforme sea capaz de decodificar flujos de bits que se ajusten a los indicadores de perfil, nivel y compatibilidad de perfil indicados en cualquiera de los SPS o subconjuntos de SPS contenidos en este registro.

La semántica de otros campos puede ser la siguiente o, si no están presentes, la definida para un RegistroConfiguraciónDecodificadorMVC: númDeConjuntosParámetrosSecuencia indica el número de SPS y subconjuntos de SPS que se utilizan para decodificar el flujo elemental 3VC. El elemento sintáctico UnidadNALConjuntoParámetroSecuencia puede contener una unidad SPS o un subconjunto SPS NAL como se especifica en el anexo I de la norma ISO/IEC 14496-10. Los SPS se presentarán en orden ascendente del identificador del conjunto de parámetros y se permitirán espacios en blanco. Los SPS de subconjunto se presentarán en orden ascendente del identificador del conjunto de parámetros y se permitirán espacios en blanco. Cualquier SPS se producirá antes de todos los SPS de subconjunto, si los hay.

El módulo de creación de archivos de vídeo 72 también puede crear un archivo multimedia para pistas de 3VC con unidades AVC, MVC y NAL de profundidad. Los dos tipos de entradas de muestra se definen como "3vc1" y "3vc2". A continuación, se ofrecen ejemplos de definiciones de entradas de muestra para MVC y 3VC. Los tipos de entrada de muestra "avc1", "avc2", "mvc1", "mvc2", "3vc1" y "3vc2" pueden estar contenidos en un cuadro de tabla de muestra ("stbl"), como el cuadro de tabla de muestra 162 de la FIG. 1. En algunos ejemplos, es obligatorio tener uno de los tipos de entrada de muestra avc1, avc2, mvc1, mvc2, 3vc1 y 3vc2 en un archivo multimedia, como el archivo multimedia de múltiples vistas con profundidad 142. En algunos ejemplos, una o más entradas de muestra pueden estar presentes en el archivo multimedia de múltiples vistas con profundidad 142.

El módulo de creación de archivos de vídeo 72 puede incluir un cuadro de configuración de AVC, un cuadro de configuración de MVC y un cuadro de configuración de 3VC. Cuando está presente, un cuadro de configuración de AVC puede documentar la información sobre el perfil, el nivel y el conjunto de parámetros correspondientes a la vista base compatible con AVC definida por un RegistroConfiguraciónDecodificadorAVC. Cuando está presente, un cuadro de configuración de MVC puede documentar la información de perfil, nivel y conjunto de parámetros correspondiente a todo el flujo de MVC, tal como se define en el RegistroConfiguraciónDecodificadorMVC. Cuando está presente, un cuadro de configuración de 3VC, como el cuadro de configuración de 3VC 168 de la FIG. 1, puede documentar la información de perfil, nivel y conjunto de parámetros correspondiente a todo el flujo de 3VC, tal como se define en el RegistroConfiguraciónDecodificador3VC, almacenado en el cuadro de configuración de 3VC 168.

Para las entradas de muestra de AVC "avc1" y "avc2", los campos de anchura y altura incluidos en la entrada de muestra pueden documentar una capa base AVC. Para las entradas de muestra de MVC "mvc1" y "mvc2", así como para las entradas de muestra de 3VC "3vc1" y "3vc2", la anchura y la altura incluidos en la entrada de muestra pueden documentar la resolución alcanzada por la decodificación de cualquier vista de textura única de todo el flujo MVC o 3VC. En el caso de las entradas de muestra de 3VC, la anchura_profundidad y la altura_profundidad incluidos en el cuadro de resolución de profundidad de 3VC 170 pueden documentar la resolución alcanzada mediante la decodificación de cualquier vista de profundidad de todo el flujo de 3VC.

En algunos ejemplos, un campo tamañoLongitudMenosUno en las configuraciones de AVC, MVC y 3VC de cualquier entrada de muestra tendrá el mismo valor.

Un identificador uniforme de recursos (URI) de asignación de prioridad proporciona un nombre (en el espacio URI) de un método utilizado para asignar valores id_prioridad. Cuando se produce en una entrada de muestra de AVC, MVC o 3VC, deberá estar presente exactamente un URI que documente las asignaciones de id_prioridad en todo el flujo de AVC, MVC o 3VC. El URI se trata aquí sólo como un nombre; debería poder quitarse la referencia, aunque esto no es necesario. Los lectores de archivos pueden ser capaces de reconocer algunos métodos y, por lo tanto, saber lo que harían las operaciones de extracción de flujos basadas en la id_prioridad.

También pueden aplicarse aquí algunos requisitos de los tipos de entrada de muestra "avc1" y "avc2" documentados en ISO/IEC 14496-15, sección 6.5.3.1.1.

Cuando están presentes en una entrada de muestra de AVC, MVC o 3VC, CuadroSEIInfoEscalabilidadVista, CuadroIdentificadorVista (como el cuadro del identificador de vista 172), CuadroParámetrosIntrínsecosCámara, CuadroParámetrosExtrínsecosCámara, CuadroSEIInformaciónEscalabilidad3VC (como el cuadro SEI de información de escalabilidad de 3VC 166), CuadroTasaBitsMPEG4 y CuadroDescriptoresExtensiónMPEG4 se aplican a todo el flujo AVC, MVC o 3VC. Los tipos de cuadro de ejemplo que pueden incluirse en una o ambas entradas de muestra "3vc1" y "3vc2" incluyen un cuadro de parámetros de cámara intrínseca ("icam"), un cuadro de parámetros de cámara extrínseca ("ecam") y un cuadro del identificador de vista ("vwid").

Para un flujo MVC, el Nombrecompresor de la clase base EntradaMuestraVisual indica el nombre del compresor utilizado, con el valor "\Codificación MVC 012" (\012 es 10, la longitud de la cadena "codificación MVC" en bytes). Del mismo modo, para un flujo de 3VC, el Nombrecompresor de la clase base EntradaMuestraVisual indica el nombre del compresor utilizado, con el valor "\Codificación VC 0123" (\012 es 10, la longitud de la cadena "Codificación 3VC" en bytes).

En algunos ejemplos, uno o más conjuntos de parámetros necesarios para decodificar una unidad NAL que está presente en los datos de muestra de un flujo de vídeo, ya sea directamente o por referencia de un extractor, estará presente en la configuración del decodificador de ese flujo de vídeo o en el flujo de conjuntos de parámetros asociado (si se utiliza).

La Tabla 2 proporciona un ejemplo de los posibles usos de las entradas de muestra para las pistas de AVC, MVC y 3VC. Es decir, la tabla 2 muestra los posibles usos de las entradas de muestra cuando un flujo elemental AVC, MVC o 3VC se almacena en una o más pistas, configuraciones y las herramientas de MVC que también pueden utilizarse (a excepción de los metadatos temporizados, que siempre se utilizan en otra pista). La tabla 2 incluye el tipo de cuadro de cuatro caracteres de las entradas de muestra, una indicación de qué registros de configuración pertenecen a las entradas de muestra y un significado de las entradas de muestra.

Tabla 2 - Uso de las entradas de muestra para las pistas de AVC, MVC y 3VC

nombre de entrada de muestra	con registros de configuración	Significado
'avc1'	Solo configuración de AVC	Una pista de AVC simple con unidades NAL de AVC únicamente; no habrá extractores, agregadores ni agrupación de capas.
'avc1'	Configuraciones de AVC y MVC	Una pista de MVC con unidades NAL de AVC y MVC; puede haber agregadores y extractores; los agregadores no contendrán, pero pueden hacer referencia a unidades NAL de AVC; puede haber agrupación de capas.
'avc1'	AVC, MVC y 3VC Configuraciones	Una pista de 3VC con unidades NAL de AVC, MVC y profundidad; puede haber agregadores y extractores; los agregadores no contendrán, pero pueden hacer referencia a unidades NAL de AVC; puede haber agrupación de capas.
'avc2'	Solo configuración de AVC	Una pista de AVC simple con unidades NAL de AVC solamente; puede haber extractores; puede haber agregadores para contener y consultar unidades NAL de AVC; puede haber agrupación de capas.
'avc2'	Configuraciones de AVC y MVC	Una pista de MVC con unidades NAL de AVC y unidades NAL de MVC; los extractores pueden estar presentes y utilizarse para consultar unidades NAL de AVC y de MVC; los agregadores pueden estar presentes para contener y consultar unidades NAL de AVC y de MVC; la agrupación de niveles puede estar presente.
'mvc1'	Solo configuración de MVC	Una pista de MVC sin unidades NAL de AVC; puede haber agregadores y agrupación de capas.
'mvc1'	Configuraciones de MVC y 3VC	Una pista de 3VC sin unidades NAL de AVC, pero con unidades NAL de MVC y profundidad; puede haber agregadores y agrupación de capas.
'mvc2'	Solo configuración de MVC	Una pista de MVC sin unidades NAL de AVC; puede haber extractores, agregadores y agrupación de capas.
'mvc2'	Configuraciones de MVC y 3VC	Una pista de 3VC sin unidades NAL de AVC, pero con unidades NAL de MVC y profundidad; puede haber extractores, agregadores y agrupación de capas.
'3vc1'	Solo configuración de 3VC	Una pista de 3VC con unidades NAL de profundidad únicamente; puede haber agregadores y agrupación de capas.
'3vc2'	Solo configuración de 3VC	Una pista de 3VC con unidades NAL de profundidad únicamente; puede haber agregadores y agrupación de capas.

A continuación, se ofrece un ejemplo de estructura sintáctica para las entradas de muestra de MVC y 3VC:

```

5  clase CuadroConfiguraciónMVC extiende Cuadro('mvcC') {
    RegistroConfiguraciónDecodificadorMVC() ConfigMVC;
  }
10  clase CuadroSEInformaciónEscalabilidadVista extiende Cuadro('vsib', tamaño)
  {
    seiinfoescalvista ent(8*tamaño-64) sin signo;
    clase CuadroResoluciónProfundidad3VC extiende Cuadro('3dpr')
    {
15     anchura profundidad ent(16) sin signo;
    altura profundidad ent(16) sin signo;
    }
    clase CuadroConfiguración3VC extiende Cuadro('3vcC') {
      3 RegistroConfiguraciónDecodificadorVC 3 ConfigVC;
      CuadroResoluciónProfundidad3VC ResProfundidad3vc; //Opcional
20    }
    clase CuadroSEInformaciónEscalabilidad3VC extiende Cuadro('3sib', tamaño)
    {
      seiinfoescal3vc ent(8*tamaño-64) sin signo;
    }
25  clase EntradaMuestraAVCMVC() extiende EntradaMuestraAVC ('avc1' o 'avc2'){
    CuadroSEInformaciónEscalabilidadVista seiinfoescalmvc; // opcional
    identificadores vista CuadroIdentificadorVista; // opcional
    CuadroConfiguraciónMVC configmvc; // opcional
    CuadroAsignaciónPrioridadVistaMVC método _prioridad_vista; // opcional
30    CuadroParámetrosIntrínsecosCámara parám_intrinsec_cámara; // opcional
    CuadroParámetrosExtrínsecosCámara parám_extrinsec_cámara; // opcional
    CuadroConfiguración3VC config3vc; // opcional
    CuadroSEInformaciónEscalabilidad3VC seiinfoescal3vc; // opcional
  }
35  clase EntradaMuestraMVC() extiende EntradaMuestraVisual ('mvc1' o 'mvc2'){

```

```

CuadroConfiguraciónMVC configmvc; // obligatorio
CuadroSEIInformaciónEscalabilidadVista seiinfoescalmvc; // opcional
identificadores vista CuadroIdentificadorVista; // obligatorio
tasabits CuadroTasaBitsMPEG4; // opcional
5 CuadroDescriptoresExtensiónMPEG4 descr; // opcional
CuadroAsignaciónPrioridadVistaMVC método_prioridad_vista; // opcional
CuadroParámetrosIntrínsecosCámara parám_intrínsec_cámara; // opcional
CuadroParámetrosExtrínsecosCámara parám_extrínsec_cámara // opcional
CuadroConfiguración3VC config3vc; // opcional
10 CuadroSEIInformaciónEscalabilidad3VC seiinfoescal3vc; // opcional
}
clase EntradaMuestra3VC() extiende EntradaMuestraVisual ('3vc1' o '3vc2'){
CuadroConfiguración3VC config3vc; // obligatorio
CuadroSEIInformaciónEscalabilidad3VC seiinfoescal3vc; // opcional
15 CuadroIdentificadorVista identificadores_vista; // obligatorio
tasabits CuadroTasaBitsMPEG4; // opcional
CuadroDescriptoresExtensiónMPEG4 descr; // opcional
CuadroAsignaciónPrioridadVistaMVC método_prioridad_vista; // opcional
CuadroParámetrosIntrínsecosCámara parám_intrínsec_cámara; // opcional
20 CuadroParámetrosExtrínsecosCámara parám_extrínsec_cámara // opcional
}

```

A continuación, se ofrecen ejemplos de elementos sintácticos para las entradas de muestra de MVC y 3VC. Un elemento sintáctico anchura_profundidad y altura_profundidad puede dar, en píxeles, los valores de la anchura y la altura, respectivamente, de los componentes de la vista de profundidad codificada en el flujo al que se aplica la entrada de muestra que contiene el CuadroResoluciónProfundidad3VC() (como el Cuadro de resolución de profundidad de 3VC 170). Un elemento sintáctico ResProfundidad3vc puede contener la anchura y la altura de los componentes de la vista de profundidad codificada en el flujo al que se aplica la entrada de muestra. Cuando no está presente, se puede inferir que la anchura y la altura de los componentes de la vista de profundidad codificados son los mismos que la anchura y la altura de los componentes de la vista de textura codificados.

El elemento sintáctico seiinfoescalmvc puede contener una unidad NAL de SEI que sólo contenga un mensaje de SEI de información sobre escalabilidad de la vista, tal como se especifica en el anexo H de la norma ISO/IEC 14496-10. El campo "tamaño" de un cuadro contenedor CuadroSEIInformaciónEscalabilidadVista no deberá ser igual a 0 ni a 1.

Otro elemento sintáctico seiinfoescal3vc puede contener una unidad NAL de SEI que sólo contenga un mensaje de SEI de información sobre escalabilidad de 3DVC, tal como se especifica en el anexo I de la norma ISO/IEC 14496-10. El campo "tamaño" del cuadro contenedor CuadroSEIInformaciónEscalabilidad3VC no deberá ser igual a 0 ni a 1.

Para los flujos de vídeo de MVC+D, el módulo de creación de archivos de vídeo 72 también puede componer un cuadro de mensaje de SEI de información de escalabilidad 3VC, como el cuadro de mensaje de SEI de información de escalabilidad de 3VC 166 de la FIG. 1. El cuadro de mensaje de SEI de información de escalabilidad de 3VC 166 puede contener un mensaje de SEI de información de escalabilidad de 3VC definido en la especificación del códec MVC+D.

El módulo de creación de archivos de vídeo 72 también puede componer un cuadro del identificador de vista, como el cuadro del identificador de vista 172 de la FIG. 1, de acuerdo con las técnicas descritas en el presente documento. El módulo de creación de archivos de vídeo 72 puede componer el cuadro del identificador de vista 172 para indicar, para cada vista incluida en la pista, la presencia de vistas de textura y profundidad en la pista y en el flujo. Además, el módulo de creación de archivos de vídeo 72 puede componer el cuadro del identificador de vistas 172 para indicar cómo se requiere una vista de textura y una vista de profundidad de una vista de referencia para decodificar una vista específica.

