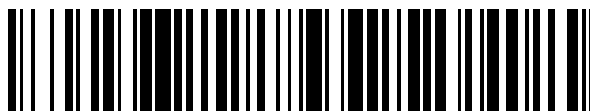


19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 711 954**

51 Int. Cl.:

H04N 19/30 (2014.01)

H04N 19/33 (2014.01)

H04N 19/463 (2014.01)

H04N 19/50 (2014.01)

H04N 19/59 (2014.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

86 Fecha de presentación y número de la solicitud internacional: **04.12.2013 PCT/US2013/073015**

87 Fecha y número de publicación internacional: **19.06.2014 WO14093079**

96 Fecha de presentación y número de la solicitud europea: **04.12.2013 E 13811078 (8)**

97 Fecha y número de publicación de la concesión europea: **28.11.2018 EP 2932718**

54 Título: **Dispositivo y procedimiento para codificación escalable de información de vídeo basándose en codificación de vídeo de alta eficiencia**

30 Prioridad:

12.12.2012 US 201261736481 P

20.02.2013 US 201361767183 P

09.10.2013 US 201314049649

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

08.05.2019

73 Titular/es:

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

72 Inventor/es:

**SEREGIN, VADIM;
CHEN, YING y
CHEN, JIANLE**

74 Agente/Representante:

FORTEA LAGUNA, Juan José

ES 2 711 954 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

Dispositivo y procedimiento para codificación escalable de información de vídeo basándose en codificación de vídeo de alta eficiencia

5

CAMPO TÉCNICO

[0001] Esta divulgación está relacionada con el campo de la codificación y compresión de vídeo, codificación escalable basada en codificación de vídeo de alta eficiencia (HEVC) y codificación de múltiples visualizaciones y 3DV basada en HEVC.

10

ANTECEDENTES

[0002] Las capacidades de vídeo digital pueden incorporarse a una amplia gama de dispositivos, incluidos televisores digitales, sistemas de difusión directa digital, sistemas de difusión inalámbrica, asistentes digitales personales (PDA), ordenadores portátiles o de sobremesa, cámaras digitales, dispositivos de grabación digitales, reproductores de medios digitales, dispositivos de videojuegos, consolas de videojuegos, teléfonos celulares o de radio por satélite, dispositivos de videoconferencia y similares. Los dispositivos de vídeo digital implementan técnicas de compresión de vídeo, tales como las descritas en las normas definidas por MPEG-2, MPEG-4, ITU-T H.263, ITU-T H.264/MPEG-4, parte 10, codificación de vídeo avanzada (AVC), la norma de codificación de vídeo de alta eficiencia (HEVC) actualmente en desarrollo y las extensiones de dichas normas. Los dispositivos de vídeo pueden transmitir, recibir, codificar, descodificar y/o almacenar información de vídeo digital más eficazmente, implementando tales técnicas de codificación de vídeo.

15

20

[0003] Las técnicas de compresión de vídeo realizan la predicción espacial (dentro de las imágenes) y/o la predicción temporal (entre imágenes) para reducir o eliminar la redundancia intrínseca a las secuencias de vídeo. Para la codificación de vídeo basada en bloques, un fragmento de vídeo (por ejemplo, una trama de vídeo, una parte de una trama de vídeo, etc.) puede dividirse en bloques de vídeo, que también pueden denominarse bloques arbolados, unidades de codificación (CU) y/o nodos de codificación. Los bloques de vídeo en un fragmento intra-codificado (I) de una imagen se codifican utilizando predicción espacial con respecto a muestras de referencia en bloques contiguos en la misma imagen. Los bloques de vídeo en un fragmento inter-codificado (P o B) de una imagen pueden usar la predicción espacial con respecto a muestras de referencia en bloques contiguos en la misma imagen o la predicción temporal con respecto a muestras de referencia en otras imágenes de referencia. Las imágenes pueden denominarse tramas, y las imágenes de referencia pueden denominarse tramas de referencia.

25

30

35

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

40

45

[0005] El documento US 2008/0165850 divulga un procedimiento en el que una unidad de codificación puede realizar una predicción entre capas para la capa de mejora con respecto a una capa base, para bloques de la capa de mejora que se superponen total o parcialmente con bloques de la capa base.

50

SUMARIO

[0006] La codificación de vídeo escalable (SVC) se refiere a la codificación de vídeo en la que se utiliza una capa base (BL), a veces denominada capa de referencia (RL), y una o más capas de mejora escalable (EL). Para SVC, la capa base puede transportar datos de vídeo con un nivel de calidad base. Las una o más capas de mejora pueden transportar datos de vídeo adicionales para soportar niveles SNR espaciales, temporales y/o de señal a ruido más altos. Las capas de mejora pueden definirse en relación con una capa previamente codificada. Por ejemplo, una capa inferior puede servir como una BL, mientras que una capa superior puede servir como una EL. Las capas medias pueden servir como EL o RL, o como ambas. Por ejemplo, una capa en el medio puede ser una EL para las capas debajo de ella, como la capa base o cualquier capa de mejora intermedia, y al mismo tiempo servir como RL para una o más capas de mejora sobre ella. De manera similar, en la extensión de múltiples visualizaciones o 3D de la norma HEVC, puede haber múltiples visualizaciones, y la información de una visualización puede utilizarse para codificar (por ejemplo, codificar o descodificar) la información de otra visualización (por ejemplo, estimación de movimiento, predicción de vector de movimiento y/u otras redundancias).

55

60

65

[0007] En la extensión HEVC, se puede predecir un bloque actual en la capa de mejora u otra visualización usando un bloque correspondiente en la capa o visualización base. Por ejemplo, los elementos sintácticos, la información de movimiento (por ejemplo, vectores de movimiento) o la información de modo (por ejemplo, modo intra) del bloque actual pueden basarse en el bloque correspondiente en la capa base. Por ejemplo, un vector de movimiento de capa base (MV) se puede utilizar como candidato en el modo de combinación de capa de mejora/listas de candidatos AMVP. Una lista de candidatos es una matriz de predictores de vectores de movimiento espacial y temporal que debe utilizar el codificador para predecir el bloque actual. Por ejemplo, en lugar de codificar y comunicar el propio vector de movimiento, un codificador de vídeo puede codificar y comunicar una diferencia de vector de movimiento (MVD) en relación con un vector de movimiento conocido (o que puede conocerse). En H.264/AVC, el vector de movimiento conocido, que se puede usar con el MVD para definir el vector de movimiento actual, se puede definir mediante un llamado predictor de vector de movimiento (MVP), que se obtiene como la media de los vectores de movimiento Asociado a bloques contiguos. Sin embargo, las técnicas MVP más avanzadas pueden permitir que el codificador de vídeo seleccione el contiguo desde el cual definir el MVP.

[0008] Sin embargo, el bloque correspondiente de la capa base se puede ubicar fuera de la trama de la capa base, por ejemplo, cuando el tamaño de trama codificada es más grande que el tamaño de trama real que puede mostrarse en un dispositivo, o cuando la capa base está codificada con otra norma de vídeo, como AVC o MPEG2. En tales situaciones, la información del bloque correspondiente no puede usarse para predecir el bloque actual en la capa o visualización de mejora porque no hay información disponible para el bloque correspondiente. En otras palabras, las partes de la capa o visualización de mejora que no corresponden a áreas dentro del tamaño de trama real o codificada de la imagen de la capa base no pueden codificarse utilizando la información de la capa o visualización base, y por lo tanto la eficiencia de la codificación puede reducirse. Al permitir que dichas partes de la capa o visualización de mejora sean predichas utilizando la información obtenida a partir de la capa o visualización base, las técnicas descritas en esta divulgación pueden mejorar la eficiencia de codificación y/o reducir la complejidad computacional asociada con un procedimiento de codificación de datos de vídeo.

[0009] 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. La invención se expone en el conjunto de reivindicaciones adjuntas; los ejemplos adicionales denominados "modos de realización" en la descripción son ejemplos ilustrativos.

[0010] En un modo de realización, un aparato configurado para codificar información de vídeo incluye una unidad de memoria y un procesador en comunicación con la unidad de memoria. La unidad de memoria está configurada para almacenar información de vídeo asociada con una capa de mejora que tiene un primer bloque y una capa base que tiene un segundo bloque, correspondiendo el segundo bloque en la capa base al primer bloque en la capa de mejora. El procesador está configurado para predecir, mediante predicción entre capas, el primer bloque en la capa de mejora basándose en la información obtenida a partir del segundo bloque en la capa base. Al menos una parte del segundo bloque está ubicada fuera de una región de referencia de la capa base, estando disponible la región de referencia para el uso para la predicción entre capas del primer bloque. La información obtenida a partir del segundo bloque en la capa base puede incluir información correspondiente a un bloque contiguo más cercano al segundo bloque. El bloque contiguo puede ubicarse recortando la información de coordenadas del segundo bloque a un rango de la región de referencia (por ejemplo, tanto horizontal como verticalmente) de la capa base. De forma alternativa, la información de coordenadas del primer bloque puede recortarse primero a un rango de una imagen actual descodificada o recortada de la capa base antes de realizar un proceso de asignación de posición para determinar el segundo bloque de la capa base. En el caso de que se determine que el segundo bloque (o una parte del mismo) está fuera de la región de referencia de la capa base, la predicción entre capas de los elementos sintácticos y la información residual se pueden desactivar para el primer bloque, o de forma alternativa, los valores sintácticos predeterminados pueden usarse para la predicción entre capas del primer bloque. En otro modo de realización, los valores sintácticos del bloque contiguo se pueden usar para la predicción entre capas del primer bloque. El tamaño de la unidad de codificación (SCU) más pequeña de la capa de mejora puede configurarse para que sea diferente de la SCU de la capa base. El tamaño de relleno de límite de cada límite de imagen de la capa de mejora puede configurarse para ser diferente del tamaño de relleno de límite de cada límite de imagen correspondiente de la capa base.

[0011] En otro modo de realización, un procedimiento para codificar información de vídeo incluye: recibir información de vídeo asociada con una capa base y una capa de mejora; y predecir, mediante la predicción entre capas, un primer bloque en la capa de mejora basado en la información obtenida a partir de un segundo bloque en la capa base, correspondiendo el segundo bloque en la capa base al primer bloque en la capa de mejora; y predecir, mediante la predicción entre capas, el primer bloque en la capa de mejora basándose en la información obtenida a partir del segundo bloque en la capa base. Al menos una parte del segundo bloque está ubicada fuera de una región de referencia de la capa base, estando disponible la región de referencia para el uso para la predicción entre capas del primer bloque. La información obtenida a partir del segundo bloque en la capa base puede incluir información correspondiente a un bloque contiguo más cercano al segundo bloque. El bloque contiguo puede ubicarse recortando la información de coordenadas del segundo bloque a un rango de la región de referencia (por ejemplo, tanto horizontal como verticalmente) de la capa base. De forma alternativa, la información de coordenadas del primer bloque puede recortarse primero a un rango de una imagen actual descodificada o recortada de la capa

base antes de realizar un proceso de asignación de posición para determinar el segundo bloque de la capa base. En el caso de que se determine que el segundo bloque (o una parte del mismo) está fuera de la región de referencia de la capa base, la predicción entre capas de los elementos sintácticos y la información residual se pueden desactivar para el primer bloque, o de forma alternativa, los valores sintácticos predeterminados pueden usarse para la predicción entre capas del primer bloque. En otro modo de realización, los valores sintácticos del bloque contiguo se pueden usar para la predicción entre capas del primer bloque. El tamaño de la unidad de codificación (SCU) más pequeña de la capa de mejora puede configurarse para que sea diferente de la SCU de la capa base. El tamaño de relleno de límite de cada límite de imagen de la capa de mejora puede configurarse para ser diferente del tamaño de relleno de límite de cada límite de imagen correspondiente de la capa base.

[0012] En otro modo de realización, un procedimiento para descodificar información de vídeo incluye: recibir elementos sintácticos extraídos de un flujo de bits de vídeo codificado, en el que los elementos sintácticos comprenden información de vídeo asociada con una capa de mejora que tiene un primer bloque y una capa base que tiene un segundo bloque, correspondiendo el segundo bloque al primer bloque; y predecir, mediante la predicción entre capas, el primer bloque en la capa de mejora basándose en la información obtenida a partir del segundo bloque en la capa base, en el que al menos una parte del segundo bloque está ubicada fuera de una región de referencia de la capa base, estando la región de referencia disponible para su uso para la predicción entre capas del primer bloque. La información obtenida a partir del segundo bloque en la capa base puede incluir información correspondiente a un bloque contiguo más cercano al segundo bloque. El bloque contiguo puede ubicarse recortando la información de coordenadas del segundo bloque a un rango de la región de referencia (por ejemplo, tanto horizontal como verticalmente) de la capa base. De forma alternativa, la información de coordenadas del primer bloque puede recortarse primero a un rango de una imagen actual descodificada o recortada de la capa base antes de realizar un proceso de asignación de posición para determinar el segundo bloque de la capa base. En el caso de que se determine que el segundo bloque (o una parte del mismo) está fuera de la región de referencia de la capa base, la predicción entre capas de los elementos sintácticos y la información residual se pueden desactivar para el primer bloque, o de forma alternativa, los valores sintácticos predeterminados pueden usarse para la predicción entre capas del primer bloque. En otro modo de realización, los valores sintácticos del bloque contiguo se pueden usar para la predicción entre capas del primer bloque. El tamaño de la unidad de codificación (SCU) más pequeña de la capa de mejora puede configurarse para que sea diferente de la SCU de la capa base. El tamaño de relleno de límite de cada límite de imagen de la capa de mejora puede configurarse para ser diferente del tamaño de relleno de límite de cada límite de imagen correspondiente de la capa base.

[0013] En otro modo de realización, un medio legible por ordenador no transitorio incluye código que, cuando se ejecuta, hace que un aparato: almacene información de vídeo asociada con una capa de mejora que tiene un primer bloque y una capa base que tiene un segundo bloque, correspondiendo el segundo bloque al primer bloque y, predecir, mediante predicción entre capas, el primer bloque en la capa de mejora basándose en la información obtenida a partir del segundo bloque en la capa base, en el que al menos una parte del segundo bloque está ubicada fuera de una región de referencia de la capa base, estando la región de referencia disponible para su uso para la predicción entre capas del primer bloque.

[0014] En otro modo de realización, un dispositivo de codificación de vídeo configurado para codificar información de vídeo incluye: medios para almacenar información de vídeo asociada con una capa de mejora que tiene un primer bloque y una capa base que tiene un segundo bloque, correspondiendo el segundo bloque al primer bloque; y medios para predecir, mediante predicción entre capas, el primer bloque en la capa de mejora basándose en la información obtenida a partir del segundo bloque en la capa base, en el que al menos una parte del segundo bloque está ubicada fuera de una región de referencia de la capa base, estando la región de referencia disponible para su uso para la predicción entre capas del primer bloque.

BREVE DESCRIPCIÓN DE LOS DIBUJOS

[0015]

La **FIG. 1** es un diagrama de bloques que ilustra un sistema de codificación y descodificación de vídeo de ejemplo que puede utilizar técnicas de acuerdo con los aspectos descritos en esta divulgación.

La **FIG. 2** es un diagrama de bloques que ilustra un ejemplo de un codificador de vídeo que puede implementar técnicas de acuerdo con los aspectos descritos en esta divulgación.

La **FIG. 3** es un diagrama de bloques que ilustra un ejemplo de un descodificador de vídeo que puede implementar técnicas de acuerdo con aspectos descritos en esta divulgación.

La **FIG. 4** un diagrama que ilustra las diversas tramas en la capa base y la capa de mejora.

La **FIG. 5** es un diagrama de flujo que ilustra un procedimiento de codificación de información de vídeo, de acuerdo con un modo de realización de la presente divulgación.

DESCRIPCIÓN DETALLADA

5 **[0016]** Ciertos modos de realización descritos en el presente documento se refieren a la predicción entre capas para la codificación de vídeo escalable en el contexto de los códecs de vídeo avanzados, como HEVC (codificación de vídeo de alta eficiencia). Más específicamente, la presente divulgación se refiere a sistemas y procedimientos para mejorar el rendimiento de la predicción entre capas en la extensión de codificación de vídeo escalable (SVC) de HEVC.

10 **[0017]** En la descripción a continuación, se describen técnicas H.264/AVC relacionadas con ciertos modos de realización; también se analizan la norma HEVC y las técnicas relacionadas. Aunque ciertos modos de realización se describen en el presente documento en el contexto de las normas HEVC y/o H.264, un experto en la técnica puede apreciar que los sistemas y procedimientos divulgados en el presente documento pueden aplicarse a cualquier norma de codificación de vídeo adecuado. Por ejemplo, los modos de realización divulgados en el presente documento pueden ser aplicables a una o más de las siguientes normas: ITU-T H.261, MPEG-1 Visual de la ISO/IEC, ITU-T H.262 o MPEG-2 Visual de la ISO/IEC, ITU-T H.263, MPEG-4 Visual de la ISO/IEC e ITU-T H.264 (también conocida como AVC del MPEG-4 de la ISO/IEC), incluyendo sus ampliaciones de codificación de vídeo escalable (SVC) y de codificación de vídeo de múltiples visualizaciones (MVC).

