

19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 754 056**

51 Int. Cl.:

H04N 19/597 (2014.01)

H04N 19/70 (2014.01)

H04N 19/30 (2014.01)

H04N 19/44 (2014.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

86 Fecha de presentación y número de la solicitud internacional: **14.10.2014 PCT/US2014/060486**

87 Fecha y número de publicación internacional: **23.04.2015 WO15057706**

96 Fecha de presentación y número de la solicitud europea: **14.10.2014 E 14795710 (4)**

97 Fecha y número de publicación de la concesión europea: **07.08.2019 EP 3058746**

54 Título: **Sistemas y procedimientos para la derivación de RPS entre capas basada en la dependencia de predicción de referencia de subcapa**

30 Prioridad:

14.10.2013 US 201361890759 P
13.10.2014 US 201414513089

45 Fecha de publicación y mención en BOPI de la traducción de la patente:
15.04.2020

73 Titular/es:

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

72 Inventor/es:

RAPAKA, KRISHNAKANTH;
WANG, YE-KUI;
CHEN, JIANLE y
HENDRY, FNU

74 Agente/Representante:

FORTEA LAGUNA, Juan José

ES 2 754 056 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

Sistemas y procedimientos para la derivación de RPS entre capas basada en la dependencia de predicción de referencia de subcapa

5

CAMPO TÉCNICO

[0001] Esta divulgación se refiere al campo de la codificación de vídeo ajustable a escala (SVC). Más en particular, está relacionada con la SVC basada en la HEVC (HEVC-SVC) y las extensiones de la HEVC.

10

ANTECEDENTES

[0002] Las capacidades del vídeo digital pueden incorporarse a una amplia gama de dispositivos, que incluye televisores digitales, sistemas de difusión digital directa, sistemas de difusión inalámbrica, asistentes digitales personales (PDA), ordenadores portátiles o de escritorio, ordenadores 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 de vídeo en tiempo real y similares. Los dispositivos de vídeo digital implementan técnicas de compresión de vídeo, tales como las que se describen en varias normas de codificación de vídeo, como se describe a continuación (por ejemplo, la codificación de vídeo de alta eficacia (HEVC)). Los dispositivos de vídeo pueden transmitir, recibir, codificar, decodificar y/o almacenar información de vídeo digital con más eficacia implementando dichas técnicas de compresión de vídeo. En CHEN J ET AL: "MV-HEVC/SHVC HLS: On Inter layer Prediction Signaling", 5. REUNIÓN JCT-3V; 27-7-2013 - 2-8-2013; VIENA se proporciona un mecanismo para evitar la señalización de predicción entre capas de encabezado de segmento en dos aspectos. El primer aspecto es cubrir dos configuraciones de codificación típicas en MV-HEVC y SHVC. El primer típico es que las imágenes de todas las capas de referencia directa especificadas por el VPS se incluyen en el subconjunto RPS entre capas. La segunda configuración es que solo la imagen de la capa de referencia directa más alta especificada por el VPS se incluye en el subconjunto RPS entre capas. Además, se propone evitar la señalización de las ID de capa de las imágenes que se incluirán en el subconjunto RPS entre capas cuando el número total es igual al número de capas de referencia directa especificadas por el VPS, como se conoce la información en este caso. El segundo aspecto es señalar un conjunto de conjuntos de imágenes de referencia de predicción entre capas candidatas (ILP-RPS) en el VPS y se puede indicar un candidato en un encabezado de corte que se va a usar para la derivación del ILP-RPS para la imagen.

15

20

25

30

35

SUMARIO

[0003] Cada uno de los sistemas, procedimientos y dispositivos de esta divulgación tiene varios aspectos innovadores, ninguno de los cuales es el único responsable de los atributos deseables divulgados en el presente documento. Los detalles de uno o más ejemplos se exponen en los dibujos adjuntos y en la siguiente descripción. Otras características, objetos y ventajas resultarán evidentes a partir de la descripción y de los dibujos, y a partir de las reivindicaciones.

40

[0004] Un aspecto de la materia objeto descrita en la divulgación proporciona un codificador de vídeo que comprende una memoria configurada para almacenar una o más imágenes de capa de referencia directa de una o más imágenes actuales en una secuencia, en el que la una o más imágenes actuales están asociadas con una capa actual, estando asociada la capa actual con una o más capas de referencia directa, y en el que una capa de referencia directa es una capa que puede usarse para la predicción entre capas de otra capa, tal como la capa actual. El codificador de vídeo comprende además un procesador en comunicación con la unidad de memoria. La unidad de memoria está configurada para determinar que una imagen de capa de referencia directa de una o más imágenes de capa de referencia directa no está restringida para su uso, de modo que la imagen de capa de referencia directa se puede usar, en la predicción entre capas, cuando la imagen de la capa de referencia directa tiene un valor de identificación temporal menor que un umbral predeterminado basado en un valor de identificación temporal máximo asociado con la capa actual. Todas las una o más imágenes de capa de referencia directa de la imagen actual que no están restringidas para su uso están disponibles para su uso en la predicción entre capas. Además, el procesador está configurado para determinar si todas las una o más imágenes de capa de referencia directa de la imagen actual que no están restringidas para uso en la predicción entre capas se incluyen en un conjunto de imágenes de referencia entre capas asociado con la imagen actual. Además, el procesador está configurado para establecer un indicador asociado con la imagen actual a un primer valor basado en la determinación de que todas las imágenes de capa de referencia directa de la imagen actual que no están restringidas para su uso en la predicción entre capas se incluyen en el conjunto de imágenes de referencia entre capas asociado con la imagen actual; o establecer el indicador en un segundo valor en función de la determinación de que no todas las imágenes de capa de referencia directa de la imagen actual que no están restringidas para su uso en la predicción entre capas se incluyen en el conjunto de imágenes de referencia entre capas asociado con la imagen actual.

45

50

55

60

65

[0005] Otro aspecto del tema en cuestión, descrito en la divulgación, proporciona un procedimiento de codificación de vídeo. El procedimiento comprende almacenar una o más imágenes de capa de referencia directa de una o más

imágenes actuales en una secuencia, en el que la una o más imágenes actuales están asociadas con una capa actual, estando asociada la capa actual con una o más capas de referencia directa, y en el que una capa de referencia directa es una capa que puede usarse para la predicción entre capas de otra capa, tal como la capa actual. El procedimiento comprende además una etapa de determinar que una imagen de capa de referencia directa de una o más imágenes de capa de referencia directa no está restringida para su uso, de modo que la imagen de capa de referencia directa se puede usar, en la predicción entre capas, cuando la imagen de la capa de referencia directa tiene un valor de identificación temporal menor que un umbral predeterminado basado en un valor de identificación temporal máximo asociado con la capa actual. Todas las una o más imágenes de capa de referencia directa de la imagen actual que no están restringidas para su uso están disponibles para su uso en la predicción entre capas. El procedimiento comprende además determinar si todas las una o más imágenes de capa de referencia directa de la imagen actual que no están restringidas para uso en la predicción entre capas se incluyen en un conjunto de imágenes de referencia entre capas asociado con la imagen actual. Además, el procedimiento implica establecer un indicador asociado con la imagen actual a un primer valor basado en la determinación de que todas las imágenes de capa de referencia directa de la imagen actual que no están restringidas para su uso en la predicción entre capas se incluyen en el conjunto de imágenes de referencia entre capas asociado con la imagen actual; o establecer el indicador en un segundo valor en función de la determinación de que no todas las imágenes de capa de referencia directa de la imagen actual que no están restringidas para su uso en la predicción entre capas se incluyen en el conjunto de imágenes de referencia entre capas asociado con la imagen actual. Otro aspecto del tema en cuestión, descrito en la divulgación, proporciona un decodificador de vídeo que comprende una unidad de memoria. La unidad de memoria está configurada para almacenar una o más imágenes de capa de referencia directa, en el que cada una de las imágenes de capa de referencia directa está asociada con una capa de referencia directa respectiva de una o más capas de referencia directa, en el que la una o más capas de referencia directa están asociadas con una capa actual, y en el que la capa actual está asociada con una imagen actual y en el que una capa de referencia directa es una capa que puede usarse para la predicción entre capas de otra capa, tal como la capa actual. El decodificador de vídeo comprende además un procesador en comunicación con la memoria. El procesador está configurado para determinar, en función de un indicador, si todas las una o más imágenes de capa de referencia directa que no están restringidas para su uso en la predicción entre capas se incluyen en un conjunto de imágenes de referencia entre capas asociado con la imagen actual. Además, el procesador está configurado para determinar una cantidad de una o más imágenes de capa de referencia directa que no están restringidas para su uso en la predicción entre capas en respuesta a la determinación de que todas las una o más imágenes de capa de referencia directa que no están restringidas para el uso en la predicción entre capas se incluyen en el conjunto de imágenes de referencia entre capas asociado con la imagen actual, en el que el número determinado de una o más imágenes de capa de referencia directa se puede usar en la predicción entre capas. El procesador está configurado además para determinar que una imagen de capa de referencia directa de una o más imágenes de capa de referencia directa no está restringida para su uso, de modo que la imagen de capa de referencia directa se puede usar, en la predicción entre capas, determinando que el valor de identificación asociado con la imagen de capa de referencia directa es menor que un umbral predeterminado basado en un valor de identificación temporal máximo asociado con la capa actual.

[0006] Otro aspecto del tema en cuestión, descrito en la divulgación, proporciona un procedimiento para decodificar vídeo. El procedimiento comprende almacenar una o más imágenes de capa de referencia directa, en el que cada una de las imágenes de capa de referencia directa está asociada con una capa de referencia directa respectiva de una o más capas de referencia directa, en el que la una o más capas de referencia directa están asociadas con una capa actual, y en el que la capa actual está asociada con una imagen actual y en el que una capa de referencia directa es una capa que puede usarse para la predicción entre capas de otra capa, tal como la capa actual. El procedimiento comprende además determinar, basándose en un indicador, si todas las una o más imágenes de capa de referencia directa que no están restringidas para uso en la predicción entre capas se incluyen en un conjunto de imágenes de referencia entre capas asociado con la imagen actual. Además, el procedimiento comprende determinar un número de una o más imágenes de capa de referencia directa que no están restringidas para su uso en la predicción entre capas en respuesta a la determinación de que todas las una o más imágenes de capa de referencia directa que no están restringidas para su uso en la predicción entre capas se incluyen en el conjunto de imágenes de referencia entre capas asociado con la imagen actual, en el que el número determinado de una o más imágenes de capa de referencia directa se puede usar en la predicción entre capas. El procedimiento comprende además determinar que una imagen de capa de referencia directa de una o más imágenes de capa de referencia directa no está restringida para su uso, de modo que la imagen de capa de referencia directa puede usarse, en la predicción entre capas, determinando que un valor de identificación temporal asociado con la imagen de la capa de referencia directa es menor que un umbral predeterminado basado en un valor de identificación temporal máximo asociado con la capa actual, en el que todas las una o más imágenes de la capa de referencia directa que no están restringidas para su uso están disponibles para su uso en predicción entre capas.

BREVE DESCRIPCIÓN DE LOS DIBUJOS

[0007]

La figura 1A es un diagrama de bloques que ilustra un sistema ejemplar de codificación y decodificación de vídeo que puede utilizar las técnicas de derivación de conjunto de imágenes de referencia (RPS) entre capas, descritas en la presente divulgación.

La figura 1B es un diagrama de bloques que ilustra otro sistema ejemplar de codificación y decodificación de vídeo que puede utilizar las técnicas de derivación de conjunto de imágenes de referencia entre capas, descritas en la presente divulgación.

5 La figura 2 es un diagrama de bloques que ilustra un codificador ejemplar de vídeo que puede implementar las técnicas de derivación de conjunto de imágenes de referencia entre capas, descritas en la presente divulgación.

10 La figura 3 es un diagrama de bloques que ilustra un decodificador ejemplar de vídeo que puede implementar las técnicas de derivación de conjunto de imágenes de referencia entre capas, descritas en la presente divulgación.

15 La figura 4 ilustra un ejemplo de información de vídeo que puede procesarse para indicar si un conjunto de imágenes de referencia entre capas incluye todas las imágenes de capa de referencia directa que no están restringidas para su uso en la predicción entre capas.

La figura 5 ilustra un diagrama de flujo de un procedimiento para un procesador de un codificador de vídeo para indicar si un conjunto de imágenes de referencia entre capas incluye todas las imágenes de capa de referencia directa que no están restringidas para su uso en la predicción entre capas.

20 La figura 6 ilustra un ejemplo de información de vídeo que puede procesarse para determinar varias imágenes de capa de referencia activa a partir de una o más imágenes de capa de referencia directa.

25 La figura 7 ilustra un diagrama de flujo de un procedimiento para que un procesador de un decodificador de vídeo determine un número de imágenes de capa de referencia activa a partir de una o más imágenes de capa de referencia directa.

DESCRIPCIÓN DETALLADA

30 **[0008]** La descripción detallada expuesta a continuación, en relación con los dibujos adjuntos, está concebida como una descripción de modos de realización ejemplares de la invención, y no está concebida para representar los únicos modos de realización en los que la invención puede llevarse a la práctica. El término "ejemplar" usado a lo largo de esta descripción significa "que sirve de ejemplo, caso o ilustración" y no debería interpretarse necesariamente como preferente o ventajoso con respecto a otros modos de realización ejemplares. La descripción detallada incluye detalles específicos con el fin de facilitar una plena comprensión de los modos de realización ejemplares de la presente invención. En algunos casos, algunos dispositivos se muestran en forma de diagrama de bloques.

40 **[0009]** Aunque, con fines de simplicidad de la explicación, las metodologías se muestran y se describen como una serie de actos, se entenderá y apreciará que las metodologías no están limitadas por el orden de los actos, ya que algunos actos, de acuerdo con uno o más aspectos, se producen en diferente orden y/o de manera concurrente con otros actos, a partir de lo que se muestra y describe en el presente documento. Por ejemplo, los expertos en la técnica entenderán y apreciarán que una metodología podría representarse de forma alternativa como una serie de estados o sucesos interrelacionados, tal como en un diagrama de estados. Además, tal vez no se requieran todos los actos ilustrados para implementar una metodología de acuerdo a uno o más aspectos.

45 **[0010]** Las normas de codificación de vídeo incluyen aquellas definidas por ITU-T H.261, ISO/IEC MPEG-1 Visual, ITU-T H.262 o ISO/IEC MPEG-2 Visual, ISO/IEC MPEG-4 Visual, ITU-T H. 263, ISO/IEC MPEG-4 Visual y ITU-T H.264 (también conocida como ISO/IEC MPEG-4 AVC), incluidas sus extensiones de codificación de vídeo ajustable a escala (SVC) y codificación de vídeo de vista múltiple (MVC), y la parte 10, Codificación de vídeo avanzada (AVC), la norma de codificación de vídeo de alta eficacia (HEVC) actualmente en desarrollo, y extensiones de dichas normas. La HEVC ha sido recientemente desarrollada por el Equipo Conjunto de Colaboración en Codificación de Vídeo (JCT-VC) del Grupo de Expertos en Codificación de Vídeo (VCEG) de la ITU-T y del Grupo de Expertos en Imágenes en Movimiento (MPEG) de la ISO/IEC. El borrador de especificación de la HEVC más reciente ("HEVC WD") está disponible en http://phenix.int-evry.fr/jct/doc_end_user/documents/14_Vienna/wg11/JCTVC-N1003-v1.zip. El JCT-3V está desarrollando la extensión de vista múltiple para la HEVC (MV-HEVC) y una extensión avanzada de codificación de vídeo tridimensional de la HEVC (3D-HEVC). Además, la extensión ajustable a escala de la codificación de vídeo para la HEVC (SHVC) también está siendo desarrollada por el JCT-VC. Un reciente Borrador de Trabajo de la MV-HEVC ("MV-HEVC WD5"), está disponible en http://phenix.it-sudparis.eu/jct2/doc_end_user/documents/5_Vienna/wg11/JCT3V-E1004-v6.zip. El más reciente borrador de trabajo de la 3D-HEVC ("3D-HEVC WD1") está disponible en http://phenix.it-sudparis.eu/jct2/doc_end_user/documents/5_Vienna/wg11/JCT3V-E1001-v3.zip. Además, un reciente borrador de trabajo de la SHVC ("SHVC WD3") está disponible en http://phenix.it-sudparis.eu/jct/doc_end_user/documents/14_Vienna/wg11/JCTVC-N1008-v3.zip.

65 **[0011]** Como se describió anteriormente, ciertos aspectos de la codificación de vídeo (por ejemplo, en extensiones HEVC) incluyen derivar un conjunto de imágenes de referencia entre capas para su uso en la predicción entre capas de imágenes actuales en las capas actuales. Por ejemplo, un codificador de vídeo puede analizar una serie de

imágenes de referencia asociadas con una imagen actual en una capa actual, de modo que el codificador de vídeo puede proporcionar información utilizable por un decodificador de vídeo que puede permitir que el decodificador de vídeo obtenga el conjunto de imágenes de referencia entre capas para la imagen actual y/o la capa actual. Por ejemplo, en el nivel del conjunto de parámetros de vídeo (VPS), el codificador de vídeo puede proporcionar una indicación (por ejemplo, un elemento de sintaxis y/o un indicador establecido en un valor de "0" o "1") que indica al decodificador de vídeo que todas las imágenes de una capa de referencia directa de la capa actual pueden usarse (por ejemplo, incluidas dentro) de un conjunto de imágenes de referencia (RPS). Al proporcionar ciertas indicaciones (por ejemplo, establecer ciertos indicadores en un valor de "0" o "1"), el codificador de vídeo puede permitir (o habilitar) el decodificador de vídeo para determinar cómo se asignará la memoria para todas las imágenes en la secuencia de vídeo asociada y puede permitir (o habilitar) el decodificador de vídeo para insertar cada una de las imágenes en el conjunto de imágenes de referencia (RPS).

[0012] Por ejemplo, los Borradores de trabajo recientes de MV-HEVC WD5 y SHVC WD3 incluyen el indicador *all_ref_layers_active_flag* en el conjunto de parámetros de vídeo (VPS), con la siguiente semántica: "*all_ref_layers_active_flag* igual a 1 especifica que para cada imagen que hace referencia al SPV, las imágenes de capa de referencia de todas las capas de referencia directa de la capa que contiene la imagen están presentes en la misma unidad de acceso que la imagen y se incluyen en el conjunto de imágenes de referencia entre capas de la imagen. *all_ref_layers_active_flag* igual a 0 especifica que la restricción anterior puede o no aplicarse". Cuando el codificador de vídeo establece *all_ref_layers_active_flag* en un valor de "1" y señala *all_ref_layers_active_flag* al decodificador de vídeo, el decodificador de vídeo puede derivar una variable que especifica un número de imágenes de referencia entre capas incluidas en el conjunto de imágenes de referencia entre capas (RPS) para una imagen asociada (por ejemplo, el decodificador de vídeo puede determinar varias imágenes de capa de referencia activa). Los borradores de trabajo recientes de MV-HEVC WD5 y SHVC WD3 incluyen la variable *NumActiveRefLayerPics* para este propósito. El siguiente código muestra cómo se deriva la variable *NumActiveRefLayerPics* en los borradores de trabajo recientes de MV-HEVC WD5 y SHVC WD3:

```

if(nuh_layer_id == 0 || NumDirectRefLayers[numh_layer_id] == 0)
    NumActiveRefLayerPics = 0;
else if(all_ref_layers_active_flag)
    NumActiveRefLayerPics = NumDirectRefLayers[numh_layer_id];
else if(! inter_layer_pred_enabled_flag)
    NumActiveRefLayerPics = 0;
else if(max_one_active_ref_layer_flag || NumDirectRefLayers[
    nuh_layer_id] == 1)
    NumActiveRefLayerPics = 1;
else
    NumActiveRefLayerPics = num_inter_layer_ref_pics_minus1 + 1;

```

[0013] Al usar *all_ref_layers_active_flag* de la manera descrita anteriormente, se puede guardar algo de sobrecarga de señalización a nivel de segmento durante la derivación de un RPS entre capas. Por ejemplo, la sobrecarga de señalización se puede guardar cuando cada una de las imágenes de capa de referencia directa (a veces denominadas imágenes de capa de referencia) de la capa actual (la capa que contiene la imagen actual) está presente en la misma unidad de acceso que la imagen actual y está incluida en el conjunto de imágenes de referencia entre capas de la imagen actual. Una capa de referencia directa es una capa que se puede usar para la predicción entre capas de otra capa, tal como la capa actual. Una imagen de capa de referencia directa (o imagen de capa de referencia) es una imagen en una capa de referencia directa que se utiliza para la predicción entre capas de la imagen actual y está en la misma unidad de acceso que la imagen actual. Tales ahorros generales de señalización pueden ocurrir en escenarios de codificación de vídeo de múltiples vistas cuando las imágenes de la capa de referencia directa especificadas en el SPV están presentes en cada unidad de acceso, las imágenes de la capa de referencia directa especificadas en el SPV se usan para la predicción entre capas, y las imágenes de capa de referencia directa especificadas en el VPS se insertan en al menos una lista de imágenes de referencia de la imagen actual. Otro escenario en el que esto puede ocurrir es durante la codificación de vídeo escalable, donde es común que cada capa de mejora tenga solo una capa de referencia directa y que las imágenes de la capa de referencia directa estén presentes en cada unidad de acceso y se usen para la predicción entre capas.