A continuación, se ofrecen más detalles sobre el cuadro del identificador de vista 172. El tipo de cuadro codificado de cuatro caracteres para el cuadro del identificador de vista 172 puede ser "vwid". El cuadro del identificador de vista 172 puede estar contenido en cualquier entrada de muestra (por ejemplo, "avc1", "avc2", "mvc1", "mvc2", "3vc1" y "3vc2") o en una entrada multigrupo. En algunos ejemplos, el cuadro del identificador de vista tiene que incluirse exactamente una vez para cada una de las entradas de muestra 'mvc1', 'mvc2', '3vc1', '3vc2', y la definición del grupo primario en las entradas del grupo de múltiples vistas. En algunos ejemplos, no se incluye el cuadro del identificador de vistas 172 para las definiciones de grupos no primarios en las entradas de grupos de múltiples vistas.

Cuando se incluye en una entrada de muestra, el cuadro del identificador de vistas 172 indica las vistas incluidas en una pista. Cuando se incluye en una entrada de grupo de múltiples vistas, el cuadro del identificador de vistas 172 indica las vistas incluidas en una capa respectiva. El cuadro del identificador de vistas 172 también indica un índice de orden de vistas para cada vista enumerada. Además, el cuadro del identificador de vistas 172 incluye los valores mínimos y máximos de id_temporal incluidos en la pista o capa cuando el cuadro del identificador de vistas 172 se

incluye en una entrada de muestra o en una entrada de grupo de múltiples vistas, respectivamente. Además, el cuadro del identificador de vista 172 puede indicar una o más vistas referenciadas necesarias para la decodificación de las vistas incluidas en la pista o capa. En el caso de los flujos de 3VC, el cuadro del identificador de vista 172 puede indicar, para cada una de las vistas incluidas en la pista, la presencia de textura y/o profundidad en la pista y en el flujo.

Un ejemplo de sintaxis para el cuadro del identificador de vista 172 es el siguiente:

```

10      {
        ent(2) sin signo reservado6 = 0;
        ent(3) sin signo id_temporal_mín
        ent(3) sin signo id_temporal_máx
        ent(16) sin signo    núm_vistas
        para (i=0; i<núm_vistas; i++) {
15            ent(6) sin signo reservado1 = 0;
            ent(10) sin signo id_vista[i];
            ent(6) sin signo reservado2 = 0;
            ent(10) sin signo índice_orden_vista;
            ent(1) sin signo    textura_en_flujo[i];
20            ent(1) sin signo    textura_en_pista[i];
            ent(1) sin signo    profundidad_en_flujo[i];
            ent(1) sin signo    profundidad_en_pista[i];
            ent(2) sin signo    tipo_vista_base;
            ent(10) sin signo    núm_vistas_ref;
25            para (j = 0; j < núm_vistas_ref; j++) {
                ent(4) sin signo    reservado5 = 0;
                ent(2) sin signo    componente_dependiente_idc[i][j];
                ent(10) sin signo    id_vista_ref[i][j];
            }
30        }
    }

```

Un ejemplo de semántica para el cuadro del identificador de vista 172 es el siguiente. Dos elementos de sintaxis `id_temporal_mín` y `id_temporal_máx` pueden tener un valor mínimo y máximo, respectivamente, de un elemento de sintaxis `id_temporal` que está presente en una extensión de la cabecera de la unidad NAL de las unidades NAL asignadas a la pista o a la capa cuando se incluye el cuadro del identificador de vista 172 en una entrada de muestra o una entrada de grupo de múltiples vistas, respectivamente. Para los flujos de AVC, `id_temporal_mín` e `id_temporal_máx` toman el valor que está, o estaría, en la unidad NAL del prefijo.

Un elemento sintáctico `núm_vistas`, cuando el cuadro del identificador de vista 172 está presente en una entrada de muestra, puede indicar el número de vistas incluidas en la pista. Cuando el cuadro del identificador de vistas 172 está presente en una entrada de grupo de múltiples vistas, `núm_vistas` puede indicar el número de vistas incluidas en la respectiva capa.

Otro elemento de sintaxis `id_vista[i]` puede indicar el valor del elemento de sintaxis `id` de vista en la extensión de la cabecera de la unidad NAL de una vista incluida en la pista o en la capa cuando el cuadro del identificador de vista 172 se incluye en una entrada de muestra o en una entrada de grupo de múltiples vistas, respectivamente.

El elemento sintáctico `índice_orden_vista` puede indicar el valor de una variable de valor de índice de orden (VOIDx), tal como se especifica en el anexo H de la norma ISO/IEC 14496-10, para una vista incluida en la pista o en la capa cuando se incluye el cuadro del identificador de vista 172 en una entrada de muestra o en una entrada de grupo de múltiples vistas, respectivamente.

Se introduce un elemento sintáctico `textura_en_flujo[i]`, donde `textura_en_flujo[i]` igual a 1 puede indicar que la vista de textura para la vista con `id_vista[i]` está presente en el flujo. El valor indica que la vista de textura para la vista con `id_vista[i]` puede no estar presente en el flujo.

Del mismo modo, se introduce un elemento sintáctico `textura_en_pista[i]`, donde `textura_en_pista[i]` igual a 1 puede indicar que la vista de textura para la vista con `id_vista[i]` está presente en la pista. El valor puede indicar que la vista de textura para la vista con `id_vista[i]` no está presente en la pista. En algunos ejemplos, cuando `textura_en_flujo[i]` es igual a 0, el valor de `textura_en_pista[i]` será igual a 0.

Se introduce otro elemento sintáctico `profundidad_en_flujo[i]`, donde `profundidad_en_flujo[i]` igual a 1 puede indicar que la vista de profundidad para la vista con `id_vista[i]` está presente en el flujo. El valor puede indicar que la vista de profundidad para la vista con `id_vista[i]` no está presente en el flujo. En algunos ejemplos, cuando `textura_en_flujo[i]` es igual a 0, el valor de `profundidad_en_flujo[i]` será igual a 1.

Se introduce un elemento sintáctico `profundidad_en_pista[i]`, donde `profundidad_en_pista[i]` igual a 1 puede indicar que la vista de profundidad para la vista con `id_vista[i]` está presente en la pista. El valor puede indicar que la vista de profundidad para la vista con `id_vista[i]` no está presente en la pista. En algunos ejemplos, cuando `profundidad_en_flujo[i]` es igual a 0, el valor de `profundidad_en_pista[i]` será igual a 0. En algunos ejemplos, cuando `textura_en_pista[i]` es igual a 0, el valor de `profundidad_en_pista[i]` será igual a 1.

Otro elemento sintáctico `tipo_vista_base` puede indicar si la vista es una vista base (virtual o no). El elemento sintáctico `tipo_vista_base` puede tomar los siguientes valores: 0 indica que la vista no es una vista base ni una vista base virtual; 1 se utilizará para etiquetar la vista base no virtual del flujo de bits de MVC; 2 es un valor reservado y no se utilizará (aunque puede utilizarse posteriormente); y 3 indica que la vista con `id_vista[i]` es una vista base virtual. La respectiva vista no base codificada independientemente con `id_vista[i]` puede residir en otra pista. Cuando `tipo_vista_base` sea igual a 3, el subsiguiente `núm_vistas_ref` deberá ser igual a 0.

Otro elemento sintáctico `núm_vistas_ref` puede indicar el número de vistas que pueden consultarse directa o indirectamente en la vista con `id_vista[i]`.

Se introduce otro elemento sintáctico `componente_dependiente_idc[i][j]` que puede indicar cómo se requieren la vista de textura y la vista de profundidad de una vista `j` de referencia para decodificar la vista con `id_vista[i]`. Si el valor es igual a 0, sólo se puede requerir la vista de textura de la vista de referencia. Si el valor es igual a 1, sólo se puede requerir la vista de profundidad de la vista de referencia. Si el valor es igual a 2, se pueden requerir tanto la vista de textura como la vista de profundidad de la vista de referencia. Se reserva el valor de 3.

Otro elemento sintáctico `id_vista_ref[i][j]` puede indicar el identificador de la vista `j` a la que puede referirse directa o indirectamente la vista con `id_vista[i]`, es decir, que puede ser necesaria para la decodificación de la vista con `id_vista[i]`. Si se requiere una vista para decodificar la vista con `id_vista[i]`, puede indicarse como una de `id_vista_ref[i][j]`. Cuando se incluye el cuadro del identificador de vista 172 en una entrada de muestra, puede recomendarse indicar las vistas a las que se hace referencia tanto para las unidades de acceso de anclaje como para las de no anclaje en la misma entrada de muestra.

Además, en algunos ejemplos, el módulo de creación de archivos de vídeo 72 puede proporcionar una indicación explícita en un cuadro de resolución de profundidad de 3VC, como el cuadro de resolución de profundidad de 3VC 170 de la FIG. 1, de las resoluciones espaciales de las vistas de profundidad. El cuadro de resolución de profundidad de 3VC 170 puede incluirse en un cuadro de configuración de 3VC, tal como el cuadro de configuración de 3VC 168 de la FIG. 1. En algunos ejemplos, el cuadro de configuración de 3VC 168 tiene que incluir un registro de configuración de decodificador 3VC y puede incluirse en una entrada de muestra, pero debe incluirse en una entrada de muestra de tipo '3vc1' o '3vc2'.

En algunos aspectos, el módulo de creación de archivos de vídeo 72 compone además archivos de vídeo en los que las vistas de textura y las vistas de profundidad se almacenan en pistas separadas. Un flujo de MVC o 3VC puede ser representado por una o más pistas de vídeo en un archivo (como los flujos de datos de vídeo 146 de la FIG. 1). Cada pista puede representar una o más vistas del flujo. Para una pista en un archivo que almacena un flujo de vídeo de 3VC, como el archivo multimedia de múltiples vistas con profundidad 142, la pista puede contener sólo la textura, sólo la profundidad, o tanto la textura como la profundidad.

En algunos ejemplos, puede haber un conjunto mínimo de una o más pistas que, cuando se toman juntas, contienen el conjunto completo de información codificada. Todas estas pistas pueden tener una indicación, "representación_completa", establecida en todas sus entradas de muestra. El grupo de pistas que forma la información codificada completa puede llamarse "subconjunto completo". La pista que tiene la bandera "representación completa" establecida y contiene unidades NAL de una vista base con `id_temporal` igual a 0 puede ser nominada como una "pista de vista base". Todas las demás pistas que formen parte del mismo flujo podrán estar vinculadas a esta pista de vista base mediante una referencia de pista de tipo "sbas" (vista base). La información codificada completa puede conservarse cuando se conservan las pistas incluidas en el "subconjunto completo"; todas las demás pistas serán subconjuntos, copias o reposiciones del subconjunto completo.

Todas las pistas que compartan la misma pista de vista base pueden también compartir la misma escala de tiempo. En algunos ejemplos con flujos de 3VC, todas las pistas que contienen la vista de textura y la vista de profundidad de una vista concreta comparten la misma escala de tiempo. La vista de textura y la vista de profundidad de una vista particular pueden tener el mismo valor de `id` de vista (por lo tanto, también el mismo valor de índice de orden de vista) y se pueden referir como asociadas entre sí.

Si una vista representada por una pista utiliza otra vista representada por otra pista como referencia de predicción entre vistas, se incluirá una referencia de pista del tipo "scal" en la pista que hace referencia a la pista de origen para la predicción entre vistas.

En algunos ejemplos para flujos de 3VC, si una vista de profundidad se almacena en una pista diferente a la que contiene la vista de textura asociada a la vista de profundidad, se incluirá una referencia de pista de tipo 'deps' en la pista que contiene la vista de profundidad, refiriéndose a la pista que contiene la vista de textura. La presencia de la referencia de pista 'deps' indica que la pista actual contiene la vista de profundidad que está asociada con una vista de textura en la pista de referencia.

En algunos ejemplos, después de que la unidad de creación de archivos de vídeo 72 genera un archivo de vídeo, la unidad de creación de archivos de vídeo 72 pasa el archivo de vídeo a la interfaz de salida de archivos de vídeo 76. La interfaz de salida de archivos de vídeo 76 emite el archivo de vídeo, por ejemplo, a la interfaz de salida 22 del dispositivo fuente 20. En algunos ejemplos, la interfaz de salida de archivos de vídeo 76 envía el archivo de vídeo a un medio de almacenamiento del dispositivo fuente 20 o del dispositivo de almacenamiento 32.

La FIG. 7 es un diagrama esquemático que ilustra un ejemplo de conjunto de dispositivos que forman parte de una red 100 que puede implementar las técnicas descritas en esta divulgación. En este ejemplo, la red 100 incluye los dispositivos de enrutamiento 104A, 104B (dispositivos de enrutamiento 104) y el dispositivo de transcodificación 106. Los dispositivos de enrutamiento 104 y el dispositivo de transcodificación 106 pretenden representar un pequeño número de dispositivos que pueden formar parte de la red 100. Otros dispositivos de red, tales como conmutadores, concentradores, puertas de acceso, cortafuegos, puentes y otros dispositivos de este tipo también pueden incluirse dentro de la red 100. Además, pueden proporcionarse dispositivos de red adicionales a lo largo de una ruta de red entre el dispositivo servidor 102 y el dispositivo cliente 108. Cualquiera de estos dispositivos puede comprender un codificador o un decodificador que puede implementar una o más técnicas descritas en el presente documento. Además, los dispositivos de red pueden comprender un MANE que no incluya necesariamente capacidades completas de codificación o decodificación, y el MANE también puede implementar una o más técnicas descritas en el presente documento. El dispositivo servidor 102 puede corresponder al dispositivo fuente 12 de la FIG. 3, mientras que el dispositivo cliente 108 puede corresponder al dispositivo de destino 14 de la FIG. 3, en algunos ejemplos.

En general, los dispositivos de enrutamiento 104 implementan uno o más protocolos de enrutamiento para intercambiar datos de red a través de la red 100. En algunos ejemplos, los dispositivos de enrutamiento 104 pueden estar configurados para realizar operaciones de proxy o caché. Por lo tanto, en algunos ejemplos, los dispositivos de enrutamiento 104 pueden denominarse dispositivos proxy. En general, los dispositivos de enrutamiento 104 ejecutan protocolos de enrutamiento para descubrir rutas a través de la red 100. Al ejecutar dichos protocolos de enrutamiento, el dispositivo de enrutamiento 104B puede descubrir una ruta de red desde sí mismo hasta el dispositivo servidor 102 a través del dispositivo de enrutamiento 104A.

A través de la red 100, el dispositivo servidor 102 puede proporcionar al dispositivo cliente 108 uno o más archivos multimedia compuestos según las técnicas descritas en el presente documento. En otros ejemplos, el dispositivo cliente 108 proporciona al dispositivo servidor 102 uno o más archivos multimedia compuestos según las técnicas descritas en el presente documento a través de la red 100.

La FIG. 8 es un diagrama de flujo que ilustra un método de ejemplo para analizar un archivo multimedia que incluye contenido de vídeo de múltiples vistas según las técnicas descritas en esta divulgación. El proceso descrito en la FIG. 8 puede ser realizado por un dispositivo de codificación de vídeo que puede incluir cualquier combinación de dispositivo de destino 20 y módulo de desencapsulación 29 descrito en el presente documento. Un dispositivo de codificación de vídeo puede analizar una pista de datos de vídeo, en la que la pista incluye una o más vistas (200). Por ejemplo, el módulo de desencapsulación 29 puede obtener un archivo multimedia que contenga múltiples vistas codificadas más profundidad, como el archivo multimedia de múltiples vistas con profundidad 142 desde el dispositivo fuente 12. El dispositivo de codificación de vídeo puede analizar la información para determinar si la pista incluye sólo vistas de textura, sólo vistas de profundidad, o ambas vistas de textura y profundidad (202). Por ejemplo, el módulo de desencapsulación 29 puede analizar la información del archivo multimedia de múltiples vistas con profundidad 142 para determinar si al menos una de las pistas incluidas en el archivo multimedia de múltiples vistas con profundidad 142 contiene vistas de textura, vistas de profundidad o ambas.

