

19



OFICINA ESPAÑOLA DE  
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 715 980**

51 Int. Cl.:

**H04N 19/503** (2014.01)

**H04N 19/70** (2014.01)

**H04N 19/58** (2014.01)

12

## TRADUCCIÓN DE PATENTE EUROPEA

T3

86 Fecha de presentación y número de la solicitud internacional: **30.11.2012 PCT/US2012/067362**

87 Fecha y número de publicación internacional: **06.06.2013 WO13082471**

96 Fecha de presentación y número de la solicitud europea: **30.11.2012 E 12798575 (2)**

97 Fecha y número de publicación de la concesión europea: **19.12.2018 EP 2786569**

54 Título: **Codificación de valores de recuento del orden de imágenes que identifican tramas de referencia a largo plazo**

30 Prioridad:

**02.12.2011 US 201161566359 P**

**20.01.2012 US 201261589295 P**

**16.04.2012 US 201261624933 P**

**29.11.2012 US 201213689559**

45 Fecha de publicación y mención en BOPI de la traducción de la patente:  
**07.06.2019**

73 Titular/es:

**QUALCOMM INCORPORATED (100.0%)**

**5775 Morehouse Drive**

**San Diego, CA 92121-1714, US**

72 Inventor/es:

**WANG, YE-KUI;**

**RAMASUBRAMONIAN, ADARSH KRISHNAN y**

**CHEN, YING**

74 Agente/Representante:

**FORTEA LAGUNA, Juan José**

ES 2 715 980 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

Codificación de valores de recuento del orden de imágenes que identifican tramas de referencia a largo plazo

### 5 CAMPO TÉCNICO

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

### ANTECEDENTES

10

[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. 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). El más reciente borrador conjunto de la MVC se describe en "Advanced video coding for generic audiovisual services" ["Codificación de vídeo avanzada para servicios audiovisuales genéricos"], Recomendación H.264 de la ITU-T, marzo de 2010.

15

20

25

[0003] Las técnicas de compresión de vídeo realizan predicción espacial y/o predicción temporal para reducir o eliminar la redundancia inherente a las secuencias de vídeo. Para la codificación de vídeo basada en bloques, una trama o un fragmento de vídeo puede dividirse en bloques. Cada bloque se puede dividir adicionalmente. Los bloques de una trama o fragmento intra-codificado (I) se codifican usando predicción espacial con respecto a bloques contiguos. Los bloques de una trama o fragmento inter-codificado (P o B) pueden usar predicción espacial con respecto a bloques contiguos de la misma trama o fragmento, o predicción temporal con respecto a otras tramas de referencia.

30

35

[0004] Se han dedicado esfuerzos para elaborar nuevas normas de codificación de vídeo basándose en H.264/AVC. Una de estas normas es la norma de codificación de vídeo ajustable a escala (SVC), que es la extensión ajustable a escala para la H.264/AVC. Otra norma es la norma de codificación de vídeo de múltiples visualizaciones (MVC), que se ha convertido en la extensión de múltiples visualizaciones para H.264/AVC. Un borrador conjunto de la MVC se describe en el documento JVT-AB204, "Joint Draft 8.0 on Multiview Video Coding" ["Borrador conjunto 8.0 sobre codificación de vídeo de múltiples visualizaciones", 28.ª reunión de JVT, Hannover, Alemania, julio de 2008, disponible en [http://wftp3.itu.int/av-arch/jvt-site/2008\\_07\\_Hannover/JVT-AB204.zip](http://wftp3.itu.int/av-arch/jvt-site/2008_07_Hannover/JVT-AB204.zip).

40

[0005] Una versión de la norma AVC se describe en JVT-AD007, "Editors' draft revision to ITU-T Rec. H.264 | ISO/IEC 14496-10 Advanced Video Coding - in preparation for ITU-T SG 16 AAP Consent (in integrated form)" ["Revisión del borrador de los editores de Rec. ITU-T. H.264 | ISO/IEC 14496-10 Codificación de vídeo avanzada: en preparación para el consentimiento de ITU-T SG 16 AAP (en forma integrada)", 30.ª reunión de JVT, Ginebra, Suiza, febrero de 2009", disponible en [http://wftp3.itu.int/av-arch/jvt-site/2009\\_01\\_Geneva/JVT-AD007.zip](http://wftp3.itu.int/av-arch/jvt-site/2009_01_Geneva/JVT-AD007.zip). Este documento integra la SVC y la MVC en la especificación de la AVC.

45

50

[0006] "AHG21: Comments on Signaling of Reference Pictures" ["AHG21: Comentarios sobre señalización de imágenes de referencia"] de Sachin Deshpande et al, 8.ª reunión de JCT-VC, San José, febrero de 2012 disponible en <http://wftp3.itu.int/av-arch/jctvc-site> se refiere a la sintaxis para señalización de imágenes de referencia a largo plazo que utilizan bits menos significativos y bits más significativos de un recuento del orden de imágenes delta.

### SUMARIO

55

[0007] En general, se describen técnicas que proporcionan diversos procedimientos para obtener un conjunto de imágenes de referencia (RPS) y señalización de imágenes de referencia a largo plazo (LTRP) para incluirlas en el RPS de una imagen codificada para usar cuando se realizan aspectos de predicción inter de codificación de vídeo.

[0008] En un ejemplo, un procedimiento de codificación de datos de vídeo se establece en una reivindicación independiente.

60

[0009] En otro ejemplo, un dispositivo de codificación de vídeo para codificar datos de vídeo comprende medios como se establece en la reivindicación independiente 4.

[0010] En otro ejemplo, un medio de almacenamiento legible por ordenador no transitorio ha almacenado en el mismo las instrucciones que se establecen en la reivindicación independiente 16.

65

[0011] En otro ejemplo, un procedimiento para descodificar datos de vídeo codificados se establece en la reivindicación independiente 5.

[0012] En otro ejemplo, un dispositivo de descodificación de vídeo para descodificar datos de vídeo codificados comprende medios como se establece en la reivindicación independiente 15.

[0013] Los detalles de uno o más aspectos de la divulgación se exponen en los dibujos adjuntos y en la descripción siguiente. La presente invención se define en las reivindicaciones adjuntas. Los modos de realización de la invención se pueden encontrar en relación con la Fig. 6 y la Fig. 7.

## BREVE DESCRIPCIÓN DE LOS DIBUJOS

[0014]

La FIG. 1 es un diagrama de bloques que ilustra un sistema de codificación y descodificación de vídeo de ejemplo que puede utilizar técnicas para codificación de datos de vídeo.

La FIG. 2 es un diagrama de bloques que ilustra un ejemplo de un codificador de vídeo que puede implementar técnicas para codificar datos de vídeo.

La FIG. 3 es un diagrama de bloques que ilustra un ejemplo de descodificador de vídeo, que descodifica una secuencia de vídeo codificada.

La FIG. 4 es un diagrama de flujo que ilustra la operación a modo de ejemplo de un codificador de vídeo al realizar el primer aspecto de las técnicas descritas en esta divulgación.

La FIG. 5 es un diagrama de flujo que ilustra el funcionamiento a modo de ejemplo de un descodificador de vídeo al realizar el primer aspecto de las técnicas descritas en esta divulgación.

La FIG. 6 es un diagrama de flujo que ilustra el funcionamiento a modo de ejemplo de un codificador de vídeo al realizar el segundo aspecto de las técnicas descritas en esta divulgación.

La FIG. 7 es un diagrama de flujo que ilustra el funcionamiento a modo de ejemplo de un descodificador de vídeo al realizar el segundo aspecto de las técnicas descritas en esta divulgación.

## DESCRIPCIÓN DETALLADA

[0015] En general, se describen técnicas que proporcionan diversos procedimientos para obtener un conjunto de imágenes de referencia (RPS) y señalización de imágenes de referencia a largo plazo (LTRP) para incluirlas en el RPS de una imagen codificada para usar cuando se realizan aspectos de predicción inter de codificación de vídeo. Por ejemplo, un codificador de vídeo (que es un término que puede referirse a uno o ambos de un codificador de vídeo y un descodificador de vídeo) incluye una memoria intermedia de imágenes descodificadas (DPB). La DPB almacena imágenes de referencia, que son imágenes que se pueden usar para realizar predicción inter de una imagen. En otras palabras, el codificador de vídeo puede predecir una imagen basándose en una o más imágenes de referencia almacenadas en la DPB.

[0016] El descodificador de vídeo también puede encargarse de construir listas de imágenes de referencia que indiquen qué imágenes de referencia se utilizan para fines de predicción inter. Dos de estas listas de imágenes de referencia se denominan Lista 0 y Lista 1, respectivamente. El descodificador de vídeo primero emplea técnicas de construcción predeterminadas para construir la Lista 0 y la Lista 1 (por ejemplo, esquemas de construcción preconfigurados para construir la Lista 0 y la Lista 1). Opcionalmente, después de construir la Lista 0 y la Lista 1 iniciales, el descodificador puede descodificar los elementos sintácticos, cuando están presentes, que ordenan al descodificador de vídeo modificar la Lista 0 y la Lista 1 iniciales.