20 **[0018]** Para fines de ilustración solamente, ciertos modos de realización divulgados en el presente documento se describen con ejemplos que incluyen solo dos capas (por ejemplo, una capa de nivel inferior tal como la capa base y una capa de nivel superior tal como la capa de mejora). Debe entenderse que dichos ejemplos pueden ser aplicables a configuraciones que incluyen múltiples capas base y/o de mejora. Además, para facilitar la explicación, la siguiente divulgación incluye los términos "tramas" o "bloques" con referencia a ciertos modos de realización. Sin embargo, estos términos no pretenden ser limitativos. Por ejemplo, las técnicas descritas a continuación se pueden usar con cualquier unidad de vídeo adecuada, como bloques (por ejemplo, CU, PU, TU, macrobloques, etc.), fragmentos, tramas, etc.

30 **[0019]** HEVC, en general, sigue el marco de las normas de codificación de vídeo anteriores en muchos aspectos. La unidad de predicción en HEVC es diferente de la de ciertos estándares de codificación de vídeo anteriores (por ejemplo, macrobloque). De hecho, el concepto de macrobloque no existe en HEVC como se entiende en ciertas normas anteriores de codificación de vídeo. Un macrobloque se reemplaza por una estructura jerárquica basada en un esquema de árbol cuaternario, lo cual puede proporcionar una gran flexibilidad, entre otros beneficios posibles. Por ejemplo, dentro del esquema HEVC, se definen tres tipos de bloques, es decir, unidad de codificación (CU), unidad de predicción (PU) y unidad de transformada (TU). La CU puede referirse a la unidad básica de división de la región. La CU puede considerarse análoga al concepto de macrobloque, pero no limita el tamaño máximo y puede permitir la división recursiva en cuatro CU de igual tamaño para mejorar la adaptabilidad del contenido. La PU puede considerarse la unidad básica de predicción inter/intra, y puede contener múltiples particiones de forma arbitraria en una única PU para codificar de manera efectiva patrones de imagen irregulares. La TU se puede considerar como la unidad básica de transformada. Se puede definir independientemente de la PU; sin embargo, su tamaño puede estar limitado a la CU a la que pertenece la TU. Esta separación de la estructura de bloques en tres conceptos diferentes puede permitir optimizar cada uno de ellos de acuerdo con su función, lo cual puede dar como resultado una mejora de la eficacia de codificación.

Normas de codificación de vídeo

45 **[0020]** Una imagen digital, como una imagen de vídeo, una imagen de TV, una imagen fija o una imagen generada por una grabadora de vídeo o un ordenador puede consistir en píxeles dispuestos en líneas horizontales y verticales. La cantidad de píxeles en una sola imagen suele ser de decenas de miles. Cada píxel contiene típicamente información de luminancia y crominancia. Sin compresión, la cantidad de información que se transmitirá desde un codificador de imágenes a un descodificador de imágenes es tan grande que haría imposible la transmisión de imágenes en tiempo real. Para reducir la cantidad de información a transmitir, se han desarrollado una serie de procedimientos de compresión diferentes, tales como las normas JPEG, MPEG y H.263.

55 **[0021]** Entre las normas de codificación de vídeo se incluyen ITU-T H.261, MPEG-1 Visual de la ISO/IEC, ITU-T H.262 o MPEG-2 Visual de la ISO/IEC, ITU-T H.263, MPEG-4 Visual de la ISO/IEC e ITU-T H.264 (también conocida como AVC del MPEG-4 de la ISO/IEC), incluyendo sus ampliaciones de codificación de vídeo escalable (SVC) y de codificación de vídeo de múltiples visualizaciones (MVC).

60 **[0022]** Además, existe una nueva norma de codificación de vídeo, concretamente la codificación de vídeo de alta eficiencia (HEVC) que está siendo desarrollada por el Equipo de Colaboración Conjunta en Codificación de Vídeo (JCT-VC) del Grupo de Expertos en Codificación de Vídeo (VCEG) de ITU-T y el Grupo de Expertos en Imágenes en Movimiento (MPEG) de ISO/IEC. Un borrador reciente de la HEVC está disponible en http://phenix.it-sudparis.eu/ict/doc_end_user/documents/12_Geneva/wg11/JCTVC-L1003-v34.zip, desde el 9 de agosto de 2013. La cita completa del borrador 10 de HEVC es el documento JCTVC-L1003, de Bross et al., titulado "High Efficiency Video Coding (HEVC) Text Specification Draft 10 [Especificación textual de la Codificación de Vídeo de Alta

Eficacia (HEVC), Borrador 10]", Equipo de colaboración conjunta en codificación de vídeo (JCT-VC) de ITU-T SG16 WP3 e ISO/IEC JTC1/SC29/WG11, 12.ª reunión: Ginebra, Suiza, del 14 de enero de 2013 al 23 de enero de 2013.

[0023] En lo sucesivo se describen de forma más detallada diversos aspectos de los sistemas, aparatos y procedimientos novedosos, con referencia a los dibujos adjuntos. Sin embargo, esta divulgación se puede realizar de muchas formas diferentes y no se debería interpretar que está limitada a cualquier estructura o función específica presentada a lo largo de esta divulgación. En su lugar, estos aspectos se proporcionan para que esta divulgación sea exhaustiva y completa, y transmita por completo el alcance de la divulgación a los expertos en la materia. Basándose en las enseñanzas en el presente documento, un experto en la materia debería apreciar que el alcance de la divulgación está concebido para abarcar cualquier aspecto de los sistemas, aparatos y procedimientos novedosos divulgados en el presente documento, ya sea implementados de forma independiente de, o en combinación con, cualquier otro aspecto de la invención. Por ejemplo, un aparato se puede implementar o un procedimiento se puede llevar a la práctica usando cualquier número de los aspectos expuestos en el presente documento.

[0024] Aunque se mencionan algunos beneficios y ventajas de los aspectos preferentes, el alcance de la divulgación no está concebido para limitarse a beneficios, usos u objetivos particulares. En cambio, los aspectos de la divulgación pretenden ser ampliamente aplicables a diferentes tecnologías inalámbricas, configuraciones de sistema, redes y protocolos de transmisión, algunos de los cuales se ilustran a modo de ejemplo en las figuras y en la siguiente descripción de los aspectos preferidos. La descripción detallada y los dibujos son meramente ilustrativos de la divulgación en lugar de ser limitativos, estando definido el alcance de la divulgación por las reivindicaciones adjuntas y equivalentes de las mismas.

[0025] Los dibujos adjuntos ilustran ejemplos. Los elementos indicados mediante números de referencia en los dibujos adjuntos corresponden a elementos indicados mediante números de referencia similares en la siguiente descripción.

Sistema de codificación de vídeo

[0026] La **FIG. 1A** es un diagrama de bloques que ilustra un sistema de codificación de vídeo 10 de ejemplo que puede utilizar técnicas de acuerdo con aspectos descritos en esta divulgación. Tal como se utiliza en el presente documento, el término "codificador de vídeo" se refiere genéricamente tanto a codificadores de vídeo como a descodificadores de vídeo. En esta divulgación, los términos "codificación de vídeo" o "codificación" pueden referirse genéricamente a la codificación de vídeo y a la descodificación de vídeo.

[0027] Como se muestra en la **FIG. 1**, el sistema de codificación de vídeo 10 incluye un dispositivo de origen 12 y un dispositivo de destino 14. El dispositivo de origen 12 genera datos de vídeo codificados. El dispositivo de destino 14 puede descodificar datos de vídeo codificados, generados por el dispositivo de origen 12. El dispositivo de origen 12 y el dispositivo de destino 14 pueden comprender una amplia gama de dispositivos, incluyendo ordenadores de sobremesa, notebooks (por ejemplo, portátiles, etc.), ordenadores tipo tablet, descodificadores, teléfonos portátiles como los denominados teléfonos "inteligentes", los denominados paneles "inteligentes", televisores, cámaras, dispositivos de visualización, reproductores de medios digitales, consolas de videojuegos, ordenadores de coche o similares. En algunos ejemplos, el dispositivo de origen 12 y el dispositivo de destino 14 pueden estar equipados para la comunicación inalámbrica.

[0028] El dispositivo de destino 14 puede recibir datos de vídeo codificados desde el dispositivo de origen 12 a través de un canal 16. El canal 16 puede comprender cualquier tipo de medio o dispositivo capaz de desplazar los datos de vídeo codificados desde el dispositivo de origen 12 al dispositivo de destino 14. En un ejemplo, el canal 16 puede comprender un medio de comunicación que permite al dispositivo de origen 12 transmitir datos de vídeo codificados directamente al dispositivo de destino 14 en tiempo real. En este ejemplo, el dispositivo de origen 12 puede modular los datos de vídeo codificados de acuerdo con una norma de comunicación, tal como un protocolo de comunicación inalámbrica, y puede transmitir los datos de vídeo modulados al dispositivo de destino 14. El medio de comunicación puede comprender un medio de comunicación inalámbrica o alámbrica, tal como un espectro de radiofrecuencia (RF) o una o más líneas de transmisión física. El medio de comunicación puede formar parte de una red basada en paquetes, tal como una red de área local, una red de área amplia o una red global tal como Internet. El medio de comunicación puede incluir routers, conmutadores, estaciones base u otros equipos que faciliten la comunicación desde el dispositivo de origen 12 al dispositivo de destino 14.

[0029] En otro ejemplo, el canal 16 puede corresponder a un medio de almacenamiento que almacena los datos de vídeo codificados generados por el dispositivo de origen 12. En este ejemplo, el dispositivo de destino 14 puede acceder al medio de almacenamiento a través del acceso al disco o del acceso a la tarjeta. El dispositivo de almacenamiento puede incluir una variedad de medios de almacenamiento de datos de acceso local, tales como discos Blu-ray, DVD, CD-ROM, memoria flash u otros medios adecuados de almacenamiento digital para almacenar datos de vídeo codificados. En un ejemplo adicional, el canal 16 puede incluir un servidor de archivos u otro dispositivo de almacenamiento intermedio que almacene el vídeo codificado generado por el dispositivo de origen 12. En este ejemplo, el dispositivo de destino 14 puede acceder a datos de vídeo codificados almacenados

en el servidor de ficheros o en otro dispositivo de almacenamiento intermedio mediante transmisión continua o descarga. El servidor de ficheros puede ser un tipo de servidor capaz de almacenar datos de vídeo codificados y transmitir los datos de vídeo codificados al dispositivo de destino 14. Entre los ejemplos de servidores de archivos se incluyen servidores web (por ejemplo, para un sitio web, etc.), servidores FTP, dispositivos de almacenamiento conectados a la red (NAS) y unidades de disco local. El dispositivo de destino 14 puede acceder a los datos de vídeo codificados a través de cualquier conexión de datos estándar, incluyendo una conexión a Internet. Entre los ejemplos de tipos de conexiones de datos pueden incluirse canales inalámbricos (por ejemplo, conexiones WiFi, etc.), conexiones alámbricas (por ejemplo, DSL, módem de cable, etc.) o combinaciones de ambos que sean adecuadas para acceder a datos de vídeo codificados almacenados en un servidor de ficheros. La transmisión de datos de vídeo codificados desde el servidor de ficheros puede ser una transmisión continua, una transmisión de descarga o una combinación de ambas.

[0030] Las técnicas de esta divulgación no están limitadas a aplicaciones o a configuraciones inalámbricas. Las técnicas pueden aplicarse a la codificación de vídeo, para soportar cualquiera de una diversidad de aplicaciones multimedia, tales como radiodifusiones de televisión por el aire, transmisiones de televisión por cable, transmisiones de televisión por satélite, transmisiones de vídeo en continuo, por ejemplo, a través de Internet (por ejemplo, transmisión adaptable dinámica por HTTP (DASH), etc.), codificación de vídeo digital para su almacenamiento en un medio de almacenamiento de datos, descodificación de vídeo digital almacenado en un medio de almacenamiento de datos, u otras aplicaciones. En algunos ejemplos, el sistema de codificación de vídeo puede configurarse para soportar la transmisión de vídeo unidireccional o bidireccional para soportar aplicaciones tales como la transmisión de vídeo, la reproducción de vídeo, la radiodifusión de vídeo y/o la videotelefonía.

[0031] En el ejemplo de la **FIG. 1**, el dispositivo de origen 12 incluye una fuente 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/desmodulador (módem) y/o un transmisor. En el dispositivo de origen 12, la fuente de vídeo 18 puede incluir una fuente tal como un dispositivo de captura de vídeo, por ejemplo, una videocámara, un archivo de vídeo que contiene datos de vídeo previamente capturados, una interfaz de alimentación de vídeo para recibir datos de vídeo desde un proveedor de contenido de vídeo y/o un sistema de gráficos por ordenador para generar datos de vídeo, o una combinación de dichas fuentes.

[0032] El codificador de vídeo 20 puede estar configurado para codificar los datos de vídeo capturados, pre-capturados o generados por ordenador. Los datos de vídeo codificados se pueden transmitir directamente al dispositivo 14 de destino por medio de la interfaz de salida 22 del dispositivo de origen 12. Los datos de vídeo codificados también pueden almacenarse en un medio de almacenamiento o en un servidor de ficheros para un acceso posterior mediante el dispositivo de destino 14 para su descodificación y/o su reproducción.

[0033] En el ejemplo de la **FIG. 1**, el dispositivo de destino 14 incluye una interfaz de entrada 28, un descodificador 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 dispositivo de destino 14 recibe los datos de vídeo codificados por el canal 16. Los datos de vídeo codificados pueden incluir una variedad de elementos sintácticos generados por el codificador de vídeo 20 que representan los datos de vídeo. Los elementos sintácticos pueden describir características y/o procesamiento de bloques y otras unidades codificadas, por ejemplo, grupos de imágenes (GOP). 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.

[0034] El dispositivo de visualización 32 puede estar integrado con, o ser externo a, el dispositivo de destino 14. En algunos ejemplos, el dispositivo de destino 14 puede incluir un dispositivo de visualización integrado y también puede estar configurado para interactuar con un dispositivo de visualización externo. En otros ejemplos, el dispositivo de destino 14 puede ser un dispositivo de visualización. En general, el dispositivo de visualización 32 muestra los datos de vídeo descodificados a un usuario. El dispositivo de visualización 32 puede comprender cualquiera entre una variedad de dispositivos de visualización, tales como una pantalla de cristal líquido (LCD), una pantalla de plasma, una pantalla de diodos orgánicos emisores de luz (OLED) u otro tipo de dispositivo de visualización.

[0035] El codificador de vídeo 20 y el descodificador de vídeo 30 pueden funcionar de acuerdo con una norma de compresión de vídeo, tal como la norma de codificación de vídeo de alta eficiencia (HEVC), actualmente en fase de elaboración, y pueden ajustarse a un modelo de prueba de HEVC (HM). De manera alternativa, el codificador de vídeo 20 y el descodificador de vídeo 30 pueden funcionar de acuerdo con otras normas de propiedad o industriales, tales como la norma ITU-T H.264, de manera alternativa llamada MPEG-4, parte 10, codificación avanzada de vídeo (AVC), o ampliaciones de dichas normas. Sin embargo, las técnicas de esta divulgación no están limitadas a ninguna norma de codificación particular. Otros ejemplos de normas de compresión de vídeo incluyen MPEG-2 e ITU-T H.263.

[0036] Aunque no se muestra en el ejemplo de la **FIG. 1**, tanto el codificador de vídeo 20 como el descodificador de vídeo 30 pueden estar integrados en un codificador y descodificador de audio, y pueden incluir unidades MUX-DEMUX adecuadas, u otro tipo de hardware y software, para ocuparse de la codificación tanto de audio como de vídeo en un flujo de datos común o en flujos de datos separados. Si procede, en algunos ejemplos, las unidades MUX-DEMUX pueden ser conformes al protocolo de multiplexador ITU H.223 o a otros protocolos, tales como el protocolo de datagramas de usuario (UDP).

[0037] Nuevamente, la **FIG. 1** es meramente un ejemplo y las técnicas de esta divulgación pueden aplicarse a configuraciones de codificación de vídeo (por ejemplo, codificación de vídeo o descodificación de vídeo) que no incluyan necesariamente ninguna comunicación de datos entre los dispositivos de codificación y descodificación. En otros ejemplos, los datos pueden recuperarse de una memoria local, transmitirse a través de una red o similar. Un dispositivo de codificación puede codificar y almacenar datos en la memoria, y/o un dispositivo de descodificación puede recuperar y descodificar datos de la memoria. En muchos ejemplos, la codificación y la descodificación se realiza mediante dispositivos que no se comunican entre sí, sino que simplemente codifican datos en la memoria y/o recuperan y descodifican datos desde la memoria.

[0038] El codificador de vídeo 20 y el descodificador de vídeo 30 pueden, cada uno, implementarse como cualquiera entre varios circuitos adecuados, tales como uno o más microprocesadores, procesadores de señales digitales (DSP), circuitos integrados específicos de la aplicación (ASIC), matrices de puertas programables in situ (FPGA), lógica discreta, hardware 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 de almacenamiento no transitorio legible por ordenador adecuado, 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 descodificador de vídeo 30 se puede incluir en uno o más codificadores o descodificadores, cualquiera de los cuales se puede integrar como parte de un codificador/descodificador (CÓDEC) combinado en un dispositivo respectivo. Un dispositivo que incluye el codificador de vídeo 20 y/o el descodificador de vídeo 30 puede comprender un circuito integrado, un microprocesador y/o un dispositivo de comunicación inalámbrica, tal como un teléfono celular.