En algunos ejemplos, el análisis de la información comprende además el análisis de un cuadro del identificador de vista de al menos una de las entradas de muestra y una entrada de grupo de múltiples vistas, en el que al menos una de las entradas de muestra y la entrada de grupo de múltiples vistas están asociadas a la pista. Por ejemplo, el módulo de desencapsulación 29 puede analizar la información del cuadro del identificador de vista 172 para determinar, para cada vista incluida en la pista, si la vista es una vista de textura o una vista de profundidad. El módulo de desencapsulación 29 puede analizar el cuadro del identificador de vista 172 para determinar además si se requiere una vista de textura o una vista de profundidad de una vista de referencia para decodificar una vista específica en la pista.

El módulo de desencapsulación 29 puede analizar otro segundo conjunto de datos de vídeo basado en los primeros datos de vídeo y analizar además la información para determinar si la primera vista de textura, la primera vista de profundidad, o ambas, la primera vista de textura y la primera vista de profundidad de los primeros datos de vídeo, se utilizan para decodificar los segundos datos de vídeo.

En otro ejemplo, el proceso incluye analizar un cuadro de mensaje de información de mejora complementaria (SEI) para determinar un mensaje de SEI de información de escalabilidad tridimensional asociado con una o más de las vistas. Es decir, el módulo de desencapsulación 29 puede analizar un cuadro de mensaje de SEI, como el cuadro de mensaje de SEI de información de escalabilidad de 3VC 166, para obtener información de un mensaje de SEI de una o más de las vistas.

En ejemplos en los que la pista contiene una o más vistas de profundidad, el proceso puede incluir además el análisis de un cuadro de resolución de profundidad de 3VC para determinar una resolución espacial de una o más vistas de profundidad. Por ejemplo, el módulo de desencapsulación 29 analiza el cuadro de resolución de profundidad de 3VC 170 para determinar una resolución espacial de una vista de profundidad incluida en el archivo multimedia de múltiples vistas con profundidad 142. En algunos ejemplos, el módulo de desencapsulación 29 analiza el cuadro de resolución de profundidad de 3VC 170 desde el cuadro de configuración de 3VC 168. En otros ejemplos, que quedan fuera de la invención reivindicada, el módulo de desencapsulación 29 analiza el cuadro de resolución de profundidad de 3VC 170 directamente desde una entrada de muestra. El módulo de desencapsulación 29 también puede analizar un registro de configuración de decodificador 3VC, RegistroConfiguraciónDecodificador3VC que incluye semántica que indica características relacionadas con un flujo de vídeo 3DV, como un flujo de vídeo de MVC+D. El proceso también puede incluir el análisis de un registro de configuración del decodificador de vídeo tridimensional (3VC) para determinar la anchura y la altura de una vista de profundidad incluida en la pista.

En algunos ejemplos, la pista es una primera pista, en la que la primera pista incluye una de las vistas de profundidad de una vista y una vista de textura de la vista, y el proceso incluye además el análisis de una segunda pista de datos de vídeo, en la que la segunda pista incluye la vista de profundidad de la vista y la vista de textura de la vista que no está incluida en la primera pista. El proceso puede incluir además el análisis de una referencia de pista para la primera pista que hace referencia a la segunda pista, en la que la referencia de pista incluye al menos un tipo de referencia de pista que indica una de las siguientes cosas: que la primera pista incluye la vista de profundidad que está asociada con la vista de textura almacenada en la segunda pista (por ejemplo, el tipo de referencia de pista es "deps"), que la primera pista depende de la vista de textura almacenada en la segunda pista (por ejemplo, el tipo de referencia de pista es "tref"), y que la primera pista depende de una vista de profundidad almacenada en la segunda pista (por ejemplo, el tipo de referencia de pista es "dref").

En los ejemplos en los que la pista es una pista de codificación de vídeo tridimensional (3VC), el proceso puede incluir además el análisis de un registro de configuración del decodificador 3VC, en el que el registro de configuración del decodificador de 3VC indica un registro de configuración para una entrada de muestra coincidente de los datos de vídeo. En otro ejemplo, cuando la pista es una pista de 3VC que incluye una o más unidades de capa de abstracción de red (NAL) de profundidad, el proceso incluye analizar al menos uno de los tipos de entrada de muestra "3vc1" que permite extractores y un tipo de entrada de muestra "3vc2" que no permite extractores. En otros ejemplos, el proceso puede contener una o más de las técnicas descritas en esta divulgación.

La FIG. 9 es un diagrama de flujo que ilustra un método de ejemplo para crear un archivo multimedia con contenido de vídeo de múltiples vistas para indicar la presencia de vistas de textura y profundidad según las técnicas descritas en esta divulgación. El proceso descrito en la FIG. 9 puede ser realizado por un dispositivo de codificación de vídeo que incluye el módulo de encapsulación 21 descrito en este documento. Un dispositivo de codificación de vídeo puede componer una pista de datos de vídeo, en la que la pista incluye una o más vistas (206). Por ejemplo, el módulo de encapsulación 21 puede obtener un dato de vídeo del codificador de vídeo 21. El dispositivo de codificación de vídeo puede componer información que indique si la pista incluye sólo vistas de textura, sólo vistas de profundidad, o tanto vistas de textura como de profundidad (208). Por ejemplo, el módulo de encapsulación 21 puede componer información en el archivo multimedia de múltiples vistas con profundidad 142 para indicar si al menos una de las pistas incluidas en el archivo multimedia de múltiples vistas con profundidad 142 contiene vistas de textura, vistas de profundidad, o tanto vistas de textura como de profundidad. El módulo de encapsulación 21 puede componer un archivo multimedia que contenga múltiples vistas codificadas más profundidad, como un archivo multimedia de múltiples vistas con profundidad 142, según las técnicas descritas en el presente documento.

En algunos ejemplos, la composición de la información comprende además la composición de un cuadro del identificador de vista en al menos una de las entradas de muestra y una entrada de grupo de múltiples vistas, en la que al menos una de las entradas de muestra y la entrada de grupo de vistas múltiples están asociadas a la pista. Por ejemplo, el módulo de encapsulación 21 puede componer la información en el cuadro del identificador de vista 172 para indicar, para cada vista incluida en la pista, si la vista es una vista de textura o una vista de profundidad. El módulo de encapsulación 21 puede componer el cuadro del identificador de vista 172 para indicar además si se requiere una vista de textura o una vista de profundidad de una vista de referencia para decodificar una vista específica en la pista.

El módulo de encapsulación 21 puede componer otro segundo conjunto de datos de vídeo basado en los primeros datos de vídeo y componer además información para indicar si la primera vista de textura, la primera vista de profundidad, o ambas, la primera vista de textura y la primera vista de profundidad de los primeros datos de vídeo, pueden utilizarse para decodificar o analizar los segundos datos de vídeo.

En otro ejemplo, el proceso incluye componer un cuadro de mensaje de información de mejora complementaria (SEI) para indicar un mensaje de SEI de información de escalabilidad tridimensional asociado a una o más de las vistas. Es decir, el módulo de encapsulación 21 puede componer un cuadro de mensaje de SEI, tal como el cuadro de mensaje de SEI de información de escalabilidad de 3VC 166, para proporcionar información en un mensaje de SEI de una o más de las vistas.

En ejemplos en los que la pista contiene una o más vistas de profundidad, el proceso puede incluir además la composición de un cuadro de resolución de profundidad de 3VC para determinar una resolución espacial de una o más vistas de profundidad. Por ejemplo, el módulo de encapsulación 21 compone el cuadro de resolución de profundidad de 3VC 170 para indicar una resolución espacial de una vista de profundidad incluida en el archivo multimedia de múltiples vistas con profundidad 142. En algunos ejemplos, el módulo de encapsulación 21 compone el cuadro de resolución de profundidad de 3VC 170 en el cuadro de configuración de 3VC 168. En otros ejemplos, el módulo de encapsulación 21 compone el cuadro de resolución de profundidad de 3VC 170 directamente en una entrada de muestra. El módulo de encapsulación 21 también puede componer un registro de configuración de decodificador 3VC, RegistroConfiguraciónDecodificador3VC que incluye semántica que indica características relacionadas con un flujo de vídeo 3DV, como un flujo de vídeo de MVC+D. El proceso también puede incluir la composición de un registro de configuración del decodificador de vídeo tridimensional (3VC) para determinar la anchura y la altura de una vista de profundidad incluida en la pista.

En algunos ejemplos, la pista es una primera pista, en la que la primera pista incluye una de las vistas de profundidad de una vista y una vista de textura de la vista, y el proceso incluye además la composición de una segunda pista de datos de vídeo, en la que la segunda pista incluye la vista de profundidad de la vista y la vista de textura de la vista que no está incluida en la primera pista. El proceso puede incluir además la composición de una referencia de pista para la primera pista que hace referencia a la segunda pista, en la que la referencia de pista incluye al menos un tipo de referencia de pista que indica una de las siguientes cosas: que la primera pista incluye la vista de profundidad que está asociada con la vista de textura almacenada en la segunda pista (por ejemplo, el tipo de referencia de pista es "deps"), que la primera pista depende de la vista de textura almacenada en la segunda pista (por ejemplo, el tipo de referencia de pista es "tref"), y que la primera pista depende de una vista de profundidad almacenada en la segunda pista (por ejemplo, el tipo de referencia de pista es "dref").

En los ejemplos en los que la pista es una pista de codificación de vídeo tridimensional (3VC), el proceso puede incluir además la composición de un registro de configuración del decodificador 3VC, en el que el registro de configuración del decodificador de 3VC indica un registro de configuración para una entrada de muestra coincidente de los datos de vídeo. En otro ejemplo, cuando la pista es una pista de 3VC que incluye una o más unidades de capa de abstracción de red (NAL) de profundidad, el proceso incluye la composición de al menos uno de los tipos de entrada de muestra "3vc1" que permite extractores y un tipo de entrada de muestra "3vc2" que no permite extractores. En otros ejemplos, el proceso puede contener una o más de las técnicas descritas en esta divulgación.

La FIG. 10 es un diagrama de flujo que ilustra un ejemplo de método de análisis de un archivo multimedia que contiene vídeo de múltiples vistas que indica la dependencia de las vistas según las técnicas descritas en esta divulgación. El proceso descrito en la FIG. 10 puede ser realizado por un dispositivo de codificación de vídeo que incluye el módulo de desencapsulación 29 descrito en este documento. Un dispositivo de codificación de vídeo puede analizar una pista de datos de vídeo, en la que la pista incluye una o más vistas (220). Por ejemplo, el módulo de desencapsulación 29 puede obtener un archivo multimedia que contenga múltiples vistas codificadas más profundidad, como el archivo multimedia de múltiples vistas con profundidad 142 desde el dispositivo fuente 12. El dispositivo de codificación de vídeo puede analizar la información para determinar si se requiere una vista de textura o una vista de profundidad de una vista de referencia para decodificar una de las vistas de la pista (202). Por ejemplo, el módulo de desencapsulación 29 puede analizar la información del archivo multimedia de múltiples vistas con profundidad 142 para determinar si se requiere una vista de textura o una vista de profundidad de una vista de referencia para decodificar una o más vistas en la pista incluida en el archivo multimedia de múltiples vistas con profundidad 142.

En algunos ejemplos, el proceso puede incluir además la decodificación de los segundos datos de vídeo basados en los primeros datos de vídeo, en los que la decodificación de los segundos datos de vídeo incluye información de decodificación que indica si se utiliza una vista de textura, una vista de profundidad o ambas vistas de textura y profundidad de los primeros datos de vídeo para decodificar los segundos datos de vídeo.

En algunos ejemplos, el análisis de la información comprende además el análisis de un cuadro del identificador de vista de al menos una de las entradas de muestra y una entrada de grupo de múltiples vistas, en el que al menos una de las entradas de muestra y la entrada de grupo de múltiples vistas están asociadas a la pista. Por ejemplo, el decodificador de vídeo 30 puede analizar la información del cuadro del identificador de vista 172 para determinar si se requiere una vista de textura o una vista de profundidad de una vista de referencia para decodificar una o más vistas de la pista incluida en el archivo multimedia de múltiples vistas con profundidad 142. En otro ejemplo, el módulo de desencapsulación 29 puede analizar la información del cuadro del identificador de vista 172 para determinar si la pista incluye sólo vistas de textura, sólo vistas de profundidad, o vistas de textura y profundidad. En otro ejemplo, el módulo de desencapsulación 29 puede analizar la información del cuadro del identificador de vista 172 para determinar, para cada vista incluida en la pista, si la vista es una vista de textura o una vista de profundidad. El módulo de

desencapsulación 29 puede analizar el cuadro del identificador de vista 172 para determinar además si se requiere una vista de textura o una vista de profundidad de una vista de referencia para decodificar una vista específica en la pista.

5 El módulo de desencapsulación 29 puede analizar otro segundo conjunto de datos de vídeo basado en los primeros datos de vídeo y analizar además la información para determinar si la primera vista de textura, la primera vista de profundidad, o ambas, la primera vista de textura y la primera vista de profundidad de los primeros datos de vídeo, se utilizan para decodificar los segundos datos de vídeo.

10 En otro ejemplo, el proceso incluye analizar un cuadro de mensaje de información de mejora complementaria (SEI) para determinar un mensaje de SEI de información de escalabilidad tridimensional asociado con una o más de las vistas. Es decir, el módulo de desencapsulación 29 puede analizar un cuadro de mensaje de SEI, como el cuadro de mensaje de SEI de información de escalabilidad de 3VC 166, para obtener información de un mensaje de SEI de una o más de las vistas.

15 En ejemplos en los que la pista contiene una o más vistas de profundidad, el proceso puede incluir además el análisis de un cuadro de resolución de profundidad de 3VC para determinar una resolución espacial de una o más vistas de profundidad. Por ejemplo, el módulo de desencapsulación 29 analiza el cuadro de resolución de profundidad de 3VC 170 para determinar una resolución espacial de una vista de profundidad incluida en el archivo multimedia de múltiples vistas con profundidad 142. En algunos ejemplos, el módulo de desencapsulación 29 analiza el cuadro de resolución de profundidad de 3VC 170 desde el cuadro de configuración de 3VC 168. En otros ejemplos, el módulo de desencapsulación 29 analiza el cuadro de resolución de profundidad de 3VC 170 directamente desde una entrada de muestra. El módulo de desencapsulación 29 también puede analizar un registro de configuración de decodificador 3VC, RegistroConfiguraciónDecodificador3VC que incluye semántica que indica características relacionadas con un flujo de vídeo 3DV, como un flujo de vídeo de MVC+D. El proceso también puede incluir el análisis de un registro de configuración del decodificador de vídeo tridimensional (3VC) para determinar la anchura y la altura de una vista de profundidad incluida en la pista.

30 En algunos ejemplos, la pista es una primera pista, en la que la primera pista incluye una de las vistas de profundidad de una vista y una vista de textura de la vista, y el proceso incluye además el análisis de una segunda pista de datos de vídeo, en la que la segunda pista incluye la vista de profundidad de la vista y la vista de textura de la vista que no está incluida en la primera pista. El proceso puede incluir además el análisis de una referencia de pista para la primera pista que hace referencia a la segunda pista, en la que la referencia de pista incluye al menos un tipo de referencia de pista que indica una de las siguientes cosas: que la primera pista incluye la vista de profundidad que está asociada con la vista de textura almacenada en la segunda pista (por ejemplo, el tipo de referencia de pista es "deps"), que la primera pista depende de la vista de textura almacenada en la segunda pista (por ejemplo, el tipo de referencia de pista es "tref"), y que la primera pista depende de una vista de profundidad almacenada en la segunda pista (por ejemplo, el tipo de referencia de pista es "dref").