[0017] El codificador de vídeo puede señalar elementos sintácticos que son indicativos del (de los) identificador(es) de las imágenes de referencia en la DPB, y el codificador de vídeo también puede señalar elementos sintácticos que incluyen índices, dentro de la Lista 0, Lista 1, o tanto la Lista 0 como la Lista 1, que indican qué imagen o imágenes de referencia usar para descodificar un bloque codificado de una imagen actual. El descodificador de vídeo, a su vez, utiliza el identificador recibido para identificar el valor o los valores de índice para una imagen de referencia o imágenes de referencia enumeradas en la Lista 0, la Lista 1 o tanto la Lista 0 como la Lista 1. Desde el (los) valor (es) de índice, así como el (los) identificador (es) de la imagen de referencia o imágenes de referencia, el codificador de vídeo recupera la imagen de referencia o las imágenes de referencia de la DPB, y descodifica el bloque codificado de la imagen actual.

[0018] En un conjunto de parámetros de imagen (PPS) o un conjunto de parámetros de secuencia (SPS) asociado con una cabecera de fragmento de una imagen actual, el codificador de vídeo puede señalar el RPS. El

RPS de una imagen actual incluye información de identificación para imágenes de referencia que se puede usar para predecir la imagen actual y las imágenes que se pueden usar para predecir imágenes después de la imagen actual en orden de decodificación. Solo las imágenes de referencia en el RPS pueden incluirse en la Lista 0 o la Lista 1.

**[0019]** La información de identificación para las imágenes de referencia puede incluir uno o más valores de recuento del orden de imágenes (POC). Los valores de POC indican el orden en el que se emiten o muestran las imágenes dentro de una secuencia de vídeo codificada (es decir, el orden de visualización de las imágenes). Por ejemplo, una imagen con un valor POC más bajo se muestra antes que una imagen con un valor POC más alto en la misma secuencia de vídeo codificada.

**[0020]** El primer aspecto de las técnicas descritas en esta divulgación se refiere a técnicas para señalar directamente imágenes de referencia a largo plazo para una imagen actual. Por ejemplo, las imágenes de referencia pueden clasificarse como imágenes de referencia a corto plazo e imágenes de referencia a largo plazo. Las imágenes de referencia a corto plazo son imágenes que típicamente están temporalmente cerca de la imagen actual en orden de salida.

**[0021]** De acuerdo con el primer aspecto de las técnicas descritas en esta divulgación, en lugar de codificar los LSB de POC delta que identifican una imagen de referencia a largo plazo que está presente en la memoria intermedia de imágenes descodificadas pero no se especifica en el SPS asociado usando codificación exponencial de Golomb, un el codificador de vídeo puede especificar (señalar) directamente los LSB de POC de tal imagen de referencia a largo plazo utilizando códigos unarios fijos variables. Es decir, el codificador de vídeo puede especificar el número de bits utilizados para señalar estos LSB de POC, que se señalan en una cabecera de fragmento asociado con la imagen actual. El número especificado de bits señala al descodificador de vídeo el número de bits que se van a analizar desde el flujo de bits y que representan los LSB de POC que identifican la imagen de referencia a largo plazo que está presente en la memoria intermedia de imágenes descodificadas pero no se especifica en el SPS asociado. Dada la distribución más amplia de valores en el contexto de dichos valores de LSB de POC, las técnicas pueden promover el ahorro de bits codificando estos LSB de POC como números binarios sin signo de longitud uniforme (y variable, en el sentido de que la longitud uniforme puede señalarse u obtenerse para cada fragmento), denotados como "u(v)" en varias normas de codificación de vídeo.

**[0022]** De esta manera, un descodificador de vídeo puede determinar una cantidad de bits utilizados para representar uno o más bits menos significativos del valor de recuento del orden de imágenes que identifica una imagen de referencia a largo plazo que se utilizará al descodificar la imagen actual. A menudo, el descodificador de vídeo puede determinar el número de bits de un elemento sintáctico especificado en uno o más de una cabecera de fragmento, conjunto de parámetros de imagen o conjunto de parámetros de secuencia. De forma alternativa, el descodificador de vídeo puede obtener el número de bits basado en otros elementos sintácticos sin recibir explícitamente un elemento sintáctico que especifique el número de bits. El descodificador de vídeo puede a continuación analizar el número determinado de bits del flujo de bits, donde este número de bits analizados representa los bits significativos del valor POC que identifica la imagen de referencia a largo plazo. El descodificador de vídeo luego puede recuperar la imagen de referencia a largo plazo de una memoria intermedia de imágenes descodificadas basándose en los bits menos significativos del valor POC que identifica la imagen de referencia a largo plazo y descodificar al menos la parte de la imagen actual usando la imagen de referencia a largo plazo recuperada.

**[0023]** De acuerdo con un segundo aspecto de las técnicas descritas en esta divulgación, el codificador de vídeo puede configurarse de forma adicional o alternativa para analizar cada imagen presente en toda la memoria intermedia de imágenes descodificadas cuando se determina si es necesario señalar información adicional, por ejemplo, parte o todos los bits más significativos (MSB), para desambiguar los LSB de POC. En otras palabras, el codificador de vídeo calcula un conjunto de uno o más LSB de POC para cada valor POC que identifica una imagen en la DPB. Al realizar el análisis, el codificador de vídeo determina si el conjunto de LSB de POC calculados para cada una de las imágenes de referencia a largo plazo señaladas en la cabecera de fragmento es único con respecto a todos los otros conjuntos de LSB de POC calculados para los valores POC que identifican a todas las otras imágenes en la DPB. Si uno de estos otros conjuntos de LSB de POC es igual al conjunto de LSB de POC de las imágenes de referencia a largo plazo actuales bajo el análisis, se determina que el conjunto de LSB de POC no es único, y el codificador de vídeo puede señalar los bits más significativos (MSB) suficientes para permitir que el descodificador de vídeo obtenga correctamente el valor POC de la imagen de referencia a largo plazo.

**[0024]** De esta manera, el descodificador de vídeo puede recibir un flujo de bits representativo de los datos de vídeo codificados del codificador de vídeo, que se ha codificado de acuerdo con varios aspectos de las técnicas descritas en esta divulgación. El descodificador de vídeo puede entonces determinar los bits menos significativos del valor POC que identifica la imagen de referencia a largo plazo que se utilizará cuando se descodifique la imagen actual. Como se señaló anteriormente, estos bits menos significativos pueden no identificar de manera única el valor de recuento del orden de imágenes de la imagen de referencia a largo plazo con respecto a los bits menos significativos de cualquier otro valor POC que identifique cualquier otra imagen presente en una memoria intermedia de imágenes descodificadas en la que la imagen de referencia a largo plazo también se almacena.

Como resultado, el descodificador de vídeo puede a continuación determinar uno o más de los bits más significativos del valor POC que identifica la imagen de referencia a largo plazo. El codificador de vídeo selecciona los bits más significativos de modo que los bits más significativos en combinación con los bits menos significativos sean suficientes para distinguir el valor POC que identifica la imagen de referencia a largo plazo de cualquier otro valor de recuento del orden de imágenes que identifique cualquier otra imagen en la memoria intermedia de imágenes descodificadas. A continuación, el descodificador de vídeo puede recuperar la imagen de referencia a largo plazo de la memoria intermedia de imágenes descodificadas basándose en la combinación inequívoca de los bits menos significativos y los bits más significativos y descodificar la imagen actual utilizando la imagen de referencia a largo plazo recuperada.

**[0025]** La FIG. 1 es un diagrama de bloques que ilustra un sistema de codificación y descodificación de vídeo de ejemplo que puede utilizar las técnicas descritas en esta divulgación para obtener RPS y LTRP de señalización. Como se muestra en la FIG. 1, el sistema 10 incluye un dispositivo de origen 12 que transmite el vídeo codificado a un dispositivo de destino 14 a través de un canal de comunicación 16. El dispositivo de origen 12 y el dispositivo de destino 14 pueden comprender uno cualquiera de una amplia gama de dispositivos. En algunos casos, el dispositivo 12 de origen y el dispositivo 14 de destino pueden comprender dispositivos de comunicación inalámbrica, tales como teléfonos inalámbricos, los denominados radioteléfonos celulares o vía satélite, o cualquier dispositivo inalámbrico que pueda transmitir información de vídeo a través de un canal 16 de comunicación, en cuyo caso el canal 16 de comunicación es inalámbrico. Las técnicas de esta divulgación no están limitadas necesariamente a aplicaciones o configuraciones inalámbricas. Por ejemplo, estas técnicas pueden aplicarse a radiodifusiones televisivas inalámbricas, a transmisiones televisivas por cable, a transmisiones televisivas vía satélite, a transmisiones de vídeo por Internet, a vídeo digital codificado que está codificado en un medio de almacenamiento y a otros contextos. Por consiguiente, el canal de comunicación 16 puede comprender cualquier combinación de medios inalámbricos, cableados o de almacenamiento adecuados para la transmisión o el almacenamiento de datos de vídeo codificados.

**[0026]** En el ejemplo de la FIG 1, el dispositivo 12 de origen incluye una fuente 18 de vídeo, un codificador 20 de vídeo, un modulador/desmodulador (módem) 22 y un transmisor 24. El dispositivo 14 de destino incluye un receptor 26, un módem 28, un descodificador 30 de vídeo y un dispositivo 32 de visualización. El codificador de vídeo 20 puede configurarse para aplicar las técnicas de esta divulgación. En otros ejemplos, un dispositivo de origen y un dispositivo de destino pueden incluir otros componentes o disposiciones. Por ejemplo, el dispositivo de origen 12 puede recibir datos de vídeo desde una fuente de vídeo externa, tal como una cámara externa. Del mismo modo, el dispositivo de destino 14 puede interactuar con un dispositivo de visualización externo, en lugar de incluir un dispositivo de visualización integrado.