[0039] Como se ha mencionado brevemente anteriormente, el codificador de vídeo 20 codifica datos de vídeo. Los datos de vídeo pueden comprender una o más imágenes. Cada una de las imágenes es una imagen fija que forma parte de un vídeo. En algunos casos, una imagen puede denominarse una "trama" de vídeo. Cuando el codificador de vídeo 20 codifica los datos de vídeo, el codificador de vídeo 20 puede generar un flujo de bits. El flujo de bits puede incluir una secuencia de bits que forman una representación codificada de los datos de vídeo. El flujo de bits puede incluir imágenes codificadas y datos asociados. Una imagen codificada es una representación codificada de una imagen.

[0040] Para generar el flujo de bits, el codificador de vídeo 20 puede llevar a cabo operaciones de codificación en cada imagen en los datos de vídeo. Cuando el codificador de vídeo 20 realiza operaciones de codificación en las imágenes, el codificador de vídeo 20 puede generar una serie de imágenes codificadas y datos asociados. Los datos asociados pueden incluir conjuntos de parámetros de secuencia, conjuntos de parámetros de imagen, conjuntos de parámetros de adaptación y otras estructuras sintácticas. Un conjunto de parámetros de secuencia (SPS) puede contener parámetros aplicables a cero o más secuencias de imágenes. Un conjunto de parámetros de imagen (PPS) puede contener parámetros aplicables a cero o más imágenes. Un conjunto de parámetros de adaptación (APS) puede contener parámetros aplicables a cero o más imágenes. Los parámetros en un APS pueden ser parámetros que tienen más probabilidades de cambiar que los parámetros en el PPS.

[0041] Para generar una imagen codificada, el codificador de vídeo 20 puede dividir una imagen en bloques de vídeo de igual tamaño. Un bloque de vídeo puede ser una matriz bidimensional de muestras. Cada uno de los bloques de vídeo está asociado con un bloque de árbol. En algunos casos, un bloque de árbol también puede denominarse una unidad de codificación más grande (LCU). Los bloques de árbol de HEVC pueden ser aproximadamente análogos a los macrobloques de normas anteriores, tales como la H.264/AVC. Sin embargo, un bloque de árbol no está limitado necesariamente a un tamaño particular y puede incluir una o más unidades de codificación (CU). El codificador de vídeo 20 puede usar la división en árbol cuádruple para dividir los bloques de vídeo de bloques de árbol en bloques de vídeo asociados con las CU, de ahí el nombre "bloques de árbol".

[0042] En algunos ejemplos, el codificador de vídeo 20 puede dividir una imagen en una pluralidad de fragmentos. Cada uno de los fragmentos puede incluir un número entero de CU. En algunos casos, un fragmento comprende un número entero de bloques de árbol. En otros casos, un límite de un fragmento puede estar dentro de un bloque de árbol.

[0043] Como parte de realizar una operación de codificación en una imagen, el codificador de vídeo 20 puede realizar operaciones de codificación en cada fragmento de la imagen. Cuando el codificador de vídeo 20 realiza una operación de codificación en un fragmento, el codificador de vídeo 20 puede generar datos codificados asociados con el fragmento. Los datos codificados asociados con el fragmento pueden denominarse un "fragmento codificado".

5 **[0044]** Para generar un fragmento codificado, el codificador de vídeo 20 puede llevar a cabo operaciones de codificación en cada bloque de árbol en un fragmento. Cuando el codificador de vídeo 20 realiza una operación de codificación en un bloque de árbol, el codificador de vídeo 20 puede generar un bloque de árbol codificado. El bloque de árbol codificado puede comprender datos que representan una versión codificada del bloque de árbol.

10 **[0045]** Cuando el codificador de vídeo 20 genera un fragmento codificado, el codificador de vídeo 20 puede realizar operaciones de codificación (es decir, codificar) en los bloques de árbol en el fragmento de acuerdo con un orden de escaneo de trama. En otras palabras, el codificador de vídeo 20 puede codificar los bloques de árbol del fragmento en un orden que avanza de izquierda a derecha en la fila más alta de los bloques de árbol en el fragmento, a continuación procede de izquierda a derecha en la siguiente fila inferior de los bloques de árbol, y así sucesivamente hasta que el codificador de vídeo 20 ha codificado cada uno de los bloques de árbol del fragmento.

15 **[0046]** Como resultado de codificar los bloques de árbol de acuerdo con el orden de escaneo de trama, los bloques de árbol de arriba y a la izquierda de un bloque de árbol determinado pueden haber sido codificados, pero los bloques de árbol de debajo y a la derecha del bloque de árbol determinado aún no han sido codificados. En consecuencia, el codificador de vídeo 20 puede tener acceso a la información generada por la codificación de los bloques de árbol de arriba y a la izquierda del bloque de árbol determinado cuando codifica el bloque de árbol determinado. Sin embargo, el codificador de vídeo 20 puede no ser capaz de acceder a la información generada por la codificación de los bloques de árbol de debajo y a la derecha del bloque de árbol determinado al codificar el bloque de árbol determinado.

20 **[0047]** Para generar un bloque de árbol codificado, el codificador de vídeo 20 puede llevar a cabo de forma recursiva la división en árbol cuaternario en el bloque de vídeo del bloque de árbol para dividir el bloque de vídeo en bloques de vídeo progresivamente más pequeños. Cada uno de los bloques de vídeo más pequeños puede estar asociado con una CU diferente. Por ejemplo, el codificador de vídeo 20 puede dividir el bloque de vídeo de un bloque de árbol en cuatro sub-bloques de igual tamaño, dividir uno o más de los sub-bloques en cuatro sub-sub-bloques de igual tamaño, etc. Una CU dividida puede ser una CU cuyo bloque de vídeo está dividido en bloques de vídeo asociados con otras CU. Una CU no dividida puede ser una CU cuyo bloque de vídeo no esté dividido en bloques de vídeo asociados con otras CU.

25 **[0048]** Uno o más elementos sintácticos en el flujo de bits puede indicar un número máximo de veces que el codificador de vídeo 20 puede dividir el bloque de vídeo de un bloque de árbol. Un bloque de vídeo de una CU puede tener forma cuadrada. El tamaño del bloque de vídeo de una CU (es decir, el tamaño de la CU) puede variar desde 8x8 píxeles hasta el tamaño de un bloque de vídeo de un bloque de árbol (es decir, el tamaño del bloque de árbol) con un máximo de 64x64 píxeles o mayor.

30 **[0049]** El codificador de vídeo 20 puede realizar operaciones de codificación (es decir, codificar) en cada CU de un bloque de árbol de acuerdo con un orden de escaneo z. En otras palabras, el codificador de vídeo 20 puede codificar una CU superior izquierda, una CU superior derecha, una CU inferior izquierda y luego una CU inferior derecha, en ese orden. Cuando el codificador de vídeo 20 realiza una operación de codificación en una CU dividida, el codificador de vídeo 20 puede codificar las CU asociadas con sub-bloques del bloque de vídeo de la CU dividida de acuerdo con el orden de escaneo z. En otras palabras, el codificador de vídeo 20 puede codificar una CU asociada con un sub-bloque superior izquierdo, una CU asociada con un sub-bloque superior derecho, una CU asociada con un sub-bloque inferior izquierdo, y luego una CU asociada con un sub-bloque inferior derecho, en ese orden.

35 **[0050]** Como resultado de la codificación de las CU de un bloque de árbol de acuerdo con un orden de escaneo z, las CU arriba, arriba y a la izquierda, arriba y a la derecha, izquierda y abajo y a la izquierda de una CU determinada pueden haber sido codificadas. Las CU de abajo y a la derecha de la CU determinada todavía no han sido codificadas. En consecuencia, el codificador de vídeo 20 puede ser capaz de acceder a la información generada codificando algunas CU que se encuentran contiguas a la CU determinada cuando codifica la CU determinada. Sin embargo, el codificador de vídeo 20 puede ser incapaz de acceder a la información generada por la codificación de otras CU que se encuentran junto a la CU determinada cuando codifica la CU determinada.

40 **[0051]** Cuando el codificador de vídeo 20 codifica una CU no dividida, el codificador de vídeo 20 puede generar una o más unidades de predicción (PU) para la CU. Cada una de las PU de la CU puede estar asociada con un bloque de vídeo diferente dentro del bloque de vídeo de la CU. El codificador de vídeo 20 puede generar un bloque de vídeo predicho para cada PU de la CU. El bloque de vídeo predicho de una PU puede ser un bloque de muestras. El codificador de vídeo 20 puede usar predicción intra o predicción inter para generar el bloque de vídeo predicho para una PU.

45 **[0052]** Cuando el codificador de vídeo 20 utiliza la predicción intra para generar el bloque de vídeo predicho de una PU, el codificador de vídeo 20 puede generar el bloque de vídeo predicho de la PU basándose en muestras descodificadas de la imagen asociada con la PU. Si el codificador de vídeo 20 utiliza la predicción intra para generar bloques de vídeo predichos de las PU de una CU, la CU es una CU con predicción intra. Si el codificador de vídeo

20 utiliza la predicción inter para generar el bloque de vídeo predicho de la PU, el codificador de vídeo 20 puede generar el bloque de vídeo predicho de la PU basándose en muestras descodificadas de una o más imágenes distintas a la imagen asociada con la PU. Si el codificador de vídeo 20 utiliza la predicción inter para generar bloques de vídeo predichos de las PU de una CU, la CU es una CU con predicción inter.

5

[0053] Además, cuando el codificador de vídeo 20 utiliza predicción inter para generar un bloque de vídeo predicho para una PU, el codificador de vídeo 20 puede generar información de movimiento para la PU. La información de movimiento para una PU puede indicar uno o más bloques de referencia de la PU. Cada bloque de referencia de la PU puede ser un bloque de vídeo dentro de una imagen de referencia. La imagen de referencia puede ser una imagen distinta de la imagen asociada con la PU. En algunos casos, un bloque de referencia de una PU también puede denominarse la "muestra de referencia" de la PU. El codificador de vídeo 20 puede generar el bloque de vídeo predicho para la PU basándose en los bloques de referencia de la PU.

10

[0054] Después de que el codificador de vídeo 20 genere bloques de vídeo predichos para una o más PU de una CU, el codificador de vídeo 20 puede generar datos residuales para la CU basándose en los bloques de vídeo predichos para las PU de la CU. Los datos residuales para la CU pueden indicar diferencias entre muestras en los bloques de vídeo predichos para las PU de la CU y el bloque de vídeo original de la CU.

15

[0055] Además, como parte de la realización de una operación de codificación en una CU no dividida, el codificador de vídeo 20 puede realizar la división recursiva en árbol cuaternario en los datos residuales de la CU para dividir los datos residuales de la CU en uno o más bloques de datos residuales (es decir, bloques de vídeo residuales) asociados con las unidades de transformada (TU) de la CU. Cada TU de una CU puede estar asociada con un bloque de vídeo residual diferente.

20

[0056] El codificador de vídeo 20 puede aplicar una o más transformadas a bloques de vídeo residuales asociados con las TU para generar bloques de coeficientes de transformada (por ejemplo, bloques de coeficientes de transformada) asociados con las TU. Conceptualmente, un bloque de coeficientes de transformada puede ser una matriz bidimensional (2D) de coeficientes de transformada.

25

[0057] Después de generar un bloque de coeficientes de transformada, el codificador de vídeo 20 puede realizar un proceso de cuantificación en el bloque de coeficientes de transformada. La cuantificación se refiere, en general, a un proceso en el que los coeficientes de transformada se cuantifican para reducir posiblemente la cantidad de datos usados para representar los coeficientes de transformada, proporcionando una compresión adicional. El proceso de cuantificación puede reducir la profundidad de bits asociada a algunos o a la totalidad de los coeficientes de transformada. Por ejemplo, un coeficiente de transformada de n bits puede redondearse a la baja hasta un coeficiente de transformada de m bits durante la cuantificación, donde n es mayor que m .

30

35

[0058] El codificador de vídeo 20 puede asociar cada CU con un valor de parámetro de cuantificación (QP). El valor QP asociado con una CU puede determinar cómo el codificador de vídeo 20 cuantifica los bloques de coeficientes de transformada asociados con la CU. El codificador de vídeo 20 puede ajustar el grado de cuantificación aplicado a los bloques de coeficientes de transformada asociados con una CU ajustando el valor QP asociado con la CU.

40

[0059] Después de que el codificador de vídeo 20 cuantifica un bloque de coeficientes de transformada, el codificador de vídeo 20 puede generar conjuntos de elementos sintácticos que representan los coeficientes de transformada en el bloque de coeficientes de transformada cuantificado. El codificador de vídeo 20 puede aplicar operaciones de codificación por entropía, tales como operaciones de codificación aritmética binaria adaptable al contexto (CABAC), a algunos de estos elementos sintácticos. También podrían usarse otras técnicas de codificación por entropía, tales como la codificación de longitud variable adaptable al contenido (CAVLC), la codificación por entropía de división de intervalo de probabilidad (PIPE) u otra codificación aritmética binaria.

45

50

[0060] El flujo de bits generado por el codificador de vídeo 20 puede incluir una serie de unidades de capa de abstracción de red (NAL). Cada una de las unidades NAL puede ser una estructura sintáctica que contiene una indicación de un tipo de datos en la unidad NAL y los bytes que contienen los datos. Por ejemplo, una unidad NAL puede contener datos que representan un conjunto de parámetros de secuencia, un conjunto de parámetros de imagen, un fragmento codificado, información de mejora suplementaria (SEI), un delimitador de unidad de acceso, datos de relleno u otro tipo de datos. Los datos en una unidad NAL pueden incluir varias estructuras sintácticas.

55

[0061] El descodificador de vídeo 30 puede recibir el flujo de bits generado por el codificador de vídeo 20. El flujo de bits puede incluir una representación codificada de los datos de vídeo codificados por el codificador de vídeo 20. Cuando el descodificador de vídeo 30 recibe el flujo de bits, el descodificador de vídeo 30 puede realizar una operación de análisis sintáctico en el flujo de bits. Cuando el descodificador de vídeo 30 realiza la operación de análisis sintáctico, el descodificador de vídeo 30 puede extraer elementos sintácticos del flujo de bits. El descodificador de vídeo 30 puede reconstruir las imágenes de los datos de vídeo basándose en los elementos sintácticos extraídos del flujo de bits. El proceso para reconstruir los datos de vídeo basados en los elementos

60

65

sintácticos puede ser en general recíproco al proceso realizado por el codificador de vídeo 20 para generar los elementos sintácticos.

5 **[0062]** Después de que el descodificador de vídeo 30 extraiga los elementos sintácticos asociados con una CU, el descodificador de vídeo 30 puede generar bloques de vídeo predichos para las PU de la CU basándose en los elementos sintácticos. Además, el descodificador de vídeo 30 puede invertir bloques de coeficientes de transformada cuantificados asociados con TU de la CU. El descodificador de vídeo 30 puede realizar transformadas inversas en los bloques de coeficientes de transformada para reconstruir los bloques de vídeo residuales asociados con las TU de la CU. Después de generar los bloques de vídeo predichos y reconstruir los bloques de vídeo
10 residuales, el descodificador de vídeo 30 puede reconstruir el bloque de vídeo de la CU basado en los bloques de vídeo predichos y los bloques de vídeo residuales. De esta manera, el descodificador de vídeo 30 puede reconstruir los bloques de vídeo de las CU basándose en los elementos sintácticos del flujo de bits.

Codificador de vídeo

15 **[0063]** La **FIG. 2** es un diagrama de bloques que ilustra un ejemplo de un codificador de vídeo que puede implementar técnicas de acuerdo con los aspectos descritos en esta divulgación. El codificador de vídeo 20 puede configurarse para realizar algunas de, o todas, las técnicas de esta divulgación. Como ejemplo, la unidad de predicción 100 puede configurarse para realizar cualquiera o todas las técnicas descritas en esta divulgación. Sin embargo, los aspectos de esta divulgación no están limitados a lo anterior. En algunos ejemplos, las técnicas descritas en esta divulgación pueden compartirse entre los diversos componentes del codificador de vídeo 20. En algunos ejemplos, adicionalmente o en lugar de, un procesador (no mostrado) puede configurarse para realizar cualquiera o todas las técnicas descritas en esta divulgación.

25 **[0064]** Para los propósitos de explicación, esta divulgación describe el codificador de vídeo 20 en el contexto de la codificación HEVC. Sin embargo, las técnicas de esta divulgación pueden ser aplicables a otras normas o procedimientos de codificación.

30 **[0065]** El codificador de vídeo 20 puede realizar intracodificación e intercodificación de bloques de vídeo dentro de segmentos de vídeo. La intracodificación se apoya en la predicción espacial para reducir o eliminar la redundancia espacial en el vídeo dentro de una trama o imagen de vídeo determinada. 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. El modo intra (modo I) puede referirse a cualquiera de varios modos de codificación de base espacial. Los modos inter, tales como la predicción unidireccional (modo P) o la bipredicción (modo B), pueden referirse a cualquiera de varios modos de codificación de base temporal.