40 En los ejemplos en los que la pista es una pista de codificación de vídeo tridimensional (3VC), el proceso puede incluir además el análisis de un registro de configuración del decodificador 3VC, en el que el registro de configuración del decodificador de 3VC indica un registro de configuración para una entrada de muestra coincidente de los datos de vídeo. En otro ejemplo, cuando la pista es una pista de 3VC que incluye una o más unidades de capa de abstracción de red (NAL) de profundidad, el proceso incluye analizar al menos uno de los tipos de entrada de muestra "3vc1" que permite extractores y un tipo de entrada de muestra "3vc2" que no permite extractores. En otros ejemplos, el proceso puede contener una o más de las técnicas descritas en esta divulgación.

50 La FIG. 11 es un diagrama de flujo que ilustra un método de ejemplo para crear un archivo multimedia con contenido de vídeo de múltiples vistas para indicar la dependencia de las vistas según las técnicas descritas en esta divulgación. El proceso descrito en la FIG. 11 puede ser realizado por un dispositivo de codificación de vídeo que incluye el módulo de encapsulación 21 descrito en este documento. Un dispositivo de codificación de vídeo puede componer una pista de datos de vídeo, en la que la pista incluye una o más vistas (226). Por ejemplo, el módulo de encapsulación 21 puede obtener un dato de vídeo del codificador de vídeo 21. El dispositivo de codificación de vídeo puede componer información que indique si se requiere una vista de textura o una vista de profundidad de una vista de referencia para decodificar una de las vistas de la pista (228). Por ejemplo, el módulo de encapsulación 21 puede componer información en el archivo multimedia de múltiples vistas con profundidad 142 para indicar si se requiere una vista de textura o una vista de profundidad de una vista de referencia para decodificar una de las pistas incluidas en el archivo multimedia de múltiples vistas con profundidad 142. El módulo de encapsulación 21 puede componer un archivo multimedia que contenga múltiples vistas codificadas más profundidad, como un archivo multimedia de múltiples vistas con profundidad 142, según las técnicas descritas en el presente documento.

65 En algunos ejemplos, la composición de la información comprende además la composición de un cuadro del identificador de vista en al menos una de las entradas de muestra y una entrada de grupo de múltiples vistas, en la que al menos una de las entradas de muestra y la entrada de grupo de vistas múltiples están asociadas a la pista. Por ejemplo, el módulo de encapsulación 21 puede componer la información en el cuadro del identificador de vista 172 para indicar, para cada vista incluida en la pista, si la vista es una vista de textura o una vista de profundidad. El módulo

de encapsulación 21 puede componer el cuadro del identificador de vista 172 para indicar además si la pista incluye sólo vistas de textura, sólo vistas de profundidad, o vistas de textura y profundidad. El módulo de encapsulación 21 puede componer el cuadro del identificador de vista 172 para indicar, además, para cada vista incluida en la pista, si la vista es una vista de textura o una vista de profundidad.

El módulo de encapsulación 21 puede componer otro segundo conjunto de datos de vídeo basado en los primeros datos de vídeo y componer además información para indicar si la primera vista de textura, la primera vista de profundidad, o ambas, la primera vista de textura y la primera vista de profundidad de los primeros datos de vídeo, se utilizan para decodificar o analizar los segundos datos de vídeo.

En otro ejemplo, el proceso incluye componer un cuadro de mensaje de información de mejora complementaria (SEI) para indicar un mensaje de SEI de información de escalabilidad tridimensional asociado a una o más de las vistas. Es decir, el módulo de encapsulación 21 puede componer un cuadro de mensaje de SEI, tal como el cuadro de mensaje de SEI de información de escalabilidad de 3VC 166, para proporcionar información en un mensaje de SEI de una o más de las vistas.

En ejemplos en los que la pista contiene una o más vistas de profundidad, el proceso puede incluir además la composición de un cuadro de resolución de profundidad de 3VC para determinar una resolución espacial de una o más vistas de profundidad. Por ejemplo, el módulo de encapsulación 21 compone el cuadro de resolución de profundidad de 3VC 170 para indicar una resolución espacial de una vista de profundidad incluida en el archivo multimedia de múltiples vistas con profundidad 142. En algunos ejemplos, el módulo de encapsulación 21 compone el cuadro de resolución de profundidad de 3VC 170 en el cuadro de configuración de 3VC 168. En otros ejemplos, el módulo de encapsulación 21 compone el cuadro de resolución de profundidad de 3VC 170 directamente en una entrada de muestra. El módulo de encapsulación 21 también puede componer un registro de configuración de decodificador 3VC, RegistroConfiguraciónDecodificador3VC que incluye semántica que indica características relacionadas con un flujo de vídeo 3DV, como un flujo de vídeo de MVC+D. El proceso también puede incluir la composición de un registro de configuración del decodificador de vídeo tridimensional (3VC) para determinar la anchura y la altura de una vista de profundidad incluida en la pista.

En algunos ejemplos, la pista es una primera pista, en la que la primera pista incluye una de las vistas de profundidad de una vista y una vista de textura de la vista, y el proceso incluye además la composición de una segunda pista de datos de vídeo, en la que la segunda pista incluye la vista de profundidad de la vista y la vista de textura de la vista que no está incluida en la primera pista. El proceso puede incluir además la composición de una referencia de pista para la primera pista que hace referencia a la segunda pista, en la que la referencia de pista incluye al menos un tipo de referencia de pista que indica una de las siguientes cosas: que la primera pista incluye la vista de profundidad que está asociada con la vista de textura almacenada en la segunda pista (por ejemplo, el tipo de referencia de pista es "deps"), que la primera pista depende de la vista de textura almacenada en la segunda pista (por ejemplo, el tipo de referencia de pista es "tref"), y que la primera pista depende de una vista de profundidad almacenada en la segunda pista (por ejemplo, el tipo de referencia de pista es "dref").

En los ejemplos en los que la pista es una pista de codificación de vídeo tridimensional (3VC), el proceso puede incluir además la composición de un registro de configuración del decodificador 3VC, en el que el registro de configuración del decodificador de 3VC indica un registro de configuración para una entrada de muestra coincidente de los datos de vídeo. En otro ejemplo, cuando la pista es una pista de 3VC que incluye una o más unidades de capa de abstracción de red (NAL) de profundidad, el proceso incluye la composición de al menos uno de los tipos de entrada de muestra "3vc1" que permite extractores y un tipo de entrada de muestra "3vc2" que no permite extractores. En otros ejemplos, el proceso puede contener una o más de las técnicas descritas en esta divulgación.

Un proceso alternativo de procesamiento de datos de vídeo según las técnicas descritas en el presente documento incluye la codificación de los primeros datos de vídeo, en los que los primeros datos de vídeo incluyen una vista de textura y una vista de profundidad. El proceso incluye además la codificación de segundos datos de vídeo basados en los primeros datos de vídeo, en los que la codificación de los segundos datos de vídeo incluye información de codificación que indica si la vista de textura, la vista de profundidad o ambas vistas de textura y profundidad de los primeros datos de vídeo se utilizan para decodificar los segundos datos de vídeo.

La FIG. 12 es un diagrama de flujo que ilustra un método, de acuerdo con la presente invención, de análisis de un archivo multimedia que contiene vídeo de múltiples vistas para determinar una resolución espacial de una vista de profundidad de acuerdo con las técnicas descritas en esta divulgación. El proceso descrito en la FIG. 12 puede ser realizado por un dispositivo de codificación de vídeo que incluye el módulo de desencapsulación 29 descrito en este documento. Un dispositivo de codificación de vídeo analiza una pista de datos de vídeo, en la que la pista incluye una o más vistas, entre ellas, al menos una vista de profundidad (240). Por ejemplo, el módulo de desencapsulación 29 puede obtener un archivo multimedia que contenga múltiples vistas codificadas más profundidad, como el archivo multimedia de múltiples vistas con profundidad 142 desde el dispositivo fuente 12. El dispositivo de codificación de vídeo analiza la información para determinar una resolución espacial asociada con la vista de profundidad, en la que la decodificación de la resolución espacial no requiere el análisis de un conjunto de parámetros de secuencia de la vista de profundidad (242). Por ejemplo, el módulo de desencapsulación 29 puede analizar la información de múltiples

vistas con el archivo multimedia de profundidad 142 para determinar una resolución espacial de la vista de profundidad almacenada en el archivo multimedia de profundidad 142.

En el método de la Fig. 12, el análisis de la información comprende además el análisis de la resolución espacial de un cuadro de resolución de profundidad de 3VC de un cuadro de configuración de 3VC. En un ejemplo que no forma parte de la invención reivindicada, el análisis de la información comprende el análisis de la resolución espacial directamente desde una entrada de muestra. Es decir, un cuadro de resolución de profundidad de 3VC 170, se almacena en un cuadro de configuración de 3VC, tal como el cuadro de configuración de 3VC 168. El proceso comprende además analizar un registro de configuración de decodificador 3VC del cuadro de Configuración de 3VC, en el que el cuadro de Configuración de 3VC se incluye en al menos una entrada de muestra de tipo '3vc1' que admite extractores y una entrada de muestra de tipo '3vc2' que no admite extractores.

En algunos ejemplos, el módulo de desencapsulación 29 puede analizar el registro de configuración del decodificador de 3VC para determinar una anchura y una altura de la vista de profundidad. El RegistroConfiguraciónDecodificador3VC puede incluir semántica que indique características relacionadas con un flujo de vídeo 3DV, como un flujo de vídeo de MVC+D. El proceso puede incluir el análisis del registro de configuración del decodificador de vídeo tridimensional (3VC) para determinar la anchura y la altura de una vista de profundidad incluida en la pista.

En algunos ejemplos, el análisis de la información comprende además el análisis de un cuadro del identificador de vista de al menos una de las entradas de muestra y una entrada de grupo de múltiples vistas para determinar al menos una de las siguientes cosas: si se requiere una vista de textura o una vista de profundidad de una vista de referencia para decodificar al menos una de las vistas de la pista, si la pista incluye sólo vistas de profundidad o vistas de textura y profundidad, y, para cada vista incluida en la pista, si la vista es una vista de textura o una vista de profundidad, en la que al menos una de las entradas de muestra y la entrada de grupo de múltiples vistas están asociadas a la pista. Por ejemplo, el módulo de desencapsulación 29 puede analizar la información del cuadro del identificador de vista 172 para determinar, para cada vista incluida en la pista, si la vista es una vista de textura o una vista de profundidad. El módulo de desencapsulación 29 puede analizar además el cuadro del identificador de vista 172 para determinar si se requiere una vista de textura o una vista de profundidad de una vista de referencia para decodificar una vista específica en la pista.

En ejemplos en los que los datos de vídeo son un primer conjunto de datos de vídeo y en los que la pista incluye además al menos una vista de textura, el módulo de desencapsulación 29 puede analizar además un segundo conjunto de datos de vídeo basado en el primer conjunto de datos de vídeo, en el que la información de análisis comprende información de análisis para determinar si la primera vista de textura, la primera vista de profundidad o ambas, la primera textura y la primera vista de profundidad del primer conjunto de datos de vídeo, se utilizan para decodificar el segundo conjunto de datos de vídeo.

En otro ejemplo, el proceso incluye analizar un cuadro de mensaje de información de mejora complementaria (SEI) para determinar un mensaje de SEI de información de escalabilidad tridimensional asociado con una o más de las vistas. Es decir, el módulo de desencapsulación 29 puede analizar un cuadro de mensaje de SEI, como el cuadro de mensaje de SEI de información de escalabilidad de 3VC 166, para obtener información de un mensaje de SEI de una o más de las vistas.

En ejemplos en los que la pista es una primera pista que contiene una vista de profundidad, el proceso puede incluir además el análisis de una segunda pista de datos de vídeo, en la que la segunda pista incluye una vista de textura correspondiente a la vista de profundidad, en la que la vista de textura no se incluye en la primera pista. El proceso puede incluir además el análisis de una referencia de pista para la primera pista que hace referencia a la segunda pista, en la que la referencia de pista incluye al menos un tipo de referencia de pista que indica una de las siguientes cosas: que la primera pista incluye la vista de profundidad que está asociada con la vista de textura almacenada en la segunda pista (por ejemplo, el tipo de referencia de pista es "deps"), que una o más vistas de la primera pista depende de una vista de textura almacenada en la segunda pista (por ejemplo, el tipo de referencia de pista es "tref"), y que una o más vistas de la primera pista depende de una vista de profundidad almacenada en la segunda pista (por ejemplo, el tipo de referencia de pista es "dref").

Otros procesos de ejemplo de procesamiento de datos de vídeo pueden incluir la decodificación de una vista de profundidad y la decodificación de una resolución espacial asociada a la vista de profundidad, en la que la decodificación de la resolución espacial no requiere el análisis de un conjunto de parámetros de secuencia de la vista de profundidad. En realizaciones de la invención, un proceso de procesamiento de datos de vídeo incluye decodificar un cuadro de resolución de profundidad para indicar una resolución espacial de vistas de profundidad, y decodificar un cuadro de configuración, en el que el cuadro de resolución de profundidad está incluido en el cuadro de configuración. El proceso de procesamiento de datos de vídeo incluye además la decodificación de un registro de configuración de decodificador de 3VC del cuadro de configuración de 3VC, en el que el cuadro de configuración de 3VC está incluido en al menos una entrada de muestra de tipo '3vc1' que admite extractores y una entrada de muestra de tipo '3vc2' que no admite extractores. En un ejemplo, el proceso de procesamiento de datos de vídeo incluye la

decodificación de una anchura y una altura de las vistas de profundidad desde el registro de configuración del decodificador de 3VC, en el que 3VC significa codificación de vídeo tridimensional.

La FIG. 13 es un diagrama de flujo que ilustra un método, de acuerdo con la invención, para componer un archivo multimedia con contenido de vídeo de múltiples vistas para indicar una resolución espacial de una vista de profundidad de acuerdo con las técnicas descritas en esta divulgación. El proceso descrito en la FIG. 13 puede ser realizado por un dispositivo de codificación de vídeo que incluye el módulo de encapsulación 21 descrito en este documento. Un dispositivo de codificación de vídeo compone una pista de datos de vídeo, en la que la pista incluye una o más vistas, entre ellas, al menos una vista de profundidad (246). Por ejemplo, el módulo de encapsulación 21 puede obtener un dato de vídeo del codificador de vídeo 21. El dispositivo de codificación de vídeo compone la información para indicar una resolución espacial asociada con la vista de profundidad, en la que la decodificación de la resolución espacial no requiere el análisis de un conjunto de parámetros de secuencia de la vista de profundidad (248). El módulo de encapsulación 21 puede componer un archivo multimedia que contenga múltiples vistas codificadas más profundidad, como un archivo multimedia de múltiples vistas con profundidad 142, según las técnicas descritas en el presente documento. Por ejemplo, el módulo de encapsulación 21 puede componer información en el archivo multimedia de múltiples vistas con profundidad 142 para indicar una resolución espacial de una vista de profundidad en la pista incluida en el archivo multimedia de múltiples vistas con profundidad 142.

En las realizaciones de la invención, la composición de la información comprende además la composición de la resolución espacial en un cuadro de resolución de profundidad de 3VC en un cuadro de configuración de 3VC. Es decir, un cuadro de resolución de profundidad de 3VC, tal como el cuadro de resolución de profundidad de 3VC 170, se almacena en un cuadro de configuración de 3VC, tal como el cuadro de configuración de 3VC 168. El proceso comprende además la composición de un registro de configuración de decodificador 3VC en el cuadro de Configuración de 3VC, en el que el cuadro de Configuración de 3VC se incluye en al menos una entrada de muestra de tipo '3vc1' que admite extractores y una entrada de muestra de tipo '3vc2' que no admite extractores.