**[0027]** El sistema 10 de la FIG 1 es simplemente un ejemplo. Las técnicas para la codificación de datos de vídeo pueden realizarse mediante varios dispositivos de codificación y/o descodificación de vídeo digital. Las técnicas de esta divulgación pueden realizarse mediante un codificador/descodificador de vídeo, típicamente denominado "CODEC". Además, las técnicas de esta divulgación también pueden realizarse mediante un preprocesador de vídeo. El dispositivo de origen 12 y el dispositivo de destino 14 son simplemente ejemplos de tales dispositivos de codificación. En algunos ejemplos, los dispositivos 12, 14 pueden funcionar de manera esencialmente simétrica, de modo que cada uno de los dispositivos 12, 14 incluya componentes de codificación y de descodificación de vídeo. Por lo tanto, el sistema 10 puede soportar una transmisión de vídeo unidireccional o bidireccional entre el dispositivo de origen 12 y el dispositivo de destino 14, por ejemplo, para la transmisión continua de vídeo, reproducción de vídeo, radiodifusión de vídeo o videotelefonía.

**[0028]** La fuente de vídeo 18 puede incluir un dispositivo de captura de vídeo, tal como una cámara de vídeo, un archivo de vídeo que contiene vídeo capturado previamente y/o una transmisión de vídeo de un proveedor de contenido de vídeo. Como otra alternativa, la fuente de vídeo 18 puede generar datos basados en gráficos por ordenador como el vídeo de origen, o una combinación de vídeo en directo, vídeo archivado y vídeo generado por ordenador. En algunos casos, si la fuente de vídeo 18 es una videocámara, el dispositivo de origen 12 y el dispositivo de destino 14 pueden formar los llamados teléfonos con cámara o videoteléfonos. En cada caso, el codificador de vídeo 20 puede codificar el vídeo captado, precaptado o generado por ordenador. La información de vídeo codificada puede modularse después mediante el módem 22, de acuerdo con una norma de comunicación, y transmitirse al dispositivo 14 de destino a través del transmisor 24. El módem 22 puede incluir diversos mezcladores, filtros, amplificadores u otros componentes diseñados para la modulación de señales. El transmisor 24 puede incluir circuitos diseñados para transmitir datos, incluidos amplificadores, filtros y una o más antenas.

**[0029]** El receptor 26 del dispositivo de destino 14 recibe información a través del canal de comunicación 16, y el módem 28 desmodula la información. La información comunicada por el canal de comunicación 16 puede incluir información sintáctica definida por el codificador de vídeo 20, que incluye elementos sintácticos que describen características y/o procesamiento de bloques y otras unidades codificadas, por ejemplo, grupos de imágenes (GOP). 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 un

tubo de rayos catódicos (CRT), una pantalla de cristal líquido (LCD), una pantalla de plasma, una pantalla de diodos orgánicos emisores de luz (OLED) u otro tipo de dispositivo de visualización.

**[0030]** En el ejemplo de la FIG 1, el canal 16 de comunicación puede comprender cualquier medio de comunicación inalámbrica o alámbrica, tal como un espectro de radiofrecuencia (RF) o una o más líneas de transmisión físicas, o cualquier combinación de medios inalámbricos y alámbricos. El canal de comunicación 16 puede formar parte de una red basada en paquetes, tal como una red de área local, una red de área extensa o una red global tal como Internet. El canal de comunicación 16 puede representar cualquier medio de comunicación adecuado, o un conjunto de medios de comunicación diferentes, para transmitir datos de vídeo desde el dispositivo de origen 12 hasta el dispositivo de destino 14, incluida cualquier combinación adecuada de medios alámbricos o inalámbricos. El canal 16 de comunicación puede incluir routers, conmutadores, estaciones base o cualquier otro equipo que pueda ser útil para facilitar la comunicación desde el dispositivo 12 de origen hasta el dispositivo 14 de destino. En otros ejemplos, el dispositivo origen 12 puede almacenar datos codificados en un medio de almacenamiento, en lugar de transmitir los datos. Asimismo, el dispositivo destino 14 puede estar configurado para recuperar datos codificados de un medio de almacenamiento.

**[0031]** 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 ITU-T H.264, de forma alternativa denominada MPEG-4, Parte 10, Codificación de vídeo avanzada (AVC), o la próxima norma de Codificación de vídeo de alta eficiencia, que también puede denominarse comúnmente H.265. Si bien las técnicas se describen en general en esta divulgación con respecto a HEVC, las técnicas también pueden implementarse o usarse de otra manera para normas de codificación de vídeo alternativas, como la norma H.264/Codificación de vídeo avanzada (AVC). La HEVC está siendo desarrollada por el Equipo de Colaboración Conjunta sobre 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 la ISO/IEC. Un reciente borrador de trabajo (WD) de HEVC, denominado y referido como HEVC WD6 en adelante, Bross et al, titulado "High efficiency video coding (HEVC) text specification draft 6" ["Borrador de especificación de texto de alta eficiencia de codificación de vídeo (HEVC) 6"], Equipo de colaboración conjunta sobre codificación de vídeo (JCT-VC) de ITU-T SG16 WP3 e ISO/CEI JTC1/SC29/WG11, 8.ª reunión, San José, CA, EE. UU., del 1 al 10 de febrero de 2012, disponible en [http://phenix.int-evry.fr/jct/doc\\_end\\_user/documents/8\\_SanJose/wg11/JCTVC-H1003-v22.zip](http://phenix.int-evry.fr/jct/doc_end_user/documents/8_SanJose/wg11/JCTVC-H1003-v22.zip). Adicionalmente, también se han aprobado otros borradores de trabajo de HEVC, con el último denominado HEVC WD9, Bross et al., titulado "High efficiency video coding (HEVC) text specification draft 9" ["Borrador 9 de memoria descriptiva textual de Codificación de Vídeo de Alta Eficiencia (HEVC)"], Equipo de colaboración conjunta sobre codificación de vídeo (JCT-VC) de ITU-T SG16 WP3 e ISO/IEC JTC1/SC29/WG11, 11.ª reunión: Shanghái, China, del 10 al 19 de octubre de 2012, JCTVC-K1003\_v9, disponible en [http://phenix.int-evry.fr/jct/doc\\_end\\_user/documents/11\\_Shanghai/wg11/JCTVC-K1003-v9.zip](http://phenix.int-evry.fr/jct/doc_end_user/documents/11_Shanghai/wg11/JCTVC-K1003-v9.zip)

**[0032]** La norma ITU-T H.264/MPEG-4 (AVC) fue formulada por el Grupo de Expertos en Codificación de Vídeo (VCEG) de la ITU-T junto con el Grupo de Expertos en Imágenes en Movimiento (MPEG) de la ISO/IEC, como el producto de una asociación colectiva conocida como Equipo Mixto de Vídeo (JVT, por sus siglas en inglés). En algunos aspectos, las técnicas descritas en esta divulgación pueden aplicarse a dispositivos que se ajustan en general a la norma H.264. La norma H.264 se describe en la recomendación H.264 de ITU-T, Codificación de vídeo avanzada para servicios audiovisuales genéricos, por el Grupo de estudio de la ITU-T, con fecha de marzo de 2005, que puede mencionarse en el presente documento como la norma H.264 o la especificación H.264, o la norma o especificación H.264/AVC. El Equipo mixto de vídeo (JVT) continúa trabajando en extensiones de la norma H.264/MPEG-4 AVC.

**[0033]** 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). El más reciente borrador conjunto de la MVC se describe en "Advanced video coding for generic audiovisual services" ["Codificación de vídeo avanzada para servicios audiovisuales genéricos"], Recomendación H.264 de la ITU-T, marzo de 2010.

**[0034]** Sin embargo, las técnicas de esta divulgación no están limitadas a ninguna norma de codificación particular. Otros ejemplos incluyen MPEG-2 e ITU-T H.263. Aunque no se muestra en la FIG 1, en algunos aspectos, tanto el codificador 20 de vídeo como el descodificador 30 de vídeo 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, las unidades MUX-DEMUX pueden ajustarse al protocolo de multiplexador H.223 de la ITU o a otros protocolos tales como el protocolo de datagramas de usuario (UDP).

**[0035]** Tanto el codificador 20 de vídeo como el descodificador 30 de vídeo pueden implementarse como cualquiera de entre una variedad de circuitos codificadores o descodificadores adecuados, según corresponda, tales como uno o más de entre microprocesadores, procesadores de señales digitales (DSP), circuitos integrados específicos de la aplicación (ASIC), matrices de puertas programables *in situ* (FPGA), circuitos de lógica discreta,

software, hardware, firmware o cualquier combinación de los mismos. El término "procesador" en general se puede usar para referirse a cualquier combinación de lo anterior y puede incluir medios legibles por ordenador no transitorios que tengan almacenados en los mismos instrucciones que, cuando se ejecutan, hacen que este procesador realice las diversas técnicas descritas en esta divulgación. Tanto el codificador de vídeo 20 como el descodificador de vídeo 30 pueden estar incluidos en uno o más codificadores o descodificadores, cada uno de los cuales puede estar integrado como parte de un codificador/descodificador (códec) de vídeo combinado. Un aparato 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, como un teléfono celular, una cámara, un ordenador, un dispositivo móvil, un dispositivo de suscripción, un dispositivo de radiodifusión, descodificador, dispositivo de juego, servidor, o similares.

