

19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 730 876**

51 Int. Cl.:

H04N 19/30 (2014.01)

H04N 19/597 (2014.01)

H04N 19/70 (2014.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

86 Fecha de presentación y número de la solicitud internacional: **30.10.2013 PCT/US2013/067537**

87 Fecha y número de publicación internacional: **08.05.2014 WO14070911**

96 Fecha de presentación y número de la solicitud europea: **30.10.2013 E 13792536 (8)**

97 Fecha y número de publicación de la concesión europea: **13.03.2019 EP 2915334**

54 Título: **Capas de salida de destino en la codificación de vídeo**

30 Prioridad:

30.10.2012 US 201261720364 P
29.10.2013 US 201314066209

45 Fecha de publicación y mención en BOPI de la traducción de la patente:
13.11.2019

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 2 730 876 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

Capas de salida de destino en la codificación de vídeo

5 **CAMPO TÉCNICO**

[0001] Esta divulgación se refiere a la codificación de vídeo.

ANTECEDENTES

10 **[0002]** Las capacidades del vídeo digital pueden incorporarse a una amplia gama de dispositivos, que incluye televisores digitales, sistemas de difusión digital directa, sistemas de difusión 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 denominados «teléfonos inteligentes», dispositivos de videoconferencia, dispositivos de transmisión de vídeo en tiempo real y similares. Los dispositivos de vídeo digital implementan técnicas de codificación 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 Avanzada de Vídeo (AVC), la norma de Codificación de Vídeo de Alta Eficacia (HEVC), actualmente en desarrollo, y las extensiones de dichas normas, tales como la Codificación de Vídeo Ajustable a Escala (SVC) y la Codificación de Vídeo de Múltiples Vistas (MVC). La versión 6 del borrador de trabajo (WD) de la HEVC, de Bross et al., "High efficiency video coding (HEVC) text specification draft 6 [Especificación textual de la Codificación de Vídeo de Alta Eficacia (HEVC), Borrador 6]", Equipo de colaboración conjunta en codificación de vídeo (JCT-VC) de ITU-T SG16 WP3 e ISO/IEC JTC1/SC29/WG11, 8ª Conferencia: San José, CA, EE. UU., 1 al 10 de febrero de 2012, está disponible en http://phenix.it-sudparis.eu/jct/doc_end_user/current_document.php?id=5096. Los dispositivos de vídeo pueden transmitir, recibir, codificar, decodificar y/o almacenar información de vídeo digital más eficazmente, implementando tales técnicas de codificación de vídeo.

30 **[0003]** Las técnicas de codificación de vídeo incluyen la predicción espacial (intra-imagen) y/o la predicción temporal (entre imágenes) para reducir o eliminar la redundancia intrínseca en las secuencias de vídeo. Para la codificación de vídeo basada en bloques, un fragmento de vídeo (por ejemplo, una trama de vídeo o una parte de una trama de vídeo) puede dividirse en bloques de vídeo, que también pueden denominarse 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 vecinos 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 vecinos en la misma imagen o la predicción temporal con respecto a muestras de referencia en otras imágenes de referencia. Las imágenes pueden denominarse tramas y las imágenes de referencia pueden denominarse tramas de referencia.

40 **[0004]** La predicción espacial o temporal da como resultado un bloque predictivo para un bloque a codificar. Los datos residuales representan diferencias de píxeles entre el bloque original que a codificar y el bloque predictivo. Un bloque intercodificado se codifica de acuerdo a 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 a una modalidad de intracodificación y a los datos residuales. Para una mayor compresión, los datos residuales pueden transformarse desde el dominio del píxel a un dominio de transformación, dando como resultado coeficientes de transformación residuales, que a continuación se pueden cuantizar. Los coeficientes de transformación cuantizados, dispuestos inicialmente en una formación bidimensional, pueden explorarse con el fin de generar un vector unidimensional de coeficientes de transformación, y puede aplicarse codificación por entropía para lograr aún más compresión.

50 **SUMARIO**

[0005] Las técnicas de esta divulgación incluyen la determinación de una o más capas de salida de destino en la codificación de vídeo de múltiples vistas y / o la codificación de vídeo ajustable a escala. Por ejemplo, un punto de operación, como se define en la especificación básica actual de la HEVC (WD 6, como se ha indicado anteriormente), puede incluir una serie de capas de datos de vídeo, de manera que un punto de operación se refiera a un conjunto de capas que tenga una o más capas de un flujo de bits de datos de vídeo. En algunos casos, solo un subconjunto de tales capas puede ser decodificado para su salida, por ejemplo, pantalla. De acuerdo a aspectos de esta divulgación, se puede definir un punto de operación de salida que identifique específicamente las capas de destino de los datos de vídeo para su salida.

[0006] Un procedimiento de decodificación de datos de vídeo se define en la reivindicación 1.

[0007] Un procedimiento para la codificación de datos de vídeo se define en la reivindicación 2.

[0008] Un dispositivo para la decodificación de datos de vídeo se define en la reivindicación 12.

[0009] Un dispositivo para la codificación de datos de vídeo se define en la reivindicación 13.

[0010] En otro ejemplo, un medio de almacenamiento no transitorio legible por ordenador tiene instrucciones almacenadas en el mismo que, al ser ejecutadas, hacen que un procesador de un dispositivo para codificar datos de vídeo ejecute el procedimiento de cualquiera de las reivindicaciones 1 a 11.

[0011] Los detalles de uno o más ejemplos se exponen en los dibujos adjuntos y en la siguiente descripción. Otras características, objetos y ventajas resultarán evidentes a partir de la descripción y de los dibujos, y a partir de las reivindicaciones.

BREVE DESCRIPCIÓN DE LOS DIBUJOS

[0012]

La figura 1 es un diagrama de bloques que ilustra un sistema ejemplar de codificación y decodificación de vídeo que puede utilizar técnicas para codificar una o más capas de salida de destino, según aspectos de esta divulgación.

La figura 2 es un diagrama de bloques que ilustra un ejemplo de un codificador de vídeo que puede implementar técnicas para determinar una o más capas de salida de destino, según aspectos de esta divulgación.

La figura 3 es un diagrama de bloques que ilustra un ejemplo de un decodificador de vídeo que puede implementar técnicas para determinar una o más capas de salida de destino, según aspectos de esta divulgación.

La figura 4 es un diagrama conceptual que ilustra un patrón ejemplar de predicción de la MVC.

La figura 5 es un diagrama conceptual que ilustra una secuencia ejemplar de codificación de vídeo ajustable a escala.

La figura 6 es un diagrama de flujo que ilustra un proceso ejemplar para codificar un punto de operación de salida, de acuerdo a aspectos de esta divulgación.

La figura 7 es un diagrama de flujo que ilustra un proceso ejemplar para decodificar un punto de operación de salida, de acuerdo a aspectos de esta divulgación.

DESCRIPCIÓN DETALLADA

[0013] Las técnicas de esta divulgación incluyen la determinación de una o más capas de salida de destino en la codificación de vídeo de múltiples vistas y / o la codificación de vídeo ajustable a escala. Actualmente, el grupo de expertos en imágenes en movimiento (MPEG) está desarrollando una norma de vídeo tridimensional (3DV) basada en la inminente norma de codificación de vídeo de alta eficacia (HEVC). Parte de los esfuerzos de estandarización también incluye la estandarización de un códec de vídeo de múltiples vistas y un códec de vídeo ajustable a escala, basado en la HEVC. Por ejemplo, un esfuerzo de estandarización incluye el desarrollo de una extensión de múltiples vistas de la HEVC, conocida como MV-HEVC, y otro es un códec de 3DV completo basado en la HEVC con profundidad realzada, denominado 3D-HEVC. Con respecto a la codificación de vídeo ajustable a escala, la ajustabilidad a escala de la vista y / o la ajustabilidad a escala espacial también pueden contribuir a los servicios de vídeo tridimensionales, ya que tales ajustabilidades a escala admiten extensiones retro-compatibles para más vistas y / o la mejora de la resolución de las vistas, de una manera que admite la decodificación por dispositivos heredados.

[0014] Para datos de vídeo tridimensional, en general, cada vista se corresponde con una perspectiva, o ángulo, diferente, en donde se capturaron los datos de vídeo correspondientes de una escena común. Las vistas codificadas pueden usarse para la visualización tridimensional (3D) de datos de vídeo. Por ejemplo, dos vistas (por ejemplo, las vistas de los ojos izquierdo y derecho de un observador humano) pueden exhibirse simultáneamente, o casi simultáneamente, utilizando diferentes polarizaciones de luz, y un espectador puede usar gafas polarizadas pasivas de tal manera que cada uno de los ojos del espectador reciba una respectiva vista. 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.

[0015] Con codificación de múltiples vistas, tal como la codificación de vídeo de múltiples vistas (MVC), según se estipula en el Anexo H de la norma UIT-T H.264 (también conocida como MPEG-4, Parte 10, norma de codificación de vídeo avanzada (AVC)), una imagen particular de una vista particular se menciona como un componente de vista. Es decir, un componente de vista de una vista corresponde a una instancia temporal particular de la vista. El vídeo de múltiples vistas puede contener una cantidad relativamente grande de dependencias estadísticas entre vistas, ya que todas las cámaras utilizadas para capturar los datos de múltiples vistas capturan la misma escena desde diferentes puntos de vista. Dichas dependencias pueden aprovecharse para la predicción combinada temporal y/o entre vistas, donde las imágenes no solo se predicen a partir de imágenes contiguas en el tiempo, sino también a partir de imágenes

correspondientes de otras vistas. Es decir, puede realizarse la predicción entre vistas entre imágenes en la misma unidad de acceso (es decir, dentro de la misma instancia temporal).

5 **[0016]** Como se ha indicado anteriormente, la codificación de vídeo ajustable a escala también se puede usar para codificar múltiples vistas, para proporcionar ajustabilidad a escala de vistas y / o ajustabilidad a escala espacial. Por ejemplo, la codificación de vídeo ajustable a escala (SVC) se describe en el Anexo G de la norma H.264 / AVC, y se puede usar para codificar múltiples capas de datos de vídeo. En algunos casos, cada capa puede corresponder a una vista particular.

10 **[0017]** Como se describe en este documento, una "capa" de vídeo puede referirse, en general, a una secuencia de imágenes que tienen al menos una característica común, tal como una vista, una velocidad de tramas, una resolución o similares. Por ejemplo, una capa puede incluir datos de vídeo asociados a una vista particular de datos de vídeo de múltiples vistas. Como otro ejemplo, una capa puede incluir datos de vídeo asociados a una capa particular de datos de vídeo ajustables a escala. Por lo tanto, esta divulgación puede referirse indistintamente a una capa y a una vista de datos de vídeo.

15 **[0018]** En cualquier caso, un codificador de vídeo (tal como un codificador de vídeo o decodificador de vídeo, como se describe con mayor detalle a continuación) puede usar uno o más conjuntos de parámetros para describir o determinar una variedad de características de datos de vídeo codificados. Con los conjuntos de parámetros, la información que cambia con poca frecuencia no necesita repetirse para cada imagen o secuencia de imágenes, aumentando por ello la eficacia de la codificación. En algunos casos, los conjuntos de parámetros pueden formar una parte de un flujo de bits de vídeo. En otros casos, los conjuntos de parámetros pueden ser recibidos por un decodificador de vídeo de otras maneras (por ejemplo, transmisión fuera de banda, codificación rígida o similares). Una secuencia de vídeo codificada, junto con sus conjuntos de parámetros asociados, puede representar una parte decodificable independientemente de un flujo de bits de vídeo.

20 **[0019]** La HEVC hereda el concepto de conjunto de parámetros de la norma H.264 / AVC, pero incluye varias modificaciones y adiciones. En general, un conjunto de parámetros de secuencia (SPS) contiene información que se aplica a todos los fragmentos de una secuencia de vídeo codificado. En la HEVC, una secuencia de vídeo codificado contiene todas las imágenes entre una primera imagen de actualización de decodificación instantánea (IDR) (incluida en la secuencia) y la siguiente imagen de IDR (no incluida en la secuencia) o el final del flujo de bits (si la primera imagen de IDR es la última en el flujo de bits). Un conjunto de parámetros de imagen (PPS) contiene información que puede cambiar de imagen a imagen.

30 **[0020]** La HEVC también incluye una estructura de conjunto de parámetros de vídeo (VPS). En general, un VPS describe las características generales de una secuencia de vídeo codificado, incluidas las dependencias entre subcapas temporales. Un VPS puede habilitar la extensibilidad compatible de la norma HEVC en una capa de sistemas. Por ejemplo, un VPS puede proporcionar información adicional sobre una estructura de flujo de bits ajustable a escala o de múltiples vistas, que puede ser relevante para un decodificador capaz de decodificar tales capas, pero que admite que un decodificador heredado decodifique una capa base e ignore dichas capas.

35 **[0021]** Cuando se incluyen varias capas de datos de vídeo en un flujo de bits, se puede usar un punto de operación para analizar sintácticamente o extraer ciertas capas del flujo de bits. Por ejemplo, un flujo de bits de múltiples capas puede incluir una pluralidad de puntos de operación para facilitar la decodificación y la visualización de un subconjunto de datos codificados recibidos. Como se describe en este documento, un "punto de operación" puede referirse generalmente a un conjunto de capas, que puede incluir una o más capas de datos de vídeo. Como se ha señalado anteriormente, una capa puede referirse generalmente a una secuencia de imágenes que tienen al menos una característica común, tal como una vista, una velocidad de tramas, una resolución o similares (por ejemplo, una vista de datos de vídeo, una capa de datos de vídeo ajustables a escala y similares). Por lo tanto, un punto de operación puede referirse a un conjunto de capas que tiene una o más vistas de datos de vídeo, una o más capas de datos de vídeo o similares. Por consiguiente, un "punto de operación" y un "conjunto de capas" pueden usarse indistintamente en este documento.

40 **[0022]** El uso de diferentes puntos de operación puede permitir que varios dispositivos clientes realicen la adaptación. Es decir, los dispositivos clientes con diferentes capacidades de representación y decodificación pueden extraer diferentes vistas para mostrar datos de vídeo bidimensionales o tridimensionales. Los dispositivos clientes también pueden negociar con un dispositivo servidor para recuperar datos de velocidades de bits variables, para adaptarse a medios de transporte de varias capacidades de ancho de banda.

50 **[0023]** En consecuencia, cada punto de operación puede representar una combinación diferente de vistas de datos de vídeo de múltiples vistas y / o una combinación diferente de capas de datos de vídeo ajustables a escala, codificados a diferentes velocidades de tramas temporales y resoluciones espaciales, por ejemplo, diferentes conjuntos de capas. En otras palabras, un punto de operación 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). En algunos casos, se puede asociar más de un punto de operación al mismo subconjunto de flujo de bits. Es decir, más de un

conjunto de capas puede estar asociado al mismo flujo de bits, de modo que un flujo de bits incluya uno o más conjuntos de capas.

[0024] Un codificador de vídeo puede definir una serie de conjuntos de capas para varias combinaciones de vistas / capas base y vistas / capas de mejora. Un conjunto de capas particular puede incluir una serie de capas de datos de vídeo, incluyendo una o más capas de salida de destino para su visualización. Por ejemplo, una vista de salida de destino puede incluir una vista de datos de vídeo de múltiples vistas que se pretende emitir, por ejemplo, visible cuando se exhibe. Del mismo modo, una capa de salida de destino puede incluir una capa de datos de vídeo ajustables a escala que se pretende emitir. Como se describe en este documento, la vista de salida de destino y la capa de salida de destino se pueden usar de manera intercambiable.

[0025] Por ejemplo, con respecto a la MVC, los puntos de operación pueden ser señalizados en un SPS. Las vistas de salida de destino de los datos de MVC pueden señalizarse junto con un nivel temporal máximo de destino de los datos de MVC. Con respecto a la SVC, una capa de salida de destino puede seleccionarse automáticamente de acuerdo a la capa más alta de un punto de operación, por ejemplo, con los máximos `id_dependencia` e `id_calidad`. Con respecto a la HEVC, un punto de operación puede especificar un subflujo de bits que puede extraerse basándose en el nivel temporal más alto del subflujo de bits. La HEVC también puede usar una lista de valores de `nuh_reservado_cero_6bits`, que pueden renombrarse como `ids_capa`.

[0026] Al adaptar la HEVC para usarla con múltiples capas de datos de vídeo, tal como en el caso de una extensión de codificación de múltiples vistas de la HEVC (MV-HEVC), los puntos de operación se pueden señalar únicamente en un VPS. Por ejemplo, las vistas que están presentes en el flujo de bits (para ser decodificadas) pueden ser señalizadas en un VPS para cada punto de operación. Con respecto a la estandarización de una extensión ajustable a escala de la HEVC (HSVC o, alternativamente, SHVC), se anticipa que la capa con el mayor Identificador de capa será la capa de salida de destino para un punto de operación en particular.

[0027] El enfoque actual de los puntos de operación en la HEVC, incluyendo el uso de la estructura de VPS, puede tener inconvenientes cuando se codifica con múltiples capas. Por ejemplo, los puntos de operación definidos en la HEVC pueden ser suficientes para fines de extracción del flujo de bits, pero el VPS no especifica las vistas / capas de salida de destino de los puntos de operación. Es decir, el VPS no identifica las capas particulares de datos de vídeo que se pretende mostrar.

[0028] Sin saber cuáles capas son las capas de salida de destino, las imágenes decodificadas de ciertas capas que no se utilizan para la salida pueden mantenerse en un almacén temporal de imágenes decodificadas (DPB) durante más tiempo del necesario, aumentando por ello los requisitos de almacenamiento para secuencias de múltiples vistas. Además, cuando se incluyen varias vistas en un subflujo de bits, un proveedor de contenido puede querer controlar cuáles subflujos de bits se decodifican y se proporcionan como salida a una pantalla para controlar la experiencia de visualización. Por ejemplo, es posible que el proveedor de contenido no desee permitir que un usuario vea imágenes de un determinado subflujo de bits o una combinación de subflujos de bits. Dicho control puede no estar disponible con el esquema de VPS actual de la HEVC.