En algunos ejemplos, el módulo de encapsulación 21 puede componer el registro de configuración del decodificador de 3VC para indicar una anchura y una altura de la vista de profundidad. El RegistroConfiguraciónDecodificador3VC puede incluir semántica que indique características relacionadas con un flujo de vídeo 3DV, como un flujo de vídeo de MVC+D.

En algunos ejemplos, la composición de la información comprende además la composición de un cuadro del identificador de vista de al menos una de las entradas de muestra y una entrada de grupo de múltiples vistas para indicar al menos una de las siguientes cosas: si se requiere una vista de textura o una vista de profundidad de una vista de referencia para decodificar al menos una de las vistas de la pista, si la pista incluye sólo vistas de profundidad o vistas de textura y profundidad, y, para cada vista incluida en la pista, si la vista es una vista de textura o una vista de profundidad, en la que al menos una de las entradas de muestra y la entrada de grupo de múltiples vistas están asociadas a la pista. Por ejemplo, el módulo de encapsulación 21 puede componer la información en el cuadro del identificador de vista 172 para indicar, para cada vista incluida en la pista, si la vista es una vista de textura o una vista de profundidad. El decodificador de vídeo 30 puede componer además el cuadro del identificador de vista 172 para indicar si se requiere una vista de textura o una vista de profundidad de una vista de referencia para decodificar una vista específica en la pista.

En ejemplos en los que los datos de vídeo son un primer conjunto de datos de vídeo y en los que la pista incluye además al menos una vista de textura, el módulo de encapsulación 21 puede componer además un segundo conjunto de datos de vídeo basado en el primer conjunto de datos de vídeo, en el que la información de composición comprende información para indicar si la primera vista de textura, la primera vista de profundidad o ambas, la primera textura y la primera vista de profundidad del primer conjunto de datos de vídeo, se utilizan para decodificar el segundo conjunto de datos de vídeo.

En otro ejemplo, el proceso incluye componer un cuadro de mensaje de información de mejora complementaria (SEI) para indicar un mensaje de SEI de información de escalabilidad tridimensional asociado a una o más de las vistas. Es decir, el módulo de encapsulación 21 puede componer un cuadro de mensaje de SEI, tal como el cuadro de mensaje de SEI de información de escalabilidad de 3VC 166, para almacenar información de un mensaje de SEI de una o más de las vistas.

En ejemplos en los que la pista es una primera pista, el proceso puede incluir además la composición de una segunda pista de datos de vídeo, en la que la segunda pista incluye una vista de textura correspondiente a la vista de profundidad, en la que la vista de textura no se incluye en la primera pista. El proceso puede incluir además la composición de una referencia de pista para la primera pista que hace referencia a la segunda pista, en la que la referencia de pista incluye al menos un tipo de referencia de pista que indica una de las siguientes cosas: que la primera pista incluye la vista de profundidad que está asociada con la vista de textura almacenada en la segunda pista (por ejemplo, el tipo de referencia de pista es "deps"), que una o más vistas de la primera pista depende de una vista de textura almacenada en la segunda pista (por ejemplo, el tipo de referencia de pista es "tref"), y que una o más vistas de la primera pista depende de una vista de profundidad almacenada en la segunda pista (por ejemplo, el tipo de referencia de pista es "dref").

Otros procesos de ejemplo de procesamiento de datos de vídeo pueden incluir la codificación de una vista de profundidad y la codificación de una resolución espacial asociada a la vista de profundidad, en la que la codificación de la resolución espacial se produce de manera que el análisis de un conjunto de parámetros de secuencia de la vista de profundidad no se utiliza para determinar la resolución espacial. En realizaciones de la invención, un proceso de procesamiento de datos de vídeo incluye codificar un cuadro de resolución de profundidad para indicar una resolución espacial de vistas de profundidad, y codificar un cuadro de configuración, en el que el cuadro de resolución de profundidad está incluido en el cuadro de configuración. El proceso de procesamiento de datos de vídeo incluye además la decodificación de un registro de configuración de decodificador de 3VC del cuadro de configuración de 3VC, en el que el cuadro de configuración de 3VC está incluido en al menos una entrada de muestra de tipo '3vc1' que admite extractores y una entrada de muestra de tipo '3vc2' que no admite extractores. En un ejemplo, el proceso de procesamiento de datos de vídeo incluye la codificación de una anchura y una altura de las vistas de profundidad en el registro de configuración del decodificador de 3VC, en el que 3VC significa codificación de vídeo tridimensional.

La FIG. 14 es un diagrama de flujo que ilustra un método de ejemplo para analizar un archivo multimedia que contiene vídeo de múltiples vistas en el que una textura y una vista de profundidad de una vista particular se almacenan en pistas separadas según las técnicas descritas en esta divulgación. El proceso descrito en la FIG. 14 puede ser realizado por un dispositivo de codificación de vídeo que incluye el módulo de desencapsulación 29 descrito en este documento. Un dispositivo de codificación de vídeo puede analizar una pista de datos de vídeo, en la que la pista incluye una o más vistas, entre ellas, sólo una de las vistas de textura de una vista particular y una vista de profundidad de la vista particular (260). Por ejemplo, el módulo de desencapsulación 29 puede analizar un archivo multimedia, como el archivo multimedia de múltiples vistas con profundidad 142 del dispositivo fuente 12, que incluye una vista de textura de una vista particular y una vista de profundidad de la vista particular almacenada en pistas separadas.

El dispositivo de codificación de vídeo puede analizar una referencia de pista para determinar una dependencia de la pista a una pista de referencia indicada en la referencia de pista (262). Por ejemplo, el módulo de desencapsulación 29 puede analizar un tipo de referencia de pista "deps" que indica que la pista incluye la vista de profundidad de la vista particular y la pista de referencia incluye la vista de textura de la vista particular. En otro ejemplo, el módulo de desencapsulación 29 puede analizar un tipo de referencia de pista "tref" que indica que la pista depende de la vista de textura de la vista particular que está almacenada en la pista de referencia. En otro ejemplo, el módulo de desencapsulación 29 puede analizar un tipo de referencia de pista "dref" que indica que la pista depende de la vista de profundidad de la vista particular que está almacenada en la pista de referencia.

En algunos ejemplos, el método puede incluir además el análisis de un cuadro del identificador de vista a partir de al menos una de las entradas de muestra y una entrada de grupo de múltiples vistas para determinar, para cada vista en la pista, si la vista es una vista de textura o una vista de profundidad, donde al menos una de las entradas de muestra y la entrada de grupo de múltiples vistas están asociadas a la pista. En otro ejemplo, el cuadro del identificador de vista puede analizarse para determinar si se requiere una vista de textura o una vista de profundidad de una vista de referencia para decodificar una vista específica en la pista.

En otro ejemplo, el proceso incluye analizar un cuadro de mensaje de información de mejora complementaria (SEI) para determinar un mensaje de SEI de información de escalabilidad tridimensional asociado con una o más de las vistas. En los ejemplos en los que la pista contiene la vista de profundidad de la vista particular, el método puede incluir además el análisis de un cuadro de resolución de profundidad de 3VC para determinar una resolución espacial la vista de profundidad de la vista particular. En otro ejemplo en el que la pista contiene la vista de profundidad de la vista particular, el método comprende el análisis de un registro de configuración del decodificador de codificación de vídeo tridimensional (3VC) para determinar una anchura y una altura de la vista de profundidad de la vista particular.

En otro ejemplo en el que la pista es una pista de codificación de vídeo tridimensional (3VC), el método puede incluir además el análisis de un registro de configuración del decodificador de 3VC, en el que el registro de configuración del decodificador de 3VC indica un registro de configuración para una entrada de muestra coincidente de los datos de vídeo.

En otro ejemplo de método de procesamiento de datos de vídeo, el método incluye la decodificación de un valor de tipo 'tref' para una pista para determinar que la pista depende de una vista de textura almacenada en una pista referenciada. En otro ejemplo de método de procesamiento de datos de vídeo, el método incluye la decodificación de un valor de tipo 'dref' para una pista para determinar que la pista depende de una vista de profundidad almacenada en una pista referenciada. En otros ejemplos, cualquiera de los procesos descritos en el presente documento puede contener una o más de las técnicas descritas en esta divulgación.

La FIG. 15 es un diagrama de flujo que ilustra un método de ejemplo para crear un archivo multimedia con contenido de vídeo de múltiples vistas en el que una textura y una vista de profundidad de una vista particular se almacenan en pistas separadas según las técnicas descritas en esta divulgación. El proceso descrito en la FIG. 15 puede ser realizado por un dispositivo de codificación de vídeo que incluye el módulo de encapsulación 21 descrito en este documento. Un dispositivo de codificación de vídeo puede componer una pista de datos de vídeo, en la que la pista incluye una o más vistas, entre ellas, sólo una de las vistas de textura de una vista particular y una vista de profundidad

de la vista particular (266). Por ejemplo, el módulo de encapsulación 21 puede componer un archivo multimedia, como el archivo multimedia de múltiples vistas con profundidad 142, que incluye una vista de textura de una vista particular y una vista de profundidad de la vista particular almacenada en pistas separadas.

El dispositivo de codificación de vídeo puede componer una referencia de pista para determinar una dependencia de la pista a una pista de referencia indicada en la referencia de pista (266). Por ejemplo, el módulo de encapsulación 21 puede componer un tipo de referencia de pista "deps" que indica que la pista incluye la vista de profundidad de la vista particular y la pista de referencia incluye la vista de textura de la vista particular. En otro ejemplo, el módulo de encapsulación 21 puede componer un tipo de referencia de pista "tref" que indica que la pista depende de la vista de textura de la vista particular que está almacenada en la pista de referencia. En otro ejemplo, el módulo de encapsulación 21 puede componer un tipo de referencia de pista "dref" que indica que la pista depende de la vista de profundidad de la vista particular que está almacenada en la pista de referencia.

En algunos ejemplos, el método puede incluir además la composición de un cuadro del identificador de vista a partir de al menos una de las entradas de muestra y una entrada de grupo de múltiples vistas para indicar, para cada vista en la pista, si la vista es una vista de textura o una vista de profundidad, donde al menos una de las entradas de muestra y la entrada de grupo de múltiples vistas están asociadas a la pista. En otro ejemplo, el cuadro del identificador de vista puede componerse para indicar si se requiere una vista de textura o una vista de profundidad de una vista de referencia para decodificar una vista específica en la pista.

En otro ejemplo, el proceso incluye componer un cuadro de mensaje de información de mejora complementaria (SEI) para indicar un mensaje de SEI de información de escalabilidad tridimensional asociado a una o más de las vistas. En los ejemplos en los que la pista contiene la vista de profundidad de la vista particular, el método puede incluir además la composición de un cuadro de resolución de profundidad de 3VC para indicar una resolución espacial la vista de profundidad de la vista particular. En otro ejemplo en el que la pista contiene la vista de profundidad de la vista particular, el método incluye la composición de un registro de configuración del decodificador de codificación de vídeo tridimensional (3VC) para indicar una anchura y una altura de la vista de profundidad de la vista particular.

En otro ejemplo en el que la pista es una pista de codificación de vídeo tridimensional (3VC), el método puede incluir además la composición de un registro de configuración del decodificador de 3VC, en el que el registro de configuración del decodificador de 3VC indica un registro de configuración para una entrada de muestra coincidente de los datos de vídeo.

En otro ejemplo de método de procesamiento de datos de vídeo, el método incluye el almacenamiento de una vista de textura y una vista de profundidad de una vista particular en pistas separadas. En otro ejemplo de método de procesamiento de datos de vídeo, el método incluye la codificación de un valor de tipo 'tref' para una pista para indicar que la pista depende de una vista de textura almacenada en una pista referenciada. En otro ejemplo de método de procesamiento de datos de vídeo, el método incluye la codificación de un valor de tipo 'dref' para una pista para indicar que la pista depende de una vista de profundidad almacenada en una pista referenciada. En otros ejemplos, cualquiera de los procesos descritos en el presente documento puede contener una o más de las técnicas descritas en esta divulgación.

En otro ejemplo, un método de procesamiento de datos de vídeo incluye la decodificación de un registro de configuración del decodificador, en el que el registro de configuración del decodificador se define como "RegistroConfiguraciónDecodificador3VC", que incluye una misma estructura sintáctica que "RegistroConfiguraciónDecodificadorMVC" y al menos alguna semántica diferente de la de "RegistroConfiguraciónDecodificadorMVC" para indicar características relacionadas con un flujo de vídeo de 3VC.

En otro ejemplo, un método de procesamiento de datos de vídeo incluye decodificar un tipo de entrada de muestra para pistas de 3VC con unidades NAL de profundidad solamente, en el que el tipo de entrada se selecciona de un grupo que incluye un primer tipo en el que no se admiten extractores y un segundo tipo en el que se admiten extractores.

En otro ejemplo, un método de procesamiento de datos de vídeo incluye la decodificación de un tipo de entrada de muestra de un grupo de tipos de entrada de muestra que incluye: 1) tipo de entrada de muestra 'avc1' que contiene configuraciones de AVC, MVC y 3VC, para pistas de 3VC con unidades NAL de AVC, MVC y profundidad; 2) tipo de entrada de muestra 'mvc1' que contiene configuraciones de MVC y 3VC, para pistas de 3VC sin unidades NAL de AVC pero con unidades NAL de MVC y profundidad en las que no se admiten extractores; y 3) tipo de entrada de muestra 'mvc2' que contiene configuraciones de MVC y 3VC, para pistas de 3VC sin unidades NAL de AVC pero con unidades NAL de MVC y profundidad en las que se admiten extractores.

En otro ejemplo, un método de procesamiento de datos de vídeo incluye la decodificación de un atributo para indicar al menos uno de los siguientes aspectos: si los puntos de operación de 3VC indicados en un cuadro de grupo de múltiples vistas están presentes, si un punto de operación indicado en un cuadro de grupo de múltiples vistas es de 3VC o de MVC, y si las vistas de salida indicadas en un cuadro de grupo de múltiples vistas contienen sólo textura, sólo profundidad, o ambas.

En otro ejemplo, un método de procesamiento de datos de vídeo incluye decodificar un cuadro de mensaje de SEI que contiene un mensaje de SEI de información de escalabilidad tridimensional definido en una especificación de códec MVC+D. En otro ejemplo, un método de procesamiento de datos de vídeo incluye decodificar un valor de nombre de compresor de "/Codificación 0123VC" para un archivo que contiene un flujo de vídeo de codificación tridimensional (3VC).

En otro ejemplo, un método de procesamiento de datos de vídeo incluye la codificación de un registro de configuración del decodificador, en el que el registro de configuración del decodificador se define como "RegistroConfiguraciónDecodificador3VC", que incluye una misma estructura sintáctica que "RegistroConfiguraciónDecodificadorMVC" y al menos alguna semántica diferente de la de "RegistroConfiguraciónDecodificadorMVC" para indicar características relacionadas con un flujo de vídeo de codificación tridimensional (3VC). En otro ejemplo, un método de procesamiento de datos de vídeo incluye la codificación de un tipo de entrada de muestra para pistas de 3VC con unidades NAL de profundidad solamente, donde el tipo de entrada se selecciona de un grupo que incluye un primer tipo en el que no se admiten extractores y un segundo tipo en el que se admiten extractores, donde NAL significa capa de abstracción de red y 3VC significa codificación de vídeo tridimensional.

En otro ejemplo, un método de procesamiento de datos de vídeo incluye la codificación de un tipo de entrada de muestra de un grupo de tipos de entrada de muestra que incluye: 1) tipo de entrada de muestra 'avc1' que contiene configuraciones de AVC, MVC y 3VC, para pistas de 3VC con unidades NAL de AVC, MVC y profundidad; 2) tipo de entrada de muestra 'mvc1' que contiene configuraciones de MVC y 3VC, para pistas de 3VC sin unidades NAL de AVC pero con unidades NAL de MVC y profundidad en las que no se admiten extractores; y 3) tipo de entrada de muestra 'mvc2' que contiene configuraciones de MVC y 3VC, para pistas de 3VC sin unidades NAL de AVC pero con unidades NAL de MVC y profundidad en las que se admiten extractores.