35 **[0066]** En el ejemplo de la **FIG. 2**, el codificador de vídeo 20 incluye una pluralidad de componentes funcionales. Los componentes funcionales del codificador de vídeo 20 incluyen una unidad de predicción 100, una unidad de generación residual 102, una unidad de transformada 104, una unidad de cuantificación 106, una unidad de cuantificación inversa 108, una unidad de transformada inversa 110, una unidad de reconstrucción 112, una unidad de filtro 113, una memoria intermedia de imágenes descodificadas 114 y una unidad de codificación por entropía 116. La unidad de predicción 100 incluye una unidad de predicción inter 121, una unidad de estimación de movimiento 122, una unidad de compensación de movimiento 124 y una unidad de predicción intra 126. En otros ejemplos, el codificador de vídeo 20 puede incluir más, menos o diferentes componentes funcionales. Además, la unidad de estimación de movimiento 122 y la unidad de compensación de movimiento 124 pueden estar altamente integradas, pero están representadas en el ejemplo de la **FIG. 2** de forma separada con fines explicativos.

40 **[0067]** El codificador de vídeo 20 puede recibir datos de vídeo. El codificador de vídeo 20 puede recibir los datos de vídeo de diversas fuentes. Por ejemplo, el codificador de vídeo 20 puede recibir los datos de vídeo de la fuente de vídeo 18 (**FIG. 1**) u otra fuente. Los datos de vídeo pueden representar una serie de imágenes. Para codificar los datos de vídeo, el codificador de vídeo 20 puede realizar una operación de codificación en cada una de las imágenes. Como parte de realizar la operación de codificación en una imagen, el codificador de vídeo 20 puede realizar operaciones de codificación en cada fragmento de la imagen. Como parte de realizar una operación de codificación en un fragmento, el codificador de vídeo 20 puede realizar operaciones de codificación en bloques de
55 árbol en el fragmento.

60 **[0068]** Como parte de realizar una operación de codificación en un bloque de árbol, la unidad de predicción 100 puede realizar la división en árbol cuaternario en el bloque de vídeo del bloque de árbol para dividir el bloque de vídeo en bloques de vídeo progresivamente más pequeños. Cada uno de los bloques de vídeo más pequeños puede estar asociado con una CU diferente. Por ejemplo, la unidad de predicción 100 puede dividir un bloque de vídeo de un bloque de árbol en cuatro sub-bloques de igual tamaño, dividir uno o más de los sub-bloques en cuatro sub-sub-bloques de igual tamaño, etc.

65 **[0069]** Los tamaños de los bloques de vídeo asociados con las CU pueden variar desde muestras de 8x8 hasta el tamaño del bloque de árbol, con un máximo de muestras de 64x64 o mayores. En esta divulgación, "NxN" y "N por N" pueden usarse indistintamente para hacer referencia a las dimensiones de muestras de un bloque de vídeo

en términos de dimensiones verticales y horizontales, por ejemplo, muestras de 16x16 o muestras de 16 por 16. En general, un bloque de vídeo de 16x16 tiene dieciséis muestras en una dirección vertical ($y = 16$) y dieciséis muestras en una dirección horizontal ($x = 16$). Asimismo, un bloque NxN presenta, en general, N muestras en una dirección vertical y N muestras en una dirección horizontal, donde N representa un valor entero no negativo.

5

[0070] Además, como parte de la realización de la operación de codificación en un bloque de árbol, la unidad de predicción 100 puede generar una estructura de datos jerárquica en árbol cuaternario para el bloque de árbol. Por ejemplo, un bloque de árbol puede corresponder a un nodo raíz de la estructura de datos en árbol cuaternario. Si la unidad de predicción 100 divide el bloque de vídeo del bloque de árbol en cuatro sub-bloques, el nodo raíz tiene cuatro nodos secundarios en la estructura de datos en árbol cuaternario. Cada uno de los nodos secundarios corresponde a una CU asociada con uno de los sub-bloques. Si la unidad de predicción 100 divide uno de los sub-bloques en cuatro sub-sub-bloques, el nodo correspondiente a la CU asociada con el sub-bloque puede tener cuatro nodos secundarios, cada uno de los cuales corresponde a una CU asociada con uno de los sub-sub-bloques.

10

15

[0071] Cada nodo de la estructura de datos en árbol cuaternario puede contener datos sintácticos (por ejemplo, elementos sintácticos) para el bloque de árbol o CU correspondiente. Por ejemplo, un nodo en el árbol cuádruple puede incluir un indicador de división, que indica si el bloque de vídeo de la CU correspondiente al nodo está dividido (es decir, partido) en cuatro sub-bloques. Los elementos sintácticos para una CU pueden definirse de manera recursiva y pueden depender de si el bloque de vídeo de la CU está dividido en sub-bloques. Una CU cuyo bloque de vídeo no está dividido puede corresponder a un nodo hoja en la estructura de datos en árbol cuaternario. Un bloque de árbol codificado puede incluir datos basados en la estructura de datos en árbol cuaternario para un bloque de árbol correspondiente.

20

25

[0072] El codificador de vídeo 20 puede realizar operaciones de codificación en cada CU no dividida de un bloque de árbol. Cuando el codificador de vídeo 20 realiza una operación de codificación en una CU no dividida, el codificador de vídeo 20 genera datos que representan una representación codificada de la CU no dividida.

30

[0073] Como parte de la realización de una operación de codificación en una CU, la unidad de predicción 100 puede dividir el bloque de vídeo de la CU entre una o más PU de la CU. El codificador de vídeo 20 y el decodificador de vídeo 30 pueden soportar varios tamaños de PU. Suponiendo que el tamaño de una CU particular sea $2N \times 2N$, el codificador de vídeo 20 y el decodificador de vídeo 30 pueden soportar tamaños de PU de $2N \times 2N$ o $N \times N$ y predicción inter en tamaños de PU simétricas de $2N \times 2N$, $2N \times N$, $N \times 2N$, $N \times N$, $2N \times nU$, $nL \times 2N$, $nR \times 2N$ o similares. El codificador de vídeo 20 y el decodificador de vídeo 30 también pueden soportar divisiones asimétricas para tamaños de PU de $2N \times nU$, $2N \times nD$, $nL \times 2N$ y $nR \times 2N$. En algunos ejemplos, la unidad de predicción 100 puede realizar una división geométrica para dividir el bloque de vídeo de una CU entre las PU de la CU a lo largo de un límite que no coincide con los lados del bloque de vídeo de la CU en ángulos rectos.

35

40

[0074] La unidad de predicción inter 121 puede realizar una predicción inter en cada PU de la CU. La predicción inter puede proporcionar compresión temporal. Para realizar la predicción inter en una PU, la unidad de estimación de movimiento 122 puede generar información de movimiento para la PU. La unidad de compensación de movimiento 124 puede generar un bloque de vídeo predicho para la PU basándose en la información de movimiento y las muestras descodificadas de imágenes distintas a la imagen asociada con la CU (es decir, imágenes de referencia). En esta divulgación, un bloque de vídeo predicho generado por la unidad de compensación de movimiento 124 se puede denominar un bloque de vídeo con predicción inter.

45

50

[0075] Los fragmentos pueden ser fragmentos I, fragmentos P o fragmentos B. El módulo de estimación de movimiento 122 y la unidad de compensación de movimiento 124 pueden realizar diferentes operaciones para una PU de una CU dependiendo de si la PU está en un fragmento I, un fragmento P o un fragmento B. En un fragmento en I, todas las PU tienen predicción inter. Por lo tanto, si la PU está en un fragmento I, la unidad de estimación de movimiento 122 y la unidad de compensación de movimiento 124 no realizan la predicción inter en la PU.

55

60

[0076] Si la PU está en un fragmento P, la imagen que contiene la PU está asociada con una lista de imágenes de referencia denominada "lista 0". Cada una de las imágenes de referencia en la lista 0 contiene muestras que pueden usarse para la predicción inter de otras imágenes. Cuando la unidad de estimación de movimiento 122 realiza la operación de estimación de movimiento con respecto a una PU en un fragmento P, la unidad de estimación de movimiento 122 puede buscar las imágenes de referencia en la lista 0 para un bloque de referencia para la PU. El bloque de referencia de la PU puede ser un conjunto de muestras, por ejemplo, un bloque de muestras, que se corresponde más estrechamente con las muestras del bloque de vídeo de la PU. La unidad de estimación de movimiento 122 puede usar una variedad de métricas para determinar cuán estrechamente un conjunto de muestras en una imagen de referencia se corresponde con las muestras del bloque de vídeo de una PU. Por ejemplo, la unidad de estimación de movimiento 122 puede determinar cuán estrechamente un conjunto de muestras en una imagen de referencia se corresponde con las muestras en el bloque de vídeo de una PU por la suma de la diferencia absoluta (SAD), la suma de diferencias al cuadrado (SSD) u otras métricas de diferencia.

65

[0077] Después de identificar un bloque de referencia de una PU en un fragmento P, la unidad de estimación de movimiento 122 puede generar un índice de referencia que indica la imagen de referencia de la lista 0 que contiene el bloque de referencia y un vector de movimiento que indica un desplazamiento espacial entre la PU y el bloque de referencia. En varios ejemplos, la unidad de estimación de movimiento 122 puede generar vectores de movimiento con diferentes grados de precisión. Por ejemplo, la unidad de estimación de movimiento 122 puede generar vectores de movimiento con una precisión de un cuarto de muestra, una precisión de un octavo de muestra o una precisión de otra fracción de muestra. En el caso de la precisión de la fracción de muestra, los valores del bloque de referencia pueden interpolarse a partir de valores de muestra de posición de entero en la imagen de referencia. La unidad de estimación de movimiento 122 puede emitir el índice de referencia y el vector de movimiento como la información de movimiento de la PU. La unidad de compensación de movimiento 124 puede generar el bloque de vídeo predicho de la PU basándose en el bloque de referencia identificado por la información de movimiento de la PU.

[0078] Si la PU está en un fragmento B, la imagen que contiene la PU puede estar asociada con dos listas de imágenes de referencia, denominadas "lista 0" y "lista 1." En algunos ejemplos, una imagen que contiene un fragmento B puede estar asociada con una combinación de listas que es una combinación de la lista 0 y la lista 1.

[0079] Además, si la PU está en un fragmento B, la unidad de estimación de movimiento 122 puede llevar a cabo la predicción unidireccional o la predicción bidireccional para la PU. Cuando la unidad de estimación de movimiento 122 realiza la predicción unidireccional para la PU, la unidad de estimación de movimiento 122 puede buscar las imágenes de referencia de la lista 0 o la lista 1 para un bloque de referencia para la PU. La unidad de estimación de movimiento 122 puede entonces generar un índice de referencia que indica la imagen de referencia en la lista 0 o la lista 1 que contiene el bloque de referencia y un vector de movimiento que indica un desplazamiento espacial entre la PU y el bloque de referencia. La unidad de estimación de movimiento 122 puede emitir el índice de referencia, un indicador de la dirección de predicción y el vector de movimiento como la información de movimiento de la PU. El indicador de dirección de predicción puede indicar si el índice de referencia indica una imagen de referencia en la lista 0 o la lista 1. La unidad de compensación de movimiento 124 puede generar el bloque de vídeo predicho de la PU basándose en el bloque de referencia indicado por la información de movimiento de la PU.

[0080] Cuando la unidad de estimación de movimiento 122 realiza la predicción bidireccional para una PU, la unidad de estimación de movimiento 122 puede buscar las imágenes de referencia de la lista 0 para un bloque de referencia para la PU y también puede buscar las imágenes de referencia de la lista 1 para otro bloque de referencia para la PU. La unidad de estimación de movimiento 122 puede entonces generar índices de referencia que indican las imágenes de referencia en la lista 0 y la lista 1 que contienen los bloques de referencia y los vectores de movimiento que indican desplazamientos espaciales entre los bloques de referencia y la PU. La unidad de estimación de movimiento 122 puede emitir los índices de referencia y los vectores de movimiento de la PU como la información de movimiento de la PU. La unidad de compensación de movimiento 124 puede generar el bloque de vídeo predicho de la PU basándose en los bloques de referencia indicados por la información de movimiento de la PU.

[0081] Como se analiza adicionalmente a continuación con referencia a la **FIG. 5**, la unidad de predicción 100 (por ejemplo, a través de la unidad de predicción inter 121) puede configurarse para predecir un bloque actual en la EL que tiene un bloque correspondiente no disponible en la BL realizando los pasos ilustrados en la **FIG. 5**.

[0082] En algunos casos, la unidad de estimación de movimiento 122 no emite un conjunto completo de información de movimiento para una PU a la unidad de codificación por entropía 116. En lugar de eso, la unidad de estimación de movimiento 122 puede señalar la información de movimiento de una PU con referencia a la información de movimiento de otra PU. Por ejemplo, la unidad de estimación de movimiento 122 puede determinar que la información de movimiento de la PU es suficientemente similar a la información de movimiento de una PU próxima. En este ejemplo, la unidad de estimación de movimiento 122 puede indicar, en una estructura sintáctica asociada con la PU, un valor que indica al descodificador de vídeo 30 que la PU tiene la misma información de movimiento que la PU contigua. En otro ejemplo, la unidad de estimación de movimiento 122 puede identificar, en una estructura sintáctica asociada con la PU, una PU próxima y una diferencia de vectores de movimiento (MVD). La diferencia de vectores de movimiento indica una diferencia entre el vector de movimiento de la PU y el vector de movimiento de la PU próxima indicada. El descodificador de vídeo 30 puede utilizar el vector de movimiento de la PU contigua indicada y la diferencia de vectores de movimiento para determinar el vector de movimiento de la PU. Haciendo referencia a la información de movimiento de una primera PU cuando se señala la información de movimiento de una segunda PU, el codificador de vídeo 20 puede ser capaz de señalar la información de movimiento de la segunda PU usando menos bits.

[0083] Como parte de la realización de una operación de codificación en una CU, la unidad de predicción intra 126 puede llevar a cabo la predicción intra en las PU de la CU. La predicción intra puede proporcionar compresión espacial. Cuando la unidad de predicción intra 126 realiza la predicción intra en una PU, la unidad de predicción intra 126 puede generar datos de predicción para la PU basados en muestras descodificadas de otras PU en la misma imagen. Los datos de predicción para la PU pueden incluir un bloque de vídeo predicho y varios elementos

5 sintácticos. La unidad de predicción intra 126 puede realizar la predicción intra en PU en fragmentos I, fragmentos P y fragmentos B.

5 **[0084]** Para realizar la predicción intra en una PU, la unidad de predicción intra 126 puede usar múltiples modos de predicción intra para generar múltiples conjuntos de datos de predicción para la PU. Cuando la unidad de predicción intra 126 utiliza un modo de predicción intra para generar un conjunto de datos de predicción para la PU, la unidad de predicción intra 126 puede extender muestras de bloques de vídeo de PU próximas a través del bloque de vídeo de la PU en una dirección y/o gradiente asociada con el modo de predicción intra. Las PU contiguas pueden estar arriba, arriba y a la derecha, arriba y a la izquierda o a la izquierda de la PU, suponiendo un orden de codificación de izquierda a derecha, de arriba a abajo para las PU, CU y bloques de árbol. La unidad de predicción intra 126 puede usar varios números de modos de predicción intra, por ejemplo, 33 modos de predicción intra direccional, dependiendo del tamaño de la PU.

15 **[0085]** La unidad de predicción 100 puede seleccionar los datos de predicción para una PU entre los datos de predicción generados por la unidad de compensación de movimiento 124 para la PU o los datos de predicción generados por la unidad de predicción intra 126 para la PU. En algunos ejemplos, la unidad de predicción 100 selecciona los datos de predicción para la PU basándose en las métricas de velocidad/distorsión de los conjuntos de datos de predicción.

20 **[0086]** Si la unidad de predicción 100 selecciona datos de predicción generados por la unidad de predicción intra 126, la unidad de predicción 100 puede señalar el modo de predicción intra que se utilizó para generar los datos de predicción para las PU, es decir, el modo de predicción intra seleccionado. La unidad de predicción 100 puede señalar el modo de predicción intra seleccionado de varias maneras. Por ejemplo, es probable que el modo de predicción intra seleccionado sea el mismo que el modo de predicción intra de una PU próxima. En otras palabras, el modo de predicción intra de la PU próxima puede ser el modo más probable para la PU actual. De este modo, la unidad de predicción 100 puede generar un elemento sintáctico para indicar que el modo de predicción intra seleccionado es el mismo que el modo de predicción intra de la PU contigua.

30 **[0087]** Después de que la unidad de predicción 100 seleccione los datos de predicción para las PU de una CU, la unidad de generación residual 102 puede generar datos residuales para la CU restando los bloques de vídeo predichos de las PU de la CU del bloque de vídeo de la CU. Los datos residuales de una CU pueden incluir bloques de vídeo residuales en 2D que corresponden a diferentes componentes de muestra de las muestras en el bloque de vídeo de la CU. Por ejemplo, los datos residuales pueden incluir un bloque de vídeo residual que corresponde a diferencias entre componentes de luminancia de muestras en los bloques de vídeo predichos de las PU de la CU y componentes de luminancia de muestras en el bloque de vídeo original de la CU. Además, los datos residuales de la CU pueden incluir bloques de vídeo residuales que corresponden a las diferencias entre componentes de crominancia de muestras en los bloques de vídeo predichos de las PU de la CU y los componentes de crominancia de las muestras en el bloque de vídeo original de la CU.