**[0036]** Una secuencia de vídeo incluye típicamente una serie de tramas de vídeo, denominadas de forma alternativa imágenes. Un grupo de imágenes (GOP) en general comprende una serie de una o más tramas de vídeo (que es otra forma comúnmente conocida de referirse a una imagen). Un GOP puede incluir datos sintácticos en una cabecera del GOP, una cabecera de una o más tramas del GOP, o en otras ubicaciones, que indican el número de tramas incluidas en el GOP. Cada trama puede incluir datos sintácticos de trama que describen un modo de codificación para la respectiva trama. Un codificador 20 de vídeo actúa típicamente sobre bloques de vídeo de tramas de vídeo individuales con el fin de codificar los datos de vídeo. Un bloque de vídeo puede corresponder a un macrobloque o una partición de un macrobloque en H.264, o una CU (que puede incluir una o más unidades de predicción (PU) y/o unidades de transformación (TU)) de HEVC. Los bloques de vídeo pueden tener tamaños fijos o variables y pueden diferir en tamaño de acuerdo con una norma de codificación especificada. Cada trama de vídeo puede incluir una pluralidad de fragmentos. Cada fragmento puede incluir una pluralidad de macrobloques (o LCU), que pueden dividirse en sub-bloques (o CU).

**[0037]** En un ejemplo, la norma ITU-T H.264 soporta la predicción intra en diversos tamaños de bloque, tales como 16 por 16, 8 por 8 o 4 por 4 para componentes de luma, y 8x8 para componentes de croma, así como la predicción inter en diversos tamaños de bloque, tales como 16x16, 16x8, 8x16, 8x8, 8x4, 4x8 y 4x4 para componentes de luma y tamaños escalados correspondientes para componentes de croma. En esta divulgación, "NxN" y "N por N" pueden usarse de manera intercambiable para referirse a las dimensiones de píxel del bloque en lo que respecta a las dimensiones vertical y horizontal, por ejemplo 16x16 píxeles o 16 por 16 píxeles. En general, un bloque de 16x16 tendrá 16 píxeles en una dirección vertical ( $y = 16$ ) y 16 píxeles en una dirección horizontal ( $x = 16$ ). Asimismo, un bloque de NxN tiene, en general, N píxeles en una dirección vertical y N píxeles en una dirección horizontal, donde N representa un valor entero no negativo. Los píxeles de un bloque se pueden disponer en filas y columnas. Además, no es necesario que los bloques tengan necesariamente el mismo número de píxeles en la dirección horizontal y en la dirección vertical. Por ejemplo, los bloques pueden comprender NxM píxeles, donde M no es necesariamente igual a N.

**[0038]** Los tamaños de bloques que son menores que 16x16 se pueden denominar divisiones de un bloque de 16x16. Los bloques de vídeo pueden comprender bloques de datos de píxel en el dominio de píxeles, o bloques de coeficientes de transformada en el dominio de transformada. En algunos casos, un bloque de vídeo puede comprender bloques de coeficientes de transformada cuantificados en el dominio de la transformada.

**[0039]** Bloques de vídeo más pequeños pueden proporcionar una mejor resolución y pueden usarse en ubicaciones de una trama de vídeo que incluyen altos niveles de detalle. En general, los bloques y las diversas particiones, denominadas en ocasiones sub-bloques, pueden considerarse bloques de vídeo. Además, un fragmento puede considerarse una pluralidad de bloques de vídeo, tales como bloques y/o sub-bloques. Cada fragmento puede ser una unidad independientemente descodificable de una trama de vídeo. De forma alternativa, las propias tramas pueden ser unidades descodificables, o pueden definirse otras partes de una trama como unidades descodificables. El término "unidad codificada" o "unidad de codificación" puede referirse a cualquier unidad independientemente descodificable de una trama de vídeo, tal como una trama completa, un fragmento de una trama, un GOP, denominado también secuencia, u otra unidad independientemente descodificable definida de acuerdo con técnicas de codificación aplicables. Un GOP también se puede denominar secuencia de vídeo codificada, por ejemplo, en H.264/AVC y HEVC.

**[0040]** El codificador de vídeo 20 puede generar unidades de capa de abstracción de red (NAL). Una unidad 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 (SPS), un conjunto de parámetros de imagen (PPS), un fragmento codificado, información de mejora suplementaria (SEI), un delimitador de unidad de acceso, datos de relleno u otro tipo de datos. Los datos de una unidad NAL pueden estar en forma de una carga útil de secuencia de bytes sin procesar (RBSP) entremezclada con bits de prevención de emulación. Una RBSP puede ser una estructura sintáctica que contenga un número entero de octetos que se encapsule dentro de una unidad de NAL.

**[0041]** Las unidades de NAL pueden clasificarse en unidades de NAL de la capa de codificación de vídeo (VCL) y unidades de NAL no de la VCL. Las unidades VCL pueden incluir datos a nivel de bloque, macrobloque y/o fragmento. Las unidades de NAL que no son de VCL pueden incluir unidades de NAL del conjunto de parámetros

y unidades de NAL de SEI, entre otras. Los conjuntos de parámetros pueden contener información de cabecera a nivel de secuencia (por ejemplo, en un SPS) y la información de cabecera a nivel de imagen, que cambia con poca frecuencia (por ejemplo, en un PPS). Con los conjuntos de parámetros (por ejemplo, PPS y SPS), la información que cambia con poca frecuencia no necesita ser repetida para cada secuencia o imagen, por lo que la eficiencia de la codificación puede mejorarse. Además, el uso de conjuntos de parámetros puede permitir la transmisión fuera de banda de la información de cabecera importante, evitando la necesidad de transmisiones redundantes, para la capacidad de recuperación de errores. En los ejemplos de transmisión fuera de banda, las unidades de NAL del conjunto de parámetros pueden transmitirse en un canal diferente al de otras unidades de NAL, tales como las unidades de NAL de SEI.

**[0042]** Los mensajes de SEI pueden contener información que no es necesaria para descodificar las muestras de imágenes codificadas a partir de las unidades de NAL de VCL, pero puede ayudar en los procesos relacionados con la descodificación, visualización, resistencia a errores y otros fines. Los mensajes de SEI pueden estar contenidos en las unidades de NAL no de VCL. Los mensajes de SEI son la parte normativa de algunas especificaciones estándar y, por lo tanto, no siempre son obligatorios para la implementación de descodificadores compatibles con las normas. Los mensajes de SEI pueden ser mensajes de SEI a nivel de secuencia o mensajes de SEI a nivel de imagen. Parte de la información a nivel de secuencia puede estar contenida en los mensajes de SEI, tales como los mensajes de SEI de información de ajustabilidad a escala en el ejemplo de la SVC y los mensajes de SEI de información de ajustabilidad a escala de visualizaciones en la MVC. Estos mensajes de SEI a modo de ejemplo pueden transmitir información, por ejemplo, sobre extracción de puntos de operación y características de los puntos de operación.

**[0043]** Como se señaló anteriormente, un codificador de vídeo (que es un término que puede referirse a uno o ambos del codificador de vídeo 20 y el descodificador de vídeo 30) incluye una memoria intermedia de imágenes descodificadas (DPB), que no se muestra en el ejemplo de la FIG. 1 para facilitar la ilustración. La DPB almacena imágenes de referencia, que son imágenes que se pueden usar para realizar predicción inter de una imagen. En otras palabras, el codificador de vídeo puede predecir una imagen basándose en una o más imágenes de referencia almacenadas en la DPB.

**[0044]** El descodificador de vídeo también se puede configurar para construir listas de imágenes de referencia que indiquen qué imágenes de referencia se utilizan con fines de predicción inter. Dos de estas listas de imágenes de referencia se denominan Lista 0 y Lista 1, respectivamente. El descodificador de vídeo 30 puede emplear primero técnicas de construcción predeterminadas para construir la Lista 0 y la Lista 1 (por ejemplo, esquemas de construcción preconfigurados para la construcción de la Lista 0 y la Lista 1, tales como los establecidos en HEVC WD 6 como un ejemplo). Opcionalmente, el codificador de vídeo 20 puede configurarse para señalar elementos sintácticos que especifican qué imágenes almacenadas en la DPB deben usarse al construir la Lista 0 y la Lista 1. Por lo tanto, después de construir la Lista 0 y la Lista 1 iniciales, el descodificador de vídeo 30 puede descodificar los elementos sintácticos, cuando están presentes, que ordenan al descodificador de vídeo 30 modificar la Lista 0 y la Lista 1 iniciales.

**[0045]** El codificador de vídeo 30 puede señalar elementos sintácticos que son indicativos del (de los) identificador(es) de imágenes de referencia en la DPB. El codificador de vídeo 30 también puede indicar elementos sintácticos que incluyen índices, dentro de la Lista 0, la Lista 1 o la Lista 0 y la Lista 1, que indican qué imagen o imágenes de referencia almacenadas en la DPB deben usarse para descodificar un bloque codificado de una imagen actual. El descodificador de vídeo 30, a su vez, utiliza el identificador recibido para identificar el valor o los valores de índice para una imagen de referencia o imágenes de referencia enumeradas en la Lista 0, la Lista 1 o la Lista 0 y la Lista 1. Desde el (los) valor(es) de índice, así como el (los) identificador(es) de la imagen de referencia o imágenes de referencia, el descodificador de vídeo 30 recupera la imagen de referencia o las imágenes de referencia de la DPB, forma la Lista 0 y la Lista 1 y descodifica el bloque codificado de la imagen actual.

**[0046]** En un conjunto de parámetros de imagen (PPS) o en un conjunto de parámetros de secuencia (SPS) asociado con una cabecera de fragmento de una imagen actual, el codificador de vídeo 20 puede señalar un conjunto de imágenes de referencia (que de nuevo puede denominarse "RPS"). El RPS de una imagen actual incluye información de identificación para imágenes de referencia que se puede usar para predecir la imagen actual y las imágenes que se pueden usar para predecir imágenes después de la imagen actual en orden de descodificación. Solo las imágenes de referencia en el RPS pueden incluirse en la Lista 0 o la Lista 1.