En otro ejemplo, un método de procesamiento de datos de vídeo incluye la codificación de un atributo para indicar al menos uno de los siguientes aspectos: si los puntos de operación de 3VC indicados en un cuadro de grupo de múltiples vistas están presentes, si un punto de operación indicado en un cuadro de grupo de múltiples vistas es de 3VC o de MVC, y si las vistas de salida indicadas en un cuadro de grupo de múltiples vistas contienen sólo textura, sólo profundidad, o ambas. En algunos ejemplos, el atributo es uno de los atributos comunes o de los atributos diferenciadores.

En otro ejemplo, un método de procesamiento de datos de vídeo incluye la codificación de un cuadro de mensaje de información de mejora complementaria (SEI) que contiene un mensaje de SEI de información de escalabilidad tridimensional definido en una especificación de códec de codificación de múltiples vistas más profundidad (MVC+D). En otro ejemplo, un método de procesamiento de datos de vídeo incluye codificar un valor de nombre de compresor de "/Codificación VC 0123" para un archivo que contiene un flujo de vídeo de codificación tridimensional (3VC).

Esta divulgación también incluye el siguiente anexo, que se basa en la última especificación integrada de 14496-15 en el documento de salida MPEG W13036, cuyo contenido completo se incorpora aquí a modo de referencia. En el siguiente anexo, las partes no modificadas no se muestran en general, sino que se sustituyen por la indicación "[Véase documento original]" dentro del anexo. La especificación original de 14496-15 en el documento de salida W13036 de MPEG está protegida por los derechos de autor de ISO/IEC.

Las técnicas de esta divulgación pueden ser implementadas por dispositivos de red tales como los dispositivos de enrutamiento 104 y el dispositivo de transcodificación 106, pero también pueden ser implementadas por el dispositivo cliente 108. De este modo, los dispositivos de enrutamiento 104, el dispositivo de transcodificación 106 y el dispositivo cliente 108 representan ejemplos de dispositivos configurados para realizar las técnicas de la presente divulgación, entre ellas, las técnicas mencionadas en la parte de las reivindicaciones de la presente divulgación. Además, los dispositivos de la FIG. 1, el codificador mostrado en la FIG. 2 y el decodificador mostrado en la FIG. 3 son también dispositivos ejemplares que pueden configurarse para llevar a cabo las técnicas de esta divulgación, que incluyen las técnicas mencionadas en la parte de las reivindicaciones de esta divulgación.

En uno o más ejemplos, las funciones descritas pueden implementarse en hardware, software, firmware o cualquiera de sus combinaciones. Si se implementan en software, las funciones pueden almacenarse o transmitirse, como una o más instrucciones o código, a través de un medio legible por ordenador y ejecutarse por una unidad de procesamiento basada en hardware. Los medios legibles por ordenador pueden incluir medios de almacenamiento legibles por ordenador, que corresponden a un medio tangible como los medios de almacenamiento de datos, o medios de comunicación que incluyen cualquier medio que facilite la transferencia de un programa informático de un lugar a otro, por ejemplo, según un protocolo de comunicación. De este modo, los medios legibles por ordenador pueden corresponder generalmente a (1) un medio de almacenamiento tangible legible por ordenador que no es transitorio o (2) un medio de comunicación como una señal u onda portadora. Los medios de almacenamiento de datos pueden ser cualquier medio disponible al que puedan acceder uno o más ordenadores o uno o más procesadores para recuperar instrucciones, código y/o estructuras de datos para la implementación de las técnicas descritas en esta divulgación. Un producto de programa de ordenador puede incluir un medio legible por ordenador.

En otros ejemplos, la presente divulgación contempla un medio legible por ordenador que comprende una estructura de datos almacenada en este, en la que la estructura de datos incluye un flujo de bits codificado que está codificado de acuerdo con la presente divulgación.

A modo de ejemplo, y no restrictivo, dichos medios legibles por ordenador pueden comprender RAM, ROM, EEPROM, CD-ROM u otro almacenamiento en disco óptico, almacenamiento en disco magnético u otros dispositivos de almacenamiento magnético, memoria flash o cualquier otro medio que se pueda utilizar para almacenar el código de programa deseado en forma de instrucciones o estructuras de datos y al que se pueda acceder mediante un ordenador. Además, cualquier conexión se denomina correctamente un medio legible por ordenador. Por ejemplo, si las instrucciones se transmiten desde un sitio web, un servidor u otra fuente remota a través de un cable coaxial, un cable de fibra óptica, un par trenzado, una línea de suscripción digital (DSL) o tecnologías inalámbricas como los infrarrojos, la radio y las microondas, el cable coaxial, el cable de fibra óptica, el par trenzado, la DSL o las tecnologías inalámbricas como los infrarrojos, la radio y las microondas se incluyen en la definición de medio. Debe entenderse, sin embargo, que los medios de almacenamiento legibles por ordenador y los medios de almacenamiento de datos no incluyen conexiones, ondas portadoras, señales u otros medios transitorios, sino que se refieren a medios de almacenamiento tangibles no transitorios. El disco, como se usa en la presente, incluye disco compacto (CD), disco láser, disco óptico, disco versátil digital (DVD), disquete y disco Blu-ray, donde los discos generalmente reproducen datos magnéticamente, mientras que los discos reproducen datos ópticamente con láser. Las combinaciones de los anteriores también deben incluirse dentro del alcance de los medios legibles por ordenador.

Las instrucciones pueden ser ejecutadas por uno o más procesadores, como uno o más procesadores de señales digitales (DSP), microprocesadores de propósito general, circuitos integrados de aplicación específica (ASIC), matrices de puertas programables en campo (FPGA), u otros circuitos lógicos integrados o discretos equivalentes. Por consiguiente, el término "procesador", como se usa en la presente, puede referirse a cualquiera de las estructuras anteriores o cualquier otra estructura adecuada para la implementación de las técnicas descritas en la presente. Además, en algunos aspectos, la funcionalidad descrita en el presente documento puede proporcionarse dentro de módulos de hardware y/o software dedicados, configurados para la codificación y decodificación, o incorporados en un códec combinado. Asimismo, las técnicas podrían implementarse completamente en uno o más circuitos o elementos lógicos.

Las técnicas de esta divulgación pueden implementarse en una amplia variedad de dispositivos o aparatos, incluido un microteléfono, un circuito integrado (IC) o un conjunto de IC (por ejemplo, un conjunto de chips). En la presente divulgación se describen diversos componentes, módulos o unidades para enfatizar los aspectos funcionales de los dispositivos configurados para llevar a cabo las técnicas divulgadas, pero no requieren necesariamente la realización de diferentes unidades de hardware. Por el contrario, como se ha descrito anteriormente, diversas unidades pueden combinarse en una unidad de hardware de códec o ser proporcionadas por una colección de unidades de hardware interoperativas, que incluyen uno o más procesadores según lo descrito con anterioridad, junto con un software y/o firmware adecuado.

REIVINDICACIONES

1. Un método de procesamiento de datos de vídeo, donde el método comprende:

5 el análisis de una pista de datos de vídeo de múltiples vistas, donde la pista incluye al menos una vista de profundidad; y se caracteriza porque comprende, además:

el análisis de información para determinar una resolución espacial asociada con la vista de profundidad, donde la determinación de la resolución espacial comprende analizar la resolución espacial de un cuadro de resolución de profundidad de codificación de vídeo tridimensional de un cuadro de configuración de codificación de vídeo tridimensional; y

el análisis de un registro de configuración del decodificador de codificación de vídeo tridimensional desde el cuadro de configuración de codificación de vídeo tridimensional,

donde el cuadro de configuración de codificación de vídeo tridimensional está incluido en al menos una de las entradas de muestra de tipo "3vc1" que admite extractores y una entrada de muestra de tipo "3vc2" que no admite extractores.

2. Un dispositivo que comprende:

una memoria configurada para almacenar datos de vídeo; y
20 uno o más procesadores configurados para:

analizar una pista de datos de vídeo de múltiples vistas, donde la pista incluye una o más vistas, entre ellas, al menos una vista de profundidad; y se caracteriza porque uno o más procesadores están configurados además para:

analizar la información para determinar una resolución espacial asociada con la vista de profundidad, donde la determinación de la resolución espacial comprende analizar la resolución espacial de un cuadro de resolución de profundidad de codificación de vídeo tridimensional de un cuadro de configuración de codificación de vídeo tridimensional; y

analizar un registro de configuración del decodificador de codificación de vídeo tridimensional desde el cuadro de configuración de codificación de vídeo tridimensional, donde el cuadro de configuración de codificación de vídeo tridimensional está incluido en al menos una de las entradas de muestra de tipo "3vc1" que admite extractores y una entrada de muestra de tipo "3vc2" que no admite extractores.

3. El dispositivo de la reivindicación 2, configurado además para:

analizar un cuadro del identificador de vista de al menos una de las entradas de muestra y una entrada de grupo de múltiples vistas para determinar al menos una de las siguientes cosas: si se requiere una vista de textura o una vista de profundidad de una vista de referencia para decodificar al menos una de las vistas de la pista, si la pista incluye sólo vistas de profundidad o vistas de textura y profundidad, y, para cada vista incluida en la pista, si la vista es una vista de textura o una vista de profundidad, en la que al menos una de las entradas de muestra y la entrada de grupo de múltiples vistas están asociadas a la pista.

4. El dispositivo de la reivindicación 2, donde los datos de vídeo de múltiples vistas son un primer conjunto de datos de vídeo de múltiples vistas y donde la pista incluye además al menos una vista de textura, también está configurado para:

analizar un segundo conjunto de datos de vídeo de múltiples vistas basado en el primer conjunto de datos de vídeo de múltiples vistas, en el que la información de análisis comprende información de análisis para determinar si la primera vista de textura, la primera vista de profundidad, o ambas, la primera textura y la primera vista de profundidad del primer conjunto de datos de vídeo de múltiples vistas, son necesarias para decodificar el segundo conjunto de datos de vídeo de múltiples vistas.

5. El dispositivo de la reivindicación 2, configurado además para:

analizar un cuadro de mensaje de información de mejora complementaria (SEI) para determinar un mensaje de SEI de información de escalabilidad tridimensional asociado con una o más de las vistas; y/o

analizar un registro de configuración de decodificación de vídeo tridimensional (3VC) para determinar una anchura y una altura de la vista de profundidad.

6. El dispositivo de la reivindicación 2, donde la pista es una primera pista, también está configurado para:

analizar una segunda pista de datos de vídeo de múltiples vistas, en la que la segunda pista incluye una vista de textura correspondiente a la vista de profundidad, en la que la vista de textura no se incluye en la primera pista.

7. Un método de procesamiento de datos de vídeo, donde el método comprende:

la composición de una pista de datos de vídeo de múltiples vistas, en la que la pista incluye una o más vistas que comprende al menos una vista de profundidad; y se caracteriza porque comprende, además:

la composición de información para indicar una resolución espacial asociada con la vista de profundidad, donde la indicación de la resolución espacial comprende la composición de la resolución espacial en un cuadro de resolución de profundidad de codificación de vídeo tridimensional en un cuadro de configuración de codificación de vídeo tridimensional; y

la composición de un registro de configuración del decodificador de codificación de vídeo tridimensional en el cuadro de configuración de codificación de vídeo tridimensional, donde el cuadro de configuración de codificación de vídeo tridimensional está incluido en al menos una de las entradas de muestra de tipo "3vc1" que admite extractores y una entrada de muestra de tipo "3vc2" que no admite extractores.

8. Un dispositivo que comprende:

una memoria configurada para almacenar datos de vídeo; y uno o más procesadores configurados para:

componer una pista de datos de vídeo de múltiples vistas, en la que la pista incluye una o más vistas, entre ellas, al menos una vista de profundidad; y se caracteriza porque uno o más procesadores están configurados además para:

componer información para indicar una resolución espacial asociada con la vista de profundidad, donde la indicación de la resolución espacial comprende la composición de la resolución espacial en un cuadro de resolución de profundidad de codificación de vídeo tridimensional en un cuadro de configuración de codificación de vídeo tridimensional; y

componer un registro de configuración del decodificador de codificación de vídeo tridimensional en el cuadro de configuración de codificación de vídeo tridimensional, donde el cuadro de configuración de codificación de vídeo tridimensional está incluido en al menos una de las entradas de muestra de tipo "3vc1" que admite extractores y una entrada de muestra de tipo "3vc2" que no admite extractores.

9. El dispositivo de la reivindicación 8, configurado además para:

componer un cuadro del identificador de vista de al menos una de las entradas de muestra y una entrada de grupo de múltiples vistas para indicar al menos una de las siguientes cosas: si se requiere una vista de textura o una vista de profundidad de una vista de referencia para decodificar al menos una de las vistas de la pista, si la pista incluye sólo vistas de profundidad o vistas de textura y profundidad, y, para cada vista incluida en la pista, si la vista es una vista de textura o una vista de profundidad, en la que al menos una de las entradas de muestra y la entrada de grupo de múltiples vistas están asociadas a la pista.

10. El dispositivo de la reivindicación 8, donde los datos de vídeo de múltiples vistas son un primer conjunto de datos de vídeo de múltiples vistas y donde la pista incluye además al menos una vista de textura, también está configurado para:

componer un segundo conjunto de datos de vídeo de múltiples vistas basado en el primer conjunto de datos de vídeo de múltiples vistas, en el que la información de composición comprende además información de composición para indicar si la primera vista de textura, la primera vista de profundidad, o ambas, la primera textura y la primera vista de profundidad del primer conjunto de datos de vídeo de múltiples vistas, son necesarias para decodificar el segundo conjunto de datos de vídeo de múltiples vistas.

11. El dispositivo de la reivindicación 8, configurado además para:

componer un cuadro de mensaje de información de mejora complementaria (SEI) para indicar un mensaje de SEI de información de escalabilidad tridimensional asociado con una o más de las vistas; y/o componer un registro de configuración del decodificador de vídeo tridimensional (3VC) para indicar una anchura y una altura de la vista de profundidad.

12. El dispositivo de la reivindicación 8, en el que la pista es una primera pista, también está configurado para:

componer una segunda pista de datos de vídeo de múltiples vistas, en la que la segunda pista incluye una vista de textura correspondiente a la vista de profundidad, en la que la vista de textura no se incluye en la primera pista.

13. Un medio de almacenamiento legible por ordenador no transitorio con instrucciones almacenadas en él que, al ejecutarse, hacen que uno o más procesadores de un dispositivo de codificación de vídeo realicen el método según la reivindicación 1 o la reivindicación 7.