[0014] Sin embargo, los procedimientos actuales de extensión HEVC para señalización y derivación de RPS descritos anteriormente tienen ciertos inconvenientes. Por ejemplo, el proceso actual no permite que el codificador de vídeo proporcione al decodificador de vídeo la indicación de ahorro de sobrecarga de señalización como se describió anteriormente si alguna de las imágenes asociadas en al menos una subcapa tiene un uso restringido en la predicción entre capas. Más específicamente, los borradores de trabajo recientes de MV-HEVC WD5 y SHVC WD3 requieren que *all_ref_layers_active_flag* se establezca en un valor de cero (es decir, una restricción que requiere que para cada imagen que se refiere al SPV, las imágenes de capa de referencia de todas las referencias directas las capas de la capa que contiene la imagen están presentes en la misma unidad de acceso que la imagen y se incluyen en el conjunto de imágenes de referencia entre capas de la imagen que puede aplicarse o no) cuando alguna de las imágenes asociadas tiene un valor de identificación temporal (TID o TemporalID) mayor que un valor de identificación temporal máximo (Max TID), por ejemplo, *max_tid_il_ref_pics_plus1* [i]. En otras palabras, una restricción de predicción

temporal entre capas de subcapa se indica mediante $max_tid_il_ref_pics_plus1 [i] - 1$, para cualquier valor de i en el rango de 0 a $vps_max_layers_minus1$, inclusive, menor o igual que el mayor valor de $TemporalId$ de todas las unidades de capa de abstracción de red (NAL) de la capa de codificación de vídeo (VCL) en las secuencias de codificación de vídeo (CVS) que se refieren al VPS. Estas limitaciones pueden causar enfoques actuales para desperdiciar recursos de codificación (por ejemplo, al tener altos costos de memoria, altos costos de señalización, etc.) y disminuir la utilidad de los indicadores y variables descritos anteriormente y/o disminuir la utilidad de cualquier otro elemento de sintaxis que de lo contrario, se puede utilizar para optimizar la memoria. En vista de tales limitaciones, sería beneficioso permitir que un codificador de vídeo proporcione un decodificador de vídeo con las indicaciones de ahorro de sobrecarga de señalización descritas anteriormente con mayor frecuencia, por ejemplo, incluso si algunas de las imágenes asociadas en al menos una subcapa está restringida en la predicción entre capas.

[0015] Por lo tanto, la presente divulgación describe técnicas para la derivación de conjuntos de imágenes de referencia entre capas (RPS) usando indicaciones y/o determinaciones relacionadas con ciertas propiedades de imágenes de capa de referencia directa para determinar varias imágenes de capa de referencia activas a partir de un conjunto de imágenes de capa de referencia directa. En un aspecto, las técnicas descritas pueden comprender varios procedimientos para mejorar los procedimientos de derivación de RPS entre capas existentes basados en la dependencia de predicción de referencia de subcapa. Por ejemplo, las técnicas pueden causar (o habilitar) un codificador de vídeo para proporcionar un decodificador de vídeo con indicaciones de ahorro de sobrecarga de señalización con más frecuencia que los procedimientos actuales. Por ejemplo, los procedimientos mejorados pueden integrarse, por ejemplo, en las extensiones de la HEVC, y aplicarse a la codificación ajustable a escala, a la codificación de múltiples vistas (por ejemplo, con o sin profundidad), y a cualquier otra extensión de la HEVC y/o a otros códecs de vídeo de múltiples capas. Cualquiera de los procedimientos y/o mecanismos de señalización descritos puede aplicarse independientemente de los demás, o en combinación.

[0016] Las ventajas de las técnicas descritas en la presente descripción, entre otras ventajas, pueden incluir la reducción de la asignación de memoria y la reducción de los costos de señalización (por ejemplo, costos de señalización de índice) durante la derivación de conjuntos de imágenes de referencia entre capas. En comparación con los procedimientos actuales en MV-HEVC WD5 y SHVC WD3, las técnicas descritas en la presente divulgación también pueden reducir los bits de encabezado de corte para la señalización del RPS entre capas durante ciertos escenarios.

[0017] La figura 1A es un diagrama de bloques que ilustra un sistema de codificación y decodificación de vídeo 10 ejemplar que puede utilizar las técnicas de derivación de conjunto de imágenes de referencia entre capas, descritas en la presente divulgación. Como se muestra en la FIG. 1A, el sistema 10 incluye un módulo de origen 12 que genera datos de vídeo codificados para ser decodificados posteriormente por un módulo de destino 14. En el ejemplo de la FIG. 1A, el módulo de origen 12 y el módulo de destino 14 están en dispositivos independientes; específicamente, el módulo de origen 12 es parte del módulo de origen, y el módulo de destino 14 es parte de un dispositivo de destino. Sin embargo, se observa que los módulos de origen y de destino 12, 14 pueden estar en el mismo dispositivo, o ser parte del mismo, como se muestra en el ejemplo de la FIG. 1B.

[0018] Con referencia una vez más a la FIG. 1A, el módulo de origen 12 y el módulo de destino 14 pueden comprender cualquiera entre una amplia gama de dispositivos, incluidos los ordenadores de escritorio, los ordenadores plegables (por ejemplo, los ordenadores portátiles), los ordenadores de tableta, los decodificadores, los equipos telefónicos de mano tales como los llamados teléfonos "inteligentes", los llamados paneles "inteligentes", televisores, cámaras, dispositivos de visualización, reproductores de medios digitales, consolas de videojuegos, dispositivos de transmisión continua de vídeo o similares. En algunos casos, el módulo de origen 12 y el módulo de destino 14 pueden estar equipados para la comunicación inalámbrica.

[0019] El módulo de destino 14 puede recibir los datos de vídeo codificados que se van a decodificar a través de un enlace 16. El enlace 16 puede comprender cualquier tipo de medio o dispositivo que pueda desplazar los datos de vídeo codificados desde el módulo de origen 12 al módulo de destino 14. En el ejemplo de la FIG. 1A, el enlace 16 puede comprender un medio de comunicación para permitir que el módulo de origen 12 transmita datos de vídeo codificados directamente al módulo de destino 14 en tiempo real. Los datos de vídeo codificados pueden modularse de acuerdo con una norma de comunicación, tal como un protocolo de comunicación inalámbrica, y transmitirse al módulo de destino 14. El medio de comunicación puede comprender cualquier medio de comunicación, inalámbrica o cableada, tal como un espectro de radiofrecuencia (RF) o una o más líneas físicas de transmisión. El medio de comunicación puede formar parte de una red basada en paquetes, tal como una red de área local, una red de área amplia o una red global tal como Internet. El medio de comunicación puede incluir encaminadores, conmutadores, estaciones base o cualquier otro equipo que pueda ser útil para facilitar la comunicación desde el módulo de origen 12 al módulo de destino 14.

[0020] De forma alternativa, los datos codificados pueden emitirse desde una interfaz de salida 22 hasta un dispositivo de almacenamiento 31 optativo. De forma similar, se puede acceder a los datos codificados desde el dispositivo de almacenamiento 31 mediante una interfaz de entrada. El dispositivo de almacenamiento 31 puede incluir cualquiera entre una diversidad de medios de almacenamiento de datos, de acceso distribuido o local, tales como una unidad de disco duro, memoria flash, memoria volátil o no volátil u otros medios adecuados cualesquiera de

almacenamiento digital para almacenar datos de vídeo codificados. En otro ejemplo, el dispositivo de almacenamiento 31 puede corresponder a un servidor de ficheros o a otro dispositivo de almacenamiento intermedio que pueda guardar el vídeo codificado generado mediante el módulo de origen 12. El módulo de destino 14 puede acceder a datos de vídeo almacenados del dispositivo de almacenamiento 31, mediante transmisión continua o descarga. El servidor de ficheros puede ser cualquier tipo de servidor que pueda almacenar datos de vídeo codificados y transmitir esos datos de vídeo codificados al módulo de destino 14. Entre los servidores de ficheros de ejemplo se incluye un servidor de la Red (por ejemplo, para una sede de la Red), un servidor del FTP, dispositivos de almacenamiento conectados en red (NAS) o una unidad de disco local. El módulo de destino 14 puede acceder a los datos de vídeo codificados mediante cualquier conexión de datos estándar, incluyendo una conexión a Internet. Esto puede incluir un canal inalámbrico (por ejemplo, una conexión de Wi-Fi), una conexión cableada (por ejemplo, DSL, módem de cable, etc.) o una combinación de ambas que sea adecuada para acceder a datos de vídeo codificados almacenados en un servidor de ficheros. La transmisión de datos de vídeo codificados desde el dispositivo de almacenamiento 31 puede ser una transmisión continua, una transmisión de descarga o una combinación de ambas.

[0021] En el ejemplo de la FIG. 1A, el módulo de origen 12 incluye un origen de vídeo 18, un codificador de vídeo 20 y una interfaz de salida 22. En algunos casos, la interfaz de salida 22 puede incluir un modulador/demodulador (módem) y/o un transmisor. En el módulo de origen 12, el origen de vídeo 18 puede incluir un origen tal como un dispositivo de captura de vídeo, por ejemplo, una videocámara, un archivo de vídeo que contiene vídeo previamente capturado, una interfaz de alimentación de vídeo para recibir vídeo desde un proveedor de contenido de vídeo y/o un sistema de gráficos por ordenador para generar datos de gráficos por ordenador como el vídeo de origen, o una combinación de dichos orígenes. Como un ejemplo, si el origen de vídeo 18 es una videocámara, el dispositivo de origen 12 y el dispositivo de destino 14 pueden formar los denominados "teléfonos con cámara" o "videoteléfonos", tal como se ilustra en el ejemplo de la FIG. 1B. Sin embargo, las técnicas descritas en esta divulgación pueden ser aplicables a la codificación de vídeo en general, y se pueden aplicar a aplicaciones inalámbricas y/o cableadas.

[0022] El vídeo capturado, precapturado o generado por ordenador puede ser codificado por el codificador de vídeo 20. Los datos de vídeo codificados pueden ser directamente transmitidos al módulo de destino 14 a través de la interfaz de salida 22 del módulo de origen 12. Los datos de vídeo codificados pueden almacenarse también (o de forma alternativa) en el dispositivo de almacenamiento 31 para un acceso posterior mediante el módulo de destino 14 u otros dispositivos, para su decodificación y/o reproducción.

[0023] El módulo de destino 14 incluye una interfaz de entrada 28, un decodificador de vídeo 30 y un dispositivo de visualización 32. En algunos casos, la interfaz de entrada 28 puede incluir un receptor y/o un módem. La interfaz de entrada 28 del módulo de destino 14 puede recibir los datos de vídeo codificados a través del enlace 16. Los datos de vídeo codificados, comunicados a través del enlace 16, o proporcionados en el dispositivo de almacenamiento 31, pueden incluir una diversidad de elementos sintácticos generados mediante el codificador de vídeo 20, para su uso mediante un decodificador de vídeo, tal como el decodificador de vídeo 30, en la decodificación de los datos de vídeo. Dichos elementos sintácticos se pueden incluir con los datos de vídeo codificados, transmitidos en un medio de comunicación, almacenados en un medio de almacenamiento o almacenados en un servidor de archivos.

[0024] El dispositivo de visualización 32 puede estar integrado con, o ser externo a, el módulo de destino 14. En algunos ejemplos, el módulo de destino 14 puede incluir un dispositivo de visualización integrado y también estar configurado para interconectarse con un dispositivo de visualización externo. En otros ejemplos, el módulo de destino 14 puede ser un dispositivo de visualización. En general, el dispositivo de visualización 32 muestra los datos de vídeo decodificados a un usuario, y puede comprender cualquiera entre varios dispositivos de visualización, tales como 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.

[0025] En aspectos relacionados, la FIG. 1B muestra un sistema ejemplar de codificación y decodificación de vídeo 10', en el que los módulos de origen y de destino 12, 14 están en, o forman parte de, un dispositivo o un dispositivo de usuario 11. El dispositivo 11 puede ser un equipo telefónico de mano, tal como un teléfono "inteligente" o similar. El dispositivo 11 puede incluir un módulo controlador/procesador 13 optativo en comunicación operativa con los módulos de origen y de destino 12, 14. El sistema 10' de la FIG. 1B puede incluir, además, una unidad de procesamiento de vídeo 21 entre el codificador de vídeo 20 y la interfaz de salida 22. En algunas implementaciones, la unidad de procesamiento de vídeo 21 es una unidad independiente, como se ilustra en la FIG. 1B; sin embargo, en otras implementaciones, la unidad de procesamiento de vídeo 21 puede implementarse como una parte del codificador de vídeo 20 y/o el módulo procesador/controlador 13. El sistema 10' también puede incluir un rastreador 29 optativo, que puede rastrear un objeto de interés en una secuencia de vídeo. El objeto de interés a rastrear puede ser segmentado por una técnica descrita en relación con uno o más aspectos de la presente divulgación. En aspectos relacionados, el rastreo puede ser realizado por el dispositivo de visualización 32, solo o junto con el rastreador 29. El sistema 10' de la FIG. 1B, y sus componentes, son similares de otro modo al sistema 10 de la FIG. 1A y sus componentes.

[0026] Como se describe con mayor detalle a continuación, los aspectos de la presente divulgación generalmente se relacionan con técnicas de derivación de conjuntos de imágenes de referencia entre capas que usan indicaciones de ciertas propiedades de las imágenes de capas de referencia para determinar varias imágenes de capas de

referencia activas a partir de un conjunto de imágenes de capas de referencia directa. Las técnicas de esta divulgación pueden ser realizadas por el codificador de vídeo 20, el decodificador de vídeo 30 y/o el módulo procesador/controlador 13.

5 **[0027]** En un ejemplo, el codificador de vídeo 20 y el decodificador de vídeo 30 pueden funcionar de acuerdo con las normas de compresión de vídeo, incluyendo la HEVC o similares. En otro ejemplo, el codificador de vídeo 20 y el decodificador de vídeo 30 pueden funcionar de acuerdo con otras normas patentadas o industriales, tales como la norma ITU-T H.264, de forma alternativa denominada MPEG-4, Parte 10, Codificación de Vídeo Avanzada (AVC), o ampliaciones de dichas normas. Otros ejemplos de normas de compresión de vídeo incluyen MPEG-2 e ITU-T H.263.
10 Sin embargo, las técnicas de esta divulgación no están limitadas a ninguna norma o técnica de codificación particular.

[0028] Aunque no se muestra en las FIG. 1A a 1B, en algunos aspectos, el codificador de vídeo 20 y el decodificador de vídeo 30 pueden estar integrados, cada uno de ellos, en un codificador y decodificador de audio, y pueden incluir unidades MUX-DEMUX adecuadas, u otro tipo de hardware y software, para gestionar la codificación tanto de audio
15 como de vídeo en un flujo de datos común o en flujos de datos diferentes. Si procede, en algunos ejemplos, las unidades MUX-DEMUX pueden ajustarse al protocolo de multiplexador ITU H.223 o a otros protocolos, tales como el protocolo de datagramas de usuario (UDP).

[0029] Cada uno entre el codificador de vídeo 20 y el decodificador de vídeo 30 puede implementarse como cualquiera entre una variedad de circuitos codificadores adecuados, tales como uno o más microprocesadores, procesadores de señales digitales (DSP), circuitos integrados específicos de la aplicación (ASIC), formaciones de compuertas programables in situ (FPGA), 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
20 en hardware usando uno o más procesadores para realizar las técnicas de esta divulgación. Cada uno entre el codificador de vídeo 20 y el decodificador de vídeo 30 puede estar incluido en uno o más codificadores o decodificadores, cualquiera de los cuales puede estar integrado como parte de un codificador/decodificador (CÓDEC) combinado en un respectivo dispositivo.

[0030] El equipo JCT-VC está trabajando en la elaboración de 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 los 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
25 tres modos de codificación mediante predicción intra.

[0031] Por lo general, el modelo de funcionamiento del HM describe que una trama o imagen de vídeo puede dividirse en una secuencia de unidades arboladas de codificación (CTU), también denominadas unidades máximas de codificación (LCU), que incluyen muestras tanto de luma como de croma. Un bloque arbolado tiene un fin similar al de un macrobloque de la norma H.264. Un fragmento incluye un número de bloques arbolados consecutivos en orden de codificación. Una trama o imagen de vídeo puede dividirse en uno o más fragmentos. Cada bloque arbolado puede dividirse en unidades de codificación (CU) de acuerdo con un árbol cuádruple. Por ejemplo, un bloque arbolado, como un nodo raíz del árbol cuádruple, puede dividirse en cuatro nodos hijos, y cada nodo hijo puede, a su vez, ser un nodo padre y dividirse en otros cuatro nodos hijos. Un nodo hijo final, no dividido, como un nodo hoja del árbol cuádruple,
35 comprende un nodo de codificación, es decir, un bloque de vídeo codificado. Los datos sintácticos asociados con un flujo de bits codificado pueden definir un número máximo de veces que puede dividirse un bloque arbolado, y también pueden definir un tamaño mínimo de los nodos de codificación.

[0032] 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 es de forma cuadrada. El tamaño de la CU puede variar desde 8x8 píxeles hasta el tamaño del bloque arbolado, con un máximo de 64x64 píxeles o más. Cada CU puede contener una o más PU y una o más TU. 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 entre si la CU está codificada en modo de salto o directa, codificada en modo de predicción intra o
40 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 cuádruple. Una TU puede tener forma cuadrada o no cuadrada.

[0033] La figura 2 es un diagrama de bloques que ilustra un codificador de vídeo 20 ejemplar que puede implementar las técnicas de derivación de conjunto de imágenes de referencia (RPS) entre capas, descritas en la presente divulgación. El codificador de vídeo 20 puede realizar la intracodificación y la intercodificación de bloques de vídeo dentro de fragmentos de vídeo. La intracodificación se basa en la predicción espacial para reducir o eliminar la redundancia espacial en el vídeo dentro de una trama o imagen de vídeo dada. La intercodificación se basa en la predicción temporal para reducir o eliminar la redundancia temporal en el vídeo dentro de tramas o imágenes adyacentes de una secuencia de vídeo. La intramodalidad (modalidad I) puede referirse a cualquiera de varias modalidades de compresión de base espacial. Las intermodalidades, tales como la predicción unidireccional
45

(modalidad P) o la bipredicción (modalidad B), pueden referirse a cualquiera de varias modalidades de compresión de base temporal.