[0029] Las técnicas de esta divulgación incluyen determinar, a partir de una pluralidad de capas de datos de vídeo que incluyen una serie de conjuntos de capas (conteniendo, cada una, una o más capas de datos de vídeo), uno o más puntos de operación de salida asociados a un conjunto de capas de la serie de conjuntos de capas y una o más capas de salida de destino. Por ejemplo, las técnicas de esta divulgación incluyen la determinación de un punto de operación de salida que identifica una o más vistas de salida de destino para datos de vídeo de múltiples vistas y / o capas de salida de destino para datos de vídeo ajustables a escala.

[0030] Las técnicas de esta descripción también incluyen la señalización, mediante un codificador de vídeo, de uno o más elementos sintácticos que indican el punto de operación de salida y / o las capas de salida de destino, y el análisis sintáctico, por un decodificador de vídeo a partir de un flujo de bits codificado, de uno o más elementos sintácticos que indican el punto de operación de salida y / o capas de salida de destino. Proporcionar y utilizar una indicación de un punto de operación de salida con una o más capas de salida de destino puede aumentar la eficacia con respecto a la gestión de almacenes temporales de imágenes decodificadas, y puede aumentar la flexibilidad con respecto a la indicación de vistas y / o capas particulares a emitir (por ejemplo, visualizar).

[0031] De acuerdo a los aspectos de esta divulgación, un punto de operación, según se define en la especificación básica actual de la HEVC (por ejemplo, WD 6, como se ha indicado anteriormente), puede corresponder a uno o más puntos de operación de salida, teniendo cada punto de operación de salida una o más vistas y / o capas de salida de destino diferentes. Por ejemplo, en algunos casos, un codificador de vídeo (tal como un codificador de vídeo o decodificador de vídeo) puede usar una serie de capas de datos de vídeo de un conjunto de capas para codificar adecuadamente los datos de vídeo. Sin embargo, solo un subconjunto de las capas decodificadas puede ser deseable para la salida, por ejemplo, la visualización. De acuerdo a aspectos de esta divulgación, el codificador de vídeo puede determinar un punto de operación de salida que identifica específicamente las capas de destino de los datos de vídeo para su salida. En algunos casos, el codificador de vídeo puede determinar más de un punto de operación de salida para un conjunto de capas, por ejemplo, para asimilar diferentes combinaciones de capas de datos de vídeo para su

salida.

[0032] En un ejemplo con fines de ilustración, supóngase que un conjunto de capas incluye una vista base (vista0), una primera vista (vista1), una segunda vista (vista2) y una tercera vista (vista3) para una aplicación estéreo (por ejemplo, dos vistas para su salida). Supóngase además que vista2 y vista3 proporcionan una salida estéreo óptima. En este ejemplo, un decodificador de vídeo puede usar la vista base como referencia para decodificar vista2 y vista3. Además, si el decodificador de vídeo cumple con ciertas normas de codificación convencionales (tales como las extensiones MVC o SVC para la H.264 / AVC indicada anteriormente), el decodificador de vídeo puede seleccionar el nivel más alto del punto de operación para su salida, es decir, la capa base, a pesar de que vista2 y vista3 son las vistas óptimas para la salida.

[0033] De acuerdo a aspectos de esta divulgación, un conjunto de capas especifica los uno o más subflujos de bits que se decodifican, y un punto de operación de salida indica un subconjunto de capas dentro del conjunto de capas que se han de emitir. Por lo tanto, en el ejemplo anterior, un decodificador de vídeo puede decodificar el conjunto de capas que tiene la vista base, la primera vista, la segunda vista y la tercera vista. Además, el decodificador de vídeo puede determinar un punto de operación de salida para el conjunto de capas, donde el punto de operación de salida incluye la vista2 y la vista3. Por lo tanto, el codificador de vídeo puede decodificar y emitir la vista2 y la vista3 para su visualización.

[0034] Un punto de operación de salida, como se ha descrito anteriormente, no se especificó previamente para la HEVC. Además, la extensión MVC para la norma H.264 / AVC no se encontró con el problema de determinar cuáles subflujos de bits de un punto de operación están concebidos para su visualización, ya que la MVC no usa el concepto de puntos de operación y todas las vistas de destino de un flujo de bits de MVC están especificadas por todas las vistas en el flujo de bits. Por consiguiente, si bien las técnicas de esta divulgación se pueden implementar con varias normas, los aspectos de esta divulgación proporcionan un mecanismo para que se especifiquen vistas de salida de destino específicas de un punto de operación utilizando la norma HEVC.

[0035] La figura 1 es un diagrama de bloques que ilustra un sistema ejemplar de codificación y decodificación de vídeo 10 que puede utilizar técnicas para determinar una o más capas de salida de destino, según aspectos de esta divulgación. Como se muestra en la figura 1, el sistema 10 incluye un dispositivo de origen 12 que proporciona datos de vídeo codificados, 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 entre una amplia gama de dispositivos, incluidos ordenadores de escritorio, ordenadores plegables (es decir, portátiles), ordenadores de tableta, decodificadores, equipos telefónicos de mano tales como los denominados teléfonos «inteligentes», los denominados paneles «inteligentes», televisores, cámaras, dispositivos de visualización, reproductores de medios digitales, consolas de videojuegos, dispositivos de transmisión de vídeo en tiempo real o similares. En algunos casos, el dispositivo de origen 12 y el dispositivo de destino 14 pueden estar equipados para la comunicación inalámbrica.

[0036] El dispositivo de destino 14 puede recibir los datos de vídeo codificados, a decodificar mediante el medio legible por ordenador 16. El medio legible por ordenador 16 puede comprender cualquier tipo de medio o dispositivo capaz de desplazar los datos de vídeo codificados desde el dispositivo de origen 12 al dispositivo de destino 14. En un ejemplo, el medio legible por ordenador 16 puede comprender un medio de comunicación para habilitar el dispositivo de origen 12 para transmitir datos de vídeo codificados directamente al dispositivo de destino 14 en tiempo real. Los datos de vídeo codificados se pueden modular de acuerdo a 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 físicas de transmisión. 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 encaminadores, conmutadores, estaciones base o cualquier otro equipo que pueda ser útil para facilitar la comunicación desde el dispositivo de origen 12 hasta el dispositivo de destino 14.

[0037] En algunos ejemplos, pueden emitirse datos codificados desde la interfaz de salida 22 hasta un dispositivo de almacenamiento. De forma similar, se puede acceder a los datos codificados del dispositivo de almacenamiento mediante una interfaz de entrada. El dispositivo de almacenamiento puede incluir cualquiera entre varios medios de almacenamiento de datos, distribuidos o de acceso 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 un ejemplo adicional, el dispositivo de almacenamiento puede corresponder a un servidor de ficheros o a otro dispositivo de almacenamiento intermedio que pueda almacenar el vídeo codificado generado por el dispositivo de origen 12. El dispositivo de destino 14 puede acceder a datos de vídeo almacenados desde el dispositivo de almacenamiento, mediante transmisión continua o descarga. El servidor de ficheros puede ser cualquier tipo de servidor capaz de almacenar datos de vídeo codificados y transmitir esos datos de vídeo codificados al dispositivo de destino 14. Los servidores ejemplares de ficheros incluyen un servidor de la Red (por ejemplo, para una sede de la Red), un servidor del 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, incluida una conexión a Internet. Esto puede incluir un canal inalámbrico (por ejemplo, una conexión

de Wi-Fi), una conexión cableada (por ejemplo, DSL, módem de cable, etc.) o una combinación de ambas que sea adecuada para acceder a datos de vídeo codificados almacenados en un servidor de ficheros. 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.

5 **[0038]** Las técnicas de esta divulgación no están limitadas necesariamente a aplicaciones o configuraciones inalámbricas. Las técnicas se pueden aplicar a la codificación de vídeo como soporte de cualquiera entre varias aplicaciones de multimedios, tales como difusiones 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
10 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 puede estar configurado para prestar soporte a una transmisión de vídeo unidireccional o bidireccional, a fin de prestar soporte a aplicaciones tales como la transmisión de vídeo en tiempo real, la reproducción de vídeo, la difusión de vídeo y/o la videotelefonía.

15 **[0039]** En el ejemplo de la figura 1, el dispositivo de origen 12 incluye un origen de vídeo 18, un codificador de vídeo 20 y una interfaz de salida 22. El dispositivo de destino 14 incluye una interfaz de entrada 28, un decodificador de vídeo 30 y un dispositivo de visualización 32. De acuerdo a esta divulgación, el codificador de vídeo 20 del dispositivo de origen 12 puede estar configurado para aplicar las técnicas para determinar una o más capas de salida de destino.
20 En otros ejemplos, un dispositivo de origen y un dispositivo de destino pueden incluir otros componentes o disposiciones. Por ejemplo, el dispositivo de origen 12 puede recibir datos de vídeo desde un origen de vídeo externo 18, tal como una cámara externa. Del mismo modo, el dispositivo de destino 14 puede interactuar con un dispositivo de visualización externo, en lugar de incluir un dispositivo de visualización integrado.

25 **[0040]** El sistema ilustrado 10 de la figura 1 es simplemente un ejemplo. Las técnicas para la determinación de una o más capas de salida de destino pueden realizarse mediante cualquier dispositivo de codificación y/o decodificación de vídeo digital. Aunque, en general, las técnicas de la presente divulgación son llevadas a cabo por un dispositivo de codificación de vídeo, las técnicas también pueden ser llevadas a cabo por un codificador/decodificador de vídeo, denominado habitualmente "CÓDEC". Además, las técnicas de la presente divulgación también se pueden llevar a cabo mediante un preprocesador de vídeo. El dispositivo de origen 12 y el dispositivo de destino 14 son simplemente
30 ejemplos de dichos dispositivos de codificación en los que el dispositivo de origen 12 genera datos de vídeo codificados para su transmisión al dispositivo de destino 14. En algunos ejemplos, los dispositivos 12, 14 pueden funcionar de manera esencialmente simétrica, de modo que cada uno de los dispositivos 12, 14 incluya componentes de codificación y decodificación de vídeo. De ahí que el sistema 10 pueda prestar soporte a la transmisión de vídeo unidireccional o bidireccional entre los dispositivos de vídeo 12, 14, por ejemplo, para la transmisión continua de vídeo, la reproducción de vídeo, la difusión de vídeo o la videotelefonía.

35 **[0041]** El origen de vídeo 18 del dispositivo de origen 12 puede incluir un dispositivo de captura de vídeo, tal como una cámara de vídeo, un archivo de vídeo que contiene vídeo capturado previamente y/o una interfaz de suministro 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 ha mencionado 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 capturado, precapturado 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.

40 **[0042]** El medio legible por ordenador 16 puede incluir medios transitorios, tales como una difusión inalámbrica o una 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, mediante 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.

45 **[0043]** 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, los GOP. El dispositivo de visualización 32 muestra los datos de vídeo decodificados a un usuario, y puede comprender cualquiera entre varios dispositivos de visualización, tales como un tubo de rayos catódicos (CRT), una pantalla de cristal líquido (LCD), una
50
55
60
65

pantalla de plasma, una pantalla de diodos orgánicos emisores de luz (OLED) u otro tipo de dispositivo de visualización.

[0044] Aunque no se muestra en la figura 1, en algunos aspectos, tanto el codificador de vídeo 20 como el decodificador de vídeo 30 pueden estar integrados con un codificador y decodificador de audio, y pueden incluir unidades MUX-DEMUX adecuadas, u otro tipo de hardware y software, para ocuparse de la codificación tanto de audio como de vídeo en un flujo de datos común o en flujos de datos separados. Si procede, las unidades MUX-DEMUX se pueden ajustar al protocolo multiplexador H.223 de la ITU o a otros protocolos tales como el protocolo de datagramas de usuario (UDP).

[0045] Tanto el codificador 20 de vídeo como el decodificador 30 de vídeo puede implementarse como cualquiera entre varios circuitos codificadores o decodificadores adecuados, según corresponda, tales como uno o más microprocesadores, procesadores de señales digitales (DSP), circuitos integrados específicos de la aplicación (ASIC), formaciones de compuertas programables in situ (FPGA), circuitos de lógica discreta, software, hardware, firmware o cualquier combinación de los mismos. Cuando las técnicas se implementan parcialmente en software, un dispositivo puede almacenar instrucciones para el software en un medio adecuado no transitorio legible por ordenador, y ejecutar las instrucciones en hardware usando uno o más procesadores para realizar las técnicas de esta divulgación. Tanto el codificador de vídeo 20 como el decodificador de vídeo 30 puede estar incluido en uno o más codificadores o decodificadores, cada uno de los cuales puede estar integrado como parte de un codificador/decodificador (CÓDEC) de vídeo combinado. Un dispositivo que incluye el codificador de vídeo 20 y/o el decodificador de vídeo 30 puede comprender un circuito integrado, un microprocesador y/o un dispositivo de comunicación inalámbrica, tal como un teléfono celular.

[0046] En el ejemplo mostrado en la figura 1, el sistema 10 también incluye el servidor / la red de entrega de contenido 34 que tiene el encaminador 36. En algunos ejemplos, el dispositivo de origen 12 se puede comunicar con el servidor / la red de entrega de contenido 34 por medio de varios medios de transmisión o almacenamiento, inalámbricos y/o cableados, como se ha descrito anteriormente. Además, si bien se muestra por separado en el ejemplo de la figura 1, en algunos ejemplos, el dispositivo de origen 12 y el servidor / la red de entrega de contenido 34 comprenden el mismo dispositivo. El servidor / la red de entrega de contenido 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 su acceso por el dispositivo de destino 14 y el decodificador de vídeo 30. En algunos ejemplos, el encaminador 36 puede ser responsable de proporcionar datos de vídeo codificados al dispositivo de destino 14 en un formato solicitado.

[0047] Esta divulgación puede referirse en general al codificador de vídeo 20 y / o al servidor / a la red de entrega de contenido 34 como "señalizadores" de cierta información a otro dispositivo, tal como el decodificador de vídeo 30. Sin embargo, debería entenderse que el codificador de vídeo 20 y / o el servidor / la red de entrega de contenido 34 pueden señalar información asociando ciertos elementos sintácticos a varias partes codificadas de datos de vídeo. Es decir, el codificador de vídeo 20 y / o el servidor / la red de entrega de contenido 34 pueden "señalizar" datos almacenando ciertos elementos sintácticos en cabeceras de varias partes codificadas de datos de vídeo. En algunos casos, dichos elementos sintácticos pueden ser codificados y almacenados (por ejemplo, almacenados en un dispositivo de almacenamiento 24) antes de ser recibidos y decodificados por el decodificador de vídeo 30. Por lo tanto, el término "señalización" puede referirse en general a la comunicación de datos sintácticos u otros, usados para decodificar datos de vídeo comprimidos, ya sea que dicha comunicación se produzca en tiempo real o casi real o en un intervalo de tiempo, tal como podría ocurrir cuando se almacenan elementos sintácticos en un medio en el momento de la codificación, que entonces pueden ser recuperados por un dispositivo de decodificación en cualquier momento después de haber sido almacenados en este medio.

[0048] El codificador de vídeo 20 y el decodificador de vídeo 30 pueden funcionar de acuerdo a una norma de compresión de vídeo, tal como la norma ITU-T H.264, denominada de forma alternativa MPEG-4, parte 10, codificación de vídeo avanzada (AVC), o extensiones de dichas normas. La norma ITU-T H.264/MPEG-4 (AVC) fue formulada por el Grupo de Expertos en Codificación de Vídeo (VCEG) de la ITU-T junto con el Grupo de Expertos en Imágenes en Movimiento (MPEG) de la ISO/IEC, como el producto de una asociación colectiva conocida como el Equipo Mixto de Vídeo (JVT, por sus siglas en inglés). 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 está descrita 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. Otros ejemplos de normas de compresión de vídeo incluyen MPEG-2 e ITU-T H.263.

[0049] Si bien las técnicas de esta divulgación no están limitadas a ninguna norma de codificación particular, las técnicas pueden ser relevantes para la norma HEVC. Más específicamente, el codificador de vídeo 20 y el decodificador de vídeo 30 pueden configurarse para codificar datos de vídeo de acuerdo a una extensión de la norma HEVC, por ejemplo, una extensión de múltiples vistas o extensión de vídeo tridimensional (3DV), incluyendo una extensión de codificación de vídeo ajustable a escala (SVC).

[0050] En general, la HEVA admite que una imagen de vídeo se divida en una secuencia de bloques arbolados o unidades máximas de codificación (LCU), que incluyen muestras tanto de luma como de croma. Los datos sintácticos dentro de un flujo de bits pueden definir un tamaño para la LCU, que es una máxima unidad de codificación en lo que respecta al número de píxeles. Un fragmento incluye una serie de unidades de árbol de codificación (CTU) consecutivas. Cada una de las CTU puede comprender un bloque arbolado de codificación de muestras de luma, dos correspondientes bloques arbolados de codificación de muestras de croma y estructuras sintácticas usadas para codificar las muestras de los bloques arbolados de codificación. En una imagen monocromática o en una imagen que tenga tres planos de color independientes, una CTU puede comprender un único bloque arbolado de codificación y estructuras sintácticas usadas para codificar las muestras del bloque arbolado de codificación.

[0051] Una imagen de vídeo puede dividirse en uno o más fragmentos. Cada bloque arbolado puede dividirse en unidades de codificación (CU) de acuerdo a un árbol cuádruple. En general, una estructura de datos de árbol cuádruple incluye un nodo por CU, con un nodo raíz correspondiente al bloque arbolado. Si una CU se divide en cuatro sub-CU, el nodo correspondiente a la CU incluye cuatro nodos hoja, cada uno de los cuales corresponde a una de las sub-CU. Una CU puede comprender un bloque de codificación de muestras de luma y dos correspondientes bloques de codificación de muestras de croma de una imagen que tenga una formación de muestras de luma, una formación de muestras de Cb y una formación de muestras de Cr, y estructuras sintácticas usadas para codificar las muestras de los bloques de codificación. En una imagen monocromática o una imagen que tenga tres planos cromáticos independientes, una CU puede comprender un único bloque de codificación y estructuras sintácticas usadas para codificar las muestras del bloque de codificación. Un bloque de codificación es un bloque de muestras de tamaño NxN.

