

19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 718 818**

51 Int. Cl.:

H04N 19/597 (2014.01)

H04N 19/176 (2014.01)

H04N 19/70 (2014.01)

H04N 19/423 (2014.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

86 Fecha de presentación y número de la solicitud internacional: **30.11.2012 PCT/US2012/067303**

87 Fecha y número de publicación internacional: **06.06.2013 WO13082431**

96 Fecha de presentación y número de la solicitud europea: **30.11.2012 E 12806754 (3)**

97 Fecha y número de publicación de la concesión europea: **02.01.2019 EP 2786573**

54 Título: **Información de nivel de secuencia para la codificación de vídeo tridimensional (3DVC) compatible con codificación de vídeo de múltiples vistas (MVC)**

30 Prioridad:

30.11.2011 US 201161565376 P

01.12.2011 US 201161565938 P

22.12.2011 US 201161579631 P

06.01.2012 US 201261584009 P

29.11.2012 US 201213689560

45 Fecha de publicación y mención en BOPI de la traducción de la patente:
04.07.2019

73 Titular/es:

QUALCOMM INCORPORATED (100.0%)

5775 Morehouse Drive

San Diego, CA 92121-1714, US

72 Inventor/es:

CHEN, YING y

WANG, YE-KUI

74 Agente/Representante:

FORTEA LAGUNA, Juan José

ES 2 718 818 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

Información de nivel de secuencia para la codificación de vídeo tridimensional (3DVC) compatible con codificación de vídeo de múltiples vistas (MVC)

5

CAMPO TÉCNICO

[0001] La presente divulgación se refiere a la codificación de vídeo y, más en particular, a la codificación de vídeo tridimensional (3DVC).

10

ANTECEDENTES

[0002] Las capacidades del vídeo digital se pueden incorporar a una amplia gama de dispositivos, incluyendo televisores digitales, sistemas de radiodifusión digital directa, sistemas de radiodifusión inalámbrica, asistentes digitales personales (PDA), ordenadores portátiles o de escritorio, ordenadores de tableta, lectores de libros electrónicos, cámaras digitales, dispositivos de grabación digital, reproductores de medios digitales, dispositivos de videojuegos, consolas de videojuegos, teléfonos celulares o de radio por satélite, los llamados "teléfonos inteligentes", dispositivos de videoconferencia, dispositivos de transmisión continua de vídeo y similares. Los dispositivos de vídeo digital implementan técnicas de compresión de vídeo, tales como las descritas en las normas definidas por MPEG-2, MPEG-4, ITU-T H.263, ITU-T H.264/MPEG-4, Parte 10, Codificación de Vídeo Avanzada (AVC), la norma de Codificación de Vídeo de Alta Eficiencia (HEVC) actualmente en desarrollo y las extensiones de dichas normas. Los dispositivos de vídeo pueden transmitir, recibir, codificar, descodificar y/o almacenar información de vídeo digital más eficazmente implementando dichas técnicas de compresión de vídeo.

15

20

25

30

[0003] Las técnicas de compresión de vídeo llevan a cabo la predicción espacial (intra-imagen) y/o la predicción temporal (inter-imagen) para reducir o eliminar la redundancia intrínseca en las secuencias de vídeo. Para la codificación de vídeo basada en bloques, un fragmento de vídeo (es decir, una imagen o una parte de una imagen) se puede dividir en bloques de vídeo, que también se pueden denominar bloques arbolados, unidades de codificación (CU) y/o nodos de codificación. Los bloques de vídeo en un fragmento intracodificado (I) de una imagen se codifican usando predicción espacial con respecto a muestras de referencia en bloques contiguos en la misma imagen. Los bloques de vídeo en un fragmento intercodificado (P o B) de una imagen pueden usar la predicción espacial con respecto a muestras de referencia en bloques contiguos en la misma imagen o la predicción temporal con respecto a muestras de referencia en otras imágenes de referencia.

35

40

45

[0004] La predicción espacial o temporal da como resultado un bloque predictivo para un bloque que se va a codificar. Los datos residuales representan diferencias de píxeles entre el bloque original que se va a codificar y el bloque predictivo. Un bloque intercodificado se codifica de acuerdo con un vector de movimiento que apunta a un bloque de muestras de referencia que forman el bloque predictivo, y los datos residuales que indican la diferencia entre el bloque codificado y el bloque predictivo. Un bloque intracodificado se codifica de acuerdo con un modo de intracodificación y los datos residuales. Para una mayor compresión, los datos residuales se pueden transformar desde el dominio de píxeles a un dominio de transformada, dando como resultado coeficientes de transformada residuales, que a continuación se pueden cuantificar. Los coeficientes de transformada cuantificados, dispuestos inicialmente en una formación bidimensional, se pueden examinar a fin de producir un vector unidimensional de coeficientes de transformada, y se puede aplicar codificación por entropía para lograr aún más compresión.

50

El documento US 2010/011183 A1 describe un procedimiento de descodificación de una señal de vídeo. Describe la búsqueda de un bloque de referencia usando información lateral, derivando la información de movimiento de un bloque actual del bloque de referencia, descodificando el bloque actual usando la información de movimiento del bloque actual. La información lateral incluye información de desplazamiento que indica una diferencia de posición entre el bloque de referencia y un bloque correspondiente y en la que el bloque correspondiente es un bloque indicado por un movimiento global que indica una diferencia de disparidad entre una vista del bloque actual y una vista del bloque de referencia.

55

Un documento relevante de la técnica anterior que describe un punto de funcionamiento es el borrador conjunto de MVC: ANÓNIMO: "Joint Draft 7.0 on Multiview Video Coding [Borrador conjunto 7.0 sobre la codificación de vídeo de múltiples vistas]", 27. REUNIÓN DEL JVT; 6-4-2008 - 10-4-2008; GINEBRA; (EQUIPO DE VÍDEO CONJUNTO DE LA ISO/IEC JTC1/SC29/WG11 E ITU-T SG.16), n.º JVT-AA209, 6 de junio de 2008 (2008-06-06), XP030007391, ISSN: 0000-0063.

60

RESUMEN

[0005] La invención se define en las reivindicaciones independientes a las que ahora se debe hacer referencia.

65

[0006] En general, esta divulgación describe técnicas para la codificación de vídeo tridimensional (3DVC). Más en particular, la presente divulgación describe técnicas para llevar a cabo 3DVC usando una extensión 3DVC de la norma H.264/Codificación de Vídeo Avanzada (AVC). La extensión 3DVC define una tecnología de codificación de vídeo para

codificar múltiples vistas de datos de vídeo con datos de profundidad. Cada vista puede corresponder a una perspectiva, o ángulo, diferente, en la que se captan los datos de vídeo correspondientes de una escena común. En el contexto de 3DVC, cada vista puede contener una vista en textura y una vista en profundidad. Una representación codificada de una vista en un caso momentáneo es un componente de vista. Un componente de vista puede contener un componente de vista en profundidad y un componente de vista en textura. Las técnicas de la presente divulgación en general se refieren a permitir tanto el tratamiento de componentes de textura como de componentes de profundidad de una vista para 3DVC cuando se codifican datos de vídeo de múltiples vistas más datos de profundidad. Las técnicas pueden promover la compatibilidad de 3DVC con MVC.

[0007] En un ejemplo, un procedimiento de codificación de datos de vídeo que incluye un componente de vista que comprende un componente de profundidad y un componente de textura comprende activar un conjunto de parámetros como un conjunto de parámetros de textura para el componente de textura del componente de vista y codificar el componente de textura del componente de vista basándose en el conjunto de parámetros de textura activados.

[0008] En otro ejemplo, un dispositivo de codificación de vídeo para codificar datos de vídeo que incluye un componente de vista compuesto de un componente de profundidad y un componente de textura comprende un procesador configurado para activar un conjunto de parámetros como un conjunto de parámetros de textura para el componente de textura del componente de vista, y codificar el componente de textura del componente de vista basándose en el conjunto de parámetros de textura activado.

[0009] En otro ejemplo, un dispositivo de codificación de vídeo para codificar datos de vídeo que incluye un componente de vista compuesto de un componente de profundidad y un componente de textura comprende medios para activar un conjunto de parámetros como un conjunto de parámetros de textura para el componente de textura del componente de vista y medios para codificar el componente de textura del componente de vista basándose en el conjunto de parámetros de textura activado.

[0010] En otro ejemplo, un medio de almacenamiento legible por ordenador no transitorio ha almacenado en el mismo instrucciones que, al ejecutarse, hacen que uno o más procesadores de un dispositivo de codificación de vídeo activen un conjunto de parámetros como un conjunto de parámetros de textura para el componente de textura del componente de vista, y codifique el componente de textura del componente de vista basándose en el conjunto de parámetros de textura activado.

[0011] En otro ejemplo, se describe un procedimiento de procesamiento de datos de vídeo que incluye un componente de vista que comprende un componente de profundidad y un componente de textura. El procedimiento comprende determinar un mensaje de información de mejora complementaria que se aplica al procesar el componente de vista de los datos de vídeo, y determinar un mensaje de información de mejora complementaria anidado que se aplica al componente de profundidad del componente de vista además del mensaje de información de mejora complementaria.

[0012] En otro ejemplo, se describe un dispositivo para procesar datos de vídeo que incluyen un componente de vista compuesto de un componente de profundidad y un componente de textura. El dispositivo comprende un procesador configurado para determinar un mensaje de información de mejora complementaria que se aplica al procesar el componente de vista de los datos de vídeo, y determinar un mensaje de información de mejora complementaria anidado que se aplica además del mensaje de información de mejora complementaria al procesar el componente de profundidad del componente de vista.

[0013] En otro ejemplo, se describe un dispositivo para procesar datos de vídeo que incluyen un componente de vista compuesto de un componente de profundidad y un componente de textura. El dispositivo comprende medios para determinar un mensaje de información de mejora complementaria que se aplica al procesar el componente de vista de los datos de vídeo, y medios para determinar un mensaje de información de mejora complementaria anidado que se aplica además del mensaje de información de mejora complementaria al procesar el componente de profundidad del componente de vista.

[0014] En otro ejemplo, un medio de almacenamiento legible por ordenador no transitorio ha almacenado en el mismo instrucciones que, al ejecutarse, hacen que uno o más procesadores de un dispositivo para procesar datos de vídeo que incluyen un componente de vista compuesto de un componente de profundidad y un componente de textura determinen un mensaje de información de mejora complementaria que se aplica al procesar un componente de vista de los datos de vídeo, en el que el componente de vista incluye tanto un componente de profundidad como un componente de textura y determinen un mensaje de información de mejora complementaria anidado que se aplica además del mensaje de información de mejora complementaria al procesar el componente de profundidad del componente de vista.

[0015] En otro ejemplo, un procedimiento para codificar vídeo comprende almacenar un componente de profundidad en una memoria intermedia de imágenes descodificadas, analizando una dependencia de la vista para determinar si el componente de profundidad se usa para la predicción entre vistas y eliminando el componente de profundidad de la memoria intermedia de imágenes descodificadas en respuesta a determinar que el componente de profundidad no se usa para la predicción entre vistas.

5 [0016] En otro ejemplo, un dispositivo de codificación de vídeo configurado para codificar datos de vídeo comprende una memoria intermedia de imágenes descodificadas y un procesador configurado para almacenar un componente de profundidad en la memoria intermedia de imágenes descodificadas, analizar una dependencia de la vista para determinar si el componente de profundidad se usa para la predicción entre vistas y eliminar el componente de profundidad de la memoria intermedia de imágenes descodificadas en respuesta a determinar que el componente de profundidad no se usa para la predicción entre vistas.

10 [0017] En otro ejemplo, un dispositivo de codificación de vídeo para codificar datos de vídeo comprende medios para almacenar un componente de profundidad en una memoria intermedia de imágenes descodificadas, medios para analizar una dependencia de la vista para determinar si el componente de profundidad se usa para la predicción entre vistas, y medios para eliminar el componente de profundidad de la memoria intermedia de imágenes descodificadas en respuesta a determinar que el componente de profundidad no se usa para la predicción entre vistas.

15 [0018] En otro ejemplo, un medio de almacenamiento legible por ordenador no transitorio ha almacenado en el mismo instrucciones que, al ejecutarse, hacen que uno o más procesadores de un dispositivo de codificación de vídeo almacenen un componente de profundidad en una memoria intermedia de imágenes descodificadas, analicen una dependencia de la vista para determinar si el componente de profundidad se usa para la predicción entre vistas, y eliminen el componente de profundidad de la memoria intermedia de imágenes descodificadas en respuesta a determinar que el componente de profundidad no se usa para la predicción entre vistas.

20 [0019] En otro ejemplo, un procedimiento de procesamiento de datos de vídeo incluye un componente de vista que comprende uno o más componentes de profundidad y uno o más componentes de textura, el procedimiento comprende determinar la información del primer nivel de secuencia que describe las características de los componentes de profundidad, y determinar la información del segundo nivel de secuencia que describe las características de un punto de funcionamiento de los datos de vídeo.

25 [0020] En otro ejemplo, se describe un dispositivo de codificación de vídeo para procesar datos de vídeo que incluyen un componente de vista que comprende uno o más componentes de profundidad y uno o más componentes de textura. El dispositivo de codificación de vídeo comprende uno o más procesadores configurados para determinar la información del primer nivel de secuencia que describe las características de los componentes de profundidad, y determinar la información del segundo nivel de secuencia que describe las características de un punto de funcionamiento de los datos de vídeo.

30 [0021] En otro ejemplo, se describe un dispositivo de codificación de vídeo para procesar datos de vídeo que incluyen un componente de vista que comprende uno o más componentes de profundidad y uno o más componentes de textura. El dispositivo de codificación de vídeo comprende medios para determinar la información del primer nivel de secuencia que describe las características de los componentes de profundidad y medios para determinar la información del segundo nivel de secuencia que describe las características de un punto de funcionamiento de los datos de vídeo.

35 [0022] En otro ejemplo, un medio de almacenamiento legible por ordenador no transitorio ha almacenado en el mismo instrucciones que, al ejecutarse, hacen que uno o más procesadores de un dispositivo de codificación de vídeo determinen la información del primer nivel de secuencia que describe las características de uno o más componentes de profundidad de datos de vídeo, en el que los datos de vídeo incluyen un componente de vista que comprende el uno o más componentes de profundidad y uno o más componentes de textura y determinen la información del segundo nivel de secuencia que describe las características de un punto de funcionamiento de los datos de vídeo.

40 [0023] Los detalles de uno o más aspectos de la divulgación se exponen en los dibujos adjuntos y la descripción siguiente. Otras características, objetos y ventajas de las técnicas descritas en la presente divulgación resultarán evidentes a partir de la descripción y de los dibujos, y a partir de las reivindicaciones.

BREVE DESCRIPCIÓN DE LOS DIBUJOS

55 [0024]

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

60 La FIG. 2 es un diagrama de bloques que ilustra el codificador de vídeo mostrado en el ejemplo de la FIG. 1 con más detalle.

La FIG. 3 es un diagrama de bloques que ilustra el descodificador de vídeo mostrado en el ejemplo de la FIG. 1 con más detalle.

65 La FIG. 4 es un diagrama de bloques que ilustra la unidad de encapsulación mostrada en el ejemplo de la FIG. 1 con más detalle.

La FIG. 5 es un diagrama conceptual que ilustra un patrón de predicción de la Codificación de Vídeo de Múltiples vistas (MVC) de ejemplo.

5 La FIG. 6 es un diagrama de flujo que ilustra el funcionamiento de un dispositivo de codificación de vídeo en la implementación de la activación del conjunto de parámetros para 3DVC compatible con MVC de acuerdo con diversos aspectos de las técnicas descritas en esta divulgación.

10 La FIG. 7 es un diagrama de flujo que ilustra el funcionamiento de ejemplo del procesamiento de datos de vídeo de múltiples vistas para generar mensajes de información de mejora complementaria (SEI) anidados para 3DVC compatible con MVC de acuerdo con las técnicas descritas en esta divulgación.

15 La FIG. 8 es un diagrama de flujo que ilustra el funcionamiento de ejemplo de un dispositivo de codificación de vídeo en la eliminación por separado de los componentes de textura y profundidad de una memoria intermedia de imágenes descodificadas para 3DVC compatible con MVC de acuerdo con las técnicas descritas en esta divulgación.

20 La FIG. 9 es un diagrama de flujo que ilustra el funcionamiento de ejemplo de un dispositivo de codificación de vídeo en la determinación de la información del nivel de secuencia para 3DVC compatible con MVC de acuerdo con las técnicas descritas en esta divulgación.

DESCRIPCIÓN DETALLADA

25 **[0025]** De acuerdo con determinados sistemas de codificación de vídeo, la estimación de movimiento y la compensación de movimiento se pueden usar para reducir la redundancia temporal en una secuencia de vídeo, con el fin de lograr la compresión de datos. En este caso, se puede generar un vector de movimiento que identifica un bloque predictivo de datos de vídeo, por ejemplo, un bloque de otra imagen o fragmento de vídeo, que se puede usar para predecir los valores del bloque de vídeo actual que se está codificando. Los valores del bloque de vídeo predictivo se restan de los valores del bloque de vídeo actual para producir un bloque de datos residuales. La información de movimiento (por ejemplo, un vector de movimiento, índices de vectores de movimiento, direcciones de predicción u otra información) se comunica desde un codificador de vídeo a un descodificador de vídeo, junto con los datos residuales. El descodificador puede ubicar el mismo bloque predictivo (basándose en el vector de movimiento) y reconstruir el bloque de vídeo codificado combinando los datos residuales con los datos del bloque predictivo.

30 **[0026]** La codificación de vídeo de múltiples vistas (MVC) es un proceso de codificación de vídeo para codificar múltiples vistas de datos de vídeo. En general, cada vista se corresponde con una perspectiva, o ángulo, diferente, en el que se captaron los datos de vídeo correspondientes de una escena común. La codificación de vídeo tridimensional (3DVC) se puede llevar a cabo usando codificación de MVC más profundidad. Una extensión 3DVC a la norma ITU-T H.264/AVC está actualmente en fase de desarrollo. Un borrador de trabajo de una enmienda a la norma H.264/AVC para añadir la extensión 3DVC se describe en "WD on MVC Extensions for Inclusion of Depth Maps [WD en extensiones MVC para la inclusión de mapas de profundidad]", ISO/IEC/JTC1/SC29/WG11/N12351, Ginebra, Suiza, con fecha de noviembre de 2011 ("3DVC WD"). La extensión 3DVC, también denominada extensiones MVC para la inclusión de mapas de profundidad, define técnicas para codificar vistas para admitir la visualización de datos de vídeo 3D.

35 **[0027]** Por ejemplo, en el vídeo 3D, dos vistas (por ejemplo, las vistas de los ojos izquierdo y derecho de un observador humano) se pueden visualizar simultáneamente, o casi simultáneamente, usando diferentes polarizaciones de luz, y un espectador puede llevar gafas polarizadas pasivas de modo que cada uno de los ojos del espectador reciba una de las vistas respectivas. De forma alternativa, el espectador puede llevar gafas activas que obturan cada ojo independientemente, y una pantalla puede alternar rápidamente entre las imágenes de cada ojo de forma sincronizada con las gafas.

40 **[0028]** Si bien cada vista (por ejemplo, las vistas de los ojos izquierdo y derecho) se puede codificar individualmente, en 3DVC, una de las vistas se reconstruye a partir de la otra vista usando un componente de profundidad de la vista. Por este motivo, esta forma de 3DVC también se puede denominar codificación de vídeo de múltiples vistas más profundidad (MVC + D). Para ilustrar, un mapa de profundidad de una imagen particular de una vista (donde esta imagen particular de una vista se puede denominar un "componente de vista" de la vista) se puede calcular como una diferencia entre la vista del ojo izquierdo y la vista del ojo derecho. El codificador puede codificar la vista del ojo izquierdo, por ejemplo, como un llamado "componente de textura" del componente de vista y el mapa de profundidad se puede codificar como un llamado "componente de profundidad" del componente de vista.

45 **[0029]** El descodificador, a continuación, puede descodificar el componente de textura del componente de vista y el componente de profundidad del componente de vista y reconstruir la vista del ojo derecho a partir del componente de textura (que representa la vista del ojo izquierdo) usando el componente de profundidad. Codificando solo una vista y un mapa de profundidad correspondiente de esta manera, 3DVC puede codificar más eficazmente la vista, tanto del

ojo izquierdo como del ojo derecho, en comparación con la codificación, tanto de la vista del ojo izquierdo como de la vista del ojo derecho, independientemente como vistas separadas de los datos de 3DVC.

[0030] Al codificar los componentes de textura y profundidad de una vista, el codificador de vídeo típicamente trata o procesa de otro modo tanto los componentes de textura como de profundidad como componentes de vista sin proporcionar ninguna manera por la que se distingue entre los componentes de textura y profundidad. Es decir, 3DVC proporciona el tratamiento o codificación de los componentes de vista sin proporcionar una manera por la que se procesan individualmente los componentes de textura por separado de los componentes de profundidad del mismo componente de vista. Esta falta de distinción en 3DVC puede dar como resultado una menor eficacia de codificación y/o una menor calidad de los datos de vídeo reconstruidos.

[0031] Por ejemplo, se puede requerir actualmente que se especifique un componente de profundidad a la misma resolución que un componente de textura correspondiente con el fin de dar cabida al tratamiento conjunto de este componente de vista. Sin embargo, un componente de profundidad con resolución más alta (en comparación con la resolución del componente de textura) puede dar como resultado una reproducción de vídeo tridimensional (3D), en la que se puede obtener un vídeo 3D mejor y más envolvente que imite mejor lo que el sistema visual de un espectador espera. Además, un componente de profundidad con resolución más baja (en comparación con la resolución del componente de textura) puede proporcionar la misma o similar experiencia 3D envolvente en determinados casos, pero consume menos bits cuando se codifica y, de este modo, incrementa la eficacia de la codificación. No permitiendo el tratamiento separado de los componentes de profundidad y textura, 3DVC puede reducir la eficacia de la codificación y/o proporcionar una menor calidad de los datos de vídeo reconstruidos (a menudo, en términos, de la experiencia visual).

[0032] Las técnicas de la presente divulgación se refieren en general a permitir el tratamiento separado o independiente de los componentes de textura y los componentes de profundidad de una vista al procesar o codificar los datos de vídeo 3DVC. Por ejemplo, la presente divulgación propone señalar el tamaño de imagen de una secuencia de mapa de profundidad en un conjunto de parámetros de secuencia (SPS). Estas técnicas de señalización se pueden aplicar por un codificador y usar por un descodificador durante el proceso de codificación y/o descodificación de vídeo. Las técnicas descritas se refieren a la codificación de imágenes de contenido de vídeo. Las imágenes codificadas pueden tener un tamaño de unidad, tal como un bloque de una altura y anchura seleccionadas, que se pueden señalar como elementos sintácticos en un conjunto de parámetros de secuencia, de acuerdo con las técnicas de la presente divulgación. Los elementos sintácticos para las secuencias de vista de textura y las secuencias de mapas de profundidad se pueden señalar en un conjunto de parámetros de secuencia.

[0033] Más específicamente, las técnicas implican señalar la información de sintaxis cuando una secuencia de mapa de profundidad tiene una resolución diferente de la secuencia de vista de textura correspondiente. 3DVC puede incluir codificar múltiples vistas con una secuencia de mapa de profundidad para cada vista. Estas secuencias de mapas de profundidad pueden tener una resolución diferente a las secuencias de vista de textura. En este caso, el componente de vista de profundidad y el componente de vista de textura no pueden compartir el mismo conjunto de parámetros de secuencia (SPS) cuando las unidades de capa de abstracción de red (NAL) de textura y profundidad simplemente se multiplexan conjuntamente. Puede que no sea posible indicar diferentes niveles, con o sin profundidad, en la extensión SPS MVC actual. En el principio de diseño de AVC, puede que no sea posible activar más de un conjunto de parámetros de secuencia, que contenga el tamaño de imagen. Por tanto, tener dos tamaños de imagen diferentes podría dar como resultado la activación de múltiples conjuntos de parámetros de secuencia.

[0034] Las técnicas de la presente divulgación se pueden usar para indicar una secuencia 3DVC basándose en estéreo AVC y MVC, cuando la secuencia 3DVC incluye un mapa de profundidad con una resolución espacial diferente a la de la vista de textura correspondiente. Permitiendo dicho tratamiento independiente, los diversos aspectos de las técnicas descritas en la presente divulgación pueden promover el ahorro de bits (o, en otras palabras, una codificación más eficaz de los datos de vídeo de múltiples vistas) y/o una mejor calidad de los datos de vídeo reconstruidos (que de nuevo se puede medir en términos de una experiencia visual percibida).

[0035] La siguiente descripción se debe entender como que está en el contexto de 3DVC, donde se entiende que la referencia a MVC es una referencia a MVC que se refiere a codificación de MVC más profundidad en la extensión 3DVC. Es decir, dado que MVC es una extensión para la H.264, y 3DVC es otra extensión de H.264 que hace uso de MVC, 3DVC incorpora o, de otro modo, se puede considerar que "hereda" todos los aspectos de MVC. 3DVC puede extender o de otro modo añadir a MVC cuando sea apropiado de la manera descrita en el presente documento para proporcionar un flujo de bits compatible con MVC que también incluya mapas de profundidad para los descodificadores de vídeo que admiten 3DVC. En otras palabras, en algunos ejemplos, las técnicas pueden proporcionar la generación de un flujo de bits de 3DVC que sea retrocompatible con MVC (o, en otras palabras, que se pueda descodificar por un descodificador de vídeo que no admita 3DVC pero que sí admita MVC). Si bien cada una de las siguientes técnicas se describe en el contexto de 3DVC, estas técnicas se pueden extender en algunos casos a otras maneras de codificación de datos de vídeo 3D que tienen tanto componentes de vista de textura como componentes de vista de profundidad.

[0036] La FIG. 1 es un diagrama de bloques que ilustra un sistema de codificación y decodificación de vídeo 10 de ejemplo que puede utilizar técnicas para predicción de vectores de movimiento en codificación de múltiples vistas. Como se muestra en la FIG. 1, el sistema 10 incluye un dispositivo de origen 12 que proporciona datos de vídeo codificados, que se van a decodificar más adelante por un dispositivo de destino 14. En particular, el dispositivo de origen 12 proporciona los datos de vídeo al dispositivo de destino 14 por medio de un medio legible por ordenador 16. El dispositivo de origen 12 y el dispositivo de destino 14 pueden comprender cualquiera de una amplia gama de dispositivos, incluyendo ordenadores de sobremesa, ordenadores plegables (es decir, portátiles), ordenadores de tableta, ordenadores de pizarra, decodificadores de televisión, equipos telefónicos de mano tales como los llamados teléfonos "inteligentes", los llamados blocs "inteligentes", televisores, cámaras, dispositivos de visualización, reproductores de medios digitales, consolas de videojuegos, dispositivo de transmisión continua de vídeo o similares. En algunos casos, el dispositivo de origen 12 y el dispositivo de destino 14 pueden estar equipados para la comunicación inalámbrica.

[0037] El dispositivo de destino 14 puede recibir los datos de vídeo codificados que se van a decodificar por medio del medio legible por ordenador 16. El medio legible por ordenador 16 puede comprender cualquier tipo de medio o dispositivo que puede transferir los datos de vídeo codificados desde el dispositivo de origen 12 al dispositivo de destino 14. En un ejemplo, el medio legible por ordenador 16 puede comprender un medio de comunicación para permitir al dispositivo de origen 12 transmitir datos de vídeo codificados directamente al dispositivo de destino 14 en tiempo real, o casi tiempo real.

[0038] Los datos de vídeo codificados se pueden modular de acuerdo con una norma de comunicación, tal como un protocolo de comunicación inalámbrica, y transmitir al dispositivo de destino 14. El medio de comunicación puede comprender cualquier medio de comunicación, inalámbrica o cableada, tal como un espectro de radiofrecuencia (RF) o una o más líneas de transmisión física. El medio de comunicación puede formar parte de una red basada en paquetes, tal como una red de área local, una red de área amplia o una red global tal como Internet. El medio de comunicación puede incluir 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.

[0039] En algunos ejemplos, se pueden emitir datos codificados desde el transmisor 24 del dispositivo de origen 24 a un dispositivo de almacenamiento. De forma similar, se puede acceder a los datos codificados desde el dispositivo de almacenamiento por el receptor 26 del dispositivo de destino 14. El dispositivo de almacenamiento puede incluir cualquiera de una variedad de medios de almacenamiento de datos, de acceso distribuido o local, tales como un disco duro, discos Blu-ray, DVD, CD-ROM, memoria flash, memoria volátil o no volátil o cualquier otro medio de almacenamiento digital adecuado para almacenar datos de vídeo codificados. En otro ejemplo, el dispositivo de almacenamiento puede corresponder a un servidor de archivos o a otro dispositivo de almacenamiento intermedio que pueda almacenar el vídeo codificado generado por el dispositivo de origen 12.

[0040] El dispositivo de destino 14 puede acceder a datos de vídeo almacenados desde el dispositivo de almacenamiento por medio de transmisión continua o descarga. El servidor de archivos puede ser cualquier tipo de servidor que pueda almacenar datos de vídeo codificados y transmitir esos datos de vídeo codificados al dispositivo de destino 14. Los servidores de archivos de ejemplo incluyen un servidor web (por ejemplo, para un sitio web), un servidor FTP, dispositivos de almacenamiento conectados en red (NAS) o una unidad de disco local. El dispositivo de destino 14 puede acceder a los datos de vídeo codificados a través de cualquier conexión de datos estándar, incluyendo una conexión a Internet. Esto puede incluir un canal inalámbrico (por ejemplo, una conexión wifi), una conexión cableada (por ejemplo, DSL, módem de cable, etc.) o una combinación de ambos que sea adecuada para acceder a datos de vídeo codificados almacenados en un servidor de archivos. La transmisión de datos de vídeo codificados desde el dispositivo de almacenamiento puede ser una transmisión continua, una transmisión de descarga o una combinación de ambas.

[0041] Las técnicas de la presente divulgación no están limitadas necesariamente a aplicaciones o contextos inalámbricos. Las técnicas se pueden aplicar a la codificación de vídeo como soporte de cualquiera de una variedad de aplicaciones de multimedia, tales como radiodifusiones de televisión por el aire, transmisiones de televisión por cable, transmisiones de televisión por satélite, transmisiones continuas de vídeo por Internet, tales como la transmisión continua adaptativa dinámica por HTTP (DASH), el vídeo digital que se codifica en un medio de almacenamiento de datos, la decodificación de vídeo digital almacenado en un medio de almacenamiento de datos, u otras aplicaciones. En algunos ejemplos, el sistema 10 se puede configurar para admitir la transmisión de vídeo unidireccional o bidireccional, para admitir aplicaciones tales como la transmisión continua de vídeo, la reproducción de vídeo, la radiodifusión de vídeo y/o la videotelefonía.

[0042] En el ejemplo de la FIG. 1, el dispositivo de origen 12 incluye un origen de vídeo 18, un codificador de vídeo 20, una unidad de encapsulación 21 y una interfaz de salida 22. El dispositivo de destino 14 incluye la interfaz de entrada 28, la unidad de desencapsulación 29, un decodificador de vídeo 30 y un dispositivo de visualización 32. En algunos casos, el dispositivo de origen 12 y el dispositivo de destino 14 pueden incluir otros componentes o disposiciones. Por ejemplo, el dispositivo de origen 12 puede recibir datos de vídeo desde un origen de vídeo externo 18, tal como una cámara externa. Asimismo, el dispositivo de destino 14 puede interactuar con un dispositivo de visualización externo, en lugar de incluir un dispositivo de visualización integrado 32.

[0043] El sistema 10 ilustrado de la FIG. 1 es simplemente un ejemplo. Las técnicas para la predicción de vectores de movimiento en la codificación de múltiples vistas (incluyendo la 3DVC) se pueden llevar a cabo por cualquier dispositivo de codificación y/o decodificación de vídeo digital. Aunque, en general, las técnicas de la presente divulgación se llevan a cabo por un dispositivo de codificación de vídeo, las técnicas también se pueden llevar a cabo por un codificador/decodificador de vídeo, denominado típicamente "CÓDEC". Además, las técnicas de la presente divulgación también se pueden llevar a cabo por un preprocesador de vídeo. El dispositivo de origen 12 y el dispositivo de destino 14 son simplemente ejemplos de dichos dispositivos de codificación en los que el dispositivo de origen 12 genera datos de vídeo codificados para su transmisión al dispositivo de destino 14. En algunos ejemplos, los dispositivos 12, 14 pueden funcionar de manera sustancialmente simétrica, de modo que cada uno de los dispositivos 12, 14 incluya componentes de codificación y decodificación de vídeo. De ahí que, el sistema 10 pueda admitir la transmisión de vídeo unidireccional o bidireccional entre los dispositivos de vídeo 12, 14, por ejemplo, para la transmisión continua de vídeo, la reproducción de vídeo, la radiodifusión de vídeo o la videotelefonía.

[0044] El origen de vídeo 18 del dispositivo de origen 12 puede incluir un dispositivo de captación de vídeo, tal como una videocámara, un archivo de vídeo que contiene vídeo captado previamente y/o una interfaz de alimentación de vídeo para recibir vídeo desde un proveedor de contenidos de vídeo. Como otra alternativa, el origen de vídeo 18 puede generar datos basados en gráficos por ordenador como el vídeo de origen, o una combinación de vídeo en directo, vídeo archivado y vídeo generado por ordenador. En algunos casos, si el origen de vídeo 18 es una videocámara, el dispositivo de origen 12 y el dispositivo de destino 14 pueden formar los llamados teléfonos con cámara o videoteléfonos. Sin embargo, como se menciona anteriormente, las técnicas descritas en la presente divulgación pueden ser aplicables a la codificación de vídeo en general, y se pueden aplicar a aplicaciones inalámbricas y/o cableadas. En cada caso, el codificador de vídeo 20 puede codificar el vídeo captado, precaptado o generado por ordenador. La información de vídeo codificada se puede emitir, a continuación, por la interfaz de salida 22 hacia un medio legible por ordenador 16.

[0045] El origen de vídeo 24 puede en general proporcionar una pluralidad de vistas de una escena al codificador de vídeo 28. El origen de vídeo 24 también puede proporcionar información indicativa de las ubicaciones de perspectivas de la cámara para las vistas. El origen de vídeo 24 puede proporcionar esta información al codificador de vídeo 28, o puede proporcionar la información directamente a la unidad de encapsulación 21.

[0046] La unidad de encapsulación 21 puede usar la información indicativa de las ubicaciones relativas de las perspectivas de la cámara para que las vistas asignen identificadores de vista a las vistas de contenido de multimedia. La unidad de encapsulación 21 puede formar una o más representaciones del contenido de multimedia, donde cada una de las representaciones puede incluir una o más vistas. En algunos ejemplos, el codificador de vídeo 20 puede codificar cada vista de diferentes maneras, por ejemplo, con diferentes velocidades de trama, diferentes velocidades de bits, diferentes resoluciones u otras de dichas diferencias. Por tanto, la unidad de encapsulación 21 puede formar diversas representaciones que tienen diversas características, por ejemplo, velocidad de bits, velocidad de tramas, resolución y similares.

[0047] Cada una de las representaciones puede corresponder a flujos de bits respectivos que se pueden recuperar por el dispositivo de destino 14. La unidad de encapsulación 21 puede proporcionar una indicación de una gama de identificadores de vista (view_ids) para vistas incluidas en cada representación, por ejemplo, dentro de una estructura de datos de descripción de presentación de medios (MPD) para el contenido de multimedia. Por ejemplo, la unidad de encapsulación 21 puede proporcionar una indicación de un identificador de vista máximo y un identificador de vista mínimo para las vistas de una representación. La MPD puede proporcionar además indicaciones de números máximos de vistas seleccionadas para su salida para cada una de una pluralidad de representaciones del contenido de multimedia. La MPD o los datos de la misma se pueden almacenar, en algunos ejemplos, en un manifiesto para la(s) representación/representaciones.

[0048] El medio legible por ordenador 16 puede incluir medios transitorios, tales como una radiodifusión inalámbrica o transmisión de red cableada, o medios de almacenamiento (es decir, medios de almacenamiento no transitorio), tales como un disco duro, una unidad de memoria flash, un disco compacto, un disco de vídeo digital, un disco Blu-ray u otros medios legibles por ordenador. En algunos ejemplos, un servidor de red (no mostrado) puede recibir datos de vídeo codificados desde el dispositivo de origen 12 y proporcionar los datos de vídeo codificados al dispositivo de destino 14, por ejemplo, por medio de transmisión por red. De forma similar, un dispositivo informático de una utilidad de producción de un medio, tal como una utilidad de grabación de discos, puede recibir datos de vídeo codificados desde el dispositivo de origen 12 y producir un disco que contiene los datos de vídeo codificados. Por lo tanto, se puede entender que el medio legible por ordenador 16 incluye uno o más medios legibles por ordenador de diversas formas, en diversos ejemplos.

[0049] La interfaz de entrada 28 del dispositivo de destino 14 recibe información desde el medio legible por ordenador 16. La información del medio legible por ordenador 16 puede incluir información sintáctica definida por el codificador de vídeo 20, que también usa el decodificador de vídeo 30, que incluye elementos sintácticos que describen características y/o procesamiento de bloques y otras unidades codificadas, por ejemplo, GOP. La unidad de desencapsulación 29 del dispositivo de destino 14 puede representar una unidad que desencapsula mensajes de SEI

de un flujo de bits (o un subconjunto de un flujo de bits denominado un "punto de funcionamiento" en el contexto de MVC). La unidad de desencapsulación 29 puede llevar a cabo operaciones en un orden inverso a las llevadas a cabo por la unidad de encapsulación 29 para desencapsular datos del flujo de bits codificado y encapsulado, tales como mensajes de SEI. El dispositivo de visualización 32 muestra los datos de vídeo descodificados a un usuario, y puede comprender cualquiera de una variedad de dispositivos de visualización, tales como un tubo de rayos catódicos (CRT), una pantalla de cristal líquido (LCD), una pantalla de plasma, una pantalla de diodos orgánicos emisores de luz (OLED) u otro tipo de dispositivo de visualización.

[0050] El codificador de vídeo 20 y el descodificador de vídeo 30 se pueden implementar, cada uno, como cualquiera de una variedad de circuitos de codificadores o descodificadores adecuados, según corresponda, tales como uno o más microprocesadores, procesadores de señales digitales (DSP), circuitos integrados específicos de la aplicación (ASIC), matrices de puertas programables (FPGA), circuitos de lógica discreta, software, hardware, firmware o cualquier combinación de los mismos. Tanto el codificador de vídeo 20 como el descodificador de vídeo 30 pueden estar incluidos en uno o más codificadores o descodificadores, cada uno de los cuales puede estar integrado como parte de un codificador/descodificador (CÓDEC) de vídeo combinado. Un dispositivo que incluye el codificador de vídeo 20 y/o el descodificador de vídeo 30 puede comprender un circuito integrado, un microprocesador y/o un dispositivo de comunicación inalámbrica, tal como un teléfono celular.