[0034] En el ejemplo de la FIG. 2, el codificador de vídeo 20 incluye una unidad de partición 35, una unidad de procesamiento de predicción 41, una memoria 64 (que también puede denominarse memoria de imágenes de referencia 64), un sumador 50, una unidad de procesamiento de transformación 52, una unidad de cuantización 54 y una unidad de codificación por entropía 56. La unidad de procesamiento de predicción 41 incluye una unidad de estimación de movimiento 42, una unidad de compensación de movimiento 44 y una unidad de procesamiento de intrapredicción 46. Para la reconstrucción de bloques de vídeo, el codificador de vídeo 20 incluye también una unidad de cuantización inversa 58, una unidad de procesamiento de transformación inversa 60 y un sumador 62. El filtro de desbloqueo 72 también se puede incluir para filtrar las fronteras del bloque para eliminar distorsiones de efecto pixelado del vídeo reconstruido. Como se muestra en la FIG. 2, el codificador de vídeo 20 también incluye filtros de bucle adicionales, incluyendo el filtro de desplazamiento adaptativo de muestras (SAO) 74 y un filtro de bucle adaptativo (ALF) 76 optativo. Aunque el filtro de desbloqueo 72, el filtro SAO 74 y el ALF 76 optativo se muestran como filtros en bucle en la FIG. 2, en algunas configuraciones, el filtro de desbloqueo 72, el filtro SAO 74 y el ALF 76 optativo pueden implementarse como filtros de pos-bucle. Además, uno o más entre el filtro de desbloqueo 72 y el ALF 76 optativo pueden omitirse en algunas implementaciones de las técnicas de esta divulgación. En particular, el ALF 76 se omitiría en las implementaciones para la HEVC, ya que el ALF 76 no existe en la HEVC.

[0035] Como se muestra en la FIG. 2, el codificador de vídeo 20 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, tejas u otras unidades mayores, así como la división de bloques de vídeo, por ejemplo, de acuerdo con una estructura de árbol cuádruple de las LCU y las CU. El codificador de vídeo 20 ilustra, en general, los componentes que codifican bloques de vídeo dentro de un fragmento de vídeo a codificar. El fragmento se puede dividir en múltiples bloques de vídeo (y, posiblemente, en conjuntos de bloques de vídeo denominados tejas). La unidad de procesamiento de predicción 41 puede seleccionar una entre una pluralidad de posibles modalidades de codificación, que pueden incluir un tamaño de partición, tal como una entre una pluralidad de modalidades de intracodificación, o una entre una pluralidad de modalidades de intercodificación, para el bloque de vídeo actual, basándose en resultados erróneos (por ejemplo, la velocidad de codificación y el nivel de distorsión). La unidad de procesamiento de predicción 41 puede proporcionar el bloque intracodificado o intercodificado resultante al sumador 50 para generar datos de bloques residuales y al sumador 62 para reconstruir el bloque codificado para su uso como imagen de referencia.

[0036] La unidad de procesamiento de intrapredicción 46, dentro de la unidad de procesamiento de predicción 41, puede realizar la codificación intrapredictiva del bloque de vídeo actual con respecto a uno o más bloques contiguos en la misma trama o fragmento que el bloque actual a codificar, para proporcionar compresión espacial. La unidad de estimación de movimiento 42 y la unidad de compensación de movimiento 44 dentro de la unidad de procesamiento de predicción 41 realizan la codificación interpredictiva del bloque de vídeo actual con respecto a uno o más bloques predictivos en una o más imágenes de referencia, para proporcionar compresión temporal.

[0037] La unidad de estimación de movimiento 42 puede estar configurada para determinar la modalidad de interpredicción para un fragmento de vídeo, de acuerdo con un patrón predeterminado para una secuencia de vídeo. El patrón predeterminado puede designar fragmentos de vídeo en la secuencia como fragmentos predichos (fragmentos P), fragmentos predichos en dos direcciones (fragmentos B) o fragmentos P/B generalizados (fragmentos GPB). La unidad de estimación de movimiento 42 y la unidad de compensación de movimiento 44 pueden estar sumamente integradas, pero se ilustran por separado con fines conceptuales. La estimación del movimiento, realizada por la unidad de estimación de movimiento 42, es el proceso de generación de vectores de movimiento, que estiman el movimiento para los bloques de vídeo. Un vector de movimiento, por ejemplo, puede indicar el desplazamiento de una PU de un bloque de vídeo dentro de una trama o imagen de vídeo actual en relación con un bloque predictivo dentro de una imagen de referencia.

[0038] Un bloque predictivo es un bloque en el que se halla una estrecha correspondencia con la PU del bloque de vídeo a codificar, en términos de diferencia de píxeles, que puede determinarse mediante una suma de diferencias absolutas (SAD), una suma de diferencias al cuadrado (SSD) u otras métricas de diferencia. En algunos ejemplos, el codificador de vídeo 20 puede calcular valores para posiciones fraccionarias de píxeles de imágenes de referencia almacenadas en la memoria de imágenes de referencia 64. Por ejemplo, el codificador de vídeo 20 puede interpolar valores de posiciones de un cuarto de píxel, posiciones de un octavo de píxel u otras posiciones fraccionarias de píxel de la imagen de referencia. Por lo tanto, la unidad de estimación de movimiento 42 puede realizar una búsqueda de movimiento relativa a las posiciones de píxel completo y las posiciones de píxel fraccionario, y enviar un vector de movimiento con una precisión de píxel fraccionario.

[0039] La unidad de estimación de movimiento 42 calcula un vector de movimiento para una PU de un bloque de vídeo en un fragmento intercodificado, comparando la posición de la PU con la posición de un bloque predictivo de una imagen de referencia. La imagen de referencia puede seleccionarse a partir de una primera lista (lista 0) de imágenes de referencia o una segunda lista (lista 1) de imágenes de referencia, cada una de las cuales identifica una o más imágenes de referencia almacenadas en una memoria de imágenes de referencia 64. La unidad de estimación

de movimiento 42 envía el vector de movimiento calculado a la unidad de codificación por entropía 56 y a la unidad de compensación de movimiento 44.

5 **[0040]** 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 mediante estimación de movimiento, realizando posiblemente interpolaciones hasta una 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 20 forma un bloque de vídeo residual restando los valores de píxeles del bloque predictivo a los valores de píxeles del bloque de vídeo actual que se está codificando, formando valores de diferencias de píxel. Los valores de diferencias de píxel forman datos residuales para el bloque, y pueden incluir componentes de diferencia, tanto de luma como de 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 con los bloques de vídeo y con el fragmento de vídeo para su uso por el decodificador de vídeo 30 en la decodificación de los bloques de vídeo del fragmento de vídeo.

20 **[0041]** La unidad de procesamiento de intrapredicción 46 puede realizar la intrapredicción en un bloque actual, como alternativa a la interpredicción realizada 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 intrapredicción 46 puede determinar una modalidad de intrapredicción a usar para codificar un bloque actual. En algunos ejemplos, la unidad de procesamiento de intrapredicción 46 puede codificar un bloque actual usando diversas modalidades de intrapredicción, por ejemplo, durante pases de codificación independientes, y la unidad de procesamiento de predicción 41 puede seleccionar una modalidad de intrapredicción o interpredicción adecuada para su uso a partir de las modalidades sometidas a prueba. Por ejemplo, la unidad de procesamiento de intrapredicción 46 puede calcular valores de velocidad-distorsión usando un análisis de velocidad-distorsión para las diversas modalidades de intrapredicción sometidas a prueba, y seleccionar la modalidad de intrapredicción que tenga las mejores características de velocidad-distorsión entre las modalidades sometidas a prueba. El análisis de velocidad-distorsión determina, en general, una magnitud de distorsión (o error) entre un bloque codificado y un bloque original no codificado que fue codificado para generar el bloque codificado, así como una velocidad de transmisión de bits (es decir, un número de bits) usada para generar el bloque codificado. La unidad de procesamiento de intrapredicción 46 puede calcular relaciones a partir de las distorsiones y velocidades para los diversos bloques codificados, para determinar qué modalidad de intrapredicción presenta el mejor valor de velocidad-distorsión para el bloque.

35 **[0042]** En cualquier caso, tras seleccionar una modalidad de intrapredicción para un bloque, la unidad de procesamiento de predicción 41 puede proporcionar información que indique la modalidad de intrapredicción seleccionada para el bloque a la unidad de codificación por entropía 56. La unidad de codificación por entropía 56 puede codificar la información que indica la modalidad de intrapredicción seleccionada de acuerdo con las técnicas de esta divulgación. El codificador de vídeo 20 puede incluir, en el flujo de bits transmitido, datos de configuración, que pueden incluir una pluralidad de tablas de índices de modo de predicción intra y una pluralidad de tablas modificadas de índices de modo de predicción intra (también denominadas tablas de correlación de palabras de código), definiciones de contextos de codificación para diversos bloques e indicaciones de un modo de predicción intra más probable, una tabla de índices de modo de predicción intra y una tabla modificada de índices de modo de predicción intra, a usar para cada uno de los contextos.

45 **[0043]** Después de que la unidad de procesamiento de predicción 41 genera el bloque predictivo para el bloque de vídeo actual, ya sea mediante la interpredicción o la intrapredicción, el codificador de vídeo 20 forma un bloque de vídeo residual restando el bloque predictivo al bloque de vídeo actual. Los datos de vídeo residuales en el bloque residual pueden estar incluidos en una o más TU y aplicarse a la unidad de procesamiento de transformación 52. La unidad de procesamiento de transformación 52 transforma los datos de vídeo residuales en coeficientes de transformación residuales usando una transformación, tal como una transformación de coseno discreta (DCT) o una transformación conceptualmente similar. La unidad de procesamiento de transformación 52 puede convertir los datos de vídeo residuales, de un dominio de píxel a un dominio de transformación, tal como un dominio de frecuencia.

55 **[0044]** La unidad de procesamiento de transformada 52 puede enviar los coeficientes de transformada resultantes a la unidad de cuantificación 54. La unidad de cuantificación 54 cuantifica los coeficientes de transformada para reducir más la velocidad 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 se puede modificar ajustando un parámetro de cuantificación. En algunos ejemplos, la unidad de cuantificación 54 puede realizar, a continuación, una exploración de la matriz que incluye los coeficientes de transformada cuantificados. De forma alternativa, la unidad de codificación por entropía 56 puede realizar la exploración.

65 **[0045]** Tras la cuantificación, la unidad de codificación por entropía 56 realiza la codificación por entropía de los coeficientes de transformada cuantificados. Por ejemplo, la unidad de codificación de entropía 56 puede realizar una codificación de longitud variable adaptativa al contexto (CAVLC), una codificación aritmética binaria adaptativa al contexto (CABAC), una codificación aritmética binaria adaptativa al contexto basada en la sintaxis (SBAC), una codificación por entropía por división de intervalos de probabilidad (PIPE) u otra metodología o técnica de codificación

por entropía. Tras la codificación por entropía por la unidad de codificación por entropía 56, el flujo de bits codificado se puede transmitir al decodificador de vídeo 30, o archivarse para su posterior transmisión o recuperación por el decodificador de vídeo 30. 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.

[0046] La unidad de cuantización inversa 58 y la unidad de procesamiento de transformación inversa 60 aplican una cuantización inversa y una transformación inversa, respectivamente, para reconstruir el bloque residual en el dominio del píxel, para su posterior uso como bloque de referencia de una imagen de referencia. La unidad de compensación de movimiento 44 puede calcular un bloque de referencia añadiendo el bloque residual a un bloque predictivo de una de las imágenes de referencia dentro de una de las listas de imágenes de referencia. La unidad de compensación de movimiento 44 también puede aplicar uno o más filtros de interpolación al bloque residual reconstruido para calcular valores fraccionarios de píxeles, para su uso en la estimación de movimiento. El sumador 62 añade el bloque residual reconstruido al bloque de predicción compensado por movimiento, generado por la unidad de compensación de movimiento 44 para generar un bloque de referencia para su almacenamiento en la memoria de imágenes de referencia 64.

[0047] Antes del almacenamiento en la memoria 64, el bloque residual reconstruido se puede filtrar mediante uno o más filtros. Si se desea, el filtro de desbloqueo 72 también se puede aplicar para filtrar los bloques residuales reconstruidos, con el fin de eliminar las distorsiones de efecto pixelado. También se pueden usar otros filtros de bucle (en el bucle de codificación o bien después del bucle de codificación) para allanar las transiciones de píxeles o mejorar de otro modo la calidad del vídeo. Un ejemplo de un filtro de bucle de ese tipo es el filtro de SAO 74. La unidad de estimación de movimiento 42 y la unidad de compensación de movimiento 44 pueden usar el bloque de referencia como bloque de referencia para realizar la interpredicción de un bloque en una trama o imagen de vídeo posterior.

[0048] El filtro de SAO 74 puede determinar valores de desplazamiento para el filtrado de SAO de una manera que mejora la calidad de la codificación de vídeo. La mejora de la calidad de codificación de vídeo, por ejemplo, puede implicar la determinación de valores de desplazamiento que hacen que una imagen reconstruida coincida más estrechamente con una imagen original. El codificador de vídeo 20, por ejemplo, puede codificar los datos de vídeo utilizando múltiples pases con diferentes valores de desplazamiento y elegir, para su inclusión en un flujo de bits codificado, los valores de desplazamiento que ofrecen una calidad de codificación deseable, según lo determinado basándose en un cálculo de velocidad-distorsión, por ejemplo.

[0049] En algunas configuraciones, el filtro de SAO 74 puede configurarse para aplicar uno o más tipos de desplazamiento, tales como el desplazamiento de bordes descrito anteriormente. El filtro de SAO 74 también puede a veces no aplicar ningún desplazamiento, lo cual puede considerarse un tercer tipo de desplazamiento en sí mismo. El tipo de desplazamiento aplicado por el filtro de SAO 74 puede señalizarse, explícita o implícitamente, a un decodificador de vídeo. Al aplicar el desplazamiento de bordes, los píxeles se pueden clasificar según la información de bordes.

[0050] El codificador de vídeo 20 de la FIG. 2 representa un ejemplo de un codificador de vídeo configurado para determinar un primer índice de borde, en donde el primer índice de borde comprende un índice de borde para un componente de luminancia de un primer píxel circundante, determinar un segundo índice de borde, en donde el segundo índice de borde comprende un índice de borde para un componente de luminancia de un segundo píxel circundante, determinar un tercer índice de borde basado en el primer índice de borde y el segundo índice de borde, en donde el tercer índice de borde comprende un índice de borde para un componente de croma de un píxel actual, seleccionar un desplazamiento basándose en el tercer índice de borde y aplicar el desplazamiento al componente de croma del píxel actual.

[0051] La figura 3 es un diagrama de bloques que ilustra un decodificador de vídeo 30 ejemplar que puede implementar las técnicas de derivación de conjunto de imágenes de referencia (RPS) entre capas, descritas en la presente divulgación. En el ejemplo de la FIG. 3, el decodificador de vídeo 30 incluye una unidad de decodificación por entropía 80, una unidad de procesamiento de predicción 81, una unidad de cuantización inversa 86, una unidad de transformación inversa 88, un sumador 90 y una memoria de imágenes de referencia 92. La unidad de procesamiento de predicción 81 incluye la unidad de compensación de movimiento 82, para la decodificación por interpredicción, y la unidad de procesamiento de intrapredicción, 84, para la decodificación por intrapredicción. En algunos ejemplos, el decodificador de vídeo 30 puede realizar un pase de decodificación que, en general, es recíproco al pase de codificación descrito con respecto al codificador de vídeo 20 de la FIG. 2.

[0052] Durante el proceso de descodificación, el descodificador de vídeo 30 recibe un flujo de bits de vídeo codificado, que representa bloques de vídeo de un fragmento de vídeo codificado y elementos sintácticos asociados, desde el codificador de vídeo 20. La unidad de decodificación por entropía 80 del decodificador de vídeo 30 decodifica por entropía el flujo de bits para generar coeficientes cuantizados, vectores de movimiento y otros elementos sintácticos. La unidad de decodificación por entropía 80 remite los vectores de movimiento y otros elementos sintácticos a la unidad de procesamiento de predicción 81. El descodificador de vídeo 30 puede recibir los elementos sintácticos en el nivel de fragmento de vídeo y/o el nivel de bloque de vídeo.

[0053] Cuando el fragmento de vídeo se codifica como un fragmento intracodificado (I), la unidad de procesamiento de intrapredicción 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 una modalidad de intrapredicción señalizada, y datos de bloques previamente decodificados de la trama o imagen actual. Cuando la trama de vídeo se codifica como un fragmento intercodificado (es decir, B, P o GPB), la unidad de compensación de movimiento 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 en otros elementos sintácticos recibidos desde la unidad de decodificación por entropía 80. Los bloques predictivos se pueden generar a partir de una de las imágenes de referencia dentro de una de las listas de imágenes de referencia. El decodificador de vídeo 30 puede construir las listas de tramas de referencia, Lista 0 y Lista 1, utilizando técnicas de construcción predeterminadas, basándose en imágenes de referencia almacenadas en la memoria de imágenes de referencia 92.

[0054] La unidad de compensación de movimiento 82 determina la información de predicción para un bloque de vídeo del fragmento de vídeo actual, analizando sintácticamente los vectores de movimiento y otros elementos sintácticos, y usa la información de predicción para generar los bloques predictivos para el bloque de vídeo actual que se está 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, un fragmento B, un fragmento P o un fragmento GPB), información de construcción para una o más de las listas de imágenes de referencia para el fragmento, vectores de movimiento para cada bloque de vídeo intercodificado del fragmento, el estado de predicción inter para cada bloque de vídeo intercodificado del fragmento y otra información para descodificar los bloques de vídeo en el fragmento de vídeo actual.

[0055] 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 20 durante la codificación de los bloques de vídeo, para calcular valores interpolados para píxeles fraccionarios de los bloques de referencia. En este caso, la unidad de compensación de movimiento 82 puede determinar los filtros de interpolación usados por el codificador de vídeo 20 a partir de los elementos sintácticos recibidos y usar los filtros de interpolación para generar bloques predictivos.

[0056] La unidad de cuantización inversa 86 cuantiza inversamente, es decir, decuantiza, los coeficientes de transformación cuantizados proporcionados en el flujo de bits y decodificados por la unidad de decodificación por entropía 80. El proceso de cuantización inversa puede incluir el uso de un parámetro de cuantización calculado por el codificador de vídeo 20 para cada bloque de vídeo en el fragmento de vídeo, con el fin de determinar un grado de cuantización y, asimismo, un grado de cuantificación inversa que se debería aplicar. La unidad de procesamiento de transformación inversa 88 aplica una transformación inversa, por ejemplo una DCT inversa, una transformación entera inversa o un proceso de transformación inversa conceptualmente similar, a los coeficientes de transformación, con el fin de generar bloques residuales en el dominio del píxel.

[0057] Una vez que la unidad de procesamiento de predicción 81 genera el bloque predictivo para el bloque de vídeo actual basándose en los vectores de movimiento y en otros elementos sintácticos, el decodificador de vídeo 30 forma un bloque de vídeo decodificado sumando los bloques residuales procedentes de la unidad de procesamiento de transformación 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. Los bloques de vídeo decodificados, formados por el sumador 90, pueden a continuación ser filtrados por un filtro de desbloqueo 93, un filtro de SAO 94 y un ALF 95 optativo. El ALF 95 optativo representa un filtro optativo que se puede excluir de algunas implementaciones. Se observa que el ALF 95 se omitiría en las implementaciones para la HEVC, ya que el ALF 95 no existe en la HEVC. Los bloques de vídeo decodificados en una trama o imagen dada se almacenan a continuación en la memoria de imágenes de referencia 92, que almacena imágenes de referencia usadas para una compensación de movimiento posterior. La memoria de imágenes de referencia 92 también almacena vídeo decodificado para una presentación posterior en un dispositivo de visualización, tal como el dispositivo de visualización 32 de las FIG. 1A a 1B. En aspectos relacionados, el filtro de SAO 94 puede configurarse para aplicar uno o más de los mismos filtros (por ejemplo, desplazamiento de borde y desplazamiento de banda) que el filtro de SAO 74 analizado anteriormente.