[0052] Cada nodo de la estructura de datos en árbol cuádruple puede proporcionar datos sintácticos para la CU correspondiente. Por ejemplo, un nodo en el árbol cuádruple puede incluir un indicador de división, que indica si la CU correspondiente al nodo está dividida en varias sub-CU. Los elementos sintácticos para una CU pueden definirse de manera recursiva y pueden depender de si la CU está dividida en varias sub-CU. Si una CU no está dividida adicionalmente, se denomina CU hoja. En esta divulgación, cuatro sub-CU de una CU hoja también se denominarán CU hojas incluso aunque no haya ninguna división explícita de la CU hoja original. Por ejemplo, si una CU con un tamaño de 16x16 no se divide adicionalmente, las cuatro sub-CU de tamaño 8x8 también se denominarán CU hojas aunque la CU de tamaño 16x16 no se haya dividido nunca.

[0053] Una CU tiene un propósito similar a un macrobloque de la norma H.264, excepto porque una CU no tiene una distinción de tamaño. Por ejemplo, un bloque arbolado puede dividirse en cuatro nodos secundarios (también denominados sub-CU) y cada nodo secundario puede a su vez ser un nodo principal y dividirse en otros cuatro nodos secundarios. Un nodo secundario final, no dividido, denominado nodo hoja del árbol cuádruple, comprende un nodo de codificación, también denominado CU hoja. Los datos sintácticos asociados a un flujo de bits codificado pueden definir un número máximo de veces en que puede dividirse un bloque arbolado, denominado profundidad de CU máxima, y también pueden definir un tamaño mínimo de los nodos de codificación. Por consiguiente, un flujo de bits también puede definir una unidad de codificación de tamaño mínimo (SCU). Esta divulgación utiliza el término "bloque" para referirse a cualquiera entre una CU, PU o TU, en el contexto de la HEVC, o a estructuras de datos similares en el contexto de otras normas (por ejemplo, macrobloques y subbloques de los mismos en la norma H.264/AVC).

[0054] Una CU incluye un nodo de codificación y unidades de predicción (PU) y unidades de transformación (TU) asociadas al nodo de codificación. Un tamaño de la CU corresponde a un tamaño del nodo de codificación y debe ser de forma cuadrada. El tamaño de la CU puede variar desde 8x8 píxeles hasta el tamaño del bloque arbolado, con un máximo de 64x64 píxeles o más. Cada CU puede contener una o más PU y una o más TU.

[0055] En general, una PU representa una zona espacial correspondiente a la totalidad, o a una parte, de la CU correspondiente, y puede incluir datos para recuperar una muestra de referencia para la PU. Además, una PU incluye datos relacionados con la predicción. Por ejemplo, cuando la PU está codificada en intramodalidad, pueden incluirse datos para la PU en un árbol cuádruple residual (RQT), que puede incluir datos que describen una modalidad de intrapredicción para una TU correspondiente a la PU. Como otro ejemplo, cuando la PU está codificada en intermodalidad, la PU puede incluir datos que definen uno o más vectores de movimiento para la PU. Un bloque de predicción puede ser un bloque rectangular (es decir, cuadrado o no cuadrado) de muestras en las que se aplique la misma predicción. Una PU de una CU puede comprender un bloque de predicción de muestras de luminancia, dos correspondientes bloques de predicción de muestras de crominancia de una imagen y estructuras sintácticas usadas para predecir las muestras de bloques de predicción. En una imagen monocromática o una imagen que tenga tres planos cromáticos independientes, una PU puede comprender un único bloque de predicción y estructuras sintácticas usadas para predecir las muestras del bloque de predicción.

[0056] Las TU pueden incluir coeficientes en el dominio de transformación después de la aplicación de una transformación, por ejemplo, una transformación de coseno discreta (DCT), una transformación entera, una transformación de ondículas o una transformación conceptualmente similar a los datos de vídeo residuales. Los datos residuales pueden corresponder a diferencias de píxeles entre píxeles de la imagen no codificada y valores de predicción correspondientes a las PU. El codificador de vídeo 20 puede formar las TU incluyendo los datos residuales para la CU y, a continuación, transformar las TU para generar coeficientes de transformación para la CU. Un bloque de transformación puede ser un bloque rectangular de muestras en el que se aplique la misma transformación. Una

unidad de transformación (TU) de una CU puede comprender un bloque de transformación de muestras de luma, dos correspondientes bloques de transformación de muestras de croma y estructuras sintácticas usadas para transformar las muestras de bloques de transformación. En una imagen monocromática o una imagen que tenga tres planos cromáticos independientes, una TU puede comprender un único bloque de transformación y estructuras sintácticas usadas para transformar las muestras del bloque de transformación.

[0057] Tras la codificación intrapredictiva o interpredictiva, usando las PU de una CU, el codificador de vídeo puede calcular datos residuales para las TU de la CU. Las PU pueden comprender datos sintácticos que describen un procedimiento o modalidad de generación de datos de píxeles predictivos en el dominio espacial (también denominado el dominio de píxeles) y las TU pueden comprender coeficientes en el dominio de transformación, tras la aplicación de una transformación, por ejemplo, una transformación de coseno discreta (DCT), una transformación entera, una transformación de ondículas o una transformación conceptualmente similar, a los datos de vídeo residuales. Los datos residuales pueden corresponder a diferencias de píxeles entre píxeles de la imagen no codificada y valores de predicción correspondientes a las PU. El codificador de vídeo 20 puede formar las TU incluyendo los datos residuales para la CU y, a continuación, transformar las TU para generar coeficientes de transformación para la CU.

[0058] Tras cualquier transformación para generar coeficientes de transformación, el codificador de vídeo 20 puede realizar la cuantización de los coeficientes de transformación. La cuantización se refiere, en general, a un proceso en el que unos coeficientes de transformación se cuantizan para reducir posiblemente la cantidad de datos usados para representar los coeficientes, proporcionando compresión adicional. El proceso de cuantización puede reducir la profundidad de bits asociada a algunos, o a la totalidad, de los coeficientes. Por ejemplo, un valor de n bits puede redondearse a la baja hasta un valor de m bits durante la cuantización, donde n es mayor que m .

[0059] Después de la cuantización, el codificador de vídeo puede recorrer los coeficientes de transformación, produciendo un vector unidimensional a partir de la matriz bidimensional que incluye los coeficientes de transformación cuantizados. El recorrido puede estar diseñado para colocar los coeficientes de energía más alta (y por lo tanto de menor frecuencia) al frente de la formación y para colocar los coeficientes de energía más baja (y por lo tanto de mayor frecuencia) al final de la formación. En algunos ejemplos, el codificador de vídeo 20 puede utilizar un orden de exploración predeterminado para explorar los coeficientes de transformación cuantizados, para producir un vector serializado que se pueda codificar por entropía. En otros ejemplos, el codificador de vídeo 20 puede realizar una exploración adaptativa. Después de recorrer los coeficientes de transformación cuantizados para formar un vector unidimensional, el codificador de vídeo 20 puede realizar la codificación por entropía del vector unidimensional, por ejemplo, de acuerdo a la codificación de longitud variable adaptable al contexto (CAVLC), la codificación aritmética binaria adaptable al contexto (CABAC), la codificación aritmética binaria adaptable al contexto basada en la sintaxis (SBAC), la codificación por entropía por división de intervalos de probabilidad (PIPE) o con otra metodología de codificación por entropía. El codificador de vídeo 20 también puede realizar la codificación por entropía de los elementos sintácticos asociados a los datos de vídeo codificados, para su uso por el decodificador de vídeo 30 en la decodificación de los datos de vídeo.

[0060] Para realizar la codificación CABAC, el codificador de vídeo 20 puede asignar un contexto, dentro de un modelo de contexto, a un símbolo a transmitir. El contexto se puede referir, por ejemplo, a si los valores contiguos del símbolo son distintos de cero o no. Para realizar la CAVLC, el codificador de vídeo 20 puede seleccionar un código de longitud variable para un símbolo a transmitir. Las palabras de código en la VLC se pueden construir de modo que los códigos relativamente más cortos correspondan a símbolos más probables, mientras que los códigos más largos correspondan a símbolos menos probables. De esta forma, el uso de la VLC puede lograr un ahorro en bits con respecto, por ejemplo, al uso de palabras de código de igual longitud para cada símbolo que se va a transmitir. La determinación de la probabilidad se puede basar en un contexto asignado al símbolo.

[0061] Además, 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 grupos de imágenes (GOP), al decodificador 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 una serie de imágenes en el GOP respectivo, y los datos sintácticos de imagen pueden indicar una modalidad de codificación/predicción usada para codificar la imagen correspondiente.

[0062] En algunos ejemplos, el codificador de vídeo 20 puede generar y el decodificador de vídeo 30 puede recibir ciertos conjuntos de parámetros, que pueden ser usados al decodificar datos de vídeo. Por ejemplo, los conjuntos de parámetros pueden incluir un SPS, PPS o VPS, que mejoran la eficacia al señalar por separado la información que cambia con poca frecuencia. 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 de capa de abstracción de red (NAL) de conjuntos de parámetros se pueden transmitir por un canal diferente al de otras unidades de NAL, tales como las unidades de NAL de información de mejora complementaria (SEI).

[0063] Las unidades de NAL de SEI (denominadas mensajes de SEI) pueden contener información que no es necesaria para la decodificación de las muestras de imágenes codificadas a partir de unidades de NAL de VCL, pero

pueden ayudar en los procesos relacionados con la decodificación, la visualización, la capacidad de recuperación tras errores y otros fines. Los mensajes de SEI pueden estar contenidos en las unidades de NAL no de VCL. Los mensajes de SEI se pueden incluir en la parte normativa de algunas especificaciones normativas y, por tanto, no siempre son obligatorios para la implementación del decodificador 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. Alguna información a nivel de secuencia puede estar contenida en los mensajes de SEI, tales como los mensajes de SEI de información de ajustabilidad a escala en el ejemplo de la SVC y los mensajes de SEI de información de ajustabilidad a escala de vistas en la MVC.

[0064] En la HEVC, el codificador de vídeo 20 puede definir puntos de operación para indicar un subflujo de bits que puede extraerse de un flujo de bits codificado en función del máximo nivel temporal del subflujo de bits, así como una lista de `nuh_reserved_zero_6bits`. Por ejemplo, el codificador de vídeo 20 puede señalar puntos de operación de acuerdo a la Tabla 1, que se muestra a continuación:

TABLA 1: Puntos de Operación

<code>operation_point_layer_ids(opIdx) {</code>	Descriptor
<code> for(i = 0; i <= vps_max_nuh_reserved_zero_layer_id ; i++)</code>	
<code> op_layer_id_included_flag[opIdx][i]</code>	<code>u(1)</code>
<code>}</code>	

[0065] En la Tabla 1 anterior, el indicador `id_capa_op_incluido [opIdx] [i]` igual a 1 puede especificar que la capa con `id_capa` igual a `i` está presente para el `opIdx`-ésimo conjunto de valores de `id_capa`. Además, el indicador `id_capa_op_incluido [opIdx] [i]` igual a 0 puede especificar que la capa con `id_capa` igual a `i` no está presente para el `opIdx`-ésimo conjunto de valores de `id_capa`.

[0066] El codificador de vídeo 20 también puede señalar el elemento sintáctico `vps_max_nuh_reservado_id_capa_cero` en el VPS, donde `vps_max_nuh_reservado_id_capa_cero` especifica el máximo valor de `nuh_reserved_zero_6bits` para todas las unidades de NAL en la secuencia de vídeo codificado en el `ConjuntoldCapaOp` de los puntos de operación a los cuales se aplica la `opIdx`-ésima estructura sintáctica `hrd_parámetros()` en el conjunto de parámetros de vídeo.

[0067] En algunos casos, el codificador de vídeo 20 puede señalar puntos de operación para su uso con la extensión de múltiples vistas para la HEVC (MV-HEVC) o la 3D-HEVC. El más reciente borrador de trabajo de la extensión de múltiples vistas de la MV-HEVC y la 3D-HEVC se especifica en el "MV-HEVC Working Draft 1" ["Borrador de Trabajo 1 de la MV-HEVC"], Equipo de Colaboración Conjunta sobre Desarrollo de Extensión de Codificación de vídeo tridimensional del UIT-T SG 16 WP 3 e ISO / IEC JTC 1 / SC 29 / WG 11, 1ª Conferencia: Estocolmo, SE, 16 a 20 de julio de 2012, JCT3V-A1004, disponible en http://phenix.it-sudparis.eu/jct2/doc_end_user/documents/1_Stockholm/wg11/JCT3V-A1004-v1.zip y "3D-HEVC Test Model 1" ["Modelo 1 de Prueba de la 3D-HEVC"], Equipo de Colaboración Conjunta sobre Desarrollo de Extensión de Codificación de vídeo tridimensional del ITU-T SG 16 WP 3 e ISO/IEC JTC 1/SC 29/WG 11, 1ª Conferencia: Estocolmo, SE, 16 a 20 de julio de 2012, JCT3V-A1005, disponible en http://phenix.int-evry.fr/jct2/doc_end_user/documents/1_Stockholm/wg11/JCT3V-A1005-v1.zip, respectivamente.

[0068] Los puntos de operación de la MV-HEVC y / o la 3D0-HEVC pueden señalizarse únicamente en el conjunto de parámetros de vídeo. Por ejemplo, el codificador de vídeo 20 puede señalar que las vistas están presentes y decodificadas para cada punto de operación.

[0069] Como se ha indicado anteriormente, los puntos de operación definidos en la HEVC pueden ser suficientes para fines de extracción del flujo de bits, pero el VPS no especifica vistas / capas de salida de destino de los puntos de operación. Es decir, el VPS no identifica las capas particulares de datos de vídeo que se pretende mostrar.

[0070] Sin saber cuáles capas son capas de salida de destino, las imágenes decodificadas de ciertas capas que no se usan para salida pueden mantenerse en un almacén temporal de imágenes decodificadas (DPB) del codificador de vídeo 20 y / o del decodificador de vídeo 30 por más tiempo que el necesario, aumentando por ello los requisitos de almacenamiento para secuencias de múltiples vistas. Además, cuando se incluyen múltiples vistas en un subflujo de bits, un proveedor de contenido asociado al servidor / red de entrega de contenido 34 puede querer controlar cuáles subflujos de bits son decodificados por el decodificador de vídeo 30 y proporcionados como una salida al dispositivo de visualización 32 para controlar la experiencia visual. Tal control puede no estar disponible con el esquema actual del VPS.

[0071] En general, el codificador de vídeo 20 y el decodificador de vídeo 30 pueden configurarse para realizar una o más de las diversas técnicas de esta divulgación, solas o en cualquier combinación. Por ejemplo, según aspectos de esta divulgación, el codificador de vídeo 20 puede indicar, con uno o más elementos sintácticos en un flujo de bits

codificado, datos indicativos de un punto de operación, por ejemplo, un conjunto de capas, que tiene una pluralidad de capas de datos de vídeo. Además, el codificador de vídeo 20 puede codificar uno o más elementos sintácticos que indican un punto de operación de salida para el punto de operación, donde el punto de operación de salida está asociado a una o más capas de salida de destino entre la pluralidad de capas de los datos de vídeo.

5 **[0072]** Del mismo modo, el decodificador de vídeo 30 puede obtener, a partir de un flujo de bits codificado, datos indicativos de un punto de operación, por ejemplo, un conjunto de capas, que tiene una pluralidad de capas de datos de vídeo. Además, el decodificador de vídeo 30 puede determinar, en función de uno o más elementos sintácticos del flujo de bits codificado, un punto de operación de salida para el punto de operación, donde el punto de operación de salida está asociado a una o más capas de salida de destino entre la pluralidad de capas de datos de vídeo.

10 **[0073]** Por ejemplo, de acuerdo a aspectos de esta divulgación, para una extensión de múltiples vistas para la HEVC (MV-HEVC) o una extensión de vídeo tridimensional para la HEVC (3DV, que puede incluir una o más capas de datos de profundidad), el codificador de vídeo 20 puede señalar capas de salida de destino para cada punto de operación de salida en un VPS o extensión de VPS. El decodificador de vídeo 30 puede recibir dicha señalización y determinar las capas de salida de destino para cada punto de operación de salida.