[0051] Aunque no se muestra en la FIG. 1, en algunos aspectos, el codificador de vídeo 20 y el descodificador de vídeo 30 se pueden integrar, cada uno, en un codificador y descodificador de audio, y pueden incluir unidades MUX-DEMUX apropiadas, u otro hardware y software, para tratar la codificación tanto de audio como de vídeo en un flujo de datos común o en flujos de datos separados. Si procede, las unidades MUX-DEMUX se pueden ajustar al protocolo de multiplexador ITU H.223 o a otros protocolos tales como el protocolo de datagramas de usuario (UDP).

[0052] En el ejemplo mostrado en la FIG. 1, el sistema 10 también incluye la red de entrega de contenido/servidor 34 que tiene un enrutador 36. En algunos ejemplos, el dispositivo de origen 12 se puede comunicar con la red de entrega de contenido/servidor 34 por medio de una variedad de medios de transmisión o almacenamiento inalámbricos y/o cableados, como se describe anteriormente. Además, si bien se muestra por separado en el ejemplo de la FIG. 1, en algunos ejemplos, el dispositivo de origen 12 y la red de entrega de contenido/servidor 34 comprenden el mismo dispositivo. La red de entrega de contenido/servidor 34 puede almacenar una o más versiones de datos de vídeo codificados (desde el codificador de vídeo 20 del dispositivo de origen 12), y puede hacer que dichos datos de vídeo codificados estén disponibles para el acceso por el dispositivo de destino 14 y el descodificador de vídeo 30. En algunos ejemplos, el enrutador 36 puede ser responsable de proporcionar datos de vídeo codificados al dispositivo de destino 14 en un formato solicitado.

[0053] El codificador de vídeo 20 y el descodificador de vídeo 30 pueden funcionar de acuerdo con una norma de codificación de vídeo, tal como la norma de codificación de vídeo de alta eficacia (HEVC), actualmente en fase de desarrollo, y se pueden ajustar al modelo de prueba de la HEVC (HM). De forma alternativa, el codificador de vídeo 20 y el descodificador de vídeo 30 pueden funcionar de acuerdo con otras normas patentadas o de la industria, tales como la norma ITU-T H.264, denominada de forma alternativa MPEG-4, parte 10, codificación avanzada de vídeo (AVC), o extensiones de dichas normas, incluyendo la extensión MVC y/o la extensión 3DVC para la H.264. Sin embargo, las técnicas de la presente divulgación no están limitadas a ninguna norma de codificación particular. Otros ejemplos de normas de codificación de vídeo incluyen MPEG-2 e ITU-T H.263.

[0054] La norma ITU-T H.264/MPEG-4 (AVC) se formuló por el Grupo de Expertos en Codificación de Vídeo (VCEG) de ITU-T, conjuntamente con el Grupo de Expertos en Imágenes en Movimiento (MPEG) de ISO/IEC, como el producto de una asociación colectiva conocida como el equipo de vídeo conjunto (JVT). En algunos aspectos, las técnicas descritas en la presente divulgación se pueden aplicar a dispositivos que se ajustan en general a la norma H.264. La norma H.264 se describe en la recomendación H.264 de ITU-T, codificación avanzada de vídeo para servicios audiovisuales genéricos, por el grupo de estudio de ITU-T, y con fecha de marzo de 2005, que se puede denominar en el presente documento la norma H.264 o la especificación H.264, o la norma o especificación H.264/AVC. La H.264/AVC incluye extensiones de codificación de vídeo escalable (SVC) y extensiones MVC. Además, existe otro desarrollo para proporcionar una extensión 3DVC, haciendo uso de MVC con la inclusión de mapas de profundidad. El equipo de vídeo conjunto (JVT) continúa trabajando en extensiones para la H.264/MPEG-4 AVC. Si bien se han descrito dentro del contexto de 3DVC, las técnicas descritas en la presente divulgación se pueden implementar con respecto a otros algoritmos de codificación de vídeo, que pueden codificar y descodificar vídeo 3D que implica tanto componentes de textura como de profundidad.

[0055] El codificador de vídeo 20 puede enviar datos sintácticos, tales como datos sintácticos basados en bloques, datos sintácticos basados en imágenes y datos sintácticos basados en GOP, al descodificador de vídeo 30, por ejemplo, en una cabecera de imagen, una cabecera de bloque, una cabecera de fragmento o una cabecera de GOP. Los datos sintácticos del GOP pueden describir un número de imágenes en el GOP respectivo, y los datos sintácticos de imagen pueden indicar un modo de codificación/predicción usado para codificar la imagen correspondiente.

[0056] En algunos ejemplos, el codificador de vídeo 20 puede generar y el descodificador de vídeo 30 puede recibir determinados conjuntos de parámetros, que se pueden usar al descodificar datos de vídeo. Por ejemplo, los conjuntos

de parámetros pueden contener información de cabecera a nivel de secuencia (en conjuntos de parámetros de secuencia (SPS)) y la información de cabecera a nivel de imagen, que cambia con poca frecuencia (en conjuntos de parámetros de imagen (PPS)). Con los conjuntos de parámetros (por ejemplo, PPS y SPS), la información que cambia con poca frecuencia no tiene por qué repetirse para cada secuencia (por ejemplo, secuencia de imágenes) o imagen; de ahí que la eficacia de codificación se pueda mejorar. Además, el uso de conjuntos de parámetros puede permitir la transmisión fuera de banda de la información de cabecera importante, evitando la necesidad de transmisiones redundantes, para la capacidad de recuperación de errores. En los ejemplos de transmisión fuera de banda, las unidades NAL de conjuntos de parámetros se pueden transmitir por un canal diferente al de otras unidades NAL, tales como las unidades NAL de información de mejora complementaria (SEI).

[0057] Las unidades NAL de SEI (denominadas mensajes de SEI) pueden contener información que no es necesaria para descodificar las muestras de imágenes codificadas a partir de las unidades NAL de la capa de codificación de vídeo (VCL), pero pueden ayudar en los procesos relacionados con la descodificación, la visualización, la capacidad de recuperación de errores y otros propósitos. Los mensajes de SEI pueden estar contenidos en las unidades NAL distintas de VCL. Los mensajes de SEI se pueden incluir en la parte normativa de algunas especificaciones estándar y, por tanto, no siempre son obligatorios para la implementación del descodificador compatible con las normas. Los mensajes de SEI pueden ser mensajes de SEI a nivel de secuencia o mensajes de SEI a nivel de imagen. Parte de la información a nivel de secuencia puede estar contenida en los mensajes de SEI, tales como los mensajes de SEI de información de escalabilidad en el ejemplo de la SVC y los mensajes de SEI de información de escalabilidad de vistas en la MVC. Los mensajes de SEI de información de escalabilidad también pueden proporcionar información a nivel de secuencia para la extensión 3DVC para la H.264. La unidad de encapsulación 21 puede formar estos mensajes de SEI y proporcionar estos mensajes de SEI a la red de entrega de contenido/servidor 34 para su uso en la negociación, como un ejemplo, de la entrega de datos de vídeo codificados.

[0058] La unidad de encapsulación 30 puede formar unidades NAL que comprenden una cabecera que identifica un programa al que pertenece la NAL, así como una carga útil, por ejemplo, datos de audio, datos de vídeo o datos que describen el flujo de transporte o de programa al que corresponde la unidad NAL. Por ejemplo, en la H.264/AVC, una unidad NAL incluye una cabecera de 1 octeto y una carga útil de tamaño variable. En un ejemplo, una cabecera de unidad NAL comprende un elemento `priority_id`, un elemento `temporal_id`, un elemento `anchor_pic_flag`, un elemento `view_id`, un elemento `non_idr_flag` y un elemento `inter_view_flag`. En la MVC convencional, se conserva la unidad NAL definida por la H.264, excepto para las unidades NAL de prefijo y las unidades NAL de fragmento codificado por MVC, que incluyen una cabecera de unidad NAL de MVC de 4 octetos y la carga útil de la unidad NAL.

[0059] En algunos ejemplos, el codificador de vídeo 20 puede codificar un flujo de bits 3DVC que se ajusta a la extensión 3DVC de H.264, usando información de MVC más profundidad. El borrador conjunto de MVC más reciente se describe en "Advanced video coding for generic audiovisual services [Codificación avanzada de vídeo para servicios audiovisuales genéricos]", Recomendación H.264 de ITU-T, marzo de 2010, mientras que un borrador de trabajo de 3DVC se describe en "WD on MVC Extensions for Inclusion of Depth Maps [WD en Extensiones MVC para inclusión de mapas de profundidad]", ISO/IEC/JTC1/SC29/WG11/N12351, Ginebra, Suiza, con fecha de noviembre de 2011, descrito anteriormente. De nuevo, la referencia dentro de la presente divulgación a MVC se debe entender como que es una referencia a MVC más profundidad dentro del contexto de 3DVC. Es decir, en la presente divulgación, se hace referencia a MVC en relación con 3DVC en el sentido de que 3DVC incorpora o "se construye a partir" de MVC.

[0060] En la extensión 3DVC para la H.264, una vista incluye textura y profundidad. La parte de textura de una vista se denomina vista de textura y la parte de profundidad de una vista se denomina vista de profundidad. La parte de textura de una vista en una unidad de acceso, es decir, una vista de textura en una unidad de acceso, se denomina componente de vista de textura. La parte de profundidad de una vista en una unidad de acceso, es decir, una vista de profundidad en una unidad de acceso, se denomina componente de vista de profundidad. El término componente de vista es una vista en una unidad de acceso y se refiere conjuntamente a tanto el componente de vista de textura como al componente de vista de profundidad en la misma unidad de acceso.

[0061] En las extensiones de H.264/AVC, se pueden añadir elementos sintácticos en la extensión de cabecera de unidad NAL para extender la cabecera de unidad NAL de un octeto a cuatro octetos para describir las características de una unidad NAL de VCL en múltiples dimensiones. Por tanto, una unidad NAL de VCL en la extensión MVC puede incluir una cabecera de unidad NAL más larga que la cabecera de unidad NAL en la norma H.264/AVC. La extensión MVC para la H.264/AVC se puede denominar en la presente divulgación "MVC/AVC".

[0062] Una unidad NAL de MVC/AVC puede contener una cabecera de unidad NAL de un octeto que incluye el tipo de unidad NAL, así como una extensión de cabecera de unidad NAL de MVC/AVC. Como ejemplo, la extensión de cabecera de unidad NAL de MVC/AVC puede incluir los elementos sintácticos en la siguiente tabla:

SINTAXIS DE EXTENSIÓN DE CABECERA DE UNIDAD NAL

nal_unit_header_extension() {	C	Descriptor
reserved_zero_bit	Todos	u(1)
idr_flag	Todos	u(1)
priority_id	Todos	u(6)
view_id	Todos	u(10)
temporal_id	Todos	u(3)
anchor_pic_flag	Todos	u(1)
inter_view_flag	Todos	u(1)
reserved_one_bit	Todos	u(1)
}		

- 5 **[0063]** En la tabla anterior, el elemento `idr_flag` puede indicar si la unidad NAL pertenece a una actualización de descodificador instantánea (IDR) o a una imagen de IDR de vista (V-IDR) que se puede usar como un punto de acceso aleatorio GOP cerrado. Por ejemplo, una imagen de IDR y todas las imágenes que suceden a la imagen IDR tanto en orden de visualización como en orden de flujo de bits, se pueden descodificar apropiadamente sin descodificar imágenes previas, en orden de flujo de bits o bien en orden de visualización.
- 10 **[0064]** El elemento `priority_id` se puede usar con un proceso de adaptación de flujo de bits que varía el flujo de bits de acuerdo con las condiciones de red cambiantes y/o las capacidades del descodificador de vídeo 30 y/o el dispositivo de visualización 32 (por ejemplo, tal como un proceso de adaptación de paso único). El elemento `view_id` se puede usar para indicar el identificador de vista para la vista a la que pertenece la unidad NAL, que se puede usar dentro de un descodificador MVC, por ejemplo, para predicción entre vistas y fuera de un descodificador, por ejemplo, para renderización. En algunos casos, el `view_id` se puede establecer igual a un id de cámara predefinida, y puede ser relativamente grande. El elemento `temporal_id` se puede usar para indicar el nivel temporal de la unidad NAL actual, que puede corresponder a una velocidad de trama particular.
- 15 **[0065]** El elemento `anchor_pic_flag` se puede usar para indicar si la unidad NAL pertenece a una imagen de anclaje que se puede usar como punto de acceso aleatorio GOP abierto. Por ejemplo, las imágenes de anclaje y todas las imágenes que suceden a la imagen de anclaje en orden de visualización se pueden descodificar apropiadamente sin descodificar las imágenes previas en el orden de descodificación (es decir, el orden de flujo de bits), y por tanto se pueden usar como puntos de acceso aleatorio. Las imágenes de anclaje y las imágenes no de anclaje pueden tener diferentes dependencias de vista, ambas de las cuales se pueden señalar en un SPS.
- 20 **[0066]** Es decir, como se describe en el presente documento, una dependencia de vista se puede referir en general a una vista de la que depende una vista que se está codificando actualmente. En otras palabras, se pueden exponer dependencias de vista a partir de las que se puede predecir una vista que se está codificando actualmente. De acuerdo con algunos ejemplos, la dependencia de vista se puede señalar en la extensión SPS MVC. En dichos ejemplos, todas las predicciones entre vistas se pueden realizar dentro del alcance especificado por la extensión SPS MVC. El elemento `inter_view_flag` se puede usar para indicar si la unidad NAL se usa para la predicción entre vistas para las unidades NAL en otras vistas.
- 25 **[0067]** Para transmitir la información de cabecera de unidad NAL (que puede ser de cuatro octetos) para la vista base de un flujo de bits MVC, una unidad NAL de prefijo se puede definir en MVC. En el contexto de MVC, la unidad de acceso de vista base puede incluir las unidades NAL de VCL de una instancia temporal actual de una vista particular, así como una unidad NAL de prefijo para la unidad de acceso de vista base, que puede contener solo la cabecera de unidad NAL. Si no se requiere la unidad NAL de prefijo para descodificar (por ejemplo, tal como descodificar una vista única), un descodificador puede ignorar y/o descartar la unidad NAL de prefijo.
- 30 **[0068]** Con respecto a una extensión MVC/AVC de SPS, el SPS de MVC puede indicar vistas que se pueden usar para propósitos de predicción entre vistas. Por ejemplo, las potenciales referencias entre vistas se pueden señalar en la extensión MVC/AVC de SPS y se pueden modificar por el proceso de construcción de la lista de imágenes de referencia, que permite el ordenamiento flexible de las referencias de predicción inter o de predicción entre vistas. Un ejemplo de SPS de MVC/AVC se expone en la siguiente tabla:
- 35
- 40
- 45

SPS de MVC DE EJEMPLO

seq_parameter_set_mvc_extension() {	C	Descriptor
num_views_minus1	0	uc(v)
for(i = 0; i <= num_views_minus1; i++)		
view_id[i]	0	uc(v)
for(i = 1; i <= num_views_minus1; i++) {		
num_anchor_refs_10[i]	0	uc(v)
for(j = 0; j < num_anchor_refs_10[i]; j++)		
anchor_ref_10[i][j]	0	uc(v)
num_anchor_refs_11[i]	0	uc(v)
for(j = 0; j < num_anchor_refs_11[i]; j++)		
anchor_ref_11[i][j]	0	uc(v)
}		
for(i = 1; i <= num_views_minus1; i++) {		
num_non_anchor_refs_10[i]	0	uc(v)

for(j = 0; j < num_non_anchor_refs_10[i]; j++)		
non_anchor_ref_10[i][j]	0	uc(v)
num_non_anchor_refs_11[i]	0	uc(v)
for(j = 0; j < num_non_anchor_refs_11[i]; j++)		
non_anchor_ref_11[i][j]	0	uc(v)
}		
num_level_values_signalled_minus1	0	uc(v)
for(i = 0; i <= num_level_values_signalled_minus1; i++) {		
level_idc[i]	0	u(8)
num_applicable_ops_minus1[i]	0	uc(v)
for(j = 0; j <= num_applicable_ops_minus1[i]; j++) {		
applicable_op_temporal_id[i][j]	0	u(3)
applicable_op_num_target_views_minus1[i][j]	0	uc(v)
for(k = 0; k <=		
applicable_op_num_target_views_minus1[i][j]; k++)		
applicable_op_target_view_id[i][j][k]	0	uc(v)
applicable_op_num_views_minus1[i][j]	0	uc(v)
}		
}		
}		

5

[0069] De acuerdo con algunos ejemplos, la dependencia de vista se puede señalar en la extensión SPS MVC. Todas las predicciones entre vistas se pueden realizar dentro del alcance especificado por la extensión SPS MVC. Es decir, el SPS puede exponer a qué vistas se puede hacer referencia para propósitos de predicción por una vista que se está codificando actualmente. En la tabla 2 anterior, el elemento num_anchor_refs_10[i] puede especificar el número de componentes de vista para la predicción entre vistas en la lista de imágenes de referencia inicializada para la Lista 0 (por ejemplo, RefPicList0).

10

[0070] Además, el elemento anchor_ref_10[i][j] puede especificar el view_id del j-ésimo componente de vista para la predicción entre vistas en la RefPicList0 inicializada. El elemento num_anchor_refs_11[i] puede especificar el número de componentes de vista para la predicción entre vistas en la lista de imágenes de referencia inicializada para la lista uno (por ejemplo, RefPicList1). El elemento anchor_ref_11[i][j] puede especificar el view_id del j-ésimo componente de vista para la predicción entre vistas en la RefPicList1 inicializada.

15

[0071] El elemento num_non_anchor_refs_10[i] puede especificar el número de componentes de vista para la predicción entre vistas en la RefPicList0 inicializada. El elemento non_anchor_ref_10[i][j] puede especificar el view_id del j-ésimo componente de vista para la predicción entre vistas en la RefPicList0 inicializada. El elemento num_non_anchor_refs_11[i] puede especificar el número de componentes de vista para la predicción entre vistas en la RefPicList1 inicializada. El elemento non_anchor_ref_11[i][j] puede especificar el view_id del j-ésimo componente de vista para la predicción entre vistas en la RefPicList inicializada.

20

25

[0072] La lista de imágenes de referencia inicializada, o "inicial", puede ser la misma o diferente de una lista de imágenes de referencia final usada para propósitos de predicción entre vistas de componentes de vista. Es decir, determinados candidatos de referencia (es decir, las imágenes de referencia que se pueden usar para la predicción entre vistas) se pueden eliminar de una lista de imágenes de referencia inicial (por ejemplo, imágenes redundantes).
 5 Además, como se describe con mayor detalle a continuación, los candidatos de referencia se pueden reordenar a partir de la lista de imágenes de referencia inicial para formar la lista de imágenes de referencia final.

[0073] En este ejemplo, de acuerdo con MVC, las dependencias de vista para imágenes de anclaje e imágenes no de anclaje se mantienen y se señalan por separado. Es decir, un codificador de vídeo puede determinar un total de cuatro listas de imágenes de referencia (por ejemplo, Lista 0, imágenes no de anclaje; Lista 1, imágenes no de anclaje; Lista 0, imágenes de anclaje; Lista 1, imágenes de anclaje). Además, como se muestra en la Tabla 2 anterior, se requiere una señalización separada para indicar una dependencia de vista al descodificador de vídeo 30. Es decir, el SPS debe incluir señalización separada de Lista 0 y Lista 1 tanto para anchor_refs como para non_anchor_refs.

[0074] Además, de acuerdo con la tabla 2, la dependencia entre vistas para componentes de vista de no anclaje es un subconjunto de la de componentes de vista de anclaje. Es decir, por ejemplo, un componente de vista de una vista de anclaje se puede predecir desde más de otra vista, tal como la vista 3 y 4. Sin embargo, una vista no de anclaje solo se puede predecir a partir de las imágenes de vista 3 (un subconjunto de la vista de anclaje). De esta manera, las dependencias de vista para los componentes de vista de anclaje y no de anclaje se pueden mantener por separado.

[0075] Además, en la Tabla 2, el num_level_values_signalled puede especificar el número de valores de nivel señalados para la secuencia de vídeo codificada. El elemento level_idc[i] puede especificar el i-ésimo valor de nivel señalado para la secuencia de vídeo codificada. El elemento num_applicable_ops_minus1[i] más 1 puede especificar el número de puntos de funcionamiento a los que se aplica el nivel indicado por level_idc1[i]. El elemento applicable_op_temporal_id[i][j] puede especificar el temporal_id del j-ésimo punto de funcionamiento al que se aplica el nivel indicado por level_idc[i].

[0076] El elemento applicable_op_num_target_views_minus1[i][j] puede especificar el número de vistas de salida de destino para el j-ésimo punto de funcionamiento al que se aplica el nivel indicado por level_idc[i]. El elemento applicable_op_target_view_id[i][j][k] puede especificar la k-ésima vista de salida de destino para el j-ésimo punto de funcionamiento al que se aplica el nivel indicado por level_idc[i]. El elemento applicable_op_num_views_minus1[i][j] puede especificar el número de vistas, incluyendo las vistas que dependen de las vistas de salida de destino pero que no pertenecen a las vistas de salida de destino, en el j-ésimo punto de funcionamiento al que se aplica el nivel indicado por level_idc[i].

[0077] En consecuencia, en el SPS para MVC, se puede señalar el número de vistas que se pueden usar para formar la lista 0 de imágenes de referencia y la lista 1 de imágenes de referencia para cada vista. Además, la relación de predicción para una imagen de anclaje, como se señala en el SPS para MVC, puede ser diferente de la relación de predicción para la imagen no de anclaje (señalada en el SPS de MVC) de la misma vista.

[0078] Como se describe en mayor detalle a continuación, el codificador de vídeo 20 y el descodificador de vídeo 30 pueden disponer de forma flexible referencias de predicción temporal y de vista al construir listas de imágenes de referencia. Permitir la disposición flexible proporciona no solo potencial ganancia de eficacia de codificación sino también capacidad de recuperación de errores, porque la sección de imágenes de referencia y los mecanismos de imagen redundantes se pueden extender a la dimensión de vista. El codificador de vídeo 20 y/o el descodificador de vídeo 30 pueden, en un ejemplo, construir una lista de imágenes de referencia de acuerdo con las siguientes etapas:

- 1) Inicializar la lista de imágenes de referencia para imágenes de referencia temporales (es decir, dentro de vistas), de modo que no se consideren imágenes de referencia de otras vistas.
- 2) Adjuntar las imágenes de referencia entre vistas al final de la lista en el orden en que aparecen las imágenes en la extensión SPS de MVC.
- 3) Aplicar un proceso de reordenamiento de la lista de imágenes de referencia (RPLR) para imágenes de referencia tanto dentro de vistas como entre vistas. Las imágenes de referencia entre vistas se pueden identificar en los comandos de RPLR por sus valores de índice, como se especifica en la extensión MVC SPS.

[0079] En MVC, el codificador de vídeo 20 y/o descodificador de vídeo 30, al descodificar datos de vídeo de múltiples vistas, pueden activar cero o un conjunto de parámetros de imagen (PPS) como un PPS activo y cero o más PPS de vista (VPPS) como el VPPS activo, donde cada VPPS activo está activo específicamente para un valor de índice de orden de vista particular menor que un valor de índice de orden de vista máximo. Es decir, el codificador de vídeo 20 y/o descodificador de vídeo 30 pueden activar un PPS que se aplica a cada componente de vista de los datos de vídeo de múltiples vistas dentro de una unidad de acceso. El codificador de vídeo 20 y/o descodificador de vídeo 30 también pueden activar un VPPS que se aplica solo a un subconjunto de los componentes de vista en la unidad de acceso. En algunos casos, el VPPS puede incluir uno o más elementos sintácticos que anulan uno o más elementos sintácticos

especificados en el PPS para aquellas vistas para las que se activó el VPPS. En algunos casos, el VPPS puede incluir uno o más elementos sintácticos que aumentan uno o más elementos sintácticos especificados en el PPS para aquellas vistas para las que se activó el VPPS.

5 **[0080]** Además, en MVC, el codificador de vídeo 20 y/o descodificador de vídeo 30 pueden, al descodificar datos de vídeo de múltiples vistas, activar cero o un conjunto de parámetros de secuencia (SPS) de MVC como el MVC SPS activo y cero o más SPS de vista (VSPS) como el VSPS activo, donde cada VSPS activo está activo específicamente para un valor de índice de orden de vista particular menor que un valor de índice de orden de vista máximo. Es decir, el codificador de vídeo 20 y/o descodificador de vídeo 30 pueden activar un SPS que se aplica a cada componente de vista de los datos de vídeo de múltiples vistas dentro de una o más unidades de acceso. El codificador de vídeo 20 y/o descodificador de vídeo 30 también pueden activar un VSPS que se aplica solo a un subconjunto de los componentes de vista en la una o más unidades de acceso. En algunos casos, el VSPS puede incluir uno o más elementos sintácticos que anulan uno o más elementos sintácticos especificados en el SPS para aquellas vistas para las que se activó el VSPS. En algunos casos, el VSPS puede incluir uno o más elementos sintácticos que aumentan uno o más elementos sintácticos especificados en el SPS para aquellas vistas para las cuales se activó el VSPS.

20 **[0081]** Adicionalmente, MVC proporciona funcionalidad para analizar o extraer las diversas llamadas partes de operación de un único flujo de bits. Cada punto de funcionamiento puede representar una combinación diferente de vistas de los datos de vídeo de múltiples vistas codificados a velocidades de tramas temporales y resoluciones espaciales variables. En otras palabras, un punto de funcionamiento se puede referir a una codificación de datos de vídeo de múltiples vistas en tres dimensiones, incluyendo la dimensión de vista (que refleja el número de vistas), la dimensión temporal (que refleja una velocidad de tramas) y la dimensión espacial (que refleja la resolución espacial).

25 **[0082]** Un punto de funcionamiento se puede identificar por un valor de temporal_id que representa el nivel temporal de destino y un conjunto de valores de view_id que representan las vistas de salida de destino. Un punto de funcionamiento está asociado con un subconjunto del flujo de bits, que consiste en las vistas de salida de destino y todas las demás vistas de las que dependen las vistas de salida de destino, que se deriva usando el proceso de extracción del subflujo de bits como se especifica en la subcláusula H.8.5.3 de la extensión de codificación de vídeo tridimensional (3DVC) de H.264/AVC con tldTarget igual al valor de temporal_id y viewIdTargetList que consiste en el conjunto de valores de view_id como entradas. Se puede asociar más de un punto de funcionamiento con el mismo subconjunto de flujo de bits. Cuando 3DVC declara que "se descodifica un punto de funcionamiento", se refiere a la descodificación de un subconjunto de flujo de bits correspondiente al punto de funcionamiento y la salida posterior de las vistas de salida de destino. En otras palabras, cada punto de funcionamiento se representa con un número de vistas de destino para la salida en un determinado nivel temporal y/o resolución o nivel espacial.

35 **[0083]** En la extensión MVC del conjunto de parámetros de secuencia, se pueden señalar diferentes puntos de funcionamiento con valores de level_idc. Los siguientes elementos sintácticos se denominan elementos sintácticos de definiciones de nivel.

num_level_values_signalled_minus1	0	ue(v)
for(i = 0; i <= num_level_values_signalled_minus1; i++) {		
level_idc[i]	0	u(8)
num_applicable_ops_minus1[i]	0	ue(v)
for(j = 0; j <= num_applicable_ops_minus1[i]; j++) {		
applicable_op_temporal_id[i][j]	0	u(3)
applicable_op_num_target_views_minus1[i][j]	0	ue(v)
for(k = 0; k <= applicable_op_num_target_views_minus1[i][j]; k++)		
applicable_op_target_view_id[i][j][k]	0	ue(v)
applicable_op_num_views_minus1[i][j]	0	ue(v)
}		

40 **[0084]** Para cada level_idc señalado en la extensión MVC SPS, se señala un nuevo level_idc_depth en la extensión SPS de vídeo 3D (3DV). Entonces, si se supone que todos los puntos de funcionamiento MVC tienen el mismo level_idc[i], después de incluir los componentes de profundidad y construir los puntos de funcionamiento 3DVC, los puntos de funcionamiento 3DVC tendrán el mismo nivel de level_idc_depth[i]. Un diseño sintáctico detallado se muestra como sigue.

seq_parameter_set_3dvc_extension() {	C	Descriptor
...		
for(i = 0; i <= num_level_values_signalled_minus1; i++)		
level_idc_depth[i]	0	u(8)
...		
}		

5 [0085] Además, MVC establece que un componente de vista se puede eliminar de la memoria intermedia de imágenes descodificadas (DPB) si el componente de vista no se usa más para la referencia entre vistas y el componente de vista no se usa para la salida, al descodificar un componente de vista en la misma unidad de acceso que el componente de vista que se va a eliminar.

10 [0086] En la actualidad, 3DVC no proporciona ninguna manera por la que distinguir entre los componentes de textura y profundidad de un componente de vista, donde el término "componente de vista" se refiere a una imagen de una vista especificada por los datos de vídeo de múltiples vistas en 3D. Como resultado, al codificar los componentes de textura y profundidad de una vista, el codificador de vídeo típicamente trata o de otro modo procesa tanto los componentes de textura como de profundidad como componentes de vista de forma similar a la descrita anteriormente con respecto a MVC sin proporcionar ninguna manera por la que distinguir entre componentes de textura y profundidad. Es decir, 3DVC proporciona actualmente el tratamiento o codificación de los componentes de vista sin proporcionar una manera por la que procesar individualmente los componentes de textura por separado de los componentes de profundidad. Esta falta de distinción en 3DVC puede dar como resultado una menor eficacia de codificación y/o una menor calidad de los datos de vídeo reconstruidos.

20 [0087] Por ejemplo, se puede requerir actualmente que se especifique un componente de profundidad a la misma resolución que un componente de textura correspondiente con el fin de dar cabida al tratamiento conjunto de este componente de vista. Sin embargo, un componente de profundidad con resolución más alta (en comparación con la resolución del componente de textura) puede dar como resultado una reproducción de vídeo tridimensional (3D) en la que se puede obtener un vídeo 3D mejor y más envolvente que imita mejor lo que espera el sistema visual de un espectador. Además, un componente de profundidad con resolución más baja (en comparación con la resolución del componente de textura) puede proporcionar la misma o similar experiencia 3D envolvente en determinados casos, pero consume menos bits cuando se codifica y, de este modo, incrementa la eficacia de la codificación. No permitiendo el tratamiento separado de los componentes de profundidad y textura, 3DVC puede reducir la eficacia de la codificación y/o proporcionar una menor calidad de los datos de vídeo reconstruidos (a menudo, en términos, de la experiencia visual).

30 [0088] Las técnicas de la presente divulgación se refieren en general a permitir el tratamiento separado o independiente de los componentes de textura y los componentes de profundidad de una vista al procesar o codificar los datos de vídeo 3D. Por ejemplo, se puede añadir una extensión 3DVC de conjunto de parámetros de secuencia (SPS) para los perfiles relacionados con 3DVC. Dentro de esta extensión SPS 3DVC, se puede señalar uno o más de los siguientes: el tamaño de imagen de las secuencias del mapa de profundidad, el tamaño de la imagen que se va a mostrar por las secuencias del mapa de profundidad y los niveles de los puntos de funcionamiento con las imágenes del mapa de profundidad tomado en consideración, donde cada uno de los niveles específicos corresponde a un nivel que ya se ha definido en la extensión SPS MVC. El tamaño de la imagen que se va a mostrar puede ser una imagen descodificada emitida por un dispositivo de visualización.

40 [0089] En un códec 3DVC, un componente de vista de cada vista de datos de vídeo en un instante de tiempo específico puede incluir un componente de vista de textura y un componente de vista de profundidad. El componente de vista de textura puede incluir componentes de luminancia (Y) y crominancia (Cb y Cr). Los componentes de luminancia (brillo) y crominancia (color) se denominan conjuntamente en el presente documento componentes de "textura". El componente de vista de profundidad puede ser de un mapa de profundidad de una imagen. En la renderización de imágenes 3D, los mapas de profundidad incluyen componentes de profundidad y se pueden usar para generar vistas virtuales desde una perspectiva visual proporcionada. Los elementos sintácticos para componentes de profundidad y componentes de textura se pueden señalar con la unidad de bloque codificado. Las unidades de bloques codificados, también denominados simplemente "bloques codificados" en la presente divulgación, pueden corresponder a macrobloques en ITU-T H.264/AVC (Codificación Avanzada de Vídeo).

55 [0090] La presente divulgación describe técnicas de señalización que se pueden aplicar por un codificador y usar por un descodificador durante la fase de predicción inter de un proceso de codificación y/o descodificación de vídeo. Las técnicas descritas se refieren a la codificación de contenido de vídeo tridimensional ("3D"). El contenido de vídeo 3D se puede representar, por ejemplo, como bloques codificados de vídeo de múltiples vistas más profundidad ("MVD"). Más específicamente, las técnicas implican la recepción de al menos una imagen bidimensional que tiene

componentes de vista de textura y componentes de vista de profundidad. Algunos componentes de vista de textura e información de profundidad se pueden codificar conjuntamente en un único bloque codificado o como bloques separados. Una imagen puede dividirse en fragmentos de imágenes. Los elementos sintácticos para codificar componentes de vista de textura y componentes de vista de profundidad se pueden señalar en un conjunto de parámetros de secuencia. Los componentes de vista de profundidad se pueden o no se pueden predecir a partir de los componentes de vista de textura.

[0091] En el contexto de las normas de codificación de vídeo, un "perfil" corresponde a un subconjunto de algoritmos, características o herramientas y restricciones que se les aplican. Como se define por la norma H.264, por ejemplo, un "perfil" es un subconjunto de toda la sintaxis del flujo de bits que está especificada por la norma H.264. Un "nivel" corresponde a las limitaciones del consumo de recursos del descodificador, tales como, por ejemplo, la memoria y el cálculo del descodificador, que se refieren a la resolución de las imágenes, la velocidad de bits y la velocidad de procesamiento de los macrobloques (MB). Un perfil se puede señalar con un valor de `profile_idc` (indicador de perfil), mientras que un nivel se puede señalar con un valor de `level_idc` (indicador de nivel).

[0092] En la norma H.264/AVC, las unidades de capa de abstracción de red (NAL) se definen para proporcionar una representación de vídeo "apta para la red" para hacer frente a aplicaciones tales como videotelefonía, almacenamiento o transmisión continua de vídeo. Las unidades NAL se pueden clasificar en unidades NAL de capa de codificación de vídeo (VCL) y unidades NAL no de VCL. Las unidades VCL pueden contener un motor de compresión de núcleo y comprender niveles de bloque, macrobloque (MB) y fragmento. Otras unidades NAL son unidades NAL no de VCL.

[0093] Para codificar vídeo 2D, cada unidad NAL puede contener una cabecera de unidad NAL de un octeto y una carga útil de tamaño variable. Se pueden usar cinco bits para especificar el tipo de unidad NAL. Se pueden usar tres bits para `nal_ref_idc`, que indica la importancia de la unidad NAL en términos de cómo otras imágenes (unidades NAL) se refieren a ella. Por ejemplo, establecer `nal_ref_idc` igual a 0 significa que la unidad NAL no se usa para la predicción inter. Como la norma H.264/AVC se puede ampliar para incluir la codificación de vídeo 3D, tal como la norma de codificación de vídeo escalable (SVC), la cabecera NAL puede ser similar a la del entorno 2D. Por ejemplo, uno o más bits en la cabecera de unidad NAL se pueden usar para indicar que la unidad NAL es una unidad NAL de cuatro componentes.

[0094] Las cabeceras de unidad NAL también se pueden usar para unidades MVC NAL. Sin embargo, en la MVC, la estructura de cabecera de unidad NAL se puede conservar excepto para las unidades NAL de prefijo y las unidades NAL de fragmento codificado por MVC. Las técnicas de la presente divulgación proporcionan un tipo de unidad NAL diferente que se puede usar para señalar un súper SPS.

[0095] La siguiente tabla proporciona un ejemplo de una sintaxis RBSP de un conjunto de parámetros de secuencia, de acuerdo con un aspecto de la presente divulgación. Solo para propósitos ilustrativos, para los ejemplos asociados con esta tabla, se puede suponer que la textura y la profundidad tienen el mismo valor de `view_id` y se usa el `nal_unit_type` para distinguir la textura y la profundidad. A continuación, los elementos en negrita representan elementos que se han añadido, actualizado o modificado de acuerdo con los ejemplos de la presente divulgación, en comparación con las versiones actuales o previas del borrador de trabajo para la extensión 3DVC para la MVC. Además, la referencia a los anexos, tal como el anexo G en la tabla a continuación, se refiere a los anexos de la norma de codificación de vídeo H.264.

	C	Descriptor
subset_seq_parameter_set_rbsp() {		
seq_parameter_set_data()	0	
if(profile_idc == 83 profile_idc == 86) {		
seq_parameter_set_svc_extension() /*especificado en el Anexo G */	0	
svc_vui_parameters_present_flag	0	u(1)
if(svc_vui_parameters_present_flag == 1)		
svc_vui_parameters_extension() /*especificado en el Anexo G */	0	
} else if(profile_idc == 118 profile_idc == 128) {		
bit_equal_to_one /* igual a 1 */	0	f(1)
seq_parameter_set_mvc_extension() /*especificado en el Anexo H */	0	
mvc_vui_parameters_present_flag	0	u(1)
if(mvc_vui_parameters_present_flag == 1)		
mvc_vui_parameters_extension() /* especificado en el Anexo H */	0	
}		
else if(profile_idc == 138 profile_idc == 148) {		
//138: perfil base 3DV ; 148: perfil asimétrico 3DV		
bit_equal_to_one /* igual a 1 */	0	f(1)
seq_parameter_set_mvc_extension()		
seq_parameter_set_3dvc_extension()		
3dvc_vui_parameter_present_flag	0	u(1)
if (3dvc_vui_parameter_present_flag) {		
3dvc_vui_parameters_extension() /*especificado en el Anexo I */		
}		
}		
additional_extension3_flag	0	u(1)
if (additional_extension3_flag)		
while(more_rbsp_data())		
additional_extension3_data_flag	0	u(1)
}		
rbsp_trailing_bits()	0	
}		

De acuerdo con la tabla mostrada anteriormente, **3dvc_vui_parameter_present_flag** igual a 1 indica que está presente 3dvc_vui_parameters_extension(). Cuando **3dvc_vui_parameter_present_flag** es igual a 0, el indicador indica que no está presente 3dvc_vui_parameters_extension(). Por ejemplo, **3dvc_vui_parameter_present_flag** indica que una extensión que incorpora algunas de las técnicas de la presente divulgación está presente en un conjunto de parámetros de secuencia.