[0058] El decodificador de vídeo 30 de la FIG. 3 representa un ejemplo de un decodificador de vídeo configurado para determinar un primer índice de borde, en donde el primer índice de borde comprende un índice de borde para un componente de luminancia de un primer píxel circundante, determinar un segundo índice de borde, en donde el segundo índice de borde comprende un índice de borde para un componente de luminancia de un segundo píxel circundante, determinar un tercer índice de borde basándose en el primer índice de borde y el segundo índice de borde, en donde el tercer índice de borde comprende un índice de borde para un componente de croma de un píxel actual, seleccionar un desplazamiento basándose en el tercer índice de borde y aplicar el desplazamiento al componente de croma del píxel actual.

[0059] De acuerdo con uno o más aspectos de la presente divulgación, se proporcionan técnicas para la derivación de conjuntos de imágenes de referencia (RPS) entre capas. Las técnicas pueden ser realizadas por un sistema o, por ejemplo, por el dispositivo 10 de la FIG. 1B. En algunas implementaciones, las técnicas pueden ser realizadas por el codificador de vídeo 20, el decodificador de vídeo 30 y/o el módulo controlador-procesador 13, solos o en combinación.

En un aspecto, las técnicas pueden implicar indicar si un conjunto de imágenes de referencia entre capas incluye todas las imágenes de capa de referencia directa que no están restringidas para su uso en la predicción entre capas, como se describe adicionalmente en relación con las figuras 4-5. En otro aspecto, las técnicas pueden implicar la determinación de varias imágenes de capa de referencia activa a partir de una o más imágenes de capa de referencia directa, como se describe adicionalmente en relación con las figuras 6-7.

[0060] La figura 4 ilustra un ejemplo de información de vídeo 400 que puede procesarse para indicar si un conjunto de imágenes de referencia entre capas incluye todas las imágenes de capa de referencia directa (también denominadas en el presente documento "imágenes de referencia" o "imágenes de capa de referencia") que no están restringidas para uso en predicción entre capas. En un aspecto, una o más de las características descritas con respecto a la figura 4 puede realizarse generalmente mediante un codificador (por ejemplo, el codificador de vídeo 20 de las figuras 1A, 1B y/o 2). En un modo de realización, el codificador de vídeo 20 puede proporcionar la indicación a un decodificador de vídeo (por ejemplo, el decodificador de vídeo 30 de las figuras 1A, 1B y/o 3). En otro modo de realización, la indicación puede indicarse o determinarse mediante un dispositivo diferente. Los diversos aspectos descritos con relación a la figura 4 pueden incluirse en una unidad de memoria (por ejemplo, el dispositivo de almacenamiento 34 de la figura 1A, la memoria 64 de la figura 2, etc.). Las diversas operaciones descritas con respecto a la figura 4 pueden realizarse mediante un procesador (por ejemplo, un procesador dentro del módulo de origen, módulo de destino, codificador de vídeo y/o decodificador de vídeo de una cualquiera o más de las figuras 1A-3, el módulo procesador/controlador 13 de la figura 1B, y/o una unidad de procesamiento de vídeo, tal como la unidad de procesamiento de vídeo 21 de la figura 1B, etc.). El codificador de vídeo, el decodificador de vídeo, la unidad de memoria y/o el procesador pueden estar en comunicación entre sí y/o conectados operativamente.

[0061] Una secuencia de vídeo 415 incluye unidades de acceso (denominadas colectivamente como unidades de acceso 420), donde cada una de las unidades de acceso 420 puede contener una o más imágenes. Las imágenes pueden estar asociadas con un conjunto de parámetros de vídeo (VPS). El VPS especifica parámetros comunes para todas las imágenes en una secuencia de vídeo, tal como la secuencia de vídeo 415. La secuencia de vídeo 415 de la figura 4 ilustra dos unidades de acceso AU1, AU2; sin embargo, la secuencia de vídeo 415 puede incluir cualquier número de unidades de acceso 420. Cada unidad de acceso AU1, AU2 puede estar asociada con una o más capas 425 de información de vídeo, como se ilustra mediante una capa base BL y tres capas de mejora EL1, EL2 y EL3. La capa o capa ilustrada más baja en el nivel más bajo de la información de vídeo 400 puede ser la capa base (BL) o una capa de referencia (RL), y la capa en el nivel más alto o en el nivel más alto de la información de vídeo 400 puede ser una capa mejorada. La "capa mejorada" puede considerarse sinónimo con una "capa de mejora", y estos términos se pueden usar indistintamente. En este ejemplo, la capa BL es la capa base y las capas EL1, EL2 y EL3 son capas de mejora primera, segunda y tercera, respectivamente. Las capas entre la capa base BL y la capa de mejora más alta EL3 pueden servir como capas de mejora y/o capas de referencia. En este ejemplo, las capas EL1, EL2 y EL3 representan capas de mejora.

[0062] Por ejemplo, una capa dada (por ejemplo, la segunda capa de mejora EL2) puede ser una capa de mejora para una capa por debajo (es decir, que precede) de la capa dada, como la capa base (por ejemplo, la capa base BL) o cualquier capa de mejora intermedia (por ejemplo, la primera capa de mejora EL1). Además, la capa dada (por ejemplo, la segunda capa de mejora EL2) también puede servir como una capa de referencia para una capa de mejora (por ejemplo, la tercera capa de mejora EL3) por encima (es decir, posterior a) de la capa dada. Cualquier capa dada entre la capa base (es decir, la capa más baja que tiene, por ejemplo, un conjunto de identificación de capa (ID) o igual a "1") y la capa superior (o la capa más alta) puede usarse como referencia para predicción entre capas por una capa más alta en relación con la capa dada y puede determinarse usando una capa más baja a la capa dada como referencia para la predicción entre capas.

[0063] La capa 425 que se está procesando actualmente puede denominarse la "capa actual". En la ilustración de ejemplo, la tercera capa de mejora EL3 representa la capa actual, aunque debe entenderse que cualquier capa 425 puede tratarse como la capa actual y procesarse de acuerdo con cualquiera de los modos de realización descritos en el presente documento. Una capa actual puede estar asociada con una o más capas de referencia directa. Como se mencionó anteriormente, una capa de referencia directa es una capa que se puede usar para la predicción entre capas de otra capa, tal como la capa actual. En el ejemplo ilustrado, la primera capa de mejora EL1 y la capa base BL representan la primera y segunda capa de referencia directa de la capa actual EL3, respectivamente. La primera capa de mejora EL1 y la capa base BL son capas de referencia directa de la capa actual porque se determina que son capas que pueden usarse para la predicción interna de la tercera capa de mejora, que en este ejemplo es la capa actual.

[0064] Cada una de las capas 425 puede comprender una o más imágenes, estando asociada cada imagen con una de las capas 425 y una de las unidades de acceso AU1, AU2. Las imágenes asociadas con la capa actual (en este ejemplo, la tercera capa de mejora EL3) pueden denominarse colectivamente como las imágenes actuales 405. Por ejemplo, en el modo de realización ilustrado, las imágenes actuales 405 incluyen una primera imagen actual CP1

ubicada dentro de la primera unidad de acceso AU1 y una segunda imagen actual CP2 ubicada dentro de la segunda unidad de acceso AU2. Se puede hacer referencia a las imágenes (por ejemplo, imágenes 410A1, 410A2, 410B1, 410B2) que están en capas de referencia directa (por ejemplo, la primera capa de mejora EL1 y la capa base BL) de la capa actual (por ejemplo, la tercera capa de mejora EL3) colectivamente como imágenes de capa de referencia directa 410 (o imágenes de capa de referencia 410). Por ejemplo, la imagen de capa de referencia directa 410B1 es la imagen en la primera capa de referencia directa (en este ejemplo, la primera capa de mejora EL1) y la segunda unidad de acceso AU2. Esta imagen de capa de referencia directa 410B1 también puede denominarse "imagen de capa de referencia directa 1 de la imagen actual 2". Como otro ejemplo, la imagen 410A2 de la capa de referencia directa es la imagen en la segunda capa de referencia directa (en este ejemplo, la capa base BL) y la primera unidad de acceso AU1. Esta imagen de capa de referencia directa 410A1 también puede denominarse "imagen de capa de referencia directa 2 de la imagen actual 1". Las imágenes que están en capas de referencia no directas de la capa actual (por ejemplo, la capa de referencia no directa EL2) pueden denominarse simplemente imágenes 408. En el modo de realización ilustrado, las imágenes 408 incluyen una imagen 408A ubicada en la primera unidad de acceso AU1 y una segunda imagen 408B ubicada en la segunda unidad de acceso AU2. En un modo de realización, el procesador (por ejemplo, el procesador 13) no puede considerar o hacer ninguna determinación basada en las imágenes 408A y 408B en la capa de referencia no directa (por ejemplo, la segunda capa de mejora EL2).

[0065] Cada una de las imágenes de capa de referencia directa 410 puede estar asociada con un número o valor de identificación temporal respectivo o "valor TID" (no mostrado). En un aspecto, cada una de las imágenes en una de las unidades de acceso AU1, AU2 puede estar asociada con el mismo valor TID (es decir, uno común). Por ejemplo, para la primera unidad de acceso AU1, la imagen actual CP1, la imagen 408A y ambas imágenes de capa de referencia directa 410A1 y 410A2 pueden estar asociadas con el mismo valor TID. En algunos aspectos, el procesador 13 puede determinar respectivamente si el valor de TID asociado de cada una de las imágenes 410 de capa de referencia directa es menor que un umbral predeterminado (por ejemplo, "valor de identificación temporal máximo", "número de identificación temporal máximo" o "TID máximo"). En algunos modos de realización, el TID máximo puede estar asociado con la capa actual y la secuencia de vídeo 415. En el ejemplo ilustrado, la imagen de capa de referencia directa 410A1 está asociada con un valor de TID que es mayor o igual al TID máximo, y la imagen de capa de referencia directa 410B1 está asociada con un valor de TID que es menor que el TID máximo. En un modo de realización, el TID máximo puede ser un valor de "4", el valor TID asociado con la imagen de capa de referencia directa 410A1 (en la primera unidad de acceso AU1) puede ser "5", y el valor TID asociado con la referencia directa la imagen de capa 410B1 (en la segunda unidad de acceso AU2) puede ser "2". En algunos modos de realización, el TID máximo puede ser un valor constante para toda la secuencia de vídeo 415. Por ejemplo, el TID máximo puede permanecer constante para toda la secuencia de vídeo 415 según un estándar de codificación de vídeo (por ejemplo, la variable `max_tid_il_ref_pics_plus1 [i]` del estándar HEVC). En otros modos de realización, el TID máximo puede variar según la unidad de acceso AU1, AU2 en la secuencia de vídeo 415. En otras palabras, cada unidad de acceso 420 respectiva puede estar asociada con un TID máximo respectivo.

[0066] El procesador 13 puede determinar diversa información usando las propiedades de cada una de las imágenes 410 de capa de referencia directa y la imagen actual asociada de cada imagen de capa de referencia directa 405. Las filas 430, 440, 450 y 460 de la figura 4 ilustran un modo de realización de dicha información. Por ejemplo, los valores en la fila 430 representan las imágenes de capa de referencia directa 410 asociadas con las unidades de acceso AU1, AU2. La figura 4 ilustra además asociaciones (mostradas como óvalos que tienen líneas discontinuas) entre cada una de las imágenes de capa de referencia directa 410 (por ejemplo, la imagen 410A1 de la capa de referencia directa asociada con la primera unidad de acceso AU1, la imagen 410A2 de la capa de referencia directa asociada con la primera unidad de acceso AU1, la imagen 410B1 de la capa de referencia directa asociada con la segunda unidad de acceso AU2, y la imagen 410B2 de la capa de referencia directa asociada con la segunda unidad de acceso AU2) y las entradas correspondientes en las filas 440 y 450, como se describe más adelante.

[0067] Para cada unidad de acceso de la secuencia de vídeo 415, un procesador 13 puede determinar si las imágenes de capa de referencia directa 410 dentro de la unidad de acceso están incluidas en un conjunto de imágenes de referencia entre capas para la imagen actual 405. En un aspecto, el conjunto de imágenes de referencia entre capas también puede incluir otras imágenes. Por ejemplo, con respecto a la primera unidad de acceso AU1, un procesador puede determinar si la primera imagen de capa de referencia directa 410A1 está incluida en un conjunto de imágenes de referencia entre capas para predecir CP1. El procesador también puede determinar si la segunda imagen de capa de referencia directa 410A2 está incluida en el conjunto de imágenes de referencia entre capas para predecir CP1. Con respecto a la segunda unidad de acceso AU2, el procesador puede determinar si la primera imagen de capa de referencia directa 410B1 está incluida en un conjunto de imágenes de referencia entre capas para predecir CP2. El procesador también puede determinar si la segunda imagen de capa de referencia directa 410B2 está incluida en el conjunto de imágenes de referencia entre capas para predecir CP2. Los resultados de estas determinaciones de ejemplo se ilustran en la fila 440. Dicha determinación puede implicar determinar si cada una de las imágenes de capa de referencia directa 410 se puede usar para la predicción entre capas de una de las imágenes actuales 405 asociadas. En algunos modos de realización, el procesador 13 puede hacer esta determinación basándose en un grado de eficiencia de predicción que puede realizarse utilizando la imagen de capa de referencia directa 410. Por ejemplo, si el procesador 13 determina que una o más de las imágenes de capa de referencia directa 410 comprenden un nivel de similitud (por ejemplo, basado en la estimación de movimiento, tal como, pero sin limitarse a, una diferencia en la estimación de movimiento entre la imagen de capa de referencia directa particular y la imagen actual asociada) a una

de las imágenes actuales 405 asociada que es mayor que una cantidad umbral, entonces el procesador 13 puede determinar que sería eficiente (por ejemplo, el procesador utilizaría menos bits en la señalización) usar dicha una o más de las imágenes de capa de referencia directa 410 para predecir la asociada de las imágenes actuales 405. Por lo tanto, el procesador 13 puede determinar si una o más de las imágenes de capa de referencia directa 410 están incluidas en el conjunto de imágenes de referencia entre capas.

[0068] Como se ilustra en la fila 440, solo las imágenes de capa de referencia directa 410A2, 410B1 y 410B2 se incluyen en el conjunto de imágenes de referencia entre capas. Alternativamente, el procesador 13 puede determinar que una o más de las imágenes de capa de referencia directa 410 no tienen similitud con la imagen actual 405 mayor que un umbral, de modo que no sería eficiente usar una o más imágenes de capa de referencia directa 410 para predecir la asociada de las imágenes actuales 405. En tal caso, el procesador 13 puede determinar que tales imágenes de la capa de referencia directa 410 no están incluidas en el conjunto de imágenes de referencia entre capas. Por ejemplo, como se ilustra en la fila 440, la imagen de capa de referencia directa 410A1 no se incluye en el conjunto de imágenes de referencia entre capas. En algunos modos de realización, el procesador 13 puede realizar las determinaciones descritas anteriormente para cada una de las imágenes de capa de referencia directa 410, como se muestra en la fila 440.

[0069] El procesador 13 también puede determinar si cada una de las imágenes de capa de referencia directa 410 está restringida de uso en la predicción entre capas. Los resultados de tales determinaciones se ilustran en la fila 450. En algunos modos de realización, el procesador 13 puede analizar solo las de las imágenes de capa de referencia directa 410 que el procesador 13 previamente determinado no se incluyeron en el conjunto de imágenes de referencia entre capas para determinar si tales imágenes de capa de referencia directa tienen un uso restringido en predicción entre capas. Por ejemplo, como se muestra en la fila 440, solo se determinó previamente que la primera imagen de capa de referencia directa 410A1 de la primera unidad de acceso AU1 no se incluyera en el conjunto de imágenes de referencia entre capas. Por lo tanto, en un modo de realización, el procesador 13 solo analiza dicha imagen (la primera imagen de capa de referencia directa 410A1) para determinar si está restringido su uso en la predicción entre capas. En otro modo de realización, el procesador 13 puede determinar si cada una de las imágenes 410 de capa de referencia directa está restringida de uso en la predicción entre capas, independientemente de si el procesador 13 ha determinado previamente si incluir tales imágenes de capa de referencia directa 410 en el conjunto de imágenes de referencia de capa. En otro modo de realización, el procesador 13 determina si una o más de las imágenes de capa de referencia directa 410 están restringidas para su uso en la predicción entre capas antes de determinar si tales imágenes de capa de referencia directa 410 están incluidas en el conjunto de imágenes de referencia entre capas. La determinación de si cada una de las imágenes de capa de referencia directa 410 está restringida para su uso en la predicción entre capas puede basarse en si cada una de las imágenes de capa de referencia directa 410 está asociada con un valor TID que es mayor o igual al TID máximo. En el ejemplo ilustrado de la figura 4, el procesador 13 determina que la primera imagen de capa de referencia directa 410A1 de la primera unidad de acceso A1 está restringida de uso en la predicción entre capas. Esta determinación se representa en la fila 450 como "Sí". El procesador 13 determina que la primera imagen de capa de referencia directa 410A1 de la primera unidad de acceso A1 está restringida para su uso en la predicción entre capas porque la imagen de capa de referencia directa 410A1 está asociada con un valor de TID mayor o igual al TID máximo (como indicado en la intersección de la primera capa de mejora EL1 y la primera unidad de acceso AU1). En un modo de realización, como se describió anteriormente (pero no se ilustra en la figura 4), debido a que las imágenes restantes de la capa de referencia directa 410A2, 410B1 y 410B2 de la secuencia de vídeo 415 se determina que están incluidas en el conjunto de imágenes de referencia entre capas (como se indica en la fila 440), el procesador 13 no necesita determinar adicionalmente si tales imágenes restantes de capa de referencia directa 410A2, 410B1 y 410B2 están restringidas de uso en la predicción entre capas. Debido a que el procesador no necesita realizar tales determinaciones adicionales, los valores en la fila 450 correspondientes a dicha determinación adicional pueden indicarse como "N/A". Sin embargo, en el modo de realización ilustrado, el procesador 13 determina además si cada una de las imágenes restantes de la capa de referencia directa 410A2, 410B1 y 410B2 está restringida de uso en la predicción entre capas en función de si cada uno de sus valores de TID respectivos es mayor o igual al TID máximo. Los resultados asociados con las imágenes restantes de la capa de referencia directa 410A2, 410B1 y 410B2, se proporcionan en la fila 450 como "Sí", "No" y "No", respectivamente. Debido a que la imagen de la capa de referencia directa 410A2 está asociada con un valor de TID mayor o igual al TID máximo, el procesador 13 determina que la imagen de la capa de referencia directa 410A2 está restringida de uso en la predicción entre capas. De manera similar, debido a que las imágenes de capa de referencia directa 410B1 y 410B2 están asociadas con un valor de TID que no es mayor o igual al TID máximo, el procesador 13 determina que las imágenes de capa de referencia directa 410B1 y 410B2 no están restringidas para su uso en predicción de capa.

[0070] El procesador 13 puede determinar si establecer una indicación en función de si las imágenes de capa de referencia directa están incluidas en un conjunto de imágenes de referencia entre capas (como se muestra en la fila 440) y si al menos las imágenes de capa de referencia directa que no están incluidas en el conjunto de imágenes de referencia entre capas (por ejemplo, la imagen de capa de referencia directa 410A1) está restringido para su uso en la predicción entre capas (como se muestra en la fila 450). La indicación establecida en un valor de "1" especifica que para cada imagen que hace referencia al SPV, las imágenes de la capa de referencia que pertenecen a todas las capas de referencia directa de la capa que contiene la imagen y que no están restringidas para usarse para la predicción entre capas, están presentes en la misma unidad de acceso que la imagen, y se incluyen en el conjunto de imágenes de referencia entre capas de la imagen. La indicación establecida en un valor de "0" especifica que la

restricción anterior puede aplicarse o no. El procesador 13 puede determinar si establecer la indicación para cada una de las imágenes actuales 405 (o colectivamente para la capa actual) y/o para una secuencia de vídeo codificado (CVS) (por ejemplo, para todas las capas dentro de un CVS). En un modo de realización, la indicación indica si todas las imágenes de capa de referencia directa 410 que no están restringidas para su uso en la predicción entre capas (en el modo de realización ilustrado, las imágenes de capa de referencia directa 410B1 y 410B2, como se indica en la fila 5 450) están incluidas en el conjunto de imágenes de referencia entre capas (como se indica en la fila 440). En un modo de realización, el procesador 13 puede establecer la indicación para todas las imágenes actuales 405 en la secuencia de vídeo 415 usando una única indicación (por ejemplo, que comprende un solo bit). Por ejemplo, la indicación puede ser un indicador de que el procesador 13 establece un valor de "1" o un valor de "0". En algunos modos de realización, el indicador que se establece en un valor de "1" puede indicar que todas las imágenes de capa de referencia directa 10 410 asociadas que no están restringidas para su uso en la predicción entre capas se incluyen en el conjunto de imágenes de referencia entre capas y el indicador que se establece en un valor de "0" puede indicar que todas las imágenes de capa de referencia directa 410 asociadas que no están restringidas para uso en la predicción entre capas no se incluyen en el conjunto de imágenes de referencia entre capas. En algunos modos de realización, el indicador 15 que se establece en un valor de "1" puede indicar que todas las imágenes de capa de referencia directa 410 asociadas que no están restringidas para su uso en la predicción entre capas se incluyen en el conjunto de imágenes de referencia entre capas y el indicador que se establece en un valor de "0" puede indicar que al menos una de las imágenes de capa de referencia directa 410 asociadas que no están restringidas para uso en la predicción entre capas no se incluye en el conjunto de imágenes de referencia entre capas. En algunos modos de realización, los valores, "0" 20 y "1", del indicador pueden proporcionar una indicación opuesta a la descrita anteriormente.