15 **[0074]** Como ejemplo, con respecto a la MV-HEVC, un punto de operación de salida puede referirse a un conjunto de capas con una lista de vistas de salida de destino identificadas que pertenecen a las vistas incluidas en el punto de operación. El codificador de vídeo 20 puede señalar el punto de operación de salida en una extensión de VPS, como se muestra en el ejemplo de la Tabla 2 a continuación:

TABLA 2 - Extensión de VPS con punto de operación de salida

vps_extension() {	Descriptor
while(!byte_aligned())	
vps_extension_byte_alignment_reserved_one_bit	u(1)
num_layer_operation_points	u(10)
num_additional_profile_level_sets	u(8)
for(i = 0; i <= vps_max_layers_minus1; i++) {	
num_types_zero_4bits[i]	u(4)
type_zero_4bits[i]	u(4)
view_id[i]	u(8)
if(i > 0)	
num_direct_ref_layers[i]	u(6)
for(j = 0; j < num_direct_ref_layers[i]; j++)	
ref_layer_id[i][j]	u(6)
}	
for(i = 0; i <= num_additional_profile_level_sets; i++)	
profile_tier_level(1, vps_max_sub_layers_minus1)	
for(i = 0; i < num_layer_operation_points; i++) {	
op_point_index	u(10)
for(j = 0; j <= vps_max_nuh_reserved_zero_layer_id; j++)	
if(op_layer_id_included_flag[op_point_index][i])	
output_layer_flag[op_point_index][j]	u(1)
if(num_additional_profile_level_sets > 0)	
profile_level_idx[i]	u(8)
}	
}	

- 5 **[0075]** En el ejemplo de la Tabla 2 anterior, num_puntos_operación_capa puede especificar el número máximo de puntos de operación de salida presentes en las secuencias de vídeo codificadas a las que se aplica el conjunto de parámetros de vídeo. Además, índice_punto_op puede identificar el punto de operación en función del cual se forma el i-ésimo punto de operación de salida actual. Además, el indicador_capa_salida [índice_punto_op][j] igual a 1 puede indicar que la capa con id_capa igual a j es el destino de salida para el i-ésimo punto de operación de salida actual. Además, el indicador_capa_salida [j] igual a 0 puede indicar que la capa con id_capa igual a j no es el destino de salida para el i-ésimo punto de operación de salida actual.
- 10 **[0076]** Como otro ejemplo, el codificador de vídeo 20 puede señalar el punto de operación de salida en un VPS, como se muestra en el ejemplo de la Tabla 3 a continuación:

TABLA 3 - VPS con punto de operación de salida

for(i = 0; i < num_layer_operation_points; i++) {	
op_point_index	u(10)
num_target_views[i]	u(6)
for (j = 0 ; j < num_target_views ; j++)	
target_layer_id[i][j]	u(6)
if (num_additional_profile_level_sets)	
profile_level_idx[i]	u(8)
}	

- 15
- [0077]** En el ejemplo de la Tabla 3 anterior, num_vistas_destino puede especificar el número de vistas de salida de destino para el i-ésimo punto de operación de salida actual. Además, el id_capa_destino [i][j] puede especificar el id_capa de la j-ésima vista de salida de destino del i-ésimo punto de operación de salida actual.
- 20 **[0078]** En otro ejemplo más, el codificador de vídeo 20 puede señalar vistas de salida de destino utilizando otras estructuras sintácticas. Por ejemplo, el codificador de vídeo 20 puede codificar datos que indiquen vistas de salida de destino en un mensaje de SEI, o utilizando otras técnicas de señalización fuera de banda.
- 25 **[0079]** Con respecto a la 3D-HEVC, un punto de operación de salida puede referirse a un punto de operación, por ejemplo, un conjunto de capas, con una lista de vistas identificadas de salida de destino, de textura o profundidad, que pertenecen a las vistas incluidas en el punto de operación. El codificador de vídeo 20 puede señalar el punto de operación de salida para la 3D-HEVC utilizando sintaxis y semántica similares a las descritas anteriormente con respecto a la MV-HEVC. Sin embargo, para la 3D-HEVC, el codificador de vídeo 20 de cada id_capa señalado como una vista de salida de destino puede corresponder a una vista de profundidad o de textura.
- 30 **[0080]** Con respecto a la HSVC, la capa de salida de destino de un punto de operación puede obtenerse como la capa con el id_capa numéricamente máximo. Por ejemplo, un punto de operación de salida puede ser un punto de operación, por ejemplo, un conjunto de capas, con solo la capa que tiene el id_capa máximo en el punto de operación que se considera para la salida. Por lo tanto, el decodificador de vídeo 30 puede decodificar un elemento sintáctico id_capa para determinar el punto de operación de salida para un punto de operación particular.
- 35 **[0081]** Por lo tanto, de acuerdo a aspectos de esta divulgación, un punto de operación, según lo definido en la especificación básica actual de la HEVC, puede corresponder a uno o más puntos de operación de salida, teniendo cada punto de operación de salida diferentes vistas y / o capas de salida de destino. Por ejemplo, el codificador de vídeo 20 puede señalar en un flujo de bits codificado (y el decodificador de vídeo 30 puede obtener al analizar sintácticamente un flujo de bits codificado) una indicación de que un subconjunto de las capas decodificadas puede ser deseable para su salida, por ejemplo, su visualización. Si bien se describe con respecto al codificador de vídeo 20 anterior, en algunos ejemplos, el servidor / la red de entrega de contenido 34 puede ser responsable de dicha señalización.
- 40 **[0082]** Por ejemplo, el servidor / la red de entrega de contenido 34 puede determinar cuáles capas de un flujo de bits son óptimas para su salida. En este ejemplo, el servidor / la red de entrega de contenido 34 puede ser responsable de generar y / o alterar un VPS, una extensión de VPS, un mensaje de SEI u otra estructura para indicar las capas de salida de destino para un punto de operación de salida.
- 45 **[0083]** Con respecto a la gestión de la memoria intermedia de imágenes decodificadas (también denominada memoria de imágenes de referencia, como se ha descrito, por ejemplo, con respecto a las Figuras 2 y 3), de acuerdo
- 50

a los aspectos de esta divulgación, cuando se está decodificando un punto de operación de salida, el codificador de vídeo 20 y / o el decodificador de vídeo 30 pueden eliminar una imagen decodificada de una capa que no está destinada a la salida desde la memoria intermedia de imágenes decodificadas, siempre que la imagen ya no sea necesaria para la predicción temporal (por ejemplo, la interpredicción dentro de una capa o una vista) o la predicción entre capas / entre vistas. Por ejemplo, para facilitar la eliminación, el codificador de vídeo 20 puede fijar un IndicadorSalidalmg, de cada componente de vista de una capa que no pertenece a una vista o capa de salida de destino, igual a 0. Por lo tanto, el codificador de vídeo 20 y / o el decodificador de vídeo 30 pueden determinar cuáles imágenes no se utilizan como referencia y no como salida, y eliminar dichas imágenes del almacén temporal de imágenes decodificadas en función de la determinación.

[0084] La figura 2 es un diagrama de bloques que ilustra un ejemplo de un codificador de vídeo 20 que puede implementar técnicas para determinar una o más capas de salida de destino, según aspectos de esta divulgación. El codificador de vídeo 20 puede realizar la intracodificación y la intercodificación de bloques de vídeo dentro de fragmentos de vídeo. La intracodificación se basa en la predicción espacial para reducir o eliminar la redundancia espacial en el vídeo dentro de una imagen de vídeo dada. La intercodificación se basa en la predicción temporal para reducir o eliminar la redundancia temporal en el vídeo dentro de tramas adyacentes de una secuencia de vídeo. La intramodalidad (modalidad I) puede referirse a cualquiera de varias modalidades de codificación de base espacial. Las intermodalidades, tales como la predicción unidireccional (modalidad P) o la bipredicción (modalidad B), pueden referirse a cualquiera de varias modalidades de codificación de base temporal.

[0085] Como se ha indicado anteriormente, el codificador de vídeo 20 puede adaptarse para realizar una codificación de vídeo de múltiples vistas y / o ajustable a escala. En algunos casos, el codificador de vídeo 20 puede configurarse para ajustarse a una norma de codificación de vídeo MV-HEVC, de modo que cada vista en una instancia temporal pueda ser procesada por un decodificador, tal como el decodificador de vídeo 30. En otros casos, el codificador de vídeo 20 puede configurarse para ajustarse a una norma de codificación de vídeo 3D-HEVC, y puede configurarse para codificar un mapa de profundidad para cada vista, además de codificar mapas de textura (es decir, valores de luma y croma) para cada vista. En otros casos más, el codificador de vídeo 20 puede configurarse para codificar múltiples capas ajustables a escala de datos de vídeo, de acuerdo a una norma de codificación de vídeo HSVC. Si bien se hace referencia a normas de codificación específicas, debería entenderse que las técnicas no son específicas para ninguna norma de codificación y que pueden implementarse con normas futuras y / o aún no desarrolladas.

[0086] En cualquier caso, como se muestra en la figura 2, el codificador de vídeo 20 recibe un bloque de vídeo actual dentro de una trama de vídeo a codificar. En el ejemplo de la figura 2, el codificador de vídeo 20 incluye una unidad de selección de modalidad 40, una memoria de imágenes de referencia 64, un sumador 50, una unidad de procesamiento de transformación 52, una unidad de cuantización 54 y una unidad de codificación por entropía 56. A su vez, la unidad de selección de modalidad 40 incluye la unidad de compensación de movimiento 44, la unidad de estimación de movimiento 42, la unidad de intrapredicción 46 y la unidad de división 48. Para la reconstrucción de bloques de vídeo, el codificador de vídeo 20 incluye también la unidad de cuantización inversa 58, la unidad de transformación inversa 60 y el sumador 62. También se puede incluir un filtro de eliminación de bloques (no mostrado en la figura 2) para filtrar fronteras de bloques, para eliminar distorsiones de efecto pixelado del vídeo reconstruido. Si se desea, el filtro de eliminación de bloques filtrará habitualmente la salida del sumador 62. También pueden usarse filtros 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 el bucle).

[0087] Durante el proceso de codificación, el codificador de vídeo 20 recibe una imagen o un fragmento de vídeo a codificar. La imagen o el fragmento pueden dividirse en múltiples bloques de vídeo. La unidad de estimación de movimiento 42 y la unidad de compensación de movimiento 44 llevan a cabo la codificación interpredictiva del bloque de vídeo recibido con respecto a uno o más bloques en una o más imágenes de referencia, para proporcionar la predicción temporal. La unidad de intrapredicción 46, de forma alternativa, puede llevar a cabo la codificación intrapredictiva del bloque de vídeo recibido, con respecto a uno o más bloques contiguos en la misma imagen o fragmento que el bloque a codificar, para proporcionar predicción espacial. El codificador de vídeo 20 puede llevar a cabo múltiples pases de codificación, por ejemplo, para seleccionar una modalidad de codificación adecuada para cada bloque de datos de vídeo.

[0088] Además, la unidad de división 48 puede dividir bloques de datos de vídeo en subbloques, basándose en la evaluación de los anteriores esquemas de división en los anteriores pases de codificación. Por ejemplo, la unidad de división 48 puede dividir inicialmente una imagen o fragmento en varias LCU, y dividir cada una de las LCU en varias sub-CU, basándose en un análisis de velocidad-distorsión (por ejemplo, optimización de velocidad-distorsión). La unidad de selección de modalidad 40 puede producir además una estructura de datos de árbol cuádruple, indicativa de la división de una LCU en las sub-CU. Las CU de nodos hoja del árbol cuádruple pueden incluir una o más PU y una o más TU.

[0089] La unidad de selección de modalidad 40 puede seleccionar una de las modalidades de codificación, intra o inter, por ejemplo, basándose en los resultados de errores, y proporciona el bloque intracodificado 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 modalidad 40 proporciona también

elementos sintácticos, tales como vectores de movimiento, indicadores de intramodalidad, información de división y otra información sintáctica de este tipo, a la unidad de codificación por entropía 56.

[0090] La unidad de estimación de movimiento 42 y la unidad de compensación de movimiento 44 pueden estar sumamente integradas, pero se ilustran por separado con fines conceptuales. La estimación del movimiento, realizada por la unidad de estimación de movimiento 42, es el proceso de generación de vectores de movimiento, que estiman el movimiento para 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 de vídeo 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). Un bloque predictivo es un bloque que se revela como estrechamente coincidente con el bloque a codificar, en lo que respecta a la diferencia de píxeles, lo cual puede determinarse mediante la suma de diferencias absolutas (SAD), suma de las diferencias al cuadrado (SSD) u otras métricas de diferencia. En algunos ejemplos, el codificador de vídeo 20 puede calcular valores para posiciones fraccionarias de píxeles de imágenes de referencia almacenadas en la memoria de imágenes de referencia 64. Por ejemplo, el codificador de vídeo 20 puede interpolar valores de posiciones de un cuarto de píxel, posiciones de un octavo de píxel u otras posiciones fraccionarias de píxel de la imagen de referencia. Por lo tanto, la unidad de estimación de movimiento 42 puede realizar una búsqueda de movimiento relativa a las posiciones de píxel completo y las posiciones de píxel fraccionario, y emitir un vector de movimiento con una precisión de píxel fraccionario.

[0091] La unidad de estimación de movimiento 42 calcula un vector 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 imagen de referencia puede seleccionarse a partir de una primera lista (lista 0) de imágenes de referencia o una segunda lista (lista 1) de imágenes de referencia, cada una de las cuales identifica una o más imágenes de referencia almacenadas en una memoria de imágenes de referencia 64. La unidad de estimación de movimiento 42 envía el vector de movimiento calculado a la unidad de codificación por entropía 56 y a la unidad de compensación de movimiento 44.

[0092] La compensación de movimiento, llevada a cabo por la unidad de compensación de movimiento 44, puede implicar capturar o generar el bloque predictivo basándose en el vector de movimiento determinado por la unidad de estimación de movimiento 42. De nuevo, la unidad de estimación de movimiento 42 y la unidad de compensación de movimiento 44 pueden integrarse 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 localizar el bloque predictivo al que apunta el vector de movimiento en una de las listas de imágenes de referencia. El sumador 50 forma un bloque de vídeo residual restando los valores de píxel del bloque predictivo a los valores de píxel del bloque de vídeo actual que se está codificando, generando valores de diferencias de píxel, como se analiza posteriormente. En general, la unidad de estimación de movimiento 42 lleva a cabo la estimación de movimiento con respecto a los componentes de luma, y la unidad de compensación de movimiento 44 utiliza los vectores de movimiento calculados basándose en los componentes de luma, tanto para los componentes de croma como para los componentes de luma. La unidad de selección de modalidad 40 también puede generar elementos sintácticos asociados a los bloques de vídeo y el fragmento de vídeo para su uso por el decodificador de vídeo 30 en la decodificación de los bloques de vídeo del fragmento de vídeo.

[0093] La unidad de intrapredicción 46 puede intrapredecir un bloque actual, como alternativa a la interpredicción llevada a cabo por la unidad de estimación de movimiento 42 y la unidad de compensación de movimiento 44, como se ha descrito anteriormente. En particular, la unidad de intrapredicción 46 puede determinar una modalidad de intrapredicción a usar para codificar un bloque actual. En algunos ejemplos, la unidad de intrapredicción 46 puede codificar un bloque actual usando diversas modalidades de intrapredicción, por ejemplo, durante pases de codificación independientes, y la unidad de intrapredicción 46 (o la unidad de selección de modalidad 40, en algunos ejemplos) puede seleccionar una modalidad de intrapredicción adecuada para usar, a partir de las modalidades probadas.

[0094] Por ejemplo, la unidad de intrapredicción 46 puede calcular valores de velocidad-distorsión usando un análisis de velocidad-distorsión para las diversas modalidades de intrapredicción probadas, y seleccionar la modalidad de intrapredicción que tenga las mejores características de velocidad-distorsión entre las modalidades probadas. El análisis de velocidad-distorsión 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 intrapredicción 46 puede calcular proporciones a partir de las distorsiones y velocidades para los diversos bloques codificados, para determinar qué modalidad de intrapredicción presenta el mejor valor de velocidad-distorsión para el bloque.

[0095] Después de seleccionar una modalidad de intrapredicción para un bloque, la unidad de intrapredicción 46 puede proporcionar información, indicativa de la modalidad de intrapredicción seleccionada para el bloque, a la unidad de codificación por entropía 56. La unidad de codificación por entropía 56 puede codificar la información que indica la modalidad de intrapredicción seleccionada. El codificador de vídeo 20 puede incluir, en el flujo de bits transmitido, datos de configuración, que pueden incluir una pluralidad de tablas de índices de modalidad de intrapredicción y una pluralidad de tablas modificadas de índices de modalidad de intrapredicción (también denominadas tablas de correlación de palabras de código), definiciones de contextos de codificación para diversos bloques e indicaciones de

una modalidad de intrapredicción más probable, una tabla de índices de modalidad de intrapredicción y una tabla modificada de índices de modalidad de intrapredicción, a usar para cada uno de los contextos.

5 **[0096]** El codificador 20 de vídeo forma un bloque de vídeo residual restando los datos de predicción de la unidad de selección de modalidad 40 del bloque de vídeo original que se está codificando. El sumador 50 representa el componente o los componentes que realizan esta operación de resta. La unidad de procesamiento de transformación 52 aplica una transformación, tal como una transformación discreta de coseno (DCT) o una transformación conceptualmente similar, al bloque residual, produciendo un bloque de vídeo que comprende valores residuales de coeficientes de transformación. La unidad de procesamiento de transformación 52 puede llevar a cabo otras transformaciones que son conceptualmente similares a la DCT. También se podrían usar transformaciones de ondulaciones, transformaciones de números enteros, transformaciones de subbandas u otros tipos de transformaciones.

15 **[0097]** En cualquier caso, la unidad de procesamiento de transformación 52 aplica la transformación al bloque residual, produciendo un bloque de coeficientes de transformación residuales. La transformación puede convertir la información residual, desde un dominio de valores de píxel a un dominio de transformación, tal como un dominio de frecuencia. La unidad de procesamiento de transformación 52 puede enviar los coeficientes de transformación resultantes a la unidad de cuantización 54. La unidad de cuantización 54 cuantiza los coeficientes de transformación para reducir más la velocidad de bits. El proceso de cuantización puede reducir la profundidad de bits asociada a algunos, o a la totalidad, de los coeficientes. El grado de cuantización se puede modificar ajustando un parámetro de cuantización. En algunos ejemplos, la unidad 54 de cuantización puede realizar, a continuación, una exploración de la matriz que incluye los coeficientes de transformación cuantizados. De forma alternativa, la unidad de codificación de entropía 56 puede realizar la exploración.

25 **[0098]** Tras la cuantización, la unidad de codificación por entropía 56 codifica por entropía los coeficientes de transformación cuantizados. 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 parte de la unidad de codificación por entropía 56, el flujo de bits codificado se puede transmitir a otro dispositivo (por ejemplo, el decodificador de vídeo 30) o archivar para su posterior transmisión o recuperación.

35 **[0099]** La unidad de cuantización inversa 58 y la unidad de transformación inversa 60 aplican la cuantización inversa y la transformación inversa, respectivamente, para reconstruir el bloque residual en el dominio del píxel, por ejemplo, para su uso posterior 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 puede ser utilizado por la unidad de estimación de movimiento 42 y la unidad de compensación de movimiento 44, como bloque de referencia para intercodificar un bloque en una imagen de vídeo posterior.

45 **[0100]** El codificador de vídeo 20 puede generar una serie de elementos sintácticos, como se ha descrito anteriormente, que pueden ser codificados 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 múltiples vistas y / o múltiples capas, como se ha descrito anteriormente.

50 **[0101]** Por ejemplo, según aspectos de esta divulgación, la unidad de codificación por entropía 56 puede codificar uno o más elementos sintácticos en un flujo de bits codificado que indica un punto de operación, por ejemplo, un conjunto de capas, con una pluralidad de capas de datos de vídeo. Además, la unidad de codificación por entropía 56 puede codificar uno o más elementos sintácticos que indican un punto de operación de salida para el conjunto de capas, donde el punto de operación de salida está asociado a una o más capas de salida de destino entre la pluralidad de capas de los datos de vídeo.