[0096] En algunos ejemplos, un SPS de subconjunto con perfiles MVC (por ejemplo, altos en estéreo o altos de múltiples vistas) puede tener el mismo id de SPS que el del SPS de subconjunto con perfiles 3DV, de manera que los componentes de vista de textura se referirán a un mismo SPS_id que tiene diferentes contenidos en diferentes perfiles. Por ejemplo, un SPS de subconjunto con MVC puede corresponder a un punto de funcionamiento que solo tiene textura. Tener el mismo id de SPS que el de un SPS de subconjunto con perfiles 3DV puede permitir que un punto de funcionamiento que solo incluía la textura se extraiga por un descodificador o dispositivo de red, tal como un elemento de red con reconocimiento de medios (MANE).

[0097] El Perfil Alto de Múltiples vistas admite un número arbitrario de vistas. El Perfil Alto en Estéreo admite vídeo estereoscópico de dos vistas. En otros ejemplos, las técnicas descritas en el presente documento se pueden aplicar a ejemplos que usan otros perfiles MVC.

5 [0098] La siguiente tabla proporciona un ejemplo de elementos sintácticos de extensión para la extensión 3DVC del conjunto de parámetros de secuencia.

seq_parameter_set_3dvc_extension() {	C	Descriptor
disable_depth_inter_view_flag	0	u(1)
depth_pic_width_in_mbs_minus1	0	ue(v)
depth_pic_height_in_map_units_minus1	0	ue(v)
depth_frame_cropping_flag	0	u(1)
if(depth_frame_cropping_flag) {		
depth_frame_crop_left_offset	0	ue(v)
depth_frame_crop_right_offset	0	ue(v)
depth_frame_crop_top_offset	0	ue(v)
depth_frame_crop_bottom_offset	0	ue(v)
}		
for(i = 0; i <= num_level_values_signalled_minus1; i++)		
level_idc_depth[i]	0	u(8)
}		

10 [0099] En un ejemplo, la extensión 3DVC del conjunto de parámetros de secuencia especifica las relaciones de dependencia entre vistas de las vistas adicionales con una resolución más alta que las otras vistas en la secuencia de vídeo codificada. Los componentes de vista de textura de las otras vistas, con una resolución más baja, pertenecen a un subflujo de bits que se ajusta al Anexo H. La extensión 3DVC del conjunto de parámetros de secuencia también especifica la dependencia de los componentes de vista de profundidad. Las vistas con resolución más alta pueden tener anchura doble, altura idéntica o altura doble, o anchura idéntica a las otras vistas. En algunos ejemplos, las vistas con resolución más alta pueden tener solo anchura doble, altura idéntica o altura doble, o anchura idéntica a las otras vistas.

20 [0100] Se infiere que un `chroma_format_idc` para componentes de vista de profundidad es igual a 4:0:0.

[0101] El elemento sintáctico **disable_depth_inter_view_flag** puede tener la misma semántica que en el indicador con el mismo nombre en m22583. El elemento sintáctico **depth_pic_width_in_mbs_minus1** más 1 especifica la anchura de cada imagen del componente de vista de profundidad en unidades de macrobloques. Una variable para la anchura de imagen de los componentes de vista de profundidad en unidades de macrobloques se puede derivar como:

$$\text{PicWidthInMbs} = \text{pic_width_in_mbs_minus1} + 1 \quad (1)$$

30 [0102] Una variable para la anchura de imagen de los componentes de vista de profundidad para un componente luma se deriva como:

$$\text{PicWidthInSamplesL} = \text{PicWidthInMbs} * 16 \quad (2)$$

35 [0103] En algunos ejemplos, el elemento sintáctico **depth_pic_height_in_map_units_minus1** más 1 especifica la altura en unidades de mapa de grupo de fragmento de una trama descodificada de los componentes de vista de profundidad. Las variables `PicHeightInMapUnits` y `PicSizeInMapUnits` se derivan como:

$$\text{PicHeightInMapUnits} = \text{pic_height_in_map_units_minus1} + 1 \quad (3)$$

$$\text{PicSizeInMapUnits} = \text{PicWidthInMbs} * \text{PicHeightInMapUnits} \quad (4)$$

[0104] En algunos ejemplos, cuando el elemento sintáctico **depth_frame_cropping_flag** es igual a 1, **depth_frame_cropping_flag** especifica que los parámetros de desplazamiento de recorte de trama siguen a continuación en el conjunto de parámetros de secuencia para los componentes de vista de profundidad. **depth_frame_cropping_flag** igual a 0 especifica que los parámetros de desplazamiento de recorte de trama no están presentes para los componentes de vista de profundidad.

[0105] Los elementos sintácticos **depth_frame_crop_left_offset**, **depth_frame_crop_right_offset**, **depth_frame_crop_top_offset**, y **depth_frame_crop_bottom_offset** pueden especificar las muestras de las imágenes de los componentes de vista de profundidad en la secuencia de vídeo codificada que se emiten del proceso de descodificación, en términos de una región rectangular especificada en las coordenadas de trama para su salida. Estos desplazamientos pueden definir una ventana de recorte para las imágenes de los componentes de vista de profundidad descodificada.

[0106] En algunos ejemplos, el elemento sintáctico **level_idc_depth[i]** especifica el valor de nivel i-ésimo señalado para la secuencia de vídeo codificada con los componentes de vista de profundidad.

[0107] Es posible extraer una o más vistas a partir de un flujo de bits de MVC para producir un subflujo de bits. Un subflujo de bits con un conjunto razonable dado de vistas de salida de destino se denomina un punto de funcionamiento. Debido a las dependencias de vista, el conjunto de vistas de salida de destino puede ser un subconjunto de las vistas contenidas en un subflujo de bits. Las vistas de salida de destino también pueden ser idénticas a las vistas contenidas en el subflujo de bits. Sin embargo, si un proveedor de servicios no tiene la intención de admitir un punto de funcionamiento con un gran número de vistas, el número de vistas de salida de destino puede ser un subconjunto de las vistas contenidas. Téngase en cuenta que las vistas contenidas son las vistas que se van a descodificar.

[0108] Un punto de funcionamiento se identifica por un valor de **temporal_id** que representa un nivel temporal de destino y un conjunto de valores de **view_id** que representan las vistas de salida de destino. Un punto de funcionamiento está asociado con un subconjunto de flujo de bits, que consiste en las vistas de salida de destino y todas las demás vistas de las que dependen las vistas de salida de destino, que se deriva usando el proceso de extracción del subflujo de bits como se especifica en la subcláusula H.8.5.3 con **tIdTarget** igual al valor de **temporal_id** y **viewIdTargetList** que consiste en el conjunto de valores de **view_id** como entradas. Se puede asociar más de un punto de funcionamiento con el mismo subconjunto de flujo de bits. Cuando el 3DVC declara que "se descodifica un punto de funcionamiento", se refiere a la descodificación de un subconjunto de flujo de bits correspondiente al punto de funcionamiento y la salida posterior de las vistas de salida de destino.

[0109] De acuerdo con las técnicas descritas en el presente documento, la construcción de los puntos de funcionamiento para la extensión 3DVC SPS pueden ser los mismos que los de la extensión MVC SPS, excepto que los componentes de vista de profundidad están asociados. En otras palabras, para cada punto de funcionamiento señalado en la extensión MVC SPS que tiene un nivel igual a **level_idc[i]**, se permite un punto de funcionamiento correspondiente en la extensión 3DVC SPS, con la misma recopilación de vistas de salida de destino y valor de **temporal_id**, pero incluyen además, para cada componente de vista que incluye un componente de profundidad, un nivel de **level_idc_depth[i]**.

[0110] De forma alternativa, los valores de **level_idc** señalados en la extensión SPS MVC del SPS actual con un perfil 3DV, son siempre para puntos de funcionamiento que contienen unidades NAL de profundidad, por tanto el elemento sintáctico **level_idc_depth[i]** no se señala en algunos casos para 3DVC. Por ejemplo, al señalar los valores de **level_idc** en el SPS con un perfil 3DV que incluye la extensión SPS MVC, de acuerdo con las técnicas descritas en el presente documento, el elemento sintáctico **level_idc_depth[i]** no se señala porque los valores de **level_idc** son para puntos de funcionamiento que contienen unidades NAL de profundidad.

[0111] La siguiente tabla proporciona la semántica de ejemplo para una extensión de los parámetros de Información de Usabilidad de Vídeo (VUI) de 3DVC.

3dvc_vui_parameters_extension() {	C	Descriptor
vui_mvc_num_ops_minus1	0	ue(v)
for(i = 0; i <= vui_mvc_num_ops_minus1; i++) {		
vui_mvc_temporal_id[i]	0	u(3)
vui_mvc_num_target_output_views_minus1[i]	5	ue(v)
depth_present_flag[i]	0	u(1)
for(j = 0; j <= vui_mvc_num_target_output_views_minus1[i]; j++)		
vui_mvc_view_id[i][j]	5	ue(v)
vui_mvc_timing_info_present_flag[i]	0	u(1)
if(vui_mvc_timing_info_present_flag[i]) {		
vui_mvc_num_units_in_tick[i]	0	u(32)
vui_mvc_time_scale[i]	0	u(32)
vui_mvc_fixed_frame_rate_flag[i]	0	u(1)
}		
vui_mvc_nal_hrd_parameters_present_flag[i]	0	u(1)
if(vui_mvc_nal_hrd_parameters_present_flag[i])		
hrd_parameters()	0	
vui_mvc_vcl_hrd_parameters_present_flag[i]	0	u(1)
if(vui_mvc_vcl_hrd_parameters_present_flag[i])		
hrd_parameters()	0	
if(vui_mvc_nal_hrd_parameters_present_flag[i] vui_mvc_vcl_hrd_parameters_present_flag[i])		
vui_mvc_low_delay_hrd_flag[i]	0	u(1)
vui_mvc_pic_struct_present_flag[i]	0	u(1)
}		
}		

En la tabla de ejemplo mostrada justo arriba, **depth_present_flag[i]** igual a 1 indica que el punto de funcionamiento actual incluye el componente de vista de profundidad para cada vista. **depth_present_flag[i]** igual a 0 indica que el punto de funcionamiento actual no incluye ningún componente de vista de profundidad.

[0112] Otros elementos sintácticos pueden tener la misma semántica que los de la extensión de parámetros MVC VUI con los mismos nombres.

[0113] De forma alternativa, se puede señalar **depth_present_idc[i]**. El elemento **depth_present_idc[i]** igual a 1 puede indicar que un punto de funcionamiento actual incluye el componente de vista de profundidad para cada vista. En algunos ejemplos, establecer **depth_present_idc[i]** igual a 0 o 2 puede indicar que el punto de funcionamiento actual incluye solo los componentes de vista de textura. En otros ejemplos, establecer **depth_present_idc[i]** igual a 0 puede indicar que el punto de funcionamiento actual no incluye ningún componente de vista de profundidad. En otros ejemplos, establecer **depth_present_idc[i]** igual a 2 puede indicar que el punto de funcionamiento actual no incluye ningún componente de vista de profundidad.

[0114] En otro ejemplo, un componente de vista de profundidad y un componente de vista de textura, incluso perteneciendo a la misma vista, pueden tener diferentes valores de **view_id** señalados. Sin embargo, los componentes de vista de profundidad todavía pueden compartir la misma dependencia de vista que la textura. En algunos ejemplos, la extensión SPS 3DVC contiene una asignación de los valores de **view_id**.

[0115] La siguiente tabla proporciona otro ejemplo de una extensión SPS 3DVC.

	C	Descriptor
seq_parameter_set_3dvc_extension() {		
for (i = 0; i <= num_views_minus1 ; i++)		
view_id_depth[i]	0	ue(v)
disable_depth_inter_view_flag	0	u(1)
depth_pic_width_in_mbs_minus1	0	ue(v)
depth_pic_height_in_map_units_minus1	0	ue(v)
depth_frame_cropping_flag	0	u(1)
if(depth_frame_cropping_flag) {		
depth_frame_crop_left_offset	0	ue(v)
depth_frame_crop_right_offset	0	ue(v)
depth_frame_crop_top_offset	0	ue(v)
depth_frame_crop_bottom_offset	0	ue(v)
}		
for(i = 0; i <= num_level_values_signalled_minus1; i++)		
level_idc_depth[i]	0	u(8)
}		

[0116] El elemento sintáctico **view_id_depth[i]** especifica el view_id de la vista de profundidad con VOIdx igual a i, pertenece a la vista teniendo la vista de textura un view_id igual a view_id[i].

5 [0117] En otro ejemplo, manteniendo iguales todos los demás aspectos, los valores de level_idc_depth se pueden señalar de la misma manera que los valores de level_idc señalados en la extensión SPS MVC.

10 [0118] En otro ejemplo, los componentes de vista de profundidad pueden compartir un seq_parameter_set_data() diferente. Como tal, se señalan una anchura y altura de los componentes de vista de profundidad. En un ejemplo, la sintaxis seq_parameter_set_data() puede incluir chorma_format_idc, que puede ser 4:0:0 en algunos ejemplos. En estos ejemplos, la relación de predicción de vista de profundidad todavía puede compartir la misma relación de predicción que la relación de predicción de vista de textura.

15 [0119] La siguiente tabla proporciona un ejemplo de elementos sintácticos para una sintaxis RBSP de un conjunto de parámetros de secuencia de subconjunto para cuando los componentes de vista de profundidad comparten el mismo seqparameter_set_data(), de acuerdo con un aspecto de la presente divulgación.

	C	Descriptor
subset_seq_parameter_set_rbsp() {		
seq_parameter_set_data()	0	
if(profile_idc == 83 profile_idc == 86) {		
seq_parameter_set_svc_extension() /*especificado en el Anexo G */	0	
svc_vui_parameters_present_flag	0	u(1)
if(svc_vui_parameters_present_flag == 1)		
seq_parameter_set_svc_extension() /* especificado en el Anexo G */	0	
} else if(profile_idc == 118 profile_idc == 128) {		
bit_equal_to_one /* igual a 1 */	0	f(1)
seq_parameter_set_mvc_extension() /*especificado en el Anexo H */	0	
mvc_vui_parameters_present_flag	0	u(1)
if(mvc_vui_parameters_present_flag == 1)		
seq_parameter_set_mvc_extension() /* especificado en el Anexo H */	0	
}		
else if (profile_idc == 138 profile_idc == 148) {		
//138: perfil base 3DV ; 148:perfil asimétrico 3DV		
bit_equal_to_one /* igual a 1 */	0	f(1)
seq_parameter_set_data()		
seq_parameter_set_mvc_extension()		
seq_parameter_set_3dvc_extension()		
3dvc_vui_parameter_present_flag	0	u(1)
if (3dvc_vui_parameter_present_flag) {		

3dvc_vui_parameters_extension() /* especificado en el Anexo I */		
}		
}		
additional_extension3_flag	0	u(1)
if (additional_extension3_flag)		
while(more_rbsp_data())		
additional_extension3_data_flag	0	u(1)
}		
rbp_trailing_bits()	0	
}		

[0120] La siguiente tabla proporciona una sintaxis de una extensión 3dVC de un conjunto de parámetros de secuencia de ejemplo para cuando los componentes de vista de profundidad comparten el mismo seq_parameter_set_data(), de acuerdo con un aspecto de la presente divulgación.

5

	C	Descriptor
seq_parameter_set_3dvc_extension() {		
disable_depth_inter_view_flag	0	u(1)
for(i = 0; i <= num_level_values_signalled_minus1; i++)		
level_idc_depth[i]	0	u(8)
}		

[0121] En algunos ejemplos, la semántica de algunos de los elementos sintácticos mostrados en algunas de las tablas anteriores puede ser similar o la misma que la de los elementos sintácticos análogos descritos con respecto a otras de estas tablas.

5 [0122] En otro ejemplo, se puede proporcionar un nuevo tipo de unidad NAL. El tipo de unidad NAL puede tener un valor, tal como 15, que se usa para señalar un súper SPS para el perfil 3DV. Las siguientes tres tablas pueden corresponder a elementos sintácticos para un súper SPS para el perfil 3DV. En algunos ejemplos, un súper SPS puede señalar tanto los componentes de vista de profundidad como los componentes de textura.

10 [0123] La siguiente tabla proporciona un ejemplo de una sintaxis Rbsp de un súper conjunto de parámetros de secuencia. La sintaxis de la extensión SPS 3DVC de la siguiente tabla se puede señalar usando el nuevo tipo de unidad NAL.

supset_seq_parameter_set_rbsp() {	C	Descriptor
seq_parameter_set_data()	0	
if (profile_idc == 138) {		
//138: perfil base 3DV;		
bit_equal_to_one /* igual a 1 */	0	f(1)
seq_parameter_set_mvc_extension()		

15

seq_parameter_set_3dvc_extension()		
3dvc_vui_parameter_present_flag	0	u(1)
if (3dvc_vui_parameter_present_flag) {		
3dvc_vui_parameters_extension() /* especificado en el Anexo I */		
}		
}		
additional_extension1_flag	0	u(1)
if (additional_extension1_flag)		
while(more_rbsp_data())		
additional_extension1_data_flag	0	u(1)
}		
rbsp_trailing_bits()	0	
}		

[0124] La siguiente tabla proporciona un ejemplo de una sintaxis de una extensión SPS 3DVC, de acuerdo con técnicas descritas en el presente documento. La sintaxis de la extensión SPS 3DVC de la siguiente tabla se puede señalar usando el nuevo tipo de unidad NAL.

20

seq_parameter_set_3dvc_extension() {	C	Descriptor
disable_depth_inter_view_flag	0	u(1)
cam_paras()		
depth_range()		
}		

[0125] La siguiente tabla proporciona un ejemplo de una sintaxis de una extensión de los parámetros 3DVC VUI, de acuerdo con técnicas descritas en el presente documento. La sintaxis de la extensión de los parámetros 3DVC VUI de la siguiente tabla se puede señalar usando el nuevo tipo de unidad NAL.

25

	C	Descriptor
3dvc_vui_parameters_extension() {		
vui_mvc_num_ops_minus1	0	ue(v)
for(i = 0; i <= vui_mvc_num_ops_minus1; i++) {		
vui_mvc_temporal_id[i]	0	u(3)
vui_mvc_num_target_output_views_minus1[i]	5	ue(v)
depth_present_flag[i]	0	u(1)
for(j = 0; j <=		
vui_mvc_num_target_output_views_minus1[i]; j++)		
vui_mvc_view_id[i][j]	5	ue(v)
vui_mvc_timing_info_present_flag[i]	0	u(1)
if(vui_mvc_timing_info_present_flag[i]) {		
vui_mvc_num_units_in_tick[i]	0	u(32)
vui_mvc_time_scale[i]	0	u(32)

vui_mvc_fixed_frame_rate_flag[i]	0	u(1)
}		
vui_mvc_nal_hrd_parameters_present_flag[i]	0	u(1)
if(vui_mvc_nal_hrd_parameters_present_flag[i])		
hrd_parameters()	0	
vui_mvc_vcl_hrd_parameters_present_flag[i]	0	u(1)
if(vui_mvc_vcl_hrd_parameters_present_flag[i])		
hrd_parameters()	0	
if(vui_mvc_nal_hrd_parameters_present_flag[i] vui_mvc_vcl_hrd_parameters_present_flag[i])		
vui_mvc_low_delay_hrd_flag[i]	0	u(1)
vui_mvc_pic_struct_present_flag[i]	0	u(1)
}		
}		

5 **[0126]** En algunos ejemplos, la semántica de algunos de los elementos sintácticos mostrados en las tablas anteriores puede ser similar o la misma que la de los elementos sintácticos análogos descritos con respecto a otros de las tablas anteriores. Los niveles definidos en la extensión MVC SPS se pueden aplicar a los puntos de funcionamiento, teniendo en cuenta las unidades NAL de profundidad. En el perfil relacionado con 3DV, se pueden activar simultáneamente un SPS, un SPS de subconjunto y un SPS de superconjunto.

10 **[0127]** Como resultado, MVC se puede extender mediante 3DVC para proporcionar un tratamiento separado de los componentes de profundidad y textura. Permitiendo dicho tratamiento independiente, los diversos aspectos de las técnicas descritas en la presente divulgación pueden promover el ahorro de bits (o, en otras palabras, una codificación más eficaz de los datos de vídeo de múltiples vistas) y/o una mejor calidad de los datos de vídeo reconstruidos (que de nuevo se puede medir en términos de una experiencia visual percibida).

15 **[0128]** Por ejemplo, diversos aspectos de las técnicas descritas en la presente divulgación pueden permitir que un dispositivo de codificación de vídeo (que puede representar un término usado en la presente divulgación para referirse a cualquier dispositivo que incluya un codificador de vídeo y/o un descodificador de vídeo, tal como el dispositivo de origen 12 y/o el dispositivo de destino 14) active un conjunto de parámetros como un conjunto de parámetros de textura para el componente de textura del componente de vista basándose al menos en un valor de índice de orden de vista asignado al componente de vista de los datos de vídeo de múltiples vistas. El índice de orden de vista puede describir el orden de decodificación de un componente de vista correspondiente de una unidad de acceso. El índice de orden de vista se puede usar con un valor de recuento de orden de imagen (POC) o un valor de trama para identificar un componente de vista de un flujo de bits. Por lo tanto, este aspecto de las técnicas puede proporcionar un conjunto de parámetros, tal como un PPS o SPS, que se activa solo para el componente de textura de un componente de vista y no se aplica al componente de profundidad correspondiente del mismo componente de vista. A este respecto, el dispositivo de codificación de vídeo puede codificar solo el componente de textura del componente de vista y no el componente de profundidad del componente de vista basándose en el conjunto de parámetros de textura activados.

[0129] Adicionalmente, el dispositivo de codificación de vídeo puede activar un conjunto de parámetros para el componente de profundidad del componente de vista basándose al menos en el valor de índice de orden de vista asignado al componente de vista de los datos de vídeo de múltiples vistas. Es decir, en algunos casos, el dispositivo de codificación de vídeo puede activar un conjunto de parámetros específicos de textura para el componente de textura de un componente de vista y a continuación activar aún otro conjunto de parámetros específicamente para el componente de profundidad del mismo componente de vista. Como resultado, el dispositivo de codificación de vídeo puede codificar el componente de profundidad del componente de vista basándose en el conjunto de parámetros activados para el componente de profundidad del componente de vista y no en el conjunto de parámetros de textura activados. El dispositivo de codificación de vídeo puede codificar el componente de textura basándose en el conjunto de parámetros de textura activados. De esta manera, las técnicas pueden promover la activación separada de conjuntos de parámetros, tales como un PPS o SPS, para los componentes de textura y profundidad de un componente de vista, proporcionando una manera de facilitar el tratamiento o procesamiento separado de los componentes de profundidad y textura en una vista 3DVC compatible con MVC.

[0130] Diversos aspectos de las técnicas descritas en la presente divulgación también pueden promover el tratamiento separado de los componentes de profundidad y textura al permitir que un dispositivo de codificación de vídeo determine un mensaje de información complementaria (SEI) anidado que se aplica tanto a los componentes de textura como de profundidad de un componente de vista o solo a un componente de profundidad del componente de vista. El término "mensaje SEI anidado" se refiere a un mensaje SEI que incorpora diversas partes de otro mensaje SEI, de modo que un mensaje SEI anidado se "anide" eficazmente dentro del alcance de otro mensaje SEI. Permitiendo que un mensaje SEI anidado se aplique solo a un componente de profundidad, estos aspectos de las técnicas pueden promover de nuevo el tratamiento separado de los componentes de textura y profundidad.

[0131] Un dispositivo de codificación de vídeo (es decir, el dispositivo de origen 12 y/o el dispositivo de destino 14 en el ejemplo de la FIG. 1) puede llevar a cabo estos aspectos de las técnicas para determinar un mensaje de información de mejora complementaria que se aplica a la codificación del componente de vista de los datos de vídeo de múltiples vistas. Este mensaje SEI se puede denominar mensaje SEI original para reflejar que se aplica a un componente de vista particular y establece el alcance de cualquier mensaje SEI anidado. El dispositivo de codificación de vídeo, a continuación, puede determinar un mensaje de información de mejora complementaria anidado que se aplica además del mensaje de información de mejora complementaria al codificar el componente de profundidad del componente de vista. Este mensaje SEI anidado puede identificar el mensaje SEI original e indicar que este mensaje SEI anidado solo se aplica al componente de profundidad del componente de vista al que se aplica el mensaje SEI original. El dispositivo de codificación de vídeo, a continuación, puede procesar el componente de profundidad del componente de vista basándose en el mensaje de información de mejora complementaria y el mensaje de información de mejora complementaria anidado.

[0132] Típicamente, el mensaje SEI anidado incluye un indicador u otro identificador que indica si el mensaje SEI anidado se aplica tanto a los componentes de textura como de profundidad del componente de vista o sólo al componente de profundidad. Por tanto, en algunos casos, el dispositivo de codificación de vídeo puede determinar que el mensaje de información de mejora complementaria anidado se aplica además del mensaje de información de mejora complementaria al procesar solo el componente de profundidad del componente de vista y procesa solo el componente de profundidad del componente de vista y no el componente de textura del componente de vista basándose en el mensaje de información de mejora complementaria y el mensaje de información de mejora complementaria anidado. De forma alternativa, el dispositivo de codificación de vídeo puede determinar que el mensaje de información de mejora complementaria anidado se aplica además del mensaje de información de mejora complementaria cuando procesa el componente de textura del componente de vista y procesa el componente de textura del componente de vista basándose en el mensaje de información de mejora complementaria y el mensaje de información de mejora complementaria anidado. De nuevo, estos aspectos de las técnicas facilitan de nuevo el tratamiento o procesamiento separado de los componentes de profundidad y textura del componente de vista.

[0133] Otros aspectos de las técnicas descritas en la presente divulgación pueden promover el tratamiento separado de los componentes de profundidad y textura permitiendo que un dispositivo de codificación de vídeo elimine por separado los componentes de textura y profundidad a partir de una memoria intermedia de imágenes descodificadas. La memoria intermedia de imágenes descodificadas se refiere a una memoria u otro medio legible por ordenador que puede almacenar imágenes de referencia para su uso en la codificación de un componente de vista actual. Debido a que las técnicas expuestas en la presente divulgación permiten el tratamiento separado de los componentes de textura y profundidad, un dispositivo de codificación de vídeo puede codificar un componente de textura de un componente de vista usando imágenes de referencia diferentes a las usadas al codificar un componente de profundidad del mismo componente de vista en contraposición con la MVC convencional donde un componente de vista (y por lo tanto los componentes de profundidad y textura de ese componente de vista) se codificó usando las mismas imágenes de referencia. Como resultado, pueden existir casos donde los componentes de profundidad de un componente de vista de referencia se puedan eliminar de la memoria intermedia de imágenes descodificadas antes de los componentes de textura del mismo componente de vista de referencia, y un componente de textura de un componente de vista de referencia se pueda eliminar de la memoria intermedia de imágenes descodificadas antes del componente de profundidad del mismo componente de vista de referencia.

[0134] En funcionamiento, un dispositivo de codificación de vídeo puede llevar a cabo las técnicas descritas en la presente divulgación almacenando un componente de profundidad en una memoria intermedia de imágenes descodificadas y analizando una dependencia de vista para determinar si el componente de profundidad se usa para la predicción entre vistas. El dispositivo de codificación de vídeo, a continuación, puede eliminar el componente de profundidad de la memoria intermedia de imágenes descodificadas en respuesta a la determinación de que el componente de profundidad no se usa para la predicción entre vistas. Es decir, el dispositivo de codificación de vídeo puede eliminar el componente de profundidad de un componente de vista de referencia de la memoria intermedia de imágenes descodificadas sin eliminar el componente de textura del mismo componente de vista de referencia en respuesta a la determinación de que el componente de profundidad no se usa para la predicción entre vistas. El dispositivo de codificación de vídeo puede, antes de eliminar el componente de profundidad, validar este componente de profundidad como que es elegible para su eliminación determinando que el componente de profundidad no pertenece a una vista de salida de destino y que el componente de profundidad está asociado con un código de identificación de referencia de capa de abstracción de red que tiene un valor igual a cero (lo que significa que el contenido de la unidad NAL que encapsula el componente de profundidad no se usa para construir imágenes de referencia para la predicción inter-imágenes).

[0135] Los aspectos adicionales de las técnicas descritas en la presente divulgación también pueden permitir o facilitar de otro modo el tratamiento separado de los componentes de profundidad y textura. Por ejemplo, en términos de puntos de funcionamiento, las técnicas pueden permitir que un dispositivo de codificación de vídeo genere un mensaje SEI que incluya un primer elemento sintáctico, un segundo elemento sintáctico y un tercer elemento sintáctico, donde el primer elemento sintáctico indica si un punto de funcionamiento contiene un componente de profundidad. Cuando el punto de funcionamiento contiene un componente de profundidad, el dispositivo de codificación de vídeo puede definir el segundo elemento sintáctico para indicar un número de componentes de profundidad de los que depende directamente una vista de salida de destino del punto de funcionamiento y definir los terceros elementos sintácticos para identificar los componentes de profundidad de los que depende directamente la vista de salida de destino del punto de funcionamiento.

[0136] En algunos casos, los puntos de funcionamiento pueden incluir un primer subconjunto de los puntos de funcionamiento y un segundo subconjunto de los puntos de funcionamiento, donde cada punto de funcionamiento en el primer subconjunto contiene sólo un componente de textura y cada punto de funcionamiento en el segundo subconjunto contiene un componente de profundidad. El dispositivo de codificación de vídeo, a continuación, puede determinar que un identificador de perfil es igual a un valor dado que está asociado con la codificación de vídeo 3D y generar, en respuesta a la determinación de que el identificador de perfil sea igual al valor dado, un SPS de subconjunto de modo que el SPS de subconjunto incluya una primera extensión SPS y una segunda extensión SPS. La primera extensión SPS puede indicar el primer subconjunto y los niveles a los que pertenecen los puntos de funcionamiento en el primer subconjunto. La segunda extensión SPS puede indicar el segundo subconjunto y los niveles a los que pertenecen los puntos de funcionamiento en el segundo subconjunto. De esta manera, las técnicas pueden permitir que los puntos de operaciones se definan por separado para los componentes de textura y profundidad y por nivel.

[0137] Este aspecto de las técnicas puede aprovechar el tratamiento separado de los componentes de profundidad y textura para permitir que los puntos de funcionamiento se extraigan más fácilmente de un flujo de bits para la codificación de vídeo de múltiples vistas. Señalando si un punto de funcionamiento incluye profundidad, un descodificador de vídeo, tal como el descodificador de vídeo 30, puede determinar qué punto de funcionamiento da cabida mejor a las capacidades del descodificador de vídeo 30. Por ejemplo, el descodificador de vídeo 30 se puede optimizar para llevar a cabo la codificación de vídeo de múltiples vistas usando vistas que no incluyen profundidad, prefiriendo datos de vídeo de múltiples vistas donde se proporcionan tanto las perspectivas del ojo izquierdo como del derecho como imágenes separadas. El descodificador de vídeo 30 también se puede optimizar, de forma alternativa, para datos de vídeo de múltiples vistas basados en profundidad, pero solo puede dar cabida a una vista única en lugar de dos o más vistas. Señalando cuándo los datos de vídeo de múltiples vistas codificados incluyen profundidad, y al incluir la profundidad, el número de componentes de profundidad de los que dependen directamente las vistas de salida de destino del punto de funcionamiento, el descodificador de vídeo 30 puede seleccionar un punto de funcionamiento que dé cabida a las capacidades u optimizaciones del descodificador de vídeo 30.

[0138] Adicionalmente, un dispositivo de codificación de vídeo puede llevar a cabo aspectos de las técnicas descritas en la presente divulgación para señalar los niveles a los que pertenecen los puntos de funcionamiento, donde al menos algunos de los puntos de funcionamiento contienen componentes de profundidad. El dispositivo de codificación de vídeo, a continuación, puede determinar que un identificador de perfil es igual a un valor dado que está asociado con la codificación de vídeo 3D y generar, en respuesta a la determinación de que el identificador de perfil sea igual al valor dado, un SPS de subconjunto de modo que el SPS de subconjunto incluya una extensión MVC SPS que señale los puntos de funcionamiento y los niveles a los que pertenecen los puntos de funcionamiento. Cada uno de los puntos de funcionamiento, en este ejemplo, puede especificar una o más vistas de salida, donde cada una de las vistas de salida contiene tanto un componente de textura como un componente de profundidad.

[0139] Aún en otros casos, el dispositivo de codificación de vídeo puede llevar a cabo las técnicas descritas en la presente divulgación para generar, en un SPS, un elemento sintáctico que indica un número de vistas que se van a descodificar para un punto de funcionamiento en un proceso MVC, donde cada una de las vistas tiene uno de un número de componentes de textura y uno de un número de componentes de profundidad. El dispositivo de codificación de vídeo puede señalar, a continuación, en el SPS, el número de componentes de textura que se van a descodificar para el punto de funcionamiento y el número de componentes de profundidad que se van a descodificar para el punto de funcionamiento. De nuevo, en virtud de facilitar el tratamiento separado de los componentes de profundidad y textura, las técnicas pueden proporcionar la señalización de cómo se utilizan estos componentes de profundidad y textura para formar puntos de funcionamiento.

[0140] La FIG. 2 es un diagrama de bloques que ilustra un codificador de vídeo de ejemplo 20 que puede implementar las técnicas descritas en la presente divulgación para codificar datos de vídeo de múltiples vistas. El codificador de vídeo 20 recibe datos de vídeo que se va a codificar. En el ejemplo de la FIG. 2, el codificador de vídeo 20 incluye una unidad de selección de modo 40, un sumador 50, una unidad de procesamiento de transformada 52, una unidad de cuantificación 54, una unidad de codificación por entropía 56 y una memoria de imágenes de referencia 64. A su vez, la unidad de selección de modo 40 incluye una unidad de estimación de movimiento/disparidad 42, una unidad de compensación de movimiento 44, una unidad de predicción intra 46 y una unidad de división 48.

[0141] Para la reconstrucción de bloques de vídeo, el codificador de vídeo 20 incluye también una unidad de cuantificación inversa 58, una unidad de procesamiento de transformada inversa 60 y un sumador 62. También se puede incluir un filtro de eliminación de bloques (no mostrado en la FIG. 2) para filtrar fronteras de bloques y eliminar distorsiones de efecto pixelado del vídeo reconstruido. Si se desea, el filtro de eliminación de bloques filtrará típicamente la salida del sumador 62. También se pueden usar filtros de bucle adicionales (en el bucle o tras el bucle), además del filtro de eliminación de bloques. Dichos filtros no se muestran por razones de brevedad pero, si se desea, pueden filtrar la salida del sumador 50 (como un filtro en bucle).

[0142] La unidad de selección de modo 40 puede recibir datos de vídeo sin procesar en forma de bloques de una o más vistas. La unidad de selección de modo 40 puede seleccionar uno de los modos de codificación, intra o inter, por ejemplo, basándose en los resultados de errores, y proporciona el bloque intra o intercodificado resultante al sumador 50 para generar datos de bloques residuales, y al sumador 62 para reconstruir el bloque codificado para su uso como una imagen de referencia. La unidad de selección de modo 40 proporciona también elementos sintácticos, tales como vectores de movimiento, indicadores de modo intra, información de división y otra información sintáctica de este tipo, a la unidad de codificación por entropía 56.

[0143] La unidad de estimación de movimiento/disparidad 42 y la unidad de compensación de movimiento 44 pueden estar altamente integradas, pero se ilustran por separado para propósitos conceptuales. La estimación del movimiento, llevada a cabo por la unidad de estimación de movimiento/disparidad 42, es el proceso de generación de vectores de movimiento, que estiman el movimiento de los bloques de vídeo. Un vector de movimiento, por ejemplo, puede indicar el desplazamiento de una PU de un bloque de vídeo dentro de una imagen actual, con respecto a un bloque predictivo dentro de una imagen de referencia (u otra unidad codificada), con respecto al bloque actual que se está codificando dentro de la imagen actual (u otra unidad codificada).

[0144] Un bloque predictivo es un bloque que se encuentra que coincide estrechamente con el bloque que se va a codificar, en términos de la diferencia de píxeles, lo que se puede determinar mediante la suma de la diferencia absoluta (SAD), suma de la diferencia de cuadrados (SSD) u otras métricas de diferencia. En algunos ejemplos, el codificador de vídeo 20 puede calcular valores para posiciones fraccionarias de píxel de imágenes de referencia almacenadas en la memoria de imágenes de referencia 64, que también se puede denominar memoria intermedia de imágenes de referencia. Por ejemplo, el codificador de vídeo 20 puede interpolar valores de posiciones de un cuarto de píxel, posiciones de un octavo de píxel u otras posiciones fraccionarias de píxel de la imagen de referencia. Por lo tanto, la unidad de estimación de movimiento/disparidad 42 puede llevar a cabo una búsqueda de movimiento con respecto a las posiciones de píxel completo y las posiciones de píxel fraccionario, y emitir un vector de movimiento con una precisión de píxel fraccionario.

[0145] La unidad de estimación de movimiento/disparidad 42 calcula unos vectores de movimiento para una PU de un bloque de vídeo en un fragmento intercodificado, comparando la posición de la PU con la posición de un bloque predictivo de una imagen de referencia. La unidad de estimación/disparidad de movimiento 42 también se puede configurar para llevar a cabo la predicción entre vistas, en cuyo caso la unidad de estimación/disparidad de movimiento 42 puede calcular vectores de desplazamiento entre los bloques de una imagen de vista (por ejemplo, la vista 0) y los bloques correspondientes de una imagen de vista de referencia (por ejemplo, la vista 1). En general, los datos para un vector de movimiento/disparidad pueden incluir una lista de imágenes de referencia, un índice en la lista de imágenes de referencia (ref_idx), un componente horizontal y un componente vertical. La imagen de referencia se puede seleccionar de una primera lista de imágenes de referencia (Lista 0), una segunda lista de imágenes de referencia (Lista 1), o una lista de imágenes de referencia combinada (Lista C), cada una de las cuales identifica una o más imágenes de referencia almacenadas en una memoria de imágenes de referencia 64. Con respecto a la lista combinada, el codificador de vídeo 20 selecciona de forma alternativa las entradas de dos listas (es decir, la Lista 0 y la Lista 1) que se van a insertar (adjuntar) en la lista combinada. Cuando una entrada ya está en la lista combinada,

comprobando el número POC, el codificador de vídeo 20 no puede insertar la entrada de nuevo. Para cada lista (es decir, la Lista 0 o la Lista 1), el codificador de vídeo 20 puede seleccionar las entradas basándose en el orden ascendente del índice de referencia.

