

19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 749 747**

51 Int. Cl.:

H04N 19/30 (2014.01)

H04N 19/70 (2014.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

86 Fecha de presentación y número de la solicitud internacional: **14.10.2014 PCT/US2014/060453**

87 Fecha y número de publicación internacional: **23.04.2015 WO15057685**

96 Fecha de presentación y número de la solicitud europea: **14.10.2014 E 14795709 (6)**

97 Fecha y número de publicación de la concesión europea: **10.07.2019 EP 3058732**

54 Título: **Sistemas y procedimientos para definir por separado las dependencias para la predicción entre capas basada en subcapas**

30 Prioridad:

14.10.2013 US 201361890864 P
13.10.2014 US 201414513113

45 Fecha de publicación y mención en BOPI de la traducción de la patente:
23.03.2020

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
KARCZEWICZ, MARTA

74 Agente/Representante:

FORTEA LAGUNA, Juan José

ES 2 749 747 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 definir por separado las dependencias para la predicción entre capas basada en subcapas

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 de tableta, lectores de libros electrónicos, cámaras digitales, dispositivos de grabación digital, reproductores de medios digitales, dispositivos de videojuegos, consolas de videojuegos, teléfonos celulares o de radio por satélite, los denominados "teléfonos inteligentes", dispositivos de videoconferencia, dispositivos de transmisión de vídeo en tiempo real y similares. Los dispositivos de vídeo digital implementan técnicas de 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.

15

20

SUMARIO

25

[0003] Sería beneficioso permitir que un codificador de vídeo proporcione indicaciones distintas e independientes de la dependencia de predicción entre capas entre una primera capa y cada una de una serie de capas de mejora asociadas con la primera capa. Dichas indicaciones distintas e independientes pueden mejorar el rendimiento de la codificación (por ejemplo, proporcionar casos de flexibilidad de compensación de recursos) en varios escenarios de codificación, por ejemplo, en escenarios de aplicaciones en tiempo real.

30

[0004] Por lo tanto, las técnicas para definir e indicar por separado las dependencias de predicción entre capas para una primera capa con respecto a cada una de varias capas de mejora asociadas con la primera capa, usando indicaciones y/o determinaciones relacionadas con ciertas propiedades de las imágenes en la primera capa y/o las imágenes en la capa de mejora asociada, se describen en la presente memoria. Por ejemplo, la presente divulgación describe técnicas para definir, indicar y/o determinar si las imágenes en la primera capa se pueden usar para la predicción entre capas de una primera capa de mejora, mientras se define, indica y/o determina por separado si las imágenes en la primera capa se pueden utilizar para la predicción entre capas de una segunda capa de mejora. Las ventajas de las técnicas descritas en la presente divulgación, entre otras ventajas, pueden incluir reducir la asignación de memoria durante la predicción entre capas y permitir la flexibilidad de compensación de recursos en ciertos escenarios de codificación, tales como la implementación de la decodificación leve (por ejemplo, decodificación de un solo bucle) para una capa mientras se implementa la decodificación de alta calidad (por ejemplo, decodificación de bucle múltiple) para otra capa.

35

40

[0005] 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.

45

50

[0006] Un aspecto del tema en cuestión, descrito en la divulgación, proporciona un codificador de vídeo que comprende una unidad de memoria configurada para almacenar una primera imagen asociada con una primera capa e imágenes de capa de mejora asociadas con una pluralidad de capas de mejora. El codificador de vídeo comprende además un procesador en comunicación con la unidad de memoria. El procesador está configurado para proporcionar una indicación distinta para cada una de las capas de mejora, que indica si la primera imagen se puede usar para la predicción entre capas de la imagen de la capa de mejora en una respectiva capa de mejora.

55

[0007] Otro aspecto del tema en cuestión, descrito en la divulgación, proporciona un procedimiento de codificación de vídeo. El procedimiento comprende almacenar una primera imagen asociada con una primera capa e imágenes de capa de mejora asociadas con una pluralidad de capas de mejora. El procedimiento comprende además proporcionar una indicación distinta para cada una de las capas de mejora, que indica si la primera imagen se puede usar para la predicción entre capas de la imagen de la capa de mejora en una respectiva capa de mejora.

60

[0008] 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 primera imagen asociada con una primera capa e imágenes de mejora asociadas con una pluralidad de capas de mejora. El

65

5 decodificador de vídeo comprende además un procesador en comunicación con la unidad de memoria. El procesador está configurado para determinar por separado para cada una de las capas de mejora, basándose en una indicación respectiva, si la primera imagen se puede usar para la predicción entre capas de la imagen de la capa de mejora en una respectiva capa de mejora.

10 **[0009]** Otro aspecto del tema en cuestión, descrito en la divulgación, proporciona un procedimiento para decodificar vídeo. El procedimiento comprende almacenar una primera imagen asociada con una primera capa e imágenes de capa de mejora asociadas con una pluralidad de capas de mejora. El procedimiento comprende además determinar por separado, para cada una de las capas de mejora, basándose en una indicación respectiva, si la primera imagen se puede usar para la predicción entre capas de una imagen de capa de mejora en una respectiva capa de mejora.

BREVE DESCRIPCIÓN DE LOS DIBUJOS

[0010]

15 La FIG. 1A es un diagrama de bloques que ilustra un sistema ejemplar de codificación y decodificación de vídeo que puede utilizar las técnicas independientes de definición de dependencia de predicción entre capas, descritas en esta divulgación.

20 La FIG. 1B es un diagrama de bloques que ilustra otro sistema ejemplar de codificación y decodificación de vídeo que puede realizar las técnicas independientes de definición de dependencia de predicción entre capas, descritas en esta divulgación.

25 La FIG. 2 es un diagrama de bloques que ilustra un codificador ejemplar de vídeo que puede implementar las técnicas independientes de definición de dependencia de predicción entre capas, descritas en esta divulgación.

La FIG. 3 es un diagrama de bloques que ilustra un decodificador ejemplar de vídeo que puede implementar las técnicas independientes de definición de dependencia de predicción entre capas, descritas en esta divulgación.

30 La FIG. 4 es un diagrama que ilustra ejemplos de unidades de acceso múltiple, teniendo cada unidad de acceso múltiples capas, teniendo cada capa múltiples imágenes.

35 La FIG. 5 ilustra un diagrama de flujo de un procedimiento para que un procesador de un codificador de vídeo indique por separado las dependencias de predicción entre capas para una primera capa con respecto a una o más segundas capas.

40 La FIG. 6 ilustra un diagrama de flujo de un procedimiento para que un procesador de un decodificador de vídeo determine por separado las dependencias de predicción entre capas, entre una primera capa y una o más segundas capas.

La FIG. 7 ilustra otro modo de realización de un procedimiento de codificación de información de vídeo.

DESCRIPCIÓN DETALLADA

45 **[0011]** 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.

55 **[0012]** 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.

60 **[0013]** 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.int-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.int-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.int-sudparis.eu/jct/doc_end_user/documents/14_Vienna/wg11/JCTVC-N1008-v3.zip. El documento de LEE H ET AL: "MV-HEVC/SHVC HLS: Signalling for Inter-layer prediction indication [MV-HEVC/SHVC HLS: señalización para indicación de predicción entre capas]", 5. CONFERENCIA DEL JCT-3V; 27-7-2013 a 2-8-2013; VIENA; (EL EQUIPO DE COLABORACIÓN CONJUNTA EN DESARROLLO DE EXTENSIONES DE CODIFICACIÓN DE VÍDEO TRIDIMENSIONAL DE ISO/IEC JTC1/SC29/WG11 E ITU-T SG.16); URL: [HTTP://PHENIX.INT-EVRY.FR/JCT2/](http://phenix.int-evry.fr/jct2/), nº JCT3V-E0091-v3, 26 de julio de 2013 (2013-07-26), describe una sintaxis de alto nivel para señalar la predicción entre capas.

[0014] Como se ha descrito anteriormente, ciertos aspectos de la codificación de vídeo incluyen definir si las imágenes en una primera capa se pueden usar para la predicción entre capas de imágenes en todas las capas de mejora asociadas con la primera capa. Por ejemplo, una secuencia de vídeo puede incluir unidades de acceso, y cada una de las unidades de acceso puede contener una o más imágenes. Cada unidad de acceso puede además estar asociada con una o más capas de información de vídeo (por ejemplo, una capa base y una o más capas de mejora). Cada una de las capas puede asociarse con un conjunto de identificaciones (ID) de capa o ser igual a un valor (por ejemplo, un valor entero). Se puede decir que una capa dada "precede" o "está debajo de" una capa diferente cuando la capa dada está asociada con una identificación de capa (ID) que es inferior a la identificación (ID) de capa asociada con la capa diferente. A la inversa, se puede decir que una capa dada es "posterior a" o que "está sobre" una capa diferente cuando la capa dada está asociada con una identificación (ID) de capa que es superior a la identificación (ID) de capa asociada con la capa diferente. Por ejemplo, una capa dada (por ejemplo, una de las una o más capas de mejora) puede ser una capa de mejora para una capa que precede a la capa dada, tal como la capa base o cualquier capa de mejora anterior. Por ejemplo, un codificador de vídeo puede analizar imágenes en la primera capa de manera tal que pueda indicar a un decodificador de vídeo si las imágenes de la primera capa se pueden usar para la predicción entre capas de las capas de mejora asociadas con la primera capa. En algunos casos, si el codificador de vídeo indica que las imágenes de la primera capa se pueden usar para la predicción entre capas de las capas de mejora asociadas con la primera capa, el codificador de vídeo también puede indicar en qué medida se pueden usar las imágenes en la primera capa para la predicción entre capas de las capas de mejora asociadas con la primera capa. Por ejemplo, el codificador de vídeo puede indicar que solo aquellas imágenes que están asociadas con un parámetro (por ejemplo, un valor de identificación temporal (TID) o "valor de TID") cuyo valor es menor que un umbral (por ejemplo, un umbral de valor de identificación temporal (por ejemplo, un "valor de umbral de TID", "TID máximo" o "TID máx") se puede usar para la predicción entre capas de las capas de mejora asociadas con la primera capa. Por ejemplo, a nivel del conjunto de parámetros de vídeo (VPS), el codificador de vídeo puede proporcionar una indicación (por ejemplo, un elemento sintáctico y/o un indicador fijado en 0 o 1) que indica a un decodificador de vídeo si las imágenes de la primera capa se pueden utilizar para la predicción entre capas de las capas de mejora asociadas con la primera capa. El codificador de vídeo luego puede proporcionar una indicación adicional (por ejemplo, un elemento sintáctico, un indicador y/o una variable que se puede fijar en un valor) que indica al decodificador de vídeo la medida en que las imágenes en la primera capa pueden ser utilizadas para la predicción entre capas de las capas de mejora asociadas con la primera capa, como se ha descrito anteriormente. Al proporcionar dichas indicaciones (por ejemplo, al fijar ciertos indicadores y variables), el codificador de vídeo puede permitir al decodificador de vídeo determinar si, y en qué medida, se usan imágenes en la primera capa para predecir imágenes entre capas en todas las capas de mejora asociadas con la primera capa.

[0015] Por ejemplo, los recientes Borradores de trabajo de MV-HEVC WD5 y SHVC WD3 incluyen el indicador *max_tid_ref_present_flag* con la siguiente semántica: El *max_tid_ref_present_flag* igual a 1 especifica que el elemento sintáctico *max_tid_il_ref_pics_plus1[i]* está presente. El *max_tid_ref_present_flag* igual a 0 especifica que el elemento sintáctico *max_tid_il_ref_pics_plus1[i]* no está presente. Además, los Borradores de trabajo recientes de MV-HEVC WD5 y SHVC WD3 incluyen el indicador *max_tid_il_ref_pics_plus1* en el conjunto de parámetros de vídeo (VPS), con la siguiente semántica: "*max_tid_il_ref_pics_plus1[i]* igual a 0 especifica que, dentro del CVS, las imágenes no IRAP con *nuh_layer_id* igual a *layer_id_in_nuh[i]* no se usan como referencia para la predicción entre capas. *max_tid_il_ref_pics_plus1[i]* mayor que 0 especifica que, dentro del CVS, las imágenes con *nuh_layer_id* igual a *layer_id_in_nuh[i]* y *TemporalId* mayor que *max_ref_pics_plus1[i] - 1* no se usan como referencia para la predicción entre capas. Cuando no está presente, se deduce que *max_tid_il_ref_pics_plus1[i]* es igual a 7". Por ejemplo, el codificador de vídeo puede señalar el *max_tid_ref_present_flag* para indicar que *max_tid_il_ref_pics_plus1[i]* está presente. Si es así, entonces el codificador de vídeo puede señalar *max_tid_il_ref_pics_plus1[i]* (por ejemplo, a un flujo de bits que puede ser almacenado y/o procesado por el decodificador de vídeo) a nivel del VPS para indicar si las imágenes pertenecientes a una subcapa temporal (por ejemplo, [i]) son utilizadas para la predicción entre capas

por imágenes de capas superiores (por ejemplo, en capas de mejora). Por ejemplo, cuando el codificador de vídeo fija *max_tid_il_ref_pics_plus1[i]* igual a un valor de "0" y lo señala al decodificador de vídeo, el decodificador de vídeo puede determinar que las imágenes en la capa [i] no se usarán para la predicción entre capas, para cualquier capa de mejora. Como otro ejemplo, cuando el codificador de vídeo fija *max_tid_il_ref_pics_plus1[i]* igual a un valor mayor que "0" (por ejemplo, X), el decodificador de vídeo puede determinar que las imágenes en la capa [i], que estén asociadas con un valor de TID mayor que X - 1, no se utilizarán para la predicción entre capas, para cualquier capa de mejora. El decodificador de vídeo puede usar esta información para descartar (por ejemplo, eliminar de un almacén temporal) aquellas imágenes que no se usan para la predicción entre capas.

[0016] Al utilizar *max_tid_ref_present_flag* y *max_tid_il_ref_pics_plus1 [i]* de la manera descrita anteriormente, se pueden lograr algunas compensaciones entre la eficacia de la codificación, el ancho de banda de la memoria y la complejidad de cálculo en las extensiones de capas múltiples de la HEVC, en varios escenarios. Por ejemplo, cuando se determina que las imágenes en la primera capa no se utilizarán para ninguna otra capa superior, las imágenes se pueden descartar para ahorrar memoria.

[0017] Sin embargo, los diseños actuales de extensiones de la HEVC tienen ciertas deficiencias. Por ejemplo, la señalización actual basada en *max_tid_il_ref_pics_plus1*, como se ha descrito anteriormente, requiere que la predicción entre capas, para imágenes que pertenecen a una subcapa temporal, esté habilitada o inhabilitada para imágenes de todas las capas superiores. Por ejemplo, si *max_tid_il_ref_pics_plus1 [i]* se fija en un valor de "2" para una primera capa (por ejemplo, [i] = "0" o "Capa 0"), entonces las imágenes en la Capa 0, asociadas con un valor de TID mayor o igual a "2", no se pueden usar como referencia entre capas para ninguna capa superior. Como otro ejemplo, si *max_tid_il_ref_pics_plus1[i]* se fija en un valor de "7" para la primera capa (por ejemplo, [i] = "0" o "Capa 0"), entonces las imágenes en la Capa 0, asociadas con un valor de TID mayor o igual que "7", no se pueden usar como referencia entre capas para ninguna capa superior. En otras palabras, dado que, según la HEVC, el mayor valor posible del valor de TID es "6", todas las imágenes en la Capa 0 se pueden usar como referencia entre capas para las capas superiores en este ejemplo.