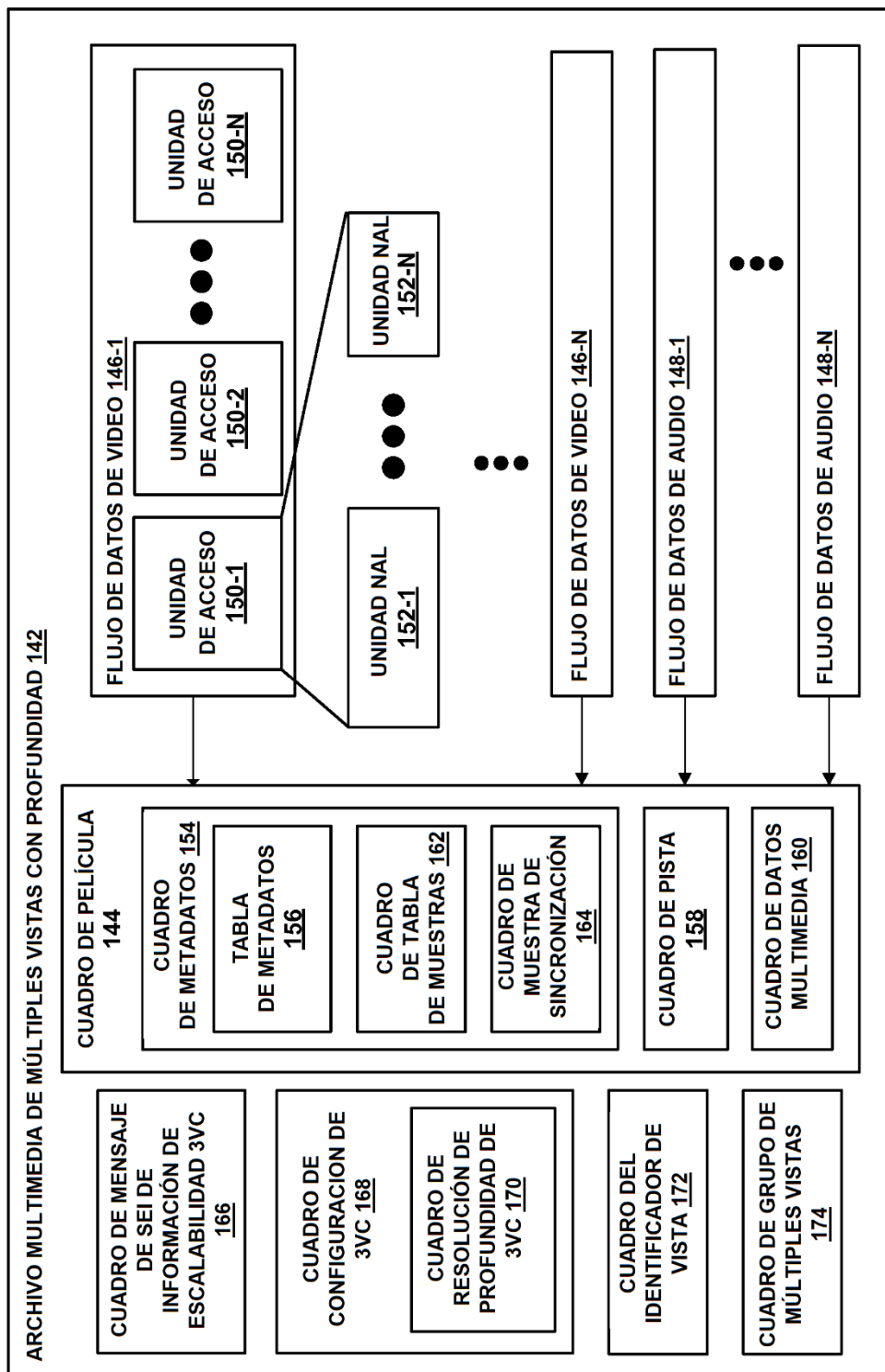


FIG. 1

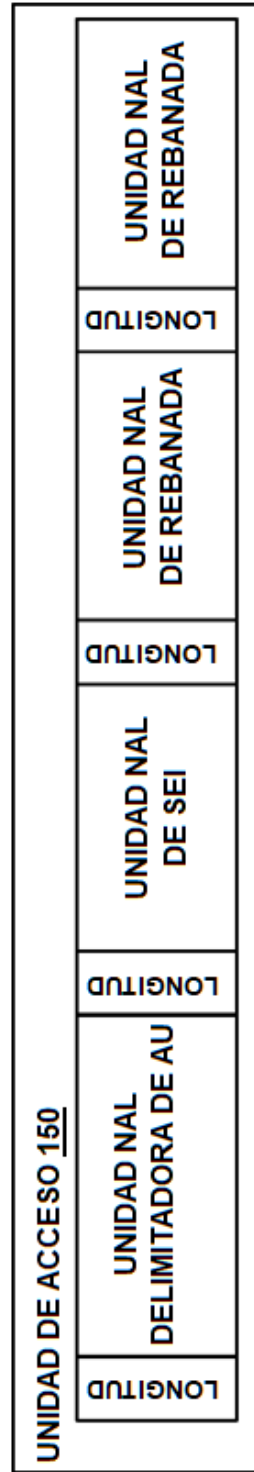


FIG. 2

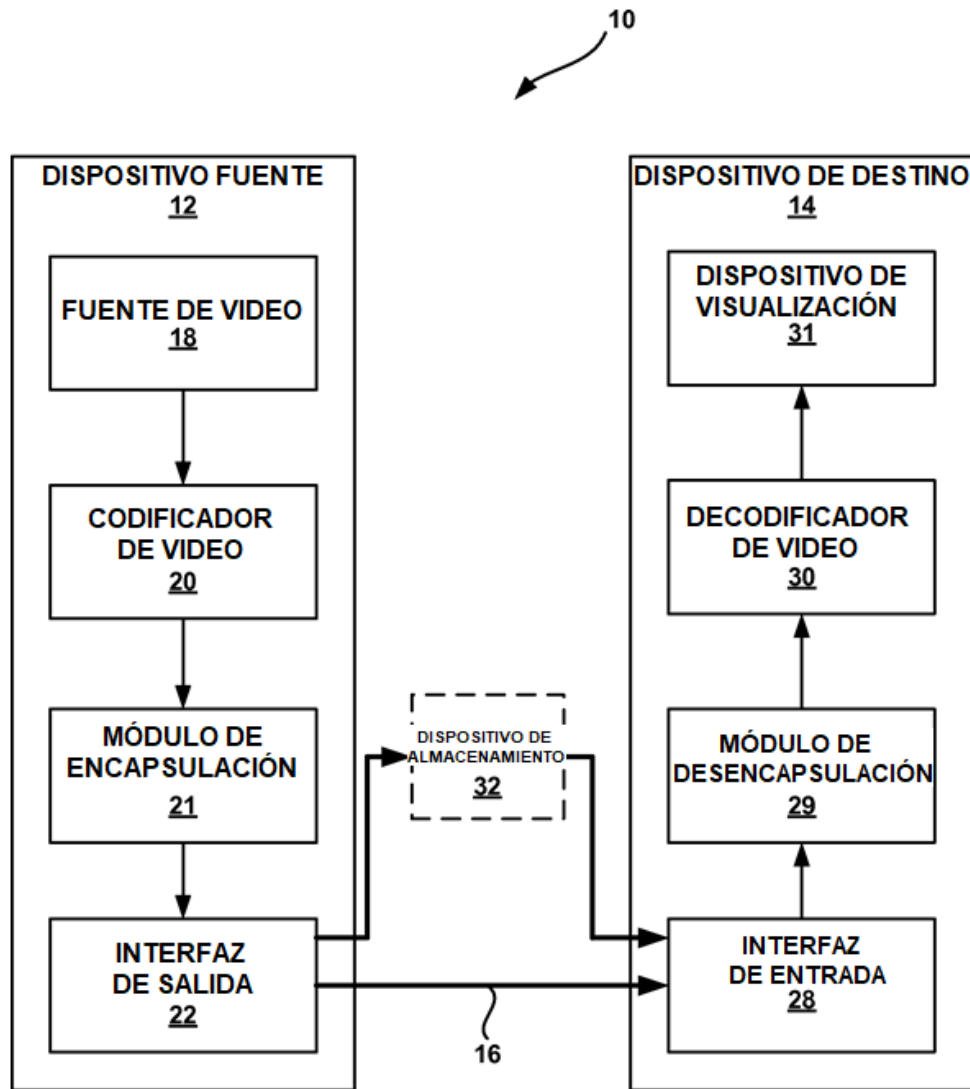


FIG. 3

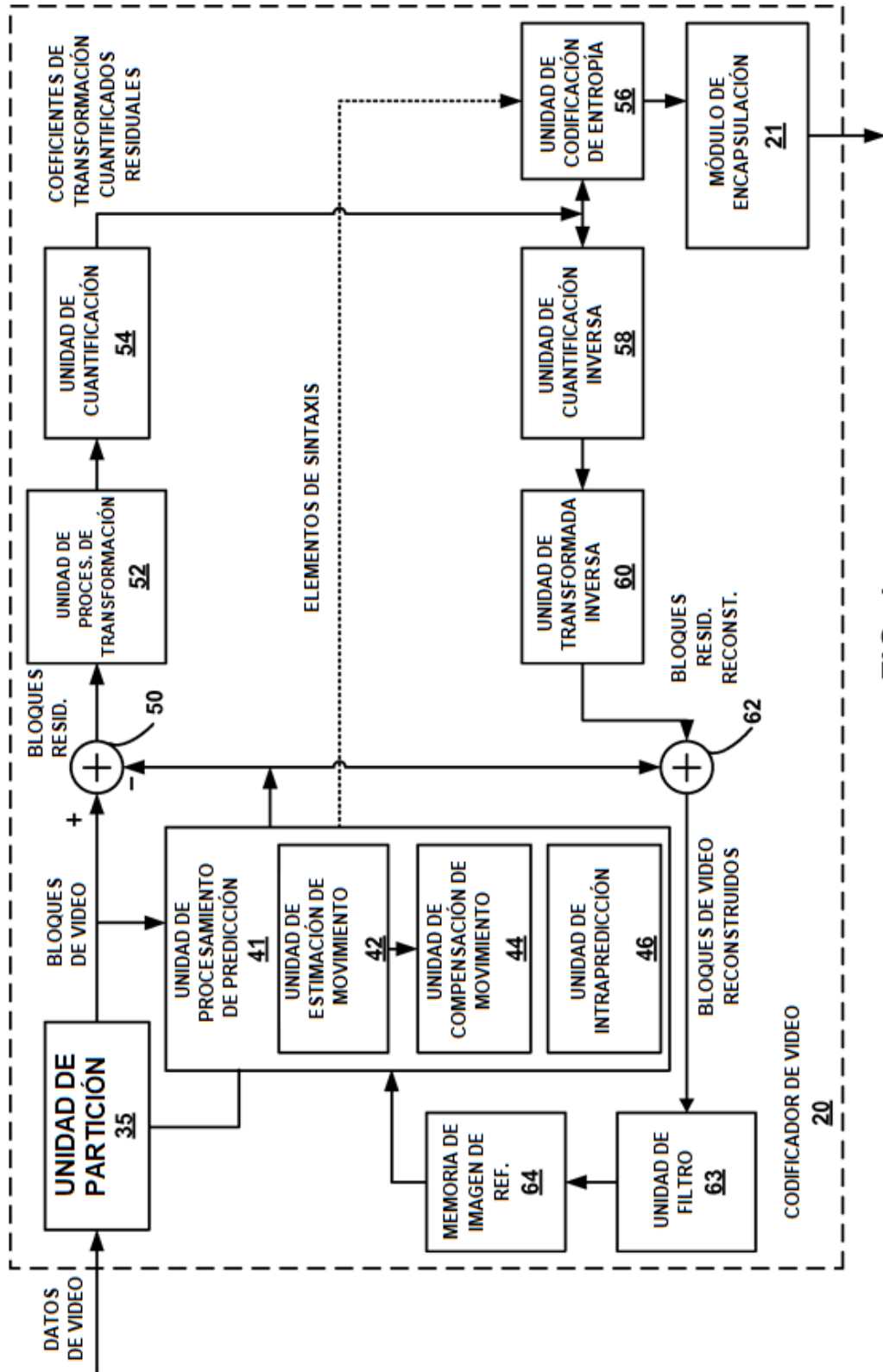


FIG. 4

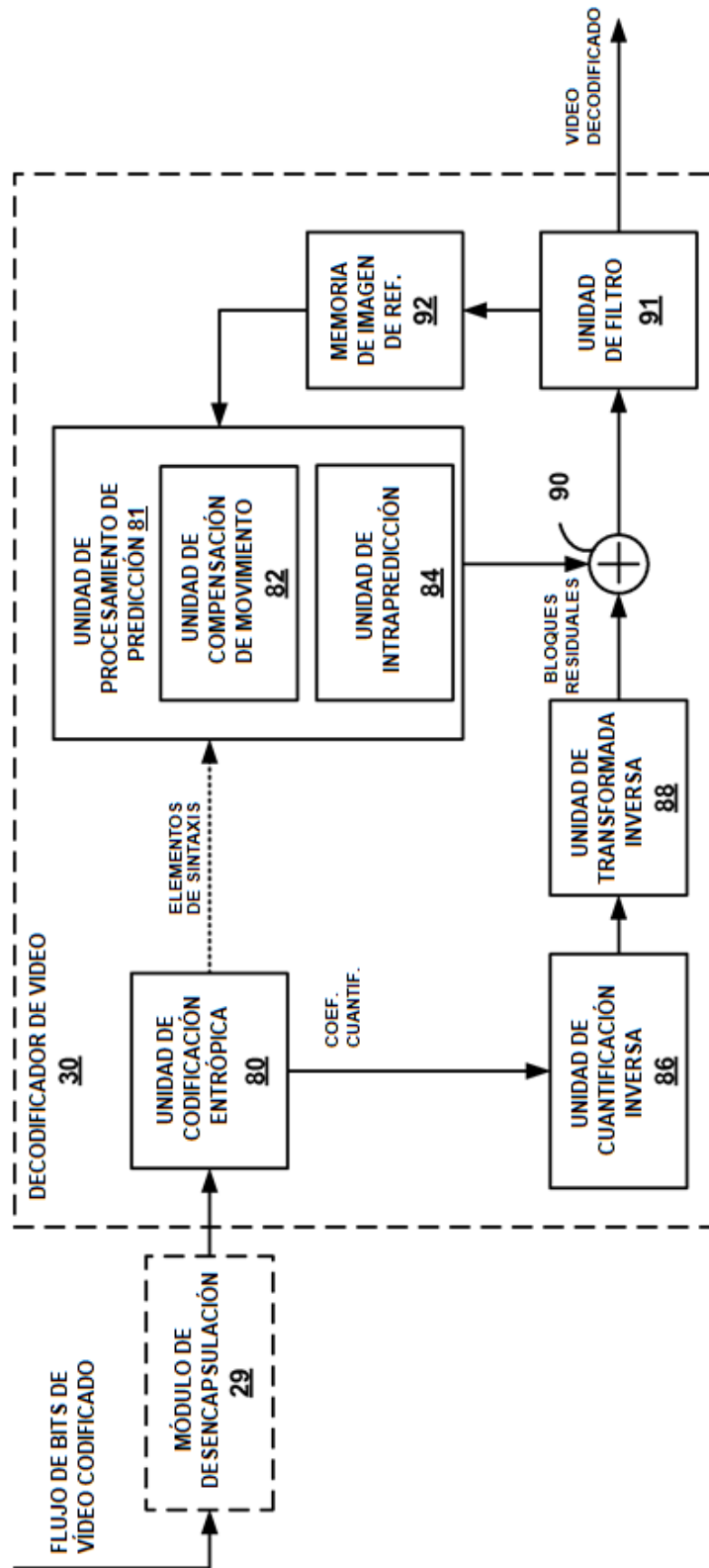


FIG. 5

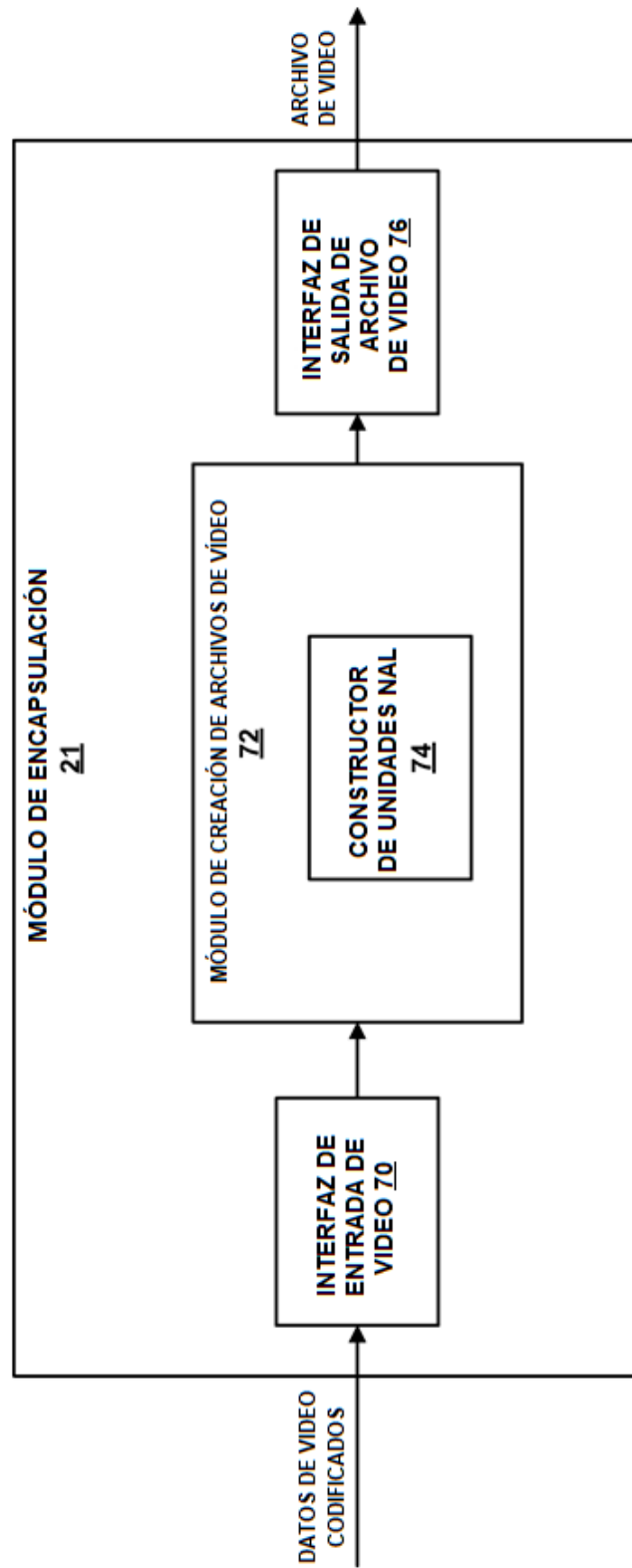


FIG. 6

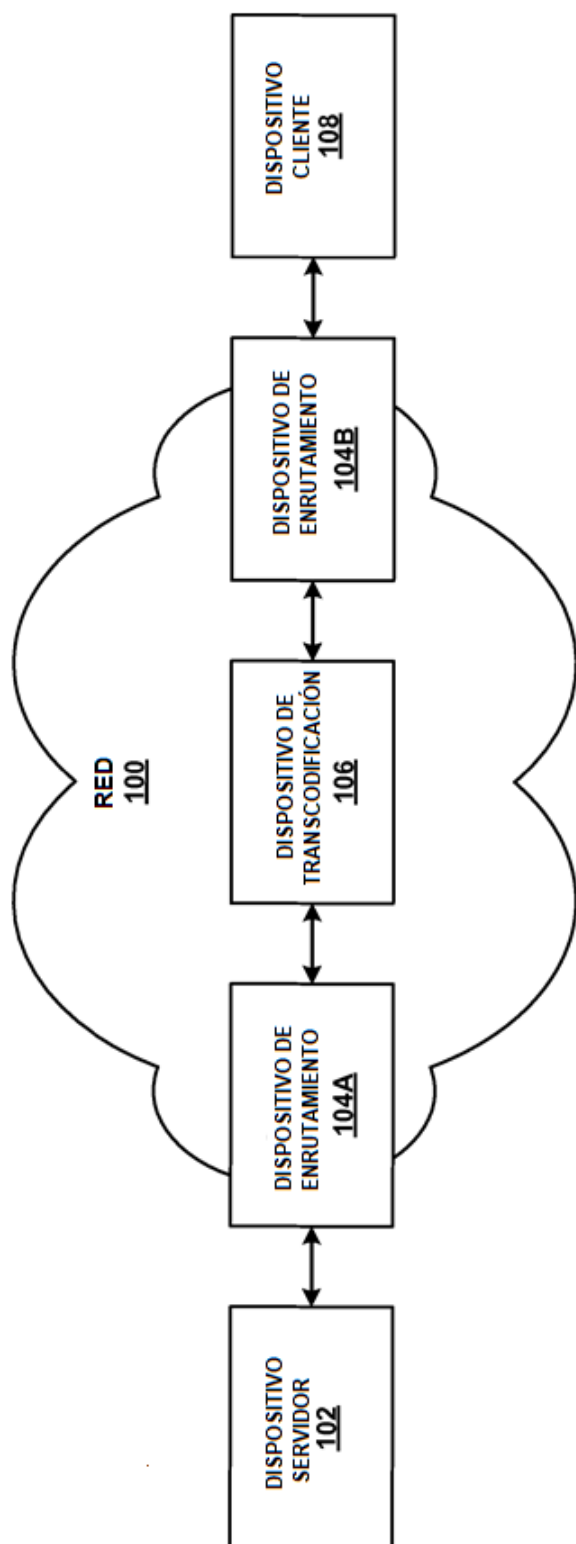


FIG. 7

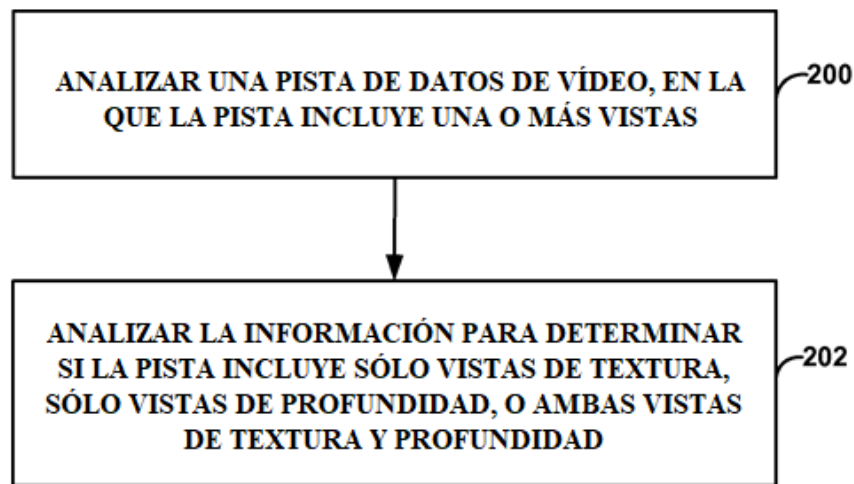