5 **[0146]** La unidad de estimación de movimiento/disparidad 42 puede generar y enviar un vector de movimiento/disparidad que identifica el bloque predictivo de la imagen de referencia a la unidad de codificación por entropía 56 y a la unidad de compensación de movimiento 44. Es decir, la unidad de estimación de movimiento/disparidad 42 puede generar y enviar datos de vectores de movimiento que identifican la lista de imágenes de referencia que contiene el bloque predictivo, un índice en la lista de imágenes de referencia que identifica la imagen del bloque predictivo y un componente horizontal y vertical para ubicar el bloque predictivo dentro de la imagen identificada.

15 **[0147]** La compensación de movimiento, llevada a cabo por la unidad de compensación de movimiento 44, puede implicar captar o generar el bloque predictivo basándose en el vector de movimiento/disparidad determinado por la unidad de estimación de movimiento/disparidad 42. De nuevo, la unidad de estimación de movimiento/disparidad 42 y la unidad de compensación de movimiento 44 se pueden integrar funcionalmente, en algunos ejemplos. Tras recibir el vector de movimiento para la PU del bloque de vídeo actual, la unidad de compensación de movimiento 44 puede ubicar el bloque predictivo al que apunta el vector de movimiento en una de las listas de imágenes de referencia.

20 **[0148]** El sumador 50 forma un bloque de vídeo residual restando los valores de píxel del bloque predictivo a los valores de píxel del bloque de vídeo actual que se está codificando, formando valores de diferencias de píxel, como se analiza a continuación. En general, la unidad de estimación de movimiento/disparidad 42 lleva a cabo la estimación de movimiento con respecto a los componentes de luma, y la unidad de compensación de movimiento 44 usa los vectores de movimiento calculados basándose en los componentes de luma, tanto para los componentes de croma como para los componentes de luma. La unidad de selección de modo 40 también puede generar elementos sintácticos asociados con los bloques de vídeo y el fragmento de vídeo para su uso por el descodificador de vídeo 30 en la descodificación de los bloques de vídeo del fragmento de vídeo.

30 **[0149]** La unidad de predicción intra 46 puede realizar predicción intra de un bloque actual, como alternativa a la predicción inter llevada a cabo por la unidad de estimación de movimiento/disparidad 42 y la unidad de compensación de movimiento 44, como se describe anteriormente. En particular, la unidad de predicción intra 46 puede determinar un modo de predicción intra para usar para codificar un bloque actual. En algunos ejemplos, la unidad de predicción intra 46 puede codificar un bloque actual usando diversos modos de predicción intra, por ejemplo, durante pasadas de codificación separadas, y la unidad de predicción intra 46 (o la unidad de selección de modo 40, en algunos ejemplos) puede seleccionar un modo de predicción intra apropiado para usar a partir de los modos probados.

40 **[0150]** Por ejemplo, la unidad de predicción intra 46 puede calcular valores de distorsión de velocidad usando un análisis de distorsión de velocidad para los diversos modos de predicción intra probados, y seleccionar el modo de predicción intra que tenga las mejores características de distorsión de velocidad entre los modos probados. El análisis de distorsión de velocidad determina en general una magnitud de distorsión (o de errores) entre un bloque codificado y un bloque original, no codificado, que se codificó para producir el bloque codificado, así como una tasa de bits (es decir, un número de bits) usada para producir el bloque codificado. La unidad de predicción intra 46 puede calcular proporciones a partir de las distorsiones y velocidades para los diversos bloques codificados para determinar qué modo de predicción intra presenta el mejor valor de velocidad-distorsión para el bloque.

45 **[0151]** Después de seleccionar un modo de predicción intra para un bloque, la unidad de predicción intra 46 puede proporcionar información indicativa del modo de predicción intra seleccionado para el bloque a la unidad de codificación por entropía 56. La unidad de codificación por entropía 56 puede codificar la información que indica el modo de predicción intra seleccionado. El codificador de vídeo 20 puede incluir datos de configuración en el flujo de bits transmitido, que pueden incluir una pluralidad de tablas de índices de modo de predicción intra y una pluralidad de tablas de índices de modo de predicción intra modificados (también denominadas tablas de asignación de palabras de código), definiciones de contextos de codificación para diversos bloques e indicaciones de un modo de predicción intra más probable, una tabla de índices de modo de predicción intra y una tabla de índices de modo de predicción intra modificados para usar para cada uno de los contextos.

55 **[0152]** El codificador de vídeo 20 forma un bloque de vídeo residual restando los datos de predicción de la unidad de selección de modo 40 del bloque de vídeo original que se está codificado. El sumador 50 representa el componente o los componentes que llevan a cabo esta operación de resta. La unidad de procesamiento de transformada 52 aplica una transformada, tal como una transformada discreta de coseno (DCT) o una transformada conceptualmente similar, al bloque residual, produciendo un bloque de vídeo que comprende valores residuales de coeficientes de transformada. La unidad de procesamiento de transformada 52 puede llevar a cabo otras transformadas que son conceptualmente similares a la DCT. También se podrían usar transformadas de ondículas, transformadas de números enteros, transformadas de subbandas u otros tipos de transformadas. En cualquier caso, la unidad de procesamiento de transformada 52 aplica la transformada al bloque residual, produciendo un bloque de coeficientes de transformada residuales. La transformada puede convertir la información residual desde un dominio de valor de píxel a un dominio de transformada, tal como un dominio de frecuencia.

- 5 **[0153]** La unidad de procesamiento de transformada 52 puede enviar los coeficientes de transformada resultantes a la unidad de cuantificación 54. La unidad de cuantificación 54 cuantifica los coeficientes de transformada para reducir más la tasa de bits. El proceso de cuantificación puede reducir la profundidad de bits asociada con algunos, o a la totalidad, de los coeficientes. El grado de cuantificación se puede modificar ajustando un parámetro de cuantificación. En algunos ejemplos, la unidad de cuantificación 54 puede llevar a cabo, a continuación, una exploración de la matriz que incluye los coeficientes de transformada cuantificados. De forma alternativa, la unidad de codificación por entropía 56 puede llevar a cabo la exploración.
- 10 **[0154]** Tras la cuantificación, la unidad de codificación por entropía 56 codifica por entropía los coeficientes de transformada cuantificados. Por ejemplo, la unidad de codificación por entropía 56 puede llevar a cabo la codificación de longitud variable adaptativa según el contexto (CAVLC), la codificación aritmética binaria adaptativa según el contexto (CABAC), la codificación aritmética binaria adaptativa según el contexto y basada en sintaxis (SBAC), la codificación por entropía mediante la división en intervalos de probabilidades (PIPE) u otra técnica de codificación por entropía. En el caso de la codificación por entropía basada en el contexto, el contexto se puede basar en bloques contiguos. Tras la codificación por entropía por la unidad de codificación por entropía 56, el flujo de bits codificado se puede transmitir a otro dispositivo (por ejemplo, el descodificador de vídeo 30) o archivar para su posterior transmisión o recuperación.
- 15 **[0155]** La unidad de cuantificación inversa 58 y la unidad de procesamiento de transformada inversa 60 aplican la cuantificación inversa y la transformación inversa, respectivamente, para reconstruir el bloque residual en el dominio del píxel, por ejemplo, para su posterior uso como bloque de referencia. La unidad 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 la memoria de imágenes de referencia 64. La unidad de compensación de movimiento 44 también puede aplicar uno o más filtros de interpolación al bloque residual reconstruido para calcular valores fraccionarios de píxel para su uso en la estimación de movimiento. El sumador 62 añade el bloque residual reconstruido al bloque de predicción compensado por movimiento producido por la unidad de compensación de movimiento 44 para producir un bloque de vídeo reconstruido para su almacenamiento en la memoria de imágenes de referencia 64. El bloque de vídeo reconstruido se puede usar por la unidad de estimación de movimiento/disparidad 42 y la unidad de compensación de movimiento 44 como un bloque de referencia para intercodificar un bloque en una imagen posterior.
- 20 **[0156]** El codificador de vídeo 20 puede generar un número de elementos sintácticos, como se describe anteriormente, que se pueden codificar por la unidad de codificación por entropía 56 u otra unidad de codificación del codificador de vídeo 20. En algunos ejemplos, el codificador de vídeo 20 puede generar y codificar elementos sintácticos para un flujo de bits de 3DVC, como se describe anteriormente, donde de nuevo este flujo de bits de 3DVC puede ser retrocompatible con MVC. Es decir, un descodificador de vídeo que admite MVC pero que no admite 3DVC todavía puede descodificar el flujo de bits de 3DVC codificado de acuerdo con las técnicas descritas en la presente divulgación. [explicar/ilustrar?]
- 25 **[0157]** El codificador de vídeo 20 puede activar un conjunto de parámetros como un conjunto de parámetros de textura para el componente de textura del componente de vista basándose al menos en un valor de índice de orden de vista asignado al componente de vista de los datos de vídeo de múltiples vistas. En otras palabras, el codificador de vídeo 20 puede activar un conjunto de parámetros, tal como un PPS o SPS, solo para el componente de textura de un componente de vista, de modo que este conjunto de parámetros de textura no se aplique al componente de profundidad correspondiente del mismo componente de vista. A este respecto, el codificador de vídeo 20 puede codificar solo el componente de textura del componente de vista y no el componente de profundidad del componente de vista basándose en el conjunto de parámetros de textura activados.
- 30 **[0158]** En particular, la unidad de selección de modo 40 del codificador de vídeo 20 puede activar un conjunto de parámetros seleccionando un modo, por ejemplo, predicción intravista (que puede implicar predicción temporal o inter-imagen donde todas las imágenes corresponden a la misma vista), predicción entre vistas o predicción intra. La unidad de selección de modo 40 puede analizar cada una de las iteraciones de codificación de una imagen, donde cada iteración puede implicar diferentes parámetros. Al seleccionar una iteración diversa de codificación de una imagen, la unidad de selección de modo 40 puede activar eficazmente un conjunto de parámetros para una secuencia particular de imágenes, vista, componente de vista, componente de profundidad y/o componente de textura. Además, la unidad de selección de modo 40 puede activar diferentes conjuntos de parámetros de esta forma para los componentes de textura y profundidad que pertenecen cada uno al mismo componente de vista de la forma que se describe a continuación con más detalle.
- 35 **[0159]** Adicionalmente, el modo de unidad de selección 40 puede activar un conjunto de parámetros para el componente de profundidad del componente de vista basándose al menos en el valor de índice de orden de vista asignado al componente de vista de los datos 3DVC. Es decir, en algunos casos, la unidad de selección de modo 40 puede activar un conjunto de parámetros específicos de textura para el componente de textura de un componente de vista y a continuación activar otro conjunto de parámetros específicamente para el componente de profundidad del mismo componente de vista. Como resultado, la unidad de selección de modo 40 puede codificar el componente de profundidad del componente de vista basándose en el conjunto de parámetros activados para el componente de
- 40
- 45
- 50
- 55
- 60
- 65

profundidad del componente de vista y no el conjunto de parámetros de textura activados. La unidad de selección de modo 40 puede codificar el componente de textura basándose en el conjunto de parámetros de textura activados. De esta manera, las técnicas pueden promover la activación separada de los conjuntos de parámetros, tal como un PPS o SPS, para los componentes de textura y profundidad de un componente de vista, proporcionando una manera de facilitar el tratamiento o procesamiento separado de los componentes de profundidad y textura en 3DVC.

[0160] El codificador de vídeo 20 también puede eliminar por separado los componentes de textura y profundidad de una memoria intermedia de imágenes descodificadas, que se muestra como memoria de imágenes de referencia 64 en el ejemplo de la FIG. 2. Debido a que las técnicas expuestas en la presente divulgación permiten el tratamiento separado de los componentes de textura y profundidad, el codificador de vídeo 20 puede codificar un componente de textura de un componente de vista usando imágenes de referencia diferentes a las usadas al codificar un componente de profundidad del mismo componente de vista en contraposición a 3DVC convencional donde un componente de vista (y por lo tanto los componentes de profundidad y textura de ese componente de vista) se codificó usando las mismas imágenes de referencia. Como resultado, pueden existir casos donde los componentes de profundidad de un componente de vista de referencia se puedan eliminar de la memoria intermedia de imágenes descodificadas antes de los componentes de textura del mismo componente de vista de referencia, y un componente de textura de un componente de vista de referencia se pueda eliminar de la memoria intermedia de imágenes descodificadas antes del componente de profundidad del mismo componente de vista de referencia.

[0161] En funcionamiento, la unidad de selección de modo 40 del codificador de vídeo 20 se puede configurar para llevar a cabo las técnicas descritas en la presente divulgación. La unidad de selección de modo 40 puede almacenar un componente de profundidad en la memoria de imagen de referencia 64 y analizar una dependencia de vista para determinar si el componente de profundidad se usa para la predicción entre vistas. La unidad de selección de modo 40, a continuación, puede eliminar el componente de profundidad de la memoria de imágenes de referencia 64 en respuesta a la determinación de que el componente de profundidad no se usa para la predicción entre vistas. Es decir, la unidad de selección de modo 40 puede eliminar el componente de profundidad de un componente de vista de referencia de la memoria de imágenes de referencia 64 sin eliminar el componente de textura del mismo componente de vista de referencia en respuesta a la determinación de que el componente de profundidad no se usa para la predicción entre vistas. La unidad de selección de modo 40, antes de eliminar el componente de profundidad, puede validar este componente de profundidad como que es elegible para su eliminación determinando que el componente de profundidad no pertenece a una vista de salida de destino y que el componente de profundidad está asociado con un código de identificación de referencia de capa de abstracción de red que tiene un valor igual a cero (lo que significa que el contenido de la unidad NAL que encapsula el componente de profundidad no se usa para construir imágenes de referencia para la predicción inter-imagen).

[0162] La FIG. 3 es un diagrama de bloques que ilustra un descodificador de vídeo 30 de ejemplo que puede implementar las técnicas descritas en la presente divulgación para descodificar un flujo de bits de múltiples vistas. En el ejemplo de la FIG. 3, el descodificador de vídeo 30 incluye una unidad de descodificación por entropía 80, una unidad de predicción 81 que tiene la unidad de compensación de movimiento 82 y la unidad de predicción intra 84, la unidad de cuantificación inversa 86, la unidad de transformación inversa 88, el sumador 90 y la memoria de imágenes de referencia 92.

[0163] Durante el proceso de descodificación, el descodificador de vídeo 30 recibe un flujo de bits de vídeo codificado, que representa los bloques de vídeo de un fragmento de vídeo codificado y elementos sintácticos asociados, desde el codificador de vídeo 20. La unidad de descodificación por entropía 80 del descodificador de vídeo 30 descodifica por entropía el flujo de bits para generar coeficientes cuantificados, vectores de movimiento y otros elementos sintácticos. La unidad de descodificación por entropía 80 reenvía los vectores de movimiento y otros elementos sintácticos a la unidad de predicción 81. El descodificador de vídeo 30 puede recibir los elementos sintácticos en el nivel de fragmento de vídeo y/o el nivel de bloque de vídeo.

[0164] Por ejemplo, el descodificador de vídeo 30 puede recibir un número de unidades NAL que tienen una cabecera de unidad NAL que identifica un tipo de datos almacenados en la unidad NAL (por ejemplo, datos de VCL y datos distintos a VCL). Los conjuntos de parámetros pueden contener la información de cabecera a nivel de secuencia, tal como un SPS, PPS u otro conjunto de parámetros descrito anteriormente.

[0165] De acuerdo con aspectos de la presente divulgación, el descodificador de vídeo 30 puede activar un conjunto de parámetros como un conjunto de parámetros de textura para el componente de textura del componente de vista basándose al menos en un valor de índice de orden de vista asignado al componente de vista de los datos de 3DVC. Más específicamente, la unidad de descodificación por entropía 80 puede analizar los conjuntos de parámetros del flujo de bits, descodificar los conjuntos de parámetros y proporcionar los conjuntos de parámetros descodificados a los otros componentes del descodificador de vídeo 30, tales como la unidad de predicción 81, la unidad de cuantificación inversa 86 y la unidad de procesamiento de transformada inversa 88. Cada uno de estos componentes se puede referir a los conjuntos de parámetros actualmente activos y, en este sentido, el descodificador de vídeo 30 activa el conjunto de parámetros para cada componente del descodificador de vídeo 30.

[0166] En cualquier caso, el índice de orden de vista, como se indica anteriormente, puede describir el orden de descodificación de un componente de vista correspondiente de una unidad de acceso. El índice de orden de vista se puede usar con un valor de recuento de orden de imagen (POC) o un valor de trama para identificar un componente de vista de un flujo de bits. Por lo tanto, este aspecto de las técnicas puede proporcionar un conjunto de parámetros, tal como un PPS o SPS, que se activa solo para el componente de textura de un componente de vista y no se aplica al componente de profundidad correspondiente del mismo componente de vista. A este respecto, el descodificador de vídeo 30 puede descodificar solo el componente de textura del componente de vista y no el componente de profundidad del componente de vista basándose en el conjunto de parámetros de textura activados.

[0167] Adicionalmente, el descodificador de vídeo 30 puede activar un conjunto de parámetros para el componente de profundidad del componente de vista basándose al menos en el valor del índice de orden de vista asignado al componente de vista de los datos de 3DV. Es decir, en algunos casos, el dispositivo de codificación de vídeo puede activar un conjunto de parámetros específicos de textura para el componente de textura de un componente de vista y a continuación activar aún otro conjunto de parámetros específicamente para el componente de profundidad del mismo componente de vista. Como resultado, el descodificador de vídeo 30 puede codificar el componente de profundidad del componente de vista basándose en el conjunto de parámetros activados para el componente de profundidad del componente de vista y no en el conjunto de parámetros de textura activados. El descodificador de vídeo 30 puede descodificar el componente de textura basándose en el conjunto de parámetros de textura activados. De esta manera, las técnicas pueden promover la activación separada de los conjuntos de parámetros, tal como un PPS o SPS, para los componentes de textura y profundidad de un componente de vista, proporcionando una manera de facilitar el tratamiento o procesamiento separado de los componentes de profundidad y textura en 3DVC.

[0168] En detalle, este aspecto de las técnicas puede permitir que un codificador de vídeo (que en general se puede referir a amos o a cualquiera de un codificador de vídeo y/o a un descodificador de vídeo) active un conjunto de parámetros separado para cada uno de un componente de profundidad y un componente de textura. El mecanismo de conjunto de parámetros de secuencia e imagen puede desacoplar eficazmente la transmisión de información que cambia con poca frecuencia de la transmisión de datos de macrobloques codificados. En algunas aplicaciones, los conjuntos de parámetros de secuencia e imagen se pueden transportar fuera de banda usando un mecanismo de transporte fiable. Una carga útil de secuencia de octetos sin procesar (o "RBSP", que es otra forma más formal de referirse a un conjunto de parámetros de imagen) del conjunto de parámetros de imagen incluye parámetros a los que pueden hacer referencia las unidades NAL de fragmento codificado de uno o más componentes de textura o profundidad de uno o más componentes de vista. Una RBSP del conjunto de parámetros de secuencia incluye parámetros a los que puede hacer referencia uno o más conjuntos de parámetros de imagen o uno o más mensajes SEI de período de memoria intermedia.

[0169] Al descodificar un flujo de bits de 3DVC que contiene vistas de profundidad, en cualquier momento dado durante el funcionamiento del proceso de descodificación, el descodificador de vídeo 30 puede activar cero o un conjuntos de parámetros de imagen activos, cero o más conjuntos de parámetros de imagen de textura activos, y cero o más conjuntos de parámetros de imagen de vista activos. Cada conjunto de parámetros de imagen de textura activo está activo específicamente para un valor de índice de orden de vista particular menor o igual al valor de índice de orden de vista máximo, y cada conjunto de parámetros de imagen de vista activo está activo específicamente para un valor de índice de orden de vista particular que es menor que el valor de índice de orden de vista máximo.

[0170] De forma similar, en un momento dado durante el funcionamiento del proceso de descodificación, el descodificador de vídeo 30 puede activar cero o un conjuntos de parámetros de secuencia de MVC activos, cero o más conjuntos de parámetros de secuencia de MVC de textura activos, y cero o más conjuntos de parámetros de secuencia de MVC de vista activos. Cada conjunto de parámetros de secuencia de MVC de textura activo está activo específicamente para un valor de índice de orden de vista particular menor o igual al valor de índice de orden de vista máximo, y cada conjunto de parámetros de secuencia de MVC de vista activo está activo específicamente para un valor de índice de orden de vista particular que es menor que el valor de índice de orden de vista máximo.

[0171] En este sentido, la activación del conjunto de parámetros para 3DVC en general es similar a la activación de los conjuntos de parámetros en MVC, excepto por las diferencias indicadas en el presente documento (que implican casos de cómo tratar componentes de vista de profundidad que en la actualidad no se admiten en MVC). Además, MVC proporciona un número de bits reservados en cada conjunto de parámetros a los que pueden utilizar elementos sintácticos adicionales dirigidos a extensiones que usan MVC, tales como 3DVC, para especificar elementos sintácticos específicos de 3DVC. Proporcionando estos bits reservados, los conjuntos de parámetros mantienen la retrocompatibilidad con MVC, aunque estos bits de reserva se pueden utilizar para especificar elementos sintácticos no relacionados o no especificados en la extensión MVC para la H.264/AVC. Por este motivo, aunque el conjunto de parámetros de MVC puede incluir elementos sintácticos adicionales no relacionados con MVC, estos conjuntos de parámetros se denominan conjuntos de parámetros MVC dados los bits reservados presentes en los conjuntos de parámetros de MVC.

[0172] El proceso de activación detallado para los conjuntos de parámetros de imagen y secuencia es como sigue. El descodificador de vídeo 30 inicialmente considera que cada RBSP del conjunto de parámetros de imagen no está activa al inicio del funcionamiento del proceso de descodificación. A lo sumo, el descodificador de vídeo 30 activa una

RBSP del conjunto de parámetros de imagen como RBSP del conjunto de parámetros de imagen activa en cualquier momento dado durante el funcionamiento del proceso de descodificación. Cuando cualquier RBSP del conjunto de parámetros de imagen particular se convierte en la RBSP del conjunto de parámetros de imagen activa, el descodificador de vídeo 30 desactiva eficazmente la RBSP del conjunto de parámetros de imagen previamente activa (si la hubiera).

[0173] Además de la RBSP del conjunto de parámetros de imagen activa, el descodificador de vídeo 30 puede activar cero o más RBSP del conjunto de parámetros de imagen específicamente para los componentes de textura (con un valor de índice de orden de vista particular ("VOIdx") menor o igual al valor de índice de orden de vista máximo ("VOIdxMax")) que pertenecen a las vistas de salida de destino o a los que se puede hacer referencia a través de la predicción entre vistas en la descodificación de los componentes de textura que pertenecen a las vistas de salida de destino. Una RBSP del conjunto de parámetros de imagen de este tipo se denomina RBSP del conjunto de parámetros de imagen de vista activa para el valor particular de VOIdx. Las restricciones sobre las RBSP del conjunto de parámetros de imagen activas también se aplican a las RBSP del conjunto de parámetros de imagen activas para un valor particular de VOIdx.

[0174] Además, el descodificador de vídeo 30 puede activar cero o más RBSP del conjunto de parámetros de imagen específicamente para los componentes de vista (con un valor particular de VOIdx menor que VOIdxMax) que pertenecen a las vistas de salida de destino o a los que se puede hacer referencia a través de la predicción entre vistas en la descodificación de los componentes de profundidad que pertenecen a las vistas de salida de destino. Una RBSP del conjunto de parámetros de imagen de este tipo se denomina RBSP del conjunto de parámetros de imagen de vista activa para el valor particular de VOIdx. Las restricciones sobre las RBSP del conjunto de parámetros de imagen activas también se aplican a las RBSP del conjunto de parámetros de imagen activas para un valor particular de VOIdx menor que VOIdxMax.

[0175] Cuando una RBSP de un conjunto de parámetros de imagen (con un valor particular de pic_parameter_set_id) no es la RBSP del conjunto de parámetros de imagen activa y se hace referencia a la RBSP del conjunto de parámetros de imagen por una unidad NAL de fragmento codificado que pertenece a un componente de profundidad y con VOIdx igual a VOIdxMax (que usa ese valor de pic_parameter_set_id), el descodificador de vídeo 30 puede activar la RBSP del conjunto de parámetros de imagen. Una unidad NAL de fragmento codificado que pertenece a un componente de profundidad puede tener un nal_unit_type igual a 21. Esta RBSP del conjunto de parámetros de imagen se llama RBSP del conjunto de parámetros de imagen activa hasta que la RBSP del conjunto de parámetros de imagen se desactiva cuando otra RBSP del conjunto de parámetros de imagen se convierte en la RBSP del conjunto de parámetros de imagen activa. Una RBSP del conjunto de parámetros de imagen, con ese valor particular de pic_parameter_set_id, estará disponible para el proceso de descodificación antes de la activación de la RBSP del conjunto de parámetros de imagen.

[0176] Cuando una RBSP de un conjunto de parámetros de imagen (con un valor particular de pic_parameter_set_id) no es el conjunto de parámetros de imagen de vista activo para un valor particular de VOIdx menor que VOIdxMax y se hace referencia a la RBSP del conjunto de parámetros de imagen por una unidad NAL de fragmento codificado que pertenece a un componente de profundidad (es decir, con nal_unit_type igual a 21) y con el valor particular de VOIdx (que usa ese valor de pic_parameter_set_id), el descodificador de vídeo 30 activa la RBSP del conjunto de parámetros de imagen para los componentes de vista con el valor particular de VOIdx. Esta RBSP del conjunto de parámetros de imagen se llama RBSP del conjunto de parámetros de imagen de vista activa para el valor particular de VOIdx hasta que la RBSP del conjunto de parámetros de imagen se desactiva cuando otra RBSP del conjunto de parámetros de imagen se convierte en la RBSP del conjunto de parámetros de imagen de vista activa para el valor particular de VOIdx. Una RBSP del conjunto de parámetros de imagen, con ese valor particular de pic_parameter_set_id, estará disponible para el proceso de descodificación antes de la activación de la RBSP del conjunto de parámetros de imagen.

[0177] Cuando una RBSP de un conjunto de parámetros de imagen (con un valor particular de pic_parameter_set_id) no es el conjunto de parámetros de imagen de textura activo para un valor particular de VOIdx menor o igual a VOIdxMax y se hace referencia a la RBSP del conjunto de parámetros de imagen por una unidad NAL de fragmento codificado que pertenece a un componente de textura (es decir, con nal_unit_type igual a 1, 5 o 20) y con el valor particular de VOIdx (que usa ese valor de pic_parameter_set_id), el descodificador de vídeo 30 activa la RBSP del conjunto de parámetros de imagen para los componentes de profundidad con el valor particular de VOIdx. Esta RBSP del conjunto de parámetros de imagen se llama RBSP del conjunto de parámetros de imagen de textura activa para el valor particular de VOIdx hasta que la RBSP del conjunto de parámetros de imagen se desactiva cuando otra RBSP del conjunto de parámetros de imagen se convierte en la RBSP del conjunto de parámetros de imagen de textura activa para el valor particular de VOIdx. Una RBSP del conjunto de parámetros de imagen, con ese valor particular de pic_parameter_set_id, estará disponible para el proceso de descodificación antes de la activación de la RBSP del conjunto de parámetros de imagen.

[0178] Cualquier unidad NAL del conjunto de parámetros de imagen que contenga el valor de pic_parameter_set_id para la RBSP del conjunto de parámetros de imagen activa para una imagen codificada tendrá el mismo contenido que el de la RBSP del conjunto de parámetros de imagen activa para esta imagen codificada a menos que la unidad NAL del conjunto de parámetros de imagen siga a la última unidad VCL NAL de esta imagen codificada y preceda a

la primera unidad VCL NAL de otra imagen codificada. Cualquier unidad NAL del conjunto de parámetros de imagen que contenga el valor de `pic_parameter_set_id` para la RBSP del conjunto de parámetros de imagen activa para un valor particular de `VOIdx` menor que `VOIdxMax` para una imagen codificada tendrá el mismo contenido que el de la RBSP del conjunto de parámetros de imagen activa para el valor particular de `VOIdx` para esta imagen codificada a menos que la unidad NAL del conjunto de parámetros de imagen siga a la última unidad VCL NAL de esta imagen codificada y preceda a la primera unidad VCL NAL de otra imagen codificada. Cualquier unidad NAL del conjunto de parámetros de imagen que contenga el valor de `pic_parameter_set_id` para la RBSP del conjunto de parámetros de imagen de textura activa para un valor particular de `VOIdx` para una imagen codificada tendrá el mismo contenido que el de la RBSP del conjunto de parámetros de imagen de textura activa para el valor particular de `VOIdx` para esta imagen codificada a menos que la unidad NAL del conjunto de parámetros de imagen siga a la última unidad VCL NAL de esta imagen codificada y preceda a la primera unidad VCL NAL de otra imagen codificada.

[0179] El descodificador de vídeo 30 determina inicialmente que cada RBSP del conjunto de parámetros de secuencia de MVC no esté activa al inicio del funcionamiento del proceso de descodificación. De nuevo, la referencia a los conjuntos de parámetros de secuencia de MVC se refiere a los conjuntos de parámetros de secuencia de MVC dentro del contexto de 3DVC, que hereda o de otro modo adopta MVC y, por tanto, incorpora conjuntos de parámetros de secuencia MVC (aunque estos conjuntos de parámetros se pueden modificar añadiendo elementos sintácticos específicos de 3DVC adicionales). A lo sumo, el descodificador de vídeo 30 determina que una RBSP del conjunto de parámetros de secuencia de MVC es la RBSP del conjunto de parámetros de secuencia de MVC activa en cualquier momento dado durante el funcionamiento del proceso de descodificación. Cuando cualquier RBSP del conjunto de parámetros de secuencia de MVC particular se convierte en la RBSP del conjunto de parámetros de secuencia de MVC activa, el descodificador de vídeo 30 desactiva la RBSP del conjunto de parámetros de secuencia de MVC previamente activa (si la hubiera).

[0180] La RBSP del conjunto de parámetros de secuencia de MVC activa es una RBSP de un conjunto de parámetros de secuencia o bien una RBSP de un conjunto de parámetros de secuencia de un subconjunto. Las RBSP del conjunto de parámetros de secuencia se activan por unidades NAL de fragmento codificado con `nal_unit_type` igual a 1 o 5 o mensajes SEI de período de almacenamiento en memoria intermedia que no están incluidos en un mensaje SEI de anidamiento escalable de MVC o un mensaje SEI de anidamiento escalable de 3DV. Los conjuntos de parámetros de secuencia de subconjunto se activan por unidades NAL de extensión MVC de fragmento codificado (`nal_unit_type` igual a 21) o mensajes SEI de período de almacenamiento en memoria intermedia que se incluyen en un mensaje SEI de anidamiento escalable de 3DV. Una RBSP de un conjunto de parámetros de secuencia y una RBSP de un conjunto de parámetros de secuencia de subconjunto pueden tener el mismo valor de `seq_parameter_set_id`.

[0181] Además de la RBSP del conjunto de parámetros de secuencia de MVC activa, el descodificador de vídeo 30 puede activar cero o más RBSP del conjunto de parámetros de secuencia de MVC específicamente para los componentes de vista (con un valor particular de `VOIdx` menor que `VOIdxMax`) que pertenecen a las vistas de salida de destino o a los que se puede hacer referencia a través de la predicción entre vistas en la descodificación de componentes de vista que pertenecen a las vistas de salida de destino. Una RBSP de un conjunto de parámetros de secuencia de MVC de este tipo se denomina RBSP del conjunto de parámetros de secuencia de MVC de vista activa para el valor particular de `VOIdx`. Las restricciones sobre las RBSP del conjunto de parámetros de la secuencia de MVC activas también se aplican a las RBSP del conjunto de parámetros de secuencia de MVC activas para un valor particular de `VOIdx` menor que `VOIdxMax`.

[0182] Además, el descodificador de vídeo 30 puede activar cero o más RBSP del conjunto de parámetros de secuencia de MVC específicamente para los componentes de textura (con un valor particular de `VOIdx` menor o igual a `VOIdxMax`) que pertenecen a las vistas de salida de destino o a los que se puede hacer referencia a través de la predicción entre vistas en la descodificación de componentes de textura que pertenecen a las vistas de salida de destino. Una RBSP de un conjunto de parámetros de secuencia de MVC de este tipo se denomina RBSP del conjunto de parámetros de secuencia de MVC de textura activa para el valor particular de `VOIdx`. Las restricciones en las RBSP del conjunto de parámetros de secuencia de MVC activas también se aplican a las RBSP del conjunto de parámetros de secuencia de MVC de textura activas para un valor particular de `VOIdx`.

[0183] Para la siguiente descripción, el mensaje SEI del período de almacenamiento en memoria intermedia de activación se especifica como sigue.

- Si `VOIdxMax` es igual a `VOIdxMin` y la unidad de acceso contiene un mensaje SEI de período de almacenamiento en memoria intermedia no incluido en un mensaje SEI de anidamiento escalable de MVC y no incluido en un mensaje SEI de anidamiento escalable de 3DVC, este mensaje SEI de período de almacenamiento en memoria intermedia es el mensaje SEI de período de almacenamiento en memoria intermedia de activación.
- De otro modo, si `VOIdxMax` no es igual a `VOIdxMin` y la unidad de acceso contiene un mensaje SEI de período de almacenamiento en memoria intermedia incluido en un mensaje SEI de anidamiento escalable de 3DVC y asociado con el punto de funcionamiento que se está descodificando, este mensaje SEI de período de almacenamiento en memoria intermedia es el mensaje SEI de período de almacenamiento en memoria intermedia de activación.

- De otro modo, la unidad de acceso no contiene un mensaje SEI de período de almacenamiento en memoria intermedia de activación.

[0184] Cuando una RBSP de un conjunto de parámetros de secuencia (`nal_unit_type` es igual a 7) con un valor particular de `seq_parameter_set_id` ya no es la RBSP del conjunto de parámetros de secuencia de MVC activa y se hace referencia a la RBSP del conjunto de parámetros de secuencia por la activación de una RBSP de un conjunto de parámetros de imagen (que usa ese valor de `seq_parameter_set_id`) y la RBSP del conjunto de parámetros de imagen se activa por una unidad NAL de fragmento codificado con `nal_unit_type` igual a 1 o 5 (la RBSP del conjunto de parámetros de imagen se convierte en la RBSP del conjunto de parámetros de imagen activa y `VOIdxMax` es igual a `VOIdxMin` y no existe ningún componente de profundidad en la unidad de acceso) y la unidad de acceso no contiene ningún mensaje SEI de período de almacenamiento en memoria intermedia de activación, el descodificador de vídeo 30 puede activar la RBSP del conjunto de parámetros de secuencia. Esta RBSP del conjunto de parámetros de secuencia se llama RBSP del conjunto de parámetros de secuencia de MVC activa hasta que la RBSP del conjunto de parámetros de secuencia se desactiva cuando otra RBSP del conjunto de parámetros de secuencia de MVC se convierte en la RBSP del conjunto de parámetros de secuencia de MVC activa. Una RBSP de un conjunto de parámetros de secuencia, con ese valor particular de `seq_parameter_set_id`, estará disponible para el proceso de descodificación antes de la activación de la RBSP del conjunto de parámetros de secuencia.

[0185] Cuando una RBSP de un conjunto de parámetros de secuencia (`nal_unit_type` es igual a 7) con un valor particular de `seq_parameter_set_id` ya no es la RBSP del conjunto de parámetros de secuencia de MVC activa y se hace referencia a la RBSP del conjunto de parámetros de secuencia por un mensaje SEI de período de almacenamiento en memoria intermedia de activación (que usa ese valor de `seq_parameter_set_id`) que no está incluido en un mensaje SEI de anidamiento escalable de 3DV y `VOIdxMax` es igual a `VOIdxMin` y no existe ningún componente de profundidad en la unidad de acceso, el descodificador de vídeo 30 activa la RBSP del conjunto de parámetros de secuencia. Esta RBSP del conjunto de parámetros de secuencia se llama RBSP del conjunto de parámetros de secuencia de MVC activa hasta que la RBSP del conjunto de parámetros de secuencia se desactiva cuando otra RBSP del conjunto de parámetros de secuencia de MVC se convierte en la RBSP del conjunto de parámetros de secuencia de MVC activa. Una RBSP de un conjunto de parámetros de secuencia, con ese valor particular de `seq_parameter_set_id`, estará disponible para el proceso de descodificación antes de la activación de la RBSP del conjunto de parámetros de secuencia.

[0186] Cuando una RBSP de un conjunto de parámetros de secuencia de subconjunto (`nal_unit_type` es igual a 15) con un valor particular de `seq_parameter_set_id` ya no es la RBSP del conjunto de parámetros de secuencia de MVC activa y se hace referencia a la RBSP del conjunto de parámetros de secuencia de subconjunto por la activación de una RBSP de un conjunto de parámetros de imagen (que usa ese valor de `seq_parameter_set_id`) y la RBSP del conjunto de parámetros de imagen se activa por una unidad NAL de fragmento codificado con `nal_unit_type` igual a 21 y con `VOIdx` igual a `VOIdxMax` (la RBSP del conjunto de parámetros de imagen se convierte en la RBSP del conjunto de parámetros de imagen activa) y la unidad de acceso no contiene un mensaje SEI de período de almacenamiento en memoria intermedia de activación, el descodificador de vídeo 30 puede activar la RBSP del conjunto de parámetros de secuencia de subconjunto. Esta RBSP del conjunto de parámetros de secuencia de subconjunto se llama RBSP del conjunto de parámetros de secuencia de MVC activa hasta que la RBSP del conjunto de parámetros de secuencia de subconjunto se desactiva cuando otra RBSP del conjunto de parámetros de secuencia de MVC se convierte en la RBSP del conjunto de parámetros de secuencia de subconjunto, con ese valor particular de `seq_parameter_set_id`, estará disponible para el proceso de descodificación antes de su activación.

[0187] Cuando una RBSP de un conjunto de parámetros de secuencia de subconjunto (`nal_unit_type` es igual a 15) con un valor particular de `seq_parameter_set_id` ya no es la RBSP del conjunto de parámetros de secuencia de MVC activa y se hace referencia a la RBSP del conjunto de parámetros de secuencia de subconjunto por un mensaje SEI de período de almacenamiento en memoria intermedia de activación (que usa ese valor de `seq_parameter_set_id`) que está incluido en un mensaje SEI de anidamiento escalable de 3DV, el descodificador de vídeo 30 activa la RBSP del conjunto de parámetros de secuencia de subconjunto. Esta RBSP del conjunto de parámetros de secuencia de subconjunto se llama RBSP del conjunto de parámetros de secuencia de MVC activa hasta que la RBSP del conjunto de parámetros de secuencia de subconjunto se desactiva cuando otra RBSP del conjunto de parámetros de secuencia de MVC se convierte en la RBSP del conjunto de parámetros de secuencia de MVC activa. Una RBSP de un conjunto de parámetros de secuencia de subconjunto, con ese valor particular de `seq_parameter_set_id`, estará disponible para el proceso de descodificación antes de la activación de la RBSP del conjunto de parámetros de secuencia de subconjunto.