**[0047]** La información de identificación para las imágenes de referencia puede incluir uno o más valores de recuento del orden de imágenes (POC). Los valores de POC indican el orden en que se general o muestran las imágenes (es decir, el orden de visualización de las imágenes). Por ejemplo, una imagen con un valor POC más bajo se muestra antes que una imagen con un valor POC más alto en la misma secuencia de vídeo codificada.

**[0048]** Los borradores de trabajo de HEVC WD 6 y posteriores, por ejemplo, HEVC WD 9, en general proporcionan dos tipos diferentes de imágenes de referencia denominadas imágenes de referencia a corto plazo (STRP) e imágenes de referencia a largo plazo (LTRP). Las imágenes de referencia a corto plazo son imágenes que típicamente están temporalmente cerca de la imagen actual en orden de salida.



**[0049]** Las imágenes de referencia a largo plazo son imágenes que típicamente están temporalmente más alejadas de la imagen actual, pero aún pueden ser útiles para fines de predicción. Por ejemplo, en la videoconferencia, unas pocas imágenes capturadas inicialmente pueden incluir contenido de imagen de fondo que puede ser útil para predecir imágenes que se capturarán más adelante en la videoconferencia. Como otro ejemplo, una imagen de referencia a largo plazo puede tratarse de manera diferente a una imagen de referencia a corto plazo en la escala del vector de movimiento, o la escala del valor de la muestra en la predicción ponderada.

**[0050]** Una imagen de referencia a corto plazo también puede estar más alejada de una imagen actual en orden de salida que una imagen de referencia a largo plazo. En este sentido, las imágenes de referencia a largo plazo pueden referirse a imágenes que pueden usarse varias veces como imágenes de referencia, mientras que las imágenes de referencia a corto plazo pueden usarse menos veces como imágenes de referencia en relación con las imágenes de referencia a largo plazo. Confiar solo en imágenes de referencia a corto plazo para la predicción puede hacer que el descodificador de vídeo 30 no sea capaz de usar imágenes de referencia a largo plazo que puedan ser útiles o deseables para la eficiencia de codificación o la capacidad de recuperación de errores.

**[0051]** Para especificar estas imágenes de referencia a largo plazo, el codificador de vídeo 20 puede proporcionar una lista de imágenes de referencia a largo plazo que pueden asociarse y usarse para descodificar una o más imágenes en el SPS. En las cabeceras de fragmentos (que también pueden denominarse "cabeceras de fragmento") de la imagen o las imágenes asociadas, el codificador de vídeo 20 puede, en algunos casos, especificar cuál de estas imágenes de referencia a largo plazo especificadas en el SPS deben utilizarse cuando se descodifica una imagen en particular. A menudo, el codificador de vídeo 20 especifica un índice que identifica cada una de las imágenes de referencia a largo plazo especificadas en el SPS que se utilizarán al descodificar una imagen en particular. Típicamente, se requiere que todas las cabeceras de fragmento de una imagen se refieran a la misma señalización de conjunto de imágenes de referencia.

**[0052]** En algunos casos, el codificador de vídeo 20 puede determinar que una o más imágenes de referencia a largo plazo que están presentes en la memoria intermedia de imágenes descodificadas, pero que no están especificadas en la lista de imágenes de referencia a largo plazo especificadas en el SPS, deben usarse al descodificar una imagen en particular. En este caso, el codificador de vídeo 20 puede especificar (o, en el contexto de la codificación de vídeo, "señalar") los valores POC que se asignan a una o más imágenes de referencia a largo plazo presentes en la DPB pero que no están especificadas en la lista de imágenes de referencia a largo plazo especificadas en el SPS.

**[0053]** Para reducir el número de bits necesarios para señalar estos valores de POC, el codificador de vídeo 20 puede determinar un valor POC en función del valor POC asignado a una de las imágenes (que puede denominarse "imagen actual" en el sentido que esta imagen es la imagen que se está descodificando actualmente) a la que se corresponde el fragmento (que puede denominarse "fragmento actual" por razones similares) a descodificar. Para ilustrar, el codificador de vídeo 20 puede restar el valor POC de la imagen de referencia a largo plazo del valor POC de la imagen actual a la que corresponde el fragmento actual para obtener un valor POC delta. En particular, el valor POC delta puede representarse por un número seleccionado de bits menos significativos (LSB) del valor POC delta. Al indicar únicamente los LSB de POC delta en una cabecera de fragmento, el codificador de vídeo 20 puede reducir el número de bits necesarios para identificar las imágenes de referencia a largo plazo que están presentes en la memoria intermedia de imágenes descodificadas pero no se especifican en el SPS asociado con el fragmento actual; estas imágenes de referencia a largo plazo también se conocen como señaladas directamente en cabeceras de fragmento.

**[0054]** Anteriormente, los LSB de POC delta para especificar un valor POC de una imagen de referencia a largo plazo se codificaban por entropía para mejorar la resistencia al error y reducir el número de bits necesarios para especificar el valor POC en el flujo de bits. En particular, estos codificadores de vídeo anteriores codificaban LSB de POC delta de imágenes de referencia a largo plazo directamente señaladas en cabeceras de fragmento utilizando una forma de codificación indicada como "ue(v)" en varias normas de codificación de vídeo, incluyendo HEVC. El término "ue(v)" se refiere a una forma de codificación denominada codificación de Golomb exponencial.

**[0055]** Cuando el valor a especificar no tiene signo (que forma la "u" de "ue(v)", mientras que la "e" se refiere a exponencial en los códigos de Golomb exponencial), la codificación de Golomb exponencial implica señalar el número de bits utilizados para codificar un valor dado especificando un número de ceros que sea igual a uno menos que el número de bits para codificar un valor dado y luego especificar el valor dado más uno. Si el valor LSB de POC delta es 0001, por ejemplo, el codificador de vídeo codificó este valor como 0 (para indicar que el valor puede señalarse utilizando un bit, es decir, 1 en este ejemplo) seguido de 10 (dado que el valor de uno más uno es dos, que en binario se especifica como 10), generando un código Golomb exponencial de 010.

**[0056]** Debido a que los códigos de Golomb exponencial son una forma de codificación de longitud variable (que se refiere al hecho de que estos códigos no son de longitud fija o uniforme y en cambio varían basándose en el valor que se codifique), la codificación de Golomb exponencial en general solo es adecuada para codificar información que tiene ciertas distribuciones estadísticas de valores. Más específicamente, la codificación de

Golomb exponencial es adecuada para la codificación de información estadísticamente probable que se especifique como un número pequeño. El cálculo de los LSB de POC delta para especificar un valor POC de imagen actual en general implicaba calcular un valor de uno (dado que el valor POC actual se restó del valor POC de imagen anterior para obtener los LSB de POC delta en el contexto de señalización del valor POC de la imagen actual). En consecuencia, se eligió la codificación de Golomb exponencial para codificar estos valores de LSB de POC delta para especificar el valor POC de la imagen actual. Sin embargo, los valores LSB de POC delta para identificar imágenes de referencia a largo plazo representan información que tiene una distribución muy diferente que no es adecuada para la codificación de Golomb exponencial.

**[0057]** De acuerdo con el primer aspecto de las técnicas descritas en esta divulgación, en lugar de codificar los LSB de POC delta que identifican una imagen de referencia a largo plazo que está presente en la memoria intermedia de imágenes descodificadas pero no se especifica en el SPS asociado usando codificación Golomb exponencial, el codificador de vídeo 20 puede especificar directamente (o, en otras palabras, señalar) los LSB de POC de dicha imagen de referencia a largo plazo utilizando códigos unarios fijos variables. Es decir, el codificador de vídeo 20 puede especificar el número de bits utilizados para señalar estos LSB de POC en una cabecera de fragmento asociada con la imagen actual. El número especificado de bits señala al descodificador de vídeo el número de bits que se van a analizar desde el flujo de bits y que representan los LSB de POC que identifican la imagen de referencia a largo plazo que está presente en la memoria intermedia de imágenes descodificadas pero no se especifica en el SPS asociado. Dada la distribución más amplia de valores en el contexto de dichos valores de LSB de POC, las técnicas pueden promover el ahorro de bits codificando estos LSB de POC como números binarios sin signo de longitud uniforme (y variable, en el sentido de que la longitud uniforme puede señalarse u obtenerse para cada fragmento), denotados como "u(v)" en varias normas de codificación de vídeo.

**[0058]** En funcionamiento, el codificador de vídeo 20 puede determinar, para una imagen actual de los datos de vídeo, una imagen de referencia a largo plazo que se utilizará cuando se codifique al menos una parte de una imagen actual de los datos de vídeo de la manera descrita anteriormente. El codificador de vídeo 20 puede luego determinar una cantidad de bits que se usarán para representar un conjunto de uno o más LSB de un valor POC que identifica la imagen de referencia a largo plazo en un flujo de bits representativo de una versión codificada de los datos de vídeo. En otras palabras, el codificador de vídeo 20 puede determinar la "v" en la codificación binaria sin signo denotada como "u(v)" con la "u", denotando nuevamente la codificación binaria sin signo.