[0018] Sin embargo, bajo las especificaciones actuales de la HEVC, no es posible indicar, por ejemplo, que las imágenes en una primera capa (por ejemplo, la Capa 0), que están asociadas con un cierto valor de TID (por ejemplo, un valor de "2"), no se pueden usar como referencia entre capas (por ejemplo, para la predicción entre capas) para una capa (por ejemplo, la Capa 1), pero se pueden usar como referencia entre capas para una capa diferente (por ejemplo, la Capa 2). Estas limitaciones crean requisitos innecesarios de complejidad de decodificación (por ejemplo, al requerir decodificación de múltiples bucles) para imágenes que, de otra manera, podría ser beneficioso decodificar usando decodificación leve (por ejemplo, decodificación de un solo bucle). Por ejemplo, a medida que aumenta el número de capas debajo de cualquier capa dada, la complejidad de la decodificación de la capa dada puede aumentar (por ejemplo, para decodificar la Capa 2, puede ser necesario decodificar la Capa 1 y la Capa 0 en algunos casos). En ciertos casos, cuando la simplicidad de la codificación es de mayor prioridad que la calidad de la imagen, puede ser más beneficioso decodificar, por ejemplo, la Capa 2 sin decodificar una o más entre la Capa 1 y la Capa 0. Por lo tanto, sería beneficioso permitir que un codificador de vídeo proporcione definiciones distintas e independientes de la dependencia de predicción entre capas para una primera capa con respecto a cada una de una serie de capas de mejora asociadas con la primera capa. Dichas configuraciones pueden ser beneficiosas (por ejemplo, proporcionar casos de flexibilidad de compensación de recursos) en varios escenarios de codificación, por ejemplo, en escenarios de aplicaciones en tiempo real.

[0019] Por lo tanto, la presente divulgación describe técnicas para la definición por separado (a veces denominada como definición independiente) de las dependencias de predicción entre capas para una primera capa con respecto a cada una de una serie de capas de mejora asociadas con la primera capa, utilizando indicaciones y/o determinaciones relacionadas con ciertas propiedades de las imágenes en la primera capa y/o las imágenes en las capas de mejora asociadas. Por ejemplo, la presente divulgación describe técnicas para definir, indicar y/o determinar si las imágenes en la primera capa se pueden usar para la predicción entre capas de una primera capa de mejora, mientras se define, indica y/o determina por separado si las imágenes en la primera capa se pueden utilizar para la predicción entre capas de una segunda capa de mejora. Además, un codificador puede adaptarse para indicar que unas imágenes (por ejemplo, imágenes que están asociadas con un valor de TID mayor que un valor de umbral de TID) no se deben usar para la predicción entre capas para una capa, pero se pueden usar para la predicción entre capas para otra capa. Siguiendo los ejemplos descritos anteriormente, la presente divulgación describe, por ejemplo, técnicas para restringir el uso de una o más de las imágenes en la Capa 0 para la predicción entre capas de imágenes en la Capa 1, sin restringir a la vez ninguna de las imágenes en la Capa 0 para ser utilizada para la predicción entre capas de imágenes en la Capa 2. Dichas realizaciones proporcionan la flexibilidad para tener una decodificación leve (por ejemplo, una decodificación de bucle único) para la Capa 1 y una decodificación de alta calidad (por ejemplo, una decodificación de múltiples bucles) para la Capa 2. En otros ejemplos, las imágenes en la Capa 0 pueden ser restringidas o no restringidas para ser usadas en la predicción entre capas para la Capa 1, la Capa 2 y/o cualquier otra capa, de diferentes maneras, como se describe en relación con las FIG. 4 a 6, como se describe a continuación.

[0020] Por lo tanto, las técnicas descritas en el presente documento pueden incluir varios procedimientos para mejorar los procedimientos de señalización de dependencia de predicción entre capas, tales como aquellos procedimientos relacionados con la señalización por separado de la dependencia de predicción entre capas de las

subcapas, para todas las capas dependientes directas. 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 descritos puede aplicarse independientemente de los demás, o en combinación.

[0021] Las ventajas de las técnicas descritas en la presente divulgación, entre otras ventajas, pueden incluir la reducción de la asignación de memoria durante la predicción entre capas. En comparación con los procedimientos actuales en MV-HEVC WD5 y SHVC WD3, las técnicas descritas en la presente divulgación también pueden permitir la flexibilidad de la compensación de recursos en ciertos escenarios de codificación, tales como la implementación de la decodificación leve (por ejemplo, la decodificación de un solo bucle) para una capa mientras se implementa la decodificación de alta calidad (por ejemplo, decodificación de bucle múltiple) para otra capa, como se ha descrito anteriormente.

[0022] La FIG. 1A es un diagrama de bloques que ilustra un sistema ejemplar de codificación y decodificación de vídeo 10 que puede utilizar las distintas técnicas de definición de dependencia de predicción entre capas, descritas en esta 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.

[0023] 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.

[0024] 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.

[0025] 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. Los servidores ejemplares de ficheros incluyen un servidor de la Red (por ejemplo, para una sede de la Red), un servidor del FTP, dispositivos de almacenamiento conectados en red (NAS) o una unidad de disco local. El 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.

[0026] 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.

[0027] 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.

[0028] 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.

[0029] 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.

[0030] 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 o 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.

[0031] Como se describe con mayor detalle a continuación, los aspectos de esta divulgación se relacionan en general con la definición por separado de las dependencias de predicción entre capas para una primera capa con respecto a cada una de una serie de capas de mejora asociadas con la primera capa. Por ejemplo, definir, indicar y/o determinar si las imágenes en la primera capa se pueden usar para la predicción entre capas de una primera capa de mejora mientras se define, indica y/o determina por separado si las imágenes en la primera capa se pueden usar para predicción entre capas de una segunda capa de mejora. 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.

[0032] 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. Sin embargo, las técnicas de esta divulgación no están limitadas a ninguna norma o técnica de codificación particular.

[0033] 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 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).

[0034] 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 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.

[0035] El equipo JCT-VC está trabajando en el desarrollo de la norma HEVC. La iniciativa de normalización de la HEVC se basa en un modelo en evolución de un dispositivo de codificación de vídeo, denominado modelo de prueba de la HEVC (HM). El HM supone varias capacidades adicionales de los dispositivos de codificación de vídeo respecto a los dispositivos existentes, de acuerdo, por ejemplo, con la norma ITU-T H.264/AVC. Por ejemplo, mientras que la norma H.264 proporciona nueve modalidades de codificación mediante intrapredicción, el HM puede proporcionar hasta treinta y tres modalidades de codificación mediante intrapredicción.

[0036] 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, 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.

[0037] Una CU incluye un nodo de codificación y unidades de predicción (PU) y unidades de transformación (TU) asociadas con el 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 con una CU pueden describir, por ejemplo, la división de la CU en una o más PU. Las modalidades de división pueden diferir entre sí la CU está codificada en modalidad de omisión o directa, codificada en modalidad de intrapredicción o codificada en modalidad de interpredicción. Las PU pueden dividirse para tener forma no cuadrada. Los datos sintácticos asociados con 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.

[0038] La FIG. 2 es un diagrama de bloques que ilustra un codificador ejemplar de vídeo 20 que puede implementar las distintas técnicas de definición de predicción entre capas, descritas en esta divulgación. El codificador de vídeo 20 puede realizar la intracodificación y la intercodificación de bloques de vídeo dentro de fragmentos de vídeo. La intracodificación se basa en la predicción espacial para reducir o eliminar la redundancia espacial en el vídeo dentro de una 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 (modalidad P) o la bipredicción (modalidad B), pueden referirse a cualquiera de varias modalidades de compresión de base temporal.

[0039] 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.

[0040] 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.

[0041] 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.

[0042] 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.

[0043] 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 con respecto a las posiciones de píxel completo y las posiciones de píxel fraccionario, y emitir un vector de movimiento con una precisión de píxel fraccionario.

[0044] 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.

[0045] 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.

[0046] 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.

[0047] 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 la modalidad de intrapredicción y una pluralidad de tablas modificadas de índices de la modalidad de intrapredicción (también denominadas tablas de correlación de palabras de código), definiciones de contextos de codificación para diversos bloques e indicaciones de una modalidad de intrapredicción más probable, una tabla de índices de la modalidad de intrapredicción y una tabla modificada de índices de la modalidad de intrapredicción, a usar para cada uno de los contextos.

[0048] 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.

[0049] La unidad de procesamiento de transformación 52 puede enviar los coeficientes de transformación resultantes a la unidad de cuantización 54. La unidad de cuantización 54 cuantiza los coeficientes de transformación para reducir más la velocidad de bits. El proceso de cuantización puede reducir la profundidad de bits asociada con algunos de, o todos, los coeficientes. El grado de cuantización se puede modificar ajustando un parámetro de cuantización. En algunos ejemplos, la unidad de cuantización 54 puede realizar, a continuación, una exploración de la matriz que incluye los coeficientes de transformación cuantizados. De forma alternativa, la unidad de codificación por entropía 56 puede realizar la exploración.

[0050] Tras la cuantización, la unidad de codificación por entropía 56 realiza la codificación por entropía de los coeficientes de transformación cuantizados. 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.

[0051] 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.

[0052] 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 (bien 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.

[0053] 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.

[0054] 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.

[0055] 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.

[0056] La FIG. 3 es un diagrama de bloques que ilustra un decodificador de vídeo 30 ejemplar que puede implementar las distintas técnicas de definición de dependencia de predicción entre capas, descritas en esta 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.

[0057] Durante el proceso de decodificación, el decodificador de vídeo 30 recibe un flujo de bits de vídeo codificado, que representa bloques de vídeo de un fragmento de vídeo codificado y elementos sintácticos asociados, desde el codificador de vídeo 20. 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 decodificador de vídeo 30 puede recibir los elementos sintácticos en el nivel de fragmento de vídeo y/o el nivel de bloque de vídeo.

[0058] 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.

[0059] 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 decodificando. Por ejemplo, la unidad de compensación de movimiento 82 usa algunos de los elementos sintácticos recibidos para determinar una modalidad de predicción (por ejemplo, intrapredicción o interpredicción) usada para codificar los bloques de vídeo del fragmento de vídeo, un tipo de fragmento de interpredicción (por ejemplo, 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 interpredicción para cada bloque de vídeo intercodificado del fragmento y otra información para decodificar los bloques de vídeo en el fragmento de vídeo actual.

[0060] 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.

[0061] 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.

[0062] 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.

[0063] 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.

[0064] De acuerdo con uno o más aspectos de la presente divulgación, se proporcionan técnicas para proporcionar (mencionadas a veces como para definir, establecer, procesar y / o determinar) por separado (por ejemplo, de manera independiente) dependencias de predicción entre capas para una primera capa con respecto a cada una entre una serie de capas de mejora asociadas con la primera capa. En algunas realizaciones, un procesador está configurado para determinar (por ejemplo, con un primer indicador) si las imágenes en la primera capa se pueden usar para la predicción entre capas de una primera capa de mejora mientras se determina por separado e independientemente (por ejemplo, con un segundo indicador) si las imágenes en la primera capa se pueden usar para la predicción entre capas de una segunda capa de mejora. 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.

[0065] La FIG. 4 es un diagrama 400 que ilustra ejemplos de unidades de acceso múltiple, teniendo cada unidad de acceso múltiples capas, teniendo cada capa múltiples imágenes. Algunas de las unidades de acceso ilustran escenarios donde se han empleado dependencias de predicción entre capas, indicadas por separado (por ejemplo, en un codificador de vídeo) para una primera capa con respecto a capas de mejora independientes. Por ejemplo, en algunas de las unidades de acceso ilustradas, un primer indicador indica que se puede usar una imagen de una primera capa (por ejemplo, Identificador de capa = 0) para predecir una imagen a partir de una primera capa de mejora,

y un segundo indicador indica que la imagen de la primera capa no se puede usar para predecir una imagen de una segunda capa de mejora. Los indicadores primero y segundo se señalizan de forma independiente (por ejemplo, señalizan dependencias de predicción independientes entre capas) como indicadores diferentes. Por ejemplo, en el diagrama ilustrado, se proporcionan indicaciones que pueden usarse para señalar o determinar si las imágenes en la primera capa se pueden usar para la predicción entre capas de la primera capa de mejora mientras se indica (o se determina) por separado si las imágenes en la primera capa se pueden utilizar para la predicción entre capas de la segunda capa de mejora. La indicación en cuanto a si cualquier capa de mejora dada depende de la primera capa para la predicción entre capas puede ser igual o diferente (por ejemplo, independiente) de la indicación en cuanto a si alguna otra capa de mejora dada depende de la primera capa para la predicción entre capas. En un aspecto, una o más de las características descritas en relación con la FIG. 4 pueden realizarse generalmente mediante un codificador (por ejemplo, el codificador de vídeo 20 de las FIG. 1B y/o 2) y/o un decodificador (por ejemplo, el decodificador de vídeo 30 de las FIG. 1B y/o 3). En un modo de realización, el codificador de vídeo 20 puede proporcionar las indicaciones de dependencia al decodificador de vídeo 30 de manera tal que el decodificador de vídeo 30 pueda realizar varias determinaciones y/u operaciones adicionales, como se describe a continuación. En otros modos de realización, las indicaciones de dependencia pueden ser procesadas, determinadas, definidas, indicadas, analizadas y/o utilizadas por uno o más dispositivos diferentes. Los diversos aspectos descritos en relación con la FIG. 4 puede incluirse en una unidad de memoria (por ejemplo, el dispositivo de almacenamiento 32 de la FIG. 1A, la memoria 64 de la FIG. 2, la memoria 92 de la FIG. 3, etc.). Las diversas operaciones descritas en relación con la FIG. 4 puede ser realizadas por un procesador (por ejemplo, el procesador 13 de la FIG. 1A, la unidad de procesamiento de vídeo 21 de la FIG. 1B, la unidad de procesamiento de predicción 41 de la FIG. 2, etc.). El codificador de vídeo 20, el decodificador de vídeo 30, el dispositivo de almacenamiento 32 y/o el procesador 13 pueden estar en comunicación y/o conectados operativamente entre sí.

[0066] En la ilustración ejemplar, se presentan dos escenarios: (1) "Escenario A" y (2) "Escenario B". Cada escenario representa unidades de acceso, capas e imágenes donde un conjunto particular de indicaciones de dependencia se señala de manera independiente. Cada una de las columnas representa una unidad de acceso en una secuencia de vídeo. Por ejemplo, las unidades de acceso 420A1 a 420A5 representan cinco unidades de acceso en una secuencia de vídeo del Escenario A. Las unidades de acceso pueden mencionarse colectivamente como las unidades de acceso 420. En un modo de realización, la secuencia de vídeo puede incluir cualquier número de unidades de acceso 420. Cada una de las unidades de acceso 420A1 a 420A5 puede contener una o más imágenes y puede estar asociada con una o más capas. Por ejemplo, en el Escenario A, cada unidad de acceso 420A1 a 420A5 incluye tres capas 410A, 415A1 y 415A2. Cada capa contiene una pluralidad de imágenes, y hay una imagen en cada unidad de acceso. La capa ilustrada más baja (410A) puede denominarse una capa base, una capa actual y/o una primera capa. Puede haber capas adicionales por encima o por debajo de la primera capa 410A (no se muestra). Las capas sobre la primera capa 410A (415A1 y 415A2) pueden considerarse capas de mejora de la primera capa 410A. En algunos aspectos, la primera capa 410A puede tener capas de mejora adicionales (no representadas). Una capa de mejora a veces se puede denominar "capa mejorada". En el ejemplo ilustrado, la primera capa 410A también se puede denominar "Capa 0", porque tiene un número de identificación de capa igual a "0". En algunas realizaciones, el procesador 13 puede usar el número de identificación de capa para diversas determinaciones, como se describe adicionalmente más adelante. En algunas realizaciones, el número de identificación de capa puede ser representado por una variable, por ejemplo, una variable basada en la HEVC, *LayerId*. De manera similar, las capas de mejora 415A1, 415A2 pueden tener, cada una, números de identificación de capa. En el ejemplo ilustrado, las capas de mejora 415A1 y 415A2 tienen un número de identificación de capa igual a 1 y 2, respectivamente y, por lo tanto, también pueden denominarse "Capa 1" y "Capa 2", respectivamente.