55 **[0102]** En algunos ejemplos, de acuerdo a aspectos de esta divulgación, el codificador de vídeo 20 puede indicar capas de salida de destino para un punto de operación de salida, de acuerdo a una norma de codificación de vídeo, tal como MV-HEVC, 3D-HEVC o HSVC. Por ejemplo, con respecto a la codificación de vídeo de múltiples vistas (por ejemplo, la MV-HEVC o la 3D-HEVC), el codificador de vídeo 20 puede codificar una extensión de VPS (por ejemplo, como la que se muestra y describe con respecto al ejemplo de la Tabla 2 anterior), un VPS (por ejemplo, como el que se muestra y describe con respecto al ejemplo de la Tabla 3 anterior), un mensaje de SEI u otro mensaje para indicar uno o más puntos de operación de salida, teniendo cada punto de operación de salida capas de salida de destino asociadas. Como se ha indicado anteriormente, cuando se codifican vistas de profundidad, un punto de operación de salida puede referirse a un punto de operación con una lista de vistas identificadas de textura o profundidad, de salida

de destino, que pertenecen a las vistas incluidas en el punto de operación.

[0103] Con respecto a la HSVC, el codificador de vídeo 20 puede codificar un identificador de capa para cada capa de datos de vídeo, para identificar cada capa de datos de vídeo (como se describe con mayor detalle, por ejemplo, con respecto a la Figura 5 a continuación). De acuerdo a aspectos de esta divulgación, un punto de operación de salida puede ser un punto de operación con solo la capa que tiene el mayor id_capa en el punto de operación que se considera para la salida.

[0104] Por lo tanto, de acuerdo a aspectos de esta divulgación, un punto de operación, según lo definido en la especificación básica actual de la HEVC, puede corresponder a uno o más puntos de operación de salida, teniendo cada punto de operación de salida diferentes vistas y / o capas de salida de destino. Por ejemplo, el codificador de vídeo 20 puede señalar en un flujo de bits codificado una indicación de que un subconjunto de las capas decodificadas puede ser deseable para la salida, por ejemplo, la visualización.

[0105] De acuerdo a algunos aspectos de esta divulgación, el codificador de vídeo 20 puede eliminar una o más imágenes de la memoria de imágenes de referencia 64 en función del punto de operación de salida determinado y las capas de salida de destino. Por ejemplo, el codificador de vídeo 20 puede eliminar imágenes de la memoria de imágenes de referencia 64 que no están destinadas a la salida y que no son necesarias para la predicción temporal (por ejemplo, la interpredicción dentro de una capa o una vista) o la predicción entre capas / entre vistas.

[0106] De esta manera, el codificador de vídeo 20 representa un ejemplo de un codificador de vídeo configurado para codificar una pluralidad de capas de datos de vídeo, incluyendo una pluralidad de conjuntos de capas, donde cada conjunto de capas contiene una o más capas de datos de vídeo entre la pluralidad de capas. El codificador de vídeo 20 también está configurado para codificar uno o más elementos sintácticos de un flujo de bits que indica uno o más puntos de operación de salida, donde cada punto de operación de salida está asociado a un conjunto de capas entre la pluralidad de conjuntos de capas y una o más capas de salida de destino entre la pluralidad de capas.

[0107] La figura 3 es un diagrama de bloques que ilustra un ejemplo de un decodificador de vídeo 30 que puede implementar técnicas para determinar una o más capas de salida de destino, según aspectos de esta divulgación. En el ejemplo de la figura 3, el decodificador de vídeo 30 incluye una unidad de decodificación por entropía 70, una unidad de compensación de movimiento 72, una unidad de intrapredicción 74, una unidad de cuantización inversa 76, una unidad de transformación inversa 78, una memoria de imágenes de referencia 82 y un sumador 80.

[0108] Como se ha indicado anteriormente, el decodificador de vídeo 30 puede adaptarse para realizar la codificación de vídeo de múltiples vistas y / o ajustable a escala. En algunos casos, el decodificador de vídeo 30 puede configurarse para ajustarse a una norma de codificación de vídeo MV-HEVC. En otros casos, el decodificador de vídeo 30 puede configurarse para ajustarse a una norma de codificación de vídeo 3D-HEVC, y puede configurarse para decodificar un mapa de profundidad para cada vista, además de codificar mapas de textura (es decir, valores de luma y croma) para cada vista. En otros casos más, el decodificador de vídeo 30 puede configurarse para decodificar múltiples capas escalables de datos de vídeo de acuerdo a una norma de codificación de vídeo HSVC. Si bien se hace referencia a normas de codificación específicas, debería entenderse que las técnicas no son específicas para ninguna norma de codificación y que pueden implementarse con normas futuras y / o aún no desarrolladas.

[0109] En algunos ejemplos, el decodificador de vídeo 30 puede realizar un pase de decodificación, generalmente recíproco al pase de codificación descrito con respecto al codificador de vídeo 20 (figura 2). La unidad de compensación de movimiento 72 puede generar datos de predicción basándose en vectores de movimiento recibidos desde la unidad de decodificación por entropía 70, mientras que la unidad de intrapredicción 74 puede generar datos de predicción basándose en indicadores de modalidad de intrapredicción, recibidos desde la unidad de decodificación por entropía 70.

[0110] Durante el proceso de decodificación, el decodificador de vídeo 30 recibe un flujo de bits de vídeo codificado, que representa bloques de vídeo de un fragmento de vídeo codificado y elementos sintácticos asociados, desde el codificador de vídeo 20. El decodificador de vídeo 30 puede recibir los elementos sintácticos en el nivel de fragmento de vídeo y/o el nivel de bloque de vídeo.

[0111] La unidad de decodificación por entropía 70 decodifica el flujo de bits para generar coeficientes cuantizados, vectores de movimiento o indicadores de modalidad de intrapredicción y otros elementos sintácticos. En algunos ejemplos, la unidad de decodificación de entropía 70 u otra unidad de decodificador de vídeo puede analizar sintácticamente y decodificar elementos sintácticos a partir de un flujo de bits de múltiples vistas y / o múltiples capas, como se ha descrito anteriormente.

[0112] De acuerdo a aspectos de esta divulgación, la unidad de decodificación por entropía 70 puede obtener, a partir de un flujo de bits codificado, datos indicativos de un punto de operación, por ejemplo, un conjunto de capas, que tiene una pluralidad de capas de datos de vídeo. Además, la unidad de decodificación por entropía 70 puede decodificar uno o más elementos sintácticos que indican un punto de operación de salida para el punto de operación, donde el punto de operación de salida está asociado a una o más capas de salida de destino entre la pluralidad de

capas de los datos de vídeo. El decodificador de vídeo 30 puede determinar, basándose en uno o más elementos sintácticos, el punto de operación de salida y las capas de salida de destino.

5 **[0113]** En algunos ejemplos, según aspectos de esta divulgación, el decodificador de vídeo 30 puede ajustarse a una norma de codificación de vídeo, tal como la MV-HEVC, la 3D-HEVC o la HSVC. Por ejemplo, con respecto a la codificación de vídeo de múltiples vistas (por ejemplo, la MV-HEVC o la 3D-HEVC), el decodificador de vídeo 30 puede analizar sintácticamente subsecuencias de bits a partir de un flujo de bits codificado, en función de un punto de operación de salida y capas de salida de destino indicadas en una extensión de VPS (por ejemplo, tal como la mostrada y descrita con respecto al ejemplo de la Tabla 2 anterior), un VPS (por ejemplo, tal como el mostrado y descrito con respecto al ejemplo de la Tabla 3 anterior), un mensaje de SEI u otro mensaje. Cuando se decodifican vistas de profundidad, un punto de operación de salida puede referirse a un punto de operación con una lista de vistas identificadas, de profundidad o textura, de salida de destino, que pertenecen a las vistas incluidas en el punto de operación.

15 **[0114]** Con respecto a la HSVC, la unidad de decodificación por entropía 70 puede decodificar un id_capa para cada capa de datos de vídeo. De acuerdo a aspectos de esta divulgación, el decodificador de vídeo 30 puede deducir un punto de operación de salida basándose en la capa que tiene el id_capa más grande en el punto de operación.

20 **[0115]** Por lo tanto, de acuerdo a aspectos de esta divulgación, un punto de operación, según lo definido en la especificación básica actual de la HEVC, puede corresponder a uno o más puntos de operación de salida, teniendo cada punto de operación de salida diferentes vistas y / o capas de salida de destino. La unidad de decodificación por entropía 70 puede decodificar un subflujo de bits que incluye las capas de salida de destino basándose en uno o más elementos sintácticos analizados sintácticamente y decodificados a partir de un flujo de bits codificado recibido.

25 **[0116]** La unidad de decodificación por entropía 70 remite los vectores de movimiento y otros elementos sintácticos a la unidad de compensación de movimiento 72. Para una capa de salida en particular, cuando el fragmento de vídeo se codifica como un fragmento intracodificado (I), la unidad de intrapredicción 74 puede generar datos de predicción para un bloque de vídeo del fragmento de vídeo actual, basándose en una modalidad de intrapredicción señalizada y datos de bloques previamente decodificados de la imagen actual. Cuando la trama de vídeo se codifica como un fragmento intercodificado (es decir, B, P o GPB), la unidad de compensación de movimiento 72 genera bloques predictivos para un bloque de vídeo del fragmento de vídeo actual basándose en los vectores de movimiento y en otros elementos sintácticos recibidos desde la unidad de decodificación por entropía 70. Los bloques predictivos se pueden generar a partir de una de las imágenes de referencia dentro de una de las listas de imágenes de referencia. El decodificador de vídeo 30 puede construir las listas de imágenes de referencia, la Lista 0 y la Lista 1, usando técnicas de construcción por omisión, basándose en las imágenes de referencia almacenadas en la memoria de imágenes de referencia 82.

30 **[0117]** La unidad de compensación de movimiento 72 determina la información de predicción para un bloque de vídeo del fragmento de vídeo actual, analizando sintácticamente los vectores de movimiento y otros elementos sintácticos, y usa la información de predicción para generar los bloques predictivos para el bloque de vídeo actual que se está decodificando. Por ejemplo, la unidad de compensación de movimiento 72 usa algunos de los elementos sintácticos recibidos para determinar una modalidad de predicción (por ejemplo, intrapredicción o interpredicción) usada para codificar los bloques de vídeo del fragmento de vídeo, un tipo de fragmento de interpredicción (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, el estado de interpredicción para cada bloque de vídeo intercodificado del fragmento y otra información para decodificar los bloques de vídeo en el fragmento de vídeo actual.

35 **[0118]** La unidad de compensación de movimiento 72 puede usar el vector de movimiento decodificado para recuperar datos desde una imagen previamente decodificada, por ejemplo, desde la memoria de imágenes de referencia 82. La unidad de compensación de movimiento 72 también puede realizar la interpolación basándose en filtros de interpolación. La unidad de compensación de movimiento 72 puede usar filtros de interpolación como los usados por el codificador de vídeo 20 durante la codificación de los bloques de vídeo, para calcular valores interpolados para píxeles fraccionarios de los bloques de referencia. En este caso, la unidad de compensación de movimiento 72 puede determinar los filtros de interpolación usados por el codificador de vídeo 20 a partir de los elementos sintácticos recibidos y usar los filtros de interpolación para generar bloques predictivos.

40 **[0119]** La unidad de cuantización inversa 76 realiza la cuantización inversa, es decir, la decuantización, de los coeficientes de transformación cuantizados proporcionados en el flujo de bits y decodificados por la unidad de decodificación por entropía 70. El proceso de cuantización inversa puede incluir el uso de un parámetro de cuantización QPy, calculado por el decodificador de vídeo 30 para cada bloque de vídeo en el fragmento de vídeo, para determinar un grado de cuantización y, asimismo, un grado de cuantización inversa que debería aplicarse. La unidad de transformación inversa 78 aplica una transformación inversa, por ejemplo, una DCT inversa, una transformación entera inversa, o un proceso de transformación inversa conceptualmente similar, a los coeficientes de transformación, con el fin de generar bloques residuales en el dominio del píxel.

[0120] Después de que la unidad de compensación de movimiento 72 genera el bloque predictivo para el bloque de vídeo actual basándose en los vectores de movimiento y otros elementos sintácticos, el decodificador de vídeo 30 forma un bloque de vídeo decodificado sumando los bloques residuales de la unidad de transformación inversa 78 a los correspondientes bloques predictivos generados por la unidad de compensación de movimiento 72. El sumador 80 representa el componente o los componentes que realizan esta operación de suma. Si se desea, también se puede aplicar un filtro de eliminación de bloques para filtrar los bloques decodificados, 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 allanar las transiciones de píxeles o mejorar de otro modo la calidad del vídeo. Los bloques de vídeo decodificados en una imagen dada se almacenan a continuación en la memoria de imágenes de referencia 82, que almacena imágenes de referencia usadas para la posterior compensación de movimiento. La memoria de imágenes de referencia 82 almacena también vídeo decodificado para su presentación posterior en un dispositivo de visualización, tal como el dispositivo de visualización 32 de la figura 1.

[0121] De acuerdo a algunos aspectos de esta divulgación, el decodificador de vídeo 30 puede eliminar una o más imágenes de la memoria de imágenes de referencia 82, en función del punto de operación de salida determinado y las capas de salida de destino. Por ejemplo, el decodificador de vídeo 30 puede eliminar imágenes de la memoria de imágenes de referencia 82 que no estén destinadas para su salida y que no sean necesarias para la predicción temporal (por ejemplo, la interpredicción dentro de una capa o una vista) o la predicción entre capas / entre vistas. El decodificador de vídeo 30 puede determinar si las imágenes de la memoria de imágenes de referencia 82 se utilizan para la predicción o se incluyen en una capa de salida de destino en función de uno o más elementos sintácticos del flujo de bits decodificado. Por ejemplo, la unidad de decodificación por entropía 70 puede decodificar un IndicadorSalidalmg para cada componente de vista de una capa que indica si la imagen pertenece a una vista de salida de destino. El decodificador de vídeo 30 puede eliminar imágenes de la memoria de imágenes de referencia 82 que no pertenecen a la vista de salida de destino y que están marcadas como "no utilizadas como referencia" durante la decodificación.

[0122] Por consiguiente, en un ejemplo, el decodificador de vídeo 30 puede decodificar un punto de operación de salida entre uno o más puntos de parte de salida, asociados a un conjunto de capas, así como decodificar un conjunto de capas decodificadas pero no de salida, que incluyen una o más capas que no son una de las una o más capas de salida de destino, sino que pertenecen al conjunto de capas. Además, el decodificador de vídeo 30 puede marcar una primera imagen decodificada, incluida en las una o más capas de salida de destino en un almacén temporal de imágenes decodificadas, como una imagen que se usará para la salida, y marcar una segunda imagen decodificada, incluida en el conjunto de capas decodificadas pero que no han de ser capas de salida, en el almacén temporal de imágenes decodificadas, como una imagen que no ha de utilizarse para su salida, de modo que la imagen que no se va a usar para su salida se elimine del almacén temporal de imágenes decodificadas antes que la imagen que se usará para su salida. Además, el decodificador de vídeo 30 puede determinar si la imagen marcada como no a utilizar para su salida se usa para una entre la interpredicción y la predicción entre capas, y eliminar la imagen marcada como no a utilizar para su salida del almacén temporal de imágenes decodificadas cuando la imagen no a utilizar para su salida no se usa para una entre la interpredicción y la predicción entre capas.

[0123] De esta manera, el decodificador de vídeo 30 representa un ejemplo de un decodificador de vídeo 30 que puede obtener, a partir de un flujo de bits de múltiples capas, una pluralidad de capas de datos de vídeo que incluyen una pluralidad de conjuntos de capas, donde cada conjunto de capas contiene una o más capas de datos de vídeo entre la pluralidad de capas. El decodificador de vídeo 30 también puede determinar, en función de uno o más elementos sintácticos del flujo de bits, uno o más puntos de operación de salida, donde cada punto de operación de salida está asociado a un conjunto de capas entre la pluralidad de conjuntos de capas y una o más capas de salida de destino entre la pluralidad de capas.

[0124] La figura 4 es un diagrama conceptual que ilustra un patrón ejemplar de predicción de la MVC. Si bien la figura 4 se describe con respecto a las normas H.264 / AVC y MVC, debería entenderse que se puede usar un patrón de predicción similar con otros esquemas de codificación de vídeo de múltiples vistas, incluyendo la MV-HEVC y la 3D-HEVC (múltiples vistas más profundidad). Por lo tanto, las referencias a la MVC a continuación se aplican a la codificación de vídeo de múltiples vistas en general, y no están restringidas a la norma H.264 / MVC.

[0125] En el ejemplo de la figura 4, se ilustran ocho vistas (que tienen identificadores de vista "S0" a "S7") y se ilustran doce ubicaciones temporales ("T0" a "T11") para cada vista. Es decir, cada fila en la figura 4 corresponde a una vista, mientras que cada columna indica una ubicación temporal.

[0126] Aunque la MVC tiene una denominada visualización de base que es decodificable mediante los decodificadores de la norma H.264/AVC y los pares de vistas estéreo podrían recibir soporte también mediante la MVC, la ventaja de la MVC es que podría dar soporte a un ejemplo que usa más de dos vistas como una entrada de vídeo tridimensional y decodifica este vídeo tridimensional representado por las múltiples vistas. Un representador de un cliente que tiene un decodificador de MVC puede esperar contenido de vídeo tridimensional con múltiples vistas.

[0127] Las imágenes en la figura 4 están indicadas en la intersección de cada fila y cada columna en la figura 4A, usando un bloque sombreado que incluye una letra, que designa si la imagen correspondiente está intracodificada (es

decir, es 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.

5 **[0128]** Al igual que con la codificación de vídeo de vista única, las imágenes de una secuencia de vídeo de codificación de vídeo de múltiples vistas pueden codificarse predictivamente con respecto a 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 de 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 en otras vistas como referencia. En la MVC, por ejemplo, la predicción entre vistas se realiza como si el componente de vista de otra vista fuera una referencia de interpredicción. Las referencias potenciales entre vistas se señalizan en la extensión de la MVC del conjunto de parámetros de secuencia (SPS) y pueden ser modificadas por el proceso de construcción de la lista de imágenes de referencia, que habilita el ordenamiento flexible de las referencias de interpredicción o de predicción entre vistas.