40 **[0088]** La unidad de predicción 100 puede realizar la división en árbol cuaternario para dividir los bloques de vídeo residuales de una CU en sub-bloques. Cada bloque de vídeo residual no dividido puede estar asociado con una TU diferente de la CU. Los tamaños y posiciones de los bloques de vídeo residuales asociados con las TU de una CU pueden o no basarse en los tamaños y posiciones de los bloques de vídeo asociados con las PU de la CU. Una estructura en árbol cuaternario conocida como "árbol cuaternario residual" (RQT) puede incluir nodos asociados con cada uno de los bloques de vídeo residuales. Las TU de una CU pueden corresponder a nodos de hoja del RQT.

50 **[0089]** La unidad de transformada 104 puede generar uno o más bloques de coeficientes de transformada para cada TU de una CU aplicando una o más transformadas a un bloque de vídeo residual asociado con la TU. Cada uno de los bloques de coeficientes de transformada puede ser una matriz en 2D de coeficientes de transformada. La unidad de transformada 104 puede aplicar varias transformadas al bloque de vídeo residual asociado con una TU. Por ejemplo, la unidad de transformada 104 puede aplicar una transformada de coseno discreta (DCT), una transformada direccional o una transformada conceptualmente similar al bloque de vídeo residual asociado con una TU.

55 **[0090]** Después de que la unidad de transformada 104 genere un bloque de coeficientes de transformada asociado con una TU, la unidad de cuantificación 106 puede cuantificar los coeficientes de transformada en el bloque de coeficientes de transformada. La unidad de cuantificación 106 puede cuantificar un bloque de coeficientes de transformada asociado con una TU de una CU basándose en un valor de QP asociado con la CU.

60 **[0091]** El codificador de vídeo 20 puede asociar un valor QP con una CU de varias maneras. Por ejemplo, el codificador de vídeo 20 puede realizar un análisis de distorsión de velocidad en un bloque de árbol asociado con la CU. En el análisis de distorsión de velocidad, el codificador de vídeo 20 puede generar múltiples representaciones codificadas del bloque de árbol realizando una operación de codificación varias veces en el bloque de árbol. El codificador de vídeo 20 puede asociar diferentes valores de QP con la CU cuando el codificador de vídeo 20 genera diferentes representaciones codificadas del bloque de árbol. El codificador de vídeo 20 puede

señalar que un valor QP dado está asociado con la CU cuando el valor QP dado está asociado con la CU en una representación codificada del bloque de árbol que tiene una velocidad de bits y una métrica de distorsión más bajas.

5 **[0092]** La unidad de cuantificación inversa 108 y la unidad de transformada inversa 110 pueden aplicar la cuantificación inversa y las transformadas inversas al bloque de coeficientes de transformada, respectivamente, para reconstruir un bloque de vídeo residual a partir del bloque de coeficientes de transformada. La unidad de reconstrucción 112 puede añadir el bloque de vídeo residual reconstruido a las muestras correspondientes de uno o más bloques de vídeo predichos generados por la unidad de predicción 100 para producir un bloque de vídeo reconstruido asociado con una TU. Mediante la reconstrucción de bloques de vídeo para cada TU de una CU de esta manera, el codificador de vídeo 20 puede reconstruir el bloque de vídeo de la CU.

15 **[0093]** Después de que la unidad de reconstrucción 112 reconstruya el bloque de vídeo de una CU, la unidad de filtro 113 puede realizar una operación de desbloqueo para reducir los artefactos de bloqueo en el bloque de vídeo asociado con la CU. Después de realizar las una o más operaciones de desbloqueo, la unidad de filtro 113 puede almacenar el bloque de vídeo reconstruido de la CU en la memoria intermedia de imágenes descodificadas 114. La unidad de estimación de movimiento 122 y la unidad de compensación de movimiento 124 pueden utilizar una imagen de referencia que contiene el bloque de vídeo reconstruido para realizar la predicción inter en las PU de las imágenes posteriores. Además, la unidad de predicción intra 126 puede usar bloques de vídeo reconstruidos en la memoria intermedia de imágenes descodificadas 114 para realizar la predicción intra en otras PU en la misma imagen que la CU.

25 **[0094]** La unidad de codificación por entropía 116 puede recibir datos desde otros componentes funcionales del codificador de vídeo 20. Por ejemplo, la unidad de codificación por entropía 116 puede recibir bloques de coeficientes de transformada de la unidad de cuantificación 106 y puede recibir elementos sintácticos de la unidad de predicción 100. Cuando la unidad de codificación por entropía 116 recibe los datos, la unidad de codificación por entropía 116 puede realizar una o más operaciones de codificación por entropía para generar datos codificados por entropía. Por ejemplo, el codificador de vídeo 20 puede realizar una operación de codificación de longitud variable adaptable al contexto (CAVLC), una operación CABAC, una operación de codificación de longitud variable a variable (V2V), una operación de codificación aritmética binaria adaptable al contexto basada en sintaxis (SBAC), una operación de codificación por entropía de división de intervalo de probabilidad (PIPE) u otro tipo de operación de codificación por entropía en los datos. La unidad de codificación por entropía 116 puede emitir un flujo de bits que incluye los datos codificados por entropía.

35 **[0095]** Como parte de la realización de una operación de codificación por entropía en los datos, la unidad de codificación por entropía 116 puede seleccionar un modelo de contexto. Si la unidad de codificación por entropía 116 está realizando una operación CABAC, el modelo de contexto puede indicar las estimaciones de las probabilidades de que un bin particular tenga unos valores particulares. En el contexto de CABAC, el término "bin" se utiliza para referirse a un bit de una versión binarizada de un elemento sintáctico.

40

DESCODIFICADOR DE VÍDEO

45 **[0096]** La **FIG. 3** es un diagrama de bloques que ilustra un ejemplo de un descodificador de vídeo que puede implementar técnicas de acuerdo con aspectos descritos en esta divulgación. El descodificador de vídeo 30 puede configurarse para realizar algunas de, o todas, las técnicas de esta divulgación. Como ejemplo, la unidad de compensación de movimiento 162 y/o la unidad de predicción intra 164 pueden configurarse para realizar cualquiera o todas las técnicas descritas en esta divulgación. Sin embargo, los aspectos de esta divulgación no están limitados a lo anterior. En algunos ejemplos, las técnicas descritas en esta divulgación pueden compartirse entre los diversos componentes del descodificador de vídeo 30. En algunos ejemplos, adicionalmente o en lugar de, un procesador (no mostrado) puede configurarse para realizar cualquiera o todas las técnicas descritas en esta divulgación.

55 **[0097]** En el ejemplo de la **FIG. 3**, el descodificador de vídeo 30 incluye una pluralidad de componentes funcionales. Los componentes funcionales del descodificador de vídeo 30 incluyen una unidad de descodificación por entropía 150, una unidad de predicción 152, una unidad de cuantificación inversa 154, una unidad de transformada inversa 156, una unidad de reconstrucción 158, una unidad de filtro 159 y una memoria intermedia de imágenes descodificadas 160. La unidad de predicción 152 incluye la unidad de compensación de movimiento 162 y una unidad de predicción intra 164. En algunos ejemplos, el descodificador de vídeo 30 puede realizar una pasada de descodificación, en general, recíproca a la pasada de codificación descrita con respecto al codificador de vídeo 20 de la **FIG. 2**. En otros ejemplos, el descodificador de vídeo 30 puede incluir más, menos o diferentes componentes funcionales.

60 **[0098]** El descodificador de vídeo 30 puede recibir un flujo de bits que comprende datos de vídeo codificados. El flujo de bits puede incluir una pluralidad de elementos sintácticos. Cuando el descodificador de vídeo 30 recibe el flujo de bits, la unidad de descodificación por entropía 150 puede realizar una operación de análisis sintáctico en el flujo de bits. Como resultado de realizar la operación de análisis sintáctico en el flujo de bits, la unidad de

65

descodificación por entropía 150 puede extraer elementos sintácticos del flujo de bits. Como parte de la realización de la operación de análisis sintáctico, la unidad de descodificación por entropía 150 puede realizar la descodificación por entropía de elementos sintácticos sometidos a codificación por entropía en el flujo de bits. La unidad de predicción 152, la unidad de cuantificación inversa 154, la unidad de transformada inversa 156, la unidad de reconstrucción 158 y la unidad de filtro 159 pueden realizar una operación de reconstrucción que genera datos de vídeo descodificados basados en los elementos sintácticos extraídos del flujo de bits.

[0099] Como se analizó anteriormente, el flujo de bits puede comprender una serie de unidades NAL. Las unidades NAL del flujo de bits pueden incluir unidades NAL de conjunto de parámetros de secuencia, unidades NAL de conjunto de parámetros de imagen, unidades SEI NAL, etc. Como parte de realizar la operación de análisis sintáctico en el flujo de bits, la unidad de descodificación por entropía 150 puede realizar operaciones de análisis sintáctico que extraen y descodifican por entropía conjuntos de parámetros de secuencia a partir de unidades NAL de conjuntos de parámetros de secuencia, conjuntos de parámetros de imagen a partir de unidades NAL de conjuntos de parámetros de imagen, datos SEI a partir de unidades NAL SEI, etc.

[0100] Además, las unidades NAL del flujo de bits pueden incluir unidades NAL de fragmentos codificados. Como parte de realizar la operación de análisis sintáctico en el flujo de bits, la unidad de descodificación por entropía 150 puede realizar operaciones de análisis sintáctico que extraen y descodifican por entropía fragmentos codificados a partir de las unidades de NAL de fragmentos codificados. Cada uno de los fragmentos codificados puede incluir una cabecera de fragmento y datos de fragmento. La cabecera de fragmento puede contener elementos sintácticos pertenecientes a un fragmento. Los elementos sintácticos en la cabecera del fragmento pueden incluir un elemento sintáctico que identifica un conjunto de parámetros de imagen asociado con una imagen que contiene el fragmento. La unidad de descodificación por entropía 150 puede realizar operaciones de descodificación por entropía, tales como operaciones de descodificación de CABAC, en los elementos sintácticos de la cabecera del fragmento codificado para recuperar la cabecera del fragmento.

[0101] Como parte de extraer los datos del fragmento de unidades NAL de fragmentos codificados, la unidad de descodificación por entropía 150 puede realizar operaciones de análisis sintáctico que extraen los elementos sintácticos de las CU codificadas en los datos del fragmento. Los elementos sintácticos extraídos pueden incluir elementos sintácticos asociados con bloques de coeficientes de transformada. La unidad de descodificación por entropía 150 puede entonces realizar operaciones de descodificación de CABAC en algunos de los elementos sintácticos.

[0102] Después de que la unidad de descodificación por entropía 150 realice una operación de análisis sintáctico en una CU no dividida, el descodificador de vídeo 30 puede realizar una operación de reconstrucción en la CU no dividida. Para realizar la operación de reconstrucción en una CU no dividida, el descodificador de vídeo 30 puede realizar una operación de reconstrucción en cada TU de la CU. Realizando la operación de reconstrucción para cada TU de la CU, el descodificador de vídeo 30 puede reconstruir un bloque de vídeo residual asociado con la CU.

[0103] Como parte de la realización de una operación de reconstrucción en una TU, la unidad de cuantificación inversa 154 puede cuantificar de forma inversa, es decir, descuantificar, un bloque de coeficientes de transformada asociado con la TU. La unidad de cuantificación inversa 154 puede cuantificar de forma inversa el bloque de coeficientes de transformada de una manera similar a los procesos de cuantificación inversa propuestos para HEVC o definidos por la norma de descodificación H.264. La unidad de cuantificación inversa 154 puede utilizar un parámetro de cuantificación QP calculado por el codificador de vídeo 20 para una CU del bloque de coeficientes de transformada para determinar un grado de cuantificación y, del mismo modo, un grado de cuantificación inversa para la unidad de cuantificación inversa 154 a aplicar.

[0104] Después de que la unidad de cuantificación inversa 154 cuantifique de forma inversa un bloque de coeficientes de transformada, la unidad de transformada inversa 156 puede generar un bloque de vídeo residual para la TU asociada con el bloque de coeficientes de transformada. La unidad de transformada inversa 156 puede aplicar una transformada inversa al bloque de coeficientes de transformada para generar el bloque de vídeo residual para la TU. Por ejemplo, la unidad de transformada inversa 156 puede aplicar una DCT inversa, una transformada de número entero inversa, una transformada de Karhunen-Loeve (KLT) inversa, una transformada de rotación inversa, una transformada direccional inversa u otra transformada inversa al bloque de coeficientes de transformada. En algunos ejemplos, la unidad de transformada inversa 156 puede determinar una transformada inversa para aplicar al bloque de coeficientes de transformada basándose en la señalización del codificador de vídeo 20. En dichos ejemplos, la unidad de transformada inversa 156 puede determinar la transformada inversa basada en una transformada señalizada en el nodo raíz de un árbol cuaternario para un bloque de árbol asociado con el bloque de coeficientes de transformada. En otros ejemplos, la unidad de transformada inversa 156 puede inferir la transformada inversa a partir de una o más características de codificación, tales como tamaño de bloque, modo de codificación o similares. En algunos ejemplos, la unidad de transformada inversa 156 puede aplicar una transformada inversa en cascada.

[0105] En algunos ejemplos, la unidad de compensación de movimiento 162 puede refinar el bloque de vídeo predicho de una PU mediante interpolación basándose en filtros de interpolación. Los identificadores para los filtros de interpolación que van a usarse para la compensación de movimiento con una precisión de submuestra pueden incluirse en los elementos sintácticos. La unidad de compensación de movimiento 162 puede usar los mismos filtros de interpolación usados por el codificador de vídeo 20 durante la generación del bloque de vídeo predicho de la PU para calcular valores interpolados para muestras de subenteros de un bloque de referencia. La unidad de compensación de movimiento 162 puede determinar los filtros de interpolación usados por el codificador de vídeo 20 de acuerdo con la información sintáctica recibida y usar los filtros de interpolación para producir el bloque de vídeo predicho.

[0106] Si se codifica una PU usando predicción intra, la unidad de predicción intra 164 puede realizar la predicción intra para generar un bloque de vídeo predicho para la PU. Por ejemplo, la unidad de predicción intra 164 puede determinar un modo de predicción intra para la PU basándose en los elementos sintácticos del flujo de bits. El flujo de bits puede incluir elementos sintácticos que la unidad de predicción intra 164 puede usar para predecir el modo de predicción intra de la PU.

[0107] En algunos casos, los elementos sintácticos pueden indicar que la unidad de predicción intra 164 tiene que utilizar el modo de predicción intra de otra PU para determinar el modo de predicción intra de la PU actual. Por ejemplo, puede ser probable que el modo de predicción intra de la PU actual sea el mismo que el modo de predicción intra de una PU próxima. En otras palabras, el modo de predicción intra de la PU próxima puede ser el modo más probable para la PU actual. Por lo tanto, en este ejemplo, el flujo de bits puede incluir un elemento sintáctico pequeño que indica que el modo de predicción intra de la PU es el mismo que el modo de predicción intra de la PU próxima. La unidad de predicción intra 164 puede entonces utilizar el modo de predicción intra para generar datos de predicción (por ejemplo, muestras predichas) para la PU basándose en los bloques de vídeo de las PU espacialmente próximas.

[0108] Como se analiza más adelante a continuación con referencia a la FIG. 5, la unidad de predicción 152 se puede configurar para predecir un bloque actual en la EL que tiene un bloque correspondiente no disponible en la BL realizando los pasos ilustrados en la FIG. 5.

[0109] La unidad de reconstrucción 158 puede utilizar los bloques de vídeo residuales asociados con las TU de una CU y los bloques de vídeo predichos de las PU de la CU, es decir, datos de predicción intra o datos de predicción inter, según corresponda, para reconstruir el bloque de vídeo de la CU. De este modo, el descodificador de vídeo 30 puede generar un bloque de vídeo predicho y un bloque de vídeo residual basándose en los elementos sintácticos en el flujo de bits y puede generar un bloque de vídeo basándose en el bloque de vídeo predicho y el bloque de vídeo residual.

[0110] Después de que la unidad de reconstrucción 158 reconstruya el bloque de vídeo de la CU, la unidad de filtro 159 puede realizar una operación de desbloqueo para reducir los artefactos de bloqueo asociados con la CU. Después de que la unidad de filtro 159 realice una operación de desbloqueo para reducir los artefactos de bloqueo asociados con la CU, el descodificador de vídeo 30 puede almacenar el bloque de vídeo de la CU en la memoria intermedia de imágenes descodificadas 160. La memoria intermedia de imágenes descodificadas 160 puede proporcionar imágenes de referencia para la posterior compensación de movimiento, la predicción intra y la presentación en un dispositivo de visualización, tal como el dispositivo de visualización 32 de la FIG. 1. Por ejemplo, el descodificador de vídeo 30 puede realizar, basándose en los bloques de vídeo en la memoria intermedia de imágenes descodificadas 160, operaciones de predicción intra o de predicción inter en las PU de otras CU.