[0067] Como se ha indicado anteriormente, cada una de las capas (410A, 415A1, 415A2) puede incluir una o más imágenes. Por lo tanto, cada imagen está asociada con una capa y una unidad de acceso. En la ilustración ejemplar, cada una de las unidades de acceso 420 contiene una imagen en cada capa; sin embargo, en otros ejemplos, cada unidad de acceso 420 puede contener más de una imagen en cada capa. En un modo de realización, una imagen asociada con la primera capa 410A puede denominarse imagen actual o "primera imagen". Las primeras imágenes en la FIG. 4 incluyen las primeras imágenes 430A1, 430A2, 430A3, 430A4 y 430A5. Las primeras imágenes 430A1, 430A2, 430A3, 430A4 y 430A5 pueden mencionarse colectivamente como primeras imágenes 430. La primera imagen 430A1 está dentro de la primera capa 410A y la unidad de acceso 420A1. Como otro ejemplo, la primera imagen 430A3 es la imagen dentro de la primera capa 410A y la unidad de acceso 420A3. Las imágenes en capas que no son la primera capa pueden denominarse imágenes de mejora. Por ejemplo, las imágenes de mejora 431A1 a 431A5 (denominadas colectivamente imagen de mejora 431) están ubicadas en la capa de mejora 415A1, y las imágenes de mejora 432A1 a 432A5 (denominadas en conjunto imágenes de mejora 432) están en la capa de mejora 415A2. Las imágenes de mejora también pueden mencionarse como "imágenes de capa de mejora". Por ejemplo, la imagen de mejora 431A2 está dentro de la capa de mejora 415A1 y la unidad de acceso 420A2. Como otro ejemplo, la imagen de mejora 432A4 está dentro de la capa de mejora 415A2 y la unidad de acceso 420A4. En un modo de realización, cada una de las imágenes de mejora 431 y 432 ubicadas dentro de la misma unidad de acceso que una primera imagen pueden ser imágenes de mejora de dicha primera imagen. Por ejemplo, las imágenes de mejora 431A3 y 432A3 son imágenes de mejora de la primera imagen 430A3.

[0068] Cada una de las imágenes 430, 431 y 432 puede asociarse con un número o valor de identificación temporal (un "valor de TID"), como se ilustra en el lado izquierdo de cada Escenario. Por ejemplo, la imagen de mejora 431A3

está asociada con un valor de TID igual a "1". Como otro ejemplo, la imagen de mejora 430A2 está asociada con un valor de TID igual a "2". Como se ilustra, cada una de las imágenes 430, 431 y 432 que están dentro de la misma de las unidades de acceso 420 están asociadas con el mismo valor de TID. En otros modos de realización, todas las imágenes dentro de la misma de las unidades de acceso 420 no necesitan asociarse con el mismo valor de TID. En algunas realizaciones, el procesador 13 puede comparar el valor de TID, asociado con una de las imágenes 430, 431 o 432, con un umbral asociado con un par de capas (por ejemplo, un "TID máximo", un "TID máx", un "Umbral de TID" o un "valor de umbral de TID"), como se describe adicionalmente más adelante en relación con las FIG. 5 y 6.

[0069] El procesador 13 puede realizar operaciones basándose en las propiedades de una de las imágenes 430 con respecto a sus imágenes de mejora asociadas (por ejemplo, las imágenes de mejora 431 y/o 432 en la misma unidad de acceso que la de las imágenes 430), como adicionalmente se describe a continuación en relación con las FIG. 5 y 6. Por ejemplo, el procesador 13 puede indicar o determinar por separado si las imágenes en la primera capa 410A pueden usarse para predecir imágenes entre capas en una o más de las capas de mejora 415. En otras palabras, el procesador 13 puede indicar o determinar por separado si cada una de las capas de mejora 415 puede usar la primera capa 410 para la predicción entre capas de cada capa de mejora 415. Si la indicación indica que las imágenes en la primera capa 410A se pueden usar para la predicción entre capas de las imágenes en una de las capas de mejora 415, entonces la respectiva capa de mejora 415 depende de la primera capa 410A. Por ejemplo, como se ha descrito anteriormente, el procesador 13 puede indicar por separado (por ejemplo, con un primer indicador) si la capa de mejora 415A1 depende de la primera capa 410A para la predicción entre capas, mientras que indica por separado (por ejemplo, en un segundo indicador) si la capa de mejora 415A2 depende de la primera capa 410A para la predicción entre capas. Las dos indicaciones de dependencia separadas y distintas pueden tener valores iguales o diferentes. Si se proporcionan capas de mejora adicionales (no representadas en la imagen), el procesador 13 puede indicar por separado, para cada una de las capas de mejora 415, si las capas de mejora adicionales también dependen de la primera capa 410.

[0070] En los Escenarios A y B ilustrados, el procesador 13 ha indicado que las primeras capas 410A (en el Escenario A) y 410B (Escenario B), incluyen imágenes que pueden usarse para la predicción entre capas. Por ejemplo, cada una de las capas de mejora 415A1 y 415A2 depende de la primera capa 410A, y cada una de las capas de mejora 415B1 y 415B2 depende de la primera capa 410B. Las flechas (sin una X) que se extienden entre las imágenes indican que las imágenes de la primera capa 410A, 410B se pueden usar para predecir imágenes en una o más de las capas de mejora 415A1, 415A2, 415B1, 415B2. En un modo de realización, incluso cuando una de las capas de mejora 415 depende de una de las primeras capas 410, algunas de las primeras imágenes 430 en la primera capa 410 pueden tener un uso restringido para la predicción entre capas de las respectivas imágenes de mejora 431, 432. Dicha restricción se indica mediante una X situada sobre la flecha que apunta desde la imagen restringida de las primeras imágenes 410 a la correspondiente de las imágenes de mejora 431 o 432. La restricción de uso puede basarse en un valor de umbral de TID asociado con la capa 410 actual y las capas de mejora 415 asociadas, como se describe adicionalmente más adelante y en relación con las FIG. 5 y 6. En algunas realizaciones, incluso aunque una primera imagen (por ejemplo, la primera imagen 430A5) no esté restringida para ser utilizada para predecir entre capas sus respectivas imágenes de mejora (por ejemplo, las imágenes de mejora 431A5, 432A5), no obstante, no se puede usar para la predicción entre capas. En dicha realización, ninguna flecha se extiende entre la primera imagen (por ejemplo, la primera imagen 430A5) y sus respectivas imágenes de mejora (por ejemplo, las imágenes de mejora 431A5, 432A5).

[0071] En un modo de realización (por ejemplo, en el Escenario A ilustrado), un procesador (por ejemplo, el procesador 13) puede determinar (por ejemplo, indicar, recibir, enviar, etc.) un primer valor de umbral de TID que se puede usar para determinar si las imágenes en la primera capa 410A pueden usarse para predecir entre capas imágenes en la primera capa de mejora 415A1 de la primera capa. El procesador puede determinar además un segundo valor de umbral de TID que se usará para determinar si las imágenes en la primera capa 410A pueden usarse para predecir entre capas imágenes en la segunda capa de mejora 415A2 de la primera capa. Los valores primero y segundo de umbral de TID pueden ser iguales o diferentes entre sí.

[0072] En los modos de realización de la FIG. 4, se especifica un primer umbral de TID para un primer par de capas, tales como la primera capa 410A y la primera capa de mejora 415A1. El primer umbral de TID se usa para determinar si las imágenes en la primera capa 410A pueden usarse para la predicción entre capas de imágenes en la primera capa de mejora 415A1. Se proporciona un segundo umbral de TID para un segundo par de capas, tales como la primera capa 410A y la segunda capa de mejora 415A2. El segundo umbral de TID se usa para determinar si las imágenes en la primera capa 410A pueden usarse para la predicción entre capas de imágenes en la segunda capa de mejora 415A2. Se pueden proporcionar umbrales de TID adicionales para pares de capas adicionales. En el modo de realización ilustrada del Escenario A, el primer valor de umbral de TID es un valor de "2". Por lo tanto, como se describe adicionalmente en relación con las FIG. 5 y 6, las imágenes en la primera capa 410A que están asociadas con un valor de TID por debajo de "2" se pueden usar para la predicción entre capas de imágenes en la primera capa de mejora 415A1 (por ejemplo, no están restringidas para su uso en la predicción entre capas). Por consiguiente, las primeras imágenes 430A2 y 430A4, que están asociadas con un valor de TID de "2" (que no está por debajo de "2") tienen un uso restringido para la predicción entre capas de las imágenes de mejora 431A2 y 431A4, respectivamente (mostrada cada restricción como una flecha con una X). A la inversa, las primeras imágenes 430A1 (que tiene un valor de TID de "0"), 430A3 (que tiene un valor de TID de "1") y 430A5 (que tiene un valor de TID de "0"), cada una de las

cuales está asociada con los valores de TID por debajo de "2" no tienen un uso restringido para la predicción entre capas de sus respectivas primeras imágenes de capa de mejora 431A1, 431A3 y 431A5. Como se ha indicado anteriormente, algunas de las imágenes permitidas se pueden usar para la predicción entre capas (como lo indica una flecha), mientras que otras no se pueden usar para la predicción entre capas (como lo indica la ausencia de una flecha). Dicha configuración proporciona flexibilidad para implementar la decodificación ligera (por ejemplo, decodificación de bucle único) para las primeras imágenes de capa de mejora 431A2 y 431A4 y la decodificación de alta calidad (por ejemplo, decodificación de bucle múltiple) para una o más de las primeras imágenes de capa de mejora 431A1, 431A3 y 431A5. Además, en algunas realizaciones, los valores de umbral de TID pueden usarse de diferentes maneras. Por ejemplo, en algunas realizaciones, un indicador puede indicar que una o más de las imágenes en la primera capa 410A, que están asociadas con un valor de TID por debajo de uno menos que un valor de umbral de TID, se pueden usar para la predicción entre capas de imágenes en capas de mejora asociadas 415. En otros modos de realización, el indicador se usa para analizar un valor de umbral de TID de una manera diferente.

[0073] Además, en el modo de realización ilustrado del Escenario A, el segundo valor de umbral de TID puede ser un valor de "4". Por lo tanto, como se describe adicionalmente en relación con las FIG 5 y 6, las imágenes en la primera capa 410A que están asociadas con un valor de TID inferior a 4 pueden usarse para la predicción entre capas de imágenes en la segunda capa de mejora 415A2. En consecuencia, todas las primeras imágenes 430A1 (que tiene un valor de TID de "0"), 430A2 (que tiene un valor de TID de "2"), 430A3 (que tiene un valor de TID de "1"), 430A4 (que tiene un valor de TID de "2"), y 430A5 (que tiene un valor de TID de "0"), cada una de los cuales está asociada con un valor de TID por debajo de "4", no están restringidas para el uso para la predicción entre capas de sus respectivas segundas imágenes de la capa de mejora 432A1, 432A2, 432A3, 432A4 y 432A5. Como se ha analizado anteriormente, algunas de las imágenes no restringidas se pueden usar para la predicción entre capas (como se indica con una flecha), mientras que otras no se pueden usar para la predicción entre capas (según lo indicado por la ausencia de una flecha).

[0074] En otro modo de realización (por ejemplo, en el Escenario B ilustrado), un procesador (por ejemplo, el procesador 13) puede determinar (por ejemplo, indicar, recibir, enviar, etc.) un primer valor de umbral de TID que se puede usar para determinar si las imágenes en la primera capa 410B pueden usarse para predecir entre capas imágenes en la primera capa de mejora 415B1 de la primera capa. El procesador puede determinar además un segundo valor de umbral de TID que se usará para determinar si las imágenes en la primera capa 410B pueden usarse para predecir entre capas imágenes en la segunda capa de mejora 415B2 de la primera capa. Los valores primero y segundo de umbral de TID pueden ser iguales o diferentes entre sí.

[0075] En el modo de realización ilustrado del Escenario B, el primer valor de umbral de TID puede ser un valor de "6". Por lo tanto, como se describe adicionalmente en relación con las FIG. 5 y 6, las imágenes en la primera capa 410B que están asociadas con un valor de TID por debajo de "6" se pueden usar para la predicción entre capas de imágenes en la primera capa de mejora 415B1 (por ejemplo, no están restringidas para su uso en la predicción entre capas). En consecuencia, todas las primeras imágenes 430B1 (que tiene un valor de TID de "0"), 430B2 (que tiene un valor de TID de "2"), 430B3 (que tiene un valor de TID de "1"), 430B4 (que tiene un valor de TID de "2") y 430B5 (que tiene un valor de TID de "0"), cada una asociada a valores de TID por debajo de "6", no están restringidos para el uso para la predicción entre capas de sus respectivas imágenes de capa de mejora 431B1, 431B2, 431B3, 431B4 y 431B5. Como se ha analizado anteriormente, algunas de las imágenes no restringidas se pueden usar para la predicción entre capas (como se indica con una flecha), mientras que otras no se pueden usar para la predicción entre capas (según lo indicado por la ausencia de una flecha).

[0076] Además, en el modo de realización ilustrada del Escenario B, el segundo valor de umbral de TID es un valor de "1". Por lo tanto, como se describe adicionalmente en relación con las FIG. 5 y 6, las imágenes en la primera capa 410B que están asociadas con un valor de TID por debajo de "1" se pueden usar para la predicción entre capas de imágenes en la segunda capa de mejora 415B2 (por ejemplo, no están restringidas para su uso en la predicción entre capas). En consecuencia, las primeras imágenes 430B2 (que tiene un valor de TID de "2"), 430B3 (que tiene un valor de TID de "1") y 430B4 (que tiene un valor de TID de "2"), cada una asociada con un valor de TID de "1" o superior, no pueden utilizarse para la predicción entre capas de las imágenes de mejora 432B2, 432B3 y 432B4, respectivamente (cada restricción se muestra como una flecha con una X). A la inversa, las primeras imágenes 430B1 y 430B5, cada una de las cuales está asociada con valores de TID de "0" (que está por debajo de un valor de "1"), no tienen un uso restringido para la predicción entre capas de sus respectivas primeras imágenes de capa de mejora 431B1 y 431B5. Como se ha analizado anteriormente, algunas de las imágenes no restringidas se pueden usar para la predicción entre capas (como se indica con una flecha), mientras que otras no se pueden usar para la predicción entre capas (según lo indicado por la ausencia de una flecha).

[0077] En otros aspectos, como se ha descrito anteriormente, el procesador 13 puede indicar por separado que ninguna de las imágenes en la primera capa puede usarse para predecir entre capas imágenes de una o más segundas capas (por ejemplo, de una o más capas de mejora). En un modo de realización de este tipo (no representada), el decodificador 30 puede descartar (por ejemplo, eliminar de un almacén temporal) tales imágenes de la primera capa. Como se describe en el presente documento, descartar tales imágenes de la primera capa puede ahorrar memoria y costes de procesamiento durante la predicción entre capas de una capa de mejora asociada. Por ejemplo, tal configuración puede permitir la flexibilidad para implementar la decodificación leve (por ejemplo,

decodificación de bucle único) para todas las imágenes de mejora asociadas con respecto a las imágenes en la primera capa.

5 **[0078]** En el contexto de los actuales procedimientos de extensión de la HEVC, descritos anteriormente, en un modo de realización, las técnicas descritas en relación con la FIG. 4 se pueden usar para modificar cierta semántica de la HEVC, a fin de permitir que un codificador de vídeo fije el *max_tid_ref_present_flag* igual a 1 para especificar que el elemento sintáctico *max_tid_il_ref_pics_plus1* está presente y asociado con una primera capa (por ejemplo, la primera capa 410A) con respecto a una segunda capa (por ejemplo, la primera capa de mejora 415A1). Por ejemplo, una nueva definición semántica del indicador *max_tid_ref_present_flag* puede incluir: "el *max_tid_ref_present_flag* igual a 1 especifica que el elemento sintáctico *max_tid_il_ref_pics_plus1[i][j]* está presente. El *max_tid_ref_present_flag* igual a 0 especifica que el elemento sintáctico *max_tid_il_ref_pics_plus1[i][j]* no está presente". En la nueva definición, el valor del índice [j] puede representar una segunda capa (por ejemplo, una capa de mejora).