10 **[0129]** En la MVC, se permite la predicción entre vistas (IVP) entre imágenes en la misma unidad de acceso (es decir, con la misma instancia de tiempo). Una unidad de acceso es, en general, una unidad de datos que incluye todos los componentes de vista (por ejemplo, todas las unidades de NAL) para una instancia temporal común. Por lo tanto, en la MVC, se permite la predicción entre vistas entre imágenes en la misma unidad de acceso. Cuando se codifica una imagen en una de las vistas no básicas, la imagen puede agregarse a una lista de imágenes de referencia, si está en una vista diferente pero con la misma instancia de tiempo (por ejemplo, el mismo valor de POC y, por lo tanto, en la misma unidad de acceso). Una imagen de referencia de predicción entre vistas puede disponerse en cualquier posición de una lista de imágenes de referencia, tal como cualquier imagen de referencia de interpredicción.

15 **[0130]** La figura 4 proporciona varios ejemplos de predicción entre vistas. Las imágenes de la vista S1, en el ejemplo de la figura 4, se ilustran como predichas a partir de imágenes en diferentes ubicaciones temporales de la vista S1, así como predichas entre vistas a partir de imágenes de las 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.

20 **[0131]** En el ejemplo de la figura 4, 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 imágenes con "B" mayúscula son relativamente más altas en la jerarquía de predicción que las imágenes con "b" minúscula. La figura 4 también ilustra 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 figura 4 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.

25 **[0132]** En general, la jerarquía de predicción se relaciona con índices del orden de visualización, en cuanto a que las imágenes relativamente más altas en la jerarquía de predicción deberían ser decodificadas antes de la decodificación de imágenes que están relativamente más bajas en la jerarquía, de tal modo que aquellas imágenes relativamente más altas en la jerarquía se puedan utilizar como imágenes de referencia durante la decodificación de las imágenes relativamente más bajas en la jerarquía. Un índice de orden de vista es un índice que indica el orden de decodificación de componentes de vista en una unidad de acceso. Los índices de orden de vista están implícitos en la ampliación de MVC del SPS, como se especifica en el anexo H de la norma H.264/AVC (la enmienda MVC). En el SPS, para cada índice i, se señala el correspondiente identificador de vista. En algunos ejemplos, la decodificación de los componentes de vista seguirá el orden ascendente del índice de orden de vista. Si se presentan todas las vistas, entonces los índices de orden de vista están en un orden consecutivo desde 0 a `núm_vistas_menos_1`.

30 **[0133]** De esta manera, las imágenes utilizadas como imágenes de referencia pueden ser decodificadas antes de la decodificación de 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 decodificación de componentes de vista en una unidad de acceso. Para cada índice i de orden de vista, se señala el `id_vista` correspondiente. La decodificación 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 puede comprender un conjunto consecutivamente ordenado desde cero a uno menos que el número total de vistas.

35 **[0134]** Para ciertas imágenes a niveles iguales de la jerarquía, el orden de decodificación entre ellas puede no importar. Por ejemplo, la trama-I de la vista S0 en la ubicación temporal T0 se utiliza como imagen de referencia para la trama-P de la vista S2 en la ubicación temporal T0, que a su vez se utiliza como imagen de referencia para la trama-P de la vista S4 en la ubicación temporal T0. En consecuencia, la trama-I de la vista S0 en la ubicación temporal T0

debería decodificarse antes de la trama-P de la vista S2 en la ubicación temporal T0, la cual debería decodificarse antes de la trama-P de la vista S4 en la ubicación temporal T0. Sin embargo, entre las vistas S1 y S3, no importa un orden de decodificación, porque las vistas S1 y S3 no se basan la una en la otra para la predicción, sino que, en cambio, se predicen solo a partir de vistas que están más altas en la jerarquía de predicción. Además, la vista S1 puede decodificarse antes que la vista S4, siempre que la vista S1 sea decodificada después de las vistas S0 y S2.

[0135] De esta manera, puede utilizarse un ordenamiento jerárquico para describir las vistas S0 a S7. Indíquese con la notación SA > SB que la vista SA debería ser decodificada antes que la vista SB. Utilizando esta notación, S0 > S2 > S4 > S6 > S7, en el ejemplo de la figura 4. Además, con respecto al ejemplo de la figura 4, S0 > S1, S2 > S1, S2 > S3, S4 > S3, S4 > S5 y S6 > S5. Es posible cualquier orden de decodificación para las vistas que no viole estos requisitos. En consecuencia, son posibles muchas órdenes de decodificación diferentes.

[0136] De acuerdo a aspectos de esta divulgación, un codificador de vídeo (tal como el codificador de vídeo 20 o el decodificador de vídeo 30) puede determinar un punto de operación de salida que tenga una o más capas de salida de destino. Por ejemplo, el codificador de vídeo 20 puede determinar una o más capas de salida de destino a partir de la estructura de múltiples vistas mostrada en la figura 4 y codificar los datos que indican las capas de salida de destino en un flujo de bits codificado. Del mismo modo, el decodificador de vídeo 30 puede determinar un punto de operación de salida que incluye las una o más capas de salida de destino basándose en datos analizados sintácticamente y decodificados a partir de un flujo de bits codificado.

[0137] Por ejemplo, el codificador de vídeo 20 puede codificar una extensión de VPS (por ejemplo, como la que se muestra y describe con respecto al ejemplo de la Tabla 2 anterior), un VPS (por ejemplo, tal como el que se muestra y describe con respecto al ejemplo de la Tabla 3 anterior), un mensaje de SEI u otro mensaje para indicar uno o más puntos de operación de salida, teniendo cada punto de operación de salida capas de salida de destino asociadas. Como se ha indicado anteriormente, cuando se codifican vistas de profundidad, un punto de operación de salida puede referirse a un punto de operación con una lista de vistas identificadas de textura o profundidad, de salida de destino, que pertenecen a las vistas incluidas en el punto de operación.

[0138] El decodificador de vídeo 30 puede recibir los datos y determinar las vistas de salida de destino para su visualización. Como ejemplo, el decodificador de vídeo 30 puede recibir un punto de operación que incluye las vistas S0, S2, S4 y S6. El decodificador de vídeo 30 puede determinar, basándose en los elementos sintácticos incluidos en el flujo de bits, que las vistas de salida de destino para un punto de operación de salida estéreo incluyen las vistas S0 y S4. El decodificador de vídeo 30 puede decodificar cada una entre S0, S2, S4 y S6, pero solo puede generar las vistas S0 y S4.

[0139] La figura 5 es un diagrama conceptual que ilustra la codificación de vídeo ajustable a escala. Si bien la figura 5 se describe con respecto a las normas H.264 / AVC y SVC, debería entenderse que las capas similares se pueden codificar aplicando otros esquemas de codificación de vídeo de múltiples capas, incluyendo la H.266 / VVC. Por lo tanto, las referencias a la SVC a continuación pueden aplicarse a la codificación de vídeo ajustable a escala en general, y no están restringidas a la norma H.264 / SVC.

[0140] En SVC, las ajustabilidades a escala se pueden habilitar en tres dimensiones que incluyen, por ejemplo, la espacial, la temporal y la cualitativa (representadas como una velocidad de bits o razón entre señal y ruido (SNR)). En general, normalmente se puede lograr una mejor representación agregando a una representación en cualquier dimensión. Por ejemplo, en el ejemplo de la figura 5, la capa 0 se codifica en un formato intermedio cuarto común (QCIF) que tiene una velocidad de tramas de 7,5 Hz y una velocidad de bits de 64 kilooctetos por segundo (KBPS). Además, la capa 1 se codifica en QCIF con una velocidad de tramas de 15 Hz y una velocidad de bits de 64 KBPS, la capa 2 se codifica en CIF con una velocidad de tramas de 15 Hz y una velocidad de bits de 256 KBPS, la capa 3 se codifica en QCIF con una velocidad de tramas de 7,5 Hz y una velocidad de bits de 512 KBPS y la capa 4 se codifica en 4CIF con una velocidad de tramas de 30 Hz y una velocidad de bits de Megaocteto por segundo (MBPS). Debería entenderse que el número, el contenido y la disposición particulares de las capas mostradas en la figura 5 se proporcionan con fines de ejemplo solamente.

[0141] En cualquier caso, una vez que un codificador de vídeo (tal como el codificador de vídeo 20) ha codificado el contenido de una tal manera ajustable a escala, un decodificador de vídeo (tal como el decodificador de vídeo 30) puede usar una herramienta extractora para adaptar el contenido efectivo entregado de acuerdo a los requisitos de aplicación. que pueden depender, por ejemplo, del cliente o del canal de transmisión.

[0142] En la SVC, las imágenes que tienen la capa espacial y cualitativa más baja son habitualmente compatibles con la norma H.264 / AVC. En el ejemplo de la figura 5, las imágenes con la capa espacial y cualitativa más bajas (imágenes en la capa 0 y capa 1, con resolución de QCIF) pueden ser compatibles con la norma H.264/AVC. Entre ellas, aquellas imágenes del nivel temporal más bajo forman la capa base temporal (capa 0). Esta capa base temporal (capa 0) se puede mejorar con imágenes de niveles temporales más altos (capa 1).

[0143] Además de la capa compatible con la norma H.264/AVC, se pueden agregar varias capas de mejora espacial y/o cualitativas para proporcionar ajustabilidades a escala, espaciales y/o cualitativas. Cada capa de mejora espacial

o cualitativa en sí misma puede ser temporalmente ajustable a escala, con la misma estructura de ajustabilidad a escala temporal que la capa compatible con la norma H.264/AVC.

5 **[0144]** De acuerdo a aspectos de esta divulgación, un codificador de vídeo (tal como el codificador de vídeo 20 o el decodificador de vídeo 30) puede determinar un punto de operación de salida que tenga una o más capas de salida de destino. Por ejemplo, el codificador de vídeo 20 puede determinar una o más capas de salida de destino y codificar datos indicando las capas de salida de destino en un flujo de bits codificado. Del mismo modo, el decodificador de vídeo 30 puede determinar un punto de operación de salida que incluye una o más capas de salida de destino basándose en datos analizados sintácticamente y decodificados a partir de un flujo de bits codificado.

10 **[0145]** En algunos ejemplos, las capas de salida de destino pueden obtenerse en función de un id_capa de un punto de operación. Por ejemplo, el decodificador de vídeo 30 puede decodificar un punto de operación que incluye la capa 0, la capa 1 y la capa 2. En este ejemplo, el decodificador de vídeo 30 puede determinar un punto de operación de salida basándose en el máximo id_capa, por ejemplo, la capa 2. En consecuencia, el decodificador de vídeo 30 puede emitir imágenes de la capa 2 como imágenes de salida de destino.

15 **[0146]** La figura 6 es un diagrama de flujo que ilustra un proceso ejemplar para codificar un punto de operación de salida, de acuerdo a aspectos de esta divulgación. Aunque se ha descrito con respecto al codificador de vídeo 20 (figuras 1 y 2), debería entenderse que pueden configurarse otros dispositivos para realizar un procedimiento similar al de la figura 6.

20 **[0147]** En este ejemplo, el codificador de vídeo 20 puede codificar una pluralidad de capas de datos de vídeo para un conjunto de capas (100). Por ejemplo, el codificador de vídeo 20 puede codificar una pluralidad de vistas de datos de vídeo para un flujo de bits de múltiples vistas (incluyendo, por ejemplo, un flujo de bits que se ajusta a la norma MV-HEVC o a la norma 3D-HEVC). En este ejemplo, el conjunto de capas puede incluir una pluralidad de vistas de datos de vídeo necesarios para decodificar el flujo de bits, así como vistas de datos de vídeo destinados a la salida, por ejemplo, la visualización por un dispositivo cliente que decodifica el flujo de bits. En otro ejemplo, el codificador de vídeo 20 puede codificar una pluralidad de capas ajustables a escala de datos de vídeo para un flujo de bits ajustable a escala (incluido, por ejemplo, un flujo de bits que se ajusta a la norma HSVC). En este ejemplo, el conjunto de capas puede incluir una pluralidad de capas de datos de vídeo necesarias solo para decodificar el flujo de bits (por ejemplo, "capas no para emitir"), así como capas de datos de vídeo destinados a la salida.

25 **[0148]** El codificador de vídeo 20 puede determinar cuáles, entre la pluralidad de capas, son capas de salida de destino (102). Las capas de salida de destino pueden incluir las capas destinadas a la salida, que proporcionan una representación visible de los datos de vídeo cuando se muestran. Las capas de salida de destino pueden seleccionarse en función de la calidad deseada de una representación, las capacidades de los dispositivos de codificación de vídeo, las consideraciones de ancho de banda u otros factores.

30 **[0149]** El codificador de vídeo 20 puede codificar datos indicativos de un punto de operación de salida que incluye las capas determinadas de salida de destino (104). Por ejemplo, el codificador de vídeo 20 puede codificar uno o más elementos sintácticos que indican las capas de salida de destino, e incluir dichos elementos sintácticos en un conjunto de parámetros. Como se ha descrito anteriormente con respecto a la Tabla 2 y la Tabla 3, el codificador de vídeo 20 puede codificar los datos indicativos del punto de operación de salida en una extensión de VPS o un VPS. En otros ejemplos, el codificador de vídeo 20 puede codificar los datos indicativos del punto de operación de salida en un SPS u otro mensaje. El codificador de vídeo 20 puede luego emitir los datos codificados por entropía (por ejemplo, para su almacenamiento o transmisión a otro dispositivo) (106).

35 **[0150]** La figura 7 es un diagrama de flujo que ilustra un proceso ejemplar para decodificar un punto de operación de salida, de acuerdo a aspectos de esta divulgación. Aunque se ha descrito con respecto al decodificador de vídeo 30 (figuras 1 y 3), debería entenderse que pueden configurarse otros dispositivos para realizar un procedimiento similar al de la figura 7.

40 **[0151]** Inicialmente, el decodificador de vídeo 30 decodifica una pluralidad de capas de datos de vídeo para un conjunto de capas (120). Por ejemplo, el decodificador de vídeo 30 puede decodificar una pluralidad de vistas de datos de vídeo de un flujo de bits de múltiples vistas (incluido, por ejemplo, un flujo de bits que se ajusta a la norma MV-HEVC o a la norma 3D-HEVC). En este ejemplo, el conjunto de capas puede incluir una pluralidad de vistas de datos de vídeo, necesarias para decodificar el flujo de bits, así como vistas de datos de vídeo destinados a la salida. En otro ejemplo, el decodificador de vídeo 30 puede decodificar una pluralidad de capas escalables de datos de vídeo de un flujo de bits ajustable a escala (incluido, por ejemplo, un flujo de bits que se ajusta a la norma HSVC). En este ejemplo, el conjunto de capas puede incluir una pluralidad de capas de datos de vídeo necesarias para decodificar el flujo de bits, así como capas de datos de vídeo destinados a la salida.

45 **[0152]** El decodificador de vídeo 30 también puede decodificar datos indicativos de un punto de operación de salida que incluye una o más capas de salida de destino (122). Como se ha señalado anteriormente, las capas de salida de destino pueden incluir las capas destinadas a la salida, que proporcionan una representación visible de los datos de vídeo cuando se muestran. En algunos ejemplos, las capas de salida de destino pueden ser un subconjunto de la

pluralidad de capas que se han decodificado.

[0153] Los datos indicativos de las capas de salida de destino pueden incluir uno o más elementos sintácticos que indican las capas de salida de destino. Estos elementos sintácticos pueden incluirse en una extensión de VPS, un VPS, un SPS u otro tipo de mensaje, como se ha descrito anteriormente.

[0154] El decodificador de vídeo 30 puede extraer las capas de salida de destino desde el flujo de bits, en función de los datos (124). Por ejemplo, el decodificador de vídeo 30 puede extraer uno o más subflujos de bits, que contienen las capas de salida de destino, del flujo de bits. El decodificador de vídeo 30 también puede decodificar uno o más subflujos de bits. El decodificador de vídeo 30 puede luego emitir las capas de salida de destino para su visualización (por ejemplo, para su exhibición por el dispositivo de visualización 32 (figura 1)) (126).

[0155] En consecuencia, el decodificador de vídeo 30 obtiene una pluralidad de capas de datos de vídeo que incluyen una pluralidad de conjuntos de capas, donde cada conjunto de capas contiene una o más capas de datos de vídeo de la pluralidad de capas, y determina, basándose en uno o más elementos sintácticos del flujo de bits, uno o más puntos de operación de salida, donde cada punto de operación de salida está asociado a un conjunto de capas entre la pluralidad de conjuntos de capas y una o más capas de salida de destino entre la pluralidad de capas.

[0156] Ha de reconocerse que, según el ejemplo, ciertos actos o sucesos de cualquiera de las técnicas descritas en el presente documento pueden realizarse en una secuencia distinta, pueden añadirse, combinarse u omitirse por completo (por ejemplo, no todos los actos o sucesos descritos son necesarios para la puesta en práctica de las técnicas). Además, en determinados ejemplos, los actos o sucesos se pueden llevar a cabo simultáneamente, por ejemplo, mediante procesamiento de múltiples hilos, procesamiento de interrupciones o múltiples procesadores, en lugar de secuencialmente.

[0157] Ciertos aspectos de esta divulgación se han descrito con respecto a la norma de HEVC en desarrollo, con fines de ilustración. Sin embargo, las técnicas descritas en esta divulgación pueden ser útiles para otros procesos de codificación de vídeo, que incluyen otros procesos de codificación de vídeo, estándar o patentados, aún no desarrollados.

[0158] Un codificador de vídeo, como se describe en esta divulgación, puede referirse a un codificador de vídeo o a un decodificador de vídeo. De manera similar, una unidad de codificación de vídeo puede referirse a un codificador de vídeo o a un decodificador de vídeo. Del mismo modo, la codificación de vídeo puede referirse a la codificación de vídeo o a la decodificación de vídeo, según corresponda.

[0159] En uno o más ejemplos, las funciones descritas pueden implementarse en hardware, software, firmware o cualquier combinación de los mismos. Si se implementan en software, las funciones pueden almacenarse en, y transmitirse por, un medio legible por ordenador, como una o más instrucciones o código, y ejecutarse mediante una unidad de procesamiento basada en hardware. Los medios legibles por ordenador pueden incluir medios de almacenamiento legibles por ordenador, que correspondan a un medio tangible tal como medios de almacenamiento de datos o medios de comunicación que incluyan cualquier medio que facilite la transferencia de un programa informático de un lugar a otro, por ejemplo, de acuerdo a un protocolo de comunicación. De esta manera, los medios legibles por ordenador pueden corresponder en general a (1) medios de almacenamiento tangibles legibles por ordenador que sean no transitorios o (2) un medio de comunicación tal como una señal o una onda portadora. Los medios de almacenamiento de datos pueden ser medios disponibles cualesquiera a los que se pueda acceder desde uno o más ordenadores o uno o más procesadores para recuperar instrucciones, código y/o estructuras de datos para la implementación de las técnicas descritas en esta divulgación. Un producto de programa informático puede incluir un medio legible por ordenador.