[0071] En algunos modos de realización, el indicador puede inicializarse a un valor de "0" o "1" antes de que el procesador 13 realice cualquiera de las determinaciones descritas anteriormente. En un modo de realización, el indicador puede ser un indicador relacionado con HEVC, por ejemplo, *all_ref_layers_active_flag*. En otro modo de 25 realización, el procesador 13 puede hacer esta indicación de una manera diferente. Por ejemplo, en un modo de realización alternativa, el procesador 13 puede proporcionar individualmente esta indicación para cada una de las imágenes actuales 405 en la secuencia de vídeo 415.

[0072] En el ejemplo ilustrado de la figura 4, las imágenes de capa de referencia directa 410 de las imágenes 30 actuales 405 son imágenes de capa de referencia directa 410A1, 410A2, 410B1 y 410B2 (como se indica en la fila 430). Las imágenes de capa de referencia directa 410A2, 410B1 y 410B2 se incluyen en el conjunto de imágenes de referencia entre capas (como se indica en la fila 440). Debido a que al menos una de las imágenes de capa de referencia directa 410 (en el modo de realización ilustrada, la imagen de capa de referencia directa 410A1) no está incluida en el conjunto de imágenes de referencia entre capas, los procedimientos anteriores determinarían que una 35 indicación debería establecerse en un valor de "0". Sin embargo, de acuerdo con diversos modos de realización descritas en el presente documento, un procesador puede determinar además si dicha imagen de capa de referencia directa 410A1 está restringida de uso en la predicción entre capas. En el ejemplo ilustrado, como se discutió anteriormente, debido a que la imagen de capa de referencia directa 410A1 tiene un valor de TID que es mayor o igual al TID máximo, el procesador 13 puede determinar que la imagen de capa de referencia directa 410A1 está restringida 40 de uso en predicción entre capas (como se indica en la fila 450). Por lo tanto, todas las imágenes de capa de referencia directa 410 para la capa actual que no están restringidas para uso en la predicción entre capas se incluyen en el conjunto de imágenes de referencia entre capas. Con respecto a la primera imagen actual CP1 en la primera unidad de acceso AU1, un procesador puede determinar que las imágenes de capa de referencia (por ejemplo, imágenes de capa de referencia directa 410A1 y 410A2) que pertenecen a todas las capas de referencia directa (por ejemplo, la 45 capa base BL y la primera capa de mejora EL1) de la capa (por ejemplo, la capa actual, que en esta realización es la tercera capa de mejora EL3) que contiene la imagen (por ejemplo, la imagen actual CP1) y que no está restringida para usarse para la predicción entre capas (por ejemplo, no hay tales imágenes, como se indica en las dos columnas izquierdas de la fila 450), están presentes en la misma unidad de acceso que la imagen y están incluidas en el conjunto de imágenes de referencia entre capas de la imagen. El procesador puede establecer la indicación en esta realización 50 porque no hay imágenes de capa de referencia directa que no tengan un uso restringido en la predicción entre capas. Además, con respecto a la segunda imagen actual CP2 en la segunda unidad de acceso AU2, el procesador también puede proporcionar la indicación (por ejemplo, la indicación en la fila 460) para indicar que las imágenes de capa de referencia (por ejemplo, imágenes de capa de referencia directa 410B1 y 410B2) que pertenecen a todas las capas de referencia directa (por ejemplo, la capa base BL y la primera capa de mejora EL1) de la capa (por ejemplo, la capa actual, que en este modo de realización es la tercera capa de mejora EL3) que contiene la imagen (por ejemplo, la 55 imagen actual CP2) y que no están restringidos para ser utilizados para la predicción entre capas (por ejemplo, como se indica en las dos columnas de la derecha de la fila 450), están presentes en la misma unidad de acceso (por ejemplo, AU2) que la imagen (por ejemplo, la imagen actual CP2) y se incluyen en el conjunto de imágenes de referencia entre capas (por ejemplo, como se indica en las dos columnas derechas de la fila 440) de la imagen (por ejemplo, la segunda imagen actual CP2). Por ejemplo, el procesador puede establecer la indicación estableciendo un 60 indicador en un valor de "1" (como se muestra en la fila 460). En otras palabras, las imágenes de capa de referencia directa que pertenecen a todas las capas de referencia directa de la capa que contiene la imagen y que no están restringidas para usarse para la predicción entre capas (según lo determinado mediante la comparación de los valores de TID asociados con un TID máximo) están presentes en la misma unidad de acceso que la imagen y puede incluirse 65 en el conjunto de imágenes de referencia entre capas de la imagen. En otro modo de realización, el indicador se inicializa a un valor de "1", y el procesador puede abstenerse de cambiar el valor del indicador cuando todas las

imágenes de capa de referencia directa 410 para la capa actual que no están restringidas para su uso en la predicción entre capas están incluidos en el conjunto de imágenes de referencia entre capas. Por ejemplo, el procesador puede establecer el indicador en un valor de "0" cuando al menos una de las imágenes de capa de referencia directa 410 para la capa actual que no está restringida para su uso en la predicción entre capas no está incluida en el conjunto de imágenes de referencia de capa.

[0073] En otro modo de realización (no representado en la figura 4), siendo todo lo demás igual a lo descrito anteriormente, el procesador 13 puede haber determinado que la imagen de capa de referencia directa 410B2 no está incluida en el conjunto de imágenes de referencia entre capas. En este ejemplo propuesto, la entrada en la fila 440 asociada con la imagen de capa de referencia directa 410B2 se rellenaría con un "No". En esta realización, el procesador 13 puede determinar si la imagen de capa de referencia directa 410B2 está restringida para su uso en la predicción entre capas (el resultado de tal determinación se indica en la entrada en la fila 450 asociada con la imagen de capa de referencia directa 410B2). Debido a que la imagen de capa de referencia directa 410B2 tiene un valor de TID menor que el TID máximo, el procesador 13 puede determinar que la imagen de capa de referencia directa 410B2 no está restringida para su uso en la predicción entre capas (como se indica en la fila 450). Por lo tanto, no todas las imágenes de capa de referencia directa 410 para la capa actual que no están restringidas para uso en la predicción entre capas se incluyen en el conjunto de imágenes de referencia entre capas. Como resultado, el procesador 13 puede proporcionar una indicación para indicar una restricción de que para cada imagen que se refiere al SPV, las imágenes de la capa de referencia que pertenecen a todas las capas de referencia directa de la capa que contiene la imagen y que no están restringidas para ser utilizadas para la predicción entre capas, están presentes en la misma unidad de acceso que la imagen, y se incluyen en el conjunto de imágenes de referencia entre capas de la imagen, puede o no aplicarse. Dicha indicación puede incluir establecer un indicador en un valor de "0". Como tal, en este ejemplo propuesto, la entrada en la fila 460 se rellenaría con un "0". En un modo de realización, el procesador 13 puede establecer el indicador en un valor de "0" para indicar que las imágenes de capa de referencia directa 410 asociadas con las imágenes actuales 405 pueden incluir o no todas las imágenes de capa de referencia directa 410 que no están restringidas para usar en la predicción entre capas. En otro modo de realización, el procesador puede inicializar el indicador para tener un valor de "0". En dicho modo de realización, el procesador 13 puede no modificar el indicador cuando no se incluyen todas las imágenes de capa de referencia directa 410 para la capa actual 425A que no están restringidas para uso en la predicción entre capas en el conjunto de imágenes de referencia entre capas. El procesador 13 puede usar cualquier otro número de posibles consideraciones y/o determinaciones para determinar un ajuste de la indicación.

[0074] En un modo de realización alternativo, incluso si el procesador determina que todas las imágenes de capa de referencia directa 410 para la capa actual que no están restringidas para su uso en la predicción entre capas se incluyen en el conjunto de imágenes de referencia entre capas, en algunos casos, el procesador aún puede establecer la indicación en un valor de "0". Por ejemplo, el procesador puede establecer la indicación en un valor de "0" cuando el procesador determina que más de un número predeterminado de las imágenes de la capa de referencia directa 410 no están incluidas en el conjunto de imágenes de referencia entre capas aunque estén restringidas de uso en la predicción entre capas. En un modo de realización, el número predeterminado puede basarse en un umbral predeterminado (por ejemplo, un porcentaje), que puede variar en función del tipo de codificación, el tipo de imagen, un valor de cuadros por segundo y/o cualquier otro parámetro de imagen. Por ejemplo, si el número de imágenes de capa de referencia directa 410 es 10, el umbral predeterminado es 4, y el procesador determina que 5 de las 10 imágenes de capa de referencia directa 410 no están incluidas en el conjunto de imágenes de referencia entre capas, entonces en un modo de realización, el procesador puede establecer el valor de indicación en "0".

[0075] Una vez establecida la indicación (como se indica en la fila 460), el procesador 13 puede habilitar un dispositivo (por ejemplo, el codificador de vídeo 20) para enviar la señal a otro dispositivo (por ejemplo, el decodificador de vídeo 30). El codificador de vídeo 20 también puede señalar otras indicaciones basadas en varias determinaciones del procesador 13. Por ejemplo, el codificador de vídeo 20 puede señalar varias imágenes de capa de referencia directa con respecto a una imagen actual 405 y/o para todas las imágenes actuales 405 en la secuencia de vídeo 415. En un aspecto, el codificador de vídeo 20 puede indicar indicaciones a nivel de segmento. En un aspecto, basado en una o más de estas indicaciones, el decodificador de vídeo 30 puede determinar varias imágenes de capa de referencia activa a partir de las imágenes de capa de referencia directa, de tal manera que el decodificador de vídeo 30 puede determinar una cantidad de imágenes de referencia entre capas incluidas en un conjunto de imágenes de referencia entre capas (RPS), como se describe adicionalmente a continuación con respecto a la figura 6.

[0076] En el contexto de los procedimientos de extensión HEVC actuales descritos anteriormente, en un modo de realización, las técnicas descritas con respecto a la figura 4 puede usarse para modificar cierta semántica de HEVC para permitir que un codificador de vídeo establezca *all_ref_layers_active_flag* en 1 incluso cuando las imágenes de algunas subcapas no se usan para la predicción entre capas. En un modo de realización, esta característica puede habilitarse si, para cada imagen que no pertenece a esas subcapas, las imágenes de capa de referencia de todas las capas de referencia directa de la capa que contiene la imagen están presentes en la misma unidad de acceso que la imagen y están incluidas en el conjunto de imágenes de referencia entre capas de la imagen. Por ejemplo, una nueva definición semántica de *all_ref_layers_active_flag* puede incluir: "*all_ref_layers_active_flag* igual a 1 especifica que para cada imagen que hace referencia al SPV, las imágenes de la capa de referencia que pertenecen a todas las capas de referencia directa de la capa que contiene la imagen y que no están restringidas para usarse para la

predicción entre capas, según lo determinado por los valores de *max_tid_il_ref_pics_plus1 [i]*, están presentes en la misma unidad de acceso que la imagen, y están incluidos en el conjunto de imágenes de referencia entre capas de la imagen. *all_ref_layers_active_flag* igual a 0 especifica que la restricción anterior puede o no aplicarse".

5 **[0077]** En un modo de realización alternativo, el indicador descrito anteriormente se puede establecer en 0 en ciertas circunstancias cuando el TID máximo es 0. Ejemplo de lenguaje adicional para esta definición puede incluir "Cuando *max_tid_il_ref_pics_plus1 [i]* es igual a 0 para cualquier valor de *i* en el rango de 0 a *vps_max_layers_minus1*, inclusive, el valor de *all_ref_layers_active_flag* será igual a 0".

10 **[0078]** En otro modo de realización alternativa, el procesador 13 puede señalar el indicador de capa de referencia descrito anteriormente (por ejemplo, *all_ref_layers_active_flag*) en una posición anterior a cualquier indicador existente que pueda indicar la presencia de un TID máximo (por ejemplo, *max_tid_ref_present_flag*). En dicho modo de realización, cuando se establece el indicador de capa de referencia (por ejemplo, igual a 1), el indicador de presencia de TID máximo puede no establecerse (por ejemplo, igual a 0).

15 **[0079]** En otro modo de realización alternativa, cualquier indicador existente que pueda indicar la presencia de un TID máximo (por ejemplo, *max_tid_ref_present_flag*) puede no estar presente (por ejemplo, o eliminado de la sintaxis en el contexto de los procedimientos de extensión HEVC descritos anteriormente). Además, el procesador 13 puede señalar el indicador de capa de referencia descrito anteriormente (por ejemplo, *all_ref_layers_active_flag*) en una posición anterior a una variable que representa un valor de TID máximo (por ejemplo, *max_tid_il_ref_pics_plus1 [i]*). En tal modo de realización, cuando se establece el indicador de capa de referencia (por ejemplo, igual a 1), la variable que representa el valor de TID máximo puede no estar presente. Un ejemplo de implementación de esta realización alternativa se muestra en la tabla de sintaxis 1 a continuación:

25 Tabla 1

vps_extension() {	Descriptor
...	
<i>all_ref_layers_active_flag</i>	u(1)
if(<i>all_ref_layers_active_flag</i>)	
para(i = 0; i < <i>vps_max_layers_minus1</i> ; i++)	
<i>max_tid_il_ref_pics_plus1[i]</i>	u(3)
...	
}	

30 **[0080]** En otro modo de realización, el procesador 13 puede señalar el indicador de capa de referencia descrito anteriormente (por ejemplo, *all_ref_layers_active_flag*) en una posición anterior a cualquier indicador existente que pueda indicar la presencia de un TID máximo (por ejemplo, *max_tid_ref_present_flag*) y anterior a una variable que representa un valor TID máximo (por ejemplo, *max_tid_il_ref_pics_plus1 [i]*). En dicho modo de realización, cuando se establece el indicador de capa de referencia (por ejemplo, igual a 1) y la variable TID máxima está presente, el valor de la variable TID máxima puede ser 7. En otras palabras, el valor del TID máximo variable puede ser igual a una variable de *subcapa* VPS máxima (por ejemplo, *vps_max_sub_layers_minus1 + 1*).

35 **[0081]** La figura 5 ilustra un diagrama de flujo de un procedimiento para un procesador (por ejemplo, el procesador 13) de un codificador de vídeo (por ejemplo, el codificador de vídeo 20) para indicar si un conjunto de imágenes de referencia entre capas incluye todas las imágenes de capa de referencia directa que no están restringidas para uso en predicción entre capas, como se describe en relación con la figura 4.

40 **[0082]** El procedimiento 500 se inicia en el bloque 505. En el bloque 510, el procesador 13 puede determinar si cada una de las imágenes de la capa de referencia directa (por ejemplo, las imágenes de la capa de referencia directa 410A1, 410A2 de la figura 4) asociadas con una primera imagen actual (por ejemplo, la primera imagen actual CP1 de la figura 4) se incluyen en un conjunto de imágenes de referencia entre capas de la primera imagen actual CP1. El procedimiento puede repetirse para cada una de las imágenes en la secuencia de vídeo. El procesador 13 puede realizar esta determinación usando cualquiera de los procedimientos descritos anteriormente. En algunos modos de realización, el procesador 13 puede hacer esta determinación de otra manera.

45 **[0083]** El procesador 13 puede determinar (en el bloque 510) que cada una de las imágenes de capa de referencia directa asociadas con la imagen actual están incluidas en el conjunto de imágenes de referencia entre capas de la imagen actual. En tal caso, en el bloque 520, el procesador puede establecer una indicación que indica que todas las imágenes de capa de referencia directa para la capa actual que no están restringidas para su uso en la predicción entre capas se incluyen en el conjunto de imágenes de referencia entre capas de la actual imagen. En un aspecto, esto puede corresponder a cada uno de los valores en la fila 440 de la figura 4 siendo "Sí", independientemente de los valores en la fila 450 de la figura 4. Por lo tanto, establecer la indicación en el bloque 520 puede corresponder a

los procedimientos descritos con respecto a establecer un indicador en un valor de "1", como se indica en la fila 460 de la figura 4.

5 **[0084]** En cambio, el procesador 13 puede determinar (en el bloque 510) que no cada una de las imágenes de capa de referencia directa asociadas con la imagen actual están incluidas en el conjunto de imágenes de referencia entre capas de la imagen actual. Por ejemplo, como se ilustra en la figura 4, la imagen de capa de referencia directa 410A1 no se incluye en el conjunto de imágenes de referencia entre capas. En tal caso, en el bloque 530, el procesador puede determinar si alguna de las imágenes de capa de referencia directa asociadas con la imagen actual que no están incluidas en el conjunto de imágenes de referencia entre capas de la imagen actual no está restringida para su uso en predicción entre capas. Por ejemplo, el procesador 13 puede hacer esta determinación con respecto a la imagen de capa de referencia directa 410A1 usando los procedimientos descritos con respecto a la fila 450 de la figura 4. En algunos aspectos, el procesador 13 puede hacer esta determinación de otra manera, como se describe adicionalmente en relación con la figura 4.

15 **[0085]** El procesador puede determinar (en el bloque 530) que ninguna de las imágenes de capa de referencia directa asociadas con la imagen actual que no están incluidas en el conjunto de imágenes de referencia entre capas de la imagen actual 405 no está restringida para su uso en la predicción entre capas (o que todas las imágenes de capa de referencia directa asociadas con la imagen actual que no están incluidas en el conjunto de imágenes de referencia entre capas de la imagen actual están restringidas para su uso en la predicción entre capas). En tal caso, en el bloque 520, el procesador puede establecer una indicación que indica que todas las imágenes de capa de referencia directa para la capa actual que no están restringidas para su uso en la predicción entre capas se incluyen en el conjunto de imágenes de referencia entre capas de la actual imagen, como se analiza anteriormente.

25 **[0086]** Alternativamente, el procesador puede determinar (en el bloque 530), que al menos una de las imágenes de capa de referencia directa asociadas con la imagen actual que no están incluidas en el conjunto de imágenes de referencia entre capas de la imagen actual 405 no está restringida para su uso en predicción entre capas. En este caso, entonces en el bloque 540, el procesador puede establecer una indicación que indica que no todas las imágenes de capa de referencia directa para la capa actual que no están restringidas para su uso en la predicción entre capas se incluyen en el conjunto de imágenes de referencia entre capas. En un aspecto, esto puede corresponder a los procedimientos descritos con respecto a establecer un indicador en un valor de "0" como se describe en relación con la figura 4. Como se discutió anteriormente, en algunos modos de realización, los valores de bandera pueden inicializarse a un valor de inicialización. En tal caso, el procesador puede no establecer el valor del indicador si el valor a establecer es igual al valor inicializado.

35 **[0087]** El procedimiento termina en el bloque 550. Como se describió anteriormente en relación con la figura 4, el proceso anterior es solo un ejemplo de cómo el procesador 13 puede determinar la indicación.