15 **[0079]** Usando la nueva definición semántica del *max_tid_ref_present_flag*, ciertas otras semánticas de la HEVC pueden modificarse para permitir la definición por separado, para cada una de las capas de mejora, de si las imágenes en la primera capa se pueden usar para la predicción entre capas de la imagen de la capa de mejora en la respectiva capa de mejora. Por ejemplo, una nueva definición semántica de *max_tid_il_ref_pics_plus1[i][j]* puede incluir: "*max_tid_il_ref_pics_plus1[i][j]* igual a 0 especifica que, dentro del CVS, las imágenes no IRAP con *nuh_layer_id* igual a *layer_id_in_nuh[i]* no se usan como referencia para la predicción entre capas para imágenes con *nuh_layer_id* igual a *layer_id_in_nuh[j]*". *max_tid_il_ref_pics_plus1[i][j]* mayor que 0 especifica que, dentro del CVS, las imágenes con *nuh_layer_id* igual a *layer_id_in_nuh[i]* y *TemporalId* mayor que *max_tid_il_ref_pics_plus1[i][j] - 1* no se usan como referencia para la predicción entre capas para las imágenes con *nuh_layer_id* igual a *layer_id_in_nuh[j]*. Cuando no está presente, se deduce que *max_tid_il_ref_pics_plus1[i][j]* es igual a 7". En la nueva definición, el valor del índice [j] y su texto asociado pueden representar una segunda capa (por ejemplo, una capa de mejora). Una implementación ejemplar de estas nuevas definiciones se muestra en la Tabla 1 de sintaxis a continuación, donde se representa el nuevo valor de índice, [j]:

Tabla 1

vps_extension () {	Descriptor
avc_base_layer_flag	u(1)
.....	
for(i = 1; i <= vps_max_layers_minus 1; i++)	
for(j = 0; j < i; j++)	
direct_dependency_flag[i][j]	u(1)
max_tid_ref_present_flag	u(1)
if (max_tid_ref_present_flag)	
for(i = 0; i < vps_max_layers_minus 1; i++)	
for(j = i + 1; j <= vps_max_layers_minus 1; j++)	
if (direct_dependency_flag [j] [i])	
max_tid_il_ref_pics_plus1[i][j]	u(3)
.....	u(1)
vps_vui_present_flag	u(1)
if (vps_vui_present_flag) {	
while(!byte_aligned())	
vps_vui_alignment_bit_equal_to_one	u(1)
vps_vui ()	
}	
}	

30 **[0080]** Además, en algunas realizaciones, otra semántica general de cabecera de segmento de fragmento se puede cambiar en función del uso del valor del índice [j], como se ha descrito anteriormente. Por ejemplo, con respecto al flujo de bits, para cada valor de i en el rango de 0 a alguna variable basada en la HEVC (por ejemplo, *NumActiveRefLayerPics - 1*), inclusive, al menos una de las dos condiciones puede ser verdadera: (1) el valor de *max_tid_il_ref_pics_plus1 [LayerIdxInVps [RefPicLayerId[i]]] [LayerIdxInVps[nuh_layer_id]* puede ser mayor que *TemporalId*; y (2) los valores de *max_tid_il_ref_pics_plus1 [LayerIdxInVps [RefPicLayerId[i]]] [LayerIdxInVps*

[*nuh_layer_id*] y *TemporalId* pueden ser ambos igual a 0 y la imagen en la unidad de acceso actual con *nuh_layer_id* igual a *RefPicLayerId[i]* pueden ser una imagen IRAP.

- 5 **[0081]** Además, en algunas realizaciones, ciertos procesos (por ejemplo, Un proceso de marcado) para imágenes de subcapa no de referencia pueden no ser necesarios para la predicción entre capas en algunos casos. Tras este resultado, el código ejemplar que utiliza el valor de índice *j* puede incluir:

```

        remainingInterLayerReferencesFlag= 0

        iLidx = LayerIdxInVps[ TargetDecLayerIdList [i]]
        for(j = latestDecldx + 1; j < numTargetDecLayers; j++) {
            jLidx = LayerIdxInVps[TargetDecLayerIdList[j]]
            if (currTid <= (max_tid_il_ref_pics_plus1 [iLidx] [jLidx] -1))
                for(k = 0; k <= NumDirectRefLayers[ TargetDecLayerIdList[j]]; k++)
                    if (TargetDecLayerIdList[j] == RefLayerId [TargetDecLayerIdList[j]] [k])
                        remainingInterLayerReferencesFlag= 1
        }

```

- 10 **[0082]** El procesador 13, el codificador de vídeo 20 y/o el decodificador de vídeo 30 pueden indicar, determinar, definir, analizar y/o utilizar las indicaciones y determinaciones descritas anteriormente de varias maneras. Las FIG. 5 y 6 proporcionan diagramas de flujo de procedimientos ejemplares que el codificador de vídeo 20 y el decodificador de vídeo 30, respectivamente, pueden realizar con respecto a la indicación y/o determinación, respectivamente. En otros modos de realización, otros dispositivos pueden realizar una o más de las etapas de procedimiento descrito.

- 15 **[0083]** La FIG. 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 por separado las dependencias de predicción entre capas para una primera capa (por ejemplo, la primera capa 410 de la FIG. 4) con respecto a una o más segundas capas (por ejemplo, las capas de mejora 415 de la FIG. 4). El procesador 13 puede indicar por separado, para cada una de las segundas capas, si las imágenes en la primera capa se pueden usar para la predicción entre capas de la imagen en una segunda capa respectiva.
- 20

- [0084]** El procedimiento comienza en el bloque 505. En el bloque 510, el procesador 13 puede determinar si se realiza el análisis de predicción entre capas de la primera capa (por ejemplo, la primera capa 410A) con respecto a una segunda capa (por ejemplo, la primera capa de mejora 415A1 de la primera capa). En un aspecto, el procesador 13 puede tomar esta determinación basándose en si el procesador 13 tiene suficiente información (por ejemplo, en una cabecera de fragmento) para realizar un análisis de predicción entre capas (por ejemplo, en el bloque 550, como se describe a continuación) de la primera capa con respecto a la segunda capa. Como ejemplo, el procesador 13 puede determinar no realizar un análisis de predicción entre capas de la primera capa con respecto a la segunda capa si la segunda capa no depende directamente de la primera capa.
- 25

- [0085]** Si, en el bloque 510, el procesador 13 determina no realizar el análisis de predicción entre capas de la primera capa con respecto a la segunda capa, entonces, en el bloque 520, el procesador 13 puede señalar que el análisis de predicción entre capas no se realiza para esta asociación de primera capa / capa de mejora. En un modo de realización, si el procesador 13 procesa capas adicionales (como se describe a continuación con respecto al bloque 530), el procesador 13 puede, en cambio, esperar hasta el final del procedimiento para proporcionar una señal única para todas las capas. En un aspecto, el procesador 13 puede señalar la indicación como un indicador en una cabecera de fragmento, por ejemplo, un indicador fijado en un valor de "0". En un aspecto, el indicador puede ser un indicador basado en la HEVC, por ejemplo, *max_tid_ref_present_flag*. En un aspecto, el procesador 13 puede señalar esta indicación a otro dispositivo (por ejemplo, el decodificador de vídeo 30 de la FIG. 3) que puede usar la indicación para determinar si una variable está presente y para realizar operaciones adicionales, como se describe en relación con la FIG. 6. En un modo de realización, el procesador 13 puede avanzar entonces al bloque 530 donde puede determinar si se procesan capas de mejora adicionales (por ejemplo, la segunda capa de mejora 415A2) con respecto a la primera capa (por ejemplo, la primera capa 410A). Si el procesador 13 decide procesar capas de mejora adicionales con respecto a la primera capa, entonces el procedimiento 500 puede regresar al bloque 510 y analizar la primera capa (por ejemplo, la primera capa 410A) con respecto a la segunda capa siguiente (por ejemplo, la segunda capa de mejora 415A2). Si el procesador 13 determina no procesar segundas capas adicionales con respecto a la primera capa, entonces el procedimiento 500 termina en el bloque 590. En el modo de realización en el que el procesador 13 espera hasta el final del procedimiento para señalar una vez para todas las distintas asociaciones de capas (por ejemplo, para cada asociación de capa distinta, de la primera capa con cada una de las segundas capas distintas) a la vez, el procesador 13 puede señalar en consecuencia y luego avanzar al bloque 590.
- 30
- 35
- 40
- 45
- 50

- [0086]** Si, en el bloque 510, el procesador 13 determina realizar un análisis de predicción entre capas de la primera capa con respecto a la segunda capa (por ejemplo, una capa de mejora) que se está considerando, entonces, en el bloque 540, el procesador 13 puede señalar que el análisis de predicción entre capas se realiza para el par primera capa / segunda capa. En un modo de realización ejemplar, el procesador 13 puede, en cambio, esperar hasta el final del procedimiento para señalar una vez para todas las asociaciones de capas distintas (por ejemplo, para cada
- 55

asociación distinta de capas, de la primera capa con cada una de las segundas capas distintas) a la vez. En un aspecto, el procesador 13 puede señalar la indicación como un indicador en una cabecera de fragmento, por ejemplo, un indicador fijado en un valor de "1". En un aspecto, el indicador puede ser un indicador basado en la HEVC, por ejemplo, *max_tid_ref_present_flag*. En un aspecto, el procesador 13 puede señalar esta indicación a otro dispositivo (por ejemplo, el decodificador de vídeo 30 de la FIG. 3) que puede usar la indicación para determinar si una variable está presente y para realizar operaciones adicionales, como se describe en relación con la FIG. 6.

[0087] En el bloque 550, el procesador 13 puede determinar si las imágenes de la primera capa (por ejemplo, las primeras imágenes 430A1-430A5) se pueden usar para predecir entre capas las imágenes de la segunda capa (por ejemplo, la primera capa de mejora 415A1). En un aspecto, el procesador 13 puede tomar esta determinación basándose en varias propiedades de las imágenes en la primera capa, varias propiedades de las imágenes en la segunda capa o cualquier otro número de factores. Algunos factores ejemplares pueden incluir: la habilitación de la codificación leve de bucle único y/o la reducción de la complejidad de la codificación, el nivel de calidad de imagen de la primera capa en comparación con el nivel de calidad de imagen de la capa de mejora, etc. Por ejemplo, si una imagen en la primera capa (por ejemplo, la primera capa 410A) es diferente a una imagen en la segunda capa en más de una magnitud de umbral (por ejemplo, basándose en una cuantización más burda), entonces el procesador 13 puede determinar que la imagen en la primera capa no se puede usar para predecir entre capas la imagen respectiva en la segunda capa (por ejemplo, la segunda capa de mejora 415A1).

[0088] Si, en el bloque 550, el procesador 13 determina que las imágenes de la primera capa 410A no se utilizarán para predecir entre capas imágenes de la segunda capa, entonces, en el bloque 560, el procesador 13 puede configurar una variable para inhabilitar el uso de las imágenes de la primera capa para la predicción entre capas de las imágenes de la segunda capa. En un modo de realización, la variable puede comprender un primer valor de índice (por ejemplo, [i]) y un segundo valor de índice (por ejemplo, [j]), donde el primer valor de índice [i] está asociado con, e identifica, la primera capa (por ejemplo, primera capa 410A), y donde el segundo valor de índice [j] está asociado con, e identifica, la segunda capa (por ejemplo, la primera capa de mejora 415A1). En un modo de realización, la variable puede comprender una variable basada en la HEVC que se fija en "0", por ejemplo, *max_tid_il_ref_pics_plus1 [i] [j] = 0*. En un ejemplo, el procesador 13 puede fijar una variable *max_tid_il_ref_pics_plus1 [0] [1]* igual a un valor de "0" (por ejemplo, para indicar que el TID máximo asociado con el par de capas [i], [j], donde i corresponde a la primera capa 410A, "Capa 0", cuando i = 0, y donde j corresponde a la primera capa de mejora 415A1, "Capa 1", cuando j = 1, es igual a un valor de "0"). En este ejemplo, la variable *max_tid_il_ref_pics_plus1 [0] [1] = 0* puede indicar que las imágenes en la primera capa 410A no se usan como referencia para la predicción entre capas para las imágenes en la primera capa de mejora 415A1. En el bloque 580, el procesador 13 puede entonces señalar esta variable (por ejemplo, en una cabecera de fragmento). Otro dispositivo (por ejemplo, el decodificador 30) puede usar la variable para realizar operaciones adicionales, como se describe en relación con la FIG. 6. En otras palabras, el procesador 13 puede indicar si las imágenes en la primera capa 410A pueden o no pueden usarse para predecir imágenes en una capa de mejora (por ejemplo, la primera capa de mejora 415A1) sin indicar si las imágenes en la primera capa 410A pueden o no pueden ser utilizadas para predecir imágenes en otra capa de mejora (por ejemplo, la segunda capa de mejora 415A2). Es decir, el procesador 13 puede indicar por separado si las imágenes en la primera capa 410A se pueden usar para predecir imágenes en la primera capa de mejora 415A1 (esta indicación puede no incluir ninguna indicación relacionada con la segunda capa de mejora 415A2) y luego puede indicar por separado si las imágenes en la primera capa 410A se pueden usar para predecir imágenes en la segunda capa de mejora 415A2 (esta indicación puede no incluir ninguna indicación relacionada con la primera capa de mejora 415A1). El procesador 13 puede entonces continuar hacia el bloque 530, como se ha descrito anteriormente. En un modo de realización, si el procesador 13 procesa múltiples capas (como se determina en el bloque 530), el procesador 13 puede, en cambio, esperar hasta el final del procedimiento para señalar una vez todas las asociaciones de capas distintas (por ejemplo, para cada asociación de capas distinta, de la primera capa con cada una de las segundas capas distintas) de inmediato. Por ejemplo, en un modo de realización, si el procesador 13 procesa múltiples capas, el procesador 13 puede esperar hasta el final del procedimiento para señalar todas las señales de asociación de capas distintas, de uno o más de los bloques 520, 540 y/o 580 de inmediato.

[0089] Si, en el bloque 550, el procesador 13 determina que las imágenes de la primera capa se pueden usar para predecir entre capas imágenes de la segunda capa, entonces, en el bloque 570, el procesador 13 puede establecer una variable para permitir el uso de las imágenes de la primera capa para la predicción entre capas de las imágenes de la segunda capa. En un modo de realización, la variable puede comprender un primer valor de índice (por ejemplo, [i]) y un segundo valor de índice (por ejemplo, [j]), donde el primer valor de índice [i] está asociado con, e identifica, la primera capa (por ejemplo, primera capa 410A), y donde el segundo valor de índice [j] está asociado con, e identifica, la segunda capa (por ejemplo, la primera capa de mejora 415A1). En un modo de realización, la variable puede comprender una variable basada en la HEVC, por ejemplo, *max_tid_il_ref_pics_plus1 [i] [j]*. En un ejemplo, el procesador 13 puede fijar una variable *max_tid_il_ref_pics_plus1 [0] [1]* (donde i corresponde a la primera capa 410A, "Capa 0", cuando i = 0, y donde j corresponde a la primera capa de mejora 415A1, "Capa 1", cuando j = 1) en un valor, como se describe adicionalmente más adelante.

[0090] En un modo de realización, el procesador 13 puede fijar la variable en un valor entero (por ejemplo, un valor entero entre "0" y "6"), que puede representar un valor de umbral de TID. Como ejemplo, puede ser beneficioso para el procesador 13 fijar la variable en un valor entero que representa un valor de umbral de TID cuando cada una de las

imágenes (por ejemplo, las imágenes de la primera capa 430 y las imágenes de la capa de mejora 431 y 432) están asociadas con un valor particular de TID. En un modo de realización, tal configuración puede permitir que el procesador 13 (o un dispositivo diferente) determine que las imágenes que están asociadas con un valor de TID particular, que es mayor que el valor de umbral de TID, no se usarán para la predicción entre capas (como se describe adicionalmente más adelante). En un modo de realización, el valor de la variable puede indicar que la primera imagen (por ejemplo, la primera imagen 430 en la primera capa 410A) puede usarse para la predicción entre capas de una segunda imagen (por ejemplo, una imagen en la primera capa de mejora 415A1) si la primera imagen está asociada con un valor de TID menor que el valor de umbral de TID, como se describe con más detalle en relación con la FIG. 6. Para ilustrar, el procesador 13 puede fijar la variable *max_tid_il_ref_pics_plus1 [0] [1]* = 3, lo que indicaría un valor de umbral de TID de "3" para las imágenes de la primera capa con respecto a las imágenes de la segunda capa (las imágenes de la segunda capa de mejora 431A). En el bloque 580, el procesador 13 puede entonces señalar esta variable (por ejemplo, en una cabecera de fragmento). Un dispositivo (por ejemplo, el decodificador 30) puede usar la variable para realizar operaciones adicionales, como se describe en relación con la FIG. 6. En otras palabras, el procesador 13 puede indicar si las imágenes en la primera capa 410A pueden o no pueden usarse para predecir imágenes en una capa de mejora (por ejemplo, la capa de mejora 415A1) sin indicar si las imágenes en la primera capa 410A pueden o no pueden usarse para predecir imágenes en otra capa de mejora (por ejemplo, la capa de mejora 415A2). El procesador 13 puede entonces continuar hacia el bloque 530, como se ha descrito anteriormente. En un modo de realización, si el procesador 13 procesa múltiples capas (como se determina en el bloque 530), el procesador 13 puede, en cambio, esperar hasta el final del procedimiento para señalar una vez todas las asociaciones de capas distintas (por ejemplo, para cada asociación de capas distinta, de la primera capa con cada una de las segundas capas distintas) de inmediato. Por ejemplo, en un modo de realización, si el procesador 13 procesa múltiples capas, el procesador 13 puede esperar hasta el final del procedimiento para señalar todas las señales de asociación de capas distintas, de uno o más de los bloques 520, 540 y/o 580 de inmediato.