[0160] A modo de ejemplo, y no de manera limitativa, tales 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 pueda usarse para almacenar código de programa deseado en forma de instrucciones o estructuras de datos y al que pueda accederse mediante un ordenador. Asimismo, cualquier conexión recibe debidamente la denominación de medio legible por ordenador. Por ejemplo, si las instrucciones se transmiten desde una sede de la Red, un servidor u otro origen remoto, usando un cable coaxial, un cable de fibra óptica, un par trenzado, una línea de abonado digital (DSL) o tecnologías inalámbricas tales como infrarrojos, radio y microondas, entonces el cable coaxial, el cable de fibra óptica, el par trenzado, la DSL o las tecnologías inalámbricas tales como infrarrojos, radio y microondas se incluyen en la definición de medio. Sin embargo, debería entenderse que los medios de almacenamiento legibles por ordenador y los medios de almacenamiento de datos no incluyen conexiones, ondas portadoras, señales u otros medios transitorios, sino que, en cambio, están orientados a medios de almacenamiento tangibles no transitorios. El término disco, como se usa en el presente documento, incluye el disco compacto (CD), el disco láser, el disco óptico, el disco versátil digital (DVD), el disco flexible y el disco Blu-ray, donde algunos discos normalmente emiten datos magnéticamente, mientras que otros discos emiten datos ópticamente con láseres. Las combinaciones de lo anterior también deberían incluirse dentro del alcance de los medios legibles por ordenador.

5 **[0161]** Las instrucciones pueden ser ejecutadas 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), formaciones de compuertas programables in situ (FPGA) u otros circuitos lógicos, integrados o discretos, equivalentes. En consecuencia, el término "procesador", como se usa en el presente documento, se puede referir a cualquiera de las estructuras anteriores o a cualquier otra estructura adecuada para la implementación de las técnicas descritas en el presente documento. Además, en algunos aspectos, la funcionalidad descrita en el presente documento se puede proporcionar dentro de módulos de hardware y/o software dedicados, configurados para la codificación y la decodificación, o incorporados en un códec combinado. Además, las técnicas se podrían implementar totalmente en uno o más circuitos o elementos lógicos.

10 **[0162]** Las técnicas de la presente divulgación se pueden implementar en una amplia variedad de dispositivos o aparatos, incluidos un equipo manual inalámbrico, un circuito integrado (CI) o un conjunto de CI (por ejemplo, un conjunto de chips). Diversos componentes, módulos o unidades se describen en esta divulgación para enfatizar aspectos funcionales de dispositivos configurados para realizar las técnicas divulgadas, pero no requieren necesariamente su realización mediante diferentes unidades de hardware. En cambio, como se ha descrito anteriormente, diversas unidades se pueden combinar en una unidad de hardware de códec, o ser proporcionadas por un grupo de unidades de hardware interoperativas, incluyendo uno o más procesadores, como se ha descrito anteriormente, conjuntamente con software y/o firmware adecuados.

20 **[0163]** Se han descrito diversos ejemplos. Estos y otros ejemplos están dentro del alcance de las siguientes reivindicaciones.

REIVINDICACIONES

1. Un procedimiento de decodificación de datos de vídeo, comprendiendo el procedimiento:

5 obtener, a partir de un flujo de bits de múltiples capas que contiene datos de vídeo de múltiples capas, una pluralidad de capas de datos de vídeo que incluyen una pluralidad de conjuntos de capas, en donde cada conjunto de capas contiene una o más capas entre la pluralidad de capas;

10 para cada conjunto de capas entre la pluralidad de conjuntos de capas, decodificar, desde el flujo de bits de múltiples capas, un indicador de capa incluida para cada capa entre la pluralidad de capas, en donde el indicador de capa incluida identifica si la capa respectiva se incluye o no en el respectivo conjunto de capas;

15 decodificar, a partir del flujo de bits de múltiples capas, un primer elemento sintáctico que identifica un número máximo de conjuntos de capas de salida, presentes en el flujo de bits de múltiples capas, en donde un conjunto de capas de salida identifica una o más capas de salida de destino, y una capa de salida de destino está concebida para ser emitida para su visualización; y

20 para cada conjunto de capas de salida del número máximo de conjuntos de capas de salida, decodificar, a partir del flujo de bits de múltiples capas, un segundo elemento sintáctico que identifica un conjunto de capas basándose en el cual se forma el respectivo conjunto de capas de salida y, para cada capa incluida en el conjunto de capas identificado, decodificar, a partir del flujo de bits de múltiples capas, un indicador de capa de salida, que identifica si la capa respectiva es o no una capa de salida de destino del respectivo conjunto de capas de salida, para determinar de ese modo las capas de salida de destino que pertenecen a cada conjunto de capas de salida.

2. Un procedimiento para codificar datos de vídeo, comprendiendo el procedimiento:

30 codificar, como un flujo de bits de múltiples capas que contiene datos de vídeo de múltiples capas, una pluralidad de capas de datos de vídeo que incluyen una pluralidad de conjuntos de capas, en donde cada conjunto de capas contiene una o más capas de datos de vídeo entre la pluralidad de capas;

35 para cada conjunto de capas entre la pluralidad de conjuntos de capas, codificar, a partir del flujo de bits de múltiples capas, un indicador de capa incluida para cada capa entre la pluralidad de capas, en donde el indicador de capa incluida identifica si la capa respectiva se incluye o no en el respectivo conjunto de capas;

40 codificar, en el flujo de bits de múltiples capas, un primer elemento sintáctico que identifica un número máximo de conjuntos de capas de salida, presentes en el flujo de bits de múltiples capas, en donde un conjunto de capas de salida identifica una o más capas de salida de destino, y una capa de salida de destino está concebida para ser emitida para su visualización; y

45 para cada conjunto de capas de salida del número máximo de conjuntos de capas de salida, codificar, en el flujo de bits de múltiples capas, un segundo elemento sintáctico que identifica un conjunto de capas en función del cual se forma el respectivo conjunto de capas de salida, y, para cada capa incluida en el conjunto de capas identificado, codificar un indicador de capa de salida, que identifica si la capa respectiva es o no una capa de salida de destino del respectivo conjunto de capas de salida.

3. El procedimiento de la reivindicación 1 o 2, en el que, para cada indicador de capa de salida, cuando el indicador de capa de salida es igual a 1, la capa respectiva es una capa de salida de destino del respectivo conjunto de capas o, cuando el indicador de capa de salida es igual a 0, la capa respectiva no es una capa de salida de destino del respectivo conjunto de capas.

55 4. El procedimiento de la reivindicación 1 o la reivindicación 2, en el que el flujo de bits de múltiples capas comprende un flujo de bits de múltiples vistas, de modo que cada capa comprenda una vista de datos de vídeo y las una o más capas de salida de destino comprendan una o más vistas de salida de destino, o en el que el flujo de bits de múltiples capas comprende una o más capas de profundidad, de modo que la determinación de las capas de salida de destino que pertenecen a cada conjunto de capas de salida comprenda además la determinación de una o más capas de profundidad de salida de destino que comprenden datos de profundidad.

60 5. El procedimiento de la reivindicación 1, en el que el flujo de bits de múltiples capas comprende un flujo de bits que tiene una pluralidad de capas ajustables a escala de datos de vídeo, de modo que determinar las capas de salida de destino que pertenecen a cada conjunto de capas de salida comprenda determinar una o más capas de salida de destino ajustables a escala.

65 6. El procedimiento de la reivindicación 5, en el que determinar el punto de operación de salida comprende determinar la capa de salida de destino ajustable a escala que tiene un id_capa numéricamente máximo.

7. El procedimiento de la reivindicación 1, que comprende además:
- 5 decodificar un conjunto de capas que se han de decodificar, pero no emitir, incluidas una o más capas que no son una de las una o más capas de salida de destino, sino que pertenecen al conjunto de capas;
- 10 marcar una primera imagen decodificada, incluida en las una o más capas de salida de destino, en un almacén temporal de imágenes decodificadas, como una imagen a utilizar para su salida; y
- 15 marcar una segunda imagen decodificada, incluida en el conjunto de capas que se decodifican pero que no se emiten, en el almacén temporal de imágenes decodificadas, como una imagen que no se ha de usar para su salida, de modo que la imagen que no se usará para su salida se elimine del almacén temporal de imágenes decodificadas antes que la imagen que se ha de utilizar para su salida.
8. El procedimiento de la reivindicación 7, que comprende además:
- 20 determinar si la imagen marcada como no a utilizar para su salida se utiliza para una entre la interpretación y la predicción entre capas; y
- 25 eliminar la imagen, marcada como no a utilizar para su salida, del almacén temporal de imágenes decodificadas cuando la imagen no a utilizar para su salida no se usa para una entre la interpretación y la predicción entre capas.
9. El procedimiento de la reivindicación 1 o la reivindicación 2, en el que los uno o más elementos sintácticos se incluyen en uno entre un conjunto de parámetros de vídeo (VPS) y una extensión de VPS del flujo de bits codificado, o en el que los uno o más elementos sintácticos se incluyen en un mensaje de información de mejora suplementaria (SEI) del flujo de bits codificado.
10. El procedimiento de la reivindicación 2, en el que el flujo de bits de múltiples capas comprende un flujo de bits de múltiples vistas, de manera que cada capa comprenda una vista de datos de vídeo y las una o más capas de salida de destino comprendan una o más vistas de salida de destino, o en el que el flujo de bits de múltiples capas comprende una o más capas de profundidad, de modo que los uno o más conjuntos de capas de salida comprendan una o más capas de profundidad de salida de destino que comprenden datos de profundidad.
11. El procedimiento de la reivindicación 2, en el que el flujo de bits de múltiples capas comprende un flujo de bits que tiene una pluralidad de capas ajustables a escala de datos de vídeo, de modo que determinar el conjunto de capas de salida comprenda determinar una o más capas de salida de destino ajustables a escala, en donde el conjunto de capas de salida comprende, preferiblemente, la capa de salida de destino ajustable a escala que tenga un id_capa numéricamente máximo.
12. Un dispositivo para decodificar datos de vídeo, comprendiendo el dispositivo:
- 45 medios para obtener, a partir de un flujo de bits de múltiples capas que contiene datos de vídeo de múltiples capas, una pluralidad de capas de datos de vídeo que incluyen una pluralidad de conjuntos de capas, en donde cada conjunto de capas contiene una o más capas entre la pluralidad de capas;
- 50 medios para decodificar, a partir del flujo de bits de múltiples capas, para cada conjunto de capas entre la pluralidad de conjuntos de capas, un indicador de capa incluida para cada capa entre la pluralidad de capas, en donde el indicador de capa incluida identifica si la capa respectiva está incluida o no en el respectivo conjunto de capas;
- 55 medios para decodificar, a partir del flujo de bits de múltiples capas, un primer elemento sintáctico que identifica un número máximo de conjuntos de capas de salida, presentes en el flujo de bits de múltiples capas, en donde un conjunto de capas de salida identifica una o más capas de salida de destino, y una capa de salida de destino está concebida para ser emitida para su visualización; y
- 60 medios para decodificar, a partir del flujo de bits de múltiples capas, para cada conjunto de capas de salida del número máximo de conjuntos de capas de salida, un segundo elemento sintáctico que identifica un conjunto de capas en base al cual se forma el respectivo conjunto de capas de salida y, para cada capa incluida en el conjunto de capas identificado, decodificar, a partir del flujo de bits de múltiples capas, un indicador de capa de salida, que identifica si la capa respectiva es o no una capa de salida de destino del respectivo punto de operación de salida, para determinar así las capas de salida de destino que pertenecen a cada conjunto de capas de salida.
13. Un dispositivo para codificar datos de vídeo, comprendiendo el dispositivo:
- 65

medios para codificar, como un flujo de bits de múltiples capas que contiene datos de vídeo de múltiples capas, una pluralidad de capas de datos de vídeo que incluyen una pluralidad de conjuntos de capas, en donde cada conjunto de capas contiene una o más capas entre la pluralidad de capas;

5 medios para codificar a partir del flujo de bits de múltiples capas, para cada conjunto de capas entre la pluralidad de conjuntos de capas, un indicador de capa incluida para cada capa entre la pluralidad de capas, en donde el indicador de capa incluida identifica si la capa respectiva está incluida o no en el respectivo conjunto de capas;

10 medios para codificar, en el flujo de bits de múltiples capas, un primer elemento sintáctico que identifica un número máximo de conjuntos de capas de salida, presentes en el flujo de bits de múltiples capas, en donde un conjunto de capas de salida identifica una o más capas de salida de destino, y una capa de salida de destino está concebida para ser emitida para su visualización; y

15 medios para codificar, en el flujo de bits de múltiples capas, para cada conjunto de capas de salida del número máximo de conjuntos de capas de salida, un segundo elemento sintáctico que identifica un conjunto de capas en función del cual se forma el respectivo conjunto de capas de salida y, para cada capa incluida en el conjunto de capas identificado, codificar un indicador de capa de salida, que identifica si la capa respectiva es o no una capa de salida de destino del respectivo conjunto de capas de salida.

20 14. Un medio de almacenamiento no transitorio legible por ordenador que tiene almacenadas en el mismo instrucciones que, al ejecutarse, hacen que un procesador de un dispositivo para codificar datos de vídeo realice el procedimiento de cualquiera de las reivindicaciones 1 a 11.