40 **[0088]** La figura 6 ilustra un ejemplo de información de vídeo 600 que puede procesarse para determinar un número de imágenes de capa de referencia activas a partir de una o más imágenes de capa de referencia directa (también denominadas en el presente documento "imágenes de referencia directa", "imágenes de referencia" o "imágenes de capa de referencia"). Las imágenes de capa de referencia activa son imágenes en capas de referencia que se están procesando actualmente mediante predicción entre capas. Las imágenes de la capa de referencia activa corresponden a las imágenes de la capa de referencia directa particular que se utilizan para decodificar la imagen actual. En un modo de realización, un conjunto de imágenes de referencia solo contiene imágenes de capa de referencia activas. Por lo tanto, en un modo de realización, los términos "conjunto de imágenes de referencia activas" y "conjunto de imágenes de referencia" pueden usarse indistintamente. En un aspecto, una o más de las características descritas con respecto a la figura 6 puede realizarse generalmente mediante un decodificador (por ejemplo, el decodificador de vídeo 30 de las figuras 1A, 1B y/o 3). En un aspecto, el decodificador de vídeo 30 puede recibir varias indicaciones desde un codificador (por ejemplo, el codificador de vídeo 20 de las figuras 1A, 1B y/o 2). En otro modo de realización, las indicaciones pueden recibirse desde un dispositivo diferente o pueden almacenarse en una unidad de memoria (por ejemplo, el dispositivo de almacenamiento 34 de la figura 1A, la memoria 92 de la figura 3, etc.). Otros modos de realización descritos con respecto a la figura 6 también se pueden incluir en la unidad de memoria. Las operaciones descritas con respecto a la figura 6 pueden realizarse mediante un procesador (por ejemplo, un procesador dentro del módulo de origen, módulo de destino, codificador de vídeo y/o decodificador de vídeo de una cualquiera o más de las figuras 1A-3, el módulo procesador/controlador 13 de la figura 1B, y/o una unidad de procesamiento de vídeo, tal como la unidad de procesamiento de vídeo 21 de la figura 1B, etc.). El decodificador de vídeo, el codificador de vídeo, la unidad de memoria y/o el procesador pueden estar en comunicación entre sí y/o conectados operativamente.

50 **[0089]** En un modo de realización, antes de determinar un número (por ejemplo, una cantidad) de imágenes de capa de referencia activa a partir de la una o más imágenes de capa de referencia directa, el decodificador de vídeo 30 puede determinar primero si todas las una o más imágenes de capa de referencia directa que no están restringidas para su uso en la predicción entre capas que se incluyen en un conjunto de imágenes de referencia entre capas asociado con una imagen actual. En un aspecto, el conjunto de imágenes de referencia entre capas puede comprender todo o una parte del conjunto de imágenes de referencia entre capas descrito en relación con las figuras 4-5. En un aspecto, el conjunto de imágenes de referencia entre capas también puede incluir otras imágenes. En un modo de realización, el decodificador de vídeo 30 puede determinar si todas las una o más imágenes de capa de referencia

directa que no están restringidas para uso en la predicción entre capas se incluyen en un conjunto de imágenes de referencia entre capas asociado con una imagen actual basada en una indicación. Por ejemplo, la indicación puede establecerse y almacenarse y/o proporcionarse al decodificador de vídeo 30 mediante el codificador de vídeo 20. En tal ejemplo, la indicación puede comprender la indicación descrita con respecto a la fila 460 de la figura 4. En un ejemplo, la indicación puede ser un indicador establecida en "1". En otros modos de realización, el decodificador de vídeo 30 puede hacer esta determinación de otras maneras. Como se describe con respecto a la figura 4, el decodificador de vídeo 30 también puede recibir una o más imágenes de capa de referencia directa del codificador de vídeo 20 o de alguna otra fuente.

5
10
15
20
[0090] Una secuencia de vídeo 615 incluye unidades de acceso (denominadas colectivamente como unidades de acceso 620), donde cada una de las unidades de acceso 620 puede contener una o más imágenes. La secuencia de vídeo 615 de la figura 6 ilustra dos unidades de acceso AU1, AU2; sin embargo, la secuencia de vídeo 615 puede incluir cualquier número de unidades de acceso 620. Cada unidad de acceso AU1, AU2 puede estar asociada con una o más capas 625 de información de vídeo, como se ilustra mediante una capa base BL y tres capas de mejora EL1, EL2 y EL3. La capa o capa ilustrada más baja más baja de la información de vídeo 600 puede ser la capa base (BL) o una capa de referencia (RL), y la capa en el nivel más alto o en el nivel más alto de la información de vídeo 600 puede ser una capa mejorada. La "capa mejorada" puede considerarse sinónimo con una "capa de mejora", y estos términos se pueden usar indistintamente. En este ejemplo, la capa BL es la capa base y las capas EL1, EL2 y EL3 son capas de mejora primera, segunda y tercera, respectivamente. Las capas entre la capa base BL y la capa de mejora más alta EL3 pueden servir como capas de mejora y/o capas de referencia. En este ejemplo, las capas EL1, EL2 y EL3 representan capas de mejora.

25
30
[0091] Por ejemplo, una capa dada (por ejemplo, la segunda capa de mejora EL2) puede ser una capa de mejora para una capa por debajo (es decir, que precede) de la capa dada, como la capa base (por ejemplo, la capa base BL) o cualquier capa de mejora intermedia (por ejemplo, la primera capa de mejora EL1). Además, la capa dada (por ejemplo, la segunda capa de mejora EL2) también puede servir como una capa de referencia para una capa de mejora (por ejemplo, la tercera capa de mejora EL3) por encima (es decir, posterior a) de la capa dada. Cualquier capa dada entre la capa base (es decir, la capa más baja que tiene, por ejemplo, un conjunto de identificación de capa (ID) o igual a "1") y la capa superior (o la capa más alta) puede usarse como referencia para predicción entre capas por una capa más alta en relación con la capa dada y puede determinarse usando una capa más baja a la capa dada como referencia para la predicción entre capas.

35
40
[0092] La capa 625 que se está procesando actualmente puede denominarse la "capa actual". En la ilustración de ejemplo, la tercera capa de mejora EL3 representa la capa actual. Una capa actual puede estar asociada con una o más capas de referencia directa. Como se mencionó anteriormente, una capa de referencia directa es una capa que se puede usar para la predicción entre capas de otra capa, tal como la capa actual. En el ejemplo ilustrado, la primera capa de mejora EL1 y la capa base BL representan la primera y segunda capas de referencia directa de la capa actual (por ejemplo, que en el modo de realización ilustrado es la tercera capa de mejora EL3), respectivamente. La primera capa de mejora EL1 y la capa base BL son capas de referencia directa de la capa actual porque se determina que son capas que pueden usarse para la predicción interna de la tercera capa de mejora, que en este ejemplo es la capa actual.

45
50
55
60
[0093] Cada una de las capas 625 puede comprender una o más imágenes, estando asociada cada imagen con una de las capas 625 y una de las unidades de acceso AU1, AU2. Las imágenes asociadas con la capa actual (en este ejemplo, la tercera capa de mejora EL3) pueden denominarse colectivamente como las imágenes actuales 605. Por ejemplo, en el modo de realización ilustrado, las imágenes actuales 605 incluyen una primera imagen actual CP1 ubicada dentro de la primera unidad de acceso AU1 y una segunda imagen actual CP2 ubicada dentro de la segunda unidad de acceso AU2. Se puede hacer referencia a las imágenes (por ejemplo, imágenes 610A1, 610A2, 610B1, 610B2) que están en capas de referencia directa (por ejemplo, la primera capa de mejora EL1 y la capa base BL) de la capa actual (por ejemplo, la tercera capa de mejora EL3) colectivamente como imágenes de capa de referencia directa 610 (o imágenes de capa de referencia 610). Por ejemplo, la imagen de capa de referencia directa 610B1 es la imagen en la primera capa de referencia directa (en este ejemplo, la primera capa de mejora EL1) y la segunda unidad de acceso AU2. Esta imagen de capa de referencia directa 610B1 también puede denominarse "imagen de capa de referencia directa 1 de la imagen actual 2". Como otro ejemplo, la imagen 610A2 de la capa de referencia directa es la imagen en la segunda capa de referencia directa (en este ejemplo, la capa base BL) y la primera unidad de acceso AU1. Esta imagen de capa de referencia directa 610A1 también puede denominarse "imagen de capa de referencia directa 2 de la imagen actual 1". Las imágenes que están en capas de referencia no directas de la capa actual (por ejemplo, la capa de referencia no directa EL2) pueden denominarse simplemente imágenes 608. En el modo de realización ilustrado, las imágenes 608 incluyen una imagen 608A ubicada en la primera unidad de acceso AU1 y una segunda imagen 608B ubicada en la segunda unidad de acceso AU2. En un modo de realización, el procesador (por ejemplo, el procesador 13) no puede considerar o hacer ninguna determinación basada en las imágenes 608A y 608B en la capa de referencia no directa (por ejemplo, la segunda capa de mejora EL2).

65
[0094] Cada una de las imágenes de capa de referencia directa 610 puede estar asociada con un número o valor de identificación temporal respectivo o "valor TID" (no mostrado). En un aspecto, cada una de las imágenes en una de las unidades de acceso AU1, AU2 puede estar asociada con el mismo valor TID (es decir, uno común). Por ejemplo,

para la primera unidad de acceso AU1, la imagen actual CP1, la imagen 608A y ambas imágenes de capa de referencia directa 610A1 y 610A2 pueden estar asociadas con el mismo valor TID. En algunos aspectos, el procesador 13 puede determinar respectivamente si el valor de TID asociado de cada una de las imágenes 610 de capa de referencia directa es menor que un umbral predeterminado (por ejemplo, "valor de identificación temporal máximo", "número de identificación temporal máximo" o "TID máximo"). En algunos modos de realización, el TID máximo puede estar asociado con la capa actual y la secuencia de vídeo 615. En el ejemplo ilustrado, la imagen de capa de referencia directa 610A1 está asociada con un valor de TID que es mayor o igual al TID máximo, y la imagen de capa de referencia directa 610B1 está asociada con un valor de TID que es menor que el TID máximo. En un modo de realización, el TID máximo puede ser un valor de "4", el valor TID asociado con la imagen de capa de referencia directa 610A1 (en la primera unidad de acceso AU1) puede ser "5", y el valor TID asociado con la referencia directa la imagen de capa 610B1 (en la segunda unidad de acceso AU2) puede ser "2". En tal caso, la imagen de capa de referencia directa 610A1 no se determinaría como una imagen de capa de referencia activa de la imagen actual porque el valor de TID asociado con la imagen de capa de referencia directa 610A1 es mayor o igual que el TID máximo. Sin embargo, en tal caso, la imagen de capa de referencia directa 610B1 se determinaría como una imagen de capa de referencia activa de la imagen actual porque el valor de TID asociado con la imagen de capa de referencia directa 610B1 no es mayor o igual que el TID máximo. En algunos modos de realización, el TID máximo puede ser un valor constante para toda la secuencia de vídeo 615. Por ejemplo, el TID máximo puede permanecer constante para toda la secuencia de vídeo 615 según un estándar de codificación de vídeo (por ejemplo, la variable `max_tid_il_ref_pics_plus1 [i]` del estándar HEVC). En otros modos de realización, el TID máximo puede variar según la unidad de acceso AU1, AU2 en la secuencia de vídeo 615. En otras palabras, cada unidad de acceso respectiva AU1, AU2 puede estar asociada con un TID máximo respectivo.

[0095] El procesador 13 puede determinar diversa información usando las propiedades de cada una de las imágenes 610 de capa de referencia directa y la imagen actual asociada de cada imagen de capa de referencia directa 605. La fila 630 de la figura 6 ilustra un modo de realización de dicha información. Por ejemplo, los valores en la fila 630 representan el número de imágenes de capa de referencia activas de una o más imágenes de capa de referencia directa 610 en cada una de las unidades de acceso AU1, AU2.

[0096] Como se describió anteriormente, para cada unidad de acceso de la secuencia de vídeo 615, un procesador 13 puede determinar varias imágenes de capa de referencia activa a partir de una o más imágenes de capa de referencia directa 610. Por ejemplo, con respecto a la primera unidad de acceso AU1, un procesador 13 puede determinar varias imágenes de capa de referencia activa a partir de las imágenes de capa de referencia directa de la imagen actual de la primera unidad de acceso 610A1 y 610A2. En un modo de realización, el procesador 13 puede determinar que el número de imágenes de la capa de referencia activa sea igual a un número de imágenes de la capa de referencia directa 610 que están asociadas con un valor de TID menor que el TID máximo. Con respecto a la unidad de acceso AU1, las dos imágenes de capa de referencia directa 610A1 y 610A2 están asociadas con un valor de TID que es mayor o igual que el TID máximo. Por lo tanto, el procesador 13 puede no contar ninguna de las imágenes de capa de referencia directa 610A1 o 610A2 como imágenes de capa de referencia activa. Por lo tanto, en el modo de realización ilustrado para la unidad de acceso AU1, el procesador 13 puede determinar que el número de imágenes de capa de referencia activa de una o más imágenes de capa de referencia directa 610 es igual a "0", como se muestra en la porción de la fila 630 correspondiente a la primera unidad de acceso AU1.

[0097] Con respecto a la segunda unidad de acceso AU2, el procesador 13 puede determinar de manera similar una serie de imágenes de capa de referencia activa a partir de las imágenes de capa de referencia directa de la imagen actual de la segunda unidad de acceso 610B1 y 610B2. Como se describió anteriormente, en un modo de realización, el procesador 13 puede determinar que el número de imágenes de la capa de referencia activa sea igual a un número de imágenes de la capa de referencia directa 610 que están asociadas con un valor de TID menor que el TID máximo. Con respecto a la unidad de acceso AU2, las dos imágenes de capa de referencia directa 610B1 y 610B2 están asociadas con un valor de TID que es menor que el TID máximo. Por lo tanto, el procesador 13 puede contar las imágenes de capa de referencia directa 610B1 y 610B2 como imágenes de capa de referencia activa. Por lo tanto, en el modo de realización ilustrado para la unidad de acceso AU2, el procesador 13 puede determinar que el número de imágenes de capa de referencia activa de una o más imágenes de capa de referencia directa 610 es igual a "2", como se muestra en la porción de la fila 630 correspondiente a la segunda unidad de acceso AU2.

[0098] El procesador 13 puede usar cualquier otro número de posibles consideraciones y/o determinaciones para determinar el número de imágenes de la capa de referencia activa. En un modo de realización alternativa, el procesador 13 puede determinar cuáles de las imágenes de capa de referencia directa 610 no son imágenes de capa de referencia activas (por ejemplo, determinando un número de imágenes de capa de referencia directa 610 que están asociadas con un valor TID mayor o igual que el TID máximo) y luego determinar una cantidad de imágenes de capa de referencia activa en función de un valor de diferencia.

[0099] Una vez determinado el número de imágenes de la capa de referencia activa (como se indica en la fila 630), el procesador 13 puede indicar dicha determinación como un valor variable según un estándar de decodificación de vídeo (por ejemplo, la variable `NumActiveRefLayerPics` del estándar HEVC). En un modo de realización, el procesador 13 puede definir un conjunto de imágenes de referencia entre capas (RPS) (por ejemplo, un "conjunto de imágenes de referencia activo" o un "conjunto de imágenes de referencia") de modo que contenga las imágenes de capa de

referencia activa asociadas. El procesador 13 puede decodificar la imagen actual asociada 605 usando predicción entre capas basada en al menos una de las imágenes de capa de referencia activa. Por ejemplo, en base al análisis descrito anteriormente, el procesador 13 puede decodificar la imagen actual 605 en AU2 usando predicción entre capas basada en una o más de las imágenes activas de capa de referencia 610B1 y/o 610B2.

[0100] En el contexto de los procedimientos de extensión HEVC actuales descritos anteriormente, en un modo de realización, ciertas definiciones de variables estándar HEVC también pueden modificarse para aprovechar las modificaciones semánticas como se describe en relación con la figura 4. Por ejemplo, ciertas modificaciones de definición variable pueden permitir que el decodificador de vídeo 30 cuente imágenes selectivamente durante la derivación de ciertas variables (por ejemplo, *NumActiveRefLayerPics*). Es decir, las imágenes de capa de referencia que pertenecen a subcapas que no se usan para la predicción entre capas (por ejemplo, como lo indica un TID máximo, que podría ser la variable *max_tid_il_ref_pics_plus1* descrita anteriormente) no pueden incluirse en el conjunto de imágenes de referencia entre capas. Por ejemplo, el nuevo código para derivar *NumActiveRefLayerPics* puede incluir lo siguiente:

```

5
15
20
25
30
    if(nuh_layer_id == 0 || NumDirectRefLayers[nuh_layer_id] == 0)
        NumActiveRefLayerPics = 0;
    else if(all_ref_layers_active_flag)
        for (i = 0, NumActiveRefLayerPics = 0; i < NumDirectRefLayers[
nuh_layer_id ]; i++) {
            layerIdx = LayerIdxInVps[RefLayerId[nuh_layer_id][ i ] ]
            if(max_tid_il_ref_pics_plus1[ layerIdx ] > TemporalId)
                }; NumActiveRefLayerPics++
    else if (!inter_layer_pred_enabled_flag)
        NumActiveRefLayerPics = 0;
    else if(max_one_active_ref_layer_flag || NumDirectRefLayers[nuh_layer_id ]
== 1)
        NumActiveRefLayerPics = 1;
    else
        NumActiveRefLayerPics = num_inter_layer_ref_pics_minus1 + 1;

```

[0101] Como se describió anteriormente en relación con la figura 4, las modificaciones semánticas y los procedimientos descritos en la presente divulgación pueden hacer que *all_ref_layers_active_flag* se establezca en 1 con mayor frecuencia que en los procedimientos actuales de extensiones HEVC. Por lo tanto, el nuevo código anterior para derivar *NumActiveRefLayerPics* puede causar que el primer 'else if' en el bloque de código se utilice con mayor frecuencia (por ejemplo, porque *all_ref_layers_active_flag* se establece en 1 con mayor frecuencia) que en los procedimientos actuales de extensiones HEVC. En un modo de realización, los parámetros del primer 'else if' en el bloque de código pueden representar una parte de los procedimientos descritos en relación con las figuras 6-7.

[0102] La figura 7 ilustra un diagrama de flujo de un procedimiento para un procesador (por ejemplo, el procesador 13) de un decodificador de vídeo (por ejemplo, el decodificador de vídeo 30) para determinar un número de imágenes de capa de referencia activa de una o más imágenes de capa de referencia directa como se describe en conexión con la figura 6.

[0103] El procedimiento 700 se inicia en el bloque 705. En el bloque 710, el procesador 13 puede determinar si una o más imágenes de capa de referencia directa que no están restringidas para su uso en la predicción entre capas se incluyen en un conjunto de imágenes de referencia entre capas asociado con una imagen actual. En un aspecto, el conjunto de imágenes de referencia entre capas puede comprender todo o una parte del conjunto de imágenes de referencia entre capas descrito en relación con las figuras 4-5. Por ejemplo, el conjunto de imágenes de referencia entre capas puede comprender cada una de las imágenes de capa de referencia directa 610A1, 610A2, 610B1 y 610B2 de la figura 6. En un modo de realización, el procesador 13 puede determinar si todas las una o más imágenes de capa de referencia directa que no están restringidas para uso en la predicción entre capas se incluyen en un conjunto de imágenes de referencia entre capas asociado con una imagen actual basada en una indicación proporcionada por un codificador de vídeo (por ejemplo, el codificador de vídeo 20), como se describe adicionalmente en relación con las figuras 4-6. En otros modos de realización, el procesador 13 puede hacer la determinación usando cualquiera de los procedimientos descritos anteriormente. En algunos modos de realización, el procesador 13 puede hacer la determinación de otra manera.