[0091] En otro modo de realización, el procesador 13 puede no fijar ningún valor para la variable. En dicho modo de realización, se puede deducir que la variable tiene un cierto valor (por ejemplo, un valor entero de "7"). Como ejemplo, puede ser beneficioso para el procesador 13 no fijar la variable en un valor entero (por ejemplo, suponiendo que el valor entero es "7") cuando cada una de las imágenes (por ejemplo, las imágenes de la primera capa 430 y las imágenes 431 y 432 de la capa de mejora) está asociada con un valor particular de TID y cada una se usa para la predicción entre capas. En un modo de realización, tal configuración puede permitir que el procesador 13 (o un dispositivo diferente) determine que cada una de las imágenes se usará para la predicción entre capas, por ejemplo, porque cada una de las imágenes está asociada con un posible valor de TID (por ejemplo, un valor de TID que es menor que "7") como se describe adicionalmente más adelante. En un modo de realización, el valor de la variable puede indicar que la primera imagen (por ejemplo, la primera imagen 430 en la primera capa 410A) puede usarse para la predicción entre capas de una segunda imagen (por ejemplo, una imagen en la primera capa de mejora 415A1), como se describe adicionalmente en relación con la FIG. 6. En el bloque 580, el procesador 13 puede entonces señalar esta variable (por ejemplo, en una cabecera de fragmento). Un dispositivo (por ejemplo, el decodificador 30) puede usar la variable para realizar operaciones adicionales, como se describe en relación con la FIG. 6. El procesador 13 puede entonces continuar hacia el bloque 530, como se ha descrito anteriormente. En un modo de realización, si el procesador 13 procesa múltiples capas (como se determina en el bloque 530), el procesador 13 puede, en cambio, esperar hasta el final del procedimiento para señalar una vez todas las asociaciones de capas distintas (por ejemplo, para cada asociación de capas distinta, de la primera capa con cada una de las segundas capas distintas) de inmediato. Por ejemplo, en un modo de realización, si el procesador 13 procesa múltiples capas, el procesador 13 puede esperar hasta el final del procedimiento para señalar todas las señales de asociación de capas distintas, de uno o más de los bloques 520, 540 y/o 580 de inmediato.

[0092] A continuación del Escenario A de la FIG. 4, junto con el procedimiento descrito en relación con la FIG. 5, el procesador 13 puede determinar realizar un análisis de predicción entre capas para una primera capa (por ejemplo, la primera capa 410A) con respecto a una segunda capa (por ejemplo, la capa de mejora 415A1) en el bloque 510. El procesador 13 puede señalar esto en el bloque 540, por ejemplo, fijando y señalizando el indicador *max_tid_ref_present_flag* igual a 1. En el bloque 550, el procesador 13 puede entonces determinar que las imágenes en la primera capa se pueden usar para predecir entre capas la imagen en la segunda capa. Por lo tanto, en el bloque 570, el procesador 13 puede fijar una variable (por ejemplo, la variable *max_tid_il_ref_pics_plus1 [i] [j]*) para permitir el uso de la primera capa para predecir entre capas las imágenes de la segunda capa. En este ejemplo, el procesador 13 puede fijar la variable igual a un valor de "2", de manera que las imágenes en la primera capa 410A, que está asociadas con un valor de TID mayor o igual a un valor de "2", no se puedan usar para predecir entre capas imágenes en la capa de mejora 415A1 (como se muestra por las flechas con X). Luego, en el bloque 580, el procesador 13 puede señalar la variable con referencia a la primera capa y la segunda capa 415A1. Por ejemplo, en este caso, el procesador 13 puede señalar *max_tid_il_ref_pics_plus1 [0] [1]* = 2. Luego, en el bloque 530, el procesador 13 puede determinar procesar capas de mejora adicionales (por ejemplo, la capa de mejora 415A2) con respecto a la primera capa 410A, y luego regresar al bloque 510. El procesador 13 puede proceder entonces de manera similar a como se ha descrito anteriormente, excepto que, en el bloque 570, el procesador 13 puede fijar la variable igual a algún valor mayor que 2 (por ejemplo, 3 a 6) o no fijar el valor en absoluto (por ejemplo, lo que implica un valor de 7). Por lo tanto, dado que todas las imágenes en la primera capa 410A están asociadas con un valor de TID de 2 o inferior, todas las imágenes en la primera capa 410A pueden usarse para predecir entre capas imágenes en la capa de mejora 415A2. Luego, en el bloque 580, el procesador 13 puede señalar la variable con referencia a la primera capa y a la segunda

capa. Por ejemplo, en este caso, el procesador 13 puede señalar *max_tid_il_ref_pics_plus1 [0] [2] = 4*. Luego, en el bloque 530, el procesador 13 puede determinar no procesar ninguna capa de mejora adicional con respecto a la primera capa 410A, y luego finalizar el procedimiento en el bloque 590. En un modo de realización, si el procesador 13 procesa múltiples capas (como se determina en el bloque 530), el procesador 13 puede, en cambio, esperar hasta el final del procedimiento para señalar una vez todas las asociaciones de capas distintas (por ejemplo, para cada asociación de capas distinta, de la primera capa con cada una de las segundas capas distintas) de inmediato. Por ejemplo, en un modo de realización, si el procesador 13 procesa múltiples capas, el procesador 13 puede esperar hasta el final del procedimiento para señalar todas las señales de asociación de capas distintas, de uno o más de los bloques 520, 540 y/o 580 de inmediato.

[0093] A continuación del Escenario B de la FIG. 4, junto con el procedimiento descrito en relación con la FIG. 5, el procesador 13 puede determinar realizar un análisis de predicción entre capas para la primera capa 410B con respecto a la capa de mejora 415B1 en el bloque 510. El procesador 13 puede señalar esto en el bloque 540, por ejemplo, al fijar y señalar el indicador *max_tid_ref_present_flag* igual a un valor de "1". En un modo de realización, el procesador 13 puede esperar para hacer tal señal hasta el final del procedimiento, como se ha descrito anteriormente. En el bloque 550, el procesador 13 puede entonces determinar que las imágenes en la primera capa 410B pueden usarse para predecir entre capas las imágenes en la capa de mejora 415B1. Por lo tanto, en el bloque 570, el procesador 13 puede fijar una variable (por ejemplo, la variable *max_tid_il_ref_pics_plus1 [i] [j]*) para permitir el uso de la primera capa 410B para predecir entre capas las imágenes de la capa de mejora 415B1. En este ejemplo, el procesador 13 puede fijar la variable igual a algún valor mayor que "2" (por ejemplo, "3" a "6") o no fijar el valor en absoluto (por ejemplo, lo que implica un valor de 7). Por lo tanto, dado que todas las imágenes en la primera capa 410B están asociadas con un valor de TID de "2" o inferior, todas las imágenes en la primera capa 410B pueden usarse para predecir entre capas las imágenes en la capa de mejora 415B1. Luego, en el bloque 580, el procesador 13 puede señalar la variable con referencia a la primera capa 410B y la capa de mejora 415B1. Por ejemplo, en este caso, el procesador 13 puede señalar *max_tid_il_ref_pics_plus1 [0] [1] = 5*. Luego, en el bloque 530, el procesador 13 puede determinar procesar capas de mejora adicionales (por ejemplo, la capa de mejora 415B2) con respecto a la primera capa 410B, y luego regresar al bloque 510. El procesador 13 puede entonces proceder de manera similar a como se ha descrito anteriormente, excepto que, en el bloque 570, el procesador 13 puede fijar la variable igual a "1", de manera que las imágenes en la primera capa 410B que están asociadas con un valor de TID mayor o igual a 1 no puedan usarse para predecir entre capas imágenes en la capa de mejora 415B2 (como lo muestran las flechas con X). Luego, en el bloque 580, el procesador 13 puede señalar la variable con referencia a la primera capa 410B y la capa de mejora 415B2. Por ejemplo, en este caso, el procesador 13 puede señalar *max_tid_il_ref_pics_plus1 [0] [2] = 1*. Luego, en el bloque 530, el procesador 13 puede determinar no procesar ninguna capa de mejora adicional con respecto a la primera capa 410B, y luego finalizar el procedimiento en el bloque 590. En un modo de realización, si el procesador 13 procesa múltiples capas (como se determina en el bloque 530), el procesador 13 puede, en cambio, esperar hasta el final del procedimiento para señalar una vez todas las asociaciones de capas distintas (por ejemplo, para cada asociación de capas distinta, de la primera capa con cada una de las segundas capas distintas) de inmediato. Por ejemplo, en un modo de realización, si el procesador 13 procesa múltiples capas, el procesador 13 puede esperar hasta el final del procedimiento para señalar todas las señales de asociación de capas distintas, de uno o más de los bloques 520, 540 y/o 580 de inmediato.

[0094] Como se ha descrito anteriormente en relación con la FIG. 4, los procesos anteriores son solo ejemplos de cómo el procesador 13 y/o el codificador de vídeo 20 pueden hacer las indicaciones descritas.

[0095] La FIG. 6 ilustra un diagrama de flujo de un procedimiento para un procesador (por ejemplo, el procesador 13) de un decodificador de vídeo (por ejemplo, el codificador de vídeo 30) para determinar por separado las dependencias de predicción entre capas, entre una primera capa (por ejemplo, la primera capa 410A de la FIG. 4) y una o más segundas capas (por ejemplo, las capas de mejora primera y segunda 415A1 y 415A2 de la FIG. 4), como se describe en relación con la FIG. 4. Por ejemplo, el procesador 13 puede determinar por separado (por ejemplo, basándose en una indicación respectiva), para cada una de las capas de mejora, si las imágenes en la primera capa se pueden usar para la predicción entre capas de una imagen de capa de mejora en una respectiva capa de mejora.

[0096] El procedimiento comienza en el bloque 605. En el bloque 610, el procesador 13 puede determinar si una variable de análisis de predicción entre capas (la "variable") está presente para una primera capa (por ejemplo, la primera capa 410A) con respecto a una segunda capa (por ejemplo, la capa de mejora 415A1). En un aspecto, el procesador 13 puede tomar esta determinación basándose en una indicación respectiva, que puede haber sido señalada desde un codificador de vídeo (por ejemplo, el codificador de vídeo 20) como se describe en relación con las FIG. 4 y 5. En un aspecto, la indicación puede ser un indicador señalado en una cabecera de fragmento. Por ejemplo, el indicador puede ser un indicador basado en la HEVC, por ejemplo, el *max_tid_ref_present_flag*, como se describe en relación con la FIG. 5.

[0097] En el bloque 610, el procesador 13 puede determinar que la variable no está fijada (por ejemplo, el *max_tid_ref_present_flag* se fija en un valor de "0"). En este caso, el procesador 13 puede avanzar luego al bloque 620, donde el procesador 13 puede determinar si se procesan capas de mejora adicionales (por ejemplo, la segunda capa de mejora 415A2) con respecto a la primera capa. Si el procesador 13 decide procesar capas de mejora adicionales con respecto a la primera capa 410A, entonces el procedimiento 600 puede volver a 610. Sin embargo,

en el bloque 610, el procesador analizará una segunda capa diferente de la que analizó anteriormente (por ejemplo, analizará la segunda capa de mejora 415A2). Si el procesador 13 determina no procesar capas de mejora adicionales con respecto a la primera capa 410A, el procedimiento termina en el bloque 690.

5 **[0098]** Alternativamente, en el bloque 610, el procesador 13 puede determinar que la variable está fijada (por ejemplo, si *max_tid_ref_present_flag* se fija en 1). En este caso, el procesador 13 puede avanzar luego al bloque 630 y determinar si la variable desactiva el uso de imágenes en la primera capa para la predicción entre capas de imágenes de la segunda capa (por ejemplo, la primera capa de mejora 415A1). Por ejemplo, si, como se ha descrito en relación con la FIG. 5, la variable comprende la variable basada en la HEVC *max_tid_il_ref_pics_plus1 [i] [j]*, y la variable se
10 fija en un valor de "0", entonces el procesador 13 puede determinar que la variable desactiva el uso de imágenes en la primera capa para la predicción entre capas de imágenes de la segunda capa (por ejemplo, la primera capa de mejora 415A1). De forma alternativa, como se ha descrito en relación con la FIG. 5, si la variable no se fija en un valor de "0" (por ejemplo, si la variable no se fija o si la variable se fija en algún valor distinto de cero, tal como un valor de "1"), entonces el procesador 13 puede determinar que la variable no desactiva el uso de imágenes en la primera capa
15 430A para la predicción entre capas de imágenes de la capa de mejora 415A1.

[0099] Como se ha descrito anteriormente, en el bloque 630, el procesador 13 puede determinar que la variable desactiva el uso de imágenes en la primera capa para la predicción entre capas de imágenes de la segunda capa (por ejemplo, la primera capa de mejora 415A1) (por ejemplo, si *max_tid_il_ref_pics_plus1 [0] [1] = 0*). En otras palabras,
20 el procesador 13 puede determinar que las imágenes en la primera capa pueden o no pueden usarse para predecir imágenes en una capa de mejora (por ejemplo, la capa de mejora 415A1) sin determinar si las imágenes en la primera capa 410A pueden o no pueden usarse para predecir imágenes en otra capa de mejora (por ejemplo, la capa de mejora 415A2). En este caso, el procesador 13 puede avanzar luego al bloque 640, donde puede descartar (por ejemplo, eliminar de un almacén temporal) las imágenes en la primera capa con respecto a la segunda capa. Por
25 ejemplo, el procesador 13 puede descartar las imágenes en la primera capa que están inhabilitadas para su uso para la predicción entre capas con respecto a la segunda capa. Como se ha descrito anteriormente, descartar las imágenes asociadas de la primera capa puede ahorrar memoria y costes de procesamiento durante la predicción entre capas de la segunda capa. Por ejemplo, como se ha descrito anteriormente, tal configuración puede admitir la flexibilidad para implementar la decodificación leve (por ejemplo, la decodificación de un solo bucle) para una capa de mejora,
30 mientras mantiene la flexibilidad para implementar la decodificación de alta calidad (por ejemplo, la decodificación de múltiples bucles) para una o más de las otras capas de mejora (por ejemplo, la capa de mejora 415A2).

[0100] De forma alternativa, como se ha descrito anteriormente, en el bloque 630, el procesador 13 puede determinar que la variable no desactiva el uso de imágenes en la primera capa para la predicción entre capas de
35 imágenes de la segunda capa (por ejemplo, si *max_tid_il_ref_pics_plus1 [0] [1] ≠ 0*). En este caso, el procesador 13 puede avanzar luego al bloque 650, donde puede determinar que las imágenes en la primera capa que están asociadas con un valor de TID menor que un umbral (por ejemplo, un valor de umbral de TID) se pueden usar para predecir entre capas imágenes de la segunda capa (por ejemplo, primera capa de mejora 415A1). Por ejemplo, el procesador 13 puede determinar que las imágenes en la primera capa 410A pueden o no pueden usarse para predecir
40 imágenes en una capa de mejora (por ejemplo, la capa de mejora 415A1) sin determinar si las imágenes en la primera capa 410A pueden o no pueden usarse para predecir imágenes en otra capa de mejora (por ejemplo, la capa de mejora 415A2). Como se ha descrito en relación con la FIG. 4, aunque algunas de las primeras imágenes 430 pueden no estar restringidas en cuanto a usarse para predecir entre capas sus respectivas imágenes de mejora 431 o 432, el procesador 13 puede, sin embargo, determinar no usar realmente una o más de las imágenes no restringidas para la predicción entre capas. Por ejemplo, el procesador 13 puede tomar esta determinación basándose en la información
45 en la cabecera del fragmento.