**[0059]** A menudo, este número de LSB se define basándose en el número de LSB necesarios para desambiguar la LTRP de otras LTRP o cualquier otra imagen de referencia almacenada en la DPB. En otras palabras, este número puede variar en un fragmento, una imagen o un GOP. En algunos casos, el número de LSB puede definirse para un flujo de bits completo, o partes del mismo. En algunos casos, el número de LSB utilizados para identificar un valor POC está predefinido, de modo que no es necesario señalar el número de LSB. En algunos casos, el número de LSB puede señalarse para cada imagen, por ejemplo, en la cabecera de fragmento. En algunos casos, el número de LSB puede señalarse para varias imágenes, por ejemplo, en el SPS o PPS. En algunos casos, el número de LSB utilizados para identificar el valor POC se puede obtener basándose en los elementos sintácticos determinados para la imagen actual y/o las imágenes codificadas previamente de los datos de vídeo, donde el codificador de vídeo 20 tal vez no necesite señalar el número de bits utilizados para representar este LSB de POC dado que el descodificador de vídeo 30 puede realizar un proceso de obtención similar para obtener el número "v" de bits utilizados para representar el LSB de POC.

**[0060]** En cualquier caso, el codificador de vídeo 20 puede especificar un conjunto de uno o más LSB del valor POC que identifica la LTRP en el flujo de bits utilizando el número determinado de bits utilizados para representar el número de un conjunto de uno o más valores LSB del POC que identifican la imagen de referencia a largo plazo. En otras palabras, el codificador de vídeo 20 puede especificar los LSB de POC anotados anteriormente en el flujo de bits utilizando el número de bits "v" como un número binario sin signo. El codificador de vídeo 20 también puede codificar al menos la parte de la imagen actual usando la LTRP. El codificador de vídeo 20 puede codificar al menos la parte de la imagen actual antes, después o simultáneamente a la especificación en el flujo de bits de los LSB de POC como un valor binario sin signo de "v" bits de longitud. El codificador de vídeo 20 puede luego transmitir este flujo de bits al descodificador de vídeo 30 o almacenar este flujo de bits para su posterior recuperación por el descodificador de vídeo 30.

**[0061]** El descodificador de vídeo 30 puede recibir el flujo de bits representativo de los datos de vídeo codificados que se han codificado de acuerdo con las técnicas descritas en esta divulgación. El módulo de descodificación por entropía 150 del descodificador de vídeo 30 puede determinar primero, para una parte de una imagen actual de los datos de vídeo codificados, el número de bits utilizados para especificar un conjunto de uno o más LSB del valor POC que identifica una LTRP para su uso en la descodificación de la imagen actual. En algunos ejemplos, el número de bits puede señalarse explícitamente en el flujo de bits como un elemento sintáctico, por ejemplo, en la cabecera de fragmento, SPS o PPS, como se describió anteriormente. En otros ejemplos, el número de bits se puede obtener en función de otros elementos sintácticos especificados en uno o más del SPS, un PPS y la cabecera de fragmento.

**[0062]** El descodificador de vídeo 30 puede luego analizar el número determinado de bits del flujo de bits, donde el número determinado de bits analizados representa el conjunto de uno o más LSB del valor POC que identifica la imagen de referencia a largo plazo que se utilizará al descodificar al menos una parte de la imagen actual. El número determinado de bits puede especificar lo que se ha denominado anteriormente como los LSB de POC. El descodificador de vídeo 30 luego puede recuperar, de la memoria intermedia de imágenes descodificadas, la LTRP para ser usada cuando se descodifique al menos la parte de la imagen actual basándose en los LSB de POC.

**[0063]** Para recuperar esta LTRP de la memoria intermedia de imágenes descodificadas, el descodificador de vídeo 30 puede comparar los bits menos significativos del valor POC que identifica la LTRP con un mismo número de bits menos significativos de un valor POC asociado con la al menos una imagen almacenada en la memoria intermedia de imágenes descodificadas. El descodificador de vídeo 30 recupera una de las al menos una imagen almacenada en la memoria intermedia de imágenes descodificadas para usar como imagen de referencia a largo plazo en la descodificación de la imagen actual que tiene los mismos LSB para el valor POC asociado con la una de al menos una imágenes almacenadas en la memoria intermedia de imágenes descodificadas como los bits menos significativos determinados de la LTRP que se utilizarán para descodificar la imagen actual. En otras palabras, el descodificador de vídeo 30 puede hacer coincidir los LSB de POC con los bits de valores de POC menos significativos para las imágenes almacenadas en la memoria intermedia de imágenes descodificadas e identificar la imagen con los LSB de POC correspondientes como la imagen de referencia a largo plazo que está presente en la memoria intermedia de imágenes descodificadas pero no especificada en el SPS asociado con la imagen actual.

**[0064]** El descodificador de vídeo 30 puede luego usar la imagen de referencia a largo plazo identificada por el valor de recuento del orden de imágenes obtenido para descodificar la parte de la imagen actual. En otras palabras, uno o más vectores de movimiento especificados para una unidad de codificación del fragmento actual de la imagen actual pueden hacer referencia a la LTRP identificada por los LSB de POC. El descodificador de vídeo 30 puede recuperar varios bloques de la LTRP identificados por estos vectores de movimiento, utilizando estos diversos bloques como bloques de referencia al realizar la compensación de movimiento. El descodificador de vídeo 30 puede agregar datos residuales codificados para bloques del fragmento actual de la imagen actual a estos bloques de referencia para generar bloques descodificados de datos de vídeo y así reconstruir los datos de vídeo originales. El descodificador de vídeo 30 puede almacenar los bloques descodificados de datos de vídeo en la memoria intermedia de imágenes descodificadas para su uso posterior en la descodificación de imágenes posteriores y/o para su visualización. El descodificador de vídeo luego puede descodificar otros fragmentos, si los hay, de la imagen actual, de manera similar a como se describió anteriormente.

**[0065]** Como se indicó anteriormente, el codificador de vídeo 20 puede especificar (o "señalar") los bits menos significativos de recuento del orden de imágenes delta ("LSB de POC delta") o los LSB de POC directos en las cabeceras de fragmento para que las imágenes actuales identifiquen las imágenes de referencia a largo plazo presentes en la memoria intermedia de imágenes descodificadas pero no especificadas en el conjunto de parámetros de secuencia (SPS) asociado con la imagen actual. Al especificar estos LSB de POC (delta), el codificador de vídeo 20 comúnmente analiza cada uno de estos LSB de POC (delta) especificados para garantizar que uno o más de estos LSB de POC (delta) no sean redundantes o iguales.

**[0066]** Para ilustrar, supongamos que el codificador de vídeo 20 es para especificar cinco imágenes de referencia a largo plazo en la cabecera de fragmento para que se codifique un fragmento u otra parte de la imagen actual. Además, suponga que dos de estas cinco imágenes de referencia a largo plazo se han especificado en el SPS asociado con la imagen actual, dejando las tres restantes de las cinco imágenes de referencia a largo plazo para que se señalen como LSB de POC directas para la imagen actual. El codificador de vídeo 20 puede señalar un índice para cada una de las dos primeras de las cinco imágenes de referencia a largo plazo, donde cada índice identifica una de las imágenes de referencia a largo plazo especificadas en el SPS asociado con la imagen actual.

**[0067]** El codificador de vídeo 20 puede, bajo estas suposiciones, especificar lo siguiente en la cabecera de fragmento para señalar las imágenes de referencia a largo plazo que se utilizarán al descodificar la imagen actual (donde típicamente se requiere que todas las cabeceras de fragmento de una imagen tengan la misma señalización de conjunto de imágenes de referencia):

- 1: índice A //identifica un valor POC de 4001<sub>10</sub>
- 2: índice B //identifica un valor POC de 4010<sub>10</sub>
- 3: POCLSB<sub>A</sub> //identifica un valor POC de 5001<sub>10</sub>
- 4: POCLSB<sub>B</sub> //identifica un valor POC de 4500<sub>10</sub>
- 5: POCLSB<sub>C</sub> //identifica un valor POC de 5005<sub>10</sub>

**[0068]** Las siguientes descripciones se basan en LSB de POC. Sin embargo, las mismas descripciones también se aplican a los LSB de POC delta.

**[0069]** Como se señaló anteriormente, cada uno de los elementos 1-5 anteriores identifica un valor POC diferente. Sin embargo, en varios casos, cuando solo se señalan los LSB de POC directos, puede haber problemas (debido a ambigüedades) que pueden impedir que el decodificador de vídeo 30 pueda determinar de manera efectiva qué imágenes de referencia a largo plazo se utilizaron para decodificar el fragmento actual. Suponiendo que solo se usen tres bits menos significativos para señalar  $POCLSB_A$ , entonces  $POCLSB_A$  es igual a  $001_2$ . Sin embargo, el índice A también identifica un valor POC de  $4001_{10}$ , lo cual significa que un valor POC de  $4001_{10}$  también está presente en la memoria intermedia de imágenes descodificadas. Dado que hay dos imágenes en la memoria intermedia de imágenes descodificadas que tienen un LSB de POC de  $001_2$ , el decodificador de vídeo 30 puede ser incapaz de recuperar cualquier LTRP para la imagen actual, ya que no está claro qué decodificador de vídeo LTRP 30 se concibió para la decodificación de imagen actual.

**[0070]** Como resultado, los codificadores de vídeo convencionales analizan los LSB de POC para cada imagen de referencia a largo plazo señalada en la cabecera de fragmento (incluidas las señaladas como un índice) para garantizar que no haya redundancias o LSB de POC señaladas ambiguamente para cada imagen de referencia a largo plazo señalada en la cabecera de fragmento. Cuando se identifica un LSB de POC delta ambiguo, estos codificadores de vídeo convencionales también codifican los bits más significativos de cada uno o al menos uno de los LSB de POC ambiguos, asegurando así que cada uno de los LSB de POC delta señalados en la cabecera de fragmento identifica de forma única un valor POC dentro del contexto de la cabecera de fragmento. Es decir, estos codificadores de vídeo pueden garantizar que cada uno de los LSB de POC delta señala de forma única un valor POC dentro del conjunto de valores POC especificados en la cabecera de fragmento.