[0111] En un modo de realización de la extensión HEVC, se puede predecir la capa de mejora u otra sintaxis de visualización, como el campo de movimiento o los modos intra, utilizando el bloque correspondiente de la capa base. Por ejemplo, un vector de movimiento de capa base (MV) se puede utilizar como candidato en las listas de candidatos del modo de combinación de capa de mejora/AMVP (predicción avanzada de vectores de movimiento). Sin embargo, puede haber una situación en la que un bloque correspondiente o colocado de la capa base esté ubicado fuera de la trama de la capa base. Esto puede suceder cuando el tamaño de trama codificada es más grande que el tamaño de trama real que se puede mostrar en un dispositivo.

[0112] Por ejemplo, una situación de este tipo puede ocurrir cuando el tamaño de relleno del límite de la imagen de mejora es mayor que el de la imagen de la capa base. La configuración del codificador específico podría introducir el diferente tamaño de relleno de límite de las capas de base y de mejora. También puede deberse a los diferentes tamaños de unidades de codificación más pequeñas (SCU) de las capas base y de mejora. Dado que en HEVC el tamaño de trama codificada real se alinea con el tamaño de la SCU (un tamaño de trama es un múltiplo entero de las SCU), si la SCU de la capa de mejora es más grande que la SCU de la capa base, no hay información de sintaxis para algunas partes de la trama de la capa base. Como se muestra en la FIG. 4, los diferentes tamaños de la SCU de la capa de mejora 422 y SCU de la capa base 432 pueden dar como resultado que una parte de la capa de mejora 400 no tenga una región correspondiente en la capa base 430 de la cual se pueda recuperar información de movimiento o información de píxeles para la predicción de capa intermedia. En otro ejemplo, la

indisponibilidad de la información de la capa base puede ocurrir cuando la capa base está codificada con otra norma de vídeo (por ejemplo, AVC o MPEG2).

[0113] En 3D-HEVC, al aplicar un vector de disparidad a un bloque de límite para ubicar un bloque en la visualización de referencia, el bloque correspondiente puede caer fuera del límite del componente de visualización de la visualización de referencia.

[0114] Esta divulgación está en general dirigida a la extensión SVC. Esta divulgación también es válida para la codificación de vídeo de múltiples visualizaciones (MVC) donde una de las visualizaciones sirve como una capa de mejora en la descripción.

[0115] La **FIG. 4** ilustra los diversas tramas de la capa de mejora 400 y la capa base 430. El rectángulo blanco 432 representa el tamaño de trama real de la imagen de la capa base (por ejemplo, recortada para ajustarse al tamaño de visualización de un dispositivo). El tamaño de trama codificada de la capa base se muestra en líneas discontinuas 436 en la parte inferior de la **FIG. 4**, y el tamaño de trama codificada de la capa de mejora se muestra en las líneas discontinuas 416 en la parte superior de la **FIG. 4**. También se muestra entre el tamaño de trama real 412 y el tamaño de trama codificado 416 de la capa de mejora en líneas discontinuas 414 un tamaño de trama codificado de la capa de base muestreada. El área 418 representa el área entre el tamaño de trama real y el tamaño de trama codificada de la capa de base muestreada de manera ascendente, y el área 420 representa el área entre el tamaño de trama codificada de la capa de base muestreada de manera ascendente y el tamaño de trama codificada de la capa de mejora.

[0116] Típicamente, para la predicción entre capas, la información de la capa base, incluido el vector de movimiento (MV) y otros elementos sintácticos, se puede obtener a partir de la ubicación de sub-bloque o píxeles correspondiente y se puede usar para predecir el bloque actual (por ejemplo, en la capa de mejora). Esta ubicación de sub-bloque o píxeles correspondiente puede estar ubicado dentro o fuera de la CU o la PU de la capa de mejora actual. Por ejemplo, la ubicación de sub-bloque o píxel puede ser una ubicación de sub-bloque o píxel de centro o de esquina. Si esta ubicación de sub-bloque o píxel correspondiente se encuentra en un área no disponible fuera de la trama de la capa base, es deseable definir el comportamiento del códec SVC o MVC para este caso. Por ejemplo, el "bloque correspondiente" (o ubicación de sub-bloque o píxel) utilizado en el presente documento puede referirse a un bloque de ubicación conjunta, un bloque contiguo o cualquier otro bloque determinado por el codificador (por ejemplo, codificador o descodificador).

[0117] Como se mencionó anteriormente, puede haber regiones de la capa de mejora que no tengan regiones correspondientes en la capa base disponibles para la predicción entre capas. Por ejemplo, los bordes inferior y derecho de la trama codificada de la capa de mejora (es decir, el área 420 en la **FIG. 4**) no tienen regiones correspondientes en la capa base desde la cual se puede obtener información de movimiento para usar en la predicción entre capas. Por lo tanto, tales partes de borde no pueden codificarse con la información obtenida de sus partes correspondientes en la capa base, y por lo tanto no se pueden usar las técnicas convencionales de predicción entre capas.

Falta de disponibilidad del bloque de referencia en BL

[0118] Se puede verificar si la ubicación del sub-bloque o píxel correspondiente está fuera de la trama de la capa base comparando las coordenadas de las esquinas del sub-bloque o la ubicación del píxel con el tamaño de trama. Por ejemplo, si las coordenadas de la esquina inferior derecha del sub-bloque están fuera del rango del tamaño de trama de la capa base (o el tamaño de trama de la BL con muestreo ascendente de acuerdo con la relación espacial de la BL y la EL), el sub-bloque correspondiente se considera no disponible para predecir el bloque actual en la EL. Una o más de las técnicas analizadas en el presente documento se pueden usar para realizar una predicción entre capas en bloques de este tipo que tienen bloques correspondientes no disponibles en la BL (por ejemplo, sintaxis de BL no disponible).

[0119] En un modo de realización, durante una predicción de movimiento entre visualizaciones/capas, se identifican las posiciones superior izquierda e inferior derecha P_0 y P_1 del bloque actual en la EL, que corresponden a las ubicaciones Q_0 y Q_1 en la visualización/capa de referencia/base. Si Q_0 o Q_1 está fuera de la imagen de visualización/capa de referencia/base, se considera que el bloque correspondiente no está disponible y no se usa para la predicción de movimiento entre capas/visualización. En tal caso, la predicción entre capas/visualización de otros elementos sintácticos o información residual puede estar deshabilitada. En otro modo de realización, se puede considerar que el bloque de visualización de capa/referencia base no está disponible solo cuando Q_0 y Q_1 están fuera de la imagen de la visualización/capa de referencia/base.

[0120] En otro modo de realización, en lugar de las posiciones superior izquierda e inferior derecha del bloque actual en la EL, el bloque actual se identifica por las posiciones superior derecha e inferior izquierda del bloque, y por lo tanto la disponibilidad del bloque de referencia en la BL se determina basándose en las posiciones asignadas de las posiciones superior derecha e inferior izquierda.

[0121] En otro modo de realización adicional, el bloque actual puede identificarse por las coordenadas horizontales de los píxeles más a la izquierda y los píxeles más a la derecha de los bloques, X_1 y X_r y las coordenadas verticales de los píxeles superiores y los píxeles inferiores, Y_1 y Y_r . Después de la asignación de dichas coordenadas a las coordenadas correspondientes en la capa/visualización de base/referencia, si alguno de los valores (por ejemplo, coordenadas horizontales y verticales) se encuentra más allá del límite de la imagen, se considera que el bloque correspondiente en la visualización de capa/referencia base no está disponible. Aquí, el límite de la imagen puede referirse al límite de la trama real o la trama codificada, como se ilustra en la **FIG. 4**.

[0122] Para un determinado elemento sintáctico (o todos los elementos sintácticos), la posición asignada en la visualización de capa/referencia base del bloque actual puede depender de la posición asignada de una coordenada específica del bloque actual. Cuando la posición asignada de la coordenada específica del bloque actual está más allá del límite de la visualización de capa/referencia base, se considera que el bloque de visualización de capa/referencia base no está disponible. La coordenada específica de un bloque podría apuntar a la posición superior izquierda, inferior derecha, central, superior derecha e inferior izquierda del bloque. La coordenada específica de un bloque podría ser la misma posición para todos los elementos sintácticos, o podrían ser diferentes posiciones para diferentes elementos sintácticos.

Región de referencia

[0123] Incluso cuando una posición asignada en la visualización de capa/referencia base está dentro de la imagen de la visualización de capa/referencia base, si está fuera de la región de referencia de la visualización de capa/referencia base, se considera que reside en una región no disponible, y el bloque identificado por tal posición se considera no disponible. En un modo de realización, la región de referencia se define como la región rectangular dentro de la imagen de la visualización de capa/referencia base, a partir de la cual se predice la capa/visualización actual (por ejemplo, disponible para uso para predicción entre capas o entre visualizaciones). La información de píxeles o bloques de la imagen fuera de la región de referencia no se utiliza para la predicción entre capas/de visualización. La "imagen" de la visualización de capa/referencia base utilizada en el presente documento puede referirse a una imagen de la capa base recortada para su visualización, que se ilustra con el rectángulo blanco 434 en la parte inferior de la **FIG. 4**. De forma alternativa, el término puede referirse a la imagen codificada (por ejemplo, descodificada) de la capa base, que se ilustra con las líneas discontinuas 436 en la parte inferior de la **FIG. 4**.

Localización del bloque disponible más cercano

[0124] En un modo de realización, cuando una ubicación de bloque o píxel se asigna a un bloque no disponible de la capa/visualización de base/referencia o se identifica fuera del límite izquierdo/derecho y/o superior/inferior de la región de referencia, la información del bloque disponible más cercano en el límite derecho/izquierdo y/o superior/inferior del bloque de referencia no disponible (es decir, el bloque correspondiente) se puede usar para la predicción entre capas/de visualización. Una forma de ubicar dicho bloque disponible más cercano es recortar las coordenadas de la ubicación de capa/visualización de base/referencia correspondiente en el rango del tamaño de la región de referencia de la capa/referencia base, donde la región de referencia puede ser la imagen de la capa base descodificada (436) o la imagen de la capa base recortada (434), como se muestra en la **FIG. 4**. Sin embargo, la región de referencia no está limitada a las representadas en la **FIG. 4**, y puede ser la región especificada por el codificador para la capa de mejora o el bloque específico en la capa de mejora.

[0125] Por ejemplo, las coordenadas horizontales de la ubicación de capa/visualización de base/referencia correspondiente pueden recortarse al rango entre 0 y (reference_region_width - 1), y las coordenadas verticales de la ubicación de capa/visualización de base/referencia correspondiente pueden recortarse a la rango entre 0 y (región de referencia altura - 1). Al recortar las coordenadas de la ubicación de capa/visualización de base/referencia correspondiente, se ubica el bloque de referencia disponible más cercano, y la información correspondiente a dicho bloque (por ejemplo, información de píxeles o información de movimiento) se puede usar para predecir el bloque actual en la capa de mejora.

[0126] La **FIG. 5** ilustra un procedimiento de ejemplo 500 para codificar información de vídeo, de acuerdo con un modo de realización de la presente divulgación. El procedimiento 500 puede ser realizado por uno o más componentes del codificador de vídeo 20 o del descodificador de vídeo 30, incluidos, entre otros, la unidad de predicción 121 y la unidad de compensación de movimiento 162. En algunos modos de realización, se pueden usar otros componentes para implementar uno o más de los pasos descritos en el presente documento. Por ejemplo, el procedimiento 500 se puede usar para realizar una o más de las acciones descritas con respecto a la **FIG. 4**. Aunque los pasos del procedimiento 500 se describen como realizados por un codificador, dichos pasos pueden ser realizados por un codificador (por ejemplo, codificador de vídeo 20) o un descodificador (por ejemplo, descodificador de vídeo 30).

[0127] El procedimiento 500 se inicia en el paso 505, por ejemplo, para codificar un bloque actual en la capa de mejora. En el paso 510, el codificador verifica si un bloque correspondiente no está disponible. Por ejemplo, el bloque correspondiente puede ser un bloque co-localizado del bloque actual en la capa de mejora, un bloque contiguo del bloque co-localizado, o cualquier otro bloque determinado por el codificador. Cualquiera de las

diversas técnicas de verificación de disponibilidad descritas en el presente documento puede usarse para determinar si dicho bloque correspondiente no está disponible. Si se determina que el bloque correspondiente no está disponible (SÍ en el paso 510), el codificador obtiene información basada en el bloque correspondiente de la capa base (paso 515). Por ejemplo, como se analiza en el presente documento, el codificador puede ubicar un bloque contiguo y puede obtenerse la información asociada con el bloque contiguo para su uso en la predicción entre capas. Basándose en la información obtenida, se realiza la predicción entre capas para predecir el bloque actual en la capa de mejora (paso 520). Por ejemplo, como se analizó anteriormente, la unidad de predicción 100 del codificador de vídeo 20 de la **FIG. 2** o la unidad de predicción 152 del descodificador de vídeo 30 de la **FIG. 2** puede realizar dicha predicción. El procedimiento 500 finaliza en el paso 525.

Elementos sintácticos e información residual

[0128] En un modo de realización, si un bloque se asigna a un bloque no disponible de capa/visualización de base/referencia o se identifica fuera del límite izquierdo/derecho y/o superior/inferior de la región de referencia, la predicción entre capas para los elementos sintácticos y/o la información residual está deshabilitada para dicho bloque. De forma alternativa, en lugar de deshabilitar la predicción entre capas para dicho bloque, los valores sintácticos predeterminados pueden usarse para la predicción entre capas. Por ejemplo, se puede usar un campo de movimiento cero (por ejemplo, un MV es igual a cero y el índice de referencia es igual a cero), y el modo intra se puede establecer en la dirección de predicción de DC.

[0129] En otro modo de realización, los elementos sintácticos de un bloque disponible contiguo pueden sustituirse por los elementos sintácticos del bloque correspondiente no disponible. En el caso de utilizar la información del bloque disponible contiguo, el campo de movimiento utilizado para los sub-bloques no disponibles puede comprimirse (por ejemplo, basándose en el tamaño de 16x16 como en HEVC), especialmente si la trama de la capa base se inserta en la lista de imágenes de referencia.

Tamaño de unidad de codificación (SCU) más pequeño

[0130] El tamaño de SCU de la capa de mejora y de base se puede restringir para evitar regiones no disponibles para la predicción de sintaxis entre capas. Por ejemplo, el tamaño de la SCU de la capa de mejora puede configurarse para que sea más pequeño o igual que el tamaño de la SCU de la capa base. Además, la relación de aspecto de escalabilidad espacial se puede tener en cuenta al definir la restricción en los tamaños de SCU.

Tamaño de relleno de límite de imagen

[0131] Los tamaños de relleno del límite de la imagen de la capa de mejora y la capa base pueden restringirse para evitar regiones no disponibles para la predicción de sintaxis entre capas. Por ejemplo, el tamaño de relleno de la capa de mejora puede configurarse para que sea más pequeño o igual que el tamaño de relleno de la capa base. Más específicamente, los tamaños de relleno de cada límite de imagen (es decir, los límites izquierdo, superior, derecho e inferior en la imagen de la capa de mejora) pueden hacerse más pequeños o iguales a los de la imagen de la capa base. Además, la relación de resolución se puede tener en cuenta al definir la restricción en los tamaños de relleno de límite de imagen para el caso de escalabilidad espacial.

Disponibilidad de información de campo de movimiento

[0132] Durante la predicción de movimiento entre visualizaciones/capas, si la esquina inferior derecha del bloque de capa de mejora está fuera de la imagen, se considera que el campo de movimiento no está disponible. La imagen como se usa en el presente documento puede referirse a la imagen de la capa base descodificada o la imagen de la capa base recortada, como se ilustra en la **FIG. 4**. En tal caso, se puede considerar que el campo de movimiento no está disponible incluso si el bloque de la capa base correspondiente está dentro de la imagen de la capa base. De forma alternativa, en otro modo de realización, si la parte representativa del bloque de capa de mejora está fuera de la imagen, pero el bloque de capa base correspondiente a la parte representativa del bloque de capa de mejora está dentro de la imagen de capa base, entonces el campo de movimiento del bloque de capa base (escalado de acuerdo con la relación de escalabilidad si es necesario) puede asignarse para el bloque de capa de mejora.

Disponibilidad de información de campo de movimiento

[0133] En otro modo de realización, en lugar de la esquina inferior derecha del bloque, se pueden utilizar otros puntos o cualquier sub-bloque dentro del bloque como el punto representativo. Por ejemplo, se puede usar el punto central de la capa base o el bloque de capa de mejora, o uno de los sub-bloques 4x4 del centro del bloque de capa base o de mejora. Si este punto o sub-bloque representativo está fuera de la imagen BL o EL correspondiente, se puede aplicar una de las técnicas descritas en la presente divulgación. La imagen aquí puede ser una imagen descodificada o una imagen recortada.

Coordenadas de recorte antes de la asignación de posición

[0134] En un modo de realización, se puede aplicar una restricción de rango a la ubicación en la imagen de la capa de mejora actual antes de realizar el proceso de asignación de posición para determinar la ubicación del bloque correspondiente en la capa/visualización de base/referencia. Por ejemplo, las coordenadas de la ubicación actual pueden recortarse en el rango de tamaños de imagen de la imagen descodificada actual antes de realizar el proceso de asignación de posición para obtener la ubicación correspondiente en la capa/visualización de base/referencia. En otro ejemplo, las coordenadas de la ubicación actual se pueden recortar en el rango de tamaños de imagen de la imagen recortada actual antes de realizar el proceso de asignación de posición para obtener la ubicación correspondiente en la capa/visualización de base/referencia.