[0101] En un modo de realización, la variable se puede fijar en un valor entero (por ejemplo, un valor entero entre "0" y "6"), que puede representar un valor de umbral de TID (por ejemplo, el valor de umbral de TID descrito en relación con el bloque 650). En un modo de realización, el valor de la variable puede indicar que la primera imagen (por
50 ejemplo, la primera imagen 430 en la primera capa 410A) puede usarse para la predicción entre capas de una segunda imagen (por ejemplo, una imagen en la primera capa de mejora 415A1) si la primera imagen está asociada con un valor de TID menor que el valor de umbral de TID, como se describe con más detalle en relación con la FIG. 6. Para ilustrar, el procesador 13 puede fijar la variable *max_tid_il_ref_pics_plus1 [0] [1] = 3*, lo que indicaría un valor de umbral de TID de "3" para las imágenes de la primera capa con respecto a las imágenes de la segunda capa (las imágenes de la segunda capa de mejora 431A).
55

[0102] En otro modo de realización, la variable no puede fijarse en ningún valor. Como se ha descrito en relación con la FIG. 5, en este caso, se puede deducir que la variable tiene un cierto valor (por ejemplo, un valor entero de "7").
60 Como se ha descrito anteriormente, en un modo de realización, tal configuración puede permitir que el procesador 13 (en el bloque 650) determine que cada una de las imágenes se usará para la predicción entre capas, por ejemplo, porque cada una de las imágenes está asociada con un posible valor de TID (por ejemplo, un valor de TID que es menor que "7").

65 **[0103]** A continuación del bloque 640 o del bloque 650, el procesador 13 puede continuar con el bloque 620, donde el procesador 13 repetirá el proceso con una segunda capa adicional (en el bloque 610) o finalizará el procedimiento

(en el bloque 690) como se ha descrito anteriormente. En un modo de realización, cuando finaliza el procedimiento, el decodificador de vídeo 30 puede proceder entonces a decodificar una o más de las imágenes de mejora (por ejemplo, las imágenes de mejora 431 y/o 432) usando la predicción entre capas, basándose en una o más de las imágenes asociadas actuales (por ejemplo, las imágenes actuales 430). En un aspecto, la una o más imágenes actuales asociadas pueden asociarse con un valor de TID que es menor que un valor umbral de TID asociado con la imagen actual y la imagen de la capa de mejora.

[0104] A continuación del Escenario A de la FIG. 4, junto con el procedimiento descrito en relación con la FIG. 6, el procesador 13 puede determinar que la variable está presente para una primera capa (por ejemplo, la primera capa 410A) con respecto a una segunda capa (por ejemplo, la capa de mejora 415A1). Por ejemplo, el procesador 13 puede determinar que el *max_tid_ref_present_flag* se fija en 1. El procesador 13 puede entonces avanzar al bloque 630 y determinar que la variable no desactiva el uso de imágenes en la primera capa para predecir entre capas imágenes de la segunda capa. Por ejemplo, el procesador 13 puede tomar dicha determinación basándose en la determinación de que *max_tid_il_ref_pics_plus1 [0] [1] ≠ 0*. Por lo tanto, el procesador 13 puede avanzar luego al bloque 650 y determinar que las imágenes en la primera capa, que están asociadas con un valor de TID menor que un valor de umbral de TID (por ejemplo, dicho valor de umbral de TID puede estar basado en el valor de *max_tid_il_ref_pics_plus1 [0] [1]*, si acaso), pueden usarse para la predicción entre capas de imágenes de la segunda capa.

[0105] En el ejemplo ilustrado del Escenario A, la variable se puede fijar en "2" (por ejemplo, *max_tid_il_ref_pics_plus1 [0] [1] = 2*), de manera que el procesador 13 pueda determinar que las imágenes en la primera capa (por ejemplo, la primera capa 410A), que están asociadas con un valor de TID mayor o igual a "2", no se pueden usar para la predicción entre capas de imágenes en la segunda capa (por ejemplo, la capa de mejora 415A1), como se muestra por las flechas con X. Luego, en el bloque 620, el procesador 13 puede determinar procesar una segunda capa adicional (por ejemplo, la capa de mejora 415A2) con respecto a la primera capa, y luego regresar al bloque 610. El procesador 13 puede luego proceder de manera similar a como se ha descrito anteriormente, excepto que, en el bloque 650, la variable se puede fijar en un valor mayor que "2" (por ejemplo, un valor de "3" a "6") o la variable no se puede fijar en absoluto (por ejemplo, lo que implica un valor de "7"). Por ejemplo, en el ejemplo ilustrado del Escenario A, la variable *max_tid_il_ref_pics_plus1 [0] [2]* se puede fijar en "4". Por lo tanto, dado que todas las imágenes en la primera capa (por ejemplo, la primera capa 410A) están asociadas con un valor de TID de "4" o menos, el procesador 13 puede determinar que todas las imágenes en la primera capa se pueden usar para la predicción entre capas de imágenes en la segunda capa (por ejemplo, la capa de mejora 415A2). Luego, en el bloque 620, el procesador 13 puede determinar no procesar ninguna capa adicional con respecto a la primera capa, y luego finalizar el procedimiento en el bloque 690. En un modo de realización, cuando finaliza el procedimiento, el decodificador de vídeo 30 puede proceder entonces a decodificar una o más de las imágenes de mejora (por ejemplo, las imágenes de mejora 431A y/o 432A) usando una predicción entre capas basada en una o más de las imágenes asociadas actuales (por ejemplo, las imágenes actuales 430A). En un aspecto, las una o más imágenes asociadas actuales pueden asociarse con un valor de TID menor que un valor de umbral de TID asociado con la imagen actual y la imagen de la capa de mejora.

[0106] A continuación del Escenario B de la FIG. 4, junto con el procedimiento descrito en relación con la FIG. 6, el procesador 13 puede determinar que la variable está presente para una primera capa (por ejemplo, la primera capa 410B) con respecto a una segunda capa (por ejemplo, la capa de mejora 415B1). Por ejemplo, el procesador 13 puede determinar que el *max_tid_ref_present_flag* se fija en 1. El procesador 13 puede entonces avanzar al bloque 630 y determinar que la variable no desactiva el uso de imágenes en la primera capa para predecir entre capas imágenes de la segunda capa. Por ejemplo, el procesador 13 puede tomar dicha determinación basándose en la determinación de que *max_tid_il_ref_pics_plus1 [0] [1] ≠ 0*. Por lo tanto, el procesador 13 puede avanzar luego al bloque 650 y determinar que las imágenes en la primera capa, que están asociadas con un valor de TID menor que un valor de umbral de TID (por ejemplo, dicho valor de umbral de TID puede estar basado en el valor de *max_tid_il_ref_pics_plus1 [0] [1]*, si acaso), pueden usarse para la predicción entre capas de imágenes de la segunda capa. En el ejemplo ilustrado del Escenario B, la variable se puede fijar en un valor mayor que "2" (por ejemplo, un valor de "3" a "6") o la variable puede no fijarse en absoluto (por ejemplo, lo que implica un valor de "7"). Por ejemplo, en el ejemplo ilustrado del Escenario B, la variable *max_tid_il_ref_pics_plus1 [0] [2]* se puede fijar en "5". Por lo tanto, como todas las imágenes en la primera capa (por ejemplo, la primera capa 410B) están asociadas con un valor de TID de "5" o menos, el procesador 13 puede determinar que todas las imágenes en la primera capa se pueden usar para la predicción entre capas de imágenes en la segunda capa (por ejemplo, la capa de mejora 415B2). Luego, en el bloque 620, el procesador 13 puede determinar procesar una segunda capa adicional (por ejemplo, la capa de mejora 415B2) con respecto a la primera capa, y luego regresar al bloque 610. El procesador 13 puede proceder de manera similar a lo descrito anteriormente, excepto que, en el bloque 650, la variable se puede fijar en un valor de "1" (por ejemplo, *max_tid_il_ref_pics_plus1 [0] [2] = 1*), de manera que el procesador 13 pueda determinar que las imágenes en la primera capa (por ejemplo, la primera capa 410B) que están asociadas con un valor de TID mayor o igual a "1" no se pueden usar para la predicción entre capas de imágenes en la segunda capa (por ejemplo, la capa de mejora 415B1), como se muestra con flechas con X. Luego, en el bloque 620, el procesador 13 puede determinar no procesar ninguna capa adicional con respecto a la primera capa, y luego finalizar el procedimiento en el bloque 690. En un modo de realización, cuando finaliza el procedimiento, el decodificador de vídeo 30 puede entonces proceder a decodificar una o más de las imágenes de mejora (por ejemplo, imágenes de mejora 431B y/o 432B) utilizando la predicción entre capas, basándose en una o más de las imágenes asociadas actuales (por ejemplo, las imágenes actuales 430B). En

un aspecto, las una o más imágenes asociadas actuales pueden asociarse con un valor de TID menor que un valor de umbral de TID asociado con la imagen actual y la imagen de la capa de mejora.

5 **[0107]** Como se ha descrito anteriormente en relación con las FIG. 4 y 5, los procesos anteriores son solo ejemplos de cómo el procesador 13 y/o el decodificador de vídeo 30 pueden tomar las determinaciones descritas.

10 **[0108]** Otra realización de un procedimiento para codificar información de vídeo se ilustra en la FIG. 7. El procedimiento 700 comienza en el bloque 710. En el bloque 710, un procesador (por ejemplo, cualquier procesador descrito en el presente documento) almacena una primera imagen asociada con una primera capa. El procesador también almacena imágenes de capa de mejora asociadas con una pluralidad de capas de mejora de la primera capa. Por ejemplo, en un modo de realización, la primera capa corresponde a la primera capa 410A, descrita anteriormente, y la pluralidad de capas de mejora corresponde a las capas de mejora primera y segunda, 415A1 y 415A2, descritas anteriormente.

15 **[0109]** En el bloque 730, el procesador proporciona una indicación distinta para cada una de las capas de mejora. La indicación indica si la primera imagen se puede usar para la predicción entre capas de la imagen de la capa de mejora en una respectiva capa de mejora. Por ejemplo, el procesador puede proporcionar un valor de TID máximo para cada primer par de capas capa-y-mejora, como se ha descrito anteriormente. En un modo de realización, el procesador proporciona un valor de TID máximo en una variable que tiene un índice [i] (donde [i] indica el valor del Identificador de Capa de la primera capa) y un índice [j] (donde [j] indica el valor del Identificador de Capa de cada capa de mejora). Por lo tanto, el procedimiento 700 proporciona al procesador la capacidad de especificar diferentes valores de TID máximos para cada capa de mejora de la capa actual. El procedimiento 700 termina en el bloque 740.

25 **[0110]** 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 transmitir por, un medio legible por ordenador, como una o más instrucciones o código, y ejecutarse mediante una unidad de procesamiento basada en hardware. Los medios legibles por ordenador pueden incluir medios de almacenamiento legibles por ordenador, que correspondan a un medio tangible tal como medios de almacenamiento de datos, o medios de comunicación que incluyan cualquier medio que facilite la transferencia de un programa informático 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 y legibles por ordenador que sean no transitorios o (2) un medio de comunicación tal como una señal o una onda portadora. Los medios de almacenamiento de datos pueden ser medios disponibles cualesquiera a los que se pueda acceder desde uno o más ordenadores o uno o más procesadores para recuperar instrucciones, código y/o estructuras de datos para la implementación de las técnicas descritas en esta divulgación. Un producto de programa informático puede incluir un medio legible por ordenador.

40 **[0111]** 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 una sede de la Red, un servidor u otro origen remoto usando un cable coaxial, un cable de fibra óptica, un par trenzado, una línea de abonado digital (DSL) o tecnologías inalámbricas tales como infrarrojos, radio y microondas, entonces el cable coaxial, el cable de fibra óptica, el par trenzado, la DSL o las tecnologías inalámbricas tales como infrarrojos, radio y microondas se incluyen en la definición de medio. Sin embargo, debería entenderse que los medios de almacenamiento legibles por ordenador y los medios de almacenamiento de datos no incluyen conexiones, ondas portadoras, señales 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.

55 **[0112]** 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 dentro de módulos de hardware y/o software dedicados, configurados para la codificación y la decodificación, o incorporados en un códec combinado. Además, las técnicas se podrían implementar totalmente en uno o más circuitos o elementos lógicos.

65 **[0113]** 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

5 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 que comprende:
- 5 una unidad de memoria configurada para almacenar imágenes asociadas con una primera capa [i] e imágenes de capa de mejora asociadas con una pluralidad de capas de mejora; y
- un procesador en comunicación con la unidad de memoria y configurado para asignar un valor a una variable `max_tid_il_ref_pics_plus1 [i] [j]`, asociada con cada combinación de la primera capa [i] con una
- 10 capa de mejora particular [j] entre la pluralidad de capas de mejora, en donde para cada una entre la pluralidad de capas de mejora, las imágenes asociadas con la primera capa [i] que tienen un valor de identificación temporal menor o igual al valor de la variable `max_tid_il_ref_pics_plus1 [i] [j]`, asociada con la capa de mejora particular [j], se pueden utilizar para la predicción entre capas de las imágenes de capa de mejora asociadas con la capa de mejora particular [j].
- 15
2. El codificador de vídeo de la Reivindicación 1, en el que el procesador está configurado además para:
- determinar para cada capa de mejora si las imágenes asociadas con la primera capa [i] se pueden usar para la predicción entre capas de la capa de mejora particular [j]; y
- 20 asignar el valor de la variable asociada con la capa de mejora particular [j], en función de dicha determinación.
3. El codificador de vídeo de la Reivindicación 1, en el que el procesador está configurado además para:
- 25 señalar un `max_tid_ref_present_flag` en la extensión del conjunto de parámetros de vídeo (VPS), en donde el `max_tid_ref_present_flag` indica la presencia de la variable `max_tid_il_ref_pics_plus1 [i] [j]`.
4. El codificador de vídeo de la Reivindicación 1, en el que el procesador está configurado para asignar a la variable `max_tid_il_ref_pics_plus1 [i] [j]` un valor de 0, basándose en las imágenes asociadas con la primera capa [i] que no se usan para la predicción entre capas de imágenes asociadas con la capa de mejora particular.
- 30
5. Un procedimiento de codificación de vídeo, comprendiendo el procedimiento:
- 35 almacenar imágenes asociadas con una primera capa [i] e imágenes de capa de mejora asociadas con una pluralidad de capas de mejora; y
- asignar un valor a una variable `max_tid_il_ref_pics_plus1 [i] [j]`, asociada con cada combinación de la primera capa [i] con una capa de mejora particular [j], entre la pluralidad de capas de mejora, en donde para cada una entre la pluralidad de capas de mejora, las imágenes asociadas con la primera capa [i] que
- 40 tienen un valor de identificación temporal menor o igual al valor de la variable `max_tid_il_ref_pics_plus1 [i] [j]`, asociada con la capa de mejora particular [j], se pueden usar para la predicción entre capas de imágenes de la capa de mejora asociadas con la capa de mejora particular [j].
- 45
6. El procedimiento de la Reivindicación 5, que comprende además:
- determinar para cada capa de mejora si las imágenes asociadas con la primera capa [i] se pueden usar para la predicción entre capas de la capa de mejora particular [j]; y
- 50 asignar el valor de la variable asociada con la capa de mejora particular [j] en función de dicha determinación.
7. El procedimiento de la Reivindicación 5, que comprende además:
- 55 señalar un `max_tid_ref_present_flag` en la extensión del conjunto de parámetros de vídeo (VPS), donde el `max_tid_ref_present_flag` indica la presencia de la variable `max_tid_il_ref_pics_plus1 [i] [j]`.
8. El procedimiento de la Reivindicación 5, que comprende además asignar a la variable `max_tid_il_ref_pics_plus1 [i] [j]` un valor de 0, basándose en las imágenes asociadas con la primera capa [i] que no se usan para la predicción entre capas de imágenes asociadas con la capa de mejora particular.
- 60
9. Un decodificador de vídeo que comprende:
- 65 una unidad de memoria configurada para almacenar imágenes asociadas con una primera capa [i] e imágenes de mejora asociadas con una pluralidad de capas de mejora; y