[0104] El procesador 13 puede determinar (en el bloque 710) que todas las una o más imágenes de capa de referencia directa que no están restringidas para uso en la predicción entre capas se incluyen en el conjunto de imágenes de referencia entre capas asociado con la imagen actual. En un ejemplo, el procesador 13 puede hacer tal determinación basándose en recibir un indicador establecido en "1" como se describe en conexión con la figura 6. En tal caso, en el bloque 720, el procesador 13 puede determinar una cantidad de una o más imágenes de capa de referencia directa que son imágenes de capa de referencia activa. En un modo de realización, el número de una o más imágenes de capa de referencia directa que son imágenes de capa de referencia activa puede ser igual a un número de imágenes de capa de referencia directa que están asociadas con un valor TID menor que un TID máximo.

El valor de TID y el TID máximo pueden ser el valor de TID y el TID máximo como se describe en relación con la figura 6. En un modo de realización, el procesador 13 puede hacer la determinación descrita en relación con el bloque 720 para cada unidad de acceso. Por ejemplo, haciendo referencia a AU1 de la figura 6, el procesador 13 puede determinar un número de imágenes de capa de referencia directa 610A1 y 610A2 que están asociadas con un valor TID menor que el TID máximo y luego determinar por separado un número de imágenes de capa de referencia directa 610B1 y 610B2 que están asociadas con un valor TID menor que el TID máximo. Como se describió anteriormente, dicha determinación también puede describirse como la determinación de varias imágenes de capa de referencia activa a partir de una o más imágenes de capa de referencia directa. Por ejemplo, dicha determinación puede corresponder a las determinaciones descritas con respecto a la fila 630 de la figura 6. Después, el procedimiento puede finalizar en el bloque 730. Como se describió anteriormente en relación con la figura 6, el proceso anterior es solo un ejemplo de cómo el procesador 13 puede hacer tal determinación.

[0105] El procesador 13 puede determinar en su lugar (en el bloque 710) que no todas las una o más imágenes de capa de referencia directa que no están restringidas para uso en la predicción entre capas se incluyen en el conjunto de imágenes de referencia entre capas asociado con la imagen actual. En un ejemplo, el procesador 13 puede hacer tal determinación basándose en recibir un indicador establecido en "0" como se describe en conexión con la figura 6. En tal caso, el procedimiento puede terminar en el bloque 730. Como se describió anteriormente en relación con la figura 6, el proceso anterior es solo un ejemplo de cómo el procesador 13 puede hacer tal determinación.

[0106] 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 se pueden almacenar en o transmitirse por, como una o más instrucciones o código, 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 corresponden a un medio tangible tal como medios de almacenamiento de datos, o medios de comunicación que incluyen cualquier medio que facilite la transferencia de un programa informático desde un lugar a otro, por ejemplo, de acuerdo a un protocolo de comunicación. De esta manera, los medios legibles por ordenador pueden corresponder, en general, a (1) medios de almacenamiento tangibles legibles por ordenador que sean no transitorios o (2) un medio de comunicación tal como una señal o una onda portadora. Los medios de almacenamiento de datos pueden ser 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.

[0107] A modo de ejemplo, y no de limitación, 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. Además, cualquier conexión recibe adecuadamente la denominación de medio legible por ordenador. Por ejemplo, si las instrucciones se transmiten desde un sitio web, un servidor u otro origen remoto usando un cable coaxial, un cable de fibra óptica, un par trenzado, una línea de abonado digital (DSL) o tecnologías inalámbricas tales como infrarrojos, radio y microondas, entonces el cable coaxial, el cable de fibra óptica, el par trenzado, la DSL o las tecnologías inalámbricas tales como infrarrojos, radio y microondas se incluyen en la definición de medio. 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 ni otros medios transitorios, sino que, en cambio, se orientan a medios de almacenamiento tangibles no transitorios. El término disco, como se usa en el presente documento, incluye el disco compacto (CD), el disco láser, el disco óptico, el disco versátil digital (DVD), el disco flexible y el disco Blu-ray, donde algunos discos normalmente reproducen datos de manera magnética, mientras que otros discos reproducen datos ópticamente con láseres. Las combinaciones de lo anterior se deberían incluir también dentro del alcance de los medios legibles por ordenador.

[0108] Las instrucciones pueden ser ejecutadas por uno o más procesadores, tales como uno o más procesadores de señales digitales (DSP), microprocesadores de propósito general, circuitos integrados específicos de la aplicación (ASIC), formaciones lógicas programables in situ (FPGA) u otros circuitos lógicos equivalentes, integrados o discretos. En consecuencia, el término "procesador", como se usa en el presente documento, se puede referir a cualquiera de las estructuras anteriores o a cualquier otra estructura adecuada para la implementación de las técnicas descritas en el presente documento. Además, en algunos aspectos, la funcionalidad descrita en el presente documento se puede proporcionar en módulos de hardware y/o software dedicados configurados para la codificación y la decodificación, o incorporados en un códec combinado. Además, las técnicas se podrían implementar totalmente en uno o más circuitos o elementos lógicos.

[0109] Las técnicas de esta divulgación se pueden implementar en una amplia variedad de dispositivos o aparatos, incluidos un equipo manual inalámbrico, un circuito integrado (IC) o un conjunto de IC (por ejemplo, un conjunto de chips). Diversos componentes, módulos o unidades se describen en esta divulgación para enfatizar aspectos funcionales de dispositivos configurados para realizar las técnicas divulgadas, pero no requieren necesariamente su realización mediante diferentes unidades de hardware. En cambio, como se ha descrito anteriormente, diversas unidades se pueden combinar en una unidad de hardware de códec o proporcionarse mediante un grupo de unidades

de hardware interoperativas, incluidos uno o más procesadores, como se ha descrito anteriormente, conjuntamente con software y/o firmware adecuados. Se han descrito diversos ejemplos. Estos y otros ejemplos están dentro del alcance de las siguientes reivindicaciones.

REIVINDICACIONES

1. Un codificador de vídeo (20), que comprende:

5 una unidad de memoria configurada para almacenar una o más imágenes de capa de referencia directa (410) de una o más imágenes actuales (405) en una secuencia (415), en el que la una o más imágenes actuales están asociadas con una capa actual (425), estando la capa actual asociada con una o más capas de referencia directa, y en el que una capa de referencia directa es una capa que puede usarse para la predicción entre capas de otra capa, tal como la capa actual; y

10 un procesador (13) en comunicación con la unidad de memoria y configurado para:

15 determinar que una imagen de capa de referencia directa de una o más imágenes de capa de referencia directa no está restringida para su uso, de modo que la imagen de capa de referencia directa se puede usar, en la predicción entre capas, cuando la imagen de la capa de referencia directa tiene un valor de identificación temporal menor que un umbral predeterminado basado en un valor de identificación temporal máximo asociado con la capa actual;

20 determinar si todas las una o más imágenes de capa de referencia directa de una imagen actual que no están restringidas para uso en la predicción entre capas se incluyen en un conjunto de imágenes de referencia entre capas asociado con la imagen actual; y

25 establecer un indicador asociado con la imagen actual a un primer valor basado en la determinación de que todas las imágenes de capa de referencia directa de la imagen actual que no están restringidas para su uso en la predicción entre capas se incluyen en el conjunto de imágenes de referencia entre capas asociado con la imagen actual; o

30 establecer el indicador en un segundo valor basado en la determinación de que no todas las imágenes de capa de referencia directa de la imagen actual que no están restringidas para su uso en la predicción entre capas se incluyen en el conjunto de imágenes de referencia entre capas asociado con la imagen actual.

35 2. El codificador de vídeo de la reivindicación 1, en el que el procesador está configurado además para establecer una indicación respectiva para cada una de las una o más imágenes actuales en la secuencia en función de la indicación asociada con la imagen actual.

40 3. El codificador de vídeo de la reivindicación 1, en el que el procesador está configurado además para determinar si cada imagen de capa de referencia directa respectiva está incluida en el conjunto de imágenes de referencia entre capas en función de una determinación de si la imagen de capa de referencia directa respectiva puede usarse para predicción de entre capas de la imagen actual.

4. Un procedimiento de codificación de vídeo, comprendiendo el procedimiento:

45 almacenar una o más imágenes de capa de referencia directa (410) de una o más imágenes actuales (405) en una secuencia (415), en el que la una o más imágenes actuales están asociadas con una capa actual (425), estando la capa actual asociada con una o más capas de referencia directa, y en el que una capa de referencia directa es una capa que puede usarse para la predicción entre capas de otra capa, tal como la capa actual; y

50 determinar que una imagen de capa de referencia directa de una o más imágenes de capa de referencia directa no está restringida para su uso, de modo que la imagen de capa de referencia directa se puede usar, en la predicción entre capas, cuando la imagen de la capa de referencia directa tiene un valor de identificación temporal menor que un umbral predeterminado basado en un valor de identificación temporal máximo asociado con la capa actual;

55 determinar si todas las una o más imágenes de capa de referencia directa de una imagen actual que no están restringidas para uso en la predicción entre capas se incluyen en un conjunto de imágenes de referencia entre capas asociado con la imagen actual; y

60 establecer un indicador asociado con la imagen actual a un primer valor basado en la determinación de que todas las imágenes de capa de referencia directa de la imagen actual que no están restringidas para su uso en la predicción entre capas se incluyen en el conjunto de imágenes de referencia entre capas asociado con la imagen actual; o

65 establecer el indicador en un segundo valor basado en la determinación de que no todas las imágenes de capa de referencia directa de la imagen actual que no están restringidas para su uso en la predicción entre capas se incluyen en el conjunto de imágenes de referencia entre capas asociado con la imagen actual.

5. El procedimiento de la reivindicación 4, que comprende además establecer una indicación respectiva para cada una de las una o más imágenes actuales en la secuencia en base a la indicación asociada con la imagen actual.
- 5 6. El procedimiento de la reivindicación 4, que comprende además determinar si cada imagen de capa de referencia directa respectiva se incluye en el conjunto de imágenes de referencia entre capas en base a una determinación de si la imagen de capa de referencia directa puede usarse para la predicción entre capas de la imagen actual.
- 10 7. Un decodificador de vídeo (30), que comprende:
- una unidad de memoria está configurada para almacenar una o más imágenes de capa de referencia directa (610), en el que cada una de las imágenes de capa de referencia directa está asociada con una capa de referencia directa respectiva de una o más capas de referencia directa, en el que la una o más capas de referencia directa están asociadas con una capa actual (625), y en el que la capa actual está asociada con una imagen actual (605), y en el que una capa de referencia directa es una capa que puede usarse para la predicción entre capas de otra capa, tal como la capa actual; y
- 15 un procesador (13) en comunicación con la memoria y configurado para:
- 20 determinar, basándose en un indicador, si todas las una o más imágenes de capa de referencia directa que no están restringidas para uso en la predicción entre capas se incluyen en un conjunto de imágenes de referencia entre capas asociado con la imagen actual; y
- 25 determinar un número de una o más imágenes de capa de referencia directa que no están restringidas para su uso en la predicción entre capas posterior a la determinación de que todas las una o más imágenes de capa de referencia directa que no están restringidas para su uso en la predicción entre capas se incluyen en el conjunto de imágenes de referencia entre capas asociado con la imagen actual, en el que el número determinado de una o más imágenes de capa de referencia directa se puede usar en la predicción entre capas,
- 30 en el que el procesador está configurado para determinar que una imagen de capa de referencia directa de una o más imágenes de capa de referencia directa no está restringida para su uso, de modo que la imagen de capa de referencia directa se puede usar, en la predicción entre capas, determinando que el valor de identificación temporal asociado con la imagen de capa de referencia directa es menor que un umbral predeterminado basado en un valor de identificación temporal máximo asociado con la capa actual.
- 35 8. Un procedimiento para decodificar vídeo, comprendiendo el procedimiento:
- 40 almacenar una o más imágenes de capa de referencia directa (610), en el que cada una de las imágenes de capa de referencia directa está asociada con una capa de referencia directa respectiva de una o más capas de referencia directa, en el que la una o más capas de referencia directa están asociadas con una capa actual (625), y en el que la capa actual está asociada con una imagen actual (605), y en el que una capa de referencia directa es una capa que puede usarse para la predicción entre capas de otra capa, tal como la capa actual;
- 45 determinar, basándose en un indicador, si todas las una o más imágenes de capa de referencia directa que no están restringidas para uso en la predicción entre capas se incluyen en un conjunto de imágenes de referencia entre capas asociado con la imagen actual; y
- 50 determinar un número de una o más imágenes de capa de referencia directa que no están restringidas para su uso en la predicción entre capas posterior a la determinación de que todas las una o más imágenes de capa de referencia directa que no están restringidas para su uso en la predicción entre capas se incluyen en el conjunto de imágenes de referencia entre capas asociado con la imagen actual, en el que el número determinado de una o más imágenes de capa de referencia directa se puede usar en la predicción entre capas,
- 55 en el que determinar que una imagen de capa de referencia directa de una o más imágenes de capa de referencia directa no está restringida para su uso, de modo que la imagen de capa de referencia directa se puede usar, en la predicción entre capas, comprende determinar que el valor de identificación asociado con la imagen de capa de referencia directa es menor que un umbral predeterminado basado en un valor de identificación temporal máximo asociado con la capa actual.
- 60