[0188] Para la siguiente descripción, el mensaje SEI de período de almacenamiento en memoria intermedia de textura de activación para un valor particular de `VOIdx` se especifica como sigue.

- Si la unidad de acceso contiene uno o más de un mensaje SEI de período de almacenamiento en memoria intermedia incluido en un mensaje SEI de anidamiento escalable de MVC y asociado con un punto de funcionamiento para el que el mayor `VOIdx` en el subconjunto de flujo de bits asociado es igual al valor particular de `VOIdx`, el primero de estos mensajes SEI de período de almacenamiento en memoria intermedia, en orden de

descodificación, es el mensaje SEI del período de almacenamiento en memoria intermedia de textura de activación para el valor particular de VOIdx.

- 5 - De otro modo, si la unidad de acceso contiene un mensaje SEI de período de almacenamiento en memoria intermedia no incluido en un mensaje SEI de anidamiento escalable de MVC o un mensaje SEI de anidamiento escalable de 3DV, este mensaje SEI de período de almacenamiento en memoria intermedia es el mensaje SEI de período de almacenamiento en memoria intermedia de textura de activación para el valor particular de VOIdx igual a VOIdxMin.
- 10 - De otro modo, la unidad de acceso no contiene un mensaje SEI de período de almacenamiento en memoria intermedia de textura de activación para el valor particular de VOIdx.

15 **[0189]** Cuando una RBSP de un conjunto de parámetros de secuencia (nal_unit_type es igual a 7) con un valor particular de seq_parameter_set_id ya no es la RBSP del conjunto de parámetros de secuencia de MVC de textura activa para VOIdx igual a VOIdxMin y se hace referencia a la RBSP del conjunto de parámetros de secuencia por la activación de una RBSP de un conjunto de parámetros de imagen (que usa ese valor de seq_parameter_set_id) y la RBSP del conjunto de parámetros de imagen se activa por una unidad NAL de fragmento codificado con nal_unit_type igual a 1 o 5 (la RBSP del conjunto de parámetros de imagen se convierte en la RBSP del conjunto de parámetros de imagen de textura activa para VOIdx igual a VOIdxMin), el descodificador de vídeo 30 puede activar la RBSP del conjunto de parámetros de secuencia para los componentes de textura con VOIdx igual a VOIdxMin. Esta RBSP del conjunto de parámetros de secuencia se llama RBSP del conjunto de parámetros de secuencia de MVC de textura activa para VOIdx igual a VOIdxMin hasta que la RBSP del conjunto de parámetros de secuencia se desactiva cuando otra RBSP del conjunto de parámetros de secuencia de MVC se convierte en la RBSP del conjunto de parámetros de secuencia de MVC de textura activa para VOIdx igual a VOIdxMin. Una RBSP de un conjunto de parámetros de secuencia, con ese valor particular de seq_parameter_set_id, estará disponible para el proceso de descodificación antes de la activación de la RBSP del conjunto de parámetros de secuencia.

30 **[0190]** Cuando una RBSP de un conjunto de parámetros de secuencia (nal_unit_type es igual a 7) con un valor particular de seq_parameter_set_id ya no es la RBSP del conjunto de parámetros de secuencia de MVC de textura activa para VOIdx igual a VOIdxMin y se hace referencia a la RBSP del conjunto de parámetros de secuencia por un mensaje SEI de período de almacenamiento en memoria intermedia de textura de activación (que usa ese valor de seq_parameter_set_id) que no está incluido en un mensaje SEI de anidamiento escalable de MVC o un mensaje SEI de anidamiento escalable de 3DV, el descodificador de vídeo 30 puede activar la RBSP del conjunto de parámetros de secuencia para los componentes de textura con VOIdx igual a VOIdxMin. Esta RBSP del conjunto de parámetros de secuencia se llama RBSP del conjunto de parámetros de secuencia de MVC de textura activa para VOIdx igual a VOIdxMin hasta que la RBSP del conjunto de parámetros de secuencia se desactiva cuando otra RBSP del conjunto de parámetros de secuencia de MVC se convierte en la RBSP del conjunto de parámetros de secuencia de MVC de textura activa para VOIdx igual a. Una RBSP de un conjunto de parámetros de secuencia, con ese valor particular de seq_parameter_set_id, estará disponible para el proceso de descodificación antes de la activación de la RBSP del conjunto de parámetros de secuencia.

45 **[0191]** Cuando una RBSP de un conjunto de parámetros de secuencia de subconjunto (nal_unit_type es igual a 15) con un valor particular de seq_parameter_set_id ya no es la RBSP del conjunto de parámetros de secuencia de MVC de textura activa para un valor particular de VOIdx menor o igual a VOIdxMax y se hace referencia a la RBSP del conjunto de parámetros de secuencia de subconjunto por activación de una RBSP de un conjunto de parámetros de imagen (que usa ese valor de seq_parameter_set_id) y la RBSP del conjunto de parámetros de imagen se activa por una unidad NAL de extensión MVC de fragmento codificado (nal_unit_type igual a 20) con el valor particular de VOIdx (la RBSP del conjunto de parámetros de imagen se convierte en la RBSP del conjunto de parámetros de imagen de textura activa para el valor particular de VOIdx), el descodificador de vídeo 30 puede activar la RBSP del conjunto de parámetros de secuencia de subconjunto para los componentes de textura con el valor particular de VOIdx. Esta RBSP del conjunto de parámetros de secuencia de subconjunto se llama RBSP del conjunto de parámetros de secuencia de MVC de textura activa para el valor particular de VOIdx hasta que la RBSP del conjunto de parámetros de secuencia de subconjunto se desactiva cuando otra RBSP del conjunto de parámetros de secuencia de MVC se convierte en la RBSP del conjunto de parámetros de secuencia de MVC de textura activa para el valor particular de VOIdx. Una RBSP de un conjunto de parámetros de secuencia de subconjunto, con ese valor particular de seq_parameter_set_id, estará disponible para el proceso de descodificación antes de la activación de la RBSP del conjunto de parámetros de secuencia de subconjunto.

60 **[0192]** Cuando una RBSP de un conjunto de parámetros de secuencia de subconjunto (nal_unit_type es igual a 15) con un valor particular de seq_parameter_set_id ya no es la RBSP del conjunto de parámetros de secuencia de MVC de textura activa para un valor particular de VOIdx menor o igual a VOIdxMax y se hace referencia a la RBSP del conjunto de parámetros de secuencia de subconjunto por un mensaje SEI de período de almacenamiento en memoria intermedia de textura de activación (que usa ese valor de seq_parameter_set_id) que está incluido en un mensaje SEI de anidamiento escalable de MVC y se asocia con el valor particular de VOIdx, el descodificador de vídeo 30 puede activar esta RBSP del conjunto de parámetros de secuencia de subconjunto para los componentes de textura con el valor particular de VOIdx. Esta RBSP del conjunto de parámetros de secuencia de subconjunto se llama RBSP del

conjunto de parámetros de secuencia de MVC de textura activa para el valor particular de VOIdx hasta que la RBSP del conjunto de parámetros de secuencia de subconjunto se desactiva cuando otra RBSP del conjunto de parámetros de secuencia de MVC se convierte en la RBSP del conjunto de parámetros de secuencia de MVC de textura activa para el valor particular de VOIdx. Una RBSP de un conjunto de parámetros de secuencia de subconjunto, con ese valor particular de seq_parameter_set_id, estará disponible para el proceso de descodificación antes de la activación de la RBSP del conjunto de parámetros de secuencia de subconjunto.

[0193] Para la siguiente especificación, el mensaje SEI de período de almacenamiento en memoria intermedia de vista de activación para un valor particular de VOIdx se especifica como sigue.

- Si la unidad de acceso contiene uno o más de un mensaje SEI de período de almacenamiento en memoria intermedia incluido en un mensaje SEI de anidamiento ajustable a escala de 3DVC y asociado a un punto de funcionamiento para el que el mayor VOIdx en el subconjunto de flujo de bits asociado es igual al valor particular de VOIdx, el primero de estos mensajes SEI de período de almacenamiento en memoria intermedia, en orden de descodificación, es el mensaje SEI de período de almacenamiento en memoria intermedia de vista de activación para el valor particular de VOIdx.
- De otro modo, la unidad de acceso no contiene un mensaje SEI de período de almacenamiento en memoria intermedia de vista de activación para el valor particular de VOIdx.

[0194] Cuando una RBSP de un conjunto de parámetros de secuencia de subconjunto (nal_unit_type es igual a 15) con un valor particular de seq_parameter_set_id ya no es la RBSP del conjunto de parámetros de secuencia de MVC de vista activa para un valor particular de VOIdx menor que VOIdxMax y se hace referencia a la RBSP del conjunto de parámetros de secuencia de subconjunto por la activación de una RBSP de un conjunto de parámetros de imagen (que usa ese valor de seq_parameter_set_id) y la RBSP del conjunto de parámetros de imagen se activa por una unidad NAL de fragmento codificado con nal_unit_type igual a 21 y con el valor particular de VOIdx (la RBSP del conjunto de parámetros de imagen se convierte en la RBSP del conjunto de parámetros de imagen de vista activa para el valor particular de VOIdx), el descodificador de vídeo 30 activa la RBSP del conjunto de parámetros de secuencia de subconjunto para los componentes de vista con el valor particular de VOIdx. Esta RBSP del conjunto de parámetros de secuencia de subconjunto se llama RBSP del conjunto de parámetros de secuencia de MVC de vista activa para el valor particular de VOIdx hasta que la RBSP del conjunto de parámetros de secuencia de subconjunto se desactiva cuando otra RBSP del conjunto de parámetros de secuencia de MVC se convierte en la RBSP del conjunto de parámetros de secuencia de MVC de vista activa para el valor particular de VOIdx o al descodificar una unidad de acceso con VOIdxMax menor o igual al valor particular de VOIdx. Una RBSP de un conjunto de parámetros de secuencia de subconjunto, con ese valor particular de seq_parameter_set_id, estará disponible para el proceso de descodificación antes de la activación de la RBSP del conjunto de parámetros de secuencia de subconjunto.

[0195] Cuando una RBSP de un conjunto de parámetros de secuencia de subconjunto (nal_unit_type es igual a 15) con un valor particular de seq_parameter_set_id ya no es la RBSP del conjunto de parámetros de secuencia de MVC de vista activa para un valor particular de VOIdx menor que VOIdxMax y se hace referencia a la RBSP del conjunto de parámetros de secuencia de subconjunto por un mensaje SEI de período de almacenamiento en memoria intermedia de vista de activación (que usa ese valor de seq_parameter_set_id) que está incluido en un mensaje SEI de anidamiento escalable de 3DV y se asocia con el valor particular de VOIdx, el descodificador de vídeo 30 activa esta RBSP del conjunto de parámetros de secuencia de subconjunto para los componentes de vista con el valor particular de VOIdx. Esta RBSP del conjunto de parámetros de secuencia de subconjunto se llama RBSP del conjunto de parámetros de secuencia de MVC de vista activa para el valor particular de VOIdx hasta que la RBSP del conjunto de parámetros de secuencia de subconjunto se desactiva cuando otra RBSP del conjunto de parámetros de secuencia de MVC se convierte en la RBSP del conjunto de parámetros de secuencia de MVC de vista activa para el valor particular de VOIdx o al descodificar una unidad de acceso con VOIdxMax menor o igual al valor particular de VOIdx. Una RBSP de un conjunto de parámetros de secuencia de subconjunto, con ese valor particular de seq_parameter_set_id, estará disponible para el proceso de descodificación antes de la activación de la RBSP del conjunto de parámetros de secuencia de subconjunto.

[0196] No se hará referencia a una RBSP de un conjunto de parámetros de secuencia de MVC que incluye un valor de profile_idc no especificado en el Anexo A o el Anexo H o el Anexo I propuesto para la H.264/AVC por la activación de una RBSP de un conjunto de parámetros de imagen como la RBSP del conjunto de parámetros de imagen activa o como RBSP del conjunto de parámetros de imagen de vista activa o como RBSP del conjunto de parámetros de imagen de textura activa (que usa ese valor de seq_parameter_set_id) o a la que se hace referencia por un mensaje SEI de período de almacenamiento de memoria intermedia (que usa ese valor de seq_parameter_set_id). Una RBSP de un conjunto de parámetros de secuencia de MVC que incluye un valor de profile_idc no especificado en el Anexo A o el Anexo H o el Anexo I propuesto para la H.264/AVC se ignora en la descodificación para los perfiles especificados en el Anexo A o el Anexo H o el Anexo I propuesto para la H.264/AVC.

[0197] Es un requisito de conformidad del flujo de bits que se cumplan las siguientes restricciones:

- Para cada valor particular de VOIdx, todas las unidades NAL de fragmento codificado (con nal_unit_type igual a 1, 5, 20 o 21) de una secuencia de vídeo codificada se referirán al mismo valor de seq_parameter_set_id (por medio de la RBSP del conjunto de parámetros de imagen a la que se hace referencia por el valor de pic_parameter_set_id).
- El valor de seq_parameter_set_id en un mensaje SEI de período de almacenamiento en memoria intermedia que no esté incluido en un mensaje SEI de anidamiento escalable de MVC será idéntico al valor de seq_parameter_set_id en la RBSP del conjunto de parámetros de imagen a la que se hace referencia por las unidades NAL de fragmento codificado con nal_unit_type igual a 1 o 5 (por medio del valor de pic_parameter_set_id) en la misma unidad de acceso.
- El valor de seq_parameter_set_id en un mensaje SEI de período de almacenamiento en memoria intermedia que esté incluido en un mensaje SEI de anidamiento escalable de MVC y esté asociado con un valor particular de VOIdx será idéntico al valor de seq_parameter_set_id en la RBSP del conjunto de parámetros de imagen a la que se hace referencia por las unidades NAL de fragmento codificado con nal_unit_type igual a 1, 5 o 20 con el valor particular de VOIdx (por medio del valor de pic_parameter_set_id) en la misma unidad de acceso.
- El valor de seq_parameter_set_id en un mensaje SEI de período de almacenamiento en memoria intermedia que esté incluido en un mensaje SEI de anidamiento escalable de 3DVC y esté asociado con un valor particular de VOIdx será idéntico al valor de seq_parameter_set_id en la RBSP del conjunto de parámetros de imagen a la que se hace referencia por las unidades NAL de fragmento codificado con nal_unit_type igual a 21 con el valor particular de VOIdx (por medio del valor de pic_parameter_set_id) en la misma unidad de acceso.

[0198] Las RBSP del conjunto de parámetros de secuencia de MVC de vista activas para diferentes valores de VOIdx pueden ser la misma RBSP del conjunto de parámetros de secuencia de MVC. La RBSP del conjunto de parámetros de secuencia de MVC activa y una RBSP de un conjunto de parámetros de secuencia de MVC de vista activa para un valor particular de VOIdx pueden ser la misma RBSP del conjunto de parámetros de secuencia de MVC.

[0199] Las RBSP del conjunto de parámetros de secuencia de MVC de textura activas para diferentes valores de VOIdx pueden ser la misma RBSP del conjunto de parámetros de secuencia de MVC. La RBSP del conjunto de parámetros de secuencia de MVC activa y una RBSP de un conjunto de parámetros de secuencia de MVC de textura activa para un valor particular de VOIdx pueden ser la misma RBSP de un conjunto de parámetros de secuencia de MVC. Una RBSP de un conjunto de parámetros de secuencia de MVC de vista activa para un valor particular de VOIdx y una RBSP de un conjunto de parámetros de secuencia de MVC de textura activa para un valor particular de VOIdx pueden ser la misma RBSP de un conjunto de parámetros de secuencia de MVC.

[0200] Cuando la RBSP del conjunto de parámetros de secuencia de MVC activa para una imagen codificada es una RBSP de un conjunto de parámetros de secuencia, cualquier RBSP del conjunto de parámetros de secuencia en la secuencia de vídeo codificada que contenga esta imagen codificada y con el valor de id de seq_parameter_set para la RBSP del conjunto de parámetros de secuencia de MVC activa tendrá el mismo contenido que el de la RBSP del conjunto de parámetros de secuencia de MVC activa.

[0201] Cuando la RBSP del conjunto de parámetros de secuencia de MVC activa para una imagen codificada sea una RBSP de un conjunto de parámetros de secuencia de subconjunto, cualquier RBSP del conjunto de parámetros de secuencia de subconjunto en la secuencia de vídeo codificada que contenga esta imagen codificada y con el valor de seq_parameter_set_id para la RBSP del conjunto de parámetros de secuencia de MVC activa tendrá el mismo contenido que el de la RBSP del conjunto de parámetros de secuencia de MVC activa.

[0202] Para cada valor particular de VOIdx, se aplica lo siguiente:

Cuando la RBSP del conjunto de parámetros de secuencia de MVC de textura activa para una imagen codificada sea una RBSP de un conjunto de parámetros de secuencia, cualquier RBSP del conjunto de parámetros de secuencia en la secuencia de vídeo codificada que contenga esta imagen codificada y con el valor de id de seq_parameter_set para la RBSP del conjunto de parámetros de secuencia de MVC de textura activa tendrá el mismo contenido que el de la RBSP del conjunto de parámetros de secuencia de MVC de textura activa.

[0203] Cuando la RBSP del conjunto de parámetros de secuencia de MVC de textura activa para una imagen codificada sea una RBSP de un conjunto de parámetros de secuencia de subconjunto, cualquier RBSP del conjunto de parámetros de secuencia de subconjunto en la secuencia de vídeo codificada que contiene esta imagen codificada y con el valor de seq_parameter_set_id para la RBSP del conjunto de parámetros de secuencia de MVC de textura activa tendrá el mismo contenido que el de la RBSP del conjunto de parámetros de secuencia de MVC de textura activa.

[0204] La RBSP del conjunto de parámetros de secuencia MVC de vista activa para una imagen codificada es una RBSP de un conjunto de parámetros de secuencia de subconjunto, y cualquier RBSP del conjunto de parámetros de secuencia de subconjunto en la secuencia de vídeo codificada que contiene esta imagen codificada y con el valor de

seq_parameter_set_id para la RBSP del conjunto de parámetros de secuencia de MVC de vista activa tendrá el mismo contenido que el de la RBSP del conjunto de parámetros de secuencia de MVC de vista activa.

[0205] Si las RBSP del conjunto de parámetros de imagen o las RBSP del conjunto de parámetros de secuencia de MVC se transmiten dentro del flujo de bits, estas restricciones imponen una restricción de orden en las unidades NAL que contienen las RBSP del conjunto de parámetros de imagen o las RBSP del conjunto de parámetros de secuencia de MVC, respectivamente. De otro modo (las RBSP del conjunto de parámetros de imagen o las RBSP del conjunto de parámetros de secuencia de MVC se transmiten por otros medios no especificados en esta Recomendación | Norma Internacional para la norma de codificación de vídeo H.264), deben estar disponibles para el proceso de decodificación de manera oportuna de modo que estas restricciones se obedezcan.

[0206] Cuando está presente, una RBSP de extensión de un conjunto de parámetros de secuencia incluye parámetros que tienen una función similar a las de una RBSP de un conjunto de parámetros de secuencia. Para los propósitos de establecimiento de restricciones en los elementos sintácticos de la RBSP de extensión del conjunto de parámetros de secuencia y para los propósitos de determinación de la activación de una RBSP de extensión de un conjunto de parámetros de secuencia, la RBSP de extensión del conjunto de parámetros de secuencia se considerará parte de la RBSP del conjunto de parámetros de secuencia anterior con el mismo valor de seq_parameter_set_id. Cuando está presente una RBSP de un conjunto de parámetros de secuencia que no está seguida de una RBSP de extensión de un conjunto de parámetros de secuencia con el mismo valor de seq_parameter_set_id antes de la activación de la RBSP del conjunto de parámetros de secuencia, la RBSP de extensión del conjunto de parámetros de secuencia y sus elementos sintácticos no se considerarán presentes para la RBSP del conjunto de parámetros de secuencia de MVC activa. El contenido de las RBSP de extensión del conjunto de parámetros de secuencia solo se aplica cuando se decodifica la vista de textura base, que se ajusta a uno o más de los perfiles especificados en el Anexo A, de una secuencia de vídeo codificada que se ajusta a uno o más perfiles especificados en el Anexo I. Las RBSP del conjunto de parámetros de secuencia de subconjunto no deben ir seguidas de una RBSP de extensión de un conjunto de parámetros de secuencia.

[0207] Las RBSP de extensión de conjuntos de parámetros de secuencia no se consideran parte de una RBSP de un conjunto de parámetros de secuencia de subconjunto y las RBSP del conjunto de parámetros de secuencia de subconjunto no deben ir seguidas de una RBSP de extensión de un conjunto de parámetros de secuencia.

[0208] Para los componentes de vista con VOIdx igual a VOIdxMax, todas las restricciones que se expresan en la relación entre los valores de los elementos sintácticos (y los valores de las variables derivadas de esos elementos sintácticos) en conjuntos de parámetros de secuencia de MVC y conjuntos de parámetros de imagen y otros elementos sintácticos son expresiones de restricciones que se aplican solo al conjunto de parámetros de secuencia de MVC activo y al conjunto de parámetros de imagen activo. Para los componentes de vista con un valor particular de VOIdx menor que VOIdxMax, todas las restricciones que se expresan en la relación entre los valores de los elementos sintácticos (y los valores de las variables derivadas de esos elementos sintácticos) en los conjuntos de parámetros de secuencia de MVC y los conjuntos de parámetros de imagen y otros elementos sintácticos son expresiones de restricciones que se aplican solo al conjunto de parámetros de secuencia de MVC de vista activo y al conjunto de parámetros de imagen de vista activo para el valor particular de VOIdx.

[0209] Si está presente alguna RBSP de un conjunto de parámetros de secuencia de MVC que tiene profile_idc igual al valor de uno de los valores de profile_idc especificados en el Anexo A, el Anexo H o el Anexo I propuesto, que nunca se activa en el flujo de bits (es decir, la RBSP del conjunto de parámetros de secuencia de MVC nunca se convierte en el conjunto de parámetros de secuencia de MVC activo o en un conjunto de parámetros de secuencia de MVC de vista activo), los elementos sintácticos de la RBSP del conjunto de parámetros de secuencia de MVC tendrán valores que se ajustarían a las restricciones especificadas si la RBSP del conjunto de parámetros de secuencia de MVC se activara por referencia en un flujo de bits que de otro modo se ajustaría. Si está presente alguna RBSP del conjunto de parámetros de imagen que nunca se activa en el flujo de bits (es decir, la RBSP del conjunto de parámetros de imagen nunca se convierte en el conjunto de parámetros de imagen activo o en un conjunto de parámetros de imagen de vista activo), los elementos sintácticos de la RBSP del conjunto de parámetros de imagen tendrán valores que se ajustarían a las restricciones especificadas si la RBSP del conjunto de parámetros de imagen se activara por referencia en un flujo de bits que de otro modo se ajustaría.

[0210] Durante el funcionamiento del proceso de decodificación, para los componentes de vista con VOIdx igual a VOIdxMax, los valores de los parámetros del conjunto de parámetros de imagen activo y del conjunto de parámetros de secuencia de MVC activo se considerarán efectivos (es decir, en un ejemplo, se aplican cuando se decodifican). Para los componentes de vista con un valor particular de VOIdx menor que VOIdxMax, los valores de los parámetros del conjunto de parámetros de imagen de vista activo y del conjunto de parámetros de secuencia de MVC de vista activo para el valor particular de VOIdx se considerarán efectivos. Para la interpretación de los mensajes SEI que se aplican a toda la unidad de acceso o al componente de vista con VOIdx igual a VOIdxMax, los valores de los parámetros del conjunto de parámetros de imagen activo y del conjunto de parámetros de secuencia de MVC activo para la misma unidad de acceso se considerarán efectivos a menos que se especifique de otro modo en la semántica del mensaje SEI. Para la interpretación de los mensajes SEI que se aplican a los componentes de vista con un valor particular de VOIdx menor que VOIdxMax, los valores de los parámetros del conjunto de parámetros de imagen de

vista activo y del conjunto de parámetros de secuencia de MVC de vista activo para el valor particular de VOIdx para la misma unidad de acceso se considerarán efectivos a menos que se especifique de otro modo en la semántica del mensaje SEI.

5 **[0211]** Para cualquier MVC SPS activo o MVC SPS de vista activo, parte de los elementos sintácticos en la extensión MVC SPS se aplica solo a las vistas de profundidad que se refieren a este SPS, mientras que algunas de otras partes de los elementos sintácticos en la extensión MVC SPS se aplican conjuntamente tanto a las vistas de profundidad que se refieren a este SPS como a las vistas de textura correspondientes. Más específicamente, la información de dependencia de vista de la extensión MVC SPS se aplica solo a las vistas de profundidad, y las definiciones de nivel se aplican conjuntamente a los puntos de funcionamiento, cada uno de los cuales contiene tanto vistas de profundidad como sus vistas de textura correspondientes.

15 **[0212]** En algunos casos, las técnicas de la presente divulgación pueden permitir que el descodificador de vídeo 30 active, durante la codificación de un flujo de bits que contiene vistas de profundidad, un conjunto de parámetros de imagen (PPS) de textura o un conjunto de parámetros de secuencia (SPS) de MVC de textura para un valor de índice de orden de vista.

20 **[0213]** En algunos casos, el descodificador de vídeo 30 puede activar, durante la codificación del flujo de bits, un PPS activo, activar, durante la codificación del flujo de bits, un PPS de vista activo para el valor de índice de orden de vista, activar, durante la codificación del flujo de bits, un MVC SPS de vista activo para el valor de índice de orden de vista, y activar, durante la codificación del flujo de bits, un MVC SPS activo.

25 **[0214]** En algunos casos, la activación del PPS de textura puede implicar la activación, para los componentes de profundidad que especifican el valor de índice de orden de vista, de una carga útil de secuencia de octetos sin procesar (RBSP) de PPS cuando la PPS RBSP no es un PPS de textura activo para el valor de índice de orden de vista y se hace referencia a la PPS RBSP por una unidad de capa de abstracción de red (NAL) de fragmento codificado que tanto pertenece a un componente de textura como que especifica el valor de índice de orden de vista, activando la PPS RBSP. El descodificador de vídeo 30 también puede activar el PPS de textura y el MVC SPS de textura como se describe anteriormente.

30 **[0215]** En cualquier caso, cuando el fragmento de vídeo se codifica como un fragmento intracodificado (I), la unidad de predicción intra 84 de la unidad de predicción 81 puede generar datos de predicción para un bloque de vídeo del fragmento de vídeo actual basándose en un modo de predicción intra señalada y datos de bloques previamente descodificados de la imagen actual en la vista de los conjuntos de parámetros activos. Cuando la imagen se codifica como un fragmento intercodificado (es decir, B, P o PB generalizado (GPB)), la unidad de compensación de movimiento 82 de la unidad de predicción 81 produce bloques predictivos para un bloque de vídeo del fragmento de vídeo actual basándose en los vectores de movimiento y otros elementos sintácticos recibidos desde la unidad de descodificación por entropía 80, de nuevo en la vista de los conjuntos de parámetros activos. Los bloques predictivos se pueden producir a partir de una de las imágenes de referencia dentro de una de las listas de imágenes de referencia. El descodificador de vídeo 30 puede construir las listas de imágenes de referencia, la lista 0 y la lista 1 (o una lista combinada, la lista c), usando técnicas de construcción por omisión, basándose en las imágenes de referencia almacenadas en la memoria de imágenes de referencia 92.

45 **[0216]** La unidad de compensación de movimiento 82 determina la información de predicción para un bloque de vídeo del fragmento de vídeo actual analizando los vectores de movimiento y otros elementos sintácticos, y usa la información de predicción para producir los bloques predictivos del bloque de vídeo actual que se está descodificando. Por ejemplo, la unidad de compensación de movimiento 82 usa algunos de los elementos sintácticos recibidos para determinar un modo de predicción (por ejemplo, predicción intra o inter) usada para codificar los bloques de vídeo del fragmento de vídeo, un tipo de fragmento de predicción inter (por ejemplo, fragmento B, fragmento P o fragmento GPB), información de construcción para una o más de las listas de imágenes de referencia para el fragmento, vectores de movimiento para cada bloque de vídeo intercodificado del fragmento, estado de predicción inter para cada bloque de vídeo intercodificado del fragmento y otra información para descodificar los bloques de vídeo en el fragmento de vídeo actual.

55 **[0217]** La unidad de cuantificación inversa 86 cuantifica de manera inversa, es decir, descuantifica, los coeficientes de transformada cuantificados proporcionados en el flujo de bits y descodificados por la unidad de descodificación por entropía 80. El proceso de cuantificación inversa puede incluir el uso de un parámetro de cuantificación calculado por el codificador de vídeo 20 para cada bloque de vídeo del fragmento de vídeo para determinar un grado de cuantificación y, asimismo, un grado de cuantificación inversa que se debería aplicar.

60 **[0218]** La unidad de procesamiento de transformada inversa 88 aplica una transformada inversa, por ejemplo una DCT inversa, una transformada inversa entera o un proceso de transformada inversa conceptualmente similar, a los coeficientes de transformada, a fin de producir bloques residuales en el dominio de píxeles. La unidad de procesamiento de transformada inversa 88 puede determinar la forma en la que las transformaciones se aplicaron a los datos residuales.

65

[0219] Después de que la unidad de compensación de movimiento 82 genera el bloque predictivo para el bloque de vídeo actual, basándose en los vectores de movimiento y otros elementos sintácticos, el descodificador de vídeo 30 forma un bloque de vídeo descodificado sumando los bloques residuales de la unidad de procesamiento de transformada inversa 88 a los correspondientes bloques predictivos generados por la unidad de compensación de movimiento 82. El sumador 90 representa el componente o componentes que llevan a cabo esta operación de suma. Si se desea, también se puede aplicar un filtro de eliminación de bloques para filtrar los bloques descodificados a fin de eliminar distorsiones de efecto pixelado. También se pueden usar otros filtros de bucle (en el bucle de codificación o bien después del bucle de codificación) para suavizar las transiciones de píxeles o mejorar de otro modo la calidad del vídeo. Los bloques de vídeo descodificados en una imagen dada se almacenan a continuación en la memoria de imágenes de referencia 92, que almacena imágenes de referencia usadas para la posterior compensación de movimiento. La memoria de imágenes de referencia 92 almacena también vídeo descodificado para su presentación posterior en un dispositivo de visualización, tal como el dispositivo de visualización 32 de la FIG. 1.

[0220] Otros aspectos de las técnicas descritas en la presente divulgación pueden promover el tratamiento separado de los componentes de profundidad y textura permitiendo que un dispositivo de codificación de vídeo elimine por separado los componentes de textura y profundidad a partir de una memoria intermedia de imágenes descodificadas. La memoria intermedia de imágenes descodificadas se refiere a una memoria u otro medio legible por ordenador que puede almacenar imágenes de referencia para su uso en la codificación de un componente de vista actual. La memoria intermedia de imágenes descodificadas se muestra como la memoria de imágenes de referencia 92 en el ejemplo de la FIG. 3. Debido a que las técnicas expuestas en la presente divulgación permiten el tratamiento separado de los componentes de textura y profundidad, el descodificador de vídeo 30 puede descodificar un componente de textura de un componente de vista usando imágenes de referencia diferentes a las usadas cuando se codifica un componente de profundidad del mismo componente de vista en contraposición a 3DVC convencional donde un componente de vista (y por lo tanto los componentes de profundidad y textura de ese componente de vista) se codificó usando las mismas imágenes de referencia dado el tratamiento de las vistas en MVC. Como resultado, pueden existir casos donde los componentes de profundidad de un componente de vista de referencia se puedan eliminar de la memoria intermedia de imágenes descodificadas antes de los componentes de textura del mismo componente de vista de referencia, y un componente de textura de un componente de vista de referencia se pueda eliminar de la memoria intermedia de imágenes descodificadas antes del componente de profundidad del mismo componente de vista de referencia.

[0221] En funcionamiento, el descodificador de vídeo 30 puede llevar a cabo las técnicas descritas en la presente divulgación almacenando un componente de profundidad en la memoria de imágenes de referencia 92 y analizando una dependencia de vista para determinar si el componente de profundidad se usa para la predicción entre vistas. El descodificador de vídeo 30, a continuación, puede eliminar el componente de profundidad de la memoria de imágenes de referencia 92 en respuesta a la determinación de que el componente de profundidad no se usa para la predicción entre vistas. Es decir, el descodificador de vídeo 30 puede eliminar el componente de profundidad de un componente de vista de referencia de la memoria de imágenes de referencia 92 sin eliminar el componente de textura del mismo componente de vista de referencia en respuesta a la determinación de que el componente de profundidad no se usa para la predicción entre vistas. El descodificador de vídeo 30 puede validar, antes de eliminar el componente de profundidad, este componente de profundidad como que es elegible para su eliminación determinando que el componente de profundidad no pertenece a una vista de salida de destino y que el componente de profundidad está asociado con un código de identificación de referencia de capa de abstracción de red que tiene un valor igual a cero (lo que significa que el contenido de la unidad NAL que encapsula el componente de profundidad no se usa para construir imágenes de referencia para la predicción entre imágenes).

[0222] Es decir, en 3DVC (que de nuevo se puede denominar MVC más profundidad o MVC+D), la parte de textura o profundidad de cada vista tiene sus propios procesos de marcado de imagen de referencia. Los componentes de textura y profundidad que pertenecen a las vistas de salida de destino se pueden emitir simultáneamente.

[0223] Para un componente de profundidad que no pertenece a una vista de salida de destino y tiene `nal_ref_idc` igual a 0, el descodificador de vídeo 30 puede eliminar el componente de profundidad de la DPB una vez que el componente de profundidad no se use nunca para referencia entre vistas, analizando la dependencia de vista señalada en la extensión MVC SPS del SPS de subconjunto que contiene un perfil 3DV y se activa actualmente como la RBSP del conjunto de parámetros de secuencia de MVC de vista activa. De nuevo, la referencia a MVC en la presente divulgación se refiere a MVC en el contexto de 3DVC, que como se indica anteriormente es una extensión propuesta de H.264.

[0224] Por ejemplo, las técnicas de la presente divulgación pueden permitir que un codificador de vídeo (que se puede referir a ambos o a cualquiera de un codificador de vídeo 20 y/o un descodificador de vídeo 30) almacene un componente de profundidad en un memoria intermedia de imágenes descodificadas, no perteneciendo el componente de profundidad a una vista de salida de destino y teniendo un `nal_ref_idc` igual a 0. Además, un codificador de vídeo puede analizar una dependencia de vista para determinar si el componente de profundidad no se usa nunca para la referencia entre vistas. Además, el codificador de vídeo puede comprender eliminar un componente de profundidad de una memoria intermedia de imágenes descodificadas en respuesta a la determinación de que el componente de profundidad no se usa nunca para la referencia entre vistas. La dependencia de vista se puede señalar en una

extensión de un conjunto de parámetros de secuencia (SPS) de MVC de un SPS de subconjunto, conteniendo el SPS de subconjunto un perfil 3DV, activándose el SPS de subconjunto actualmente como un conjunto de parámetros de secuencia de MVC de vista activo. El codificador de vídeo puede codificar el componente de profundidad. De forma alternativa, el codificador de vídeo puede descodificar el componente de profundidad.

[0225] Si bien se describen anteriormente diversos aspectos de las técnicas con respecto a uno del codificador de vídeo 20 o descodificador de vídeo 30, las técnicas se pueden implementar, en algunos casos, tanto por el codificador de vídeo 20 como por el descodificador de vídeo 30 de la misma manera o de una manera recíproca. Como resultado, se pueden implementar las técnicas, en general, por un codificador de vídeo, que puede representar un término usado en la presente divulgación para referirse a ambos o solo a uno del codificador de vídeo 20 y descodificador de vídeo 30. Por lo tanto, estas técnicas no se deben limitar a este respecto, pero en general se pueden implementar por cualquier codificador de vídeo.

[0226] La FIG. 4 es un diagrama de bloques que ilustra la unidad de encapsulación 21 con más detalle. En el ejemplo de la FIG. 4, la unidad de encapsulación 21 incluye una interfaz de entrada de vídeo 100, una interfaz de entrada de audio 102, una unidad de creación de archivos de vídeo 104 y una interfaz de salida de archivos de vídeo 106. La unidad de creación de archivos de vídeo 104, en este ejemplo, incluye una unidad de generación de mensajes de información de mejora complementaria (SEI) 108, una unidad de asignación de identificador (ID) de vista 110, una unidad de creación de representación 112 y una unidad de creación de puntos de funcionamiento 114.

[0227] La interfaz de entrada de vídeo 100 y la interfaz de entrada de audio 102 reciben datos de vídeo y audio codificados, respectivamente. Si bien no se muestra en el ejemplo de la FIG. 1, el dispositivo de origen 12 también puede incluir un origen de audio y un codificador de audio para generar datos de audio y codificar datos de audio, respectivamente. La unidad de encapsulación 21, a continuación, puede encapsular los datos de audio codificados y los datos de vídeo codificados para formar un archivo de vídeo. La interfaz de entrada de vídeo 100 y la interfaz de entrada de audio 102 pueden recibir datos de vídeo y audio codificados a medida que se codifican los datos, o puede recuperar datos de vídeo y audio codificados de un medio legible por ordenador. Tras recibir datos de vídeo y audio codificados, la interfaz de entrada de vídeo 100 y la interfaz de entrada de audio 102 pasan los datos de vídeo y audio codificados a la unidad de creación de archivos de vídeo 104 para ensamblarlos en un archivo de vídeo.

[0228] La unidad de creación de archivos de vídeo 104 puede corresponder a una unidad de control que incluye hardware, software y/o firmware configurada para llevar a cabo las funciones y procedimientos que se le atribuyen. La unidad de control puede llevar a cabo además las funciones atribuidas a la unidad de encapsulación 21 en general. Para ejemplos en los que la unidad de creación de archivos de vídeo 104 está incorporada en software y/o firmware, la unidad de encapsulación 21 puede incluir un medio legible por ordenador que comprende instrucciones para la unidad de creación de archivos de vídeo 104 y una unidad de procesamiento para ejecutar las instrucciones. Cada una de las subunidades de la unidad de creación de archivos de vídeo 104 (unidad de generación de mensajes SEI 108, unidad de asignación de ID de vista 110, unidad de creación de representación 112 y unidad de creación de puntos de funcionamiento 114, en este ejemplo) se puede implementar como unidades de hardware y/o módulos de software individuales, y se pueden integrar funcionalmente o separar adicionalmente en subunidades adicionales.