- 5 un procesador en comunicación con la unidad de memoria y configurado para determinar, para cada combinación de la primera capa [i] con una capa de mejora particular [j] entre la pluralidad de capas de mejora, que pueden usarse imágenes asociadas con la primera capa [i] para la predicción entre capas de la capa de mejora particular [j], basándose en imágenes asociadas con la primera capa [i] que tienen un valor de identificación temporal menor o igual al valor de una variable `max_tid_il_ref_pics_plus1 [i] [j]`, asociada con la combinación de la primera capa [i] con la capa de mejora particular [j].
- 10 **10.** El decodificador de vídeo de la Reivindicación 9, en el que el procesador está adicionalmente configurado para recibir un `max_tid_ref_present_flag` en la extensión del conjunto de parámetros de vídeo (VPS), en el que el `max_tid_ref_present_flag` indica la presencia de la variable `max_tid_il_ref_pics_plus1 [i] [j]`.
- 11.** El decodificador de vídeo de la reivindicación 9, en el que el procesador está configurado para descartar las imágenes asociadas con la primera capa [i], basándose en un valor de la variable que es 0.
- 15 **12.** El decodificador de vídeo de la Reivindicación 10, en el que el procesador está configurado para determinar que todas las imágenes asociadas con la primera capa [i] se pueden usar para la predicción entre capas de las imágenes asociadas con la capa de mejora particular [j], basándose en el `max_tid_ref_present_flag` que tiene un valor de 0, estando el procesador configurado además para asignar a la variable un valor de 7.
- 20 **13.** Un procedimiento de decodificación de vídeo, comprendiendo el procedimiento:
 almacenar imágenes asociadas con una primera capa [i] e imágenes de capa de mejora asociadas con una pluralidad de capas de mejora; y
- 25 para cada combinación de la primera capa [i] con una capa de mejora particular [j] entre la pluralidad de capas de mejora, las imágenes asociadas con la primera capa [i] se pueden usar para la predicción entre capas de la capa de mejora particular [j], basándose en imágenes asociadas con la primera capa [i] que tienen un valor de identificación temporal menor o igual que el valor de una variable `max_tid_il_ref_pics_plus1 [i] [j]`, asociada con la combinación de la primera capa [i] con la capa de mejora particular [j].
- 30 **14.** El procedimiento de la Reivindicación 13, que comprende además recibir un `max_tid_ref_present_flag` en la extensión del conjunto de parámetros de vídeo (VPS), en donde el `max_tid_ref_present_flag` indica la presencia de la variable `max_tid_il_ref_pics_plus1 [i] [j]`.
- 35 **15.** El procedimiento de la Reivindicación 13, en el que el procedimiento comprende además descartar las imágenes asociadas con la primera capa [i], basándose en que un valor de la variable sea 0.