Imagen recortada e imagen descodificada

[0135] En una visualización más detallada, la región fuera de la trama puede dividirse en dos partes: la primera parte (por ejemplo, el área 418 en la **FIG. 4**) que está relacionada con una parte extendida de la trama de la capa base debido a la alineación con el tamaño de la SCU de la capa base 432 mostrada en la **FIG. 4**, y la segunda parte (por ejemplo, el área 420 en la **FIG. 4**) que no tiene ninguna parte correspondiente en la trama de la capa base, que se muestra con el sombreado en la **FIG. 4**.

[0136] Las técnicas descritas anteriormente se pueden aplicar a las dos áreas 418 y 420. De forma alternativa, se puede considerar que el área 418 está disponible, ya que la información se puede obtener a partir de la trama de la capa base extendida (es decir, la trama codificada 436 de la capa base), y por lo tanto, las técnicas descritas en esta divulgación solo se pueden aplicar al área 420.

[0137] En un modo de realización, la indisponibilidad de un bloque se puede verificar directamente de acuerdo con la ubicación asignada en la imagen de la capa/visualización de base/referencia. Como se muestra en la **FIG. 4**, una técnica que se puede usar es considerar que cualquier posición asignada dentro del tamaño de trama real de la imagen de capa/visualización de base/referencia esté disponible y la parte más allá del tamaño de trama real no esté disponible. De forma alternativa, cualquier posición asignada dentro del tamaño de trama codificada de la imagen de capa/visualización de base/referencia puede considerarse disponible, y la parte más allá del tamaño de trama codificada puede considerarse como no disponible.

Otras aplicaciones

[0138] En la descripción anterior, aunque se utilizó la predicción entre capas de movimiento como un ejemplo, los mismos principios y técnicas también son aplicables a otros elementos sintácticos como el modo intra, modo de predicción, indicador de división, indicador de omisión, etc.

[0139] De forma alternativa o adicional, los procedimientos mencionados anteriormente pueden aplicarse a la predicción de píxeles entre capas (incluida la predicción Intra-BL o Texture-BL, la predicción de residuos entre capas, la predicción intra del dominio de diferencia, la predicción inter de dominio de diferencia, etc.). Por ejemplo, para aquellos bloques o parte de los bloques que se encuentran en una región no disponible, la predicción de píxeles entre capas se puede desactivar, o los píxeles se pueden inicializar con valores predeterminados, por ejemplo, 0 o $(1 \ll (\text{profundidad de bits} - 1))$, donde profundidad de bits puede ser 8, 10, 12 bits y así sucesivamente.

[0140] De forma alternativa, los píxeles se pueden rellenar utilizando los píxeles adyacentes disponibles. El procedimiento de relleno de píxeles, por ejemplo, puede ser el mismo procedimiento que se utiliza para el relleno de píxeles para la interpolación. Sin embargo, la cantidad de píxeles que deben extenderse puede ser diferente de los necesarios para fines de interpolación. Otros procedimientos de relleno de píxeles también son aplicables sin restricción.

[0141] Para el caso de predicción de píxeles entre capas, el procedimiento de comprobación de indisponibilidad mencionado anteriormente se puede realizar píxel a píxel. Es decir, para cada píxel, su posición colocada en la imagen de capa/visualización de base/referencia se utiliza para verificar la disponibilidad con la regla mencionada anteriormente. Cuando un píxel se asigna a una posición no disponible de la capa/visualización de base/referencia, el píxel de la posición disponible más cercana al límite de la derecha/izquierda y/o superior/inferior de la posición asignada se usa para la predicción entre capas/de visualización.

[0142] Como se muestra en la **FIG. 4**, la esquina superior izquierda con coordenadas (0,0) está alineada para las capas base y de mejora. Sin embargo, la solución se puede aplicar en una situación en la que no existe tal alineación de esquina. En tal situación, hay otra área no disponible fuera de la trama de la capa de mejora desde los lados izquierdo y superior similar a las homólogas derecha e inferior mostradas en la **FIG. 4**.

[0143] La información y las señales divulgadas en el presente documento pueden representarse usando cualquiera entre varias tecnologías y técnicas diferentes. Por ejemplo, los datos, las instrucciones, los comandos, la información, las señales, los bits, los símbolos y los chips que puedan haberse mencionado a lo largo de la

descripción anterior pueden representarse mediante tensiones, corrientes, ondas electromagnéticas, campos o partículas magnéticos, campos o partículas ópticos o cualquier combinación de los mismos.

5 **[0144]** Los diversos bloques lógicos, unidades, módulos, circuitos y pasos de algoritmo ilustrativos, descritos en relación con los modos de realización divulgados en el presente documento pueden implementarse como hardware electrónico, software informático o combinaciones de ambos. Para ilustrar claramente esta intercambiabilidad de hardware y software, anteriormente se han descrito, en general, diversos componentes, bloques, módulos, circuitos y pasos ilustrativos en términos de su funcionalidad. Que dicha funcionalidad se implemente como hardware o software depende de la aplicación particular y de las restricciones de diseño impuestas en el sistema global. Los expertos en la técnica pueden implementar la funcionalidad descrita de formas diversas para cada aplicación particular, pero no debería interpretarse que dichas decisiones de implementación suponen apartarse del alcance de la presente invención.

15 **[0145]** Las técnicas descritas en el presente documento pueden implementarse en hardware, software, firmware o en cualquier combinación de lo anterior. Dichas técnicas pueden implementarse en cualquiera entre una variedad de dispositivos tales como ordenadores de propósito general, equipos manuales de dispositivos de comunicación inalámbrica o dispositivos de circuitos integrados que tienen múltiples usos, incluyendo su aplicación en equipos manuales de dispositivos de comunicación inalámbrica y otros dispositivos. Cualquier característica descrita como unidades, módulos o componentes puede implementarse conjuntamente en un dispositivo lógico integrado o por separado, como dispositivos lógicos discretos pero interoperables. Si se implementan en software, las técnicas pueden realizarse, al menos en parte, mediante un medio de almacenamiento de datos legible por ordenador que comprenda código de programa que incluye instrucciones que, cuando se ejecutan, realizan uno o más de los procedimientos descritos anteriormente. El medio de almacenamiento de datos legible por ordenador puede formar parte de un producto de programa informático, que puede incluir materiales de embalaje. El medio legible por ordenador puede comprender memoria o medios de almacenamiento de datos, tales como memoria de acceso aleatorio (RAM), tal como memoria de acceso aleatorio dinámica síncrona (SDRAM), memoria de solo lectura (ROM), memoria de acceso aleatorio no volátil (NVRAM), memoria de solo lectura programable y borrable eléctricamente (EEPROM), memoria FLASH, medios de almacenamiento de datos magnéticos u ópticos, y similares. Las técnicas se pueden realizar, adicionalmente o de forma alternativa, al menos en parte, por un medio de comunicación legible por ordenador que transporta o comunica código de programa en forma de instrucciones o estructuras de datos y a las que se puede acceder, leer y/o ejecutar por medio de un ordenador, tales como señales u ondas propagadas.

35 **[0146]** El código de programa puede ser ejecutado por un procesador, que puede incluir uno o más procesadores, tales como uno o más procesadores de señales digitales (DSP), microprocesadores de propósito general, circuitos integrados específicos de la aplicación (ASIC), matrices de puertas programables en el terreno (FPGA) u otros circuitos lógicos equivalentes, integrados o discretos. Un procesador de este tipo puede estar configurado para realizar cualquiera de las técnicas descritas en esta divulgación. Un procesador de propósito general puede ser un microprocesador, pero, de forma alternativa, el procesador puede ser cualquier procesador, controlador, micro-controlador o máquina de estados convencional. Un procesador también puede implementarse como una combinación de dispositivos informáticos, por ejemplo, una combinación de un DSP y un microprocesador, una pluralidad de microprocesadores, uno o más microprocesadores junto con un núcleo de DSP o cualquier otra configuración de este tipo. Por consiguiente, el término "procesador", como se usa en el presente documento, puede referirse a cualquier estructura anterior, cualquier combinación de la estructura anterior, o cualquier otra estructura o aparato adecuados para la implementación de las técnicas descritas en el presente documento. Además, en algunos aspectos, la funcionalidad descrita en el presente documento puede proporcionarse dentro de módulos de software o módulos de hardware dedicados configurados para la codificación y la descodificación, o incorporados en un codificador-descodificador de vídeo combinado (CODEC). También, las técnicas se podrían implementar totalmente en uno o más circuitos o elementos lógicos.

50 **[0147]** Las técnicas de la presente divulgación se pueden implementar en una amplia variedad de dispositivos o aparatos, incluidos un equipo manual inalámbrico, un circuito integrado (CI) o un conjunto de CI (por ejemplo, un conjunto de chips). Diversos componentes, módulos o unidades se describen en esta divulgación para enfatizar aspectos funcionales de dispositivos configurados para realizar las técnicas divulgadas, pero no requieren necesariamente su realización mediante diferentes unidades de hardware. En cambio, como se ha descrito anteriormente, diversas unidades se pueden combinar en una unidad de hardware de códec, o ser proporcionadas por un grupo de unidades de hardware interoperativas, incluyendo uno o más procesadores, como se ha descrito anteriormente, conjuntamente con software y/o firmware adecuados.

REIVINDICACIONES

1. Un procedimiento de codificación de información de vídeo, el procedimiento que comprende:
 - 5 recibir información de vídeo asociada con una capa de mejora (400) y una capa base (430), la capa base que comprende una región de referencia y una región fuera de la región de referencia, en el que la región de referencia es una región recortada de la capa base, en el que la región de referencia está disponible para la predicción entre capas de la capa de mejora, en el que la región fuera de la región de referencia no está disponible para la predicción entre capas de la capa de mejora, y en la que la capa de mejora comprende bloques que corresponden a bloques respectivos en la capa base;
 - 10 codificar la capa base; y
 - 15 codificar la capa de mejora, en el que codificar la capa de mejora comprende:
 - usar la región de referencia de la capa base para la predicción entre capas de la capa de mejora;
 - 20 asignar coordenadas que identifican un bloque de capa de mejora actual a las coordenadas correspondientes de una posición asignada en la capa base;
 - determinar si las coordenadas asignadas están fuera de la región de referencia de la capa base; y
 - 25 en respuesta a la determinación de que las coordenadas asignadas están fuera de la región de referencia de la capa base, determinar que el bloque de capa base correspondiente no está disponible para predecir el bloque actual, y deshabilitar la predicción de elementos sintácticos entre capas para el bloque de capa de mejora actual.

2. Un procedimiento de decodificación de información de vídeo, el procedimiento que comprende:
 - 30 recibir un flujo de bits de vídeo codificado que comprende información de vídeo codificada asociada con una capa de mejora (400) y una capa base (430);
 - 35 descodificar la capa base, en el que la capa base descodificada comprende una región de referencia y una región fuera de la región de referencia, en el que la región de referencia es una región recortada de la capa base, en el que la región de referencia está disponible para la predicción entre capas de la capa de mejora, en el que la región fuera de la región de referencia no está disponible para la predicción entre capas de la capa de mejora; y
 - 40 descodificar la capa de mejora, la capa de mejora que tiene bloques que corresponden a bloques respectivos en la capa base, en el que la decodificación de la capa de mejora comprende:
 - 45 usar la región de referencia de la capa base para la predicción entre capas de la capa de mejora;
 - asignar coordenadas que identifican un bloque de capa de mejora actual a las coordenadas correspondientes de una posición asignada en la capa base;
 - 50 determinar si las coordenadas asignadas están fuera de la región de referencia de la capa base; y
 - en respuesta a la determinación de que las coordenadas asignadas están fuera de la región de referencia de la capa base, determinar que el bloque de capa base correspondiente no está disponible para predecir el bloque actual, y deshabilitar la predicción de elementos sintácticos entre capas para el bloque de capa de mejora actual.

3. El procedimiento según la reivindicación 2, que comprende además recortar la capa base descodificada para ajustarse al tamaño de visualización de un dispositivo, en el que la región de referencia de la capa base es la capa base recortada.

4. El procedimiento según la reivindicación 1 o la reivindicación 2, en el que un tamaño de una unidad de codificación más pequeña (SCU) de la capa de mejora es diferente del tamaño de una SCU de la capa base.

5. El procedimiento según la reivindicación 1 o la reivindicación 2, en el que un tamaño de relleno de límite de cada límite de imagen de la capa de mejora es diferente de un tamaño de relleno de límite de cada límite de imagen correspondiente de la capa base.

6. Un medio no transitorio legible por ordenador que tenga instrucciones almacenadas en el mismo que, al ejecutarse, hacen que un aparato realice un procedimiento de acuerdo con una cualquiera de las reivindicaciones 1-7.
- 5 7. Un dispositivo de codificación de vídeo configurado para codificar información de vídeo, el dispositivo de codificación de vídeo que comprende:
- 10 medios para almacenar información de vídeo asociada con una capa de mejora (400) y una capa base (430), la capa base que comprende una región de referencia y una región fuera de la región de referencia, en el que la región de referencia es una región recortada de la capa base, en el que la región de referencia está disponible para la predicción entre capas de la capa de mejora, en el que la región fuera de la región de referencia no está disponible para la predicción entre capas de la capa de mejora, y en la que la capa de mejora comprende bloques que corresponde a bloques respectivos en la capa base;
- 15 medios para codificar la capa base; y
- medios para codificar la capa de mejora, en el que los medios para codificar la capa de mejora comprenden:
- 20 medios para usar la región de referencia de la capa base para la predicción entre capas de la capa de mejora;
- 25 medios para asignar coordenadas que identifican un bloque de capa de mejora actual a las coordenadas correspondientes de una posición asignada en la capa base;
- medios para determinar si las coordenadas asignadas están fuera de la región de referencia de la capa base; y
- 30 medios para determinar que el bloque de capa base correspondiente no está disponible para predecir el bloque actual, y deshabilitar la predicción entre capas de elementos sintácticos para el bloque de capa de mejora actual, en respuesta a la determinación de que las coordenadas asignadas están fuera de la región de referencia de la capa base.
- 35 8. Un dispositivo de descodificación de vídeo configurado para descodificar información de vídeo, el dispositivo de descodificación de vídeo que comprende:
- 40 medios para recibir un flujo de bits de vídeo codificado que comprende información de vídeo asociada con una capa de mejora (400) y una capa base (430);
- 45 medios para descodificar la capa base, en el que la capa base descodificada comprende una región de referencia y una región fuera de la región de referencia, en el que la región de referencia es una región recortada de la capa base, en el que la región de referencia está disponible para la predicción entre capas de la capa de mejora, en la que la región fuera de la región de referencia no está disponible para la predicción entre capas de la capa de mejora; y
- medios para descodificar la capa de mejora, la capa de mejora que tiene bloques que corresponden a bloques respectivos en la capa base, en el que los medios para descodificar la capa de mejora comprenden:
- 50 medios para usar la región de referencia de la capa base para la predicción entre capas de la capa de mejora;
- 55 medios para asignar coordenadas que identifican un bloque de capa de mejora actual a las coordenadas correspondientes de una posición asignada en la capa base;
- medios para determinar si las coordenadas asignadas están fuera de la región de referencia de la capa base; y
- 60 medios para determinar que el bloque de capa base correspondiente no está disponible para predecir el bloque actual, y deshabilitar la predicción entre capas de elementos sintácticos para el bloque de capa de mejora actual, en respuesta a la determinación de que las coordenadas asignadas están fuera de la región de referencia de la capa base.