**[0071]** Sin embargo, asegurarse de que los valores POC señalados en la cabecera de fragmento se señalen de manera única entre sí no es suficiente para permitir de forma adecuada al decodificador de vídeo recuperar estas imágenes de referencia a largo plazo de la memoria intermedia de imágenes descodificadas. Para ilustrar, considere el ejemplo anterior, pero también suponga que hay otra imagen de referencia a largo plazo en la memoria intermedia de imágenes descodificadas identificada por un valor POC de  $4005_{10}$ . Al verificar que no haya conflicto con ningún otro valor POC especificado en la cabecera de fragmento, el codificador de vídeo determina que este valor no señala de manera ambigua el valor POC de  $5005_{10}$  en el contexto de los otros valores POC señalados en la cabecera de fragmento y especificando  $POCLSB_C$  como  $101_2$  (que es simplemente  $5_{10}$ ) sin especificar ninguno de los bits más significativos.

**[0072]** Bajo estas suposiciones, el decodificador de vídeo 30 analizaría  $POCLSB_C$  del flujo de bits para el fragmento actual e intentará recuperar la imagen de referencia a largo plazo identificada por  $POCLSB_C$  ( $101_2$  o  $5_{10}$ ). Sin embargo, dado que hay dos imágenes en la memoria intermedia de imágenes descodificadas que tienen un LSB de POC de  $101_2$ , el decodificador de vídeo 30 puede ser incapaz de recuperar cualquier LTRP para la imagen actual, ya que no está claro qué tipo de decodificador de vídeo LTRP 30 se pretendía utilizar al decodificar la imagen actual.

**[0073]** De acuerdo con las técnicas descritas en esta divulgación, el codificador de vídeo 20 está configurado para analizar cada imagen presente en toda la memoria intermedia de imágenes descodificadas al determinar si es necesario señalar la información adicional de bits más significativos (MSB) para desambiguar los LSB de POC delta. En otras palabras, el codificador de vídeo 20 determina los LSB de POC (o los LSB de POC delta dependiendo de la implementación) para cada valor POC que identifique una imagen en la memoria intermedia de imágenes descodificadas. Al realizar el análisis de los LSB de POC señalados en la cabecera de fragmento, el codificador de vídeo 20 determina si los LSB de POC determinados para cada una de las imágenes de referencia a largo plazo señaladas en la cabecera de fragmento son únicos con respecto a cada uno de los LSB de POC determinados para cada valor de POC que identifica una imagen en la memoria intermedia de imágenes descodificadas. Si se determina que uno de estos LSB de POC señalados en la cabecera de fragmento no es único, el codificador de vídeo 20 puede entonces señalar suficientes bits más significativos (MSB) para el valor POC con el fin de habilitar el valor POC de la imagen de referencia a largo plazo para que sea correctamente identificado por el decodificador de vídeo 30. En algunos casos, un número de uno o más bits menos significativos cuando se suma a un número de uno o más bits más significativos da como resultado un total que es menor que un número total de bits para especificar el valor de recuento del orden de imágenes.

**[0074]** En funcionamiento, el codificador de vídeo 20 puede determinar, para una imagen actual de los datos de vídeo, uno o más bits menos significativos de un valor de recuento del orden de imágenes que identifica una imagen de referencia a largo plazo que se utilizará al codificar al menos una parte de la imagen actual. El codificador de vídeo 20 puede entonces determinar si estos bits menos significativos son suficientes para identificar inequívocamente la imagen de referencia a largo plazo de cualquier otra imagen de referencia almacenada en la memoria intermedia de imágenes descodificadas en la que también se almacena la imagen de referencia a largo plazo (o se almacenará dependiendo en el orden en que se almacena la imagen de referencia a largo plazo en comparación con cuando se realiza esta determinación).

**[0075]** Cuando uno o más bits menos significativos de cualquier otro valor de recuento del orden de imágenes que identifican cualquier otra imagen almacenada en esta memoria intermedia de imágenes descodificadas en la

que también se almacena la imagen de referencia a largo plazo son los mismos que el uno o más bits menos significativos determinados del valor de recuento del orden de imágenes que identifica la imagen de referencia a largo plazo, el codificador de vídeo 20 puede determinar uno o más bits más significativos del valor de recuento del orden de imágenes que identifica la imagen de referencia a largo plazo que se utilizará al codificar al menos la parte de la imagen actual. El codificador de vídeo 20 puede entonces especificar tanto el uno o más bits menos significativos determinados como el uno o más bits más significativos determinados del valor de recuento del orden de imágenes que identifica la imagen de referencia a largo plazo que se utilizará cuando se codifique al menos la parte de la imagen actual en un flujo de bits representativo de una versión codificada de los datos de vídeo. El codificador de vídeo 20 puede entonces codificar al menos la parte de la imagen actual usando la imagen de referencia a largo plazo.

**[0076]** En otras palabras, en lugar de solo determinar que los LSB de POC que identifican la imagen de referencia a largo plazo son suficientes para identificar de manera única la imagen de referencia a largo plazo de cualquier otra imagen de referencia a largo plazo especificada en la cabecera de SPS, PPS y/o fragmento. el codificador de vídeo 20 puede determinar que los LSB de POC que identifican la imagen de referencia a largo plazo son suficientes para identificar de manera única la imagen de referencia a largo plazo de cualquier otra imagen de referencia a largo plazo almacenada en toda la DPB. Nuevamente, un número del uno o más bits menos significativos cuando se suma a un número del uno o más bits más significativos da como resultado un total que es menor que un número total de bits para especificar el valor de recuento del orden de imágenes. A este respecto, el codificador de vídeo 20 puede evitar aquellos casos que den lugar a problemas señalados anteriormente en los que el descodificador de vídeo no puede identificar correctamente una imagen de referencia a largo plazo mediante los LSB de POC señalados cuando dos o más imágenes de referencia a largo plazo que tienen el mismo LSB de POC se almacenan en la DPB pero solo una de estas imágenes de referencia a largo plazo se especificó en la cabecera de SPS, PPS y/o fragmento. Como resultado, al implementar estas técnicas, el codificador de vídeo 20 puede codificar de manera más robusta los datos de vídeo en comparación con los codificadores de vídeo convencionales que solo identifican cuándo codificar de manera más robusta la imagen de referencia a largo plazo señalada con respecto a las imágenes de referencia a largo plazo señaladas en la cabecera de SPS, PPS y/o fragmento.

**[0077]** El descodificador de vídeo 30 puede recibir este flujo de bits representativo de la versión codificada de los datos de vídeo (que también puede denominarse "datos de vídeo codificados"). El descodificador de vídeo 30 luego puede determinar, para una imagen actual de estos datos de vídeo codificados, uno o más bits menos significativos de un recuento del orden de imágenes que identifica una imagen de referencia a largo plazo que se utilizará cuando se descodifique al menos una parte de la imagen actual. Estos uno o más bits menos significativos pueden no identificar de manera única el valor de recuento del orden de imágenes de la imagen de referencia a largo plazo con respecto al uno o más bits menos significativos de cualquier otro valor de recuento del orden de imágenes que identifique cualquier otra imagen presente en una memoria intermedia de imágenes descodificadas en el que también se almacena la imagen de referencia a largo plazo. Como se señaló anteriormente, los bits menos significativos del valor de recuento del orden de imágenes pueden identificar de manera única el valor de recuento del orden de imágenes de la imagen de referencia a largo plazo con respecto a los bits menos significativos de cualquier otro valor de recuento del orden de imágenes calculado con respecto al valor de recuento del orden de imágenes valor que identifica cualquier otra imagen de referencia a largo plazo que se utilizará como candidato para descodificar la parte de la imagen actual.

**[0078]** En otras palabras, el análisis previo realizado por los codificadores de vídeo convencionales que no implementan las técnicas descritas en esta divulgación no habría identificado a los LSB de POC como ambiguos, dado que todos los LSB de POC para cualquier otra imagen de referencia a largo plazo señalada en la cabecera de fragmento identificaron de forma única sus respectivas imágenes de referencia a largo plazo. Sin embargo, esto no significa que los LSB de POC identifiquen de forma única las imágenes de referencia a largo plazo con respecto a la memoria intermedia de imágenes descodificadas completa. Debido a que el codificador de vídeo 20 se ha modificado para implementar las técnicas descritas en esta divulgación, de acuerdo con un ejemplo de esta divulgación, el codificador de vídeo 20 extiende este análisis para considerar todas y cada una de las imágenes almacenadas en la memoria intermedia de imágenes descodificadas. Como resultado, el codificador de vídeo 20 puede, como se describió anteriormente, señalar adicionalmente los MSB de POC incluso aunque los LSB de POC sean únicos para un subconjunto dado de imágenes almacenadas en la memoria intermedia de imágenes descodificadas (por ejemplo, el subconjunto señalado en la cabecera de fragmento).