FIG. 8

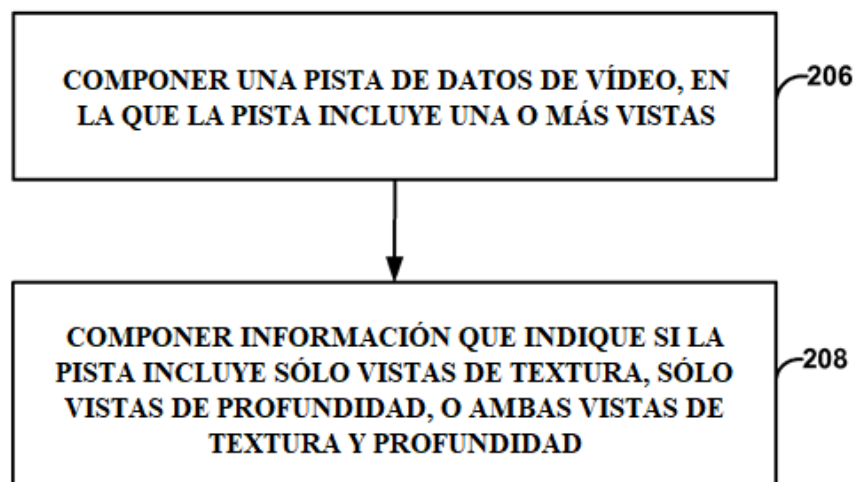


FIG. 9

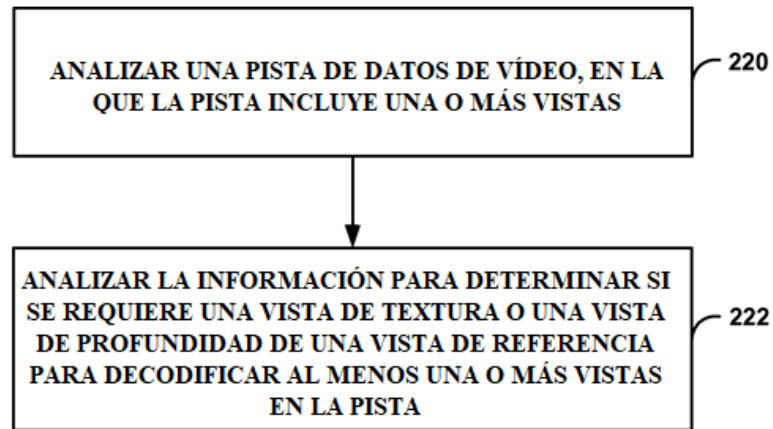


FIG. 10

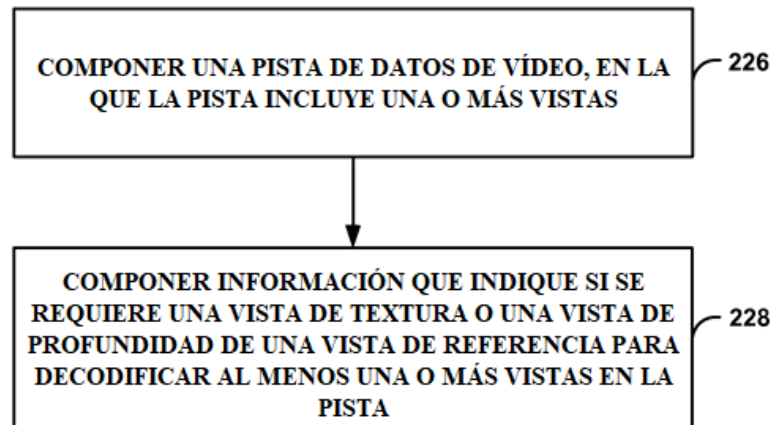


FIG. 11

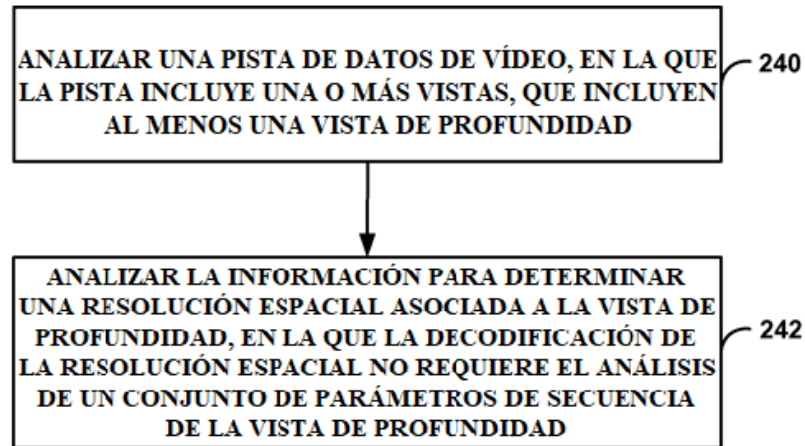


FIG. 12

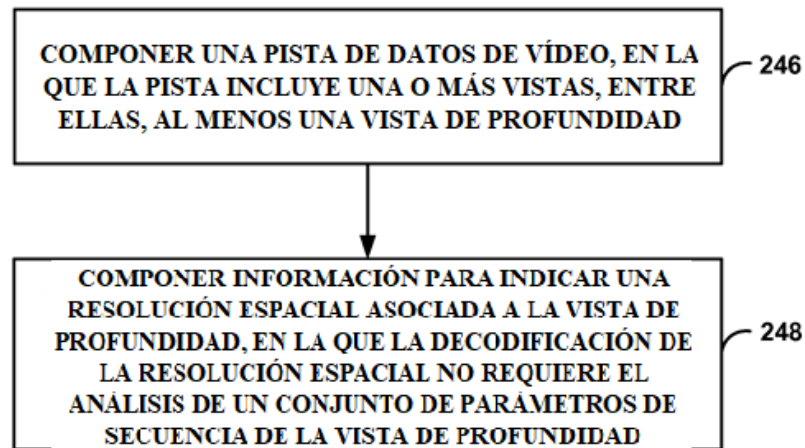


FIG. 13

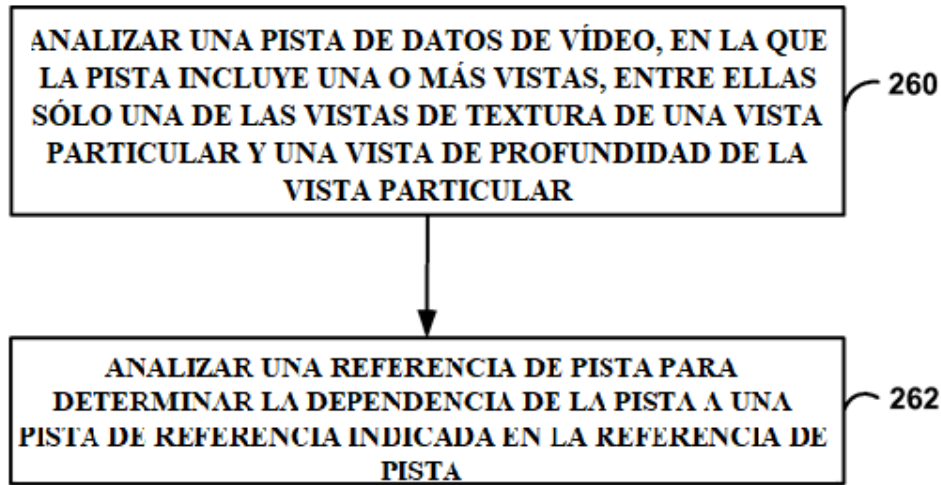


FIG. 14

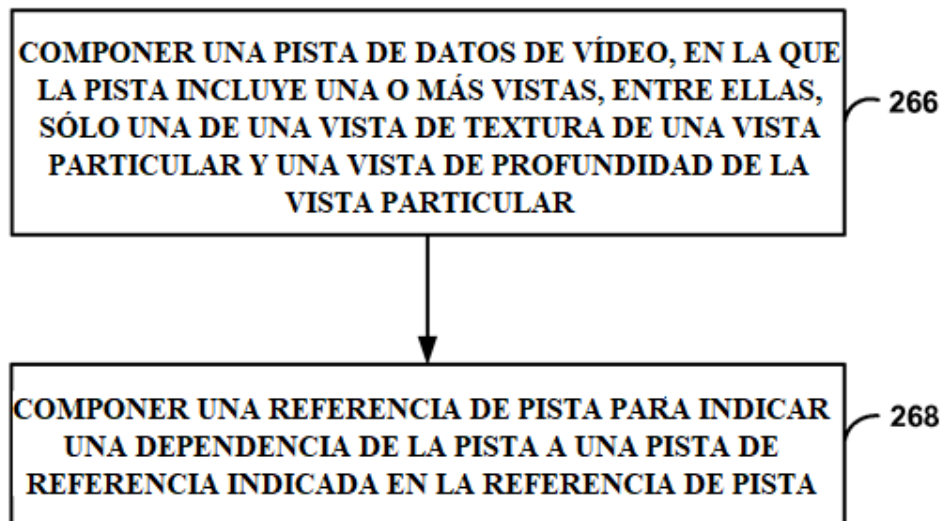


FIG. 15