[0229] El módulo de creación de archivos de vídeo 104 puede corresponder a cualquier unidad de procesamiento o circuitos de procesamiento adecuados, tales como, por ejemplo, uno o más microprocesadores, circuitos integrados específicos de la aplicación (ASIC), matriz de puertas programables (FPGA), procesadores de señales digitales (DSP) o cualquier combinación de los mismos. La unidad de creación de archivos de vídeo 104 puede incluir además un medio legible por ordenador no transitorio que almacena instrucciones para cualquiera o la totalidad de la unidad de generación de mensajes SEI 108, la unidad de asignación de ID de vista 110, la unidad de creación de representación 112 y la unidad de creación de puntos de funcionamiento 114, así como un procesador para ejecutar las instrucciones.

[0230] En general, la unidad de creación de archivos de vídeo 104 puede crear uno o más archivos de vídeo incluyendo los datos de audio y vídeo recibidos. La unidad de creación de archivos de vídeo 104 puede construir una descripción de presentación de medios (MPD) para contenido de multimedia que incluye dos o más vistas. En otros ejemplos, la unidad de creación de archivos de vídeo 104 puede crear un manifiesto que almacena datos similares a los de la MPD para el contenido de multimedia. La unidad de generación de mensajes SEI 108 puede representar una unidad que genera mensajes SEI. La unidad de asignación de ID de vista 110 puede asignar identificadores de vista a cada una de las vistas del contenido de multimedia. La unidad de creación de representación 112 puede construir una o más representaciones para el contenido de multimedia, cada una de las cuales puede incluir una o más de las vistas para el contenido de multimedia. En algunos ejemplos, la unidad de asignación de ID de vista 110 puede incluir datos en la MPD y/o las representaciones (por ejemplo, datos de cabecera para las representaciones) que indican un máximo y un mínimo de los identificadores de vista para las vistas incluidas en las representaciones. Además, la unidad de creación de representación 112 puede proporcionar información en la MPD que indica si los ID de vista más grandes corresponden a vistas que tienen perspectivas de cámara a la derecha o a la izquierda de las perspectivas de cámara para vistas que tienen ID de vista más pequeños.

[0231] En algunos ejemplos, se puede codificar la misma vista usando diversas características de codificación, tales como diferentes velocidades de trama, diferentes tasas de bits, diferentes esquemas de codificación u otras

diferencias. La unidad de creación de representación 112 puede garantizar que cada vista incluida en una representación común se codifique de acuerdo con las mismas características de codificación. De esta forma, la MPD y/o los datos de cabecera para la representación pueden señalar un conjunto de características (o atributos) para la representación que se aplica a todas las vistas en la representación. Además, la unidad de creación de representación 112 puede crear múltiples representaciones que incluyen las mismas vistas, aunque con características de codificación potencialmente diferentes. En algunos ejemplos, la unidad de creación de representación 112 puede encapsular cada vista de contenido de multimedios en representaciones individuales. En dichos ejemplos, para emitir más de una vista, el dispositivo de destino 14 puede solicitar dos o más representaciones del contenido de multimedios.

[0232] La unidad de creación de puntos de funcionamiento 114 puede crear puntos de funcionamiento para una o más representaciones del contenido de multimedios. En general, un punto de funcionamiento corresponde a un subconjunto de vistas en una representación que están destinadas para su salida, donde cada una de las vistas comparte un nivel temporal común. Como se define por la especificación H.264/AVC: Un punto de funcionamiento se identifica por un valor de temporal_id que representa el *nivel temporal de destino* y un conjunto de valores de view_id que representan las *vistas de salida de destino*. Un *punto de funcionamiento* está asociado con un *subconjunto de flujo de bits*, que consiste en las *vistas de salida de destino* y todas las demás *vistas* de las que dependen las *vistas de salida de destino*, que se deriva usando el proceso de extracción del *subflujo de bits* como se especifica en la subcláusula H.8.5.3 con tIdTarget igual al valor temporal_id y consistiendo viewIdTargetList en el conjunto de valores de id de vista como entradas. Se puede asociar más de un *punto de funcionamiento* con el mismo *subconjunto de flujo de bits*. Cuando la especificación declara que "se descodifica un *punto de funcionamiento*" se refiere a la descodificación de un *subconjunto de flujo de bits* correspondiente al *punto de funcionamiento* y la salida posterior de las *vistas de salida de destino*.

[0233] La unidad de generación de mensajes SEI 108, de acuerdo con las técnicas descritas en la presente divulgación, puede determinar un mensaje de información complementaria (SEI) anidado que se aplica tanto a los componentes de textura como de profundidad de un componente de vista o solo a un componente de profundidad del componente de vista. Permitiendo que un mensaje SEI anidado se aplique solo a un componente de profundidad, estos aspectos de las técnicas pueden promover de nuevo el tratamiento separado de los componentes de textura y profundidad.

[0234] La unidad de generación de mensajes SEI 108 puede llevar a cabo estos aspectos de las técnicas para determinar un mensaje de información de mejora complementaria que se aplica a la codificación del componente de vista de los datos de vídeo de múltiples vistas. La unidad de generación de mensajes SEI 108, a continuación, puede determinar un mensaje de información de mejora complementaria anidado que se aplica además del mensaje de información de mejora complementaria cuando procesa el componente de profundidad del componente de vista. La unidad de generación de mensajes SEI 108, a continuación, puede codificar el componente de profundidad del componente de vista basándose en el mensaje de información de mejora complementaria y el mensaje de información de mejora complementaria anidado.

[0235] Típicamente, el mensaje SEI anidado incluye un indicador u otro identificador que indica si el mensaje SEI anidado se aplica tanto a los componentes de textura como de profundidad del componente de vista o sólo al componente de profundidad. Por tanto, en algunos casos, la unidad de generación de mensajes SEI 108 puede determinar que el mensaje de información de mejora complementaria anidado se aplica además del mensaje de información de mejora complementaria cuando solo procesa el componente de profundidad del componente de vista y procesa solo el componente de profundidad del componente de vista y no el componente de textura del componente de vista basándose en el mensaje de información de mejora complementaria y el mensaje de información de mejora complementaria anidado. De forma alternativa, la unidad de generación de mensajes SEI 108 puede determinar que el mensaje de información de mejora complementaria anidado se aplica además del mensaje de información de mejora complementaria cuando procesa el componente de textura del componente de vista y codifica el componente de textura del componente de vista basándose en el mensaje de información de mejora complementaria y el mensaje de información de mejora complementaria anidado. De nuevo, estos aspectos de las técnicas facilitan de nuevo el tratamiento o procesamiento separado de los componentes de profundidad y textura del componente de vista.

[0236] Este indicador se puede definir para un mensaje SEI denominado "mensaje SEI de anidamiento escalable de 3DV", que puede representar un mensaje SEI para anidar mensajes SEI para puntos de funcionamiento que contienen componentes de profundidad. La siguiente tabla ilustra una sintaxis de mensajes SEI de anidamiento escalable de 3DV de ejemplo.

	C	Descriptor
3dv_scalable_nesting(payloadSize) {		
operation_point_flag	5	u(1)
if (!operation_point_flag) {		
all_view_components_in_au_flag	5	u(1)
if(!all_view_components_in_au_flag) {		
num_view_components_minus1	5	ue(v)
for(i = 0; i <= num_view_components_minus1; i++) {		
sei_view_id[i]	5	u(10)
sei_view_applicability_flag[i]	5	u(1)
}		
}		
} else {		
num_view_components_op_minus1	5	ue(v)
for(i = 0; i <= num_view_components_op_minus1; i++) {		
sei_op_view_id[i]	5	u(10)
sei_op_view_applicability_flag[i]	5	u(1)
}		
sei_op_temporal_id	5	u(3)
}		
while(!byte_aligned())		
sei_nesting_zero_bit /* equal to 0 */	5	f(1)
sei_message()	5	
}		

5 [0237] En la tabla anterior, **sei_view_applicability_flag[i]** igual a 1 indica que el mensaje SEI anidado se aplica tanto al componente de textura como al componente de profundidad de la vista con id de vista igual a sei_view_id[i]. sei_view_applicability_flag[i] igual a 0 indica que el mensaje SEI anidado se aplica solo al componente de profundidad de la vista con view_id igual a sei_view_id[i].

10 [0238] En la tabla anterior, **sei_op_view_applicability_flag[i]** igual a 1 indica que el mensaje SEI anidado se aplica tanto al componente de textura como al componente de profundidad de la vista con view_id igual a sei_op_view_id[i]. sei_op_view_applicability_flag[i] igual a 0 indica que el mensaje SEI anidado se aplica solo al componente de profundidad de la vista con id de vista igual a sei_op_view_id[i].

15 [0239] La semántica de los otros elementos sintácticos del mensaje SEI de anidamiento escalable de 3DV es la misma que la semántica del mensaje SEI de anidamiento escalable de MVC.

20 [0240] De forma alternativa, cuando un mensaje SEI se anida para un componente de vista, el mensaje SEI se puede anidar tanto para el componente de textura como para el componente de profundidad de este componente de vista. En este caso, la unidad de generación de mensajes SEI 108 define la sintaxis del mensaje SEI de anidamiento escalable de 3DV similar a la sintaxis del mensaje SEI de anidamiento escalable de MVC, y la semántica del mensaje SEI de anidamiento escalable de 3DV similar a la semántica del mensaje SEI de anidamiento escalable de MVC. En la semántica del mensaje SEI de anidamiento ajustable a escala de 3DV, como se define previamente, un componente de vista puede incluir (y puede consistir en) un componente de textura y un componente de profundidad. En el contexto del Anexo I propuesto, que puede hacer referencia a la extensión 3DVC para la H.264/AVC, la semántica del mensaje SEI de anidamiento escalable de MVC se cambia como sigue. La semántica del mensaje SEI de anidamiento escalable de MVC se aplica con todos los casos de "componentes de vista" que se reemplazan por "componentes de textura". La siguiente tabla presenta la sintaxis del mensaje SEI de anidamiento escalable de MVC.

	C	Descriptor
<code>mvc_scalable_nesting(payloadSize) {</code>		
<code> operation_point_flag</code>	5	u(1)
<code> if (!operation_point_flag) {</code>		
<code> all_view_components_in_au_flag</code>	5	u(1)
<code> if(!all_view_components_in_au_flag) {</code>		
<code> num_view_components_minus1</code>	5	uc(v)
<code> for(i = 0; i <= num_view_components_minus1; i++)</code>		
<code> sei_view_id[i]</code>	5	u(10)
<code> } else {</code>		
<code> num_view_components_op_minus1</code>	5	uc(v)
<code> for(i = 0; i <= num_view_components_op_minus1; i++)</code>		
<code> sei_op_view_id[i]</code>	5	u(10)
<code> sei_op_temporal_id</code>	5	u(3)
<code> }</code>		
<code> while(!byte_aligned())</code>		
<code> sei_nesting_zero_bit /* equal to 0 */</code>	5	f(1)
<code> sei_message()</code>	5	
<code> }</code>		

5 [0241] Para facilitar aún más el tratamiento separado de los componentes de profundidad y textura, la unidad de generación de mensajes SEI 108 puede generar un mensaje SEI denominado mensaje SEI de escalabilidad de vista 3D que incluye un primer elemento sintáctico, un segundo elemento sintáctico y un tercer elemento sintáctico, donde el primer elemento sintáctico indica si un punto de funcionamiento contiene un componente de profundidad. Cuando el punto de funcionamiento contiene un componente de profundidad, la unidad de generación de mensajes SEI 108 puede definir el segundo elemento sintáctico para indicar un número de componentes de profundidad de los que depende directamente la vista de salida de destino del punto de funcionamiento y definir el tercer elemento sintáctico para identificar los componentes de profundidad de los que depende directamente la vista de salida de destino del punto de funcionamiento.

15 [0242] En algunos casos, los puntos de funcionamiento pueden incluir un primer subconjunto de los puntos de funcionamiento y un segundo subconjunto de los puntos de funcionamiento, donde cada punto de funcionamiento en el primer subconjunto contiene sólo un componente de textura y cada punto de funcionamiento en el segundo subconjunto contiene un componente de profundidad. La unidad de generación de mensajes SEI 108, a continuación, puede determinar que un identificador de perfil es igual a un valor dado que está asociado con la codificación de vídeo 3D y generar, en respuesta a la determinación de que el identificador de perfil es igual al valor dado, un SPS de subconjunto de modo que el SPS de subconjunto incluya una primera extensión SPS y una segunda extensión SPS. 20 La primera extensión SPS puede indicar el primer subconjunto y los niveles a los que pertenecen los puntos de funcionamiento en el primer subconjunto. La segunda extensión SPS puede indicar el segundo subconjunto y los niveles a los que pertenecen los puntos de funcionamiento en el segundo subconjunto. De esta manera, las técnicas pueden permitir que los puntos de operaciones se definan por separado para los componentes de textura y profundidad y por nivel.

25 [0243] De esta forma, un dispositivo de codificación de vídeo puede determinar la primera información de nivel de secuencia que describe las características de los componentes de profundidad y determinar la segunda información de nivel de secuencia que describe las características de un punto de funcionamiento de los datos de vídeo. En algunos ejemplos, la primera información de nivel de secuencia comprende un conjunto de parámetros de secuencia de codificación de vídeo tridimensional que especifica una dependencia de vista de los componentes de profundidad. Además, el dispositivo de codificación de vídeo puede determinar una lista de imágenes de referencia que identifica una o más imágenes de referencia para los componentes de profundidad indicados en el conjunto de parámetros de secuencia de codificación de vídeo tridimensional.

35 [0244] En algunos ejemplos, la segunda información de nivel de secuencia incluye un conjunto de parámetros de secuencia de codificación de vídeo tridimensional que describe, para el punto de funcionamiento, una lista de vistas de salida de destino, un número de vistas de textura que se van a descodificar al descodificar el punto de funcionamiento y un número de vistas de profundidad que se van a descodificar al descodificar el punto de funcionamiento, donde la serie de vistas de textura que se van a descodificar es diferente de la serie de vistas de profundidad. En algunos ejemplos, el dispositivo de codificación de vídeo puede seleccionar además, para cada una 40

de las vistas de salida de destino especificadas en la lista de vistas de salida de destino, el uno o más componentes de profundidad cuando estén disponibles.

5 **[0245]** En algunos ejemplos, el dispositivo de codificación de vídeo puede especificar un mensaje SEI de información de escalabilidad de codificación de vídeo tridimensional como parte de los datos de vídeo, en el que el mensaje SEI de información de escalabilidad de codificación de vídeo tridimensional incluye una descripción del punto de funcionamiento sin los componentes de profundidad o una descripción del punto de funcionamiento con los componentes de profundidad. En algunos casos, el mensaje SEI incluye una indicación de si el punto de funcionamiento incluye los componentes de profundidad. Además, en algunos ejemplos, el mensaje SEI incluye una
10 indicación de si el punto de funcionamiento incluye los componentes de profundidad y las vistas de profundidad directamente dependientes con vistas de textura directamente dependientes para la descodificación del punto de funcionamiento.

15 **[0246]** Este aspecto de las técnicas puede aprovechar el tratamiento separado de los componentes de profundidad y textura para permitir que los puntos de funcionamiento se extraigan más fácilmente de un flujo de bits para la codificación de vídeo de múltiples vistas. Señalando si un punto de funcionamiento incluye profundidad, un descodificador de vídeo, tal como el descodificador de vídeo 30, puede determinar qué punto de funcionamiento da cabida mejor a las capacidades del descodificador de vídeo 30. Por ejemplo, el descodificador de vídeo 30 se puede optimizar para llevar a cabo la codificación de vídeo de múltiples vistas usando vistas que no incluyen profundidad, prefiriendo datos de vídeo de múltiples vistas donde se proporcionan tanto las perspectivas del ojo izquierdo como del derecho como imágenes separadas. El descodificador de vídeo 30 también se puede optimizar, de forma alternativa, para datos de vídeo de múltiples vistas basados en profundidad, pero solo puede dar cabida a una vista única en lugar de dos o más vistas. Señalando cuándo los datos de vídeo de múltiples vistas codificados incluyen profundidad, y al incluir la profundidad, el número de componentes de profundidad de los que dependen directamente las vistas de
20 salida de destino del punto de funcionamiento, el descodificador de vídeo 30 puede seleccionar un punto de funcionamiento que dé cabida a las capacidades u optimizaciones del descodificador de vídeo 30.

25 **[0247]** En funcionamiento, un mensaje SEI de información de escalabilidad de vista de 3D, en un ejemplo, puede tener elementos sintácticos especificados como se muestra en la siguiente tabla.

30

3d_view_scalability_info(payloadSize) {	C	Descriptor
num_operation_points_minus1	5	ue(v)
for(i = 0; i <= num_operation_points_minus1; i++) {		
operation_point_id[i]	5	ue(v)
priority_id[i]	5	u(5)
temporal_id[i]	5	u(3)
num_target_output_views_minus1[i]	5	ue(v)
op_with_depth_flag[i]		
for(j = 0; j <= num_target_output_views_minus1[i]; j++)		
view_id[i][j]	5	ue(v)
profile_level_info_present_flag[i]	5	u(1)
bitrate_info_present_flag[i]	5	u(1)

frm_rate_info_present_flag[i]	5	u(1)
if(!num_target_output_views_minus1[i])		
view_dependency_info_present_flag[i]	5	u(1)
parameter_sets_info_present_flag[i]	5	u(1)
bitstream_restriction_info_present_flag[i]	5	u(1)
if(profile_level_info_present_flag[i])		
op_profile_level_idc[i]	5	u(24)
if(bitrate_info_present_flag[i]) {		
avg_bitrate[i]	5	u(16)
max_bitrate[i]	5	u(16)
max_bitrate_calc_window[i]	5	u(16)
}		
if(frm_rate_info_present_flag[i]) {		
constant_frm_rate_idc[i]	5	u(2)
avg_frm_rate[i]	5	u(16)
}		
if(view_dependency_info_present_flag[i]) {		
num_directly_dependent_views[i]	5	ue(v)
for(j = 0; j < num_directly_dependent_views[i]; j++) {		
directly_dependent_view_id[i][j]	5	ue(v)
if(op_with_depth_flag[i]) {		
num_directly_dependent_depth_views[i]		
for(j = 0; j < num_directly_dependent_views[i]; j++) {		
directly_dependent_depth_view_id[i][j]	5	ue(v)
}		
} else		
view_dependency_info_src_op_id[i]	5	ue(v)
if(parameter_sets_info_present_flag[i]) {		
num_seq_parameter_set_minus1[i]	5	ue(v)
for(j = 0; j <= num_seq_parameter_set_minus1[i]; j++)		
seq_parameter_set_id_delta[i][j]	5	ue(v)
num_subset_seq_parameter_set_minus1[i]	5	ue(v)
for(j = 0; j <= num_subset_seq_parameter_set_minus1[i]; j++)		
subset_seq_parameter_set_id_delta[i][j]	5	ue(v)
num_pic_parameter_set_minus1[i]	5	ue(v)
for(j = 0; j <= num_init_pic_parameter_set_minus1[i]; j++)		
pic_parameter_set_id_delta[i][j]	5	ue(v)
} else		

parameter_sets_info_src_op_id[i]	5	uc(v)
if(bitstream_restriction_info_present_flag[i]) {		
motion_vectors_over_pic_boundaries_flag[i]	5	u(1)
max_bytes_per_pic_denom[i]	5	uc(v)
max_bits_per_mb_denom[i]	5	uc(v)
log2_max_mv_length_horizontal[i]	5	uc(v)
log2_max_mv_length_vertical[i]	5	uc(v)
num_reorder_frames[i]	5	uc(v)
max_dec_frame_buffering[i]	5	uc(v)
}		
}		
}		

[0248] De forma alternativa, los mensajes SEI de información de escalabilidad de vista de 3D pueden tener elementos sintácticos como se define en la siguiente tabla.

5

3d_view_scalability_info(payloadSize) {	C	Descriptor
//same syntax elements as in view_scalability_info()		
for(i = 0; i <= num_operation_points_minus1; i++) {		
op_with_depth_flag[i]	5	u(1)
if(view_dependency_info_present_flag[i] && op_with_depth_flag[i]) {		
num_directly_dependent_depth_views[i]		
for(j = 0; j < num_directly_dependent_views[i]; j++) {		
directly_dependent_depth_view_id[i][j]	5	uc(v)
}		
}		
}		

[0249] La semántica de los otros elementos sintácticos del mensaje de escalabilidad de vista de 3DV es la misma que la semántica del mensaje SEI de información de escalabilidad de vista de MVC. Los elementos sintácticos en la SEI de información de escalabilidad de vista tienen la misma semántica que en el Anexo H, pero se aplican a puntos de funcionamiento que potencialmente pueden contener componentes de profundidad.

10

[0250] En las tablas anteriores, op_with_depth_flag[i] igual a 1 indica que el punto de funcionamiento actual contiene vistas de profundidad. op_with_depth_flag[i] igual a 0 indica que el punto de funcionamiento actual no contiene ninguna vista de profundidad. num_directly_dependent_views[i] y directly_dependent_view_id[i][j] se aplican solo a los componentes de textura de un punto de funcionamiento si el punto de funcionamiento contiene tanto textura como profundidad (op_with_depth_flag[i] igual a 0). De forma alternativa, los nombres de los dos elementos sintácticos se cambian a num_directly_dependent_texture_views[i] y directly_dependent_texture_view_id[i][j], respectivamente, con la misma semántica que anteriormente. La semántica de los otros elementos sintácticos del mensaje de escalabilidad de vista de 3DV es la misma que la semántica del mensaje SEI de información de escalabilidad de vista de MVC.

15

20

[0251] En las tablas anteriores, num_directly_dependent_depth_views[i] especifica el número de vistas de profundidad de las que depende directamente la vista de salida de destino del punto de funcionamiento actual dentro de la representación del punto de funcionamiento actual. El valor de num_directly_dependent_depth_views[i] estará en el intervalo de 0 a 16, inclusive.

25

[0252] En las tablas anteriores, directly_dependent_depth_view_id[i][j] especifica el id de vista de la j-ésima vista de profundidad de la que la vista de profundidad de la vista de salida de destino del punto de funcionamiento actual depende directamente dentro de la representación del punto de funcionamiento actual. El valor de directly_dependent_depth_view_id[i][j] estará en el intervalo de 0 a 1023, inclusive.

30

[0253] En otros ejemplos, los puntos de funcionamiento de solo profundidad se pueden señalar en este mensaje SEI y por tanto se añade una indicación op_texture_depth_idc[i], como se muestra en la siguiente tabla.

3d_view_scalability_info(payloadSize) {	C	Descriptor
view_scalability_info(payloadSize)		
for(i = 0; i <= num_operation_points_minus1; i++) {		
op_texture_depth_idc[i]	5	ue(v)
if(view_dependency_info_present_flag[i] && op_with_depth_flag[i] == 1) {		
num_directly_dependent_depth_views[i]		
for(j = 0; j < num_directly_dependent_views[i]; j++) {		
directly_dependent_depth_view_id[i][j]	5	ue(v)
}		
}		
}		

5 [0254] En la tabla anterior, op_texture_depth_idc[i] igual a 0 indica que los puntos de funcionamiento actuales sólo contienen componentes de textura. El elemento sintáctico op_texture_depth_idc[i] igual a 1 indica que el punto de funcionamiento actual contiene tanto componentes de textura como componentes de profundidad. El elemento sintáctico op_texture_depth_idc[i] igual a 2 indica que el punto de funcionamiento actual solo contiene componentes de profundidad. Num_directly_dependent_views[i] y directly_dependent_view_id[i][j] se aplican solo a los componentes de textura de un punto de funcionamiento si el punto de funcionamiento contiene textura (op_texture_depth_idc[i] igual a 0 o 1). Num_directly_dependent_views[i] y directly_dependent_view_id[i][j] se aplican solo a los componentes de profundidad de un punto de funcionamiento si el punto de funcionamiento solo contiene textura (op_texture_depth_idc[i] igual a 2).

15 [0255] En algunos ejemplos, de acuerdo con las técnicas de la divulgación, un codificador de vídeo puede generar un mensaje SEI que incluye un primer elemento sintáctico, un segundo elemento sintáctico y terceros elementos sintácticos. El primer elemento sintáctico indica si un punto de funcionamiento contiene un componente de profundidad. Si el punto de funcionamiento contiene una vista de profundidad, el segundo elemento sintáctico indica un número de componentes de profundidad de los que depende directamente una vista de salida de destino del punto de funcionamiento, y los terceros elementos sintácticos identifican los componentes de profundidad de los que depende directamente la vista de salida de destino del punto de funcionamiento. Un codificador de vídeo puede codificar los componentes de profundidad.

25 [0256] En algunos de dichos ejemplos, el codificador de vídeo puede generar el mensaje SEI de modo que el mensaje SEI incluya un cuarto elemento sintáctico y quintos elementos sintácticos. Si el punto de funcionamiento tiene tanto componentes de textura como de profundidad, el cuarto elemento sintáctico indica cuántos componentes de textura están asociados con el punto de funcionamiento y los quintos elementos sintácticos identifican los componentes de textura de los que depende directamente la vista de salida de destino del punto de funcionamiento.

30 [0257] De forma alternativa, el codificador de vídeo puede señalar puntos de funcionamiento de solo profundidad en el mensaje SEI. Por ejemplo, el codificador de vídeo puede generar el mensaje SEI de modo que el primer elemento sintáctico indique si el punto de funcionamiento contiene solo componentes de textura, contiene tanto componentes de textura como componentes de profundidad, o contiene solo componentes de profundidad.

35 [0258] Además, en algunos ejemplos, las técnicas de la presente divulgación pueden proporcionar un descodificador de vídeo que recibe un mensaje SEI que incluye un primer elemento sintáctico, un segundo elemento sintáctico y terceros elementos sintácticos. El primer elemento sintáctico indica si un punto de funcionamiento contiene un componente de profundidad. Si el punto de funcionamiento contiene una vista de profundidad, el segundo elemento sintáctico indica un número de componentes de profundidad de los que depende directamente una vista de salida de destino del punto de funcionamiento y el tercer elemento sintáctico identifica los componentes de profundidad de los que depende directamente la vista de salida de destino del punto de funcionamiento. El descodificador de vídeo puede descodificar los componentes de profundidad.

45 [0259] Además, el mensaje SEI puede incluir un cuarto elemento sintáctico y quintos elementos sintácticos. Si el punto de funcionamiento tiene tanto componentes de textura como de profundidad, el cuarto elemento sintáctico indica cuántos componentes de textura están asociados con el punto de funcionamiento y los quintos elementos sintácticos identifican los componentes de textura de los que depende directamente la vista de salida de destino del punto de funcionamiento.

50 [0260] De forma alternativa, el mensaje SEI puede señalar puntos de funcionamiento de solo profundidad en el mensaje SEI. Por ejemplo, el primer elemento de sintaxis puede indicar si el punto de funcionamiento contiene solo

componentes de textura, contiene tanto componentes de textura como componentes de profundidad, o solo contiene componentes de profundidad.

5 **[0261]** Con respecto a los puntos de funcionamiento y la adaptación o modificación de los puntos de funcionamiento para proporcionar potencialmente un tratamiento separado de los componentes de textura y profundidad, diversos aspectos de las técnicas descritas en la presente divulgación pueden proporcionar la señalización de niveles de puntos de funcionamiento que contienen componentes de profundidad de cuatro maneras diferentes: 1) reutilizando los elementos sintácticos SPS, 2) usando elementos sintácticos similares en la extensión 3DVC SPS, 3) añadiendo más elementos sintácticos en la extensión 3DVC SPS y 4) usando un conjunto alternativo de más elementos sintácticos en la extensión 3DVC SPS.

15 **[0262]** Con respecto a la primera manera, es decir, reutilizar los elementos sintácticos SPS, se usan elementos sintácticos de definiciones de nivel en la extensión MVC SPS para indicar los puntos de funcionamiento que contienen componentes de profundidad. Cada punto de funcionamiento señalado en la extensión MVC SPS como parte del SPS de subconjunto contiene un `profile_idc` que pertenece a uno de los perfiles de vídeo 3D. Cada punto de funcionamiento especifica un número de vistas de salida, cada una de las que contiene tanto un componente de textura como un componente de profundidad. La unidad de creación de puntos de funcionamiento 114 puede implementar esta primera manera para facilitar el tratamiento separado de los componentes de profundidad y textura.

20 **[0263]** Con respecto a la segunda manera, es decir, usar elementos sintácticos similares en la extensión 3DV SPS, los elementos sintácticos de definiciones de nivel en la extensión MVC SPS todavía se pueden usar para indicar los puntos de funcionamiento que contienen sólo los componentes de textura. En una extensión 3DVC SPS del SPS de subconjunto, se añaden definiciones de nivel similares, que indican los puntos de funcionamiento que contienen componentes de profundidad. La unidad de creación de puntos de funcionamiento 114 puede implementar esta segunda manera para facilitar el tratamiento separado de los componentes de profundidad y textura. Las siguientes tablas sintácticas ilustran los elementos sintácticos similares.

	C	Descriptor
<code>subset_seq_parameter_set_rbsp() {</code>		
<code> seq_parameter_set_data()</code>	0	
<code> if(profile_idc == 83 profile_idc == 86) {</code>		
<code> seq_parameter_set_svc_extension() /*especificado en el Anexo G</code>	0	
<code> */</code>		
<code> svc_vui_parameters_present_flag</code>	0	u(1)
<code> if(svc_vui_parameters_present_flag == 1)</code>		
<code> svc_vui_parameters_extension() /*especificado en el Anexo G</code>	0	
<code> */</code>		
<code> } else if(profile_idc == 118 profile_idc == 128) {</code>		

bit_equal_to_one /* igual a 1 */	0	f(1)
seq_parameter_set_mvc_extension() /* especificado en el Anexo H */	0	
mvc_vui_parameters_present_flag	0	u(1)
if(mvc_vui_parameters_present_flag == 1)		
mvc_vui_parameters_extension() /* especificado en el Anexo H */	0	
}		
else if(profile_idc == 138) {		
bit_equal_to_one /* igual a 1 */	0	f(1)
seq_parameter_set_mvc_extension()		
mvc_vui_parameters_present_flag	0	u(1)
if(mvc_vui_parameters_present_flag == 1)		
mvc_vui_parameters_extension() /* especificado en el Anexo H */	0	
seq_parameter_set_3dv_extension()		
...		
}		
}		
additional_extension3_flag	0	u(1)
if(additional_extension3_flag)		
while(more_rbsp_data())		
additional_extension3_data_flag	0	u(1)
}		
rbsp_trailing_bits()	0	
}		

	C	Descriptor
seq_parameter_set_3dv_extension() {		
for(i = 0; i <= num_level_values_signalled_minus1; i++)		
num_level_values_signalled_minus1	0	uc(v)
for(i = 0; i <= num_level_values_signalled_minus1; i++) {		
level_idc[i]	0	u(8)
num_applicable_ops_minus1[i]	0	uc(v)
for(j = 0; j <= num_applicable_ops_minus1[i]; j++) {		
applicable_op_temporal_id[i][j]	0	u(3)
applicable_op_num_target_views_minus1[i][j]	0	uc(v)
for(k = 0; k <= applicable_op_num_target_views_minus1[i][j]; k++)		
applicable_op_target_view_id[i][j][k]	0	uc(v)
applicable_op_num_views_minus1[i][j]	0	uc(v)
}		
}		

}		
...		
}		

5 **[0264]** En la tabla sintáctica anterior, cada punto de funcionamiento tiene `num_target_views_minus1[i][j]+1` vistas, cada una de las que contiene tanto un componente de textura como un componente de profundidad. Así que, la semántica de los elementos sintácticos es la misma que la de la extensión SPS MVC, excepto que cada vista contiene tanto un componente de textura como un componente de profundidad, o de forma equivalente, cada componente de vista contiene tanto un componente de textura como un componente de profundidad.

10 **[0265]** Con respecto a la tercera manera, es decir, añadir más elementos sintácticos en extensión 3DV SPS, la extensión 3DV SPS puede añadir más puntos de funcionamiento que contienen componentes de textura que es necesario señalar en la parte superior de los puntos de funcionamiento señalados en la extensión MVC SPS. En este caso, la unidad de creación de puntos de funcionamiento 114 puede añadir un indicador adicional para indicar si el punto de funcionamiento actual contiene un componente de profundidad o no, como se muestra en la siguiente tabla sintáctica.

	C	Descriptor
<code>seq_parameter_set_3dvc_extension() {</code>		
<code>for(i = 0; i <= num_level_values_signalled_minus1; i++)</code>		
<code>num_level_values_signalled_minus1</code>	0	uc(v)
<code>for(i = 0; i <= num_level_values_signalled_minus1; i++) {</code>		
<code>level_idc[i]</code>	0	u(8)
<code>num_applicable_ops_minus1[i]</code>	0	uc(v)
<code>for(j = 0; j <= num_applicable_ops_minus1[i]; j++) {</code>		
<code>applicable_op_temporal_id[i][j]</code>	0	u(3)
<code>applicable_op_num_target_views_minus1[i][j]</code>	0	uc(v)
<code>applicable_op_with_depth_flag[i][j]</code>	0	u(1)
<code>for(k = 0; k <=</code> <code>applicable_op_num_target_views_minus1[i][j]; k++)</code>		
<code>applicable_op_target_view_id[i][j][k]</code>	0	uc(v)
<code>applicable_op_num_views_minus1[i][j]</code>	0	uc(v)
<code>}</code>		
<code>}</code>		
<code>...</code>		
<code>}</code>		

15 **[0266]** En la tabla anterior, `applicable_op_with_depth_flag[i][j]` igual a 1 indica que en el punto de funcionamiento actual (identificado por `i` y `j`), cada vista de salida de destino contiene tanto un componente de textura como un componente de profundidad, este indicador igual a 0 indica que el punto de funcionamiento actual contiene solo textura para cada vista de salida de destino. La semántica de los otros elementos sintácticos del mensaje SEI de anidamiento escalable de 3DVC es la misma que la semántica del mensaje SEI de anidamiento escalable de MVC.

20

25 **[0267]** Con respecto a la cuarta manera, es decir, usar un conjunto alternativo de más elementos sintácticos en la extensión 3DV SPS, podría haber puntos de funcionamiento de solo profundidad. En este caso, la unidad de creación de puntos de funcionamiento 114 puede añadir una indicación de 2 bits para señalar tres casos diferentes para cada punto de funcionamiento: solo textura, solo profundidad y tanto textura como profundidad. La siguiente tabla ilustra estos nuevos elementos sintácticos que se van a añadir a la extensión SPS 3DVC:

	C	Descriptor
seq_parameter_set_3dvc_extension() {		
for(i = 0; i <= num_level_values_signalled_minus1; i++)		
num_level_values_signalled_minus1	0	ue(v)
for(i = 0; i <= num_level_values_signalled_minus1; i++) {		
level_idc[i]	0	u(8)
num_applicable_ops_minus1[i]	0	ue(v)
for(j = 0; j <= num_applicable_ops_minus1[i]; j++) {		
applicable_op_temporal_id[i][j]	0	u(3)
applicable_op_num_target_views_minus1[i][j]	0	ue(v)
applicable_op_with_texture_depth_idc[i][j]	0	u(2)
for(k = 0; k <=		
applicable_op_num_target_views_minus1[i][j]; k++)		
applicable_op_target_view_id[i][j][k]	0	ue(v)
applicable_op_num_views_minus1[i][j]	0	ue(v)
}		
}		
...		
}		

5 [0268] En la tabla anterior, `applicable_op_with_texture_depth_idc[i][j]` igual a 0 indica que en el punto de funcionamiento actual (identificado por *i* y *j*), cada vista de salida de destino contiene tanto un componente de textura como un componente de profundidad, `applicable_op_with_texture_depth_idc[i][j]` igual a 1 indica que el punto de funcionamiento actual contiene solo textura para cada vista de salida de destino, `applicable_op_with_texture_depth_idc[i][j]` igual a 2 indica que el punto de funcionamiento actual contiene solo profundidad para cada vista de salida de destino. De forma alternativa, `applicable_op_with_texture_depth_idc[i][j]` se puede codificar como `ue(v)`.

10 [0269] En algunos casos, para un punto de funcionamiento de 3DV, se pueden señalar dos dependencias de vista diferentes, por ejemplo, en un MVC SPS de textura activo y un MVC SPS de vista activo porque las vistas de profundidad pueden tener diferentes dependencias entre vistas. Un ejemplo es que podría no haber dependencia entre vistas en absoluto para la profundidad. En este caso, la semántica de `applicable_op_num_views_minus1 [i][j]` no está clara. En el contexto de MVC, `applicable_op_num_views_minus1[i][j]` especifica el número de vistas que se van a descodificar para un punto de funcionamiento específico.

20 [0270] En algunos ejemplos, se puede suponer la misma dependencia de vista para la textura y la profundidad. Si las dependencias de vista de textura y profundidad son las mismas, la semántica de este elemento sintáctico (es decir, `applicable_op_num_views_minus1[i][j]`) se puede mantener alineada con el número de vistas que se van a descodificar, sin embargo, está implícito que cada vista contiene textura y profundidad.

25 [0271] En algunos ejemplos, se señalan números exactos de vistas que se van a descodificar tanto para textura como para profundidad. La siguiente tabla ilustra los elementos sintácticos de una extensión 3DVC del conjunto de parámetros de secuencia que incorpora este aspecto de las técnicas:

	C	Descriptor
seq_parameter_set_3dvc_extension() {		
for(i = 0; i <= num_level_values_signalled_minus1; i++)		
num_level_values_signalled_minus1	0	uc(v)
for(i = 0; i <= num_level_values_signalled_minus1; i++) {		
level_idc[i]	0	u(8)
num_applicable_ops_minus1[i]	0	uc(v)
for(j = 0; j <= num_applicable_ops_minus1[i]; j++) {		
applicable_op_temporal_id[i][j]	0	u(3)
applicable_op_num_target_views_minus1[i][j]	0	uc(v)
for(k = 0; k <=		
applicable_op_num_target_views_minus1[i][j]; k++)		
applicable_op_target_view_id[i][j][k]	0	uc(v)
applicable_op_num_views_minus1[i][j]	0	uc(v)
applicable_op_num_views_depth_minus1[i][j]	0	uc(v)
}		
}		
...		
}		

[0272] En la tabla anterior, `applicable_op_num_views_minus1[i][j]` más 1 especifica el número de vistas de textura requeridas para descodificar las vistas de salida de destino correspondientes al j-ésimo punto de funcionamiento al que se aplica el nivel indicado por `level_idc[i]`. `Applicable_op_num_views_depth_minus1[i][j]` más 1 especifica el número de vistas de profundidad requeridas para descodificar las vistas de salida de destino correspondientes al j-ésimo punto de funcionamiento al que se aplica el nivel indicado por `level_idc[i]`. De forma alternativa, `applicable_op_num_views_depth_minus1[i][j]` no se añade, pero se cambia la semántica de `applicable_op_num_views_minus1[i][j]`. `Applicable_op_num_views_minus1[i][j]` más 1 especifica el número de vistas de textura más el número de vistas de profundidad requeridas para descodificar las vistas de salida de destino correspondientes al j-ésimo punto de funcionamiento al que se aplica el nivel indicado por `level_idc[i]`.