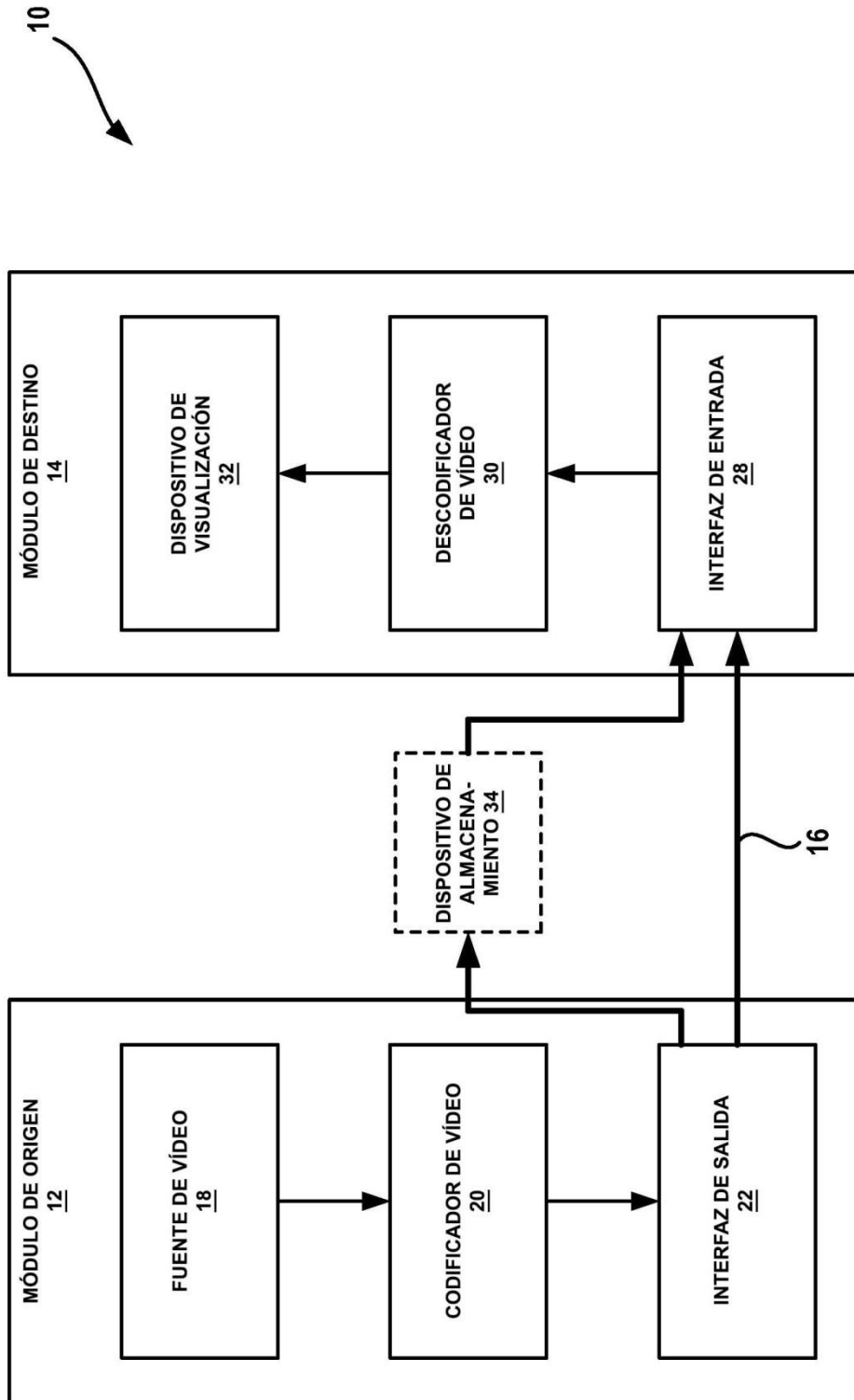


FIG. 1A

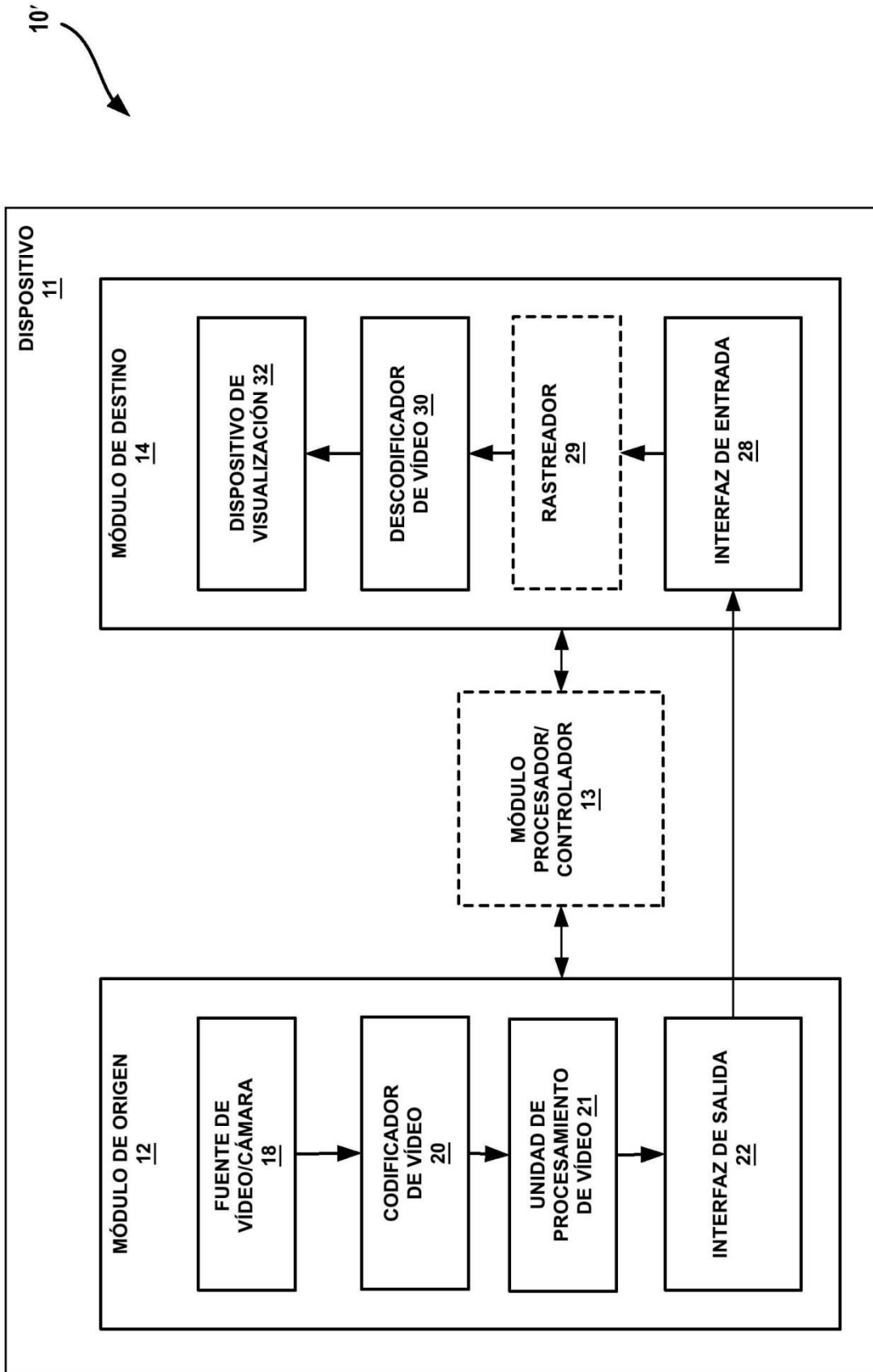
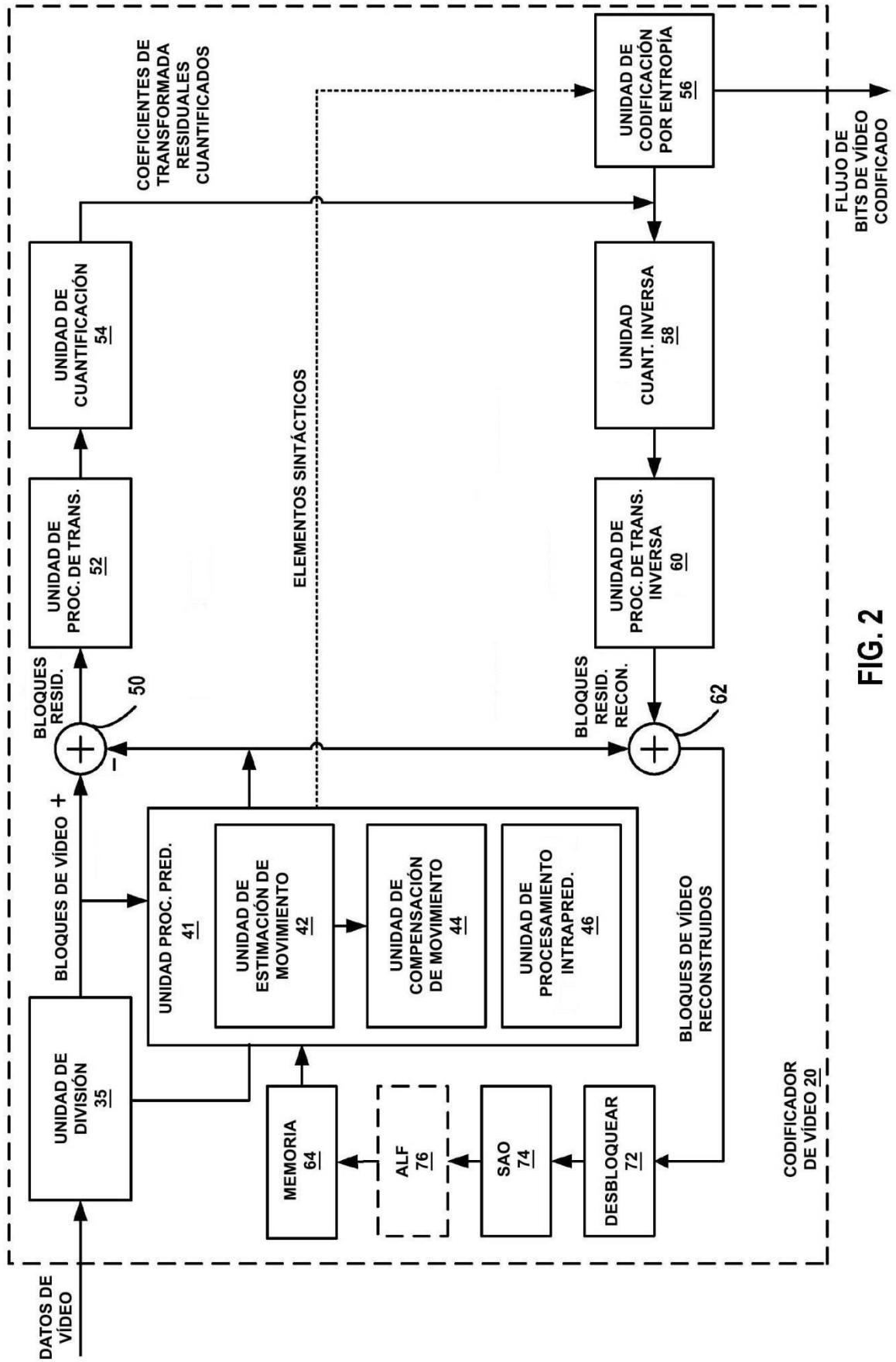


FIG. 1B



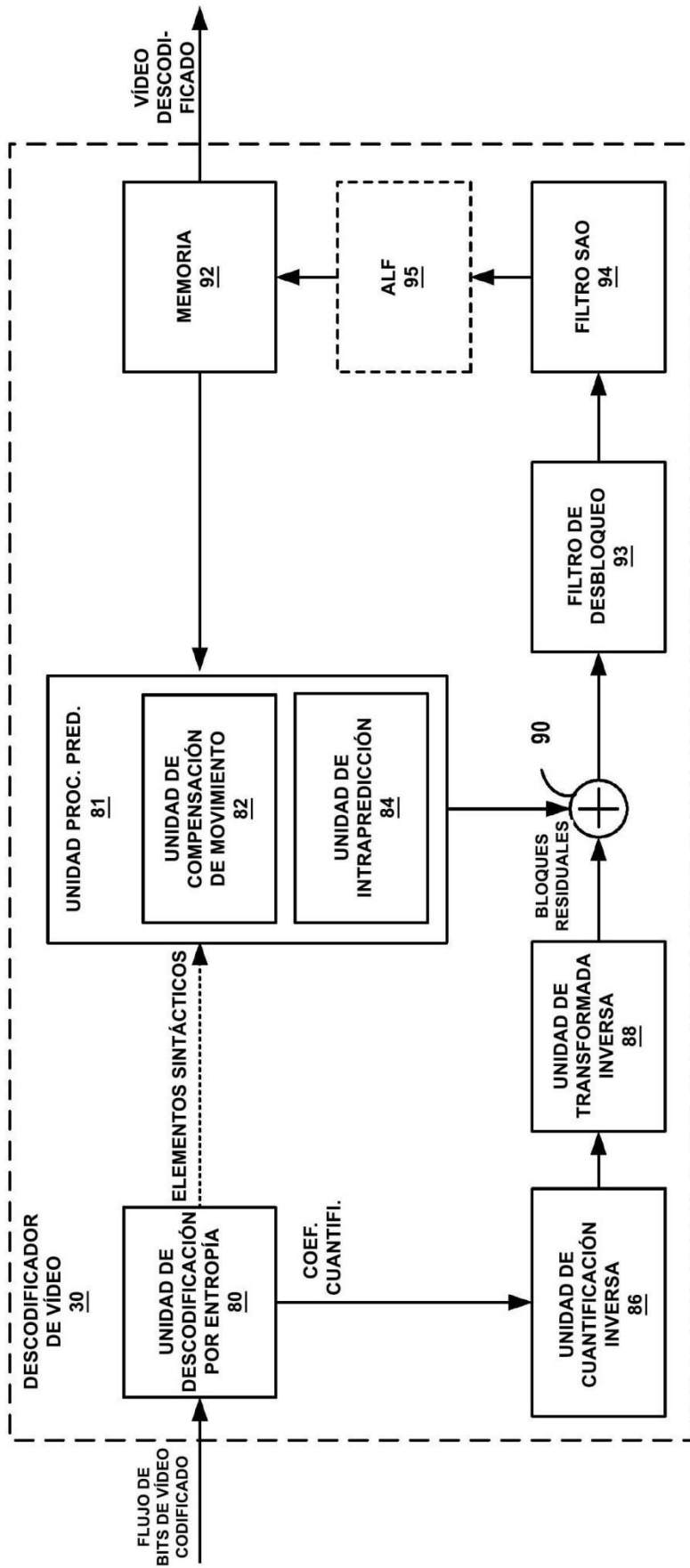


FIG. 3

400

	AU1	AU2
CP1	<input type="checkbox"/>	CP2 <input type="checkbox"/>
405	↔	
408A	↔	408B <input type="checkbox"/>
DRLP1	<input type="checkbox"/>	DRLP1 <input type="checkbox"/>
410A1	↔	410B1 <input type="checkbox"/>
TID < Max TID		TID < Max TID
DRLP2	<input type="checkbox"/>	DRLP2 <input type="checkbox"/>
410A2	↔	410B2 <input type="checkbox"/>
TID < Max TID		TID < Max TID

■ ■ ■ ■

415

425

Capa actual

EL3

425

EL2

Capa de referencia directa #1 de la capa actual

425

EL1

Capa de referencia directa #2 de la capa actual

425

BL

Para la secuencia de la capa actual:

Imagen de capa de referencia directa

¿Se incluye DRLP en el conjunto de imágenes de referencia entre capas?

¿Se restringe el uso de DRLP en la predicción entre capas?

Indicación:

410A1	410A2	410B1	410B2
No	Sí	Sí	Sí
Sí	Sí	No	No
1			

430

440

450

460

FIG. 4

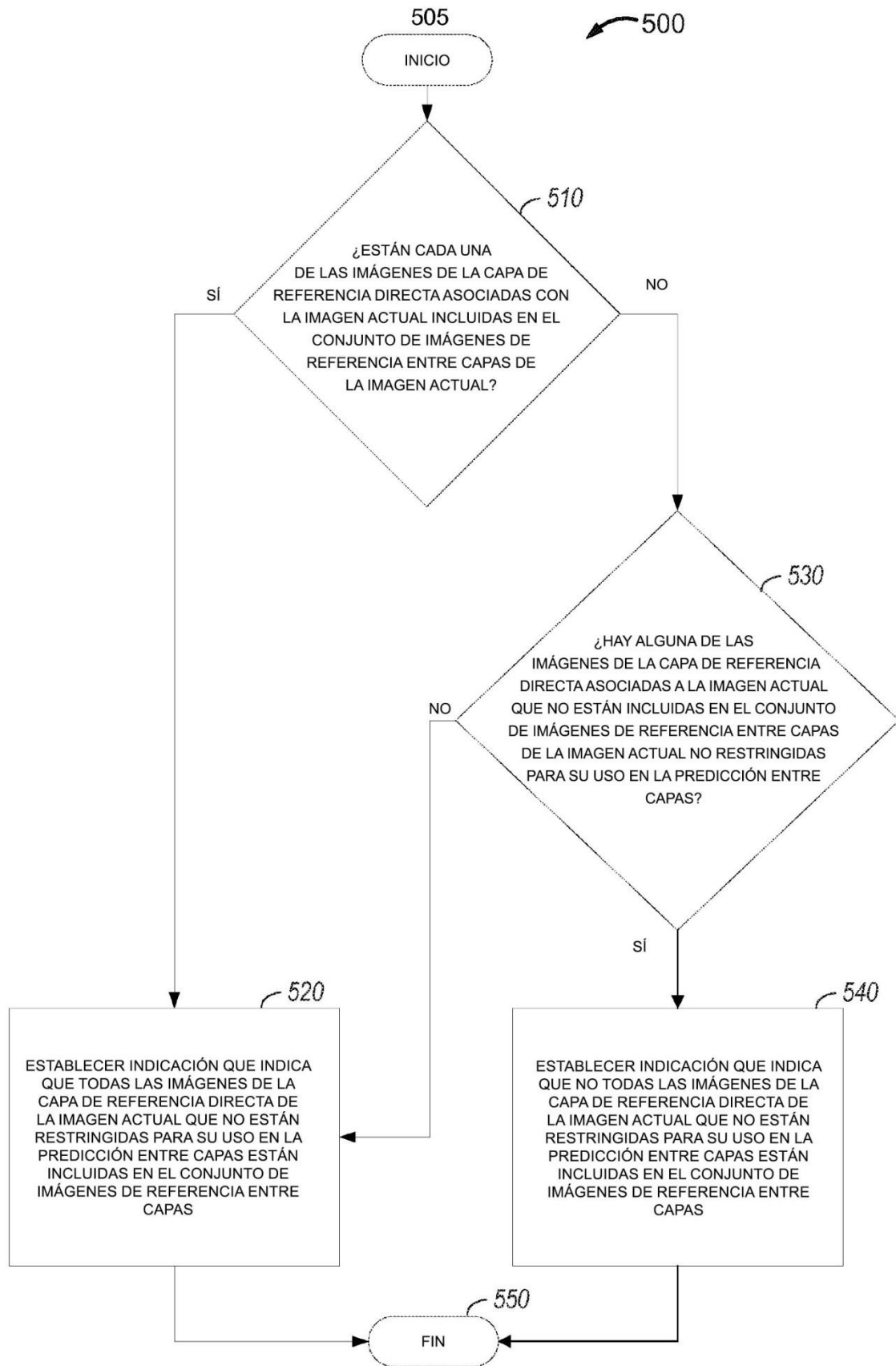


FIG. 5

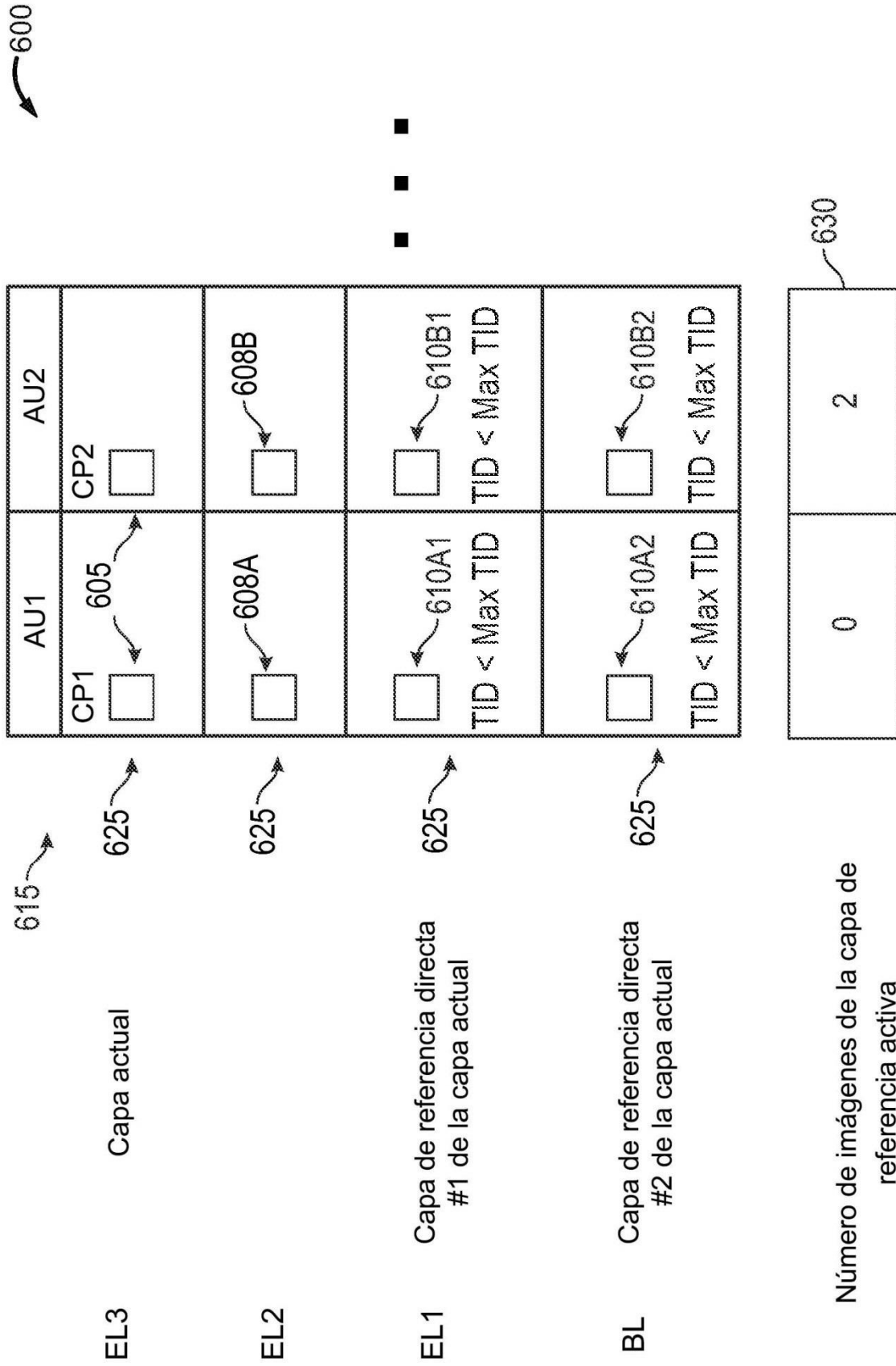


FIG. 6

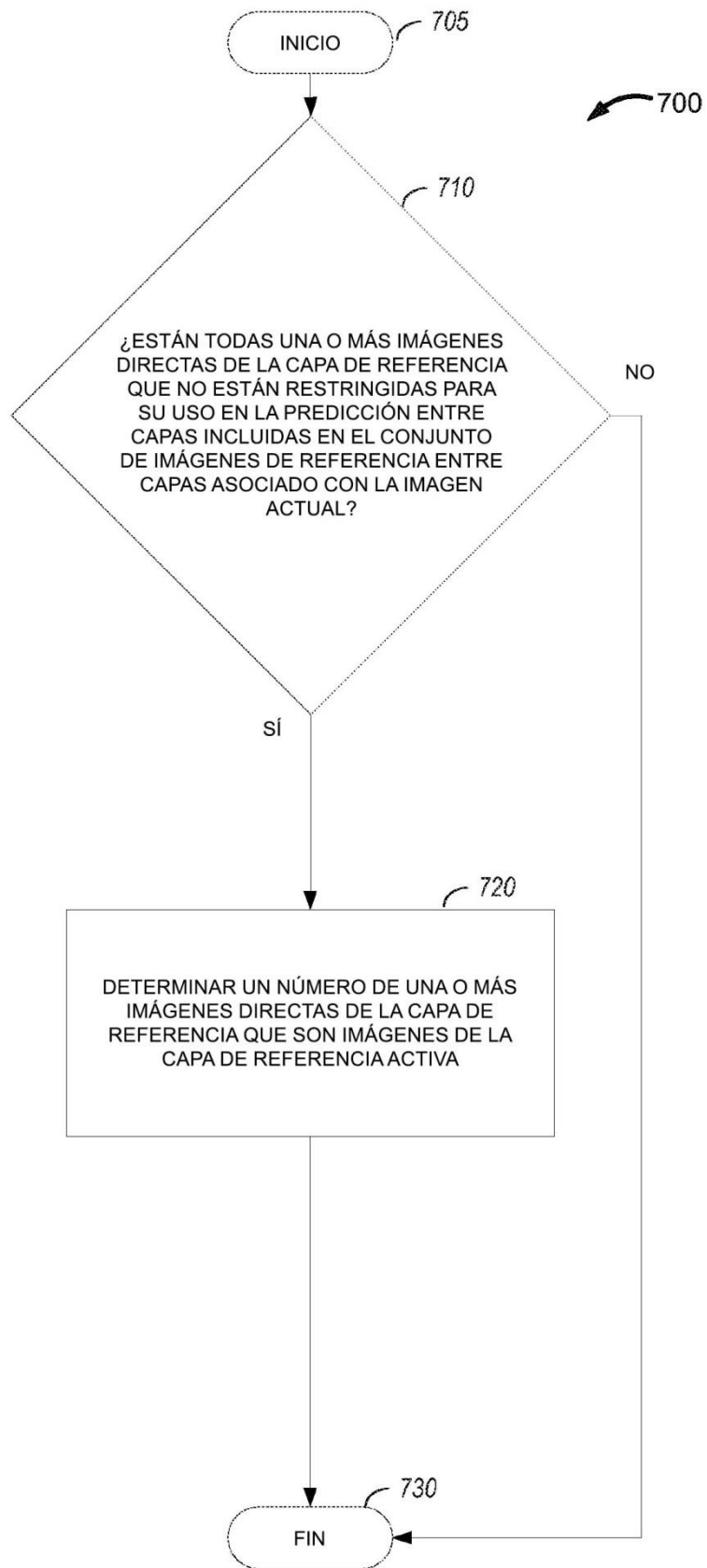


FIG. 7