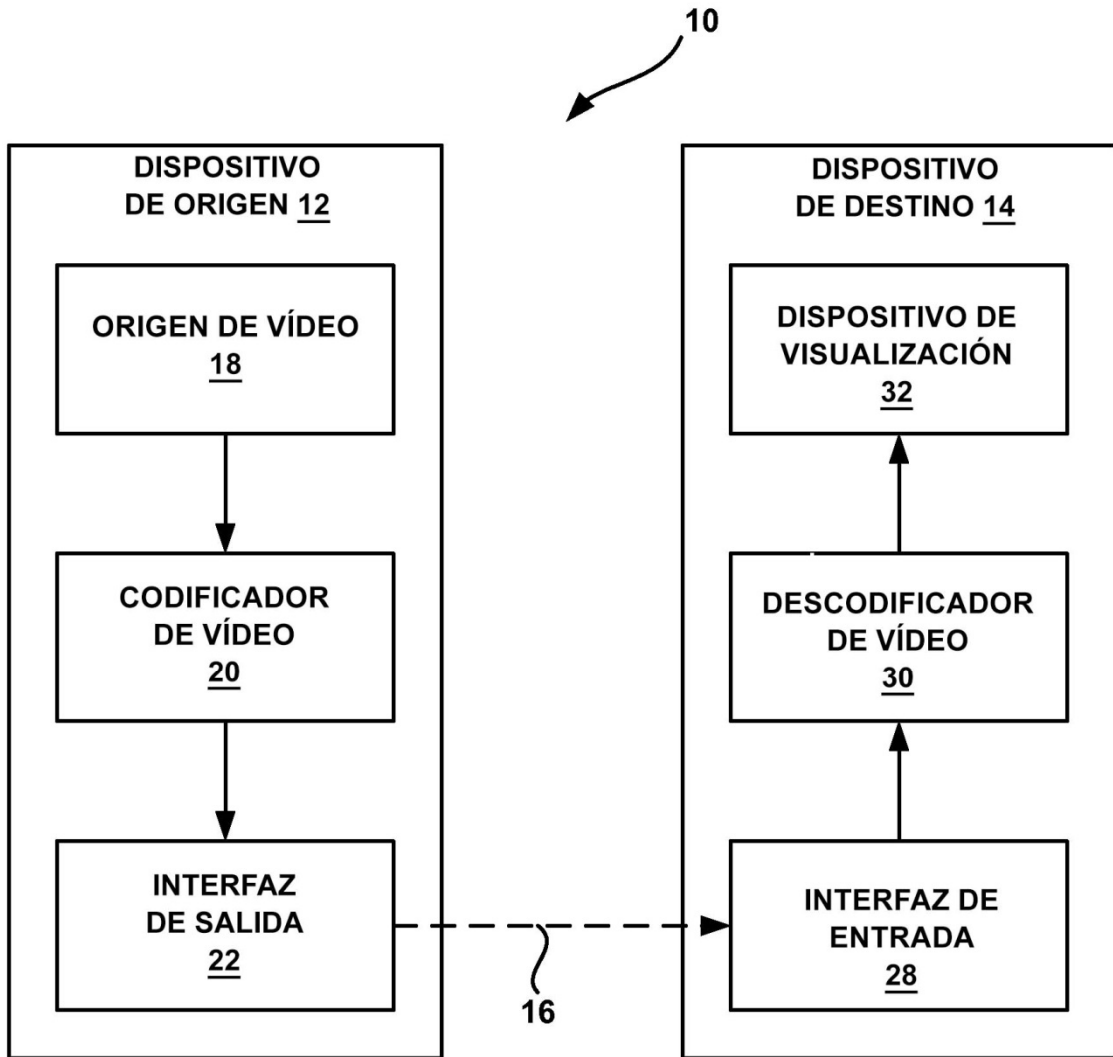


FIG. 1

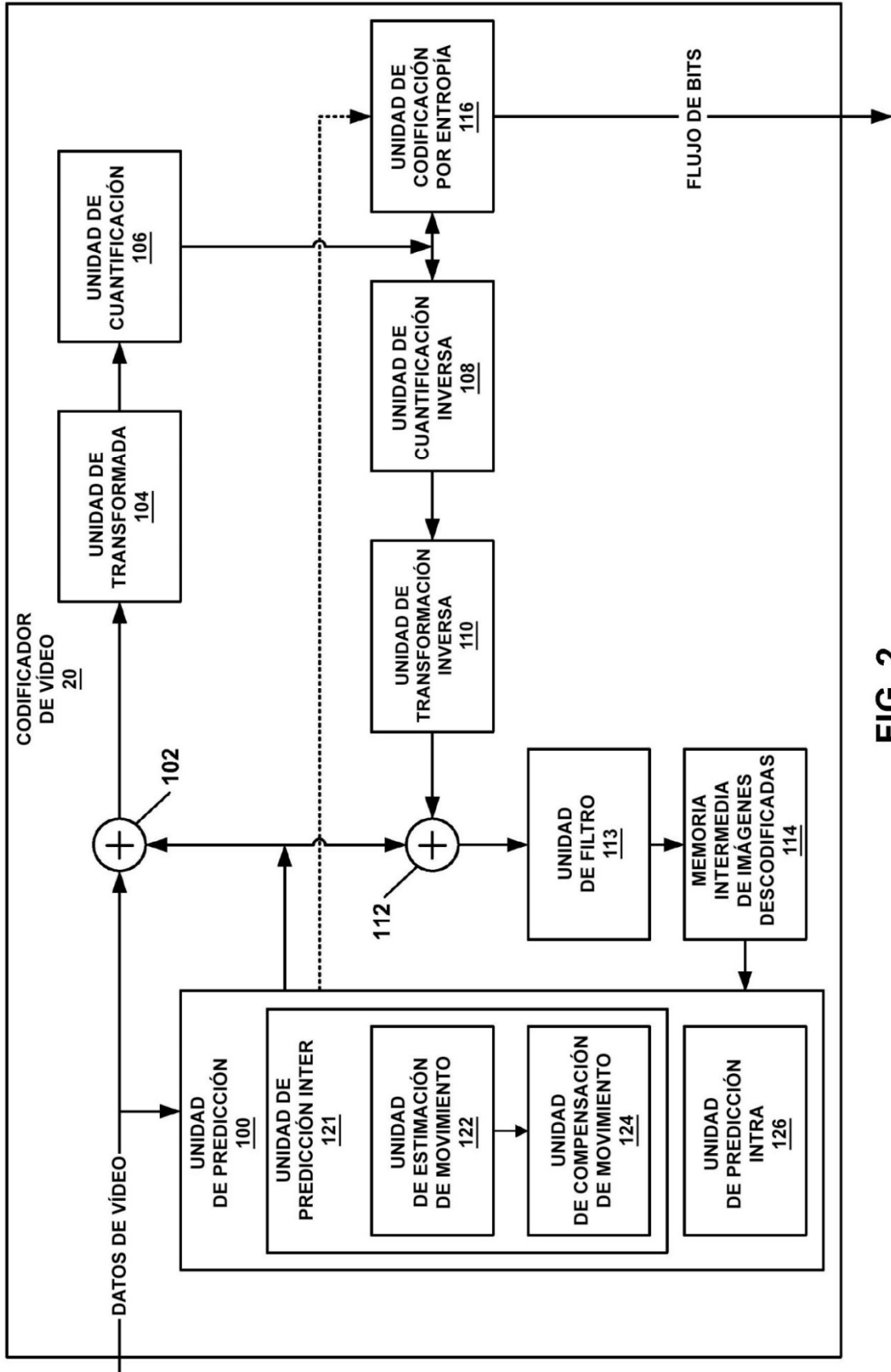


FIG. 2

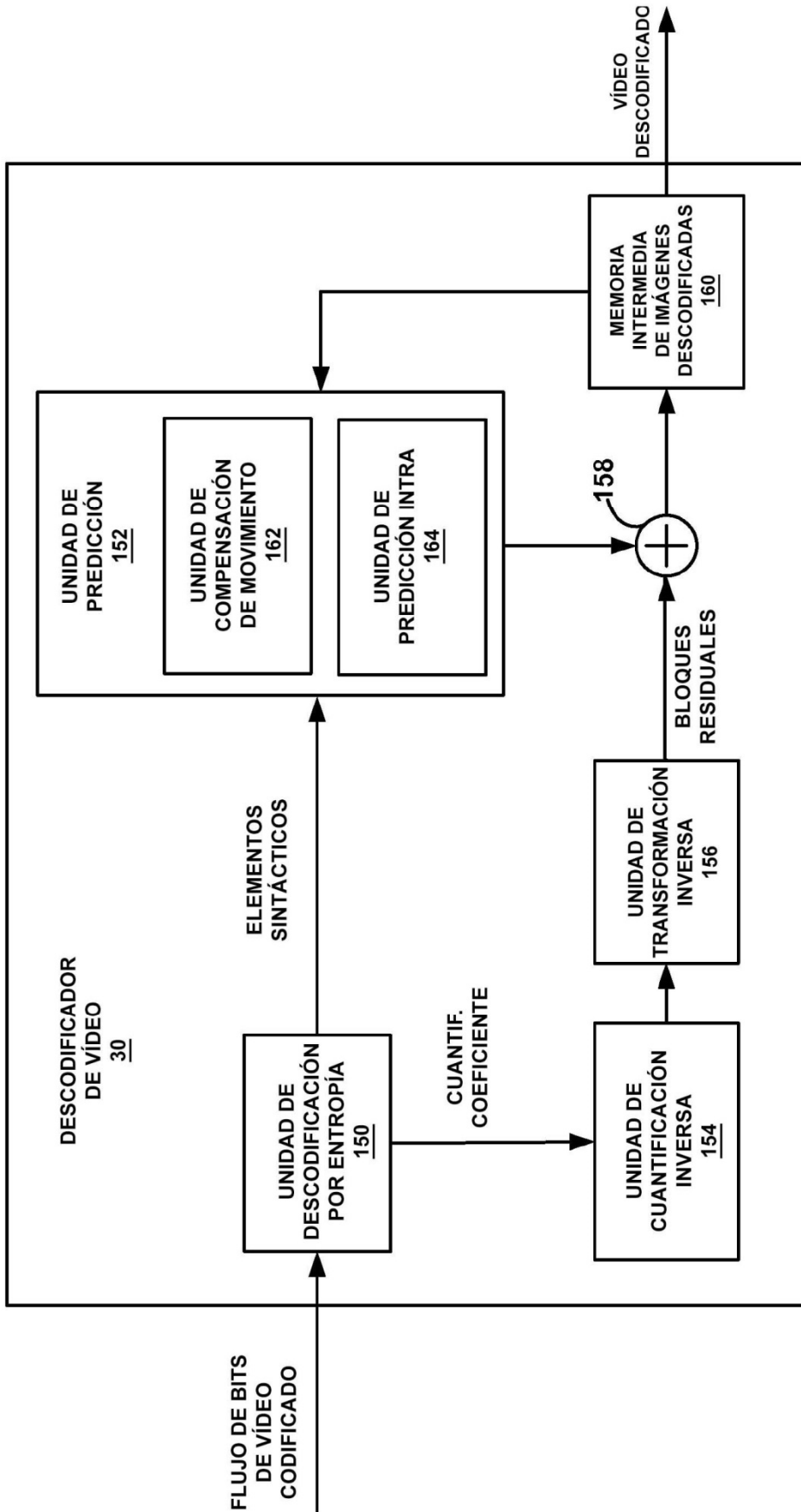


FIG. 3

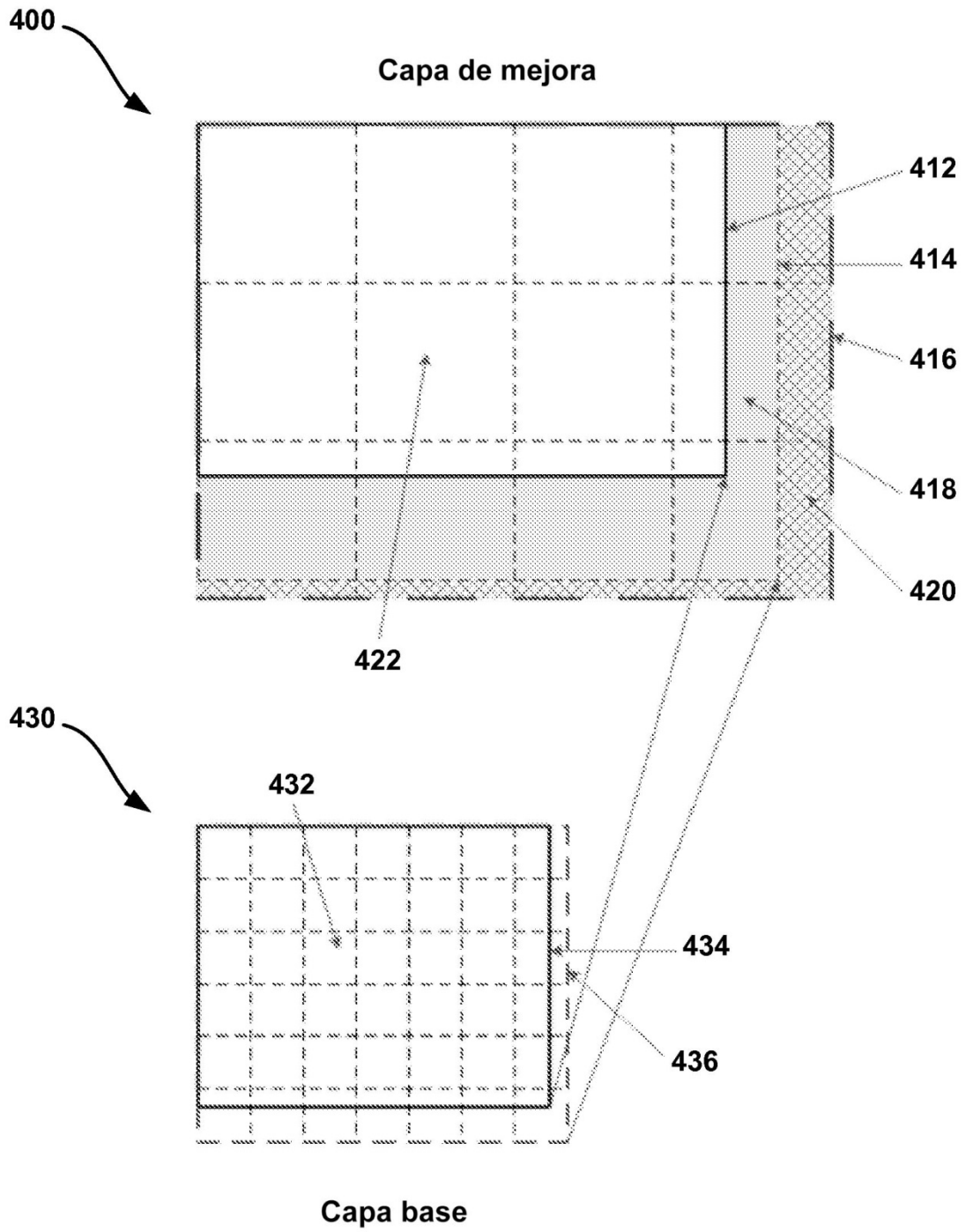


FIG. 4

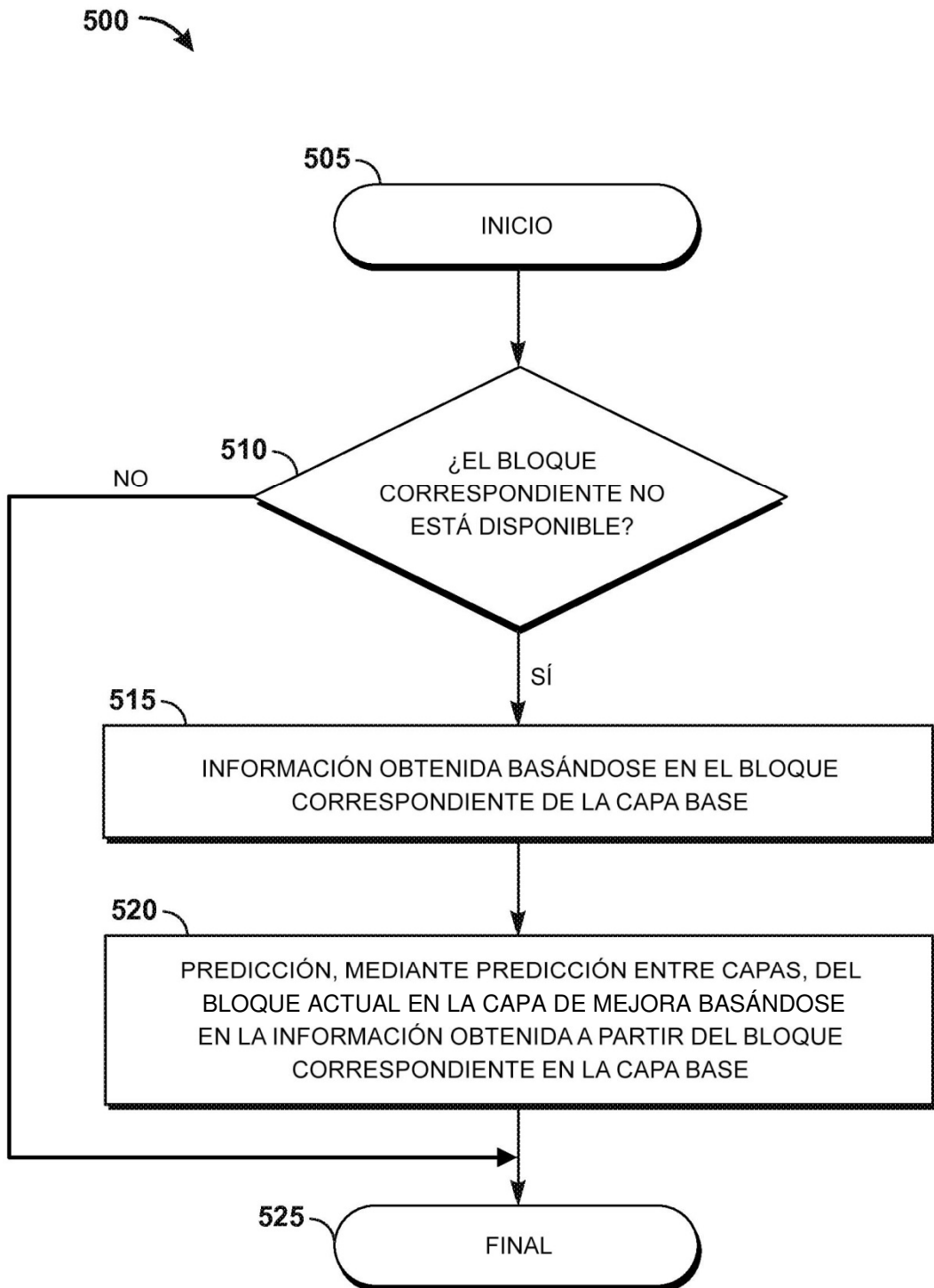


FIG. 5