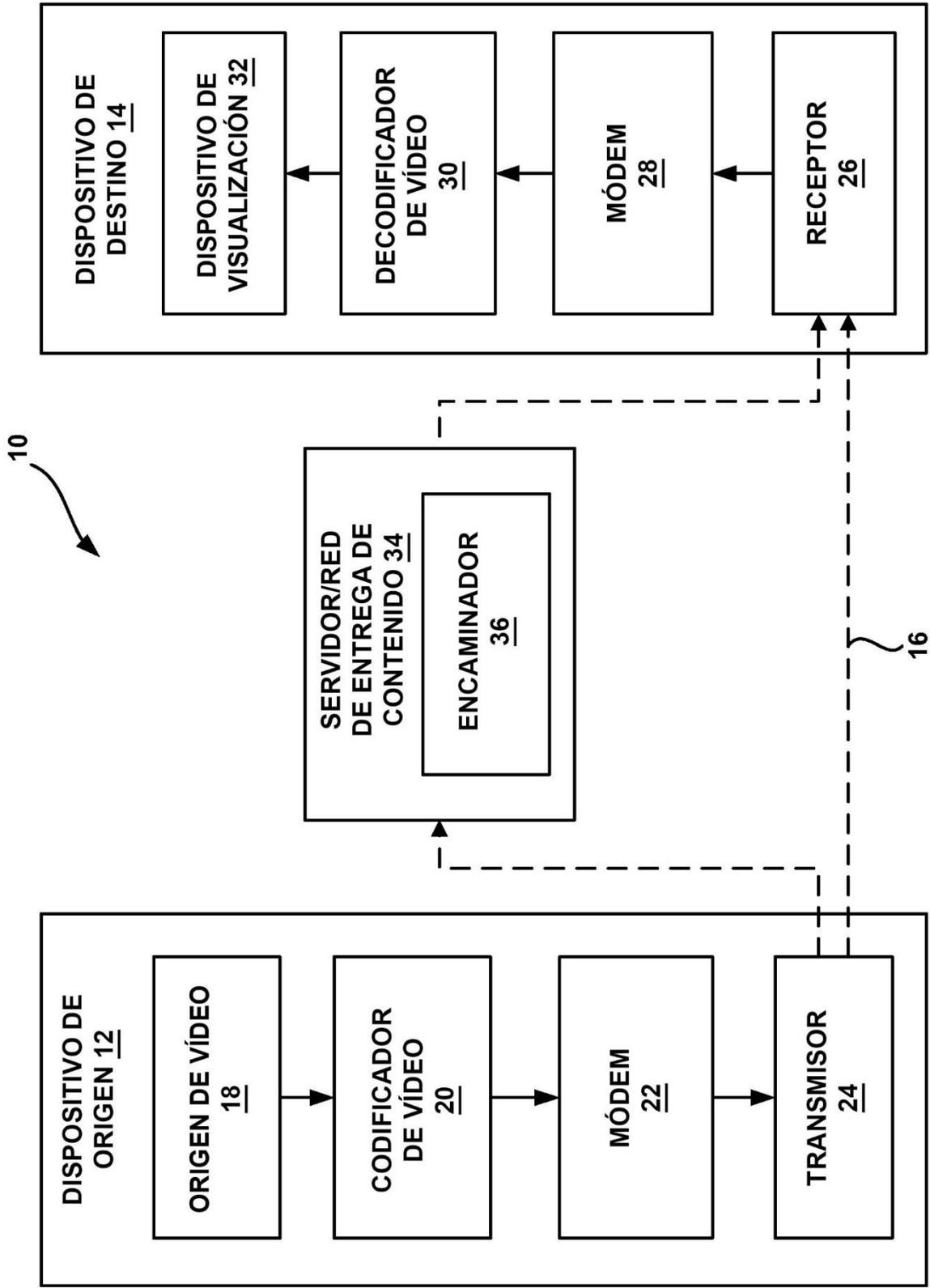


FIG. 1

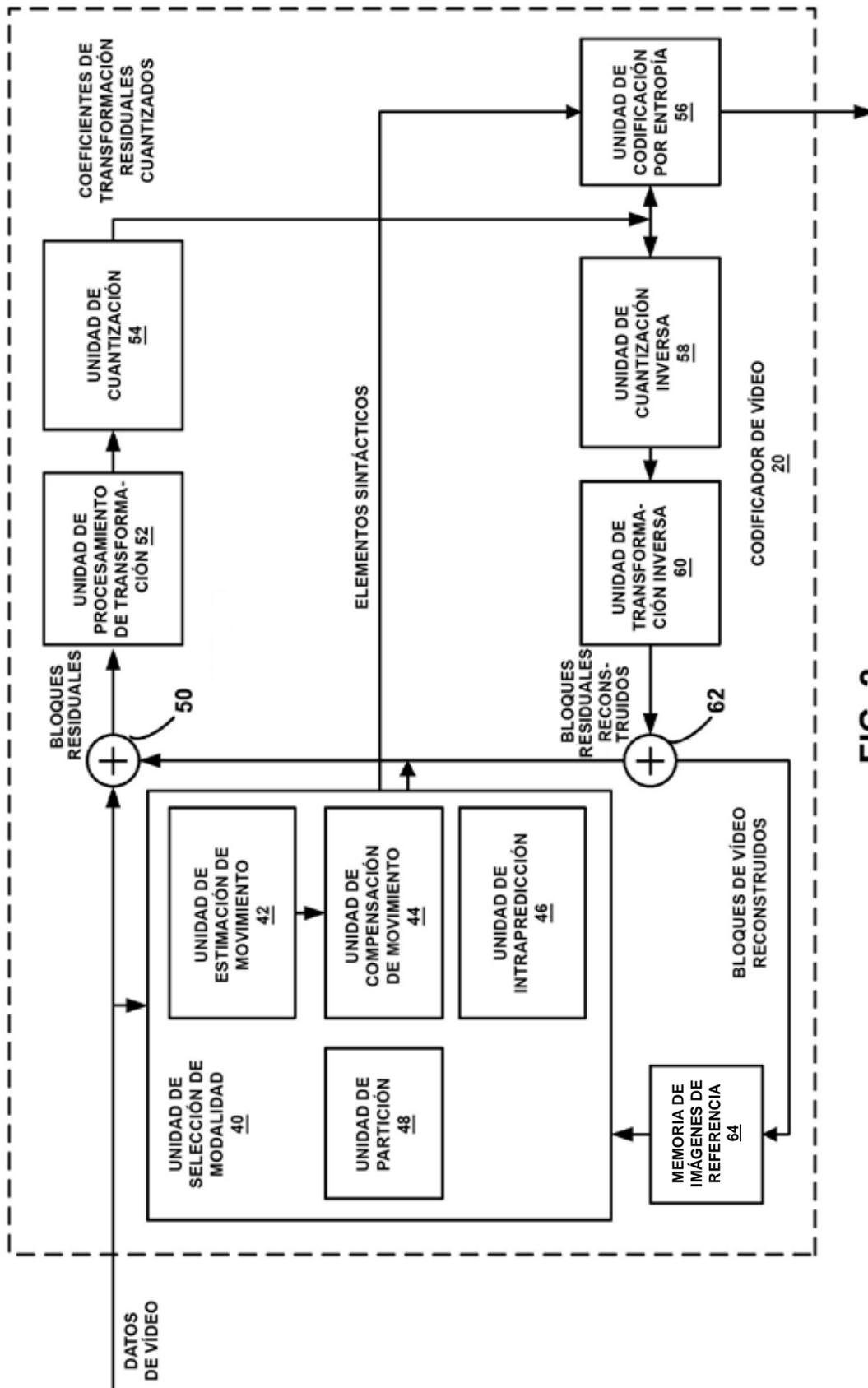


FIG. 2

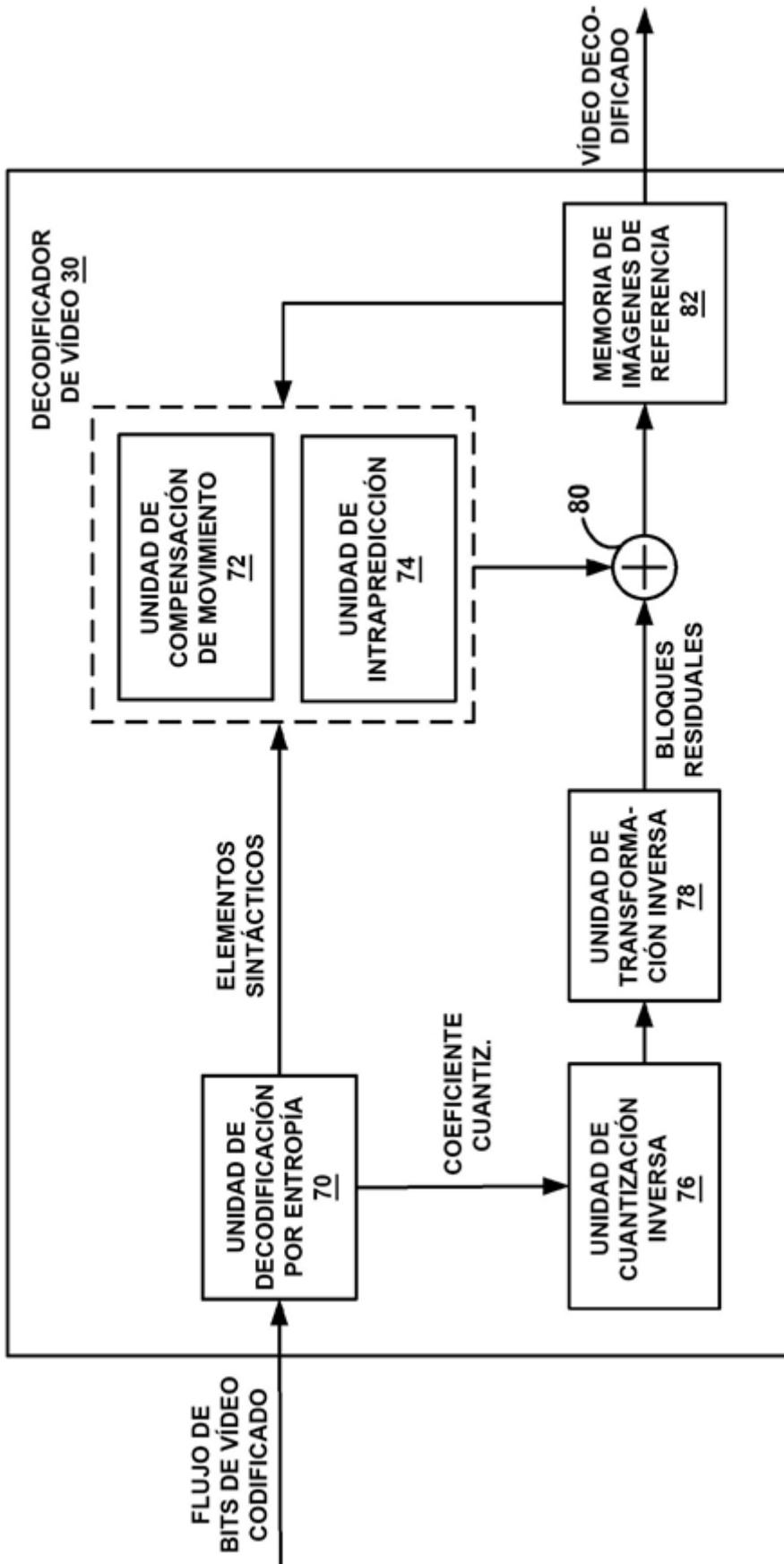


FIG. 3

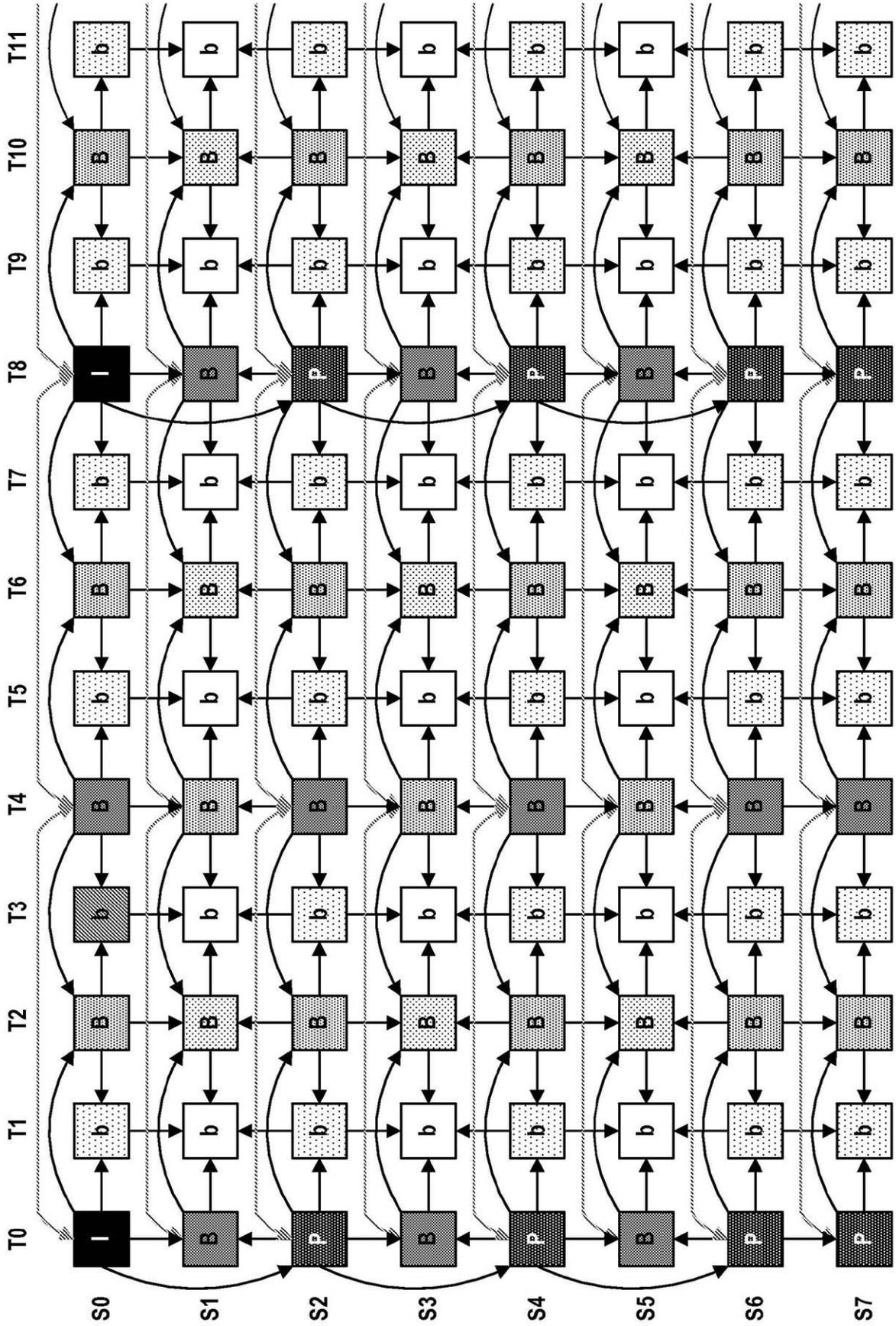


FIG. 4

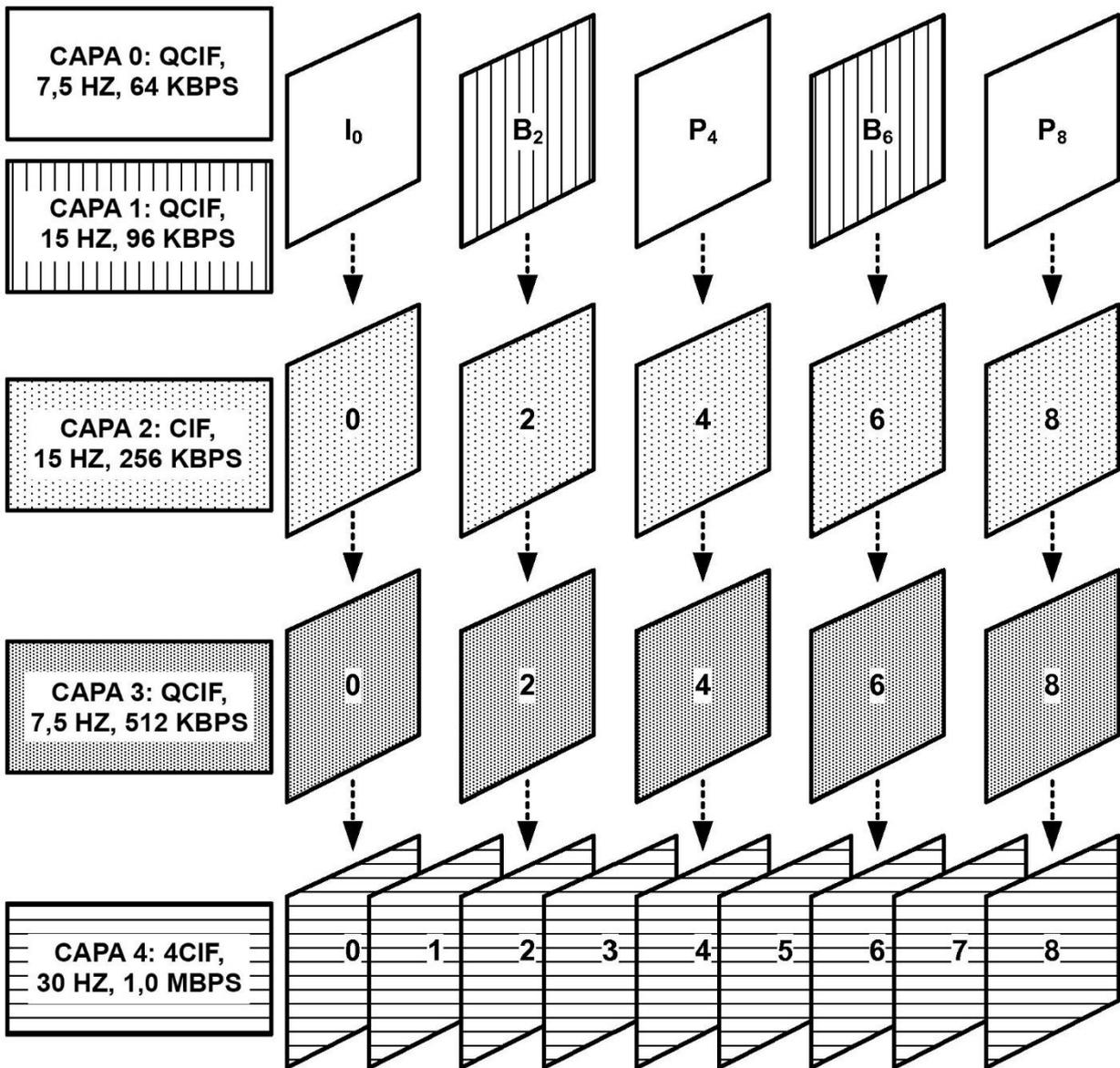


FIG. 5

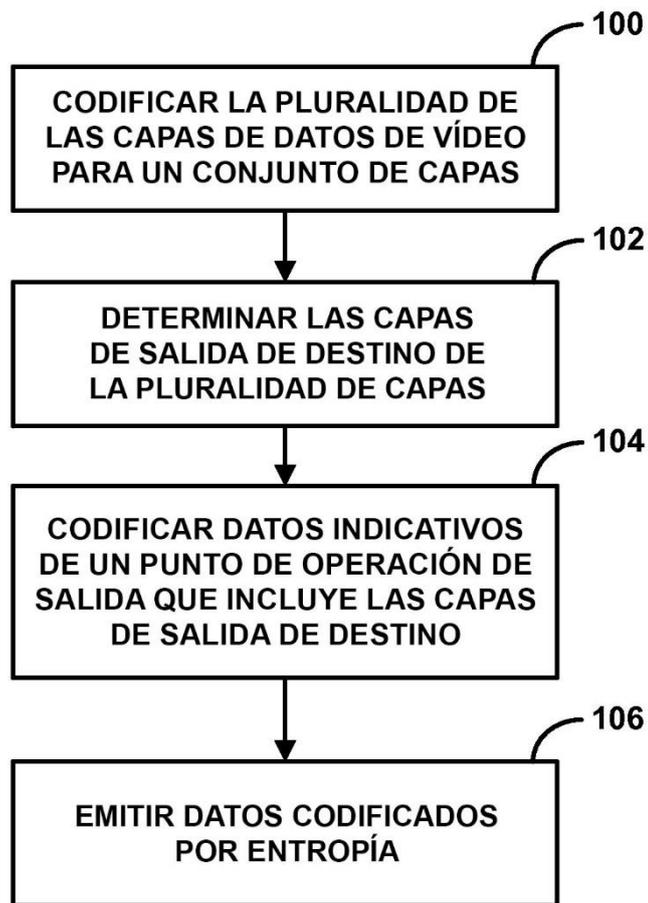


FIG. 6

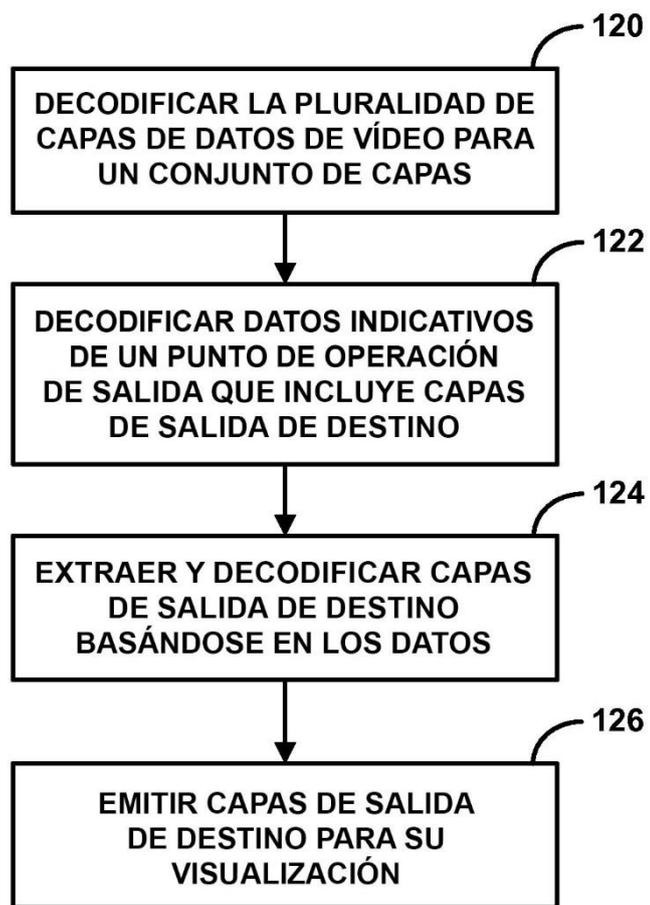


FIG. 7