**[0079]** A este respecto, incluso cuando los LSB de POC son únicos para un subconjunto dado de las imágenes almacenadas en la memoria intermedia de imágenes descodificadas, el descodificador de vídeo 30 determina uno o más bits más significativos del valor de recuento del orden de imágenes que identifica la imagen de referencia a largo plazo. El uno o más bits más significativos en combinación con uno o más bits menos significativos determinados es suficiente para distinguir el valor de recuento del orden de imágenes (POC) que identifica la imagen de referencia a largo plazo de cualquier otro valor de recuento del orden de imágenes que identifica cualquier otra imagen en la memoria intermedia de imágenes descodificadas. El descodificador de vídeo 30 puede luego recuperar la imagen de referencia a largo plazo de la memoria intermedia de imágenes descodificadas

basándose en uno o más de los bits menos significativos determinados del valor de recuento del orden de imágenes y el uno o más bits más significativos determinados del valor de recuento del orden de imágenes.

[0080] El descodificador de vídeo 30 puede luego usar la imagen de referencia a largo plazo identificada por el valor de recuento del orden de imágenes obtenido para descodificar la parte de la imagen actual. En otras palabras, uno o más vectores de movimiento especificados para una unidad de codificación del fragmento actual de la imagen actual pueden hacer referencia a la imagen de referencia a largo plazo identificada por el valor de recuento del orden de imágenes obtenido. El descodificador de vídeo puede recuperar varios bloques de la imagen de referencia a largo plazo identificada por estos vectores de movimiento, utilizando estos diversos bloques como bloques de referencia al realizar la compensación de movimiento. El descodificador de vídeo puede agregar datos residuales codificados para bloques del fragmento actual de la imagen actual a estos bloques de referencia para generar bloques descodificados de datos de vídeo. El descodificador de vídeo puede almacenar los bloques descodificados de datos de vídeo en la memoria intermedia de imágenes descodificadas para su uso posterior en la descodificación de imágenes posteriores y/o para su visualización. El descodificador de vídeo luego puede descodificar otras partes, si las hay, de la imagen actual, de manera similar a como se describió anteriormente.

[0081] La FIG. 2 es un diagrama de bloques que ilustra una configuración de ejemplo de codificador de vídeo 20 que está configurado para implementar las técnicas de esta divulgación. La FIG. 2 se proporciona para propósitos de explicación y no debería considerarse limitadora de las técnicas tales como las ampliamente ejemplificadas y descritas en esta divulgación. 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, como se ha indicado anteriormente, ser aplicables a otras normas o procedimientos de codificación.

[0082] 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 un módulo de predicción 100, un módulo de generación residual 102, un módulo de transformada 104, un módulo de cuantificación 106, un módulo de cuantificación inversa 108, un módulo de transformada inversa 110, un módulo de reconstrucción 112, una memoria intermedia de imágenes descodificadas 114, y un módulo de codificación por entropía 116. El módulo de predicción 100 incluye un módulo de estimación de movimiento 122, un módulo de compensación de movimiento 124 y un módulo de predicción intra 126.

[0083] En otros ejemplos, el codificador de vídeo 20 puede incluir más, menos o diferentes componentes funcionales. Por ejemplo, el codificador de vídeo 20 puede incluir un filtro de desbloqueo para filtrar la salida del módulo de reconstrucción 112 para eliminar las distorsiones de bloqueo del vídeo reconstruido. Además, el módulo de estimación de movimiento 122 y el módulo de compensación de movimiento 124 pueden estar altamente integrados, pero están representados en el ejemplo de la FIG. 4 de forma separada con fines explicativos.

[0084] El codificador de vídeo 20 puede recibir datos de vídeo. En varios ejemplos, 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 secuencias de imágenes. Las imágenes pueden incluir visualizaciones de textura y visualizaciones en profundidad. Para codificar los datos de vídeo, el codificador de vídeo 20 puede realizar una operación de codificación en cada secuencia de imágenes. Como parte de realizar la operación de codificación en una secuencia de imágenes, el codificador de vídeo 20 puede realizar operaciones de codificación en cada imagen dentro de la secuencia de 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. Cuando el codificador de vídeo 20 realiza una operación de codificación en un fragmento, el codificador de vídeo 20 genera un fragmento codificado. El fragmento codificado es el fragmento en su forma codificada. El fragmento codificado puede incluir una cabecera de fragmento y datos de fragmento. La cabecera de fragmento puede contener elementos sintácticos asociados con el fragmento.

[0085] Como parte de realizar una operación de codificación en un fragmento, el codificador de vídeo 20 puede realizar operaciones de codificación en bloques de árbol en el 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. En otras palabras, el bloque de árbol codificado puede ser un bloque de árbol en su forma codificada. Las técnicas, aunque se describieron anteriormente con respecto a H.264/AVC, pueden también aplicarse como se describe a continuación con respecto a HEVC. A este respecto, las técnicas no deben limitarse a H.264 o HEVC, sino que pueden aplicarse tanto en el contexto de H.264/AVC como en HEVC.

[0086] Como parte de realizar una operación de codificación en un bloque de árbol, el módulo de predicción 100 puede realizar la división en árbol cuaternario en el bloque de árbol para dividir el bloque de árbol en CU progresivamente más pequeños. Por ejemplo, el módulo de predicción 100 puede dividir un bloque de árbol cuaternario sub-CU de igual tamaño, dividir una o más de las sub-CU en cuatro sub-sub-CU de igual tamaño, etc.

[0087] Los tamaños de las CU pueden variar desde 8x8 píxeles hasta el tamaño del bloque de árbol, con un máximo de 64x64 píxeles o más. En la presente divulgación, "NxN" y "N por N" pueden usarse indistintamente para

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

**[0088]** Además, como parte de la realización de la operación de codificación para un bloque de árbol, el módulo 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 el módulo de predicción 100 divide el bloque de árbol cuaternario sub-CU, el nodo raíz tiene cuatro nodos secundarios en la estructura de datos en árbol cuaternario. Cada uno de los nodos secundarios corresponde a una de las sub-CU. Si el módulo de predicción 100 divide una de las sub-CU en cuatro sub-sub-CU, el nodo correspondiente a la sub-CU puede tener cuatro nodos secundarios, cada uno de los cuales corresponde a una de las sub-sub-CU.

**[0089]** Cada nodo de la estructura de datos en árbol cuaternario puede proporcionar datos sintácticos para la CU correspondiente. Por ejemplo, un nodo en el árbol cuaternario puede incluir un indicador de división, que indica si la CU correspondiente al nodo está dividida (es decir, partida) en cuatro sub-CU. Los elementos sintácticos para una CU pueden definirse de manera recursiva y pueden depender de si la CU está dividida en varias sub-CU. Una CU que no está dividida puede corresponder a un nodo hoja en la estructura de datos en árbol cuaternario. Un nodo de hoja en la estructura de datos de árbol cuaternario puede denominarse "nodo de codificación". Un bloque de árbol codificado puede incluir datos basados en la estructura de datos en árbol cuaternario para un bloque de árbol correspondiente. Un bloque de árbol codificado es un bloque de árbol en su forma codificada. Un bloque de árbol codificado corresponde a un bloque de árbol cuando el bloque de árbol codificado es el bloque de árbol en su forma codificada.

**[0090]** El codificador de vídeo 20 puede realizar operaciones de codificación en cada CU no dividida del 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 versión codificada de la CU no dividida.

**[0091]** Como parte de la realización de una operación de codificación en una CU, el módulo de estimación de movimiento 122 y el módulo de compensación de movimiento 124 pueden realizar predicción inter en la CU. En otras palabras, el módulo de estimación de movimiento 122 y el módulo de compensación de movimiento 124 pueden generar datos de predicción para la CU basados en muestras descodificadas de imágenes de referencia distintas de la imagen que contiene la CU. La predicción inter puede proporcionar compresión temporal.

**[0092]** Para realizar la predicción inter en una CU, el módulo de estimación de movimiento 122 puede dividir la CU en una o más unidades de predicción (PU). El codificador de vídeo 20 y el descodificador 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 descodificador 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 descodificador de vídeo 30 también pueden soportar división asimétrica para tamaños de PU de  $2N \times nU$ ,  $2N \times nD$ ,  $nL \times 2N$  y  $nR \times 2N$ . En algunos ejemplos, el módulo de estimación de movimiento 122 puede dividir una CU en PU a lo largo de un límite que no se encuentra con los lados de la CU en ángulos rectos.

**[0093]** El módulo de estimación de movimiento 122 puede realizar una operación de estimación de movimiento con respecto a cada PU de una CU. Cuando el módulo de estimación de movimiento 122 realiza una operación de estimación de movimiento con respecto a una PU, el módulo de estimación de movimiento 122 puede generar uno o más vectores de movimiento para la PU. Por ejemplo, los fragmentos pueden ser fragmentos I, fragmentos P o fragmentos B. El módulo de estimación de movimiento 122 y el módulo de compensación de movimiento 124 pueden realizar diferentes operaciones para una PU de una CU dependiendo de si la CU está en un fragmento I, un fragmento P o un fragmento B. En un fragmento en I, todas las CU tienen predicción intra. Por lo tanto, si la CU está en un fragmento I, el módulo de estimación de movimiento 122 y el módulo de compensación de movimiento 124 no realizan predicción inter en la PU.

**[0094]** Si la CU está en un fragmento P, la imagen que contiene la CU 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 imágenes posteriores en orden de descodificación. Cuando el módulo de estimación de movimiento 122 realiza la operación de estimación de movimiento con respecto a una PU en un fragmento P, el módulo de estimación de movimiento 122 busca las imágenes de referencia en la lista 0 para una muestra de referencia para la PU. La muestra de referencia de la PU puede ser un conjunto de valores de píxel que se corresponda más aproximadamente a los valores de píxel de la PU. El módulo de estimación de movