

19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 703 342**

51 Int. Cl.:

H04N 19/105 (2014.01)

H04N 19/164 (2014.01)

H04N 19/172 (2014.01)

H04N 19/573 (2014.01)

H04N 19/58 (2014.01)

H04N 19/577 (2014.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

86 Fecha de presentación y número de la solicitud internacional: **02.12.2014 PCT/US2014/068031**

87 Fecha y número de publicación internacional: **11.06.2015 WO15084777**

96 Fecha de presentación y número de la solicitud europea: **02.12.2014 E 14816531 (9)**

97 Fecha y número de publicación de la concesión europea: **03.10.2018 EP 3078193**

54 Título: **Selección de imagen de referencia**

30 Prioridad:

02.12.2013 US 201361910766 P

01.12.2014 US 201414557049

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

08.03.2019

73 Titular/es:

**QUALCOMM INCORPORATED (100.0%)
International IP Administration
5775 Morehouse Drive
San Diego, California 92121-1714, US**

72 Inventor/es:

**TAO, SIPING;
COBAN, MUHAMMED ZEYD y
KARCZEWICZ, MARTA**

74 Agente/Representante:

FORTEA LAGUNA, Juan José

ES 2 703 342 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

Selección de imagen de referencia

5 **CAMPO TÉCNICO**

[0001] Esta divulgación se refiere a la codificación de vídeo (es decir, codificación de vídeo o descodificación de vídeo).

10 **ANTECEDENTES**

[0002] Las capacidades de vídeo digital pueden incorporarse en una amplia gama de dispositivos, incluyendo televisores digitales, sistemas de difusión directa digital, sistemas de difusión inalámbrica, asistentes digitales personales (PDA), ordenadores portátiles o de escritorio, ordenadores tipo tablet, 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 continua de vídeo y similares. Los dispositivos de vídeo digitales implementan técnicas de compresión de vídeo, tales como las descritas en las normas definidas por MPEG-2, MPEG-4, ITU-T H.263, ITU-T H.264/MPEG-4, parte 10, codificación de vídeo avanzada (AVC), la norma de codificación de vídeo de alta eficiencia (HEVC) actualmente en desarrollo y las ampliaciones de dichas normas. Los dispositivos de vídeo pueden transmitir, recibir, codificar, descodificar y/o almacenar información de vídeo digital con más eficiencia implementando dichas técnicas de compresión de vídeo.

[0003] Las técnicas de compresión de vídeo realizan 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 de árbol, unidades de codificación (CU) y/o nodos de codificación. Los bloques de vídeo en un fragmento con codificación intra (I) de una imagen se codifican utilizando predicción espacial con respecto a muestras de referencia en bloques contiguos en la misma imagen. Los bloques de vídeo en un fragmento con codificación inter (P o B) de una imagen pueden usar la predicción espacial con respecto a muestras de referencia en bloques contiguos en la misma imagen o la predicción temporal con respecto a muestras de referencia en otras imágenes de referencia. Las imágenes pueden denominarse tramas y las imágenes de referencia pueden denominarse tramas de referencia.

[0004] La predicción espacial o temporal da como resultado un bloque predictivo para un bloque que se va a codificar. Los datos residuales representan diferencias de píxeles entre el bloque original que se va a codificar y el bloque predictivo. Un bloque con codificación inter se codifica de acuerdo con un vector de movimiento que apunta a un bloque de muestras de referencia que forman el bloque predictivo, y los datos residuales que indican la diferencia entre el bloque codificado y el bloque predictivo. Un bloque con codificación intra se codifica de acuerdo con un modo de codificación intra y los datos residuales. Para una mayor compresión, los datos residuales pueden transformarse desde el dominio del píxel hasta un dominio de transformada, dando como resultado unos coeficientes de transformada residuales, que a continuación se pueden cuantificar. Los coeficientes de transformada cuantificados, dispuestos inicialmente en una matriz bidimensional, pueden examinarse con el fin de generar un vector unidimensional de coeficientes de transformada, y puede aplicarse codificación por entropía para lograr aún más compresión.

[0005] El documento WO 2013/126284 divulga cómo seleccionar a partir de un conjunto de imágenes de referencia confirmadas aquellas que se pueden esperar que estén dentro de una memoria intermedia de imágenes de referencia en el descodificador para codificar una imagen actual.

[0006] La publicación de la conferencia "Selección de imagen de referencia en ventana para recuperación de error de transmisión H.264" (26. Simposio sobre codificación, imágenes, Lisboa, 7-11-2007) divulga el concepto de mantener siempre, en el codificador en la memoria intermedia de imágenes de referencia, una imagen que ha sido confirmada por el descodificador.

[0007] La publicación de la conferencia "Error en la codificación de vídeo resistente usando tramas de referencia flexibles" (Comunicaciones visuales y procesamiento de imágenes, Pekín, 12-07-2005) divulga conceptos para usar imágenes de referencia que fueron confirmados por un descodificador para predecir una imagen actual.

60 **SUMARIO**

[0008] Esta divulgación describe técnicas de ejemplo para determinar imágenes a partir de las cuales se seleccionan una o más imágenes de referencia para codificar y descodificar una imagen actual basándose en la información que emite un descodificador de vídeo. Por ejemplo, un descodificador de vídeo puede emitir información que identifica una imagen previamente descodificada almacenada en una memoria intermedia del descodificador de vídeo a un codificador de vídeo. El codificador de vídeo recibe la información de salida del descodificador de vídeo y puede determinar las imágenes que están disponibles para la codificación basándose en: la imagen identificada, las imágenes

que se determinaron como imágenes que podrían haber sido imágenes de referencia para la imagen identificada y/o las imágenes siguientes a la imagen identificada en orden de codificación. Luego, el codificador de vídeo puede utilizar una o más de las imágenes que se determinó que están disponibles para la codificación con el fin de realizar predicción inter de una imagen actual.

5
 [0009] Por ejemplo, no todas las imágenes que se determinaron como imágenes que podrían haber sido imágenes de referencia para la imagen identificada y las imágenes que siguen a la imagen identificada en el orden de codificación pueden estar disponibles para la codificación. Sin embargo, existe una certeza relativamente alta de que si las imágenes, que se determinaron como imágenes que podrían haber sido imágenes de referencia para la imagen identificada e imágenes que siguen a la imagen identificada en orden de codificación, todavía están disponibles para codificar una imagen actual mediante el codificador de vídeo, entonces esas imágenes también estarán disponibles para descodificar la imagen actual mediante el descodificador de vídeo. De esta manera, el codificador de vídeo puede ser capaz de seleccionar a partir de una pluralidad de imágenes, en lugar de solo la imagen identificada, para codificar una imagen actual con una certeza relativamente alta de que las imágenes seleccionadas estarán disponibles en el descodificador de vídeo para descodificar la imagen actual.

10
 [0010] En un ejemplo, la divulgación describe un procedimiento para codificar datos de vídeo, con el procedimiento que comprende recibir, de un dispositivo que comprende un descodificador de vídeo, información que identifica una imagen previamente descodificada que se almacena en una memoria intermedia de imágenes descodificadas (DPB) del descodificador de vídeo en el momento en que el descodificador de vídeo transmitió la información, determinar una o más imágenes de referencia candidatas que están disponibles para codificar una imagen actual basándose en la imagen identificada, en el que la una o más imágenes de referencia candidatas incluyen una o más imágenes que previamente se determinó que son utilizables para la codificación mediante predicción inter de una o más de las imágenes o imágenes identificadas que siguen a la imagen identificada en orden de codificación, y que aún se almacenan en una DPB de un codificador de vídeo, seleccionar una o más imágenes de referencia para la imagen actual a partir de la una o más imágenes de referencia candidatas determinadas y la imagen identificada, y codificar mediante predicción inter de la imagen actual basándose en la una o más imágenes de referencia seleccionadas.

20
 [0011] En un ejemplo, la divulgación describe un procedimiento para descodificar datos de vídeo, con el procedimiento que comprende enviar, a un dispositivo que comprende un codificador de vídeo, información que identifica una imagen previamente descodificada almacenada en una memoria intermedia de imágenes descodificadas (DPB) de un descodificador de vídeo en el momento de emitir la información y codificar mediante predicción inter una imagen actual con una o más imágenes de referencia candidatas, en el que la imagen o imágenes de referencia candidatas incluyen una o más imágenes que el codificador de vídeo determinó previamente para que se puedan utilizar para codificar mediante predicción inter de la imagen identificada.

30
 [0012] En un ejemplo, la divulgación describe un dispositivo para codificar datos de vídeo, comprendiendo el dispositivo una primera memoria intermedia de imágenes descodificadas (DPB) que comprende una memoria configurada para almacenar una o más imágenes, y un codificador de vídeo que comprende uno o más procesadores. El codificador de vídeo está configurado para recibir, desde un dispositivo que comprende un descodificador de vídeo, información que identifica una imagen previamente descodificada que se almacena en una segunda DPB del descodificador de vídeo en el momento en que el descodificador de vídeo transmitió la información, determinar una o más imágenes de referencia candidatas que están disponibles para codificar una imagen actual basándose en la imagen identificada, en el que la una o más imágenes de referencia candidatas incluyen una o más imágenes que previamente se determinó que se pueden usar para la codificación mediante predicción inter de una o más de las imágenes o imágenes identificadas después de imagen identificada en el orden de codificación, y que aún se almacenan en la primera DPB, seleccionar una o más imágenes de referencia para la imagen actual a partir de la una o más imágenes de referencia candidatas determinadas y la imagen identificada, y codificar mediante predicción inter de la imagen actual basándose en la una o más imágenes de referencia seleccionadas.

40
 [0013] En un ejemplo, la divulgación describe un dispositivo para descodificar datos de vídeo, comprendiendo el dispositivo una memoria intermedia de imágenes descodificadas (DPB) que comprende una memoria configurada para almacenar una o más imágenes, y un descodificador de vídeo que comprende uno o más procesadores. El descodificador de vídeo está configurado para emitir, a un dispositivo que comprende un codificador de vídeo, información que identifica una imagen previamente descodificada almacenada en la DPB del descodificador de vídeo en el momento de emitir la información, y descodificar mediante predicción inter una imagen actual con una o más imágenes de referencia candidatas, en el que una o más imágenes de referencia candidatas incluyen una o más imágenes que el codificador de vídeo determinó previamente para que se puedan utilizar para la codificación mediante predicción inter de la imagen identificada.

50
 [0014] Los detalles de uno o más ejemplos de la divulgación se exponen en los dibujos adjuntos y la descripción siguiente. Otras características, objetivos y ventajas resultarán evidentes a partir de la descripción, de los dibujos y de las reivindicaciones.

60
 65

BREVE DESCRIPCIÓN DE LOS DIBUJOS

[0015]

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

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

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

La FIG. 4 es un diagrama de flujo que ilustra un procedimiento de ejemplo para codificar datos de vídeo.

La FIG. 5 es un diagrama de flujo que ilustra un procedimiento de ejemplo para descodificar datos de vídeo.

La FIG. 6 es un diagrama de flujo que ilustra otro procedimiento de ejemplo para codificar datos de vídeo.

La FIG. 7 es un diagrama de flujo que ilustra otro procedimiento de ejemplo para descodificar datos de vídeo.

DESCRIPCIÓN DETALLADA

[0016] Las técnicas descritas en esta divulgación están en general relacionadas con la codificación de vídeo, incluidos ejemplos de aplicaciones de vídeo en tiempo real. Por ejemplo, las técnicas descritas en esta divulgación pueden ser aplicables a la videotelefonía o videoconferencia, donde el vídeo se codifica y descodifica, en tiempo real, a cada lado de la videotelefonía o conferencia. Como otro ejemplo, las técnicas descritas en esta divulgación pueden ser aplicables a la transmisión continua, donde el vídeo del evento en vivo se codifica en tiempo real para su transmisión. Sin embargo, las técnicas descritas en esta divulgación no deben considerarse limitadas a la videotelefonía, la videoconferencia o la codificación de vídeo en tiempo real de un evento en vivo. Las técnicas descritas en esta divulgación pueden ser aplicables a la codificación de vídeo donde un descodificador de vídeo es capaz de transmitir información a un codificador de vídeo que codifica los datos de vídeo que se transmiten al descodificador de vídeo.

[0017] En la codificación de vídeo en tiempo real, un descodificador de vídeo puede configurarse para comunicarse con un codificador de vídeo (por ejemplo, usando un protocolo de transporte en tiempo real (RTP)) para que el codificador de vídeo determine las imágenes disponibles para la codificación desde la cual el codificador de vídeo selecciona una o más imágenes de referencia para la codificación mediante predicción inter de una imagen actual. Puede haber varias razones por las que el descodificador de vídeo puede comunicarse con el codificador de vídeo para determinar las imágenes disponibles para la codificación, y las técnicas descritas en esta divulgación no están limitadas a ninguna razón en particular por la cual el descodificador de vídeo puede comunicarse con el codificador de vídeo para determinar las imágenes que están disponibles para la codificación.

[0018] En algunos ejemplos, el descodificador de vídeo puede comunicarse con el codificador de vídeo para que el codificador de vídeo determine las imágenes de referencia candidatas (es decir, las imágenes que están disponibles para la codificación), de modo que el codificador de vídeo seleccione las imágenes de referencia para la codificación mediante predicción inter de una imagen actual para la cual existe es un nivel relativamente alto de certeza de que estas imágenes de referencia estarán disponibles en una memoria intermedia (por ejemplo, memoria intermedia de imágenes descodificadas (DPB)) del descodificador de vídeo. A modo de ejemplo únicamente con fines ilustrativos, el descodificador de vídeo puede comunicarse con el codificador de vídeo para que el codificador de vídeo determine las imágenes de referencia candidatas con el fin de limitar la propagación de errores.

[0019] En algunos casos, es posible que un descodificador de vídeo no reciba algunos o todos los datos de vídeo necesarios para reconstruir una imagen. Una imagen que el descodificador de vídeo no puede reconstruir se conoce como una imagen perdida. Además de que el descodificador de vídeo no puede emitir la imagen perdida, si la imagen perdida se usa como imagen de referencia para una imagen subsiguiente, los errores de descodificación se propagan porque el descodificador de vídeo puede no ser capaz de reconstruir esta imagen subsiguiente.

[0020] Para limitar los errores de propagación, el descodificador de vídeo puede enviar información al codificador de vídeo que el codificador de vídeo usa para seleccionar imágenes de referencia, para realizar predicción inter de una imagen actual, para lo cual existe una probabilidad relativamente alta de que las imágenes de referencia se almacenen en la DPB. del descodificador de vídeo. En algunos casos, puede garantizarse virtualmente que las imágenes seleccionadas están en la DPB del descodificador de vídeo. De esta manera, al usar imágenes con una probabilidad relativamente alta de existir en la DPB del descodificador de vídeo, el codificador de vídeo puede no usar la imagen perdida como imagen de referencia, y el descodificador de vídeo podría reconstruir la imagen actual sin error.

[0021] Debe entenderse que la limitación de la propagación de errores se describe con fines ilustrativos, y la

comunicación de información desde el descodificador de vídeo al codificador de vídeo también se puede realizar por otras razones. El descodificador de vídeo puede generar información que el codificador de vídeo usa para seleccionar imágenes de referencia por razones adicionales o independientes a la limitación de la propagación de errores. Por ejemplo, el descodificador de vídeo puede emitir información que el codificador de vídeo usa para seleccionar imágenes de referencia en ejemplos donde no había una imagen perdida.

[0022] En las técnicas descritas en esta divulgación, el descodificador de vídeo puede transmitir información al codificador de vídeo, en el que la información identifica una imagen previamente descodificada que actualmente está almacenada en la DPB del descodificador de vídeo en el caso (por ejemplo, cuando) el descodificador de vídeo transmite la información al codificador de vídeo. La imagen identificada puede corresponder a una posible imagen de referencia candidata (es decir, la imagen identificada puede ser una de las imágenes disponibles para la codificación). Sin embargo, de acuerdo con las técnicas descritas en esta divulgación, además de la imagen identificada, una o más imágenes que el codificador de vídeo había determinado como posibles imágenes de referencia para la imagen identificada también son imágenes de referencia candidatas adicionales (es decir, una o más imágenes que el codificador de vídeo había determinado como posibles imágenes de referencia para la imagen identificada son imágenes que están disponibles para la codificación y, en algunos ejemplos, también incluyen posibles imágenes de referencia de las imágenes de referencia de la imagen identificada). Cabe señalar que una o más de las posibles imágenes de referencia para la imagen identificada pueden haber sido imágenes de referencia reales para la imagen identificada. En otras palabras, las "posibles imágenes de referencia para la imagen identificada" no deben interpretarse como excluyentes de las imágenes de referencia que realmente se usaron para realizar predicción inter de la imagen identificada.

[0023] En algunos ejemplos, al codificar una imagen, el codificador de vídeo puede determinar las imágenes que podrían haberse utilizado para codificar mediante predicción inter de la imagen (incluidas las imágenes que se utilizan realmente para codificar mediante predicción inter de la imagen). El codificador de vídeo puede almacenar información que indica las imágenes que podrían haberse utilizado para codificar mediante predicción inter las imágenes respectivas y las imágenes siguientes. En respuesta a la recepción de información que identifica una imagen particular del descodificador de vídeo, el codificador de vídeo puede recuperar la información almacenada que indica las imágenes que podrían haber sido usadas para codificar mediante predicción inter de la imagen identificada y las siguientes imágenes. El codificador de vídeo puede determinar qué imágenes indicadas en la información recuperada aún están almacenadas en una memoria intermedia del codificador de vídeo, y luego puede seleccionar una o más imágenes de referencia de las imágenes determinadas para codificar mediante predicción inter una imagen actual.

[0024] Por ejemplo, la norma de codificación de vídeo de alta eficiencia (HEVC) define una pluralidad de subconjuntos de imágenes de referencia que el codificador de vídeo y el descodificador de vídeo construyen para cada imagen. Para una imagen en particular, tres de los subconjuntos de imágenes de referencia indican imágenes de referencia que potencialmente se pueden usar para realizar predicción inter de la imagen en particular y las imágenes que siguen a la imagen en particular en orden de descodificación. Dos de los subconjuntos de imágenes de referencia indican imágenes de referencia que no se pueden usar para realizar predicción inter de la imagen en particular, pero potencialmente se pueden usar para realizar predicción inter de las imágenes que siguen a la imagen en particular en orden de descodificación. El término "conjunto de imágenes de referencia" se refiere a la combinación de la pluralidad de subconjuntos de imágenes de referencia.

[0025] Un codificador de vídeo y un descodificador de vídeo configurados de acuerdo con el la norma HEVC pueden explotar los subconjuntos de imágenes de referencia definidos en el la norma HEVC para implementar las técnicas de ejemplo descritas en esta divulgación. Sin embargo, para un codificador de vídeo y un descodificador de vídeo configurados de acuerdo con el la norma H.264/AVC, o de otra manera, es posible que no tenga subconjuntos de imágenes de referencia predefinidos. Para tales codificadores de vídeo y descodificadores de vídeo, el codificador de vídeo puede configurarse para mantener una lista de posibles imágenes de referencia que el codificador de vídeo podría haber usado para la predicción inter de imágenes.

[0026] Para facilitar la comprensión, el término "conjunto de imágenes" de una imagen en particular, como se usa en esta divulgación, se refiere a las imágenes que podrían haberse utilizado para realizar predicción inter de la imagen en particular y las imágenes que siguen a la imagen identificada en orden de codificación. Como ejemplo, en el contexto de HEVC, el término "conjunto de imágenes" de una imagen en particular se refiere a las imágenes, pero no necesariamente a todas las imágenes, en un conjunto de imágenes de referencia de la imagen en particular. Como otro ejemplo, en el contexto de H.264/AVC o de otra manera, el término "conjunto de imágenes" de una imagen en particular se refiere a las imágenes indicadas en una lista de imágenes mantenidas que el codificador de vídeo podría haber utilizado para realizar predicción inter de la imagen en particular.

[0027] En las técnicas descritas en esta divulgación, en respuesta a recibir información que identifica una imagen almacenada en la DPB del descodificador de vídeo, el codificador de vídeo puede determinar qué imágenes en el conjunto de imágenes de la imagen identificada aún están disponibles en la DPB del codificador de vídeo (es decir, actualmente almacenadas en la DPB del codificador de vídeo), y en algunos ejemplos, el codificador de vídeo también puede determinar qué imágenes en el conjunto de imágenes de las imágenes en el conjunto de imágenes de la imagen identificada todavía están disponibles en la DPB del codificador de vídeo. Las imágenes consideradas como imágenes

que están disponibles para codificar una imagen actual (es decir, imágenes de referencia candidatas para una imagen actual) pueden incluir: imágenes que se determina que están en el conjunto de imágenes de la imagen identificada, imágenes en el conjunto de imágenes de las imágenes en el conjunto de imágenes de la imagen identificada que todavía están disponibles en la DPB del codificador de vídeo, y la imagen identificada en sí. El codificador de vídeo puede seleccionar una o más imágenes de referencia de las imágenes que están disponibles para codificar la imagen actual, y realizar predicción inter de la imagen actual a partir de la(s) imagen(es) de referencia seleccionada(s).

[0028] En general, puede haber una certeza muy alta (p. ej., virtualmente garantizada) de que una imagen se almacena en la DPB del codificador de vídeo, que podría haber sido una imagen de referencia para la imagen identificada y las imágenes que siguen a la imagen identificada en el orden de codificación, en el momento en que el codificador de vídeo codifica la imagen actual también se almacena en la DPB del descodificador de vídeo en el momento en que el descodificador de vídeo descodifica la imagen actual. Esta certeza muy alta puede deberse al hecho de que el descodificador de vídeo puede eliminar una imagen de su DPB después de que el codificador de vídeo elimine la imagen de su DPB. En otras palabras, si hay una imagen en la DPB que podría haber sido utilizada para realizar predicción inter de la imagen identificada y las imágenes que siguen a la imagen identificada en orden de codificación, y el codificador de vídeo aún no ha eliminado esa imagen de su DPB, entonces es muy probable que el descodificador de vídeo no haya eliminado esa imagen de su DPB. Por lo tanto, aunque no todas las imágenes en la DPB del codificador de vídeo tienen que estar en la DPB del descodificador de vídeo, y viceversa, si hay una imagen en la DPB del codificador de vídeo que podría haberse utilizado para realizar predicción inter de la imagen identificada y las imágenes que siguen a la imagen identificada en el orden de codificación, que el descodificador de vídeo identificó como una imagen almacenada en su DPB, entonces esa imagen está virtualmente garantizada en la DPB del descodificador de vídeo.

[0029] En las técnicas descritas en esta divulgación, el codificador de vídeo puede ser capaz de determinar una pluralidad de imágenes, que se sabe que están disponibles para descodificación mediante el descodificador de vídeo, que se pueden usar para realizar predicción inter de imagen actual. Por ejemplo, puede ser posible que el descodificador de vídeo transmita información que identifica una pluralidad de imágenes en su DPB, y teniendo el codificador de vídeo entonces seleccione una o más de estas imágenes identificadas como imágenes de referencia para la imagen actual. Sin embargo, hacer que el descodificador de vídeo transmita información que identifica una pluralidad de imágenes en su DPB puede requerir un uso intensivo del ancho de banda. Sin embargo, hacer que el descodificador de vídeo identifique solo una imagen y que el codificador de vídeo use esa imagen para realizar predicción inter de la imagen actual puede dar como resultado una eficiencia de codificación deficiente, lo que a su vez resulta en un mayor uso de ancho de banda.

[0030] Las técnicas descritas en esta divulgación pueden permitir que el descodificador de vídeo identifique solo unas pocas imágenes (por ejemplo, una o más imágenes) en su DPB, y permitir que el codificador de vídeo determine imágenes adicionales (por ejemplo, una pluralidad de imágenes) basadas en la(s) imagen(es) identificada(s) (p. ej., imágenes en el conjunto de imágenes de referencia de la(s) imagen(s) identificada(s) que todavía están disponibles en la DPB del codificador de vídeo) que se pueden usar para realizar predicción inter de una imagen actual. De esta manera, las técnicas descritas en esta divulgación requieren un menor uso del ancho de banda (por ejemplo, el descodificador de vídeo puede emitir información que identifica una imagen), pero proporciona una mayor eficiencia de codificación porque el codificador de vídeo puede seleccionar de una pluralidad de imágenes las imágenes que brindarán la mejor eficiencia de codificación para realizar predicción inter de la imagen actual.

[0031] Por ejemplo, un codificador de vídeo puede recibir, desde un dispositivo que incluye un descodificador de vídeo, información que identifica una imagen previamente descodificada que se almacena en una DPB del descodificador de vídeo en el momento en que el descodificador de vídeo transmitió la información (por ejemplo, cuando el descodificador de vídeo transmita información identificando una imagen, esa imagen se almacena en la DPB del descodificador de vídeo). El codificador de vídeo puede determinar una o más imágenes de referencia candidatas que están disponibles para codificar una imagen actual basándose en la imagen identificada. La una o más imágenes de referencia candidatas incluyen una o más imágenes que previamente se determinó que eran utilizables para la codificación mediante predicción inter de una o más de la imagen o imágenes identificadas que siguen a la imagen identificada en el orden de codificación, y que aún se almacenan en una DPB del codificador de vídeo. El codificador de vídeo puede seleccionar una o más imágenes de referencia para la imagen actual a partir de una o más imágenes de referencia candidatas determinadas y la imagen identificada. El codificador de vídeo puede entonces codificar mediante predicción inter de la imagen actual basándose en una o más imágenes de referencia seleccionadas.

[0032] Un descodificador de vídeo, en las técnicas de ejemplo descritas en esta divulgación, puede enviar, a un dispositivo que incluye un codificador de vídeo, información que identifica una imagen previamente descodificada almacenada en su DPB en el momento de emitir la información. El descodificador de vídeo puede descodificar mediante predicción inter una imagen actual con una o más imágenes de referencia candidatas. La una o más imágenes de referencia candidatas incluyen una o más imágenes que el codificador de vídeo determinó previamente para que se puedan utilizar para codificación mediante predicción inter de la imagen identificada.

[0033] La FIG. 1 es un diagrama de bloques que ilustra un ejemplo de sistema 10 de codificación y descodificación

de vídeo que puede utilizar las técnicas descritas en esta divulgación. Como se muestra en la FIG. 1, el sistema 10 incluye un dispositivo 12 y un dispositivo 14. En algunos ejemplos, el dispositivo 12 y el dispositivo 14 pueden ser dispositivos utilizados para videotelefonía o videoconferencia en los que el vídeo se codifica y descodifica en "tiempo real" (por ejemplo, como capturado, codificado y transmitido y recibido, descodificado y mostrado con muy poco retardo). Para facilitar la descripción, las técnicas se describen con respecto a la videotelefonía o videoconferencia, pero las técnicas no son tan limitadas. En algunos ejemplos, las técnicas pueden utilizarse para la transmisión continua. Las técnicas descritas en esta divulgación no necesitan limitarse a aplicaciones en tiempo real.

[0034] El dispositivo 12 y el dispositivo 14 pueden comprender cualquiera de entre una amplia gama de dispositivos, incluidos ordenadores de escritorio, ordenadores plegables (es decir, portátiles), ordenadores tipo tablet, descodificadores, teléfonos tales como los denominados teléfonos «inteligentes», los denominados paneles «inteligentes», televisores, cámaras, dispositivos de visualización, reproductores de medios digitales, consolas de videojuegos, un dispositivo de transmisión continua de vídeo o similares. En algunos casos, el dispositivo 12 y el dispositivo 14 pueden estar equipados para la comunicación inalámbrica.

[0035] Para el vídeo en tiempo real, el dispositivo 12 puede transmitir el vídeo que el dispositivo 14 descodifica, y el dispositivo 14 puede transmitir el vídeo que el dispositivo 12 descodifica. En este sentido, en algunos casos, el dispositivo 12 puede funcionar como un dispositivo de origen y como un dispositivo de destino, y el dispositivo 14 puede funcionar como un dispositivo de destino y como un dispositivo de origen.

[0036] Por ejemplo, el dispositivo 12 se ilustra incluyendo componentes como el codificador de vídeo 20A para transmitir vídeo y el dispositivo 14 se ilustra como componentes que incluyen un descodificador de vídeo 30B para recibir datos de vídeo. Sin embargo, dado que el dispositivo 12 también puede necesitar recibir vídeo, el dispositivo 12 también puede incluir el descodificador de vídeo 30A. Además, dado que el dispositivo 14 también puede necesitar transmitir vídeo, el dispositivo 14 también puede incluir el codificador de vídeo 20B.

[0037] Para facilitar la descripción, las técnicas se describen con respecto al dispositivo 12 que transmite el vídeo que recibe el dispositivo 14. Sin embargo, debe entenderse que el dispositivo 14 también transmite vídeo, en algunos ejemplos, y puede configurarse, en general, para transmitir datos al dispositivo 12.

[0038] Por ejemplo, en algunos ejemplos, el dispositivo 12 y el dispositivo 14 pueden estar en comunicación entre sí, donde ambos dispositivos están configurados para transmitir datos y recibir datos entre sí. En una aplicación de vídeo en tiempo real, el dispositivo 14 puede necesitar transmitir datos (por ejemplo, su propio vídeo codificado u otros datos, como la información que el dispositivo 12 usa para codificar el vídeo que recibe el dispositivo 14). Por lo tanto, como se ilustra en la FIG. 1, el dispositivo 14 puede incluir componentes para transmitir datos al dispositivo 12 (por ejemplo, la fuente de vídeo 18B, el codificador de vídeo 20B y la interfaz de entrada/salida 22B que transmiten el vídeo codificado del codificador de vídeo 20B al dispositivo 12). Del mismo modo, como se ilustra en la FIG. 1, el dispositivo 12 puede incluir componentes para recibir datos del dispositivo 14 (por ejemplo, la interfaz de entrada/salida 22A, el descodificador de vídeo 30A, el dispositivo de visualización 31A).

[0039] En otras palabras, el dispositivo 12 y el dispositivo 14 pueden incluir componentes sustancialmente similares para permitir la comunicación bidireccional. Por ejemplo, la fuente de vídeo 18A y la fuente de vídeo 18B pueden capturar vídeos respectivos. El codificador de vídeo 20A codifica el vídeo de la fuente de vídeo 18A, y el codificador de vídeo 20B codifica el vídeo de la fuente de vídeo 18B. La interfaz de entrada/salida 22A transmite los datos de vídeo del codificador de vídeo 20A a través del enlace 16 a la interfaz de entrada/salida 22B, y la interfaz de entrada/salida 22B transmite los datos de vídeo del codificador de vídeo 20B a través del enlace 16 a la interfaz de entrada/salida 22A. La interfaz de entrada/salida 22B puede transmitir el vídeo al descodificador de vídeo 30B para la descodificación de vídeo, y luego al dispositivo de visualización 31B, desde el descodificador de vídeo 30B, para su visualización. La interfaz de entrada/salida 22A puede transmitir el vídeo al descodificador de vídeo 30A para la descodificación de vídeo, y luego al dispositivo de visualización 31A, desde el descodificador de vídeo 30A, para su visualización.

[0040] Aunque los ejemplos anteriores de la aplicación de vídeo en tiempo real se describen en el contexto de la videotelefonía o videoconferencia donde el dispositivo 12 y el dispositivo 14 transmiten y reciben vídeo, las técnicas descritas en esta divulgación no están tan limitadas. En algunos ejemplos, el dispositivo 12 puede transmitir datos de vídeo de un evento en vivo al dispositivo 14. En este ejemplo, el dispositivo 12 puede no necesitar recibir vídeo del dispositivo 14 (y no necesita incluir el descodificador de vídeo 30A o el dispositivo de pantalla 31A); sin embargo, las técnicas descritas en esta divulgación son aplicables a tales ejemplos de aplicaciones de vídeo en tiempo real.

[0041] El dispositivo 12 y el dispositivo 14 pueden transmitir y recibir datos de vídeo codificados para ser descodificados a través de un enlace 16. El enlace 16 puede comprender cualquier tipo de medio o dispositivo capaz de mover los datos de vídeo codificados entre el dispositivo 12 y el dispositivo 14. En un ejemplo, el enlace 16 puede comprender un medio de comunicación para permitir que el dispositivo 12 y el dispositivo 14 transmitan datos de vídeo codificados directamente entre sí en tiempo real. Los datos de vídeo codificados pueden modularse de acuerdo con un la norma de comunicación, como un protocolo de comunicación inalámbrica, y transmitirse. El medio de comunicación puede comprender cualquier medio de comunicación, inalámbrica o cableada, tal como un espectro de

radiofrecuencia (RF) o una o más líneas de transmisión física. El medio de comunicación puede formar parte de una red basada en paquetes, tal como una red de área local, una red de área amplia o una red global tal como Internet. El medio de comunicación puede incluir routers, conmutadores, estaciones base o cualquier otro equipo que pueda ser útil para facilitar la comunicación del dispositivo 12 al dispositivo 14. De esta manera, el enlace 16 funciona como un medio por el cual el dispositivo 12 y el dispositivo 14 realizan comunicación bidireccional.

[0042] En algunos ejemplos descritos en el presente documento, un flujo de bits de vídeo se puede comunicar a través de un protocolo de dos vías. Por ejemplo, los protocolos de ejemplo pueden estar basados en el Protocolo de Internet (IP), como el Protocolo de transporte en tiempo real (RTP). En particular, algunos ejemplos incluyen mensajes de protocolo adicionales para RTP y protocolos similares para mejorar la comunicación continua como se analiza en el presente documento. RTP se utiliza en general para videotelefonía y videoconferencia, así como para transmitir contenido continuo.

[0043] Por ejemplo, en la FIG. 1, el descodificador de vídeo 30A, el codificador de vídeo 20A, el descodificador de vídeo 30B y el codificador de vídeo 20B se ilustran con flechas dobles a la respectiva interfaz de entrada/salida 22A y 22B. En algunos ejemplos, además de recibir datos de vídeo, el descodificador de vídeo 30A puede transmitir información que es utilizada por el codificador de vídeo 20B para la codificación de vídeo y, además de recibir datos de vídeo, el descodificador de vídeo 30B puede transmitir información que es utilizada por el codificador de vídeo 20A para la codificación de vídeo. Dicha transmisión de información desde el descodificador de vídeo 30A al codificador de vídeo 20B y desde el descodificador de vídeo 30B al codificador de vídeo 20A puede estar de acuerdo con el RTP (por ejemplo, mensajes de protocolo adicionales para RTP) o pueden ser protocolos similares para permitir la transmisión de información en tiempo real desde un descodificador de vídeo a un codificador de vídeo.

[0044] De forma alternativa, los datos codificados pueden enviarse desde la interfaz de entrada/salida 22A a un dispositivo de almacenamiento 32 para acceder por la interfaz de entrada/salida 22B. De manera similar, los datos codificados pueden enviarse desde la interfaz de entrada/salida 22B al dispositivo de almacenamiento 32 para acceder a través de la interfaz de entrada/salida 22A. El dispositivo de almacenamiento 32 puede incluir cualquiera de una diversidad de medios de almacenamiento de datos de acceso distribuido o local, tales como unidad de 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 32 puede corresponder a un servidor de archivos u otro dispositivo de almacenamiento intermedio que pueda contener el vídeo codificado generado por el dispositivo 12 o el dispositivo 14 para su posterior acceso por parte del otro. El servidor de archivos puede ser cualquier tipo de servidor capaz de almacenar datos de vídeo codificados y transmitir esos datos de vídeo codificados. Entre los ejemplos de servidores de archivos se incluyen un servidor web (por ejemplo, para un sitio web), un servidor FTP, dispositivos de almacenamiento conectado en red (NAS) o una unidad de disco local. El dispositivo 12 y el dispositivo 14 pueden acceder a los datos de vídeo codificados a través de cualquier conexión de datos la norma, incluida una conexión a Internet. Esto puede incluir un canal inalámbrico (por ejemplo, una conexión wifi), una conexión alámbrica (por ejemplo, DSL, módem de cable, etc.) o una combinación de ambos que sea adecuada para acceder a datos de vídeo codificados almacenados en un servidor de archivos. La transmisión de datos de vídeo codificados desde el dispositivo 32 de almacenamiento puede ser una transmisión en continuo, una transmisión de descarga o una combinación de ambas.

[0045] Las técnicas de esta divulgación no están limitadas necesariamente a aplicaciones o configuraciones inalámbricas. Las técnicas pueden aplicarse a la codificación de vídeo, como apoyo a cualquiera de una diversidad de aplicaciones de multimedia, tales como radiodifusiones de televisión por aire, transmisiones de televisión por cable, transmisiones de televisión por satélite, transmisiones de vídeo en continuo, por ejemplo, mediante Internet, codificación de vídeo digital para su almacenamiento en un medio de almacenamiento de datos, descodificación de vídeo digital almacenado en un medio de almacenamiento de datos, u otras aplicaciones. En algunos ejemplos, el sistema 10 puede estar configurado para soportar la transmisión de vídeo unidireccional o bidireccional para soportar aplicaciones tales como la transmisión en continuo de vídeo, la reproducción de vídeo, la radiodifusión de vídeo y/o la videotelefonía.

[0046] En algunos casos, cada una de las interfaces de entrada/salida 22A y 22B puede incluir un modulador/desmodulador (módem) y/o un transmisor. En el dispositivo 12 y el dispositivo 14, la fuente de vídeo 18A y 18B, respectivamente, puede incluir una fuente como un dispositivo de captura de vídeo (por ejemplo, una cámara de vídeo), un archivo de vídeo que contiene vídeo capturado previamente, una interfaz de alimentación de vídeo para recibir vídeo de un proveedor de contenido de vídeo y/o un sistema de gráficos de ordenador para generar datos de gráficos de ordenador como fuente de vídeo, o una combinación de tales fuentes. Como ejemplo, si las fuentes de vídeo 18A y 18B son cámaras de vídeo, el dispositivo 12 y el dispositivo 14 pueden formar los llamados teléfonos con cámara o teléfonos con vídeo. Sin embargo, las técnicas descritas en esta divulgación pueden ser aplicables a la codificación de vídeo en general, y pueden aplicarse a aplicaciones inalámbricas y/o cableadas.

[0047] El vídeo capturado, precapturado o generado por ordenador puede codificarse mediante los codificadores de vídeo 20A y 20B respectivos. Los datos de vídeo codificados pueden transmitirse directamente desde el dispositivo 12 al dispositivo 14, y viceversa, a través de la respectiva interfaz de entrada/salida 22A y 22B. Los datos de vídeo codificados también pueden (o de forma alternativa) almacenarse en el dispositivo de almacenamiento 32 para un

acceso posterior para la descodificación y/o reproducción.

[0048] Los datos de vídeo codificados comunicados a través del enlace 16, o proporcionados en el dispositivo de almacenamiento 32, pueden incluir una variedad de elementos sintácticos generados por el codificador de vídeo 20A para su uso mediante el descodificador de vídeo 30B en la descodificación de los datos de vídeo, y pueden incluir una variedad de elementos sintácticos generados por el codificador de vídeo 20B para uso mediante el descodificador de vídeo 30A para descodificar los datos de vídeo. Dichos elementos sintácticos pueden incluirse con los datos de vídeo codificados, transmitidos en un medio de comunicación, almacenados en un medio de almacenamiento o almacenados en un servidor de ficheros.

[0049] El dispositivo de visualización 31A y 31B puede estar con un componente del dispositivo 12 o dispositivo 14 o puede ser un componente independiente que sea externo al dispositivo 12 o al dispositivo 14. En algunos ejemplos, el dispositivo 12 y el dispositivo 14 pueden incluir un componente de dispositivo de visualización, y también pueden configurarse para interactuar con otro dispositivo de visualización externo. En otros ejemplos, el dispositivo 12 y el dispositivo 14 pueden ser dispositivos de visualización. En general, los dispositivos de visualización 31A y 31B muestran los datos de vídeo descodificados a un usuario y pueden comprender cualquiera de una variedad de dispositivos de visualización, tales como una pantalla de cristal líquido (LCD), una pantalla de plasma, una pantalla de diodos orgánicos emisores de luz (OLED) u otro tipo de dispositivo de visualización.

[0050] Los codificadores de vídeo 20A y 20B (colectivamente "codificadores de vídeo 20") y los descodificadores de vídeo 30A y 30B (colectivamente "descodificadores de vídeo 30") pueden funcionar de acuerdo con una norma de compresión de vídeo, como la norma de codificación de vídeo de alta eficiencia (HEVC), y pueden estar conforme con el modelo de prueba HEVC (HM). Además, los codificadores de vídeo 20 y los descodificadores de vídeo 30 pueden funcionar de acuerdo con las extensiones de la norma HEVC (denominada norma de extensión de rango HEVC) actualmente en desarrollo. Los codificadores de vídeo 20 y los descodificadores de vídeo 30 también pueden funcionar de acuerdo con otras normas de propiedad o de la industria, como la norma ITU-T H.264, conocida como MPEG-4, parte 10, Codificación de vídeo avanzada (AVC) (es decir, H.264/AVC), o extensiones de dichas normas. Sin embargo, las técnicas de esta divulgación no están limitadas a ninguna norma de codificación particular. Otros ejemplos de normas de compresión de vídeo incluyen MPEG-2 e ITU-T H.263.

[0051] Aunque no se muestra en la FIG. 1, en algunos aspectos, tanto los codificadores de vídeo 20 como los descodificadores de vídeo 30 pueden estar integrados en un codificador y descodificador de audio, y pueden incluir unidades MUX-DEMUX adecuadas, u otro tipo de hardware y software, para ocuparse de la codificación tanto de audio como de vídeo en un flujo de datos común o en flujos de datos separados. Si procede, en algunos ejemplos, las unidades MUX-DEMUX pueden ser conformes al protocolo de multiplexador ITU H.223 o a otros protocolos, tales como el protocolo de datagramas de usuario (UDP).

[0052] Los codificadores de vídeo 20 y los descodificadores de vídeo 30 pueden implementarse como cualquiera de una variedad de circuitos de codificadores adecuados, tales como uno o más microprocesadores, procesadores de señales digitales (DSP), circuitos integrados específicos de la aplicación (ASIC), matrices de puertas programables *in situ* (FPGA), lógica discreta, software, hardware, firmware o cualquier combinación de los mismos. Cuando las técnicas se implementan parcialmente en software, un dispositivo puede almacenar instrucciones para el software en un medio adecuado no transitorio legible por ordenador, y ejecutar las instrucciones en hardware mediante uno o más procesadores para realizar las técnicas de esta divulgación.

[0053] Cada uno del codificador de vídeo 20 y el descodificador de vídeo 30 puede estar incluido en uno o más codificadores o descodificadores, cualquiera de los cuales puede estar integrado como parte de un codificador/descodificador (CÓDEC) combinado en un dispositivo respectivo. En algunos ejemplos, cuando esta divulgación describe codificadores de vídeo 20 y descodificadores de vídeo 30 que forman parte de un dispositivo, el dispositivo puede ser un circuito integrado (IC) o un microprocesador (por ejemplo, el codificador de vídeo 20A y el descodificador de vídeo 30A pueden integrarse juntos y posiblemente con otros componentes en un circuito integrado o microprocesador, y el codificador de vídeo 20B y el descodificador de vídeo 30B se pueden integrar juntos y posiblemente con otros componentes en un circuito integrado o microprocesador). En algunos ejemplos, el dispositivo puede ser un dispositivo inalámbrico, como un dispositivo de comunicación inalámbrica como el dispositivo 12 o el dispositivo 14.

[0054] La siguiente es una breve descripción de la norma HEVC para facilitar la comprensión. Sin embargo, las técnicas descritas en esta divulgación no están limitadas a la norma HEVC, y son aplicables a otras normas como H.264/AVC, o aplicables a la codificación de vídeo no basada en normas. Para facilitar la descripción, se describe lo siguiente con respecto al codificador de vídeo 20A y al descodificador de vídeo 30B. Sin embargo, el codificador de vídeo 20B y el descodificador de vídeo 30A pueden configurarse de manera similar al codificador de vídeo 20A y el descodificador de vídeo 30B, respectivamente. En algunos ejemplos, puede ser posible que el codificador de vídeo 20A y el codificador de vídeo 20B se configuren de manera diferente. Por ejemplo, el descodificador de vídeo 30A y el descodificador de vídeo 30B pueden configurarse de manera diferente, pero el codificador de vídeo 20A puede generar datos de vídeo que son descodificables por el descodificador de vídeo 30B y el codificador de vídeo 20B puede generar datos de vídeo que son descodificables por el descodificador de vídeo 30A. En otras palabras, el

codificador de vídeo 20A y el codificador de vídeo 20B pueden configurarse de manera similar, y el descodificador de vídeo 30A y el descodificador de vídeo 30B pueden configurarse de manera similar. Sin embargo, esto no es un requisito, y el codificador de vídeo 20A y el codificador de vídeo 20B pueden configurarse de manera diferente, y el descodificador de vídeo 30A y el descodificador de vídeo 30B pueden configurarse de manera diferente.

[0055] El JCT-VC ha desarrollado la norma HEVC. La iniciativa de normalización HEVC se basa en un modelo en evolución de un dispositivo de codificación de vídeo denominado modelo de prueba HEVC (HM). El HM supone varias capacidades adicionales de los dispositivos de codificación de vídeo respecto a dispositivos existentes de acuerdo con, por ejemplo, la norma ITU-T H.264/AVC. Por ejemplo, mientras que la norma H.264 proporciona nueve modos de codificación mediante predicción intra, el HM puede proporcionar hasta treinta y tres modos de codificación mediante predicción intra.

[0056] Un borrador reciente de la norma HEVC, denominado "Borrador 10 de trabajo de la HEVC" o "WD10", se describe en el documento JCTVC-H1003, de Bross et al., titulado "High Efficiency Video Coding (HEVC) Text Specification Draft 10 (for FDIS & Last Call)" ["Memoria descriptiva textual de la Codificación de Vídeo de Alta Eficiencia (HEVC), Borrador 10 (para Norma Internacional de Borrador Final y Última Llamada)"], Equipo de Colaboración Conjunta en Codificación de Vídeo (JCT-VC) de ITU-T SG16 WP3 e ISO/IEC JTC1/SC29/WG11, décimo segunda conferencia: Ginebra, Suiza, 14-23 de enero de 2013, que se puede descargar en: http://phenix.int-evry.fr/jct/doc_end_user/documents/12_Geneva/wg11/JCTVC-L1003-v34.zip.

[0057] Otro borrador de la norma HEVC, se denomina en el presente documento "revisiones WD10" descritas en Bross et al., "Correcciones propuestas por los editores a la versión 1 de HEVC," Equipo de colaboración conjunta sobre codificación de vídeo (JCT-VC) de ITU-T SG16 WP3 e ISO/IEC JTC1/SC29/WG11, 13.ª reunión, Incheon, Corea, abril de 2013, disponible en: http://phenix.int-evry.fr/jct/doc_end_user/documents/13_Incheon/wg11/JCTVC-M0432-v3.zip.

[0058] En general, el modelo de explotación del HM describe que una trama o imagen de vídeo puede dividirse en una secuencia de bloques de árbol o unidades de codificación de máximo tamaño (LCU), que incluyen muestras tanto de luma como de croma. Un bloque de árbol tiene un fin similar al de un macrobloque de la norma H.264. Un fragmento incluye un número de bloques de árbol consecutivos en orden de codificación. Una trama o imagen de vídeo puede dividirse en uno o más fragmentos. Cada bloque de árbol puede dividirse en unidades de codificación (CU) de acuerdo con un árbol cuaternario. Por ejemplo, un bloque de árbol, como un nodo raíz del árbol cuaternario, puede dividirse en cuatro nodos hijo, y cada nodo hijo puede, a su vez, ser un nodo padre y dividirse en otros cuatro nodos hijo. Un nodo hijo final no dividido, como un nodo hoja del árbol cuaternario, comprende un nodo de codificación, es decir, un bloque de vídeo codificado. Los datos sintácticos asociados a un flujo de bits codificado pueden definir un número máximo de veces que puede dividirse un bloque de árbol, y también pueden definir un tamaño mínimo de los nodos de codificación.

[0059] Una CU incluye un nodo de codificación y unidades de predicción (PU) y unidades de transformada (TU) asociadas al nodo de codificación. Un tamaño de la CU corresponde a un tamaño del nodo de codificación y debe ser de forma cuadrada. El tamaño de la CU puede variar desde 8x8 píxeles hasta el tamaño del bloque de árbol, con un máximo de 64x64 píxeles o más. Cada CU puede contener una o más PU y una o más TU. Los datos sintácticos asociados a una CU pueden describir, por ejemplo, la división de la CU en una o más PU. Los modos de división pueden diferir dependiendo de si la CU está codificada en modo de salto o directo, codificada en modo de predicción intra o codificada en modo de predicción inter. Las PU pueden dividirse para tener forma no cuadrada. Los datos sintácticos asociados a una CU también pueden describir, por ejemplo, la división de la CU en una o más TU de acuerdo con un árbol cuaternario. Una TU puede tener forma cuadrada o no cuadrada.

[0060] La norma HEVC admite transformadas de acuerdo con las TU, que pueden ser diferentes para diferentes CU. El tamaño de las TU típicamente se basa en el tamaño de las PU de una CU dada, definida para una LCU dividida, aunque puede que no sea siempre así. Las TU son típicamente del mismo tamaño o de un tamaño más pequeño que las PU. En algunos ejemplos, las muestras residuales correspondientes a una CU pueden subdividirse en unidades más pequeñas mediante una estructura de árbol cuaternario conocida como «árbol cuaternario residual» (RQT). Los nodos hoja del RQT pueden denominarse TU. Los valores de diferencia de píxeles asociados a las TU pueden transformarse para generar coeficientes de transformada, que pueden cuantificarse.

[0061] En general, una PU incluye datos relacionados con el proceso de predicción. Por ejemplo, cuando la PU se codifica mediante el modo intra, la PU puede incluir datos que describen un modo de predicción intra para la PU. En otro ejemplo, cuando la PU está codificada en modo de predicción inter, la PU puede incluir datos que definen un vector de movimiento para la PU. Los datos que definen el vector de movimiento para una PU pueden describir, por ejemplo, un componente horizontal del vector de movimiento, un componente vertical del vector de movimiento, una resolución para el vector de movimiento (por ejemplo, precisión de un cuarto de píxel o precisión de un octavo de píxel), una imagen de referencia a la que apunta el vector de movimiento y/o una lista de imágenes de referencia (por ejemplo, Lista 0 (RefPicList0), Lista 1 (RefPicList1), o Lista C) para el vector de movimiento.

[0062] En general, se usa una TU para los procesos de transformada y cuantificación. Una CU dada que tiene una

o más PU también puede incluir una o más TU. Tras la predicción, el codificador de vídeo 20A puede calcular valores residuales correspondientes a la PU. Los valores residuales comprenden valores de diferencia de píxeles que se pueden transformar en coeficientes de transformada, cuantificar y examinar mediante las TU, para generar coeficientes de transformada en serie para la codificación por entropía. Esta divulgación usa típicamente el término «bloque de vídeo» para referirse a un nodo de codificación de una CU. En algunos casos específicos, esta divulgación también puede usar el término «bloque de vídeo» para referirse a un bloque de árbol, es decir, una LCU o una CU, que incluye un nodo de codificación y unas PU y TU.

[0063] Una secuencia de vídeo incluye típicamente una serie de tramas o imágenes de vídeo. Un grupo de imágenes (GOP) comprende, en general, una serie de una o más de las imágenes de vídeo. Un GOP puede incluir datos sintácticos en una cabecera del GOP, en una cabecera de una o más de las imágenes, o en otras ubicaciones, que describen un número de imágenes incluidas en el GOP. Cada fragmento de una imagen puede incluir datos sintácticos de fragmento que describen un modo de codificación para el fragmento respectivo. Un codificador de vídeo 20A actúa típicamente sobre bloques de vídeo dentro de fragmentos de vídeo individuales con el fin de codificar los datos de vídeo. Un bloque de vídeo puede corresponder a un nodo de codificación dentro de una CU. Los bloques de vídeo pueden tener tamaños fijos o variables y pueden diferir en tamaño de acuerdo con una norma de codificación especificada.

[0064] En un ejemplo, el HM soporta predicción en diversos tamaños de PU. Si se supone que el tamaño de una CU particular es $2N \times 2N$, el HM soporta predicción intra en tamaños de PU de $2N \times 2N$ o $N \times N$ y predicción inter en tamaños de PU simétricos de $2N \times 2N$, $2N \times N$, $N \times 2N$ o $N \times N$. El HM también soporta la división asimétrica para la predicción inter en tamaños de PU de $2N \times nU$, $2N \times nD$, $nL \times 2N$ y $nR \times 2N$. En la división asimétrica, una dirección de una CU no está dividida, mientras que la otra dirección está dividida en el 25 % y el 75 %. La parte de la CU correspondiente a la división del 25 % está indicada por una "n" seguida de una indicación de "arriba", "abajo", "izquierda" o "derecha". Así pues, por ejemplo, « $2N \times nU$ » se refiere a una CU $2N \times 2N$ que está dividida horizontalmente con una PU $2N \times 0,5N$ encima y una PU $2N \times 1,5N$ debajo.

[0065] En esta divulgación, « $N \times N$ » y «N por N» pueden usarse indistintamente para hacer referencia a las dimensiones de píxeles de un bloque de vídeo en términos de dimensiones verticales y horizontales, por ejemplo, 16×16 píxeles o 16 por 16 píxeles. En general, un bloque de 16×16 tendrá 16 píxeles en una dirección vertical ($y = 16$) y 16 píxeles en una dirección horizontal ($x = 16$). Asimismo, un bloque de $N \times N$ tiene, en general, N píxeles en una dirección vertical y N píxeles en una dirección horizontal, donde N representa un valor entero no negativo. Los píxeles de un bloque pueden disponerse en filas y columnas. Además, no es necesario que los bloques presenten necesariamente el mismo número de píxeles en la dirección horizontal y en la dirección vertical. Por ejemplo, los bloques pueden comprender $N \times M$ píxeles, donde M no es necesariamente igual a N.

[0066] Tras la codificación mediante predicción intra o inter mediante las PU de una CU, el codificador de vídeo 20A puede calcular datos residuales para las TU de la CU. Las PU pueden comprender datos de píxeles en el dominio espacial (también denominado dominio del píxel) y las TU pueden comprender coeficientes en el dominio de la transformada tras la aplicación de una transformada, por ejemplo, una transformada discreta de coseno (DCT), una transformada de enteros, una transformada de *wavelet* o una transformada conceptualmente similar a los datos de vídeo residuales. Los datos residuales pueden corresponder a diferencias de píxel entre píxeles de la imagen no codificada y los valores de predicción correspondientes a las PU. El codificador de vídeo 20A puede formar las TU incluyendo los datos residuales para la CU y, a continuación, transformar las TU para generar coeficientes de transformada para la CU.

[0067] Tras cualquier transformada para generar coeficientes de transformada, el codificador de vídeo 20A puede realizar la cuantificación de los coeficientes de transformada. La cuantificación se refiere, en general, a un proceso en el que los coeficientes de transformada se cuantifican para reducir posiblemente la cantidad de datos usados para representar los coeficientes, proporcionando compresión adicional. El proceso de cuantificación puede reducir la profundidad de bits asociada a algunos, o a la totalidad, de los coeficientes. Por ejemplo, un valor de n bits puede redondearse hacia abajo hasta un valor de m bits durante la cuantificación, donde n es mayor que m.

[0068] En algunos ejemplos, el codificador de vídeo 20A puede utilizar un orden de escaneado predefinido para escanear los coeficientes de transformada cuantificados con el fin de generar un vector en serie que se pueda someter a codificación por entropía. En otros ejemplos, el codificador de vídeo 20A puede realizar un escaneado adaptativo. Después de escanear los coeficientes de transformada cuantificados para formar un vector unidimensional, el codificador de vídeo 20A puede realizar la codificación por entropía del vector unidimensional, por ejemplo, de acuerdo con 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 basada en la sintaxis (SBAC), la codificación por entropía de división de intervalo de probabilidad (PIPE) u otros procedimientos de codificación por entropía. El codificador de vídeo 20A también puede realizar la codificación por entropía de los elementos sintácticos asociados a los datos de vídeo codificados, para su uso mediante el descodificador 30B de vídeo en la descodificación de los datos de vídeo.

[0069] Para realizar la CABAC, el codificador de vídeo 20A puede asignar un contexto de un modelo de contexto a

un símbolo que se va a transmitir. El contexto puede indicar, por ejemplo, si los valores contiguos del símbolo son distintos de cero o no. Para realizar la CAVLC, el codificador de vídeo 20A puede seleccionar un código de longitud variable para un símbolo que se va a transmitir. Las palabras de código en la VLC pueden construirse de tal manera que los códigos relativamente más cortos correspondan a símbolos más probables, y los códigos más largos correspondan a símbolos menos probables. De esta manera, 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 puede basarse en un contexto asignado al símbolo.

[0070] La FIG. 2 es un diagrama de bloques que ilustra un codificador de vídeo de ejemplo 20A que puede implementar las técnicas descritas en esta divulgación. Como se describió anteriormente, el codificador de vídeo 20B puede ser similar al codificador de vídeo 20A; sin embargo, las técnicas descritas en esta divulgación no están tan limitadas. El codificador de vídeo 20A puede realizar codificación intra o inter de bloques de vídeo dentro de fragmentos de vídeo. La codificación intra se basa en la predicción espacial para reducir o eliminar la redundancia espacial en el vídeo dentro de una trama o imagen de vídeo determinada. La codificación inter se basa en la predicción temporal para reducir o eliminar la redundancia temporal en el vídeo dentro de tramas o imágenes adyacentes de una secuencia de vídeo. El modo intra (modo I) puede referirse a cualquiera de varios modos de compresión espacial. Los modos inter, tales como la predicción unidireccional (modo P) o la predicción bidireccional (modo B), pueden referirse a cualquiera de varios modos de compresión temporal.

[0071] En el ejemplo de la FIG. 2, el codificador de vídeo 20A incluye una memoria de datos de vídeo 39, una unidad de partición 35, una unidad de procesamiento de predicción 41, una memoria intermedia de imágenes descodificadas (DPB) 64, un sumador 50, una unidad de procesamiento de transformada 52, una unidad de cuantificación 54 y unidad de codificación por entropía 56. La unidad de procesamiento de predicción 41 incluye la unidad de estimación de movimiento 42, la unidad de compensación de movimiento 44 y la unidad de procesamiento de predicción intra 46. Para la reconstrucción de bloques de vídeo, el codificador de vídeo 20A incluye también una unidad de cuantificación inversa 58, una unidad de procesamiento de transformada inversa 60 y un sumador 62. Puede incluirse, asimismo, un filtro de eliminación de bloques (no mostrado en la FIG. 2) para filtrar fronteras de bloques, con el fin de eliminar distorsiones de efecto pixelado del vídeo reconstruido. Si se desea, el filtro de eliminación de bloques filtrará típicamente la salida del sumador 62. También pueden usarse filtros de bucle adicionales (en bucle o tras el bucle), además del filtro de eliminación de bloques.

[0072] Como se muestra en la FIG. 2, la memoria de datos de vídeo 39 recibe datos de vídeo que se usan para codificar un bloque de vídeo actual dentro de una imagen de vídeo. La memoria de datos de vídeo 39 puede almacenar datos de vídeo para ser codificados por los componentes del codificador de vídeo 20A (por ejemplo, configurado para almacenar datos de vídeo) o almacenar datos de vídeo que se usarán para codificar imágenes de vídeo. Por ejemplo, la memoria de datos de vídeo 39 puede almacenar conjuntos de parámetros para cada imagen que codifica el codificador de vídeo 20A (p. ej., almacenar información que indique imágenes que podrían haberse utilizado para realizar predicción inter de imágenes respectivas). De acuerdo con las técnicas descritas en esta divulgación, la memoria de datos de vídeo 39 también puede almacenar información, recibida desde el descodificador de vídeo 30B, identificando una imagen previamente descodificada por el descodificador de vídeo 30B. Por ejemplo, el codificador de vídeo 20A puede configurarse para realizar predicción inter una imagen actual basada en una o más de una imagen identificada por el descodificador de vídeo 30B, imágenes que podrían haberse utilizado para realizar predicción inter de la imagen identificada por el descodificador de vídeo 30B, o imágenes que se podrían usar para realizar predicción inter de las imágenes que se podrían haber utilizado para realizar predicción inter de la imagen identificada por el descodificador de vídeo 30B. En general, si el codificador de vídeo 20A selecciona imágenes de un grupo determinado de imágenes determinadas, existe una alta probabilidad de que las imágenes seleccionadas estén disponibles para descodificación mediante el descodificador de vídeo 30B.

[0073] En algunos ejemplos, los datos de vídeo almacenados en la memoria de datos de vídeo 39 pueden obtenerse, por ejemplo, de la fuente de vídeo 18A. La DPB 64 almacena datos de vídeo de referencia para su uso en la codificación de datos de vídeo mediante el codificador de vídeo 20A (por ejemplo, en modos de codificación intra o modos de codificación inter. La memoria de datos de vídeo 39 y la DPB 64 pueden estar formadas por cualquiera entre varios dispositivos de memoria, tales como memoria dinámica de acceso aleatorio (DRAM), incluyendo DRAM síncrona (SDRAM), RAM magnetorresistiva (MRAM), RAM resistiva (RRAM) u otros tipos de dispositivos de memoria. La memoria de datos de vídeo 39 y la DPB 64 pueden ser proporcionadas por el mismo dispositivo de memoria o por dispositivos de memoria independientes. En varios ejemplos, la memoria de datos de vídeo 39 puede estar en un chip con otros componentes del codificador de vídeo 20A, o fuera de chip con respecto a esos componentes.

[0074] Como se representa en la FIG. 2, el codificador de vídeo 20A recibe datos de vídeo, y la unidad de división 35 divide los datos en bloques de vídeo. Esta división también puede incluir la división en fragmentos, mosaicos u otras unidades mayores, así como la división de bloques de vídeo, por ejemplo, de acuerdo con una estructura de árbol cuaternario de LCU y CU. El codificador de vídeo 20A ilustra, en general, los componentes que codifican bloques de vídeo de un fragmento de vídeo que se va a codificar. El fragmento puede dividirse en múltiples bloques de vídeo (y, posiblemente, en conjuntos de bloques de vídeo denominados mosaicos). La unidad de procesamiento de predicción 41 puede seleccionar una entre una pluralidad de posibles modos de codificación, tal como una entre una pluralidad de modos de codificación intra, o una de una pluralidad de modos de codificación inter, para el bloque de

vídeo actual, basándose en resultados de errores (por ejemplo, la velocidad de codificación y el nivel de distorsión).

5 **[0075]** La unidad de procesamiento de predicción 41 puede configurarse para implementar las técnicas de esta divulgación para realizar la predicción inter basada en la imagen de referencia identificada por el descodificador de vídeo 30B. La unidad de procesamiento de predicción 41 puede proporcionar el bloque con codificación intra o codificación inter 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.

10 **[0076]** La unidad de procesamiento de predicción intra 46, dentro de la unidad de procesamiento de predicción 41, puede realizar la codificación mediante predicción intra del bloque de vídeo actual con respecto a uno o más bloques contiguos en la misma trama o fragmento que el bloque que va a codificarse, para proporcionar compresión espacial. La unidad de estimación del movimiento 42 y la unidad de compensación del movimiento 44 de la unidad de procesamiento de predicción 41 realizan la codificación mediante predicción inter del bloque de vídeo actual en relación con uno o más bloques predictivos de una o más imágenes de referencia, para proporcionar una compresión temporal.

15 **[0077]** La unidad de estimación de movimiento 42 puede estar configurada para determinar el modo de predicción inter para un fragmento de vídeo de acuerdo con un patrón por defecto para una secuencia de vídeo. El patrón predeterminado puede designar fragmentos de vídeo de la secuencia como fragmentos P, fragmentos B o fragmentos GPB. La unidad 42 de estimación de movimiento y la unidad 44 de compensación de movimiento pueden estar sumamente integradas, pero se ilustran por separado con fines conceptuales. La estimación de movimiento, realizada por la unidad de estimación de movimiento 42, es el proceso de generación de vectores de movimiento, que estiman el movimiento de los bloques de vídeo. Un vector de movimiento, por ejemplo, puede indicar el desplazamiento de una PU de un bloque de vídeo de una trama o imagen de vídeo actual a un bloque predictivo de una imagen de referencia.

20 **[0078]** Un bloque predictivo es un bloque que se observa que corresponde estrechamente con la PU del bloque de vídeo que se va a codificar en términos de diferencia de píxel, que puede determinarse mediante una suma de una diferencia absoluta (SAD), una suma de diferencia al cuadrado (SSD) u otras métricas de diferencia. En algunos ejemplos, el codificador de vídeo 20A puede calcular valores para posiciones de píxeles sub-enteras de imágenes de referencia almacenadas en la DPB 64. Por ejemplo, el codificador de vídeo 20A puede interpolar valores de posiciones de un cuarto de píxel, posiciones de un octavo de píxel u otras posiciones de píxel fraccionario de la imagen de referencia. Por lo tanto, la unidad de estimación de movimiento 42 puede realizar una búsqueda de movimiento en relación con las posiciones de píxeles completas y las posiciones de píxeles fraccionarias, y generar un vector de movimiento con una precisión de píxel fraccionaria.

25 **[0079]** La unidad de estimación de movimiento 42 calcula un vector de movimiento para una PU de un bloque de vídeo en un fragmento con codificación inter comparando la posición de la PU con la posición de un bloque predictivo de una imagen de referencia. La imagen de referencia se puede seleccionar de una primera lista de imágenes de referencia (RefPicList0) o una segunda lista de imágenes de referencia (RefPicList1), cada una de las cuales identifica una o más imágenes de referencia almacenadas en la DPB 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.

30 **[0080]** La compensación de movimiento, realizada por la unidad de compensación de movimiento 44, puede implicar extraer o generar el bloque predictivo basándose en el vector de movimiento determinado por la estimación de movimiento, realizando posiblemente interpolaciones hasta la precisión de subpíxel. 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 codificador de vídeo 20A forma un bloque de vídeo residual restando los valores de píxel del bloque predictivo de los valores de píxel del bloque de vídeo actual que se está codificando, generando valores de diferencia de píxel. Los valores de diferencia de píxel forman datos residuales para el bloque, y pueden incluir componentes de diferencia de luma y croma. El sumador 50 representa el componente o los componentes que realizan esta operación de resta. La unidad de compensación de movimiento 44 también puede generar elementos sintácticos asociados a los bloques de vídeo y al fragmento de vídeo para su uso mediante el descodificador de vídeo 30B en la descodificación de los bloques de vídeo del fragmento de vídeo.

35 **[0081]** La unidad de procesamiento de predicción intra 46 puede realizar predicción intra de un bloque actual, como alternativa a la predicción inter llevada a cabo por la unidad de estimación de movimiento 42 y la unidad de compensación de movimiento 44, como se ha descrito anteriormente. En particular, la unidad de procesamiento de predicción intra 46 puede determinar un modo de predicción intra a usar para codificar un bloque actual. En algunos ejemplos, la unidad de procesamiento de predicción intra 46 puede codificar un bloque actual usando varios modos de predicción intra, por ejemplo, durante diferentes pases de codificación, y la unidad de procesamiento de predicción intra 46 (o la unidad de procesamiento de predicción 41, en algunos ejemplos) puede seleccionar un modo adecuado de predicción intra para utilizar a partir de los modos probados. Por ejemplo, la unidad de procesamiento de predicción intra 46 puede calcular valores de velocidad-distorsión usando un análisis de velocidad-distorsión para los diversos

modos de predicción intra probados, y seleccionar el modo de predicción intra que tenga las mejores características de velocidad-distorsión entre los modos probados. El análisis de velocidad-distorsión determina, en general, una cantidad de distorsión (o error) entre un bloque codificado y un bloque original no codificado que se ha codificado para generar el bloque codificado, así como una velocidad de bits (es decir, un número de bits) usada para generar el bloque codificado. La unidad de procesamiento de predicción intra 46 puede calcular proporciones a partir de las distorsiones y velocidades para los diversos bloques codificados, para determinar qué modo de predicción intra presenta el mejor valor de velocidad-distorsión para el bloque.

[0082] En cualquier caso, tras seleccionar un modo de predicción intra para un bloque, la unidad de procesamiento de predicción intra 46 puede proporcionar información que indica el modo de predicción intra 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 el modo de predicción intra seleccionado de acuerdo con las técnicas de esta divulgación. El codificador de vídeo 20A puede incluir datos de configuración en el flujo de bits transmitido, que pueden incluir una pluralidad de tablas de índices de modos de predicción intra y una pluralidad de tablas de índices de modos de predicción intra modificadas (también denominadas tablas de correlación de palabras de código), definiciones de contextos de codificación para varios bloques e indicaciones del modo de predicción intra más probable, una tabla de índices de modos de predicción intra y una tabla de índices de modos de predicción intra modificadas a utilizar para cada uno de los contextos.

[0083] Después de que la unidad de procesamiento de predicción 41 genere el bloque predictivo para el bloque de vídeo actual, ya sea mediante la predicción inter o la predicción intra, el codificador de vídeo 20A forma un bloque de vídeo residual restando el bloque predictivo del bloque de vídeo actual. Los datos de vídeo residuales del bloque residual pueden incluirse en una o más TU y aplicarse a la unidad de procesamiento de transformada 52. La unidad de procesamiento de transformada 52 transforma los datos de vídeo residuales en coeficientes de transformada residuales mediante una transformada, tal como una transformada de coseno discreta (DCT) o una transformada similar desde un punto de vista conceptual. La unidad de procesamiento de transformada 52 puede convertir los datos de vídeo residuales de un dominio de píxeles a un dominio de las transformadas, tal como un dominio de frecuencia.

[0084] La unidad de procesamiento de transformadas 52 puede enviar los coeficientes de transformada resultantes a la unidad de cuantificación 54. La unidad de cuantificación 54 cuantifica los coeficientes de transformada para reducir adicionalmente la velocidad de transmisión de bits. El proceso de cuantificación puede reducir la profundidad de bits asociada a algunos, o a la totalidad, de los coeficientes. El grado de cuantificación puede modificarse ajustando un parámetro de cuantificación. En algunos ejemplos, la unidad de cuantificación 54 puede realizar, a continuación, un recorrido de la matriz que incluye los coeficientes de transformada cuantificados. De forma alternativa, la unidad de codificación por entropía 56 puede llevar a cabo el recorrido.

[0085] Tras la cuantificación, la unidad de codificación por entropía 56 codifica por entropía los coeficientes de transformada cuantificados. Por ejemplo, la unidad de codificación por entropía 56 puede realizar una codificación de longitud variable adaptativa según el contexto (CAVLC), una codificación aritmética binaria adaptativa según el contexto (CABAC), una codificación aritmética binaria adaptativa según el contexto basándose en la sintaxis (SBAC), una codificación por entropía por división de intervalos de probabilidad (PIPE) u otros procedimientos o técnicas de codificación por entropía. Tras la codificación por entropía realizada por la unidad de codificación por entropía 56, el flujo de bits codificado puede transmitirse al descodificador de vídeo 30B, o archivarse para su posterior transmisión o recuperación mediante el descodificador de vídeo 30B. La unidad de codificación por entropía 56 también puede realizar la codificación por entropía de los vectores de movimiento y los otros elementos sintácticos para el fragmento de vídeo actual que se está codificando.

[0086] La unidad de cuantificación inversa 58 y la unidad de procesamiento de transformada inversa 60 aplican una cuantificación inversa y una transformada inversa, respectivamente, para reconstruir el bloque residual en el dominio de píxel, para su posterior uso como un bloque de referencia de una imagen de referencia. La unidad de compensación de movimiento 44 puede calcular un bloque de referencia sumando el bloque residual a un bloque predictivo de una de las imágenes de referencia de una de las listas de imágenes de referencia. La unidad 44 de compensación de movimiento también puede aplicar uno o más filtros de interpolación al bloque residual reconstruido para calcular valores de píxeles de subentero para su uso en la estimación de movimiento. El sumador 62 agrega el bloque residual reconstruido al bloque de predicción compensado por movimiento producido por la unidad de compensación de movimiento 44 para producir un bloque de referencia para almacenamiento en la DPB 64. La unidad de estimación de movimiento 42 y la unidad de compensación de movimiento 44 pueden usar el bloque de referencia como un bloque de referencia para realizar la predicción inter de un bloque en una imagen o trama de vídeo subsiguiente.

[0087] Como se describe con más detalle a continuación, el codificador de vídeo 20A puede configurarse para recibir información, desde el descodificador de vídeo 30B, indicativa de imágenes de referencia disponibles para la descodificación en el descodificador de vídeo 30B. El codificador de vídeo 20A puede codificar mediante predicción inter una imagen actual basada en una o más de las imágenes de referencia disponibles para descodificar en el descodificador de vídeo 30B. En algunos ejemplos, el codificador de vídeo 20A puede recibir información indicativa de imágenes de referencia disponibles para descodificar en el descodificador de vídeo 30B si se produjo un error en la reconstrucción de una imagen de referencia mediante el descodificador de vídeo 30B. En algunos ejemplos, para

recibir información indicativa de imágenes de referencia disponibles para descodificación en el descodificador de vídeo 30B, el codificador de vídeo 20A puede recibir información que identifica una imagen identificada en un conjunto de imágenes de referencia. El codificador de vídeo 20A puede configurarse para determinar las imágenes de referencia incluidas en el conjunto de imágenes de referencia como las imágenes de referencia disponibles para descodificar en el descodificador de vídeo 30B.

[0088] De esta manera, el codificador de vídeo 20A puede configurarse para recibir, desde el dispositivo 14 que incluye el descodificador de vídeo 30B, información que identifica una imagen previamente descodificada que se almacena en la DPB 92 (FIG. 3) del descodificador de vídeo 30B en el momento en que el descodificador de vídeo 30B transmitió la información. El codificador de vídeo 20A puede determinar una o más imágenes de referencia candidatas que están disponibles para codificar una imagen actual basándose en la imagen identificada. La una o más imágenes de referencia candidatas incluyen una o más imágenes que previamente se determinó que eran utilizables para la codificación mediante predicción inter de una o más de las imágenes o imágenes identificadas que siguen a la imagen identificada en orden de codificación (por ejemplo, basándose en el conjunto de parámetros para la imagen identificada que incluye uno o más de los subconjuntos de imágenes de referencia que forman un conjunto de imágenes de referencia), y que aún se almacenan en la DPB 64 del codificador de vídeo 20A. El codificador de vídeo 20A (por ejemplo, a través de la unidad de procesamiento de predicción 41) puede seleccionar una o más imágenes de referencia para la imagen actual de la una o más imágenes de referencia candidatas determinadas y la imagen identificada.

[0089] El codificador de vídeo 20A puede codificar mediante predicción inter de la imagen actual basándose en una o más imágenes de referencia seleccionadas. Por ejemplo, el codificador de vídeo 20A puede seleccionar una imagen de las imágenes de referencia candidatas determinadas y la imagen identificada, y realizar una predicción unidireccional para realizar predicción inter de la imagen actual. En algunos ejemplos, el codificador de vídeo 20A puede seleccionar dos imágenes para la imagen actual a partir de las imágenes de referencia candidatas determinadas y la imagen identificada, y codificar mediante predicción bidireccional la imagen actual (por ejemplo, usar dos imágenes de referencia, una de cada lista de imágenes de referencia, para codificar mediante predicción inter de la imagen actual).

[0090] En algunos casos, el codificador de vídeo 20A puede determinar si la imagen identificada está disponible para codificar la imagen actual (por ejemplo, almacenada en DPB 64). El codificador de vídeo 20A puede codificar mediante predicción intra, en lugar de codificar mediante predicción inter, la imagen actual basándose en la determinación de que la imagen identificada no está disponible para codificar la imagen actual.

[0091] El codificador de vídeo 20A también puede recibir información que identifica una imagen perdida. El codificador de vídeo 20A puede excluir la imagen perdida de una o más imágenes de referencia candidatas determinadas.

[0092] La FIG. 3 es un diagrama de bloques que ilustra un descodificador de vídeo de ejemplo 30B que puede implementar las técnicas descritas en esta divulgación. Como se describió anteriormente, el descodificador de vídeo 30A puede ser similar al descodificador de vídeo 30B; sin embargo, las técnicas descritas en esta divulgación no están tan limitadas. En el ejemplo de la FIG. 3, el descodificador de vídeo 30B incluye una unidad de descodificación por entropía 80, unidad de procesamiento de predicción 81, unidad de cuantificación inversa 86, unidad de transformada inversa 88, sumador 90 y memoria intermedia de imágenes descodificadas (DPB) 92. La unidad de procesamiento de predicción 81 incluye la unidad de compensación de movimiento 82 y la unidad de procesamiento de predicción intra 84. En algunos ejemplos, el descodificador de vídeo 30B puede realizar una pasada de descodificación que en general es recíproca a la pasada de codificación descrita con respecto al codificador de vídeo 20 de la FIG. 2.

[0093] En el ejemplo de la FIG. 3, la memoria de datos de vídeo 79 recibe vídeo codificado. La memoria de datos de vídeo 79 puede almacenar datos de vídeo (por ejemplo, configurados para almacenar datos de vídeo), como un flujo de bits de vídeo codificado, para ser descodificados por los componentes del descodificador de vídeo 30B. La memoria de datos de vídeo 79 también puede almacenar datos de vídeo que son utilizados por los componentes del descodificador de vídeo 30B para descodificar los datos de vídeo y reconstruir una imagen.

[0094] Los datos de vídeo almacenados en la memoria de datos de vídeo 79 pueden obtenerse de una fuente de vídeo local, como una cámara, a través de la comunicación de datos de vídeo a través de una red alámbrica o inalámbrica, o al acceder a medios físicos de almacenamiento de datos. La memoria de datos de vídeo 79 puede formar una memoria intermedia de imágenes codificadas (CPB) que almacena datos de vídeo codificados a partir de un flujo de bits de vídeo codificado.

[0095] DPB 92 es un ejemplo de memoria que almacena datos de vídeo de referencia para su uso en la descodificación de datos de vídeo mediante el descodificador de vídeo 30B (por ejemplo, en el modo de codificación intra y el modo de codificación inter). La memoria de datos de vídeo 79 y la DPB 92 pueden estar formadas por cualquiera entre varios dispositivos de memoria, tales como memoria dinámica de acceso aleatorio (DRAM), incluyendo DRAM síncrona (SDRAM), RAM magnetorresistiva (MRAM), RAM resistiva (RRAM) u otros tipos de dispositivos de memoria. La memoria de datos de vídeo 79 y la DPB 92 pueden ser proporcionadas por el mismo

dispositivo de memoria o por dispositivos de memoria independientes. En diversos ejemplos, la memoria de datos de vídeo 79 puede estar en un chip con otros componentes del descodificador de vídeo 30B, o fuera de chip con respecto a esos componentes.

5 **[0096]** Durante el proceso de descodificación, el descodificador de vídeo 30B 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 20A. La unidad de descodificación por entropía 80 del descodificador de vídeo 30B realiza la descodificación por entropía del flujo de bits para generar coeficientes cuantificados, vectores de movimiento y otros elementos sintácticos. La unidad de descodificación por entropía 80 remite los vectores de movimiento y otros elementos sintácticos a la unidad de procesamiento de predicción 81. El descodificador de vídeo 30B puede recibir los elementos sintácticos en el nivel del fragmento de vídeo y/o el nivel del bloque de vídeo.

15 **[0097]** Cuando el fragmento de vídeo se codifica como un fragmento con codificación intra (I), la unidad de procesamiento de predicción intra 84 de la unidad de procesamiento de predicción 81 puede generar datos de predicción para un bloque de vídeo del fragmento de vídeo actual, basándose en un modo de predicción intra señalado, y datos de bloques previamente descodificados de la trama o imagen actual. La unidad de procesamiento de predicción 81 puede configurarse para implementar las técnicas de esta divulgación para indicar qué imagen de referencia debe usarse para la predicción inter, por ejemplo, cuando no se recibe una imagen de referencia. Cuando la trama de vídeo es codificada como un fragmento con codificación inter (es decir, B o P), la unidad de compensación de movimiento 20 82 de la unidad de procesamiento de predicción 81 genera bloques predictivos para un bloque de vídeo del fragmento de vídeo actual, basándose en los vectores de movimiento y otros elementos sintácticos recibidos desde la unidad de descodificación por entropía 80. Los bloques predictivos pueden generarse a partir de una de las imágenes de referencia dentro de una de las listas de imágenes de referencia. El descodificador de vídeo 30B puede construir las listas de tramas de referencia, RefPicList0 y RefPicList1, usando técnicas de construcción predeterminadas basadas en imágenes de referencia almacenadas en DPB 92.

30 **[0098]** La unidad de compensación de movimiento 82 determina la información de predicción para un bloque de vídeo del fragmento de vídeo actual, analizando los vectores de movimiento y otros elementos sintácticos, y usa la información de predicción para generar los bloques predictivos del bloque de vídeo actual que se está descodificando. Por ejemplo, la unidad de compensación de movimiento 82 usa algunos de los elementos sintácticos recibidos para determinar un modo de predicción (por ejemplo, predicción intra o predicción inter) usada para codificar los bloques de vídeo del fragmento de vídeo, un tipo de fragmento de predicción inter (por ejemplo, fragmento B, fragmento P o fragmento GPB), información de construcción para una o más de las listas de imágenes de referencia del fragmento, vectores de movimiento para cada bloque de vídeo con codificación inter del fragmento, el estado de predicción inter para cada bloque de vídeo con codificación inter del fragmento y otra información para descodificar los bloques de vídeo en el fragmento de vídeo actual.

40 **[0099]** La unidad de compensación de movimiento 82 también puede realizar la interpolación basándose en filtros de interpolación. La unidad de compensación de movimiento 82 puede usar filtros de interpolación como los usados por el codificador de vídeo 20A durante la codificación de los bloques de vídeo, para calcular valores interpolados para píxeles fraccionarios de bloques de referencia. En este caso, la unidad de compensación de movimiento 82 puede determinar los filtros de interpolación utilizados por el codificador de vídeo 20A a partir de los elementos sintácticos recibidos y utilizar los filtros de interpolación para generar bloques predictivos.

45 **[0100]** La unidad de cuantificación inversa 86 cuantifica de manera inversa, es decir, descuantifica, los coeficientes de transformada cuantificados, proporcionados en el flujo de bits y descodificados por la unidad de descodificación por entropía 80. El proceso de cuantificación inversa puede incluir el uso de un parámetro de cuantificación calculado por el codificador de vídeo 20A para cada bloque de vídeo del fragmento de vídeo, con el fin de determinar un grado de cuantificación y, asimismo, un grado de cuantificación inversa que debería aplicarse. La unidad de procesamiento de transformada inversa 88 aplica una transformada inversa, por ejemplo una DCT inversa, una transformada inversa entera o un proceso de transformada inversa conceptualmente similar, a los coeficientes de transformada, con el fin de generar bloques residuales en el dominio de píxeles.

55 **[0101]** Después de que la unidad de compensación de movimiento 82 genera el bloque predictivo para el bloque de vídeo actual, basándose en los vectores de movimiento y otros elementos sintácticos, el descodificador de vídeo 30B forma un bloque de vídeo descodificado sumando los bloques residuales procedentes de la unidad de procesamiento de transformada inversa 88 a los correspondientes bloques predictivos generados por la unidad de compensación de movimiento 82. El sumador 90 representa el componente o los componentes que realizan esta operación de suma. Si se desea, también puede aplicarse un filtro de eliminación de bloques para filtrar los bloques descodificados con el fin de eliminar distorsiones de efecto pixelado. También pueden utilizarse otros filtros de bucle (ya sea en el bucle de codificación o después del bucle de codificación) para allanar las transiciones de píxeles o mejorar de otro modo la calidad del vídeo. Los bloques de vídeo descodificados en una trama o imagen dada son a continuación almacenados en la DPB 92, que almacena imágenes de referencia usadas para la posterior compensación de movimiento. La DPB 92 almacena también vídeo descodificado para su presentación posterior en un dispositivo de visualización, tal como el dispositivo de visualización 31B de la FIG. 1.

[0102] El codificador de vídeo 20A, el descodificador de vídeo 30B, el codificador de vídeo 20B y el descodificador de vídeo 30A pueden configurarse para implementar las técnicas descritas en esta divulgación. De nuevo, para facilitar la descripción, las técnicas se describen con respecto al codificador de vídeo 20A y el descodificador de vídeo 30B, y el codificador de vídeo 20B y el descodificador de vídeo 30A pueden configurarse para implementar técnicas similares. Sin embargo, las técnicas no deben considerarse limitadas a requerir que el codificador de vídeo 20B y el descodificador de vídeo 30A realicen las mismas técnicas que las del codificador de vídeo 20A y el descodificador de vídeo 30B.

[0103] Por ejemplo, las técnicas descritas anteriormente se describen con respecto a un codificador en un lado emisor y un descodificador en el lado del receptor. El codificador de vídeo 20A puede ser un ejemplo de dicho codificador en el lado del emisor, y el descodificador de vídeo 30B puede ser un ejemplo de dicho descodificador en el lado del receptor. En algunos ejemplos, el descodificador de vídeo 30B puede configurarse para emitir el mensaje de retroalimentación de indicación de selección de conjunto de imágenes de referencia (RPSSI) (por ejemplo, como parte de un mensaje RTP). A partir del mensaje de retroalimentación RPSSI, el codificador de vídeo 20A puede configurarse para determinar qué imágenes de referencia están disponibles en la DPB 92 del descodificador de vídeo 30B y puede codificar una imagen basada en una de las imágenes que se sabe están en la DPB 92 del descodificador de vídeo 30B.

[0104] Como se describió anteriormente, las técnicas descritas en esta divulgación pueden estar relacionadas con aplicaciones de vídeo en tiempo real. Para aplicaciones de vídeo en tiempo real bajo las mejores condiciones de la red, los mensajes de retroalimentación son importantes para informar al remitente (por ejemplo, el codificador de vídeo 20A o el dispositivo 12 que incluye el codificador de vídeo 20A) sobre lo que sucedió en el receptor (por ejemplo, el descodificador de vídeo 30B o el dispositivo 14 que incluye descodificador de vídeo 30B). Debe entenderse que en videotelefonía o videoconferencia, el remitente puede ser el dispositivo 12 o el dispositivo 14, y el receptor puede ser el dispositivo 12 o el dispositivo 14. Para facilitar la descripción, esta divulgación describe al dispositivo 12 como el remitente y al dispositivo 14 como el receptor, entendiendo que el dispositivo 14 también puede funcionar como el remitente, y el dispositivo 12 puede funcionar como el receptor.

[0105] Por ejemplo, en los casos en que el descodificador de vídeo 30B no recibió una imagen del flujo de bits señalizado por el codificador de vídeo 20A, puede ser beneficioso que el descodificador de vídeo 30B transmita un mensaje de retroalimentación para informar al codificador de vídeo 20A que hay una imagen perdida. Para el códec HEVC existente, cuando el descodificador de vídeo 30B pierde una de las imágenes de referencia, el descodificador de vídeo 30B puede usar los mensajes de Indicación de selección de imagen de referencia (RPSI) y de Indicación de pérdida de imagen específica (SPLI), que se definen en "Formato de carga útil RTP para codificación de vídeo de alta eficiencia" por Wang et al. del Grupo de trabajo de la red, con fecha del 6 de septiembre de 2013, para informar al codificador de vídeo 20A que las imágenes se perdieron.

[0106] La información sobre el formato de carga útil RTP para codificación de vídeo de alta eficiencia se puede encontrar, a partir del 2 de diciembre de 2013, en <http://tools.ietf.org/html/draft-ietf-payload-rtp-h265-01>. Por ejemplo, la información sobre el formato de carga útil RTP describe que el formato de carga útil RTP permite la paquetización de una o más unidades de Capa de abstracción de red (NAL) en cada carga útil de paquetes RTP, así como la fragmentación de una unidad NAL en múltiples paquetes RTP. Además, el formato de carga útil RTP soporta la transmisión de un flujo HEVC a través de un único, así como múltiples flujos RTP. El formato de carga útil tiene una amplia aplicabilidad en videoconferencia, transmisión de vídeo por Internet y vídeo con calidad de entretenimiento de alta velocidad de bits, entre otros.

[0107] Puede haber varias formas en que el descodificador de vídeo 30B puede determinar que hubo una imagen perdida. Como ejemplo, el descodificador de vídeo 30B puede recibir elementos sintácticos en el flujo de bits codificado que dan instrucciones al descodificador de vídeo 30B para utilizar la imagen perdida como imagen de referencia para la descodificación. En este caso, el descodificador de vídeo 30B determina que la imagen que se va a usar como imagen de referencia no se almacena en la DPB 92 del descodificador de vídeo 30B porque se perdió la imagen, y el descodificador de vídeo 30B determina que se perdió una imagen en el flujo de bits. Puede haber otras causas por las que el descodificador de vídeo 30B determina que se perdió una imagen y otras formas en que el descodificador de vídeo 30B determina que se perdió una imagen. Las técnicas descritas en esta divulgación no se limitan a ninguna forma particular en la que el descodificador de vídeo 30B determine que se perdió una imagen.

[0108] La siguiente descripción explica la funcionalidad de los mensajes RPSI para el contexto en la comprensión de las técnicas descritas en esta divulgación. Los mensajes SPLI se describen mucho más adelante.

[0109] Para RPSI, cuando ocurre un error, el descodificador de vídeo 30B elige una imagen de referencia descodificada correctamente anterior de la DPB 92, y solicita que el codificador de vídeo 20A use esa imagen, en general una imagen de referencia a largo plazo, para referencia cuando se codifica la siguiente imagen (es decir, imagen actual que se está codificando). Si esta imagen de referencia especificada aún existe en la DPB 64 del codificador de vídeo 20A, el codificador de vídeo 20A puede codificar la siguiente imagen con esta referencia para detener la propagación de errores.

[0110] En RPSI, el descodificador de vídeo 30B puede especificar la imagen de referencia utilizada para la codificación. Sin embargo, la información que indica exactamente todas las imágenes que están almacenadas en la DPB 92 del codificador de vídeo 20A puede no estar disponible para el descodificador de vídeo 30B. En cambio, la información que el descodificador de vídeo 30B puede usar para determinar las imágenes almacenadas en la DPB 64 del codificador de vídeo 20A en el momento en que el descodificador de vídeo 30B transmitió la información que especifica la imagen de referencia puede ser inexacta y estar desactualizada.

[0111] Por consiguiente, en algunos ejemplos, el descodificador de vídeo 30B puede dar como resultado la especificación de una imagen de referencia que el codificador de vídeo 20A debería usar para la predicción parcial, cuyo codificador de vídeo 20A ya se eliminó de la DPB 64. En este caso, el codificador de vídeo 20A puede no ser capaz de usar la imagen identificada que el descodificador de vídeo 30B identificó como una imagen de referencia para la codificación, ya que la imagen identificada se eliminó de la DPB 64. En algunos ejemplos, si la imagen identificada no se encuentra en la DPB 64 del codificador de vídeo 20A, entonces el codificador de vídeo 20A codifica una imagen actual como un punto de acceso aleatorio intra (IRAP) para detener el error. Sin embargo, codificar la imagen actual como IRAP puede ser costoso en términos de eficiencia de codificación.

[0112] Para solucionar el problema de que el descodificador de vídeo 30B identifica una imagen que se usará como imagen de referencia que no está almacenada en la DPB 64 del codificador de vídeo 20A, en algunos ejemplos, el codificador de vídeo 20A y el descodificador de vídeo 30B pueden negociar el uso de imágenes de referencia a largo plazo (LTR) (por ejemplo, mediante algún tipo de comunicación entre el codificador de vídeo 20A y el descodificador de vídeo 30B), que se mantendrán por más tiempo en la DPB 64 del codificador de vídeo 20A y la DPB 92 del descodificador de vídeo 30B. En algunos ejemplos, en lugar de negociar el uso de imágenes LTR, el descodificador de vídeo 30B puede preconfigurarse para identificar una imagen LTR como la imagen de referencia que debe usar el codificador de vídeo 20A. En cualquiera de los dos ejemplos, el descodificador de vídeo 30B puede especificar una imagen LTR que se almacena actualmente en la DPB 92 del descodificador de vídeo 30B como la imagen que se va a usar para codificar la imagen actual a punto de ser codificada por el codificador de vídeo 20A. Sin embargo, confiar en imágenes LTR puede afectar negativamente a la eficiencia de la codificación porque puede haber una menor correlación temporal, ya que las imágenes LTR pueden estar más distanciadas temporalmente de la imagen actual que se codifica en comparación con otras imágenes (por ejemplo, imágenes de referencia a corto plazo). En otras palabras, en algunos casos (aunque no es un requisito), las imágenes de referencia a largo plazo tienden a almacenarse durante más tiempo en DPB 64 y DPB 92, en comparación con las imágenes de referencia a corto plazo, y por lo tanto, pueden estar más distanciadas temporalmente de la imagen actual en comparación con una imagen de referencia a corto plazo.

[0113] Puede haber otros problemas con las técnicas existentes del RPSI. Por ejemplo, incluso si el descodificador de vídeo 30B no identifica una imagen LTR e identifica una imagen que se encuentra en la DPB 64 del codificador de vídeo 20A, en algunas otras técnicas, se requiere que el codificador de vídeo use la imagen identificada mediante el descodificador de vídeo 30B para realizar predicción inter de la imagen actual (es decir, la imagen que está a punto de ser codificada). El codificador de vídeo 20A usa la imagen identificada, asumiendo que está en la DPB 64 del codificador de vídeo 20A, porque se garantiza que la imagen identificada existe en la DPB 92 del descodificador de vídeo 30B ya que el descodificador de vídeo 30B seleccionó la imagen de la DPB 92. Sin embargo, puede haber otras imágenes además de la imagen que el descodificador de vídeo 30B identificó que son más adecuadas para la predicción inter de la imagen actual. Además, el hecho de estar limitado a usar solo la imagen identificada mediante el descodificador de vídeo 30B hace que el codificador de vídeo 20A solo utilice la predicción única ya que la predicción bidireccional utiliza dos imágenes. Esto tiene un impacto negativo adicional en la eficiencia de la codificación mediante predicción inter.

[0114] Las técnicas descritas en esta divulgación pueden abordar uno o más de los problemas descritos anteriormente con el RPSI. Por ejemplo, en las técnicas descritas en esta divulgación, el descodificador de vídeo 30B puede generar información que identifica una imagen previamente descodificada en DPB 92; sin embargo, el codificador de vídeo 20A puede no estar limitado a usar solo la imagen identificada para realizar predicción inter de la imagen actual.

[0115] En lugar de eso, el codificador de vídeo 20A puede determinar una o más imágenes que están disponibles tanto en la DPB 64 del codificador de vídeo 20A como en la DPB 92 del descodificador de vídeo 30B de la imagen identificada. El codificador de vídeo 20A puede seleccionar una o más de las imágenes determinadas y potencialmente la imagen identificada, también, para realizar predicción inter de la imagen actual. De esta manera, el codificador de vídeo 20A puede configurarse para seleccionar entre un gran conjunto de imágenes, aparte de una sola imagen, y determinar qué imágenes proporcionan una mejor eficiencia de codificación y seleccionar esas imágenes para realizar predicción inter de la imagen actual.

[0116] Además, el codificador de vídeo 20A puede configurarse para determinar las imágenes que están disponibles en las respectivas DPB (por ejemplo, DPB 64 y DPB 92) tanto del codificador de vídeo 20A como del descodificador de vídeo 30B con señal mínima del descodificador de vídeo 30B. Por ejemplo, con la señalización de solo unas pocas imágenes (por ejemplo, una o unas pocas imágenes seleccionadas), el codificador de vídeo 20A puede configurarse para determinar más imágenes (por ejemplo, una pluralidad de imágenes) que están disponibles en ambos DPB 64

del codificador de vídeo 20A y DPB 92 del descodificador de vídeo 30B. Con fines de ilustración, esta divulgación describe el descodificador de vídeo 30B que identifica una imagen, pero la divulgación no debe considerarse tan limitante.

5 **[0117]** En las técnicas descritas en esta divulgación, el codificador de vídeo 20A puede tomar una decisión más óptima para la recuperación de errores (por ejemplo, para el caso de ejemplo donde hay una imagen perdida) una vez que el codificador de vídeo 20A determina las imágenes de referencia en la DPB 92 del descodificador de vídeo 30B. Por ejemplo, el descodificador de vídeo 30B debe informar al codificador de vídeo 20A de todas las imágenes de referencia en la DPB 92 del descodificador de vídeo 30B. En algunos ejemplos, el codificador de vídeo 20A puede
10 determinar las imágenes de referencia exactas en la DPB 92 del descodificador de vídeo 30B y puede determinar un modo de recuperación de error (por ejemplo, qué imágenes usar para realizar predicción inter de la imagen actual) basándose en las imágenes de referencia determinadas en DPB 92 del descodificador de vídeo 30B.

15 **[0118]** Para que el codificador de vídeo 20A determine las imágenes de referencia exactas en la DPB 92 del descodificador de vídeo 30B, el descodificador de vídeo 30B puede señalar todas las imágenes de referencia en la DPB 92 del descodificador de vídeo 30B al codificador de vídeo 20A. Sin embargo, dicha señalización puede ser ineficiente en el ancho de banda y, en las técnicas descritas en esta divulgación, el descodificador de vídeo 30B puede señalar una imagen desde la cual el codificador de vídeo 20A puede determinar una pluralidad de imágenes disponibles para codificar o descodificar la imagen actual.

20 **[0119]** Un ejemplo de manera en que el codificador de vídeo 20A puede determinar qué imagen almacenada en la DPB 64 también se almacena en la DPB 92 del descodificador de vídeo 30B es utilizando los subconjuntos de imágenes de referencia definidos en la norma HEVC. Las imágenes almacenadas en la DPB 64 del codificador de vídeo 20A que también están determinadas para ser almacenadas en la DPB 92 del descodificador de vídeo 30B se
25 conocen como imágenes disponibles para codificación de vídeo o descodificación de vídeo. De nuevo, debe entenderse que las imágenes que se determinó que están disponibles para la codificación o descodificación de vídeo no necesitan incluir todas las imágenes en las respectivas DPB. En lugar de eso, el codificador de vídeo 20A puede implementar ciertos procedimientos para determinar qué imágenes se almacenan en las DPB tanto del codificador de vídeo 20A como del descodificador de vídeo 30B (por ejemplo, tanto DPB 64 como DPB 92).

30 **[0120]** La norma HEVC define cinco subconjuntos de imágenes de referencia que juntos forman un conjunto de imágenes de referencia. En la norma HEVC, cada uno del codificador de vídeo 20A y el descodificador de vídeo 30B pueden construir los subconjuntos de imágenes de referencia para cada imagen. Por ejemplo, para codificar o descodificar una imagen, la unidad de procesamiento de predicción 41 del codificador de vídeo 20A y la unidad de
35 procesamiento de predicción 81 del descodificador de vídeo 30B pueden construir listas de imágenes de referencia (por ejemplo, RefPicList0 y RefPicList1) basadas en los subconjuntos de imágenes de referencia construidos para esa imagen, y codificar o descodificar mediante predicción inter esa imagen basándose en las listas de imágenes de referencia construidas.

40 **[0121]** Los subconjuntos y conjuntos de imágenes de referencia, para la norma HEVC, se describen con más detalle en la SERIE H: SISTEMAS AUDIOVISUALES Y MULTIMEDIA, Infraestructura de servicios audiovisuales - Codificación de vídeo en movimiento, recomendación ITU-T H.265 de codificación de vídeo de alta eficiencia, con fecha 4 de abril de 2013 y denominada en el presente documento como "Codificación de vídeo de alta eficiencia, Rec. ITU-T H.265 e ISO/IEC 23008-2 ", y se pueden encontrar en <http://www.itu.int/rec/T-RECH.265-201304-1>.
45

[0122] Por ejemplo, para codificar o descodificar una imagen en particular, el codificador de vídeo 20A y el descodificador de vídeo 30B pueden construir los cinco subconjuntos de imágenes de referencia, que incluyen:

50 RefPicSetStCurrBefore, RefPicSetStCurrAfter, RefPicSetStFoll, RefPicSetLtCurr y RefPicSetLtFoll. RefPicSetStCurrBefore, RefPicSetStCurrAfter y RefPicSetLtCurr incluyen todas las imágenes que pueden usarse para descodificar la imagen en particular. En algunos ejemplos, RefPicSetStCurrBefore puede incluir imágenes de referencia determinadas como imágenes de referencia a corto plazo que se muestran antes de la imagen en particular, y

55 RefPicSetStCurrAfter puede incluir cualquier imagen de referencia determinada como imágenes de referencia a corto plazo que se muestran después de la imagen en particular. RefPicSetLtCurr puede incluir imágenes de referencia a largo plazo. RefPicSetStFoll y RefPicSetLtFoll incluyen todas las imágenes de referencia que no se utilizan para codificar o descodificar la imagen en particular, pero se pueden usar para las imágenes que siguen a la imagen en particular en orden de descodificación. RefPicSetStFoll puede incluir cualquier imagen de referencia determinada como imágenes de referencia a corto plazo, y RefPicSetLtFoll puede incluir cualquier imagen de
60 referencia determinada como imágenes de referencia a largo plazo. En algunos ejemplos, las imágenes en los conjuntos pueden ser exclusivas (por ejemplo, una imagen en uno de los conjuntos puede no estar en cualquier otro conjunto).

65 **[0123]** En algunos ejemplos, la construcción de los cinco conjuntos puede ser explícita. Por ejemplo, la mera existencia de una imagen en la DPB no debe considerarse como una imagen de referencia que pertenece a uno de

estos cinco conjuntos. En lugar de eso, una imagen identificada en uno de estos conjuntos es una imagen determinada específicamente como una imagen que se puede usar para codificar o descodificar la imagen particular y las imágenes que siguen a la imagen particular en orden de descodificación, o solo las imágenes que siguen a la imagen particular en orden de descodificación.

[0124] En algunos ejemplos, el codificador de vídeo 20A y el descodificador de vídeo 30B pueden construir una(s) lista(s) de imágenes de referencia basadas en el conjunto de imágenes de referencia. Los índices en la(s) lista(s) de imágenes de referencia pueden identificar la(s) imagen(es) de referencia que se usan para codificar mediante predicción inter o descodificar mediante predicción inter de la imagen particular para la cual el codificador de vídeo 20A y el descodificador de vídeo 30B construyeron el conjunto de imágenes de referencia. Por lo tanto, se determina específicamente que las imágenes de referencia en los cinco subconjuntos de ejemplo son imágenes que pertenecen a estos subconjuntos porque es a partir de estas imágenes que se forman las listas de imágenes de referencia.

[0125] Una imagen determinada como una imagen de referencia a corto plazo o una imagen de referencia a largo plazo no significa que dicha imagen pertenezca a uno de los subconjuntos de imágenes de referencia de ejemplo. Nuevamente, los subconjuntos de imágenes de referencia incluyen imágenes específicamente determinadas a pertenecer a esos subconjuntos de imágenes de referencia, que especifican características en cuanto a su facilidad de uso para la codificación o descodificación mediante predicción inter de la imagen o imágenes en particular que siguen la imagen en particular en orden de descodificación.

[0126] En algunos ejemplos, las técnicas descritas en esta divulgación utilizan los subconjuntos de imágenes de referencia para determinar cuáles de los almacenados en la DPB 64 del codificador de vídeo 20A también se almacenan en la DPB 92 del descodificador de vídeo 30B. Como se describió anteriormente, en el mensaje RPSI, el descodificador de vídeo 30B transmite información que identifica una imagen, y se requiere que el codificador de vídeo 20A use la imagen identificada para realizar predicción inter de la imagen actual, suponiendo que la imagen identificada se almacena en la DPB 64 del codificador de vídeo 20A en el momento en que el codificador de vídeo 20A está realizando predicción inter de la imagen actual. Para evitar confusiones, esta divulgación utiliza el término mensaje de respuesta de Indicación de selección de conjunto de imágenes de referencia (RPSSI), para indicar que el codificador de vídeo 20A está utilizando las técnicas descritas en esta divulgación, que están separadas de la RPSI. Sin embargo, el uso del término "RPSSI" es simplemente para evitar confusiones. Puede ser posible redefinir el término "RPSI" para incluir las técnicas descritas en esta divulgación.

[0127] En algunos ejemplos, el codificador de vídeo 20A puede recibir información que identifica una imagen del descodificador de vídeo 30B. El codificador de vídeo 20A puede determinar qué imágenes de uno o más subconjuntos de imágenes de referencia de la imagen identificada aún se almacenan en la DPB 64. Por ejemplo, el codificador de vídeo 20A puede almacenar información que indica los subconjuntos de imágenes de referencia contruidos para cada imagen que codifica el codificador de vídeo 20A. El codificador de vídeo 20A puede recuperar la información que indica las imágenes que podrían haberse utilizado para realizar predicción inter de la imagen identificada a partir de la información almacenada, y determinar qué imágenes que podrían haberse utilizado para realizar predicción inter de la imagen identificada todavía se almacenan en la DPB 64.

[0128] En general, si una imagen, que podría haberse utilizado para realizar predicción inter de la imagen identificada y las imágenes que siguen a la imagen identificada en el orden de codificación, todavía se almacena en la DPB 64 del codificador de vídeo 20A, existe una probabilidad muy alta (prácticamente garantizada) de que la imagen también se almacenará en la DPB 92 del descodificador de vídeo 30B. Tal como se utiliza en esta divulgación, una imagen que podría haberse utilizado para realizar predicción inter de la imagen identificada y las imágenes que siguen a la imagen identificada en orden de codificación significa una imagen que podría haber sido una imagen de referencia de la imagen identificada y las imágenes que siguen a la imagen identificada en orden de codificación. Además, una imagen que podría haberse utilizado para realizar predicción inter de la imagen identificada y las imágenes que siguen a la imagen identificada en orden de codificación incluye cualquier imagen que se haya utilizado realmente para realizar predicción inter de la imagen identificada.

[0129] En algunos ejemplos, el descodificador de vídeo 30B puede borrar una imagen de la DPB 92 solo después de que el codificador de vídeo 20A borra la imagen de la DPB 64. Por lo tanto, si el codificador de vídeo 20A no borró una imagen que pudo haber sido una imagen de referencia para la imagen identificada (por ejemplo, se usó para realizar predicción inter de la imagen identificada), existe una probabilidad muy alta de que el descodificador de vídeo 30B no haya borrado esa imagen de la DPB 92.

[0130] Debe entenderse que no hay garantía de que las DPB del codificador de vídeo 20A y del descodificador de vídeo 30B sean exactamente iguales (por ejemplo, las imágenes almacenadas en DPB 64 y DPB 92 deben ser exactamente el mismo conjunto de imágenes). Sin embargo, si una imagen que podría haberse utilizado para realizar predicción inter de la imagen identificada se almacena en la DPB 64 del codificador de vídeo 20A, entonces esa imagen existe en la DPB 92 del descodificador de vídeo 30B, incluso si hay imágenes en la DPB 64 del codificador de vídeo 20A que no están en la DPB 92 del descodificador de vídeo 30B e incluso si hay imágenes en la DPB 92 del descodificador de vídeo 30B que no están en la DPB 64 del codificador de vídeo 20A.

[0131] Una forma en que el codificador de vídeo 20A determina qué imágenes se podrían haber utilizado para realizar predicción inter de la imagen identificada es a través de los subconjuntos de imágenes de referencia de la imagen identificada. Por ejemplo, el codificador de vídeo 20A puede determinar si las imágenes en los subconjuntos de imágenes de referencia de RefPicSetStCurrBefore, RefPicSetStCurrAfter, RefPicSetLtCurr, RefPicSetLtFoll y RefPicSetStFoll de la imagen identificada están todavía disponibles en la DPB 64 del codificador de vídeo 20A (por ejemplo, basándose en información almacenada de la imagen identificada durante el tiempo que el codificador de vídeo 20A codificó la imagen identificada). Como se describió anteriormente, los subconjuntos RefPicSetStCurrBefore, RefPicSetStCurrAfter y RefPicSetLtCurr incluyen imágenes que podrían haberse usado para realizar predicción inter de la imagen identificada y las imágenes que siguen a la imagen identificada en orden de codificación. Los subconjuntos RefPicSetLtFoll y RefPicSetStFoll incluyen imágenes que podrían haberse utilizado para realizar predicción inter de imágenes que siguen a la imagen identificada en orden de codificación, pero no de la imagen identificada.

[0132] Las imágenes en uno o más de los subconjuntos de imágenes de referencia de la imagen identificada que se almacenan en la DPB 64 del codificador de vídeo 20A se conocen como imágenes disponibles para codificación (o imágenes de referencia candidatas). En algunos ejemplos, el codificador de vídeo 20A puede determinar las imágenes que se encuentran en cualquiera de los subconjuntos de imágenes de referencia de la imagen identificada como imágenes de referencia candidatas que se pueden usar para realizar predicción inter de la imagen actual (es decir, que están disponibles para realizar predicción inter de la imagen actual o están disponibles como imágenes de referencia para la imagen actual). En algunos ejemplos, el codificador de vídeo 20A puede determinar adicionalmente las imágenes que están en cualquiera de los subconjuntos de imágenes de referencia de las imágenes en los subconjuntos de imágenes de referencia de la imagen identificada que se almacenan en la DPB 64 del codificador de vídeo 20A.

[0133] En algunos ejemplos, el codificador de vídeo 20A puede determinar las imágenes que están solo en uno de los subconjuntos RefPicSetStCurrBefore, RefPicSetStCurrAfter y RefPicSetLtCurr de la imagen identificada como imágenes de referencia candidatas que están disponibles para realizar predicción inter de la imagen actual (es decir, que están disponibles como imágenes de referencia para la imagen actual). Por ejemplo, en algunos ejemplos, debido a la extracción del flujo de bits secundario que puede ocurrir antes de que el descodificador de vídeo 30B reciba el flujo de bits, las imágenes en el subconjunto RefPicSetStFoll y RefPicSetLtFoll (es decir, las imágenes que no podrían haberse usado para realizar predicción inter de las imágenes identificadas, y solo las imágenes que siguen a la imagen identificada) pueden eliminarse del flujo de bits.

[0134] En este caso especial, la DPB 64 del codificador de vídeo 20A puede incluir imágenes del conjunto de imágenes de referencia de la imagen identificada que no están disponibles en la DPB 92 del descodificador de vídeo 30B. Debido a que el codificador de vídeo 20A puede no ser capaz de determinar si la extracción de sub-flujo de bits está habilitada, para evitar problemas donde las imágenes de referencia no están disponibles, el codificador de vídeo 20A puede determinar las imágenes que se pueden usar para realizar predicción inter de la imagen identificada, y no solo las imágenes que siguen a la imagen identificada, como imágenes de referencia candidatas que se pueden usar para realizar predicción inter de la imagen actual. Nuevamente, los subconjuntos RefPicSetStCurrBefore, RefPicSetStCurrAfter y RefPicSetLtCurr de la imagen identificada incluyen imágenes que podrían haberse utilizado para realizar predicción inter de la imagen identificada, y por lo tanto, el codificador de vídeo 20A puede determinar las imágenes que están disponibles para realizar predicción inter de la imagen actual solo desde las imágenes en los subconjuntos RefPicSetStCurrBefore, RefPicSetStCurrAfter y RefPicSetLtCurr de la imagen identificada que todavía están disponibles en la DPB 64 del codificador de vídeo 20A.

[0135] En algunos ejemplos, el codificador de vídeo 20A puede determinar las imágenes que se encuentran en los subconjuntos de imágenes de referencia a corto plazo que podrían haberse utilizado para realizar predicción inter de la imagen identificada (es decir, RefPicSetStCurrBefore y RefPicSetStCurrAfter). Esto puede deberse a que la correlación temporal de las imágenes de referencia a largo plazo es relativamente deficiente, y para reducir el tiempo de procesamiento, puede tener poco o ningún beneficio si el codificador de vídeo 20A evalúa las imágenes de referencia a largo plazo para fines de predicción inter.

[0136] Para reducir aún más el tiempo de procesamiento, en algunos ejemplos, el codificador de vídeo 20A puede determinar imágenes que son imágenes de referencia a corto plazo que podrían haberse utilizado para realizar predicción inter de la imagen identificada que se muestran antes que la imagen identificada (es decir, las imágenes en el subconjunto RefPicSetStCurrBefore) como imágenes de referencia candidatas que se pueden utilizar para realizar predicción inter de la imagen actual. Como otro ejemplo, el codificador de vídeo 20A puede determinar las imágenes que son imágenes de referencia a corto plazo que podrían haberse utilizado para realizar predicción inter de la imagen identificada que se muestran después de la imagen identificada (es decir, las imágenes en el subconjunto RefPicSetStCurrFoll) como imágenes de referencia candidatas que se pueden utilizar para realizar predicción inter de la imagen actual.

[0137] Después de que el codificador de vídeo 20A determina qué imágenes están disponibles para realizar predicción inter de la imagen actual basándose en las imágenes de referencia candidatas que podrían haberse utilizado para realizar predicción inter de la imagen identificada y las imágenes que siguen a la imagen identificada, el

codificador de vídeo 20A puede seleccionar una o más de las imágenes determinadas imágenes para realizar predicción inter de la imagen actual. Además de las imágenes que se podrían haber utilizado para realizar predicción inter de la imagen identificada, la imagen identificada mediante el descodificador de vídeo 30B también es una imagen de referencia candidata que se puede usar para realizar predicción inter de la imagen actual. El codificador de vídeo 20A puede incluir las imágenes que se usan para realizar predicción inter de la imagen actual en el conjunto de imágenes de referencia (RPS) de la imagen actual, y construir listas de imágenes de referencia que incluyan las imágenes seleccionadas, y codificar mediante predicción inter la imagen actual basándose en la lista de imágenes de referencia construida.

[0138] Debido a que las imágenes seleccionadas están en el RPS de la imagen actual, el descodificador de vídeo 30B puede construir de manera similar listas de imágenes de referencia que son idénticas a las construidas por el codificador de vídeo 20A, y por lo tanto, incluir las imágenes seleccionadas en las listas de imágenes de referencia. El descodificador de vídeo 30B puede entonces descodificar mediante predicción inter de la imagen actual basándose en las listas de imágenes de referencia.

[0139] En algunos ejemplos, el descodificador de vídeo 30B puede identificar una imagen que no está almacenada en la DPB 64 del codificador de vídeo 20A. Por ejemplo, es posible que el codificador de vídeo 20A ya haya descargado la imagen identificada, antes de que el descodificador de vídeo 30B haya descargado la imagen identificada. En los ejemplos en los que la DPB 64 del codificador de vídeo 20A no almacena la imagen identificada, el codificador de vídeo 20A puede realizar predicción intra de la imagen actual como un IRAP (punto de acceso aleatorio de actualización interna).

[0140] En algunos ejemplos, incluso si la imagen que identificó el descodificador de vídeo 30B está disponible en la DPB 64 del codificador de vídeo 20A, el codificador de vídeo 20A aún puede realizar predicción intra de la imagen actual. Por ejemplo, si la eficiencia de la codificación es menor (o no mejor) al usar una o más de las imágenes de referencia candidatas para realizar predicción inter de la imagen actual que para realizar predicción intra de imagen actual, el codificador de vídeo 20A puede realizar predicción intra la imagen actual, posiblemente como un IRAP.

[0141] Por consiguiente, en las técnicas descritas en esta divulgación, el descodificador de vídeo 30B puede informar al codificador de vídeo 20A a través de un identificador único para una imagen para describir el conjunto preferido de imágenes de referencia disponibles en la DPB 92 del descodificador de vídeo 30B que puede ser utilizado por el codificador de vídeo 20A para determinar las imágenes de referencia disponibles tanto en el codificador de vídeo 20A como en el descodificador de vídeo 30B (por ejemplo, determinar las imágenes de referencia disponibles almacenadas en DPB 64 y DPB 92). El identificador único (por ejemplo, la información que identifica una imagen) puede comprender un valor de recuento de orden de imagen (POC) de la imagen. Por ejemplo, cada imagen puede asociarse con un valor de POC único que identifica el orden en el que se muestra o emite la imagen. Una imagen con un valor de POC más pequeño se muestra o emite antes que una imagen con un valor de POC más grande. El valor de POC es solo un ejemplo de un identificador único utilizado mediante el descodificador de vídeo 30B para identificar la imagen, y otras formas de identificar la imagen pueden ser posibles.

[0142] De esta manera, las imágenes de referencia comunes disponibles tanto para el codificador de vídeo 20A como para el descodificador de vídeo 30B se pueden usar para establecer una referencia común entre el codificador de vídeo 20A y el descodificador de vídeo 30B (es decir, una vez que el codificador de vídeo 20A determina las imágenes que están disponibles en la DPB 64 del codificador de vídeo 20A y en la DPB 92 del descodificador de vídeo 30B, el codificador de vídeo 20A puede determinar qué imágenes usar para realizar predicción inter de la imagen actual). En algunos ejemplos, las imágenes de referencia asociadas con un identificador de imagen (es decir, imagen identificada) se describen como todas las imágenes de referencia disponibles después de descodificar la imagen cuyo identificador se envía (es decir, todas las imágenes que podrían haber sido utilizadas para realizar predicción inter de la imagen identificada y las imágenes que siguen a la imagen identificada en orden de codificación).

[0143] En algunos ejemplos, el indicador puede indicar todas las imágenes de referencia disponibles antes de descodificar la imagen indicada (es decir, las imágenes en el subconjunto RefPicSetStCurrBefore de la imagen identificada), o las imágenes después de descodificar la imagen indicada (es decir, las imágenes en el subconjunto RefPicSetStCurrAfter de la imagen identificada). Una vez que el codificador de vídeo 20A establece el estado DPB de la DPB 92 del descodificador de vídeo 30B con las imágenes de referencia asociadas en el lado del descodificador de vídeo 30B (receptor), el codificador de vídeo 20A puede hacer una elección de referencias para el establecimiento de un conjunto de referencia común. De forma alternativa, el codificador de vídeo 20A puede enviar una imagen IRAP (por ejemplo, una imagen con predicción intra tal como una imagen de punto de acceso aleatorio de actualización intra).

[0144] Como se describió anteriormente, el codificador de vídeo 20A también puede evitar referirse a las imágenes de RefPicSetStFoll y RefPicSetLtFoll en las imágenes de referencia señalizadas, ya que estas imágenes pueden no estar disponibles en el lado del receptor (por ejemplo, el descodificador de vídeo 30B) debido a la posible extracción del flujo de bits secundario en el flujo de bits original. En algunos ejemplos, un dispositivo intermedio como el dispositivo de elemento de red de acceso a medios (MANE) puede realizar una extracción de sub-flujo de bits en la que se extraen imágenes que solo pueden usarse para realizar predicción inter de imágenes después de la imagen actual en orden

de descodificación.

[0145] Como también se describió anteriormente, el indicador (es decir, la información que identifica la imagen) puede ser el PicOrderCntVal (valor de POC) especificado en la especificación HEVC ("Codificación de vídeo de alta eficiencia, Rec. ITU-T H.265 e ISO/IEC 23008-2") (valor con signo de 32 bits) o algo que es equivalente a la identificación única del valor de POC de la imagen. Además de PicOrderCntVal, el identificador de la capa nuh, como se especifica en la especificación HEVC, puede señalarse para flujos de bits multicapa similares a los de la sintaxis actual de RPSI, SPLI en el borrador de la especificación de carga útil RTP de HEVC (es decir, "Formato de carga útil RTP para codificación de vídeo de alta eficiencia").

[0146] Por consiguiente, en uno o más ejemplos, el codificador de vídeo 20A puede recibir, desde el dispositivo 14 que incluye el descodificador de vídeo 30B, información que identifica una imagen previamente descodificada que se almacena en la DPB 64 del descodificador de vídeo 30B en el momento en que el descodificador de vídeo 30B transmitió la información. El codificador de vídeo 20A puede determinar una o más imágenes de referencia candidatas que están disponibles para codificar una imagen actual basándose en la imagen identificada. La una o más imágenes de referencia candidatas incluyen una o más imágenes que previamente se determinó que eran utilizables para la codificación mediante predicción inter de una o más de la imagen identificada o imágenes que siguen a la imagen identificada en el orden de codificación, y que aún se almacenan en DPB 64 de codificador de vídeo 20A en el momento de codificar la imagen actual. El codificador de vídeo 20A puede seleccionar una o más imágenes de referencia para la imagen actual a partir de una o más imágenes de referencia candidatas determinadas y la imagen identificada. El codificador de vídeo 20A puede codificar mediante predicción inter de la imagen actual basándose en una o más imágenes de referencia seleccionadas.

[0147] Las una o más imágenes que se determinaron previamente para ser utilizadas para la codificación mediante predicción inter de una o más de la imagen identificada o las imágenes que siguen a la imagen identificada en orden de codificación incluyen imágenes de uno o más subconjuntos de imágenes de referencia que el codificador de vídeo 20A construyó durante la codificación de la imagen identificada. Nuevamente, la imagen identificada es una imagen previamente descodificada mediante el descodificador de vídeo 30B, lo cual significa que el codificador de vídeo 20A codificó la imagen identificada antes de codificar la imagen actual.

[0148] Debido a que las técnicas descritas en esta divulgación permiten que el codificador de vídeo 20A determine una pluralidad de imágenes de referencia candidatas, el codificador de vídeo 20A puede codificar mediante predicción bidireccional la imagen actual basándose en dos imágenes de las imágenes de referencia seleccionadas. Dicha codificación de bi-predicción puede permitir opciones de predicción inter adicionales que de otra forma no estarían disponibles si el codificador de vídeo 20A se limitara a usar solo la imagen identificada para la predicción inter.

[0149] Para la descodificación, el descodificador de vídeo 30B puede emitir, al dispositivo 12 que incluye el codificador de vídeo 20A, información que identifica una imagen previamente descodificada almacenada en la DPB 92 del descodificador de vídeo 30B en el momento de emitir la información. El descodificador de vídeo 30B puede descodificar mediante predicción inter una imagen actual con una o más imágenes de referencia candidatas. La una o más imágenes de referencia candidatas incluyen una o más imágenes que se determinaron previamente, mediante el codificador de vídeo 20A, que se pueden utilizar para la codificación mediante predicción inter de la imagen identificada. En algunos ejemplos, el descodificador de vídeo 30B puede determinar que se perdió una imagen (por ejemplo, el descodificador de vídeo 30B no recibió información para reconstruir completa o correctamente la imagen perdida). El descodificador de vídeo 30B puede configurarse para emitir información basándose en una determinación de que la imagen se perdió. Además, debido a que el codificador de vídeo 20A puede ser capaz de codificar mediante predicción bidireccional la imagen actual a partir de dos imágenes, el descodificador de vídeo 30B puede configurarse para descodificar mediante predicción bidireccional la imagen actual basándose en dos imágenes de las imágenes de referencia candidatas.

[0150] El descodificador de vídeo 30B puede comunicar la información que identifica una imagen previamente descodificada almacenada en la DPB 92 del descodificador de vídeo 30B en el momento de emitir la información a través del enlace 16. En algunos ejemplos, el descodificador de vídeo 30B puede comunicar la información a través del enlace o protocolo de comunicación utilizado para comunicar el flujo de bits de vídeo desde el codificador de vídeo 20A al descodificador de vídeo 30B (por ejemplo, un protocolo de señalización de dos vías, incluido un RTP). Por ejemplo, la información puede ser comunicada desde el descodificador de vídeo 30B al codificador de vídeo 20A a través de un mensaje o mecanismo de señalización nuevo o existente.

[0151] En algunos ejemplos, las técnicas introducen un mensaje de retroalimentación de Indicación de Selección de Conjunto de Imágenes de Referencia (RPSSI), que en algunos ejemplos comunica la información que identifica las imágenes previamente descodificadas almacenadas en la DPB 92 del descodificador de vídeo 30B con el codificador de vídeo 20A. En un ejemplo, descrito más detalladamente a continuación, el mensaje RPSSI se señala como parte de una sesión de protocolo RTP que comunica el flujo de bits de vídeo codificado. Por ejemplo, el mensaje de retroalimentación RPSSI puede reemplazar al mensaje RPSI, puede usarse además del mensaje RPSI para diferenciar cuándo usar las técnicas existentes y técnicas descritas en esta divulgación, o el mensaje RPSI puede modificarse para indicar que deben implementarse las técnicas descritas en esta divulgación. Para facilitar la

descripción, a continuación se describe el mensaje como el mensaje de retroalimentación RPSSI, pero es simplemente para facilitar la comprensión y no debe considerarse limitante ni limitado al ejemplo específico descrito a continuación.

5 **[0152]** El mensaje de retroalimentación RPSSI se puede identificar por PT=PSFB y FMT=9. Puede haber una RPSSI (y en algunos ejemplos debe ser exactamente una RPSSI) contenida en el campo FCI. La tabla 1 a continuación ilustra el ejemplo del formato de la RPSSI.



10 **TABLA 1 El formato PCI de la RPSSI.**

15 **[0153]** En el ejemplo de la Tabla 1, PB es de 8 bits. PB puede definir el número de bits no utilizados necesarios para rellenar la longitud del mensaje RPSSI a un múltiplo de 32 bits. En la tabla 1, 0 es 1 bit. Este bit puede ponerse a cero (por ejemplo, debe ponerse a cero) en la transmisión y puede ignorarse en la recepción. En la tabla 1, el tipo de carga útil es de 7 bits. El tipo de carga útil puede indicar el tipo de carga útil RTP en el contexto en el que la cadena de bits RPSSI original puede (por ejemplo, debe) interpretarse. En la Tabla 1, la cadena de bits RPSSI original es de longitud variable. El RPSSI original puede indicar la información SPLI como se define de forma original mediante el códec de vídeo. En la Tabla 1, el relleno puede ser #PB bits. Una cantidad de bits se establece en cero para llenar el contenido del mensaje RPSSI hasta el siguiente límite de 32 bits. El número de bits de relleno puede (por ejemplo, debe) estar indicado por el campo PB.

25 **[0154]** En algunos ejemplos, se aplican las mismas reglas de tiempo que para el mensaje RPSI, como se define en [RFC4585], para el mensaje RPSSI. El campo "Cadena de bits RPSSI original definida por códec" puede ser una representación base16 [RFC4648] de los 8 bits que consta de 2 bits más significativos iguales a 0 y 6 bits de id de capa nuh, como se define en HEVC, seguidos de los 32 bits representa el valor de PicOrderCntVal, como se define en HEVC, para la imagen que se indica para señalar la presencia de imágenes de referencia disponibles en el lado del receptor.

30 **[0155]** Aunque las técnicas de ejemplo anteriores se describen con respecto a la norma HEVC, las técnicas descritas en esta divulgación no están necesariamente limitadas. Por ejemplo, las técnicas se pueden usar con otros códecs, como el códec compatible con H.264/AVC, así como otras normas de compresión de vídeo o incluso técnicas de compresión de vídeo patentadas. A continuación se describe una forma de ejemplo en la que se pueden utilizar las técnicas con el códec H.264/AVC.

35 **[0156]** En cualquiera de las técnicas de ejemplo, una ventaja puede ser que al señalar un único identificador de imagen, se puede señalar un conjunto completo de imágenes de referencia que consiste en múltiples tramas/imágenes. En otras palabras, el descodificador de vídeo 30B muestra información que identifica algunas imágenes (por ejemplo, una imagen) y, a partir de esa información, el codificador de vídeo 20A puede determinar más imágenes (por ejemplo, una pluralidad de imágenes) que se pueden usar para realizar predicción inter de la imagen original. Esto permite una utilización eficiente del ancho de banda mediante el descodificador de vídeo 30B, mientras que promueve la eficiencia de la codificación de vídeo mediante el codificador de vídeo 20A.

45 **[0157]** En algunas técnicas de vídeo, aparte de las técnicas de vídeo de acuerdo con HEVC, el codificador de vídeo 20A puede no estar preconfigurado para determinar y mantener una lista de imágenes que podrían usarse para realizar predicción inter de una imagen y las imágenes que siguen a la imagen en el orden de codificación. Por ejemplo, si el codificador de vídeo 20A se configuró basándose en la norma H.264/AVC, entonces el codificador de vídeo 20A puede configurarse para determinar las imágenes que podrían haberse utilizado para realizar predicción inter de la imagen identificada (por ejemplo, como parte del codificador de vídeo de proceso 20A implementado al codificar anteriormente la imagen identificada), pero no se puede configurar para mantener una lista de imágenes que podrían haberse utilizado para realizar predicción inter de la imagen identificada mediante el descodificador de vídeo 30B. Por ejemplo, después de codificar la imagen identificada, el codificador de vídeo 20A puede descartar información que indica qué imágenes se podrían haber utilizado para realizar predicción inter de la imagen identificada. En otras palabras, el codificador de vídeo 20A, configurado de acuerdo con H.264/AVC, no puede construir subconjuntos de imágenes de referencia de la manera en que el codificador de vídeo 20A tendría si estuviera configurado de acuerdo con HEVC.

55 **[0158]** Para ejemplos en los que el codificador de vídeo 20A no está configurado de acuerdo con HEVC, el codificador de vídeo 20A puede configurarse para (por ejemplo, diseñarse para en este caso) mantener la lista de imágenes para cada imagen que podría haber utilizado para realizar predicción inter de una imagen, como como en

la memoria del codificador de vídeo 20A, para que la información esté disponible para la imagen identificada. De esta manera, incluso para el codificador de vídeo 20A configurado de acuerdo con H.264/AVC, puede ser posible que el codificador de vídeo 20A implemente las técnicas de ejemplo descritas en esta divulgación.

5 **[0159]** De manera similar, algunos ejemplos de un descodificador de vídeo que están configurados de acuerdo con H.264/AVC pueden no soportar la información de salida que identifica una imagen previamente descodificada. Para tales ejemplos, el descodificador de vídeo 30B puede configurarse específicamente para soportar la información de salida que identifica una imagen previamente descodificada para implementar las técnicas de ejemplo descritas en esta divulgación.

10 **[0160]** La descripción anterior describió técnicas de ejemplo en las que el codificador de vídeo 20A puede determinar una pluralidad de imágenes de referencia candidatas que estarán disponibles para la descodificación mediante el descodificador de vídeo 30B para la codificación mediante predicción inter de una imagen actual basada en la información emitida mediante el descodificador de vídeo 30B. Sin embargo, esta divulgación no se limita a las técnicas de ejemplo anteriores. A continuación se describen algunas técnicas de ejemplo adicionales que se pueden usar junto con las técnicas de ejemplo anteriores o independientemente de las técnicas de ejemplo anteriores.

15 **[0161]** En algunos ejemplos, para cada imagen de referencia perdida en el lado del descodificador (por ejemplo, para cada imagen que el descodificador de vídeo 30B pierde), el descodificador de vídeo 30B envía un identificador único de esta imagen (por ejemplo, imagen perdida), como su valor de POC al codificador de vídeo 20A como un mensaje SPLI (el mensaje SPLI se describe con más detalle a continuación). Además (por ejemplo, adicionalmente), el descodificador de vídeo 30B puede enviar un identificador único de la imagen que el descodificador de vídeo 30B está actualmente descodificando al codificador de vídeo 20A. La imagen que el descodificador de vídeo 30B está actualmente descodificando es una imagen que el codificador de vídeo 20A previamente codificó. Con el identificador de la imagen que el descodificador de vídeo 30B estaba descodificando en el momento en que el descodificador de vídeo 30B emitió la información que identifica la imagen, el codificador de vídeo 20A puede determinar qué imágenes de referencia se deben almacenar en la DPB 92 del descodificador de vídeo 30B.

20 **[0162]** Por ejemplo, como se describió anteriormente, como parte de la descodificación de una imagen, el descodificador de vídeo 30B construye los subconjuntos de imágenes de referencia para la imagen y construye las listas de imágenes de referencia para esa imagen a partir de los subconjuntos de imágenes de referencia. En este ejemplo, el codificador de vídeo 20A recibe información de la imagen que el descodificador de vídeo 30B estaba descodificando en el momento en que el descodificador de vídeo 30B transmitió la información, y determina los subconjuntos de imágenes de referencia (o, más en general, el parámetro establecido para cubrir H.264/AVC) para determinar qué imágenes están disponibles en la DPB 92 del descodificador de vídeo 30B. Además, el codificador de vídeo 20A puede determinar que la imagen que el descodificador de vídeo 30B indicó como perdida no se encuentra en la DPB 92 del descodificador de vídeo 30B. Por lo tanto, a excepción de las imágenes de referencia perdidas informadas desde el descodificador de vídeo 30B, las imágenes de referencia restantes deben existir todas en DPB 92 del descodificador de vídeo 30B. En general, debido a que solo se pierden unas pocas imágenes de referencia, el descodificador de vídeo 30B puede generar información que identifique las imágenes perdidas solo unas pocas veces (por ejemplo, se pierden pocas imágenes de referencia, solo se requiere enviar pocos mensajes). En consecuencia, esta técnica de ejemplo puede ser relativamente eficiente en ancho de banda.

30 **[0163]** Como otro ejemplo, similar al ejemplo anterior, el descodificador de vídeo 30B puede generar información que identifica una imagen perdida, y el codificador de vídeo 20A puede recibir dicha información que identifica la imagen perdida. El codificador de vídeo 20A puede excluir la imagen perdida de una o más imágenes de referencia candidatas determinadas. De esta manera, el descodificador de vídeo 30B puede descodificar mediante predicción inter la imagen actual con una o más imágenes de referencia candidatas, excepto la imagen perdida.

35 **[0164]** Como otra técnica de ejemplo que se puede usar con las técnicas anteriores o por separado de las técnicas anteriores, para cada imagen de referencia descodificada correctamente en la DPB 92 del descodificador de vídeo 30B, el descodificador de vídeo 30B emite al codificador de vídeo 20A un identificador único de esta imagen, tal como como su valor de POC, mediante un mensaje de retroalimentación. Esta técnica de ejemplo puede ser similar a enviar varios mensajes RPSI juntos. Sin embargo, la diferencia es que el codificador de vídeo 20A puede determinar qué imágenes de referencia usar, en lugar de que el descodificador de vídeo 30B especifique una imagen de referencia para la imagen a utilizar para codificar mediante predicción inter la imagen actual.

40 **[0165]** En este ejemplo, el descodificador de vídeo 30B puede generar información que identifique cada imagen que el descodificador de vídeo 30B pudo reconstruir correctamente que está almacenada en la DPB 92 del descodificador de vídeo 30B. El codificador de vídeo 20A puede determinar cuál de las imágenes, para la cual el codificador de vídeo 20A recibió dicha información, todavía está disponible en la DPB 64 del codificador de vídeo 20A. El codificador de vídeo 20A puede usar una o más de estas imágenes para la codificación mediante predicción inter de la imagen actual. Cuando hay muchas imágenes de referencia en la DPB 92 del descodificador de vídeo 30B, el descodificador de vídeo 30B puede emitir muchos mensajes de retroalimentación juntos al codificador de vídeo 20A. En consecuencia, esta técnica de ejemplo puede ser relativamente ineficiente en cuanto a ancho de banda, pero con una mejor garantía de que el codificador de vídeo 20A codificó mediante predicción inter la imagen actual con imágenes de referencia

disponibles en la DPB 92 del descodificador de vídeo 30B.

[0166] De acuerdo con una o más técnicas de esta divulgación, el descodificador de vídeo 30B puede transmitir información al codificador de vídeo 20A, indicativa de imágenes de referencia disponibles para descodificación. El descodificador de vídeo 30B puede entonces recibir datos de vídeo para una imagen que está codificada mediante predicción con una o más de las imágenes de referencia. El descodificador de vídeo 30B puede descodificar mediante predicción inter la imagen basándose en los datos de vídeo.

[0167] En algunos ejemplos, el descodificador de vídeo 30B puede transmitir información al codificador de vídeo 20A indicativa de imágenes de referencia disponibles para descodificar si se produjo un error en la reconstrucción de una imagen de referencia durante la descodificación. En algunos ejemplos, para transmitir información al codificador de vídeo 20A indicativa de las imágenes de referencia disponibles para descodificación, el descodificador de vídeo 30B puede transmitir información que identifica una imagen identificada en un conjunto de imágenes de referencia a partir del cual las imágenes de referencia incluidas en el conjunto de referencia pueden determinarse como las imágenes de referencia disponibles para descodificación en descodificador de vídeo 30B.

[0168] De acuerdo con uno o más ejemplos, el descodificador de vídeo 30B puede emitir, al dispositivo 12 que incluye el codificador de vídeo 20A, información que identifica una imagen previamente descodificada almacenada en la DPB 92 del descodificador de vídeo 30B en el momento de emitir la información. El descodificador de vídeo 30B puede descodificar mediante predicción inter una imagen actual con una o más imágenes de referencia candidatas. La una o más imágenes de referencia candidatas incluyen una o más imágenes que se determinaron previamente, mediante el codificador de vídeo 20A, que se pueden utilizar para la codificación mediante predicción inter de la imagen identificada. En algunos ejemplos, el descodificador de vídeo 30B puede determinar que se perdió una imagen. El descodificador de vídeo 30B puede generar información que identifique una imagen previamente descodificada basándose en una determinación de que la imagen se perdió.

[0169] El codificador de vídeo 20A puede seleccionar dos imágenes de las imágenes de referencia candidatas seleccionadas para codificar mediante predicción bidireccional la imagen actual. Por consiguiente, el descodificador de vídeo 30B puede descodificar mediante predicción bidireccional la imagen actual basándose en dos imágenes de las imágenes de referencia candidatas.

[0170] Además, aparte de emitir información que identifica una imagen previamente descodificada, el descodificador de vídeo 30B puede emitir información que identifica una imagen perdida. En tales ejemplos, el descodificador de vídeo 30B puede descodificar mediante predicción inter la imagen actual con una o más imágenes de referencia candidatas, excepto la imagen perdida.

[0171] La FIG. 4 es un diagrama de flujo que ilustra un procedimiento de ejemplo para codificar datos de vídeo. Con fines de ilustración, el ejemplo se describe con respecto al codificador de vídeo 20A. El codificador de vídeo 20A puede determinar las imágenes que se pueden usar para codificar mediante predicción inter cada imagen y las imágenes siguientes en orden de codificación (100). Por ejemplo, como parte de la codificación de cada imagen, el codificador de vídeo 20A puede configurarse para determinar un conjunto de parámetros para la imagen (por ejemplo, el uno o más subconjuntos de imágenes de referencia definidos en HEVC o mantener el conjunto de parámetros en técnicas de codificación de vídeo que no están necesariamente de acuerdo con HEVC). El codificador de vídeo 20A puede almacenar información que indica imágenes determinadas (por ejemplo, imágenes del conjunto de parámetros) para cada imagen (102).

[0172] El codificador de vídeo 20A puede recibir información que identifica una imagen en particular (por ejemplo, por su valor de POC) (104). El codificador de vídeo 20A puede determinar una o más imágenes que previamente se determinó que se pueden usar para la codificación mediante predicción inter de la imagen identificada a partir de la información almacenada (106). En algunos ejemplos, el codificador de vídeo 20A también puede determinar una o más imágenes que previamente se determinó que eran utilizables para la codificación mediante predicción inter de las imágenes que previamente se determinó que eran utilizables para la codificación mediante predicción inter de la imagen identificada.

[0173] A partir de las imágenes que el codificador de vídeo 20A determinó que se pueden usar para la codificación mediante predicción inter de la imagen identificada (y posiblemente se determinó que se puede usar para la codificación mediante predicción inter de las imágenes determinadas para usarse para la codificación mediante predicción inter de la imagen identificada), el codificador de vídeo 20A puede determinar cuál de estas imágenes aún está almacenada en la DPB 64 del codificador de vídeo 20A (108). El codificador de vídeo 20A puede seleccionar una o más imágenes de referencia (por ejemplo, una imagen de referencia para la predicción unidireccional de una lista de imágenes de referencia o dos imágenes de referencia para la predicción bidireccional de imágenes de referencia en cada una de las dos listas de imágenes de referencia) de las imágenes de referencia candidatas determinadas (110). El codificador de vídeo 20A puede codificar mediante predicción inter la imagen actual basándose en las imágenes de referencia seleccionadas (112).

[0174] La FIG. 5 es un diagrama de flujo que ilustra un procedimiento de ejemplo para descodificar datos de vídeo. Para facilitar la descripción, el ejemplo se describe con respecto al descodificador de vídeo 30B. El descodificador de vídeo 30B puede emitir, al dispositivo 12 que incluye el codificador de vídeo 20A, información que identifica una imagen previamente descodificada almacenada en la DPB 92 del descodificador de vídeo 30B en el momento de emitir la información (200). En algunos ejemplos, el descodificador de vídeo 30B puede determinar que se perdió una imagen. El descodificador de vídeo 30B puede configurarse para emitir la información que identifica la imagen previamente descodificada basándose en una determinación de que la imagen se perdió.

[0175] El descodificador de vídeo 30B puede descodificar mediante predicción inter una imagen actual con una o más imágenes de referencia candidatas (202). La una o más imágenes de referencia candidatas incluyen una o más imágenes que se determinaron previamente, mediante el codificador de vídeo 20A, para ser utilizadas para la codificación mediante predicción inter de la imagen identificada y/o imágenes identificadas para ser utilizadas para la codificación mediante predicción inter de las imágenes que se identificaron para ser utilizadas para codificación mediante predicción inter de la imagen identificada. La imagen identificada también puede ser parte de una o más imágenes de referencia candidatas.

[0176] Por ejemplo, si el codificador de vídeo 20A codificó mediante predicción bidireccional la imagen actual, el descodificador de vídeo 30B puede descodificar mediante predicción bidireccional la imagen actual de una manera sustancialmente similar. Además, en algunos ejemplos, el descodificador de vídeo 30B también puede generar información que identifique una imagen perdida. En tales ejemplos, el descodificador de vídeo 30B puede descodificar mediante predicción inter la imagen actual con una o más imágenes de referencia candidatas, excepto la imagen perdida.

[0177] Las técnicas descritas anteriormente relacionadas con RPSI y RPSSI (de nuevo, la definición de RPSI puede cambiarse para incluir la definición de RPSSI). En algunos ejemplos, el codificador de vídeo 20A y el descodificador de vídeo 30B pueden utilizar los mensajes específicos de indicación de pérdida de imagen (SPLI). Para SPLI, cuando ocurre un error, el descodificador de vídeo 30B solo informa de imágenes de referencia perdidas o parcialmente perdidas. El codificador de vídeo 20A rastrea los errores de acuerdo con los mensajes SPLI, y una vez que el codificador de vídeo 20A puede inferir una imagen de referencia que está correctamente descodificada y aún existe en la DPB 92 del descodificador de vídeo 30B, el codificador de vídeo 20A puede usar esta imagen de referencia para codificar la imagen actual para detener la propagación del error.

[0178] Para los mensajes SPLI, dado que el descodificador de vídeo 30B solo puede informar qué imágenes se pierden, entonces el codificador de vídeo 20A solo puede inferir qué imágenes de referencia están disponibles en la DPB 92 del descodificador de vídeo 30B basándose en los mensajes recibidos. Sin embargo, el codificador de vídeo 20A puede no ser capaz de determinar cuándo ocurrió el error en el lado del descodificador de vídeo 30B (es decir, el codificador de vídeo 20A puede no determinar el estado exacto de la DPB 92 del descodificador de vídeo 30B cuando se informó de los errores). En algunos casos, puede ser complicado para el codificador de vídeo 20A inferir exactamente las imágenes almacenadas en la DPB 92 del descodificador de vídeo 30B, como se ilustra en el ejemplo de la Tabla 2.

TABLA 2 ESTRUCTURA DE REFERENCIA

	#	#	#	#	#
	#	#	#	#	#
	#	#	#	#	#
POC	t	t+1	t+2	t+3	t+4
DPB		(t)	(t+1, t*)	(t, t+1*)	(t+1, t+3)

[0179] La tabla 2 indica una estructura de referencia. En la Tabla 2, * significa que esta imagen se mantiene en DPB, pero no se usa para la descodificación de la imagen actual, y POC significa recuento de orden de imágenes, que indica cuándo se debe mostrar una imagen. Por ejemplo, una imagen con un valor de POC menor se muestra antes que una imagen con un valor de POC mayor. El orden de visualización de las imágenes puede ser diferente del orden de codificación o descodificación de las imágenes.

[0180] En la Tabla 2, el codificador de vídeo 20A usa la estructura de referencia ilustrada. Hay dos casos: Caso 1: solo se pierde la imagen t+1, el descodificador nota esta pérdida al descodificar la imagen t+2, y luego envía un mensaje SPLI para informar al codificador de que la imagen t+1 se pierde, y Caso 2: tanto la imagen t+1 como la imagen t+2 se pierden, el descodificador nota esta pérdida involuntaria al descodificar la imagen t+4, y luego el descodificador envía un mensaje SPLI para informar al codificador que se perdió la imagen t+1.

[0181] En algunos ejemplos, el descodificador de vídeo 30B puede no informar de la pérdida de la imagen de referencia cuando la imagen de referencia perdida no se usa para la descodificación de la imagen actual, porque esta

imagen de referencia puede ser descartada intencionalmente. Comparando estos dos casos, el codificador de vídeo 20A puede no ser capaz de distinguirlos solo por los mensajes SPLI recibidos. Es decir, el codificador de vídeo 20A no puede inferir si la imagen t+3 está disponible o no en la DPB del descodificador de vídeo 30B.

5 **[0182]** En las técnicas de ejemplo anteriores, el descodificador de vídeo 30B genera información que identifica algunas imágenes (por ejemplo, una imagen o algunas imágenes seleccionadas) desde las cuales el codificador de vídeo 20A determina más imágenes (por ejemplo, una pluralidad de imágenes) que se pueden usar para codificar mediante predicción inter la imagen actual. Sin embargo, en algunos ejemplos, el descodificador de vídeo 30B que emite información que identifica algunas imágenes puede no ser suficiente para activar el codificador de vídeo 20A para determinar las imágenes que el codificador de vídeo 20A puede usar para codificar mediante predicción inter una imagen actual. En tales ejemplos, además de emitir información que identifica una o más imágenes en la DPB 92 del descodificador de vídeo 30B, el descodificador de vídeo 30B también puede emitir información que indica que se ha perdido una imagen, e información que identifica la imagen perdida. En respuesta, el codificador de vídeo 20A puede codificar mediante predicción inter una imagen actual basándose en imágenes identificadas como que se encuentran en la DPB 92 del descodificador de vídeo 30B o imágenes que previamente se determinaron como utilizables para realizar predicción inter de las imágenes identificadas o imágenes que siguen a las imágenes identificadas en orden de codificación.

20 **[0183]** La FIG. 6 es un diagrama de flujo que ilustra otro procedimiento de ejemplo para codificar datos de vídeo. Como anteriormente, el ejemplo se describe con respecto al codificador de vídeo 20A. Como se ilustra en la FIG. 6, el codificador de vídeo 20A puede recibir información que identifica una imagen perdida del descodificador de vídeo 30B (300). El codificador de vídeo 20A también puede recibir información que identifica una o más imágenes que actualmente están almacenadas en la DPB 92 del descodificador de vídeo 30B (302). En algunos ejemplos, el codificador de vídeo 20A puede recibir información que identifica una o más imágenes que actualmente están almacenadas en la DPB 92 del descodificador de vídeo 30B solo en los casos en que el descodificador de vídeo 30B determina que se perdió una imagen.

30 **[0184]** Por ejemplo, en algunos ejemplos, el descodificador de vídeo 30B puede emitir información que identifica una imagen descodificada correctamente con el codificador de vídeo 20A después de cada instancia de descodificación correcta de la imagen, y el codificador de vídeo 20A puede a continuación utilizar una o más de estas imágenes identificadas para codificar mediante predicción inter una imagen actual. En algunos ejemplos, el descodificador de vídeo 30B puede generar información que identifique una o más imágenes almacenadas en la DPB 64 (por ejemplo, una o más imágenes descodificadas correctamente) solo cuando el descodificador de vídeo 30B determina que una imagen se perdió (por ejemplo, una imagen no se descodificó, no se reconstruyó con precisión, o no se reconstruyó completamente). En estos ejemplos, el descodificador de vídeo 30B también puede generar información que identifique la imagen perdida (por ejemplo, el valor de POC de la imagen perdida).

40 **[0185]** Como se ilustra en la FIG. 6, la información de recepción que identifica una imagen perdida y la información de recepción que identifica una o más imágenes en la DPB 92 del descodificador de vídeo 30B se ilustran por separado. Sin embargo, las técnicas descritas en esta divulgación no están limitadas a esto. En algunos ejemplos, el codificador de vídeo 20A puede recibir información que identifica una imagen perdida e información que identifica una o más imágenes en la DPB 92 del descodificador de vídeo 30B en un mensaje de retroalimentación agrupado. Como ejemplo, un solo mensaje de retroalimentación puede incluir información que identifica la imagen perdida e información que identifica una o más imágenes almacenadas actualmente en la DPB 92 del descodificador de vídeo 30B.

45 **[0186]** El codificador de vídeo 20A puede determinar imágenes de referencia candidatas (304). Las imágenes candidatas pueden estar limitadas a solo una o más imágenes en la DPB 92 del descodificador de vídeo 30B que el descodificador de vídeo 30B identificó en el mensaje de retroalimentación. En algunos ejemplos, las imágenes candidatas pueden ser imágenes, aún almacenadas en la DPB 64 del codificador de vídeo 20A, que podrían haberse utilizado para codificar mediante predicción inter el descodificador de vídeo de imágenes 30B identificado.

50 **[0187]** El codificador de vídeo 20A puede seleccionar la imagen de referencia de las imágenes de referencia candidatas para codificar mediante predicción inter la imagen actual (306). El codificador de vídeo 20A puede codificar mediante predicción inter la imagen actual basándose en las imágenes de referencia seleccionadas (308).

55 **[0188]** La FIG. 7 es un diagrama de flujo que ilustra otro procedimiento de ejemplo para descodificar datos de vídeo. Como anteriormente, para mayor facilidad, la descripción se realiza con respecto al descodificador de vídeo 30B. Como se ilustra, el descodificador de vídeo 30B puede determinar que se perdió una imagen (400). Por ejemplo, el descodificador de vídeo 30B puede determinar que se perdió una imagen cuando esa imagen se va a usar para realizar predicción inter de otra imagen, y el descodificador de vídeo 30B determina que la imagen no está disponible en la DPB 92 del descodificador de vídeo 30B. Como otro ejemplo, el descodificador de vídeo 30B puede determinar que se perdió una imagen cuando el descodificador de vídeo 30B emitirá esa imagen, y la imagen no está disponible en la DPB 92 del descodificador de vídeo 30B.

65 **[0189]** El descodificador de vídeo 30B puede emitir información que identifica la imagen perdida al codificador de vídeo 20A (402). El descodificador de vídeo 30B también puede mostrar información que identifica una o más

imágenes almacenadas actualmente en la DPB 92 del descodificador de vídeo 30B (404). Por ejemplo, el descodificador de vídeo 30B puede generar valores de recuento de orden de imágenes (POC) que identifiquen la imagen perdida y la una o más imágenes almacenadas actualmente en la DPB 92 del descodificador de vídeo 30B. En algunos ejemplos, el descodificador de vídeo 30B puede emitir información que identifica la imagen perdida e información que identifica la una o más imágenes almacenadas actualmente en la DPB 92 en un mensaje de retroalimentación agrupado. Como ejemplo, puede haber solo un mensaje que el descodificador de vídeo 30B debe enviar para identificar la imagen perdida y la o las imágenes almacenadas actualmente en la DPB 92. De esta manera, el descodificador de vídeo 30B puede transmitir un mensaje de tipo SPLI que identifica una imagen perdida y un mensaje de tipo RPSI que identifica una o más imágenes almacenadas actualmente en la DPB 92 del descodificador de vídeo 30B.

[0190] En algunos ejemplos, el descodificador de vídeo 30B puede generar información que identifique la imagen perdida y una o más imágenes almacenadas actualmente en la DPB 92 en más de un mensaje. En algunos ejemplos, el descodificador de vídeo 30B puede generar información que identifique cada imagen que descodificó correctamente al codificador de vídeo 20A después de descodificar la imagen. En algunos ejemplos, el descodificador de vídeo 30B puede generar información que identifique una o más imágenes que actualmente están almacenadas en la DPB 92 solo cuando el descodificador de vídeo 30B determina que se perdió una imagen.

[0191] El descodificador de vídeo 30B puede descodificar mediante predicción inter de una imagen actual de una o más imágenes de referencia (406). Las imágenes de referencia pueden incluir solo las imágenes identificadas. En algunos ejemplos, las imágenes de referencia pueden incluir las imágenes identificadas y las imágenes que podrían haberse utilizado para realizar predicción inter de las imágenes identificadas. En algunos ejemplos, las imágenes de referencia solo pueden incluir imágenes que podrían haberse utilizado para realizar predicción inter de las imágenes identificadas.

[0192] 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, como una o más instrucciones o código, pueden almacenarse en, o transmitirse por, un medio legible por ordenador 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. De esta manera, los medios legibles por ordenador pueden corresponder en general a medios de almacenamiento legibles por ordenador tangibles que no sean transitorios. Los medios de almacenamiento de datos pueden ser cualquier medio disponible al que se pueda acceder desde uno o más ordenadores o uno o más procesadores para recuperar instrucciones, código y/o estructuras de datos para la implementación de las técnicas descritas en esta divulgación. Un producto de programa informático puede incluir un medio legible por ordenador.

[0193] 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. 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, se orientan a medios de almacenamiento tangibles no transitorios. El término disco, tal como se utiliza en el presente documento, incluye un disco compacto (CD), un disco láser, un disco óptico, un disco versátil digital (DVD), un disco flexible y un disco Blu-ray, donde algunos discos habitualmente 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.

[0194] 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), matrices de puertas programables *in situ* (FPGA) u otros circuitos lógicos integrados o discretos equivalentes. En consecuencia, el término «procesador», como se usa en el presente documento, puede referirse 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 puede proporcionarse dentro de módulos de hardware y/o software dedicados, configurados para la codificación y la descodificación, o incorporados en un códec combinado. Asimismo, las técnicas podrían implementarse por completo en uno o más circuitos o elementos lógicos.

[0195] 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 pueden combinarse en una unidad de hardware de códec o proporcionarse por medio de un grupo de unidades de hardware interoperativas, incluyendo uno o más procesadores, como se ha descrito anteriormente, conjuntamente con software y/o firmware adecuados.

[0196] Se han descrito diversos ejemplos. Estos y otros ejemplos están dentro del alcance de las siguientes reivindicaciones.

REIVINDICACIONES

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

5 almacenar un conjunto de parámetros para cada imagen que está codificada, el conjunto de parámetros de una imagen particular que incluye uno o más subconjuntos de imágenes de referencia que forman un conjunto de imágenes de referencia de la imagen particular;

10 recibir, desde un dispositivo que comprende un descodificador de vídeo, información que identifica una imagen previamente descodificada que se almacena en una memoria intermedia de imágenes descodificadas (DPB) del descodificador de vídeo en el momento en que el descodificador de vídeo transmitió la información;

15 determinar, basándose en el conjunto de parámetros de la imagen identificada, una o más imágenes de referencia candidatas que están disponibles para codificar una imagen actual, en el que la una o más imágenes de referencia candidatas incluyen una o más imágenes que están:

20 en el conjunto de imágenes de referencia de la imagen identificada, en el que el conjunto de imágenes de referencia identifica las imágenes que se determinaron previamente, durante la codificación de la imagen identificada, como imágenes que podrían haberse utilizado para codificar mediante predicción inter la imagen identificada o como imágenes que podrían haberse utilizado para la codificación mediante predicción inter de las imágenes que siguen a la imagen identificada en orden de codificación; y

25 que todavía se almacenan en una DPB de un codificador de vídeo;

seleccionar una o más imágenes de referencia para la imagen actual de una o más imágenes de referencia candidatas determinadas y la imagen identificada; y

30 codificar mediante predicción inter la imagen actual basándose en una o más imágenes de referencia seleccionadas.

2. El procedimiento según la reivindicación 1, que comprende además:

35 construir, durante la codificación de una imagen disponible que luego es identificada mediante el descodificador de vídeo y antes de la codificación de la imagen actual, uno o más de:

40 un primer subconjunto de imágenes de referencia que incluye imágenes de referencia a corto plazo que preceden a la imagen disponible en el orden de visualización y que se determinó que se pueden usar para realizar predicción inter de la imagen disponible y se determinó que se pueden usar para realizar predicción inter de la una o más imágenes que siguen a la imagen disponible en orden de codificación;

45 un segundo subconjunto de imágenes de referencia que incluye imágenes de referencia a corto plazo que siguen a la imagen disponible en el orden de visualización y que se determinó que se pueden usar para realizar predicción inter de la imagen disponible y se determinó que se pueden usar para realizar predicción inter una o más imágenes que siguen a la imagen disponible en orden de codificación;

50 un tercer subconjunto de imágenes de referencia que incluye cualquier imagen de referencia a largo plazo que se determinó que era utilizable para realizar predicción inter de la imagen disponible y que se determinó que era utilizable para realizar predicción inter de la una o más imágenes que siguen a la imagen disponible en orden de codificación;

55 un cuarto subconjunto de imágenes de referencia que incluye imágenes de referencia a corto plazo que se determinó que se pueden usar para realizar predicción inter de la una o más imágenes que siguen a la imagen disponible en orden de codificación, y se determinó que no se pueden usar para realizar predicción inter de la imagen disponible ; y

60 un quinto subconjunto de imágenes de referencia que incluye imágenes de referencia a largo plazo que se determinó que se pueden usar para realizar predicción inter de la una o más imágenes que siguen a la imagen disponible en orden de codificación, y se determinó que no se pueden usar para realizar predicción inter de las imágenes disponibles,

65 en el que las imágenes en uno o más de los primeros, segundo, tercero, cuarto y quinto subconjuntos de imágenes de referencia comprenden imágenes que previamente se determinó que eran utilizables para la codificación mediante predicción inter de la imagen disponible o se determinó previamente que eran utilizables para la codificación mediante predicción inter de imágenes que siguen a la imagen disponible en

orden de codificación, y

en el que determinar una o más imágenes de referencia candidatas comprende determinar una o más imágenes de referencia candidatas de uno o más del primer subconjunto de imágenes de referencia, el segundo subconjunto de imágenes de referencia, el tercer subconjunto de imágenes de referencia, el cuarto subconjunto de imágenes de referencia y el quinto subconjunto de imágenes de referencia.

3. El procedimiento según la reivindicación 2, en el que determinar una o más imágenes de referencia candidatas comprende la determinación de una o más imágenes de referencia candidatas de los subconjuntos de imágenes de referencia primero, segundo y tercero, y no de los subconjuntos de imágenes de referencia cuarto y quinto.

4. El procedimiento según la reivindicación 2, en el que determinar una o más imágenes de referencia candidatas comprende determinar una o más imágenes de referencia candidatas a partir del primer y segundo subconjuntos de imágenes de referencia, y no del tercer, cuarto y quinto subconjuntos de imágenes de referencia.

5. El procedimiento según la reivindicación 1, que comprende además:

determinar si la imagen identificada está disponible para codificar la imagen actual; y

codificar mediante predicción intra la imagen actual basándose en una determinación de que la imagen identificada no está disponible para codificar la imagen actual.

6. El procedimiento según la reivindicación 1, que comprende además:

recibir información que identifica una imagen perdida; y

excluir la imagen perdida de la una o más imágenes de referencia candidatas determinadas.

7. El procedimiento según la reivindicación 1, en el que la codificación mediante predicción inter de la imagen actual basándose en la una o más imágenes de referencia seleccionadas comprende una codificación mediante predicción bidireccional de la imagen actual basándose en dos imágenes de las imágenes de referencia seleccionadas.

8. Un dispositivo para codificar datos de vídeo, comprendiendo el dispositivo:

una primera memoria intermedia de imágenes descodificadas (DPB) que comprende una memoria configurada para almacenar una o más imágenes; y

un codificador de vídeo que comprende uno o más procesadores, en el que el codificador de vídeo está configurado para:

almacenar un conjunto de parámetros para cada imagen que está codificada, el conjunto de parámetros de una imagen particular que incluye uno o más subconjuntos de imágenes de referencia que forman un conjunto de imágenes de referencia de la imagen particular;

recibir, desde un dispositivo que comprende un descodificador de vídeo, información que identifica una imagen previamente descodificada que se almacena en una segunda DPB del descodificador de vídeo en el momento en que el descodificador de vídeo transmitió la información;

determinar, basándose en el conjunto de parámetros de la imagen identificada, una o más imágenes de referencia candidatas que están disponibles para codificar una imagen actual, en el que las una o más imágenes de referencia candidatas incluyen una o más imágenes que son:

en el conjunto de imágenes de referencia de la imagen identificada, en el que el conjunto de imágenes de referencia identifica las imágenes que se determinaron previamente, durante la codificación de la imagen identificada, como imágenes que podrían haberse utilizado para codificar mediante predicción inter la imagen identificada o como imágenes que podrían haberse utilizado para la codificación mediante predicción inter de las imágenes que siguen a la imagen identificada en orden de codificación; y

que aún se almacenan en la primera DPB;

seleccionar una o más imágenes de referencia para la imagen actual de una o más imágenes de referencia candidatas determinadas y la imagen identificada; y

codificar mediante predicción inter la imagen actual basándose en la una o más imágenes de referencia seleccionadas.

5 **9.** El dispositivo de la reivindicación 8, en el que el codificador de vídeo está configurado para:

construir, durante la codificación de una imagen disponible que luego es identificada mediante el descodificador de vídeo y antes de la codificación de la imagen actual, uno o más de:

10 un primer subconjunto de imágenes de referencia que incluye imágenes de referencia a corto plazo que preceden a la imagen disponible en el orden de visualización y que se determinó que se pueden usar para realizar predicción inter de la imagen disponible y se determinó que se pueden usar para realizar predicción inter de la una o más imágenes que siguen a la imagen disponible en orden de codificación;

15 un segundo subconjunto de imágenes de referencia que incluye imágenes de referencia a corto plazo que siguen a la imagen disponible en el orden de visualización y que se determinó que se pueden usar para realizar predicción inter de la imagen disponible y se determinó que se pueden usar para realizar predicción inter una o más imágenes que siguen a la imagen disponible en orden de codificación;

20 un tercer subconjunto de imágenes de referencia que incluye cualquier imagen de referencia a largo plazo que se determinó que era utilizable para realizar predicción inter de la imagen disponible y que se determinó que era utilizable para realizar predicción inter de la una o más imágenes que siguen a la imagen disponible en orden de codificación;

25 un cuarto subconjunto de imágenes de referencia que incluye imágenes de referencia a corto plazo que se determinó que se pueden usar para realizar predicción inter de la una o más imágenes que siguen a la imagen disponible en orden de codificación, y se determinó que no se pueden usar para realizar predicción inter de la imagen disponible ; y

30 un quinto subconjunto de imágenes de referencia que incluye imágenes de referencia a largo plazo que se determinó que se pueden usar para realizar predicción inter de la una o más imágenes que siguen a la imagen disponible en orden de codificación, y se determinó que no se pueden usar para realizar predicción inter de las imágenes disponibles,

35 en el que las imágenes en uno o más de los primeros, segundo, tercero, cuarto y quinto subconjuntos de imágenes de referencia comprenden imágenes que previamente se determinó que eran utilizables para la codificación mediante predicción inter de la imagen disponible o se determinó previamente que eran utilizables para la codificación mediante predicción inter de imágenes que siguen a la imagen disponible en orden de codificación, y

40 en el que, para determinar una o más imágenes de referencia candidatas, el codificador de vídeo está configurado para determinar una o más imágenes de referencia candidatas de una o más del primer subconjunto de imágenes de referencia, el segundo subconjunto de imágenes de referencia, el tercer subconjunto de imágenes de referencia, la cuarta referencia subconjunto de imágenes, y el quinto subconjunto de imágenes de referencia.

50 **10.** El dispositivo según la reivindicación 9, en el que, para determinar una o más imágenes de referencia candidatas, el codificador de vídeo está configurado para determinar una o más imágenes de referencia candidatas del primer, segundo y tercer subconjuntos de imágenes de referencia, y no del cuarto y quinto subconjuntos de imágenes de referencia.

55 **11.** El dispositivo de la reivindicación 9, en el que, para determinar una o más imágenes de referencia candidatas, el codificador de vídeo está configurado para determinar una o más imágenes de referencia candidatas de los primer y segundo subconjuntos de imágenes de referencia, y no del tercer, cuarto y quinto subconjuntos de imágenes de referencia.

12. El dispositivo de la reivindicación 8, en el que el codificador de vídeo está configurado para:

60 determinar si la imagen identificada está disponible para codificar la imagen actual; y

codificar mediante predicción intra la imagen actual basándose en la determinación de que la imagen identificada no está disponible para codificar la imagen actual.

65 **13.** El dispositivo de la reivindicación 8, en el que el codificador de vídeo está configurado para:

recibir información identificando una imagen perdida; y

excluir la imagen perdida de la una o más imágenes de referencia candidatas determinadas.

- 5 **14.** El dispositivo de la reivindicación 8, en el que para codificar mediante predicción inter la imagen actual basándose en la una o más imágenes de referencia seleccionadas, el codificador de vídeo está configurado para codificar mediante predicción bidireccional la imagen actual basándose en dos imágenes de las imágenes de referencia seleccionadas.

10

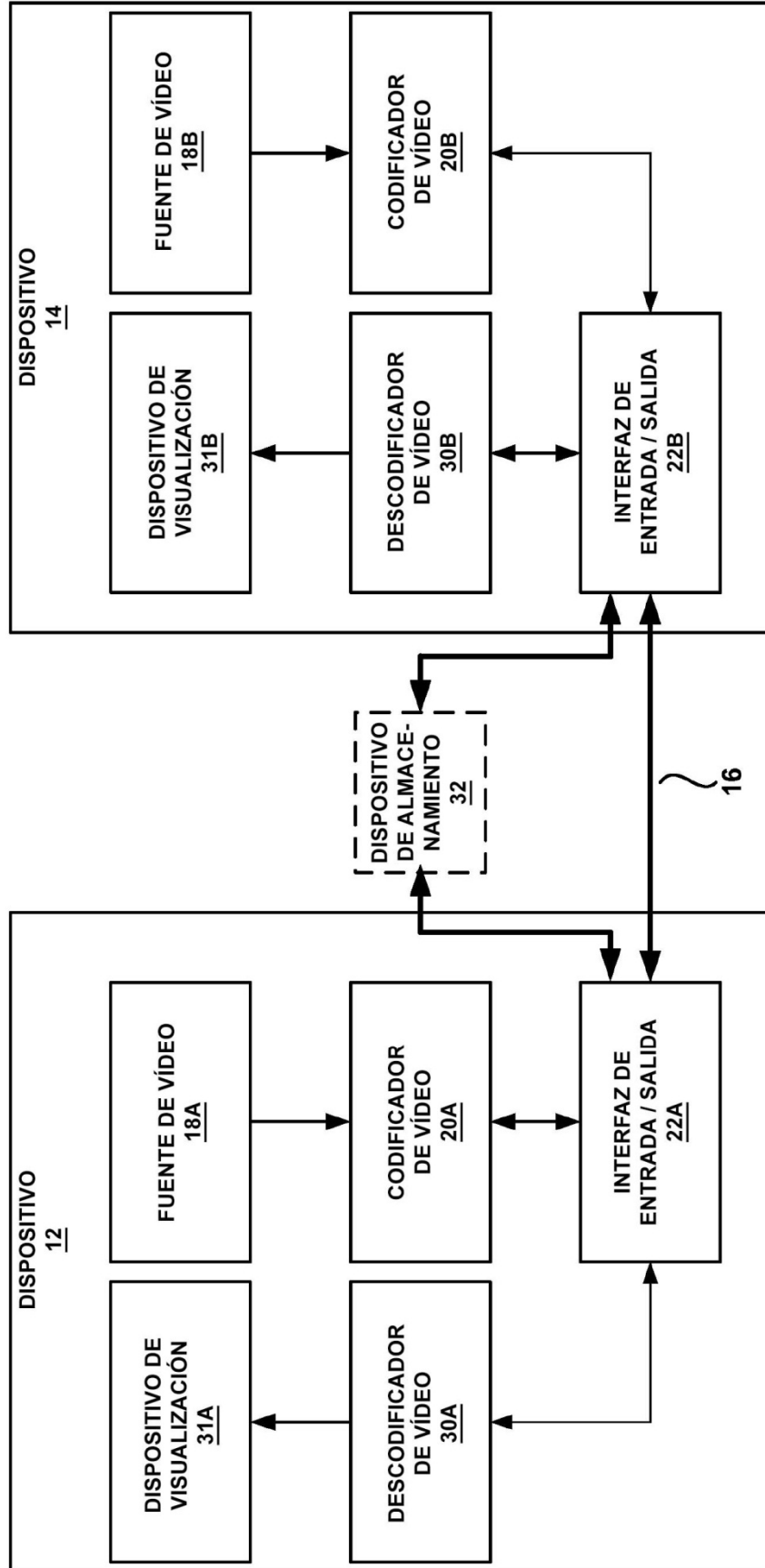


FIG. 1

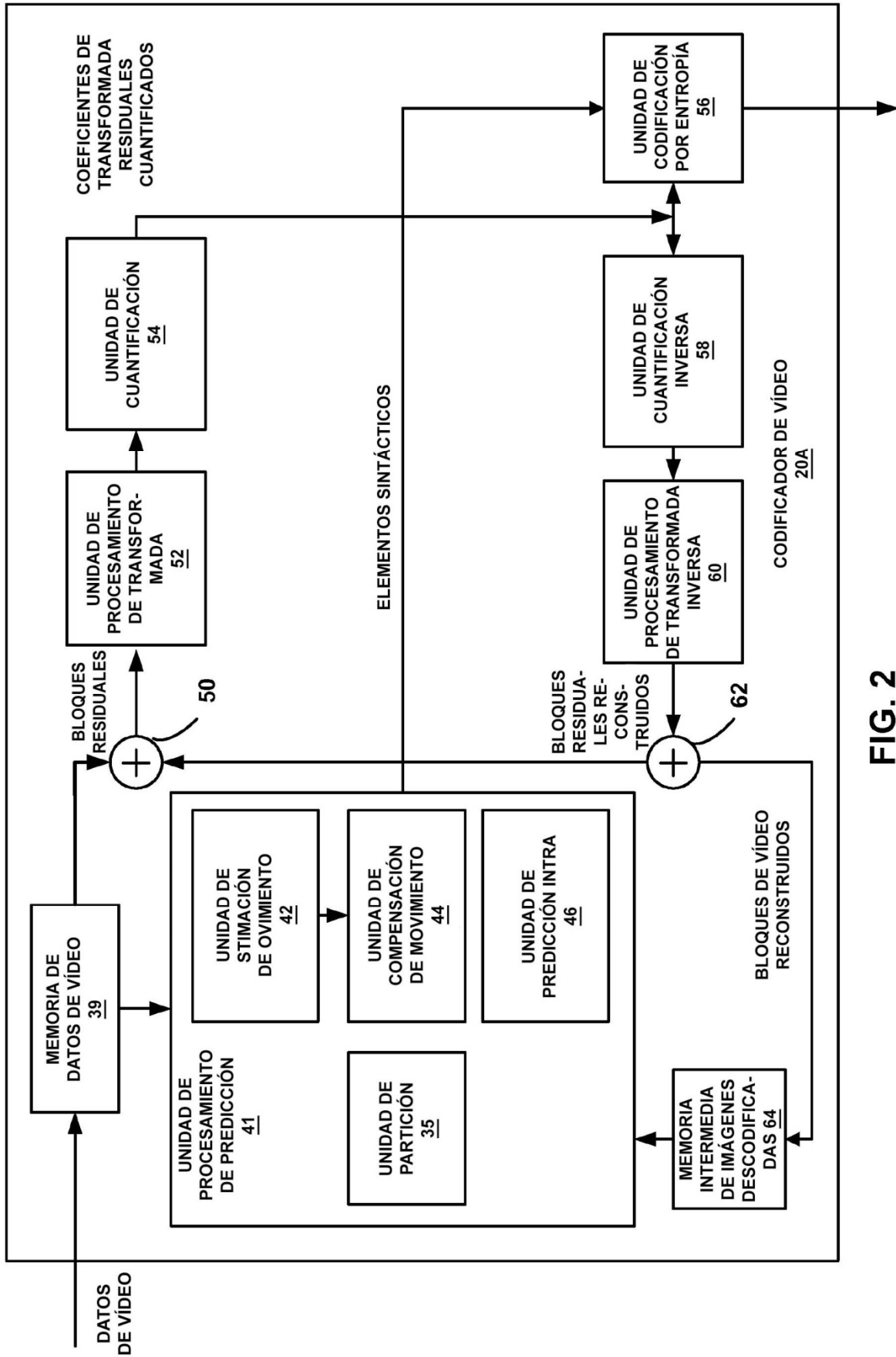


FIG. 2

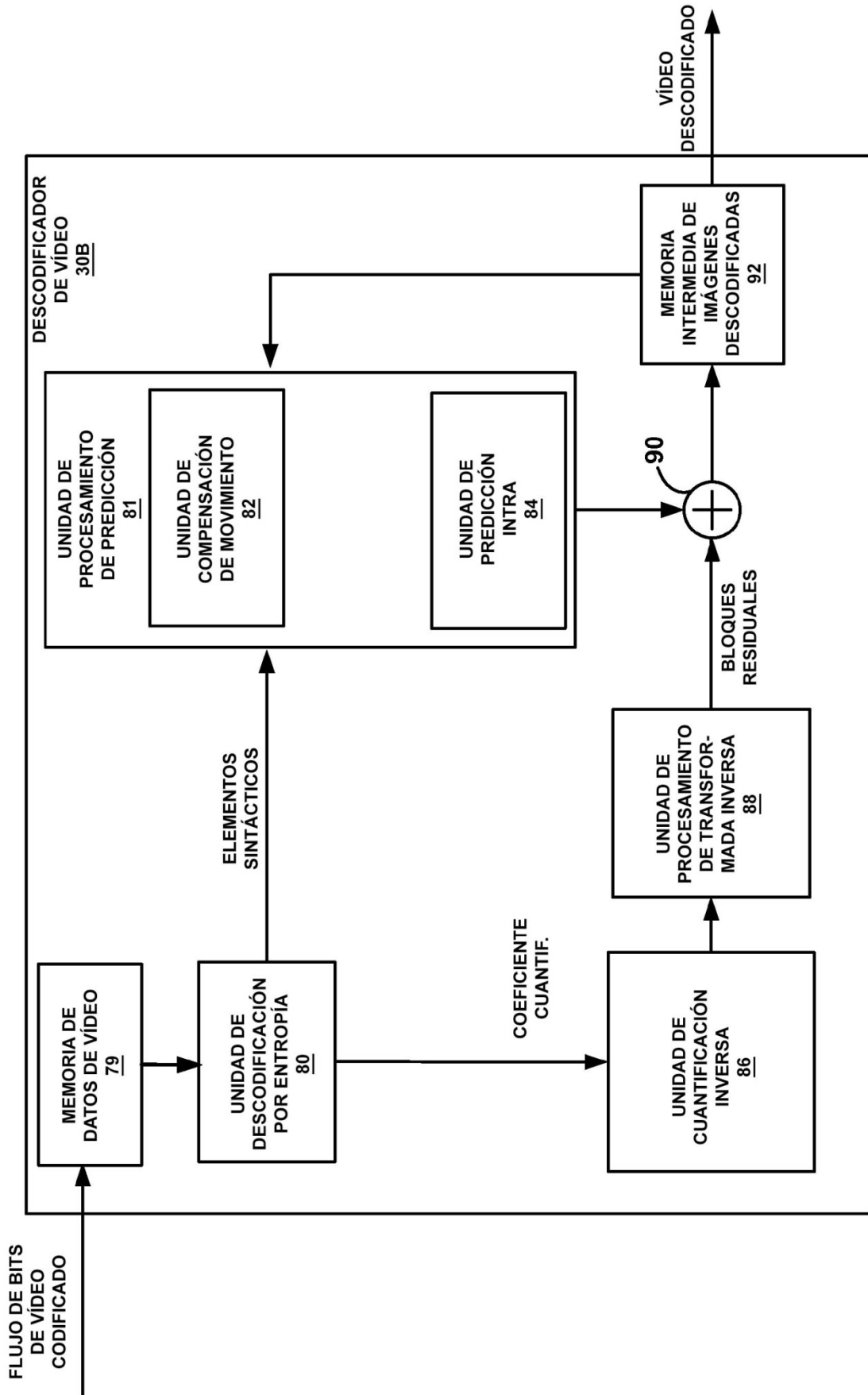


FIG. 3

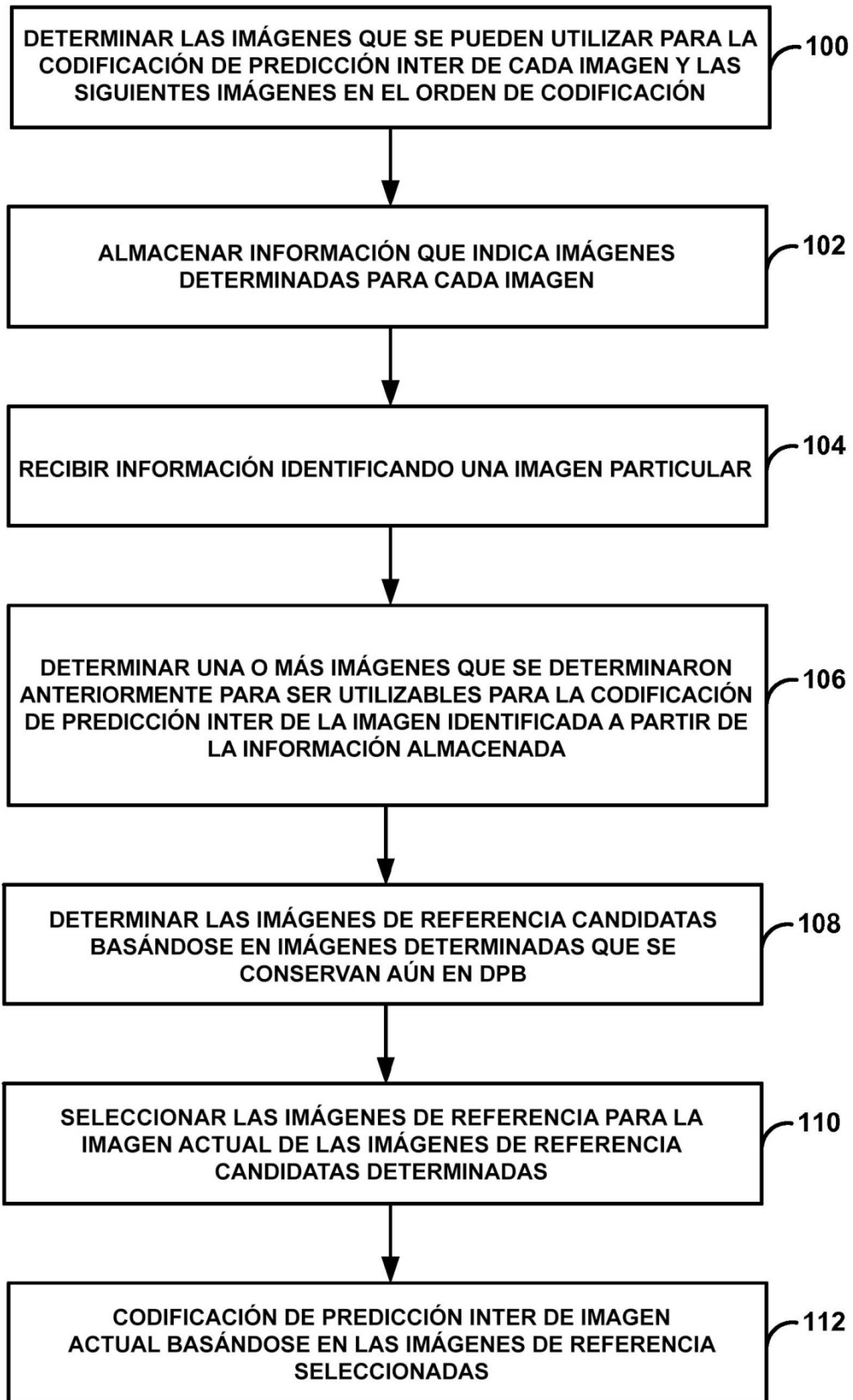


FIG. 4

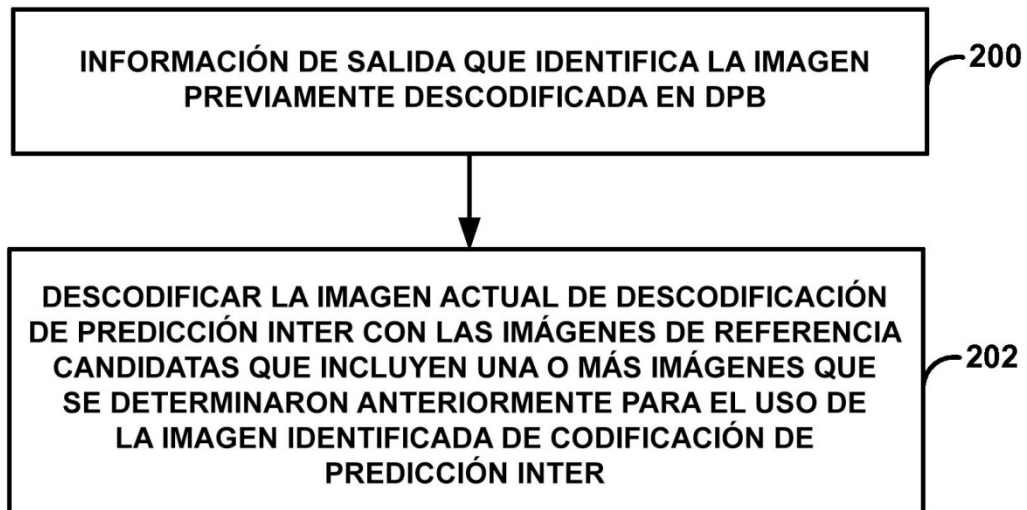


FIG. 5

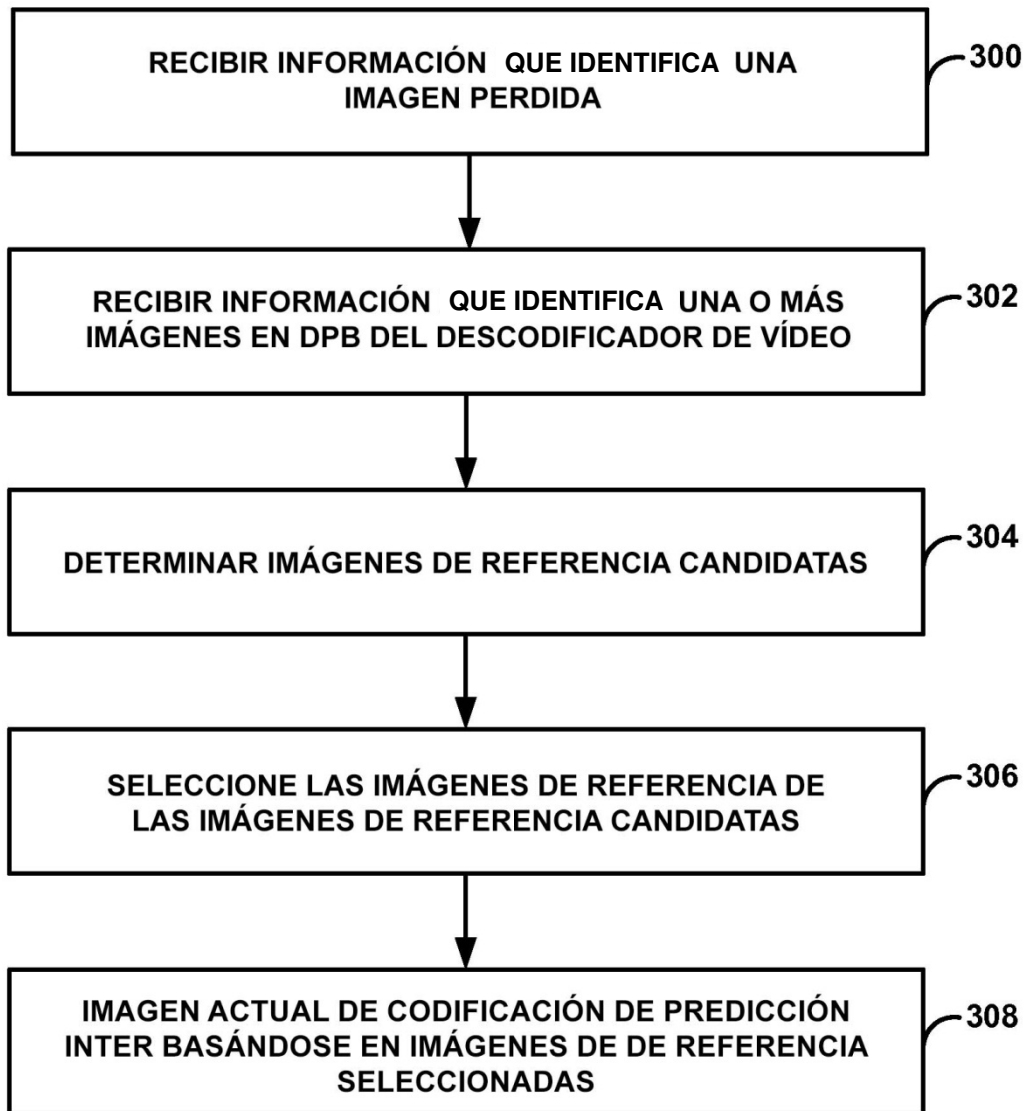


FIG. 6

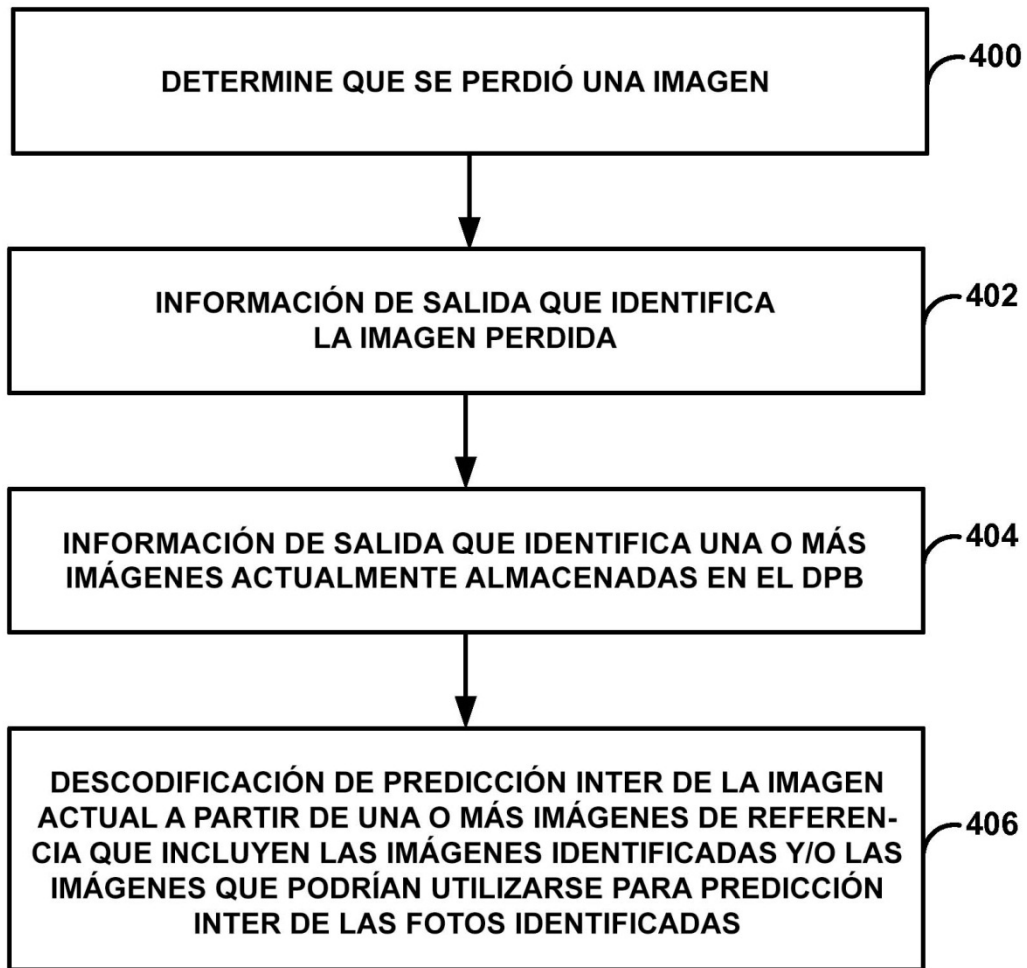


FIG. 7