[0273] En algunos casos, las técnicas de la presente divulgación proporcionan la señalización, en un conjunto de parámetros de secuencia, de un número de vistas de textura que se van a descodificar para un punto de funcionamiento y un número de vistas de profundidad que se van a descodificar para el punto de funcionamiento. La señalización del número de vistas de textura y el número de vistas de profundidad puede comprender generar, en el conjunto de parámetros de secuencia, un elemento sintáctico que indica un número de vistas que se van a descodificar para el punto de funcionamiento, teniendo cada una de las vistas una de las vistas de textura y una de las vistas de profundidad. La señalización del número de vistas de textura y el número de vistas de profundidad puede comprender generar, en el conjunto de parámetros de secuencia, un primer elemento sintáctico y un segundo elemento sintáctico, especificando el primer elemento sintáctico el número de vistas de textura que se van a descodificar para el punto de funcionamiento, y especificando el segundo elemento sintáctico el número de vistas de profundidad que se va a descodificar para el punto de funcionamiento. La señalización del número de vistas de textura y el número de vistas de profundidad puede comprender generar, en el conjunto de parámetros de secuencia, un elemento sintáctico que indica el número de vistas de textura que se van a descodificar para el punto de funcionamiento más el número de vistas de profundidad que se van a descodificar para la punto de funcionamiento.

[0274] En algunos ejemplos, las técnicas de la presente divulgación pueden proporcionar la recepción de un conjunto de parámetros de secuencia que señala un número de vistas de textura que se van a descodificar para un punto de funcionamiento y un número de vistas de profundidad que se van a descodificar para el punto de funcionamiento. En dichos ejemplos, el conjunto de parámetros de secuencia puede incluir un elemento sintáctico que indica un número de vistas que se van a descodificar para el punto de funcionamiento, teniendo cada una de las vistas una de las vistas de textura y una de las vistas de profundidad. El conjunto de parámetros de secuencia puede incluir un primer elemento sintáctico y un segundo elemento sintáctico, especificando el primer elemento sintáctico el número de vistas de textura que se van a descodificar para el punto de funcionamiento, y especificando el segundo elemento sintáctico el número de vistas de profundidad que se van a descodificar para el punto de funcionamiento. El conjunto de parámetros de secuencia puede incluir un elemento sintáctico que indica el número de vistas de textura que se van a descodificar para el punto de funcionamiento más el número de vistas de profundidad que se van a descodificar para el punto de funcionamiento.

[0275] Si bien se describen anteriormente diversos aspectos de las técnicas como que se llevan a cabo por la unidad de encapsulación 21, las técnicas se pueden llevar a cabo de la misma forma o de forma recíproca por la unidad de desencapsulación 29. Además, diversos aspectos de las técnicas descritas anteriormente como que se llevan a cabo por la unidad de encapsulación 21 también se pueden llevar a cabo, en otras implementaciones, por el codificador de vídeo 20 (llevando a cabo el descodificador de vídeo 30 las mismas operaciones u operaciones recíprocas que las que lleva a cabo el codificador de vídeo 20). Por tanto, si bien se muestra una implementación particular en los ejemplos de la FIG. 1-3, las técnicas no se deben limitar a este ejemplo, sino que se pueden aplicar a otras implementaciones que proporcionan un tratamiento separado de los componentes de textura y profundidad de un componente de vista especificado por datos de vídeo de múltiples vistas.

[¿Debería haber una fig. de desencapsulación que sea recíproca a la fig. 4? ¿O no es necesario?]

[0276] La FIG. 5 es un diagrama conceptual que ilustra un patrón de predicción de MVC de ejemplo. La siguiente descripción de MVC de nuevo se refiere a 3DVC en el sentido de que 3DVC incorpora MVC más profundidad. En consecuencia, la descripción de MVC y las diversas predicciones proporcionadas por MVC se describen para proporcionar un contexto para entender 3DVC y las técnicas descritas en la presente divulgación que se refieren a 3DVC.

[0277] En el ejemplo de la FIG. 5, se ilustran ocho vistas y se ilustran doce ubicaciones temporales para cada vista. En general, cada fila en la FIG. 5 corresponde a una vista, mientras que cada columna indica una ubicación temporal. Cada una de las vistas se puede identificar usando un identificador de vista ("view_id"), que se puede usar para indicar una ubicación de cámara relativa con respecto a las otras vistas. En el ejemplo mostrado en la FIG. 5, los ID de vista se indican como de "S0" a "S7", aunque también se pueden usar ID de vista numéricos. Además, cada una de las ubicaciones temporales se puede identificar usando un valor de recuento de orden de imágenes (POC), que indica un orden de visualización de las imágenes. En el ejemplo mostrado en la FIG. 4, los valores de POC se indican como de "T0" a "T11".

[0278] Aunque MVC tiene una llamada vista base que se puede descodificar por descodificadores H.264/AVC y el par de vistas estéreo se puede admitir por MVC, MVC puede admitir más de dos vistas como entrada de vídeo 3D. En consecuencia, un renderizador de un cliente que tiene un descodificador de MVC puede esperar contenido de vídeo 3D con múltiples vistas.

[0279] Las imágenes de la FIG. 5 se indican usando un bloque sombreado que incluye una letra, que designa si la imagen correspondiente está intracodificada (es decir, una trama I), o intercodificada en una dirección (es decir, como una trama P) o en múltiples direcciones (es decir, como una trama B). En general, las predicciones se indican mediante flechas, donde la imagen a la que se apunta usa el objeto desde el que se apunta como referencia de predicción. Por ejemplo, la trama P de la vista S2 en la ubicación temporal T0 se predice a partir de la trama I de la vista S0 en la ubicación temporal T0. Cada una de las imágenes mostradas en la FIG. 5 se puede denominar un componente de vista.

[0280] Como con la codificación de vídeo de vista única, las imágenes de una secuencia de vídeo de múltiples vistas se pueden codificar predictivamente con respecto a las imágenes en diferentes ubicaciones temporales. Por ejemplo, la trama b de la vista S0 en la ubicación temporal T1 tiene una flecha apuntando a la misma desde la trama I desde la vista S0 en la ubicación temporal T0, que indica que la trama b se predice a partir de la trama I. Adicionalmente, sin embargo, en el contexto de la codificación de vídeo de múltiples vistas, las imágenes se pueden predecir entre vistas. Es decir, un componente de vista puede usar los componentes de vista de otras vistas como referencia. En la MVC, por ejemplo, la predicción entre vistas se realiza como si el componente de vista de otra vista es una referencia de predicción inter. Las posibles referencias entre vistas se pueden señalar en la extensión SPS MVC y se pueden modificar por el proceso de construcción de la lista de imágenes de referencia, que permite el ordenamiento flexible de las referencias de predicción inter o de predicción entre vistas.

[0281] La FIG. 5 proporciona diversos ejemplos de predicción entre vistas. Las imágenes de la vista S1, en el ejemplo de la FIG. 5, se ilustran como que se predicen a partir de imágenes en diferentes ubicaciones temporales de la vista S1, así como entre vistas predichas a partir de imágenes de las imágenes de vistas S0 y S2 en las mismas ubicaciones temporales. Por ejemplo, la trama b de la vista S1 en la ubicación temporal T1 se predice a partir de cada una de las tramas B de la vista S1 en las ubicaciones temporales T0 y T2, así como las tramas b de las vistas S0 y S2 en la ubicación temporal T1.

[0282] En el ejemplo de la FIG. 5, la "B" mayúscula y la "b" minúscula están destinadas a indicar diferentes relaciones jerárquicas entre imágenes, en lugar de diferentes metodologías de codificación. En general, las tramas con "B" mayúscula están relativamente más altas en la jerarquía de predicción que las tramas con "b" minúscula. La FIG. 5 también ilustra las variaciones en la jerarquía de predicción usando diferentes niveles de sombreado, donde las imágenes con una mayor cantidad de sombreado (es decir, relativamente más oscuras) están más altas en la jerarquía de predicción que aquellas imágenes que tienen menos sombreado (es decir, relativamente más claras). Por ejemplo, todas las tramas I en la FIG. 5 se ilustran con sombreado completo, mientras que las tramas P tienen un sombreado algo más claro, y las tramas B (y las tramas b minúscula) tienen diversos niveles de sombreado entre sí, pero siempre más claro que el sombreado de las tramas P y las tramas I.

[0283] En general, la jerarquía de predicción se refiere a los índices de orden de vista, en cuanto a que las imágenes relativamente más altas en la jerarquía de predicción se deberían descodificar antes de descodificar imágenes que están relativamente más bajas en la jerarquía, de modo que esas imágenes relativamente más altas en la jerarquía se puedan usar como imágenes de referencia durante la descodificación de las imágenes relativamente más bajas en la jerarquía. Un índice de orden de vista es un índice que indica el orden de descodificación de componentes de vista en una unidad de acceso. Los índices de orden de vista pueden estar implícitos en un conjunto de parámetros, tal como un SPS.

[0284] De esta forma, las imágenes usadas como imágenes de referencia se pueden descodificar antes de descodificar las imágenes que se codifican con referencia a las imágenes de referencia. Un índice de orden de vista es un índice que indica el orden de descodificación de componentes de vista en una unidad de acceso. De acuerdo con MVC/AVC, para cada índice *i* de orden de vista, se señala el id de vista correspondiente. La descodificación de los componentes de vista sigue el orden ascendente de los índices de orden de vista. Si se presentan todas las vistas, entonces el conjunto de índices de orden de vista comprende un conjunto consecutivamente ordenado de cero a uno menos que el número total de vistas.

[0285] En algunos casos, se puede extraer un subconjunto de un flujo de bits completo para formar un subflujo de bits que todavía se ajuste a MVC. Existen muchos posibles subflujos de bits que las aplicaciones específicas pueden requerir, basándose, por ejemplo, en un servicio proporcionado por un servidor, la capacidad, el soporte y las capacidades de los descodificadores de uno o más clientes y/o la preferencia de uno o más clientes. Por ejemplo, un cliente podría requerir solo tres vistas, y podría haber dos entornos. En un ejemplo, un cliente puede requerir una experiencia visual sin fisuras y podría preferir vistas con valores de *view_id* S0, S1 y S2, mientras que otro cliente puede requerir escalabilidad de las vistas y preferir vistas con valores de *view_id* S0, S2 y S4. Ambos de estos subflujos de bits se pueden descodificar como flujos de bits de MVC independientes y se pueden admitir simultáneamente.

[0286] Si bien la FIG. 5 muestra ocho vistas (S0-S7), como se indica anteriormente, la extensión MVC/AVC admite hasta 1024 vistas y usa un *view_id* en una cabecera de unidad NAL para identificar la vista a la que pertenece una unidad NAL. Un índice de orden de vista se puede señalar como parte de una cabecera de unidad NAL. Es decir, para propósitos de comparación, un índice de orden de vista puede reemplazar el *view_id* que se señala en la cabecera de unidad NAL de la extensión MVC/AVC. El orden de vista en general describe el ordenamiento de las vistas en una unidad de acceso, y un índice de orden de vista identifica una vista particular en el orden de vista de la unidad de acceso. Es decir, un índice de orden de vista describe el orden de descodificación de un componente de vista correspondiente de una unidad de acceso.

[0287] En consecuencia, un SPS puede proporcionar una relación entre *view_ids* para las vistas y los índices de orden de vista para las vistas. Usando el índice de orden de vista y los datos en el SPS, el codificador de vídeo 20 y el descodificador de vídeo 30 pueden reemplazar el id de vista de 10 bits de MVC/AVC en la cabecera de unidad NAL por el índice de orden de vista, que puede dar lugar a un ahorro de bits con respecto al esquema MVC/AVC.

[0288] Un SPS de ejemplo que proporciona una relación entre *view_ids* para las vistas y los índices de orden de vista se proporciona en la tabla a continuación:

EXTENSIÓN MVC DEL CONJUNTO DE PARÁMETROS DE SECUENCIA

seq_parameter_set_mvc_extension() {	C	Descriptor
num_views_minus1	0	ue(v)
for(i = 0; i <= num_views_minus1; i++) {		
view_id[i]	0	ue(v)
view_level[i]	0	ue(v)
}		
for(i = 1; i <= num_views_minus1; i++) {		
num_ref_views[i]	0	ue(v)
for(j = 0; j < num_ref_views[i]; j++)		
ref_view_idx[i][j]	0	ue(v)
}		
num_level_values_signalled_minus 1	0	ue(v)
for(i = 0; i <= num_level_values_signalled_minus 1; i++) {		
level_idc[i]	0	u(8)

	C	Descriptor
seq_parameter_set_mvc_extension() {		
num_applicable_ops_minus1[i]	0	ue(v)
for(j = 0; i <= num_applicable_ops_minus1[i]; j++) {		
applicable_op_temporal_id[i][j]	0	u(3)
applicable_op_num_target_views_minus1[i][j]	0	ue(v)
for(k = 0; k <= applicable_op_num_target_views_minus1[i][j]; k++		
)		
applicable_op_target_view_idx[i][j][k]	0	ue(v)
applicable_op_num_views_minus1[i][j]	0	ue(v)
}		
}		
}		

[0289] El SPS mostrado en la Tabla anterior especifica las relaciones de dependencia entre vistas para una secuencia de vídeo codificada. El SPS también especifica valores de nivel para un subconjunto de los puntos de funcionamiento para la secuencia de vídeo codificada. Todos los SPS a los que hace referencia una secuencia de vídeo codificada deben ser idénticos. Sin embargo, algunas vistas identificadas por view_id[i] pueden no estar presentes en la secuencia de vídeo codificada. Además, algunas vistas o subconjuntos temporales descritos por el SPS se pueden haber eliminado de la secuencia de vídeo codificada original, y por tanto pueden no estar presentes en la secuencia de vídeo codificada. Sin embargo, la información en el SPS, siempre se puede aplicar a las vistas y los subconjuntos temporales restantes.

[0290] En la tabla anterior, el elemento sintáctico num_views_minus1 más 1 especifica el número máximo de vistas codificadas en la secuencia de vídeo codificada. El valor de num_view_minus1 puede estar en el intervalo de 0 a 31, inclusive. En algunos casos, el número real de vistas en la secuencia de vídeo codificada puede ser menor que num_views_minus1 más 1. El elemento view_id[i] especifica el identificador de vista de la vista con un índice de orden de vista igual a i. El elemento view_level[i] especifica el view_level de la vista con un índice de orden de vista igual a i. En algunos ejemplos, todos los componentes de vista con un view_level de hasta un valor predefinido (VL) se pueden descodificar sin descodificar ningún componente de vista con un nivel de vista más grande que VL.

[0291] El elemento num_ref_views[i] especifica el número de componentes de vista para la predicción entre vistas en la lista de imágenes de referencia inicial RefPicList0 y RefPicList1 al descodificar componentes de vista con un índice de orden de vista igual a i. El valor del elemento num_ref_views[i] no puede ser mayor que Min(15, num_views_minus1). El valor de num_ref_views[0] puede ser igual a 0. Además, el elemento ref_view_idx[i][j] puede especificar el índice de orden de vista del j-ésimo componente de vista para la predicción entre vistas en la lista de imágenes de referencia inicial RefPicList0 y RefPicList1 al descodificar un componente de vista con índice de orden de vista igual a i. El valor de ref_view_idx[i][j] puede estar en el intervalo de 0 a 31, inclusive.

[0292] La FIG. 6 es un diagrama de flujo que ilustra el funcionamiento de un dispositivo de codificación de vídeo en la implementación de la activación del conjunto de parámetros de acuerdo con diversos aspectos de las técnicas descritas en la presente divulgación. Inicialmente, un dispositivo de codificación de vídeo, tal como el descodificador de vídeo 30, puede determinar un valor de índice de orden asignado a un componente de vista de datos de vídeo de múltiples vistas (130). Este valor de índice de orden se describe anteriormente con más detalle con respecto a la FIG. 5. Como se describe anteriormente, el descodificador de vídeo 30, a continuación, puede activar un conjunto de parámetros como un conjunto de parámetros de textura para el componente de textura del componente de vista basándose al menos en el valor de índice de orden de vista determinado asignado al componente de vista de los datos de vídeo de múltiples vistas (132). Después de activar este conjunto de parámetros de textura, el codificador de vídeo 30, a continuación, puede codificar el componente de textura del componente de vista basándose en el conjunto de parámetros de textura activado (134).

[0293] En algunos casos, el descodificador de vídeo 30 codifica sólo el componente de textura del componente de vista y no el componente de profundidad del componente de vista basándose en el conjunto de parámetros de textura activado. En algunos casos, el descodificador de vídeo 30 activa otro conjunto de parámetros para el componente de profundidad del componente de vista basándose al menos en el valor de índice de orden de vista asignado al componente de vista de los datos de vídeo de múltiples vistas y descodifica el componente de profundidad del componente de vista basándose en el conjunto de parámetros activado para el componente de profundidad del componente de vista y no el conjunto de parámetros de textura activado de la forma descrita anteriormente. Por tanto, si bien no se muestra explícitamente en el ejemplo de la FIG. 6, se pueden activar individualmente conjuntos de parámetros adicionales para el componente de profundidad, que el descodificador de vídeo 30 puede usar, a continuación, para descodificar el componente de profundidad pero no el componente de textura.

5 **[0294]** En algunos de los casos descritos anteriormente, el descodificador de vídeo 30 puede activar un conjunto de parámetros de secuencia para codificar el componente de vista de los datos de vídeo de múltiples vistas basándose al menos en el valor de índice de orden de vista asignado al componente de vista de los datos de vídeo de múltiples vistas. Además, como se indica anteriormente, el descodificador de vídeo 30 puede activar un conjunto de parámetros de secuencia como el conjunto de parámetros de secuencia de textura para el componente de textura del componente de vista basándose al menos en el valor de índice de orden de vista asignado al componente de vista de los datos de vídeo de múltiples vistas de modo que uno o más elementos sintácticos especificados por el conjunto de parámetros de secuencia de textura anulen o bien aumenten uno o más elementos sintácticos especificados por el conjunto de parámetros de secuencia activado.

15 **[0295]** En algunos de los ejemplos descritos anteriormente, el descodificador de vídeo 30 puede activar un conjunto de parámetros de imagen como el conjunto de parámetros de imagen de textura para el componente de textura del componente de vista basándose al menos en el valor de índice de orden de vista asignado al componente de vista de los datos de vídeo de múltiples vistas de modo que uno o más elementos sintácticos especificados por el conjunto de parámetros de imagen de textura anulen o bien aumenten uno o más elementos sintácticos especificados por el conjunto de parámetros de secuencia activado.

20 **[0296]** En algunos de los ejemplos descritos anteriormente, el descodificador de vídeo 30 puede activar un conjunto de parámetros de imagen como un conjunto de parámetros de imagen de vista para el componente de vista de los datos de vídeo de múltiples vistas basándose al menos en el valor de índice de orden de vista asignado al componente de vista de los datos de vídeo de múltiples vistas. Adicionalmente, el descodificador de vídeo 30, a continuación, puede activar el conjunto de parámetros como el conjunto de parámetros de textura activando un conjunto de parámetros de imagen como el conjunto de parámetros de imagen de textura para el componente de textura del componente de vista basándose al menos en el valor de índice de orden de vista asignado al componente de vista de los datos de vídeo de múltiples vistas de modo que uno o más elementos sintácticos especificados por el conjunto de parámetros de imagen de textura anulen o bien aumenten uno o más elementos sintácticos especificados por el conjunto de parámetros de imagen de vista activado.

30 **[0297]** En algunos ejemplos, el descodificador de vídeo 30, a continuación, puede activar el conjunto de parámetros como el conjunto de parámetros de textura activando uno o más de un conjunto de parámetros de secuencia y un parámetro de imagen como un conjunto de parámetros de secuencia de textura y un conjunto de parámetros de imagen de textura, respectivamente, para el componente de textura del componente de vista basándose al menos en el valor de índice de orden de vista asignado al componente de vista de los datos de vídeo de múltiples vistas.

35 **[0298]** Se debe entender también que las etapas mostradas y descritas con respecto a la FIG. 6 se proporcionan simplemente como un ejemplo. Es decir, no es necesario que las etapas mostradas en el ejemplo de la FIG. 6 se lleven a cabo en el orden mostrado en la FIG. 6, y se pueden llevar a cabo menos etapas, o etapas adicionales o alternativas. Además, si bien las técnicas se describen anteriormente con respecto a un codificador de vídeo particular, es decir, el descodificador de vídeo 30, las técnicas se pueden implementar por cualquier codificador de vídeo, incluyendo el codificador de vídeo 20.

40 **[0299]** La FIG. 7 es un diagrama de flujo que ilustra el funcionamiento de ejemplo del procesamiento de datos de vídeo de múltiples vistas para generar mensajes SEI anidados de acuerdo con las técnicas descritas en la presente divulgación. Si bien las técnicas se describen con respecto a una unidad particular a continuación, es decir, la unidad de encapsulación 21, en general se pueden implementar por cualquier dispositivo de codificación de vídeo, incluyendo el dispositivo de destino 14 y más en particular la unidad de desencapsulación 29 del dispositivo de destino 14.

50 **[0300]** Inicialmente, la unidad de encapsulación 21 recibe datos de vídeo de múltiples vistas codificados y solicita la unidad de generación de mensajes SEI 108, que determina que un mensaje de información de mejora complementaria (SEI) se aplique al procesar el componente de vista de los datos de vídeo de múltiples vistas de la forma descrita anteriormente (140). La unidad de generación de mensajes SEI 108, como se describe anteriormente, también puede determinar si un mensaje SEI anidado se aplica tanto a los componentes de textura como de profundidad del componente de vista o solo al componente de profundidad del componente de vista (141). Tras determinar que el mensaje SEI anidado se aplica solo al componente de profundidad ("NO" 141), la unidad de generación de mensajes SEI 108 puede determinar que un mensaje de información de mejora complementaria anidado se aplique solo al componente de profundidad del componente de vista además del mensaje de información de mejora complementaria (142). Sin embargo, tras determinar que el mensaje SEI anidado se aplica tanto a los componentes de profundidad como de textura ("SÍ" 141), la unidad de generación de mensajes SEI 108 puede determinar que un mensaje de información de mejora complementaria anidado se aplique tanto al componente de textura como al componente de profundidad del componente de vista además del mensaje de información de mejora complementaria (143). La unidad de creación de archivos de vídeo 104, a continuación, puede generar un archivo de vídeo que incluya tanto el mensaje SEI como el mensaje SEI anidado (144).

65 **[0301]** En algunos casos, la unidad de encapsulación 21 genera este archivo de vídeo y proporciona este archivo de vídeo al enrutador 36 u otro dispositivo, tal como un servidor de vídeo de la red de entrega de contenido 34 (ambos

de los cuales se describen anteriormente con respecto al ejemplo de la FIG. 1). A continuación, este dispositivo puede procesar el componente de profundidad del componente de vista basándose en el mensaje de información de mejora complementaria y el mensaje de información de mejora complementaria anidado.

5 **[0302]** En algunos ejemplos, la unidad de encapsulación 21 determina que el mensaje de información de mejora complementaria anidado se aplique además del mensaje de información de mejora complementaria original, al codificar sólo el componente de profundidad del componente de vista. En este caso, el enrutador 36, a continuación, puede procesar solo el componente de profundidad del componente de vista y no el componente de textura del componente de vista basándose en el mensaje de información de mejora complementaria original y el mensaje de información de mejora complementaria anidado.

15 **[0303]** Como se indica anteriormente, la unidad de encapsulación 21 puede determinar si el mensaje de información de mejora complementaria anidado se aplica además del mensaje de información de mejora complementaria al codificar el componente de textura del componente de vista. Como resultado, el enrutador 36, en este caso, puede procesar el componente de textura del componente de vista basándose en el mensaje de información de mejora complementaria y el mensaje de información de mejora complementaria anidado basándose en la determinación de si el mensaje de información de mejora complementaria anidado se aplica además del mensaje de información de mejora complementaria al codificar el componente de textura del componente de vista.

20 **[0304]** Al determinar que el mensaje de mejora complementaria anidado se aplica además del mensaje de información de mejora complementaria al procesar el componente de textura del componente de vista, el enrutador 36 puede identificar un indicador en el mensaje de información de mejora complementaria anidado que especifica si se aplica o no el mensaje de información de mejora complementaria anidado solo al componente de profundidad o tanto al componente de vista como al componente de profundidad y determina que el mensaje de mejora complementaria anidado se aplica además del mensaje de información de mejora complementaria cuando se codifica el componente de textura del componente de vista basándose en el indicador identificado. El indicador puede comprender uno o más de un indicador de aplicabilidad de vista de información de mejora complementaria y un indicador de aplicabilidad de vista operacional de información de mejora complementaria.

30 **[0305]** En algunos casos, al determinar si el mensaje de información de mejora complementaria anidado se aplica además del mensaje de información de mejora complementaria al procesar el componente de textura del componente de vista, el enrutador 36 puede determinar que se aplique el mensaje de información de mejora complementaria anidado, además del mensaje de información de mejora complementaria, al procesar tanto el componente de profundidad del componente de vista como el componente de textura del componente de vista. El enrutador 36, en este caso, puede procesar el componente de textura del componente de vista basándose en el mensaje de información de mejora complementaria y el mensaje de información de mejora complementaria anidado.

40 **[0306]** En algunos casos, al determinar si el mensaje de información de mejora complementaria anidado se aplica además del mensaje de información de mejora complementaria al codificar el componente de textura del componente de vista, el enrutador 36 puede determinar que el mensaje de información de mejora complementaria anidado sólo se aplique además del mensaje de información de mejora complementaria al codificar el componente de profundidad del componente de vista y no el componente de textura del componente de vista. Como resultado, el enrutador 36 puede procesar el componente de textura del componente de vista basándose en el mensaje de información de mejora complementaria y no basándose en el mensaje de información de mejora complementaria anidado.

45 **[0307]** También se debe entender que la FIG. 7 simplemente proporciona un ejemplo. Es decir, no es necesario que las etapas mostradas en el ejemplo de la FIG. 7 se lleven a cabo en el orden mostrado en la FIG. 7, y se pueden llevar a cabo menos etapas, o etapas adicionales o alternativas.

50 **[0308]** La FIG. 8 es un diagrama de flujo que ilustra el funcionamiento de ejemplo de un dispositivo de codificación de vídeo en la eliminación por separado de los componentes de textura y profundidad de una memoria intermedia de imágenes descodificadas de acuerdo con las técnicas descritas en la presente divulgación. El ejemplo mostrado en la FIG. 8 se describe en general como que se lleva a cabo por el descodificador de vídeo 30 (FIG. 1 y 3). Sin embargo, se debe entender que el proceso descrito con respecto a la FIG. 8 se puede realizar por una variedad de otros procesadores, unidades de procesamiento, unidades de codificación basadas en hardware tales como codificadores/descodificadores (CÓDEC) y similares.

60 **[0309]** En general, el descodificador de vídeo 30 puede almacenar un componente de profundidad y un componente de textura de un componente de vista (especificado por los datos de vídeo de múltiples vistas) en una memoria intermedia de imágenes descodificadas, que se muestra como memoria de imágenes de referencia 92 en el ejemplo de la FIG. 3 (150). El descodificador de vídeo 30, a continuación, puede analizar una dependencia de vista para determinar si el componente de profundidad se usa para la predicción entre vistas (donde dicha dependencia de vista se puede reflejar anteriormente con respecto al ejemplo de la FIG. 5; 152). El descodificador de vídeo 30 puede analizar esta dependencia de vista solo para el componente de profundidad y analizar por separado una dependencia de vista al intentar eliminar el componente de textura asociado de la memoria de imágenes de referencia 92. En algunos casos, la dependencia de vista se señala en una extensión de conjunto de parámetros de secuencia de

codificación de vídeo de múltiples vistas de un conjunto de parámetros de secuencia de subconjunto. El conjunto de parámetros de secuencia de subconjunto puede contener un perfil de vídeo tridimensional y se activa como un conjunto de parámetros de secuencia de codificación de vídeo de múltiples vistas de vista activo al analizar la dependencia de vista. En cualquier caso, el descodificador de vídeo 30, a continuación, puede eliminar el componente de profundidad de la memoria de imágenes de referencia 92 sin eliminar el componente de textura en respuesta a la determinación de que el componente de profundidad no se usa para la predicción entre vistas (154).

[0310] Antes de eliminar este componente de profundidad, el descodificador de vídeo 30 puede calificar el componente de profundidad como elegible para su eliminación al determinar que el componente de profundidad no pertenece a una vista de salida de destino y está asociado con un código de identificación de referencia de capa de abstracción de red que tiene un valor igual a cero. Como resultado de proporcionar potencialmente el tratamiento separado de los componentes de profundidad y textura, el descodificador de vídeo 30 puede determinar una o más imágenes de referencia para el componente de profundidad que son diferentes del componente de textura. Es decir, el descodificador de vídeo 30 puede determinar la una o más imágenes de referencia para el componente de profundidad, al menos, determinando una lista de imágenes de referencia que identifica la una o más imágenes de referencia para el componente de profundidad y determinando una lista de imágenes de referencia que identifica una o más imágenes de referencia adicionales para el componente de textura, en el que la una o más imágenes de referencia adicionales determinadas para el componente de textura son diferentes de la una o más imágenes de referencia determinadas para el componente de profundidad.

[0311] En la determinación de la lista de imágenes de referencia para el componente de profundidad, el descodificador de vídeo 30 puede determinar la lista de imágenes de referencia para el componente de profundidad y el componente de textura basándose en marcas de los componentes de profundidad y textura por el codificador de vídeo 20, que se usan para identificar la lista de imágenes de referencia para el componente de profundidad por separado del marcado del componente de textura. En algunos casos, se pueden usar uno o más elementos sintácticos para marcar por separado el componente de profundidad con la lista de imágenes de referencia determinada para el componente de profundidad y se usan uno o más elementos sintácticos para marcar por separado el componente de textura con la lista de imágenes de referencia determinada para el componente de textura.

[0312] En algunos casos, el descodificador de vídeo 30 puede determinar que el componente de profundidad y el componente de textura pertenecen a una vista de salida de destino que se va a emitir para su visualización y simultáneamente o casi simultáneamente emitir el componente de profundidad y el componente de textura.

[0313] Se debe entender que las etapas mostradas y descritas con respecto al ejemplo de la FIG. 8 se proporcionan simplemente como un ejemplo. Es decir, no es necesario que las etapas mostradas en el ejemplo de la FIG. 8 se lleven a cabo en el orden mostrado en la FIG. 8, y se pueden llevar a cabo menos etapas, o etapas adicionales o alternativas.

[0314] La FIG. 9 es un diagrama de flujo que ilustra el funcionamiento de ejemplo de un dispositivo de codificación de vídeo en la determinación de la información de nivel de secuencia para 3DVC compatible con MVC de acuerdo con las técnicas descritas en la presente divulgación. El ejemplo mostrado en la FIG. 8 se describe en general como que se lleva a cabo por el descodificador de vídeo 30 (FIG. 1 y 3). Sin embargo, se debe entender que el proceso descrito con respecto a la FIG. 8 se puede realizar por una variedad de otros procesadores, unidades de procesamiento, unidades de codificación basadas en hardware tales como codificadores/descodificadores (CÓDEC) y similares.

[0315] Como se describe anteriormente, el descodificador de vídeo 30 puede determinar la primera información de nivel de secuencia que describe características de los componentes de profundidad y determinar la segunda información de nivel de secuencia que describe características de un punto de funcionamiento de los datos de vídeo (160, 162). Además, el descodificador de vídeo 30 puede determinar una lista de imágenes de referencia que identifica una o más imágenes de referencia para los componentes de profundidad indicados en el conjunto de parámetros de secuencia de codificación de vídeo tridimensional.

[0316] En algunos ejemplos, la segunda información de nivel de secuencia incluye un conjunto de parámetros de secuencia de codificación de vídeo tridimensional que describe, para el punto de funcionamiento, una lista de vistas de salida de destino, un número de vistas de textura que se van a descodificar al descodificar el punto de funcionamiento y un número de vistas de profundidad que se van a descodificar al descodificar el punto de funcionamiento, donde la serie de vistas de textura que se van a descodificar es diferente de la serie de vistas de profundidad. En algunos ejemplos, el descodificador de vídeo 30 puede seleccionar además, para cada una de las vistas de salida de destino especificadas en la lista de vistas de salida de destino, el uno o más componentes de profundidad cuando estén disponibles.

[0317] En algunos ejemplos, el descodificador de vídeo 30 puede especificar un mensaje SEI de información de escalabilidad de codificación de vídeo tridimensional como parte de los datos de vídeo, en el que el mensaje SEI de información de escalabilidad de codificación de vídeo tridimensional incluye una descripción del punto de funcionamiento sin los componentes de profundidad o una descripción del punto de funcionamiento con los componentes de profundidad. En algunos casos, el mensaje SEI incluye una indicación de si el punto de

funcionamiento incluye los componentes de profundidad. Además, en algunos ejemplos, el mensaje SEI incluye una indicación de si el punto de funcionamiento incluye los componentes de profundidad y las vistas de profundidad directamente dependientes con vistas de textura directamente dependientes para la descodificación del punto de funcionamiento.

5 **[0318]** Se debe entender que las etapas mostradas y descritas con respecto al ejemplo de la FIG. 9 se proporcionan simplemente como un ejemplo. Es decir, no es necesario que las etapas mostradas en el ejemplo de la FIG. 9 se lleven a cabo en el orden mostrado en la FIG. 9, y se pueden llevar a cabo menos etapas, o etapas adicionales o alternativas.

10 **[0319]** Si bien se ha proporcionado nombres de ejemplo a determinados elementos sintácticos descritos con respecto a la presente divulgación para propósitos explicativos, se debe entender que los conceptos descritos en la presente divulgación son más aplicables en general a cualquier elemento sintáctico, independientemente del nombre. Por ejemplo, si bien determinados aspectos se refieren a un "índice de orden de vista", "view_order_index" o "view_idx", se debe entender que se le puede dar a dicho elemento sintáctico un nombre alternativo en una futura norma de codificación.

15 **[0320]** Si bien se describen determinadas técnicas de la presente divulgación con respecto a la norma H.264, se debe entender que las técnicas no se limitan necesariamente a una norma de codificación particular. Es decir, las técnicas se refieren más en general a lograr eficacias de codificación en 3DVC, por ejemplo, a través de unidades NAL y conjuntos de parámetros más cortos y/o menos complejos, como se describe anteriormente.

20 **[0321]** Se debe entender que, dependiendo del ejemplo, determinados actos o acontecimientos de cualquiera de los procedimientos descritos en el presente documento se pueden llevar a cabo en una secuencia diferente, se pueden añadir, fundir u omitir por completo (por ejemplo, no todos los actos o acontecimientos descritos son necesarios para la puesta en práctica del procedimiento). Además, en determinados ejemplos, los actos o acontecimientos se pueden llevar a cabo simultáneamente, por ejemplo, a través de procesamiento de múltiples hilos, procesamiento de interrupciones o múltiples procesadores, en lugar de secuencialmente. Además, si bien determinados aspectos de la presente divulgación se describen como que se llevan a cabo por un único módulo o unidad para propósitos de claridad, se debe entender que las técnicas de la presente divulgación se pueden llevar a cabo por una combinación de unidades o módulos asociados con un codificador de vídeo.

25 **[0322]** En uno o más ejemplos, las funciones descritas se pueden implementar en hardware, software, firmware o cualquier combinación de los mismos. Si se implementan en software, las funciones se pueden almacenar en o transmitir por un medio legible por ordenador, como una o más instrucciones o código, y ejecutar por una unidad de procesamiento basada en hardware. Los medios legibles por ordenador pueden incluir medios de almacenamiento legibles por ordenador, que correspondan a un medio tangible tal como medios de almacenamiento de datos o medios de comunicación que incluyan cualquier medio que facilite la transferencia de un programa informático desde un lugar a otro, por ejemplo, de acuerdo a un protocolo de comunicación.

30 **[0323]** De esta forma, los medios legibles por ordenador pueden corresponder en general a (1) medios de almacenamiento legibles por ordenador tangibles que sean no transitorios o (2) un medio de comunicación tal como una señal u onda portadora. Los medios de almacenamiento de datos pueden ser cualquier medio disponible al que se pueda acceder desde uno o más ordenadores o uno o más procesadores para recuperar instrucciones, código y/o estructuras de datos para la implementación de las técnicas descritas en la presente divulgación. Un producto de programa informático puede incluir un medio legible por ordenador.

35 **[0324]** A modo de ejemplo, y no de limitación, dichos medios de almacenamiento legibles por ordenador pueden comprender RAM, ROM, EEPROM, CD-ROM u otro almacenamiento de disco óptico, almacenamiento de disco magnético u otros dispositivos de almacenamiento magnético, memoria flash o cualquier otro medio que se pueda usar para almacenar un código de programa deseado en forma de instrucciones o estructuras de datos y al que se pueda acceder por un ordenador. Además, cualquier conexión recibe debidamente la denominación de medio legible por ordenador. Por ejemplo, si las instrucciones se transmiten desde un sitio web, un servidor u otro origen remoto usando un cable coaxial, un cable de fibra óptica, un par trenzado, una línea de abonado digital (DSL) o tecnologías inalámbricas tales como infrarrojos, radio y microondas, entonces el cable coaxial, el cable de fibra óptica, el par trenzado, la DSL o las tecnologías inalámbricas tales como infrarrojos, radio y microondas se incluyen en la definición de medio.

40 **[0325]** Sin embargo, se debe entender que los medios de almacenamiento legibles por ordenador y los medios de almacenamiento de datos no incluyen conexiones, ondas portadoras, señales u otros medios transitorios, sino que, en cambio, se dirigen a medios de almacenamiento tangibles no transitorios. Los discos, como se usa en el presente documento, incluyen disco compacto (CD), disco láser, disco óptico, disco versátil digital (DVD), disco flexible y disco Blu-ray, donde algunos discos reproducen normalmente los datos magnéticamente, mientras que otros discos reproducen los datos ópticamente con láseres. Las combinaciones de lo anterior también se deben incluir dentro del alcance de los medios legibles por ordenador.

5 **[0326]** Las instrucciones se pueden ejecutar por uno o más procesadores, tales como uno o más procesadores de señales digitales (DSP), microprocesadores de propósito general, circuitos integrados específicos de la aplicación (ASIC), matrices lógicas programables (FPGA) u otros circuitos lógicos integrados o discretos equivalentes. En consecuencia, el término "procesador", como se usa en el presente documento, se puede referir a cualquiera de las estructuras anteriores o a cualquier otra estructura adecuada para la implementación de las técnicas descritas en el presente documento. Además, en algunos aspectos, la funcionalidad descrita en el presente documento se puede proporcionar dentro de módulos de hardware y/o software dedicados, configurados para codificar y decodificar, o incorporados en un códec combinado. Además, las técnicas se podrían implementar totalmente en uno o más circuitos o elementos lógicos.

10 **[0327]** Las técnicas de la presente divulgación se pueden implementar en una amplia variedad de dispositivos o aparatos, incluyendo un teléfono inalámbrico, un circuito integrado (IC) o un conjunto de IC (por ejemplo, un conjunto de chips). Diversos componentes, módulos o unidades se describen en la presente divulgación para enfatizar aspectos funcionales de dispositivos configurados para llevar a cabo las técnicas divulgadas, pero no requieren necesariamente su realización por diferentes unidades de hardware. En cambio, como se describe anteriormente, diversas unidades se pueden combinar en una unidad de hardware de códec, o proporcionar por un grupo de unidades de hardware interoperativas, incluyendo uno o más procesadores, como se describe anteriormente, junto con software y/o firmware adecuados.

15 **[0328]** Se han descrito diversos aspectos de la divulgación. Estos y otros aspectos están dentro del alcance de las siguientes reivindicaciones.

REIVINDICACIONES

- 5 1. Un procedimiento de codificación de datos de vídeo de múltiples vistas tridimensional que incluye una secuencia de vídeo codificada que comprende uno o más componentes de vista, cada uno de los cuales comprende un componente de profundidad y un componente de textura, comprendiendo el procedimiento:
- 10 determinar (162) una primera información de nivel de secuencia que describa características de los componentes de textura o los componentes de profundidad del uno o más componentes de vista incluidos dentro de la secuencia de vídeo codificada;
- 15 determinar (164) una segunda información de nivel de secuencia que describa características de un punto de funcionamiento de los datos de vídeo, identificándose el punto de funcionamiento por un valor que representa una velocidad de trama de destino y un conjunto de valores que representan vistas de salida de destino, y en la que el punto de funcionamiento define una combinación de vistas a extraer de los datos de vídeo, y
- 20 en el que la etapa de determinar (164) la segunda información de nivel comprende una etapa de selección del punto de funcionamiento de un primer subconjunto de puntos de funcionamiento o un segundo subconjunto de puntos de funcionamiento, en la que cada punto de funcionamiento en el primer subconjunto contiene solo un componente de textura y en la que cada punto de funcionamiento en el segundo subconjunto contiene un componente de profundidad, en la que el punto de funcionamiento seleccionado identifica los componentes de textura y los componentes de profundidad de los componentes de vista de modo que los componentes de textura y los componentes de profundidad se pueden extraer por separado de los componentes de vista; y
- 25 codificar un subconjunto de flujo de bits asociado con el punto de funcionamiento seleccionado que comprende las vistas de salida de destino y todas las demás vistas de las que dependen las vistas de salida de destino, y emitir las vistas de salida de destino asociadas con el subconjunto de flujo de bits.
- 30 2. El procedimiento de la reivindicación 1, en el que la primera información de nivel de secuencia comprende un conjunto de parámetros de secuencia de codificación de vídeo tridimensional que especifica una dependencia de vista del componente de profundidad y, preferentemente, que comprende además determinar una lista de imágenes de referencia que identifique una o más imágenes de referencia para el componente de profundidad indicado en el conjunto de parámetros de secuencia de codificación de vídeo tridimensional.
- 35 3. El procedimiento de la reivindicación 2,
- 40 en el que la segunda información de nivel de secuencia incluye un conjunto de parámetros de secuencia de codificación de vídeo tridimensional que describe, para el punto de funcionamiento, una lista de vistas de salida de destino, un número de vistas de textura a descodificar cuando se descodifique el punto de funcionamiento y un número de vistas de profundidad a descodificar cuando se descodifique el punto de funcionamiento,
- 45 en el que el número de vistas de textura a descodificar es diferente del número de vistas de profundidad.
- 50 4. El procedimiento de la reivindicación 3, que comprende además seleccionar, para cada una de las vistas de salida de destino especificadas en la lista de vistas de salida de destino, el componente de profundidad cuando esté disponible.
- 55 5. El procedimiento de la reivindicación 1, que comprende además especificar un mensaje de información de mejora complementaria (SEI) de información de escalabilidad de codificación de vídeo tridimensional como parte de los datos de vídeo, en el que el mensaje SEI de información de escalabilidad de codificación de vídeo tridimensional incluye una descripción del punto de funcionamiento sin el componente de profundidad o una descripción del punto de funcionamiento con el componente de profundidad y, preferentemente, en el que los datos de vídeo se ajustan a una extensión de codificación de vídeo tridimensional para una norma H.264/Codificación de Vídeo Avanzada y es retrocompatible con una extensión de codificación de vídeo de múltiples vistas para la norma H.264/Codificación de Vídeo Avanzada.
- 60 6. El procedimiento de la reivindicación 5, en el que el mensaje SEI incluye una indicación de si el punto de funcionamiento incluye el componente de profundidad y, preferentemente, en el que el mensaje SEI incluye una indicación de si el punto de funcionamiento incluye el componente de profundidad y vistas de profundidad directamente dependientes con vistas de textura directamente dependientes para la descodificación del punto de funcionamiento.
- 65 7. Un dispositivo de codificación de vídeo (20, 30) para codificar datos de vídeo de múltiples vistas tridimensional que incluye una secuencia de vídeo codificada que comprende uno o más componentes de vista, cada uno de los cuales comprende un componente de profundidad y un componente de textura, comprendiendo el dispositivo de codificación de vídeo:

medios para determinar una primera información de nivel de secuencia que describa características de los componentes de textura o los componentes de profundidad del uno o más componentes de vista incluidos dentro de la secuencia de vídeo codificada; y

5 medios para determinar una segunda información de nivel de secuencia que describa características de un punto de funcionamiento de los datos de vídeo, identificándose el punto de funcionamiento por un valor que representa una velocidad de trama de destino y un conjunto de valores que representan vistas de salida de destino, y en el que el punto de funcionamiento define una combinación de vistas a extraer de los datos de vídeo, y

10 en el que los medios para determinar la segunda información de nivel seleccionan el punto de funcionamiento de un primer subconjunto de puntos de funcionamiento o un segundo subconjunto de puntos de funcionamiento, en el que cada punto de funcionamiento en el primer subconjunto contiene solo un componente de textura y en el que cada punto de funcionamiento en el segundo subconjunto contiene un componente de profundidad, en el que el punto de funcionamiento seleccionado identifica los componentes de textura y los componentes de profundidad de los componentes de vista de modo que los componentes de textura y los componentes de profundidad se puedan extraer por separado de los componentes de vista; y

20 un codificador para codificar un subconjunto de flujo de bits asociado con el punto de funcionamiento seleccionado que comprende las vistas de salida de destino y todas las demás vistas de las que dependen las vistas de salida de destino, y emitir las vistas de salida de destino asociadas con el subconjunto de flujo de bits.

25 **8.** El dispositivo de codificación de vídeo de la reivindicación 7, en el que la primera información de nivel de secuencia comprende un conjunto de parámetros de secuencia de codificación de vídeo tridimensional que especifica una dependencia de vista del componente de profundidad.

30 **9.** El dispositivo de codificación de vídeo de la reivindicación 8, que comprende además medios para determinar una lista de imágenes de referencia que identifica una o más imágenes de referencia para el componente de profundidad indicado en el conjunto de parámetros de secuencia de codificación de vídeo tridimensional.

35 **10.** El dispositivo de codificación de vídeo de la reivindicación 8, en el que la segunda información de nivel de secuencia incluye un conjunto de parámetros de secuencia de codificación de vídeo tridimensional que describe, para el punto de funcionamiento, una lista de vistas de salida de destino, un número de vistas de textura a descodificar cuando se descodifique el punto de funcionamiento y un número de vistas de profundidad a descodificar cuando se descodifique el punto de funcionamiento, en el que el número de vistas de textura a descodificar es diferente del número de vistas de profundidad.

40 **11.** El dispositivo de codificación de vídeo de la reivindicación 10, que comprende además medios para seleccionar, para cada una de las vistas de salida de destino especificadas en la lista de vistas de salida de destino, el componente de profundidad cuando esté disponible.

45 **12.** El dispositivo de codificación de vídeo de la reivindicación 10, que comprende además medios para especificar un mensaje de información de mejora complementaria (SEI) de información de escalabilidad de codificación de vídeo tridimensional como parte de los datos de vídeo, en el que el mensaje SEI de información de escalabilidad de codificación de vídeo tridimensional incluye una descripción del punto de funcionamiento sin el componente de profundidad o una descripción del punto de funcionamiento con el componente de profundidad.

50 **13.** El dispositivo de codificación de vídeo de la reivindicación 12, en el que el mensaje SEI incluye una indicación de si el punto de funcionamiento incluye el componente de profundidad.

55 **14.** El dispositivo de codificación de vídeo de la reivindicación 12, en el que el mensaje SEI incluye una indicación de si el punto de funcionamiento incluye el componente de profundidad y vistas de profundidad directamente dependientes con vistas de textura directamente dependientes para la descodificación de los y, preferentemente, en el que los datos de vídeo se ajustan a una extensión de codificación de vídeo tridimensional para una norma H.264/Codificación de Vídeo Avanzada y es retrocompatible con una extensión de codificación de vídeo de múltiples vistas para la norma H.264/Codificación de Vídeo Avanzada.

60 **15.** Un medio de almacenamiento legible por ordenador no transitorio que tiene almacenadas en el mismo instrucciones que, cuando se ejecutan, hacen que uno o más procesadores de un dispositivo de codificación de vídeo ejecuten el procedimiento de una cualquiera de las reivindicaciones 1 a 6.

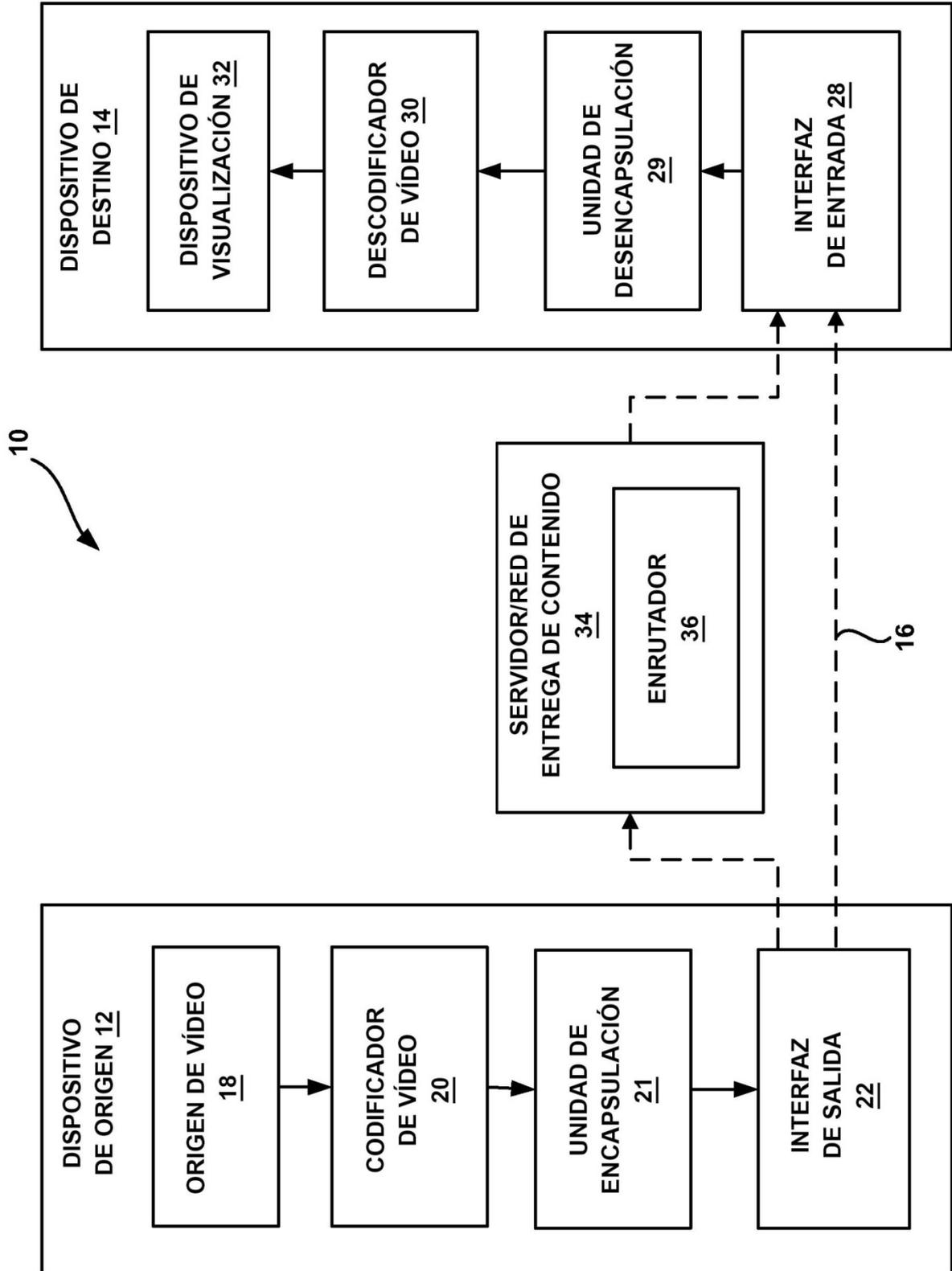


FIG. 1

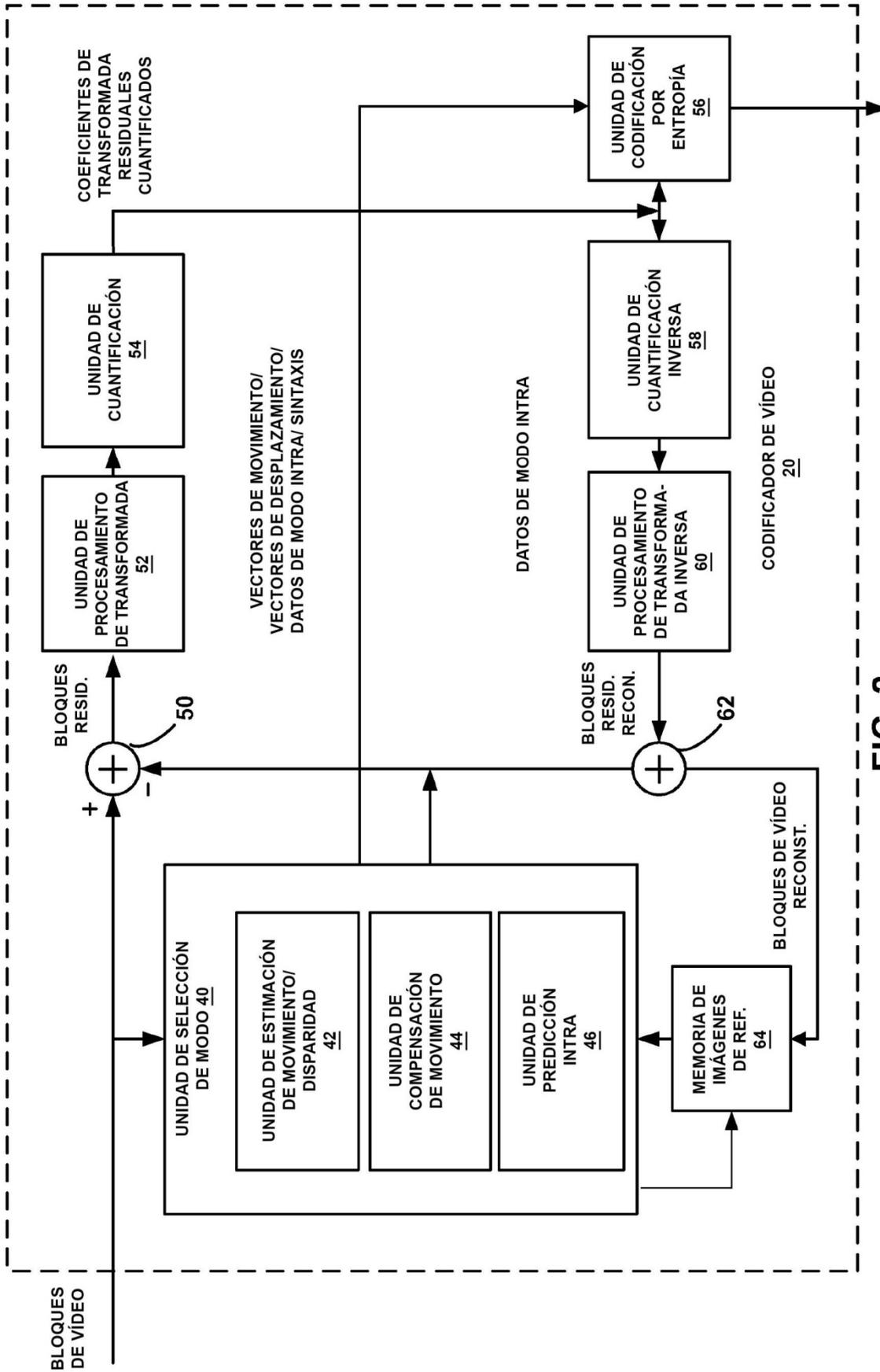


FIG. 2

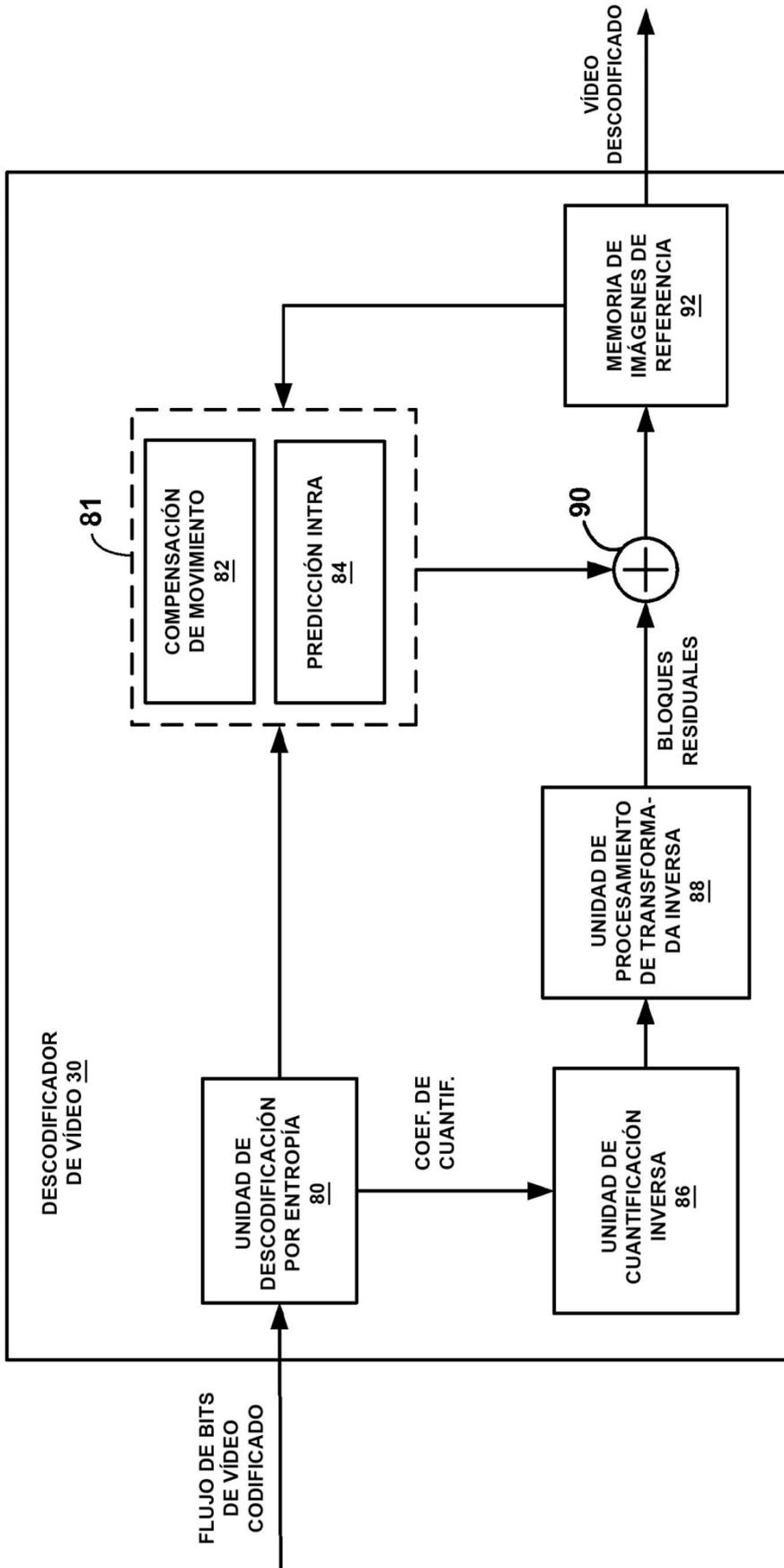


FIG. 3

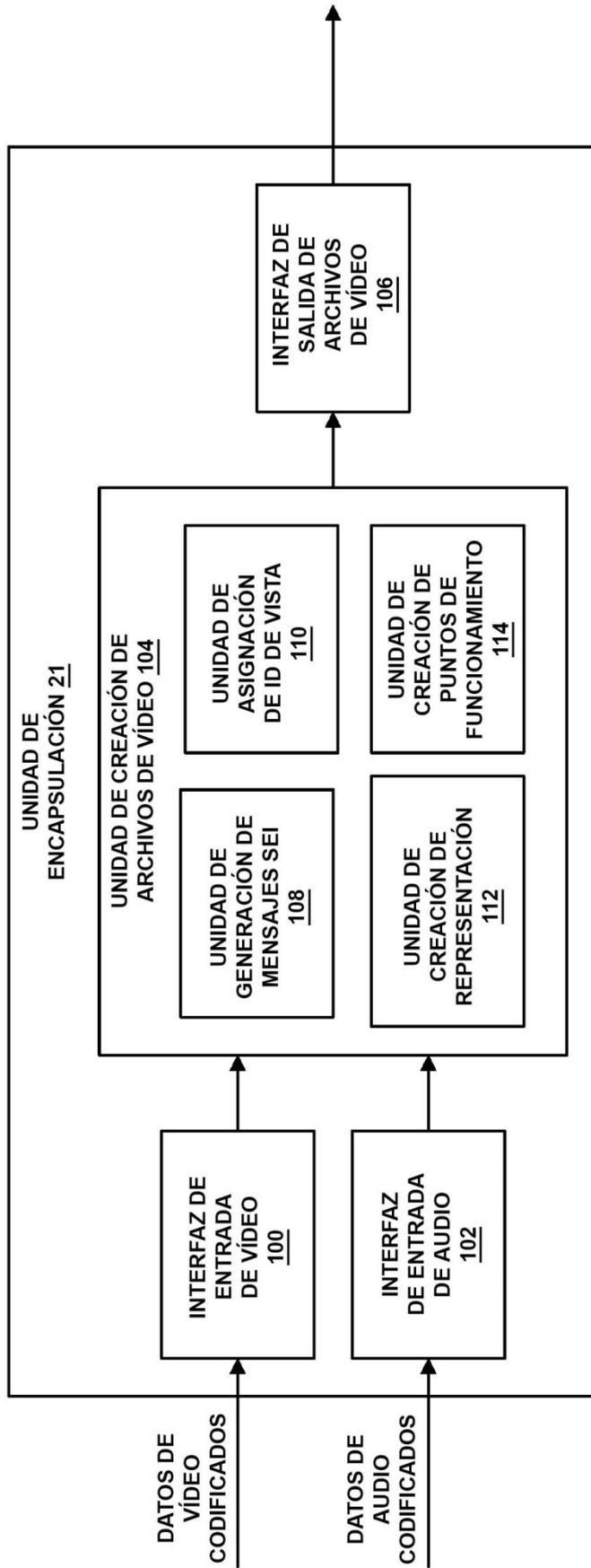


FIG. 4

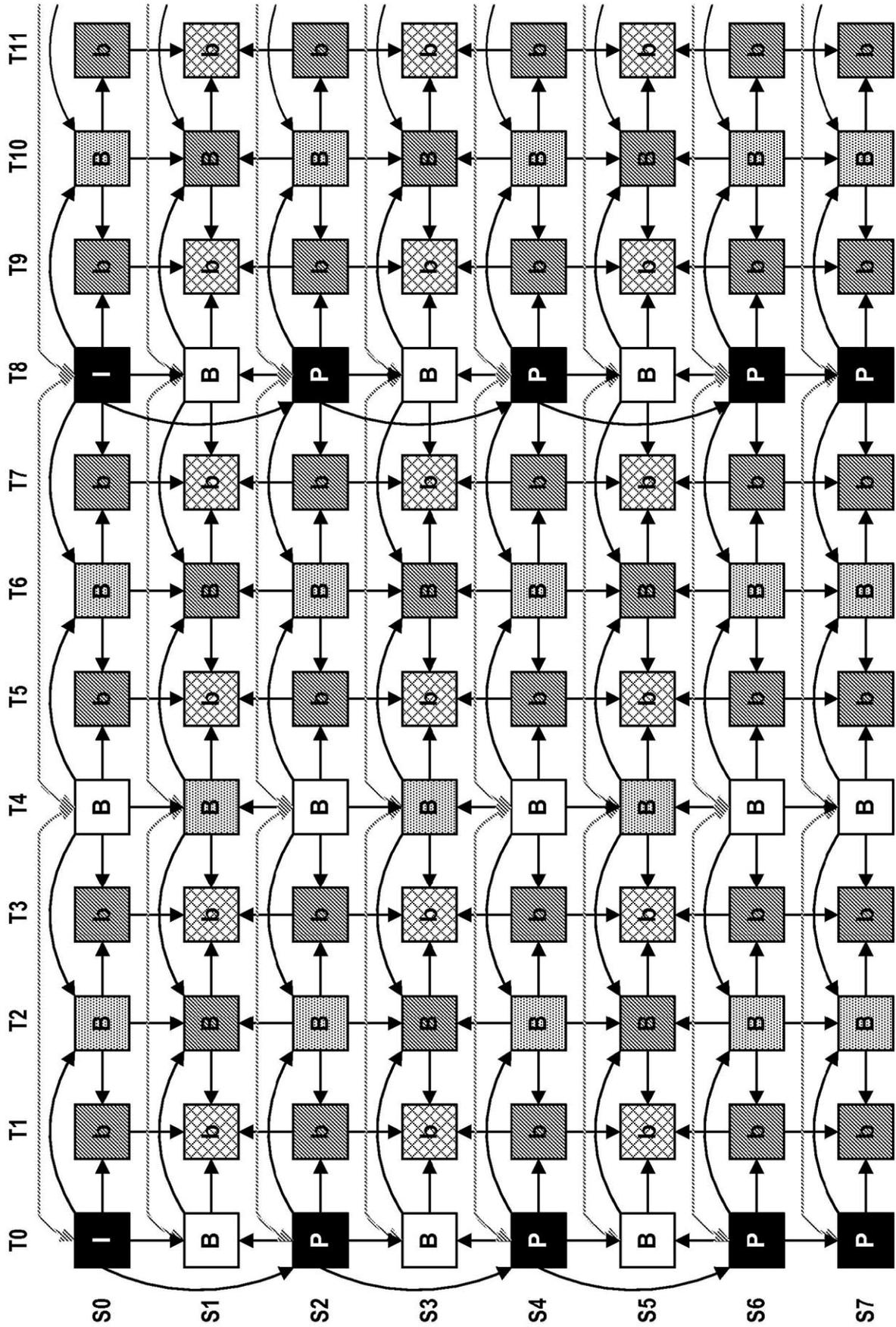


FIG. 5

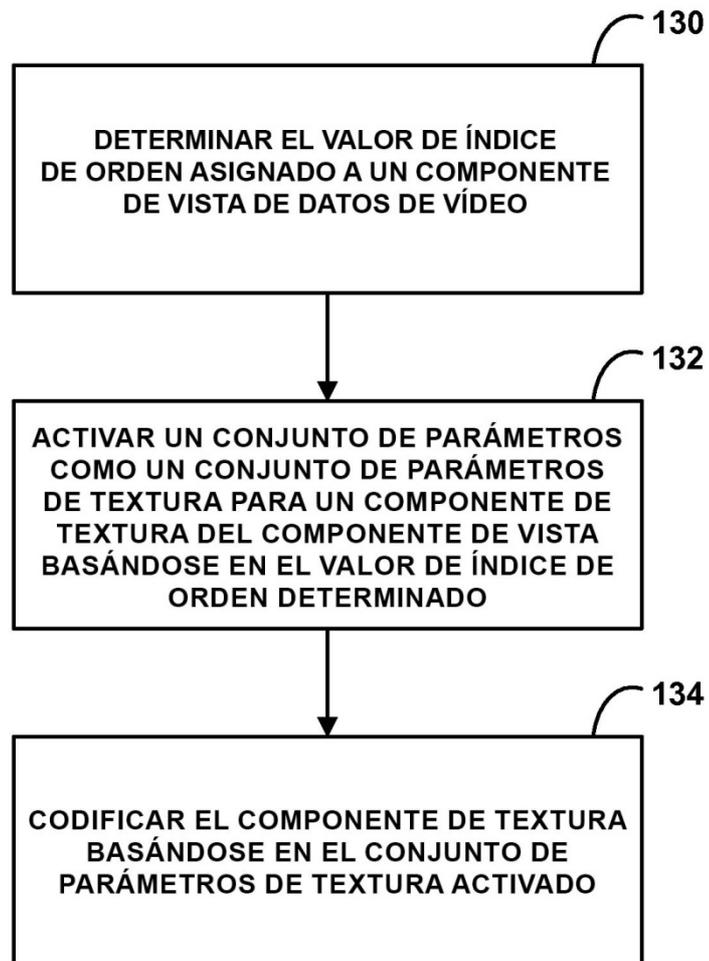


FIG. 6

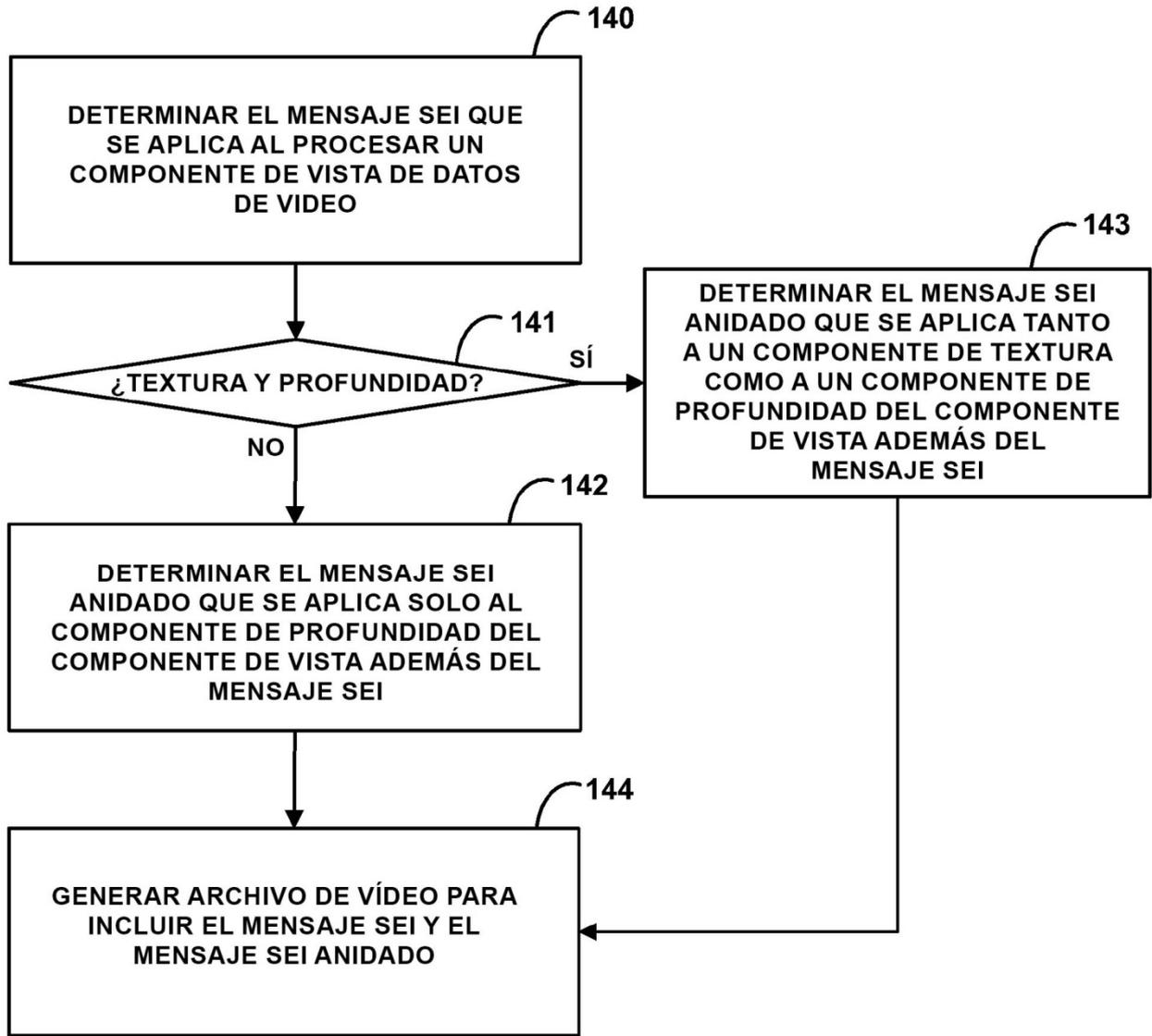


FIG. 7

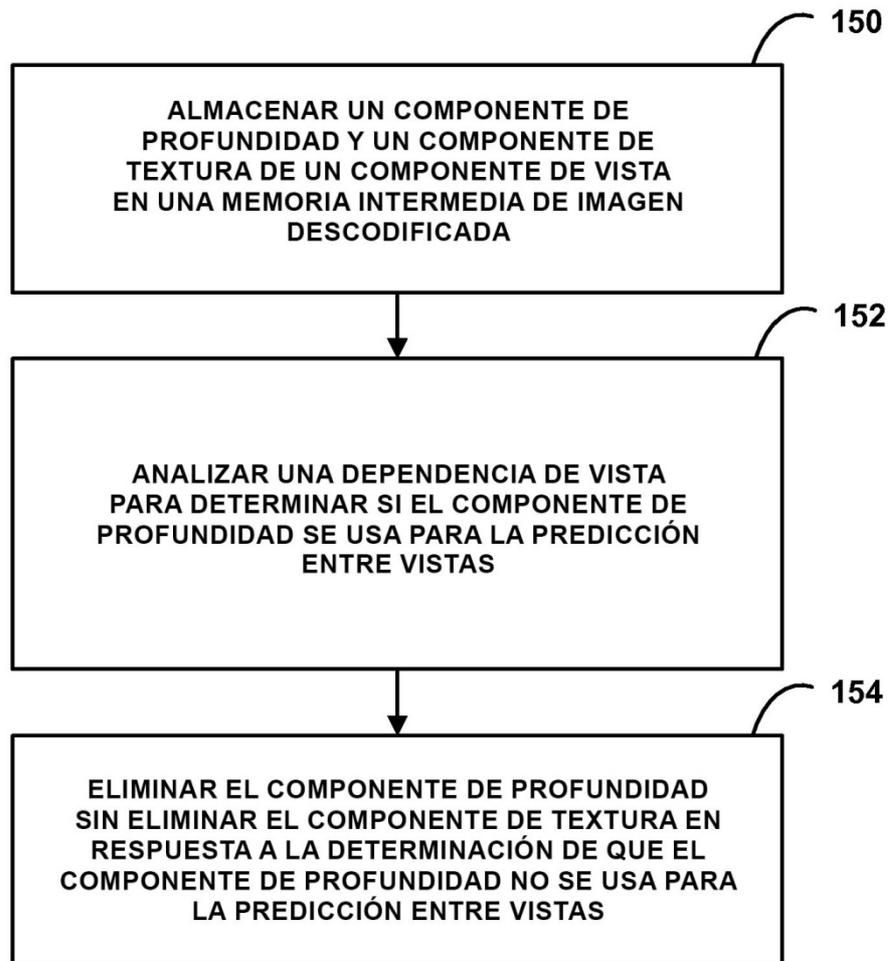


FIG. 8

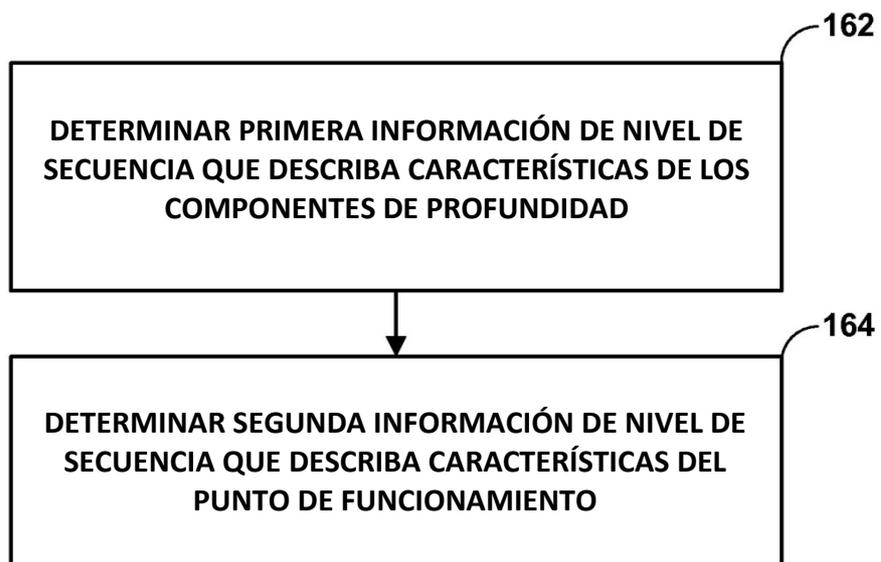


FIG. 9