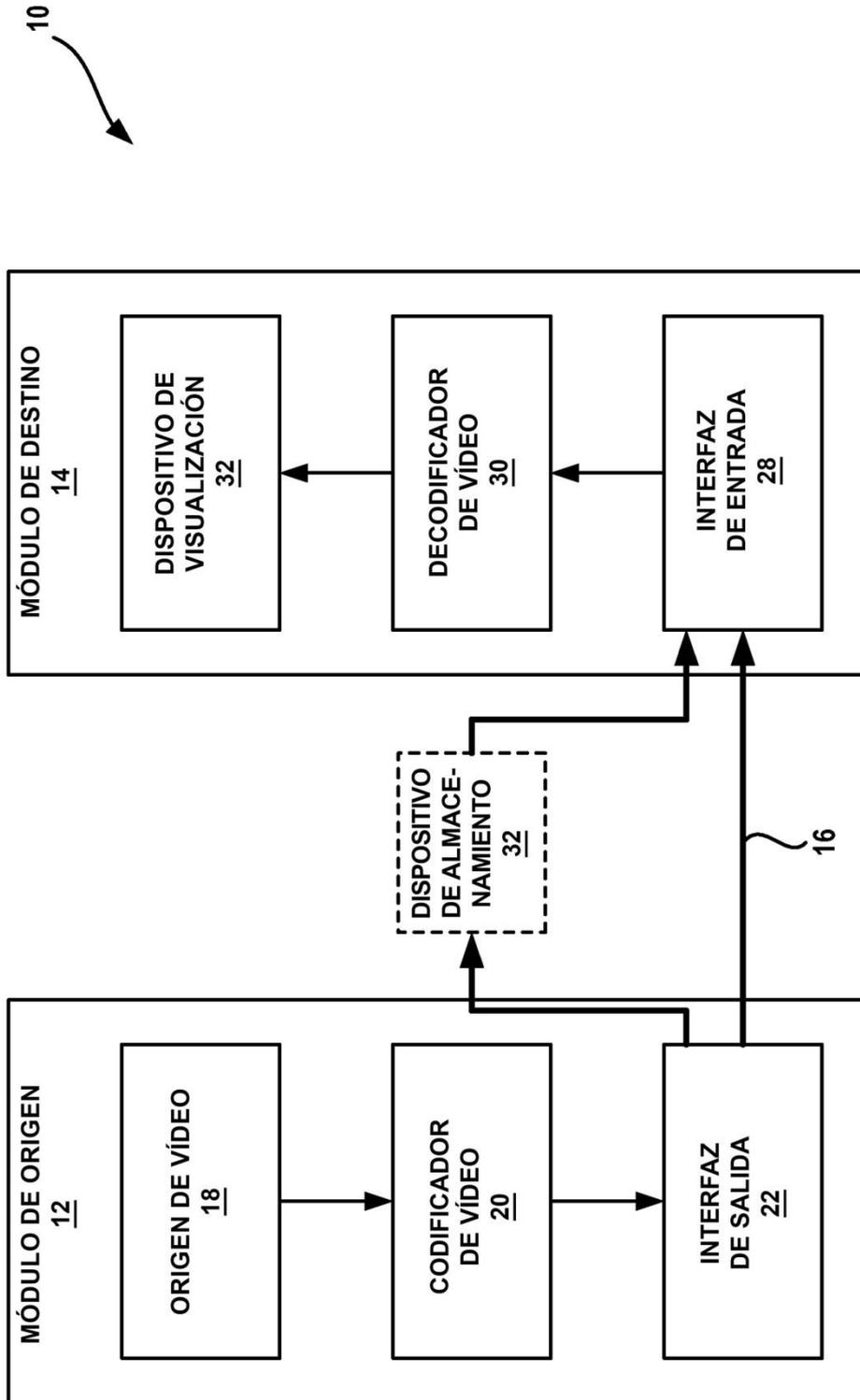


FIG. 1A

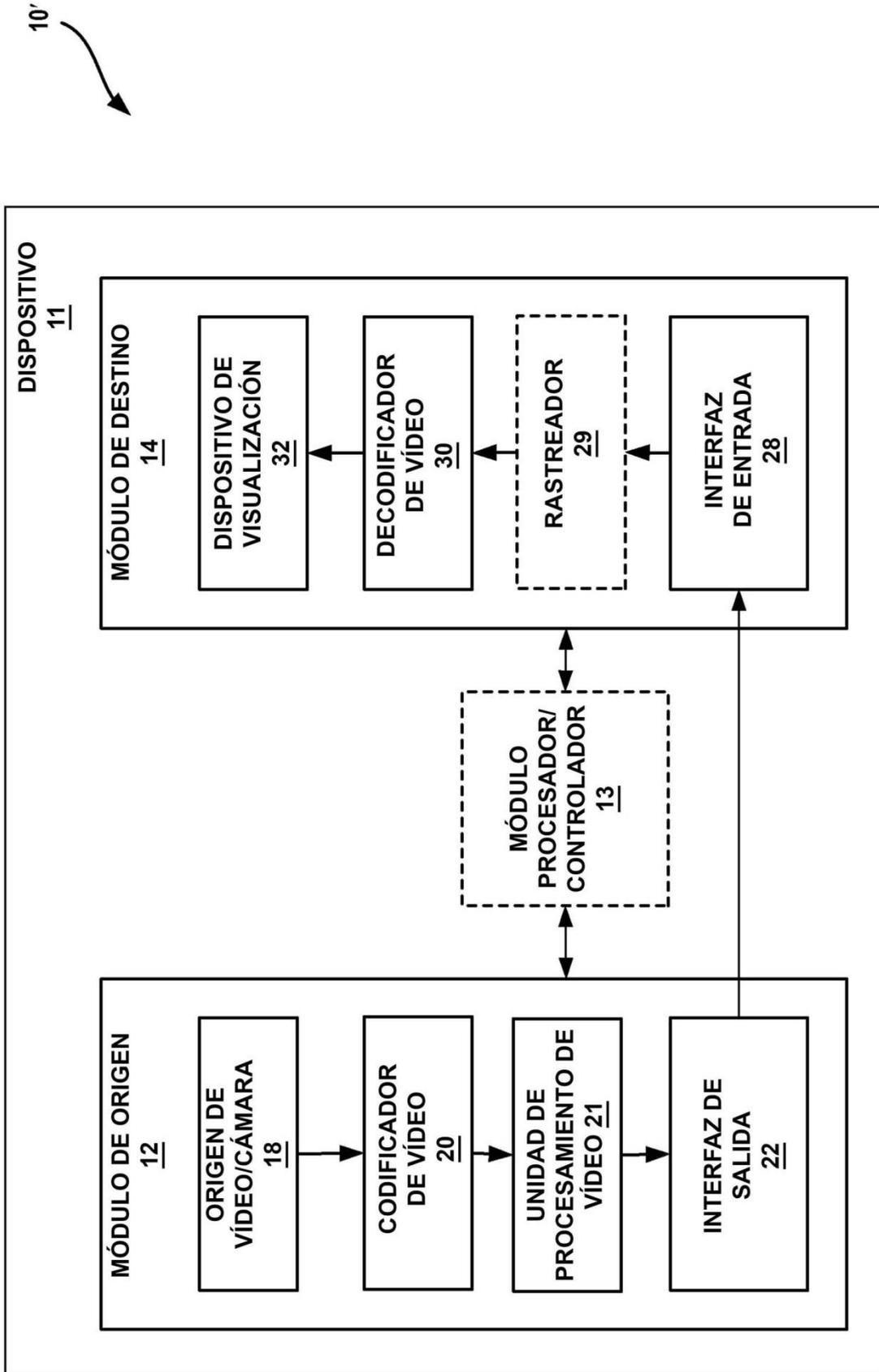


FIG. 1B

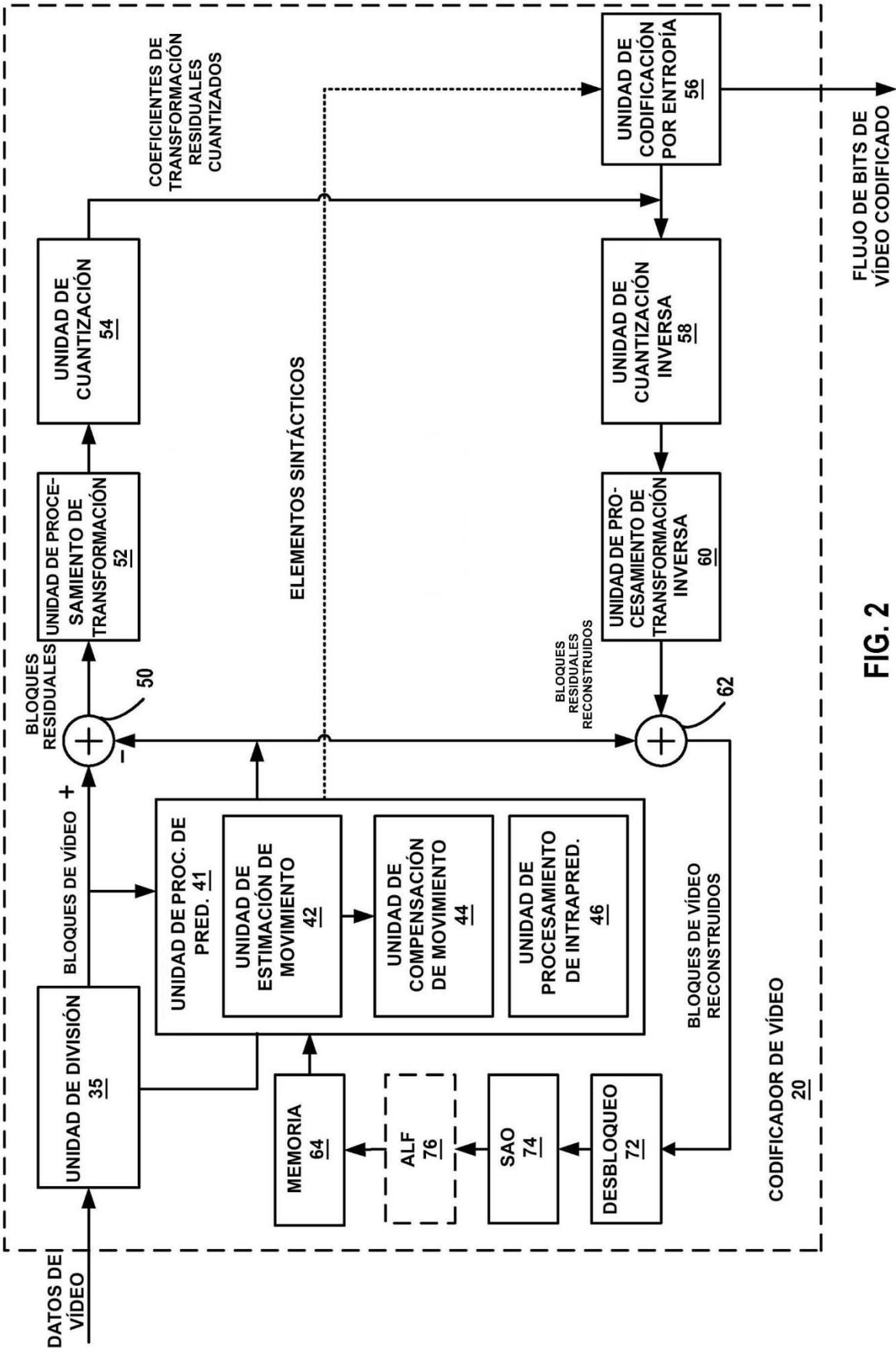


FIG. 2

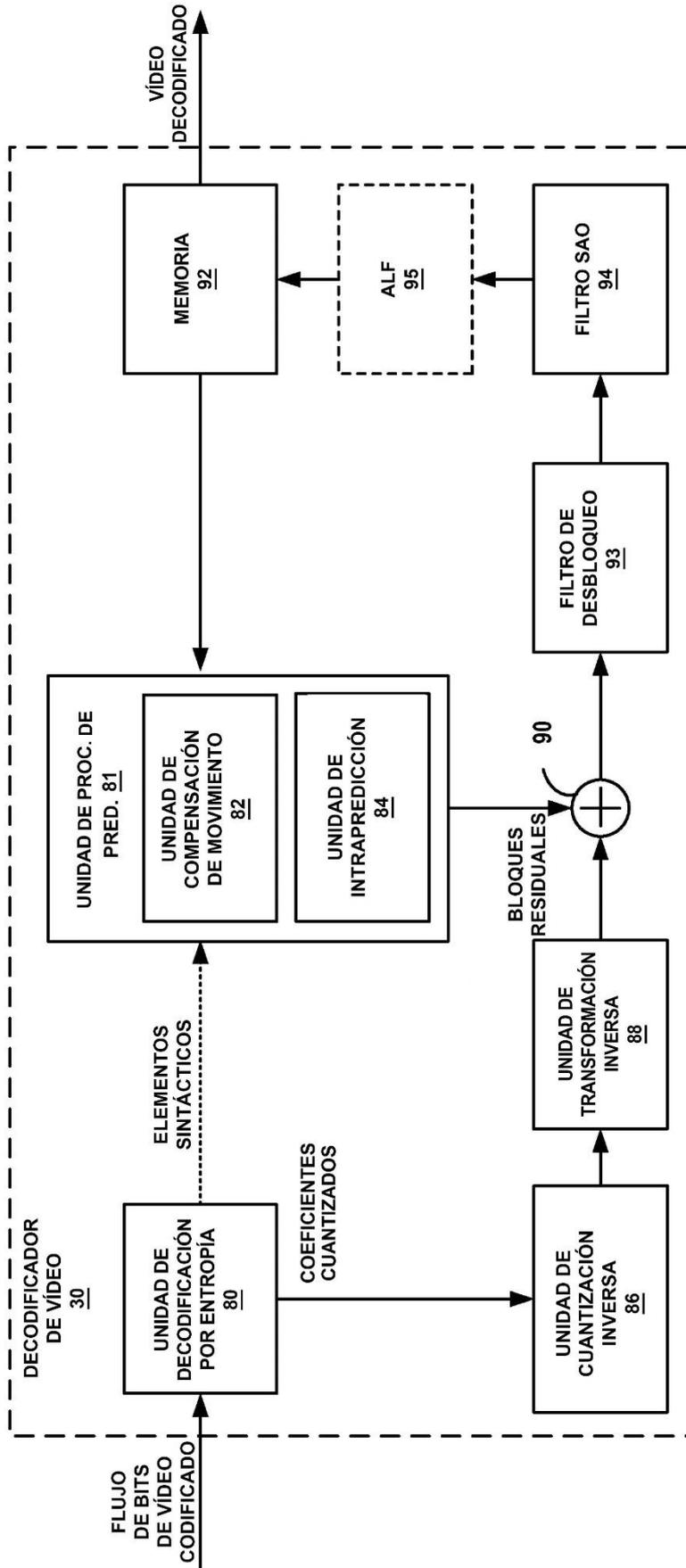


FIG. 3

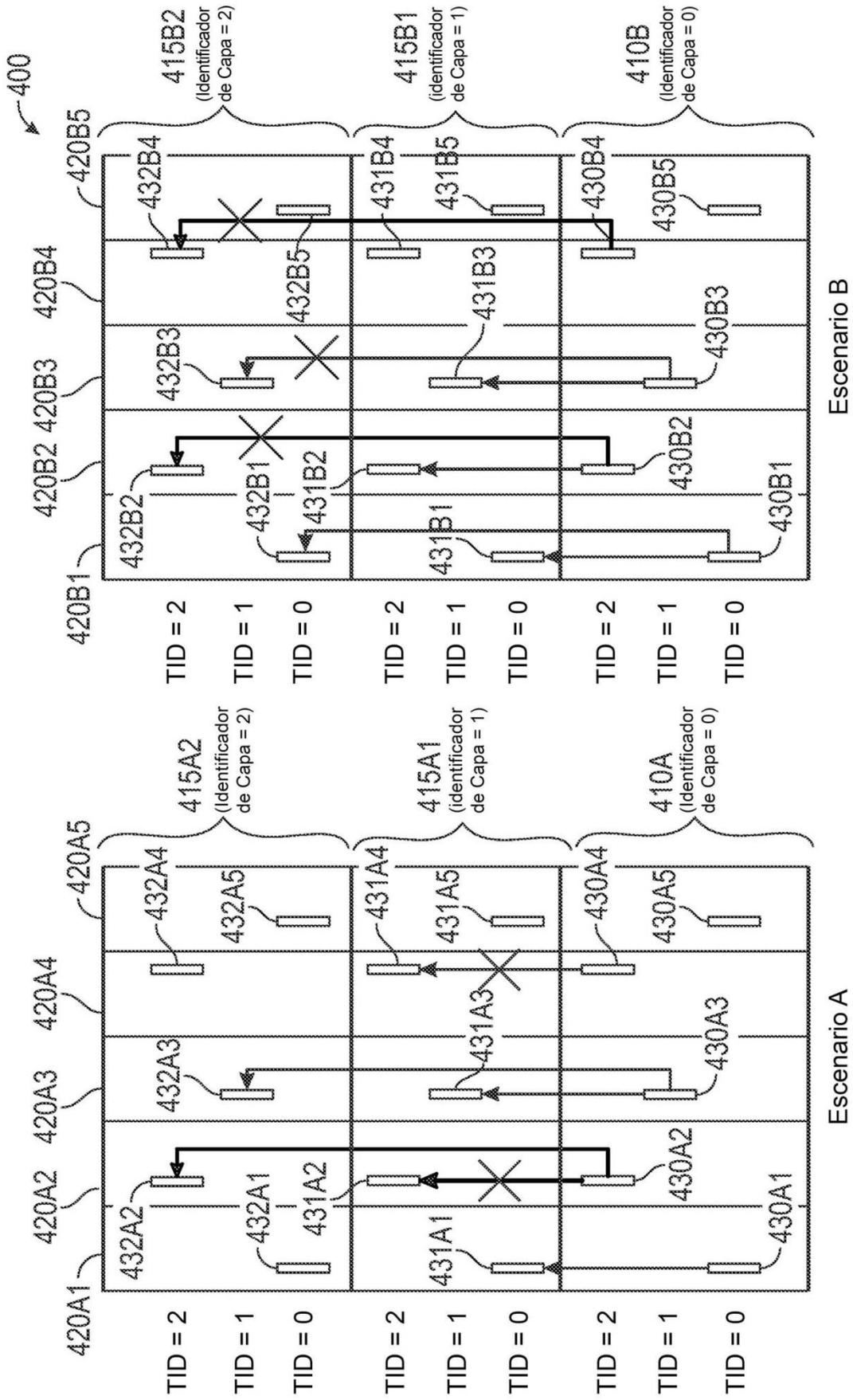


FIG. 4

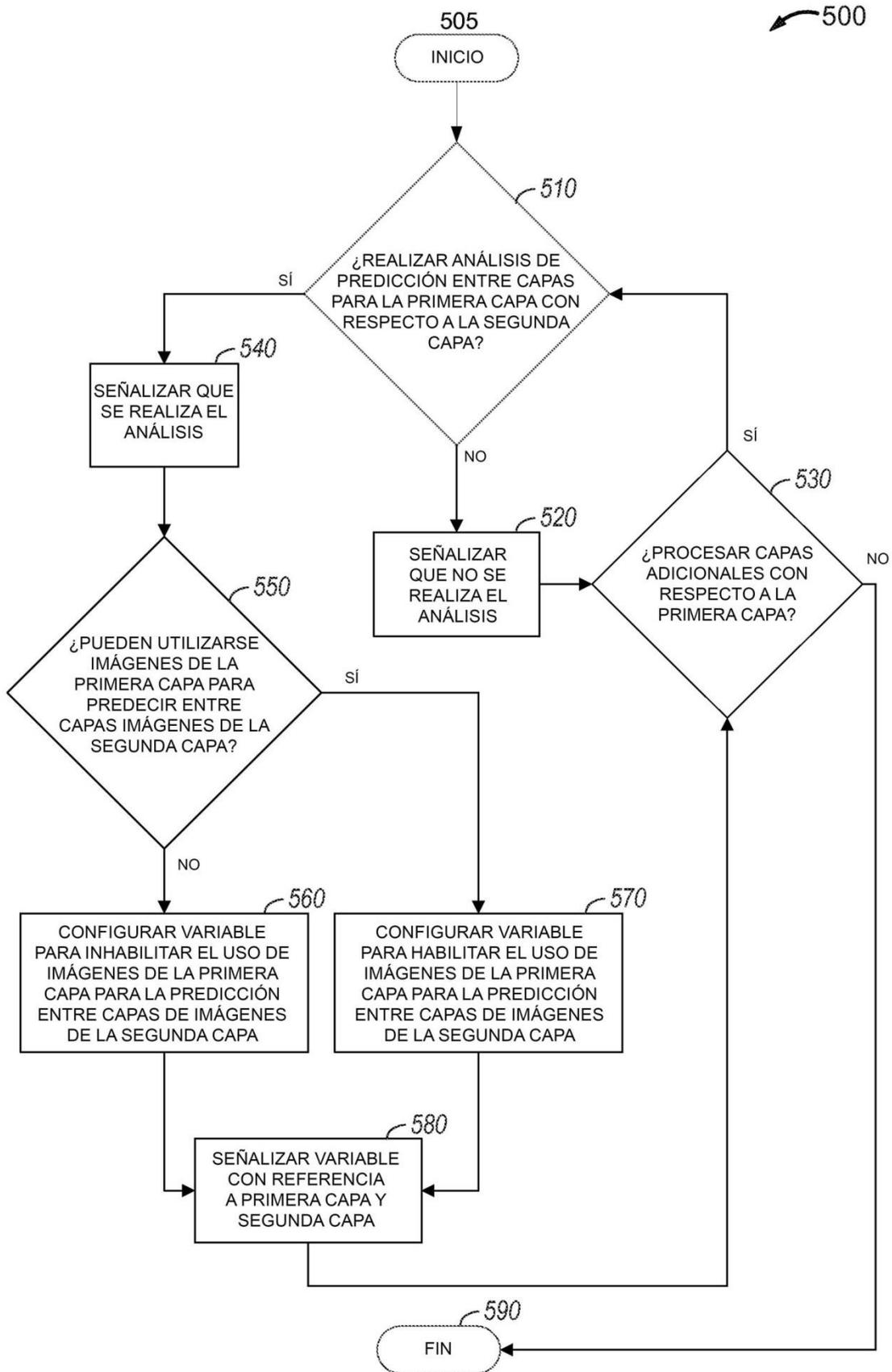


FIG. 5

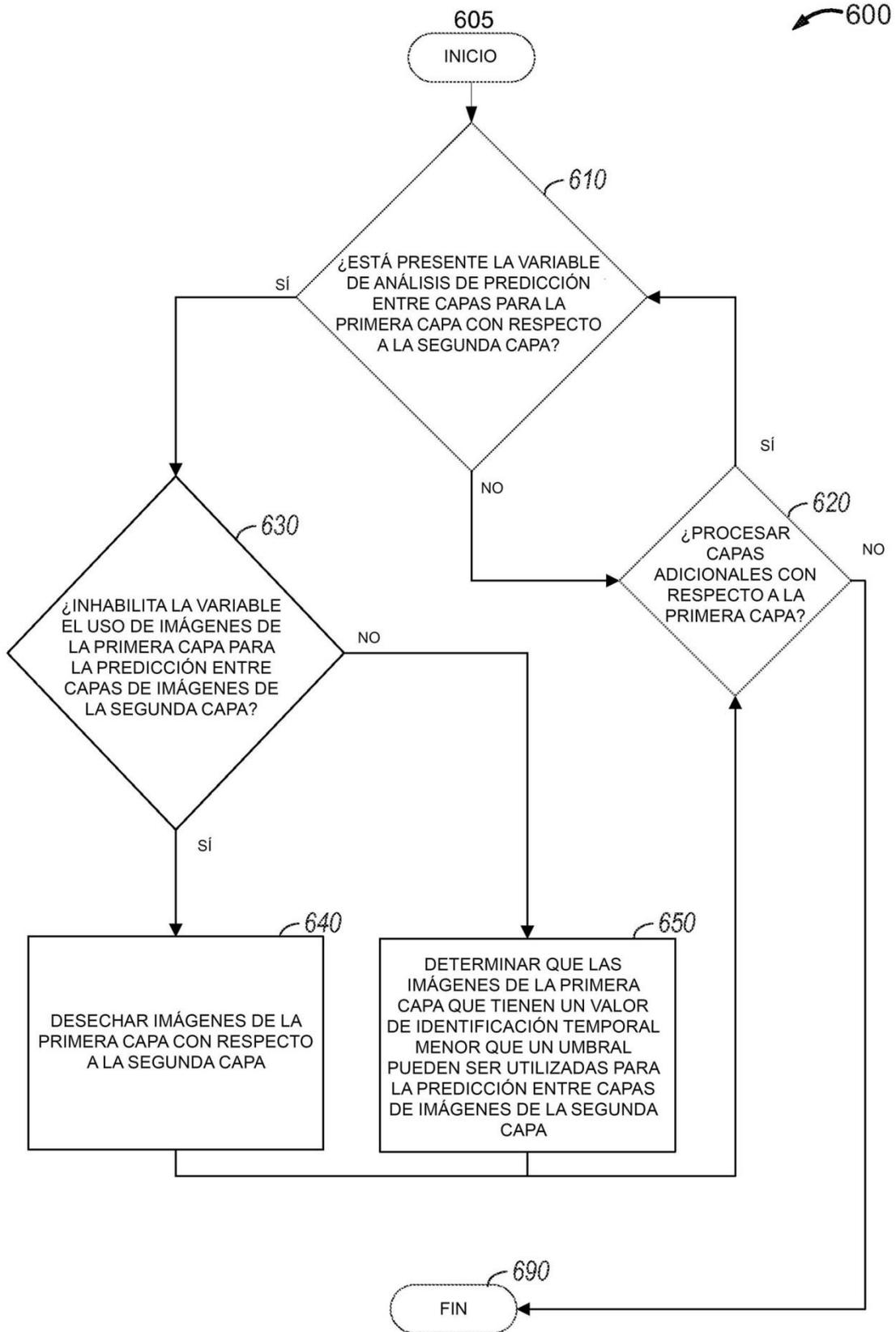


FIG. 6

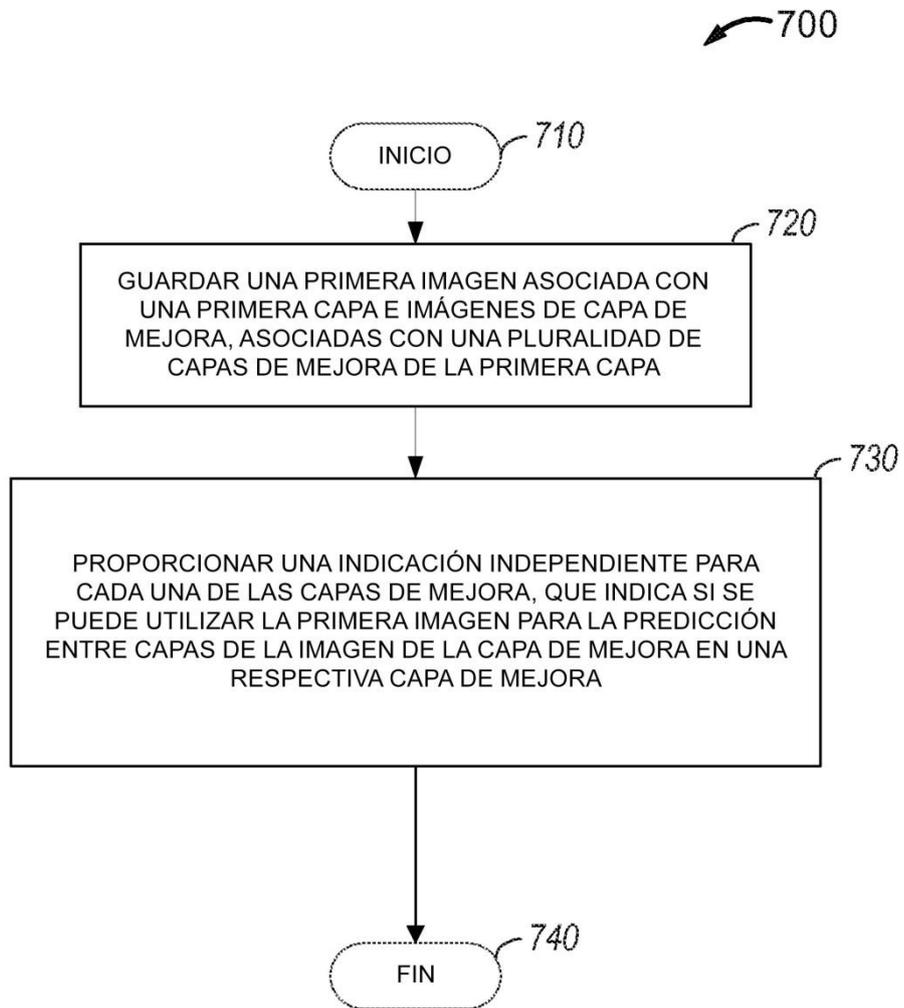


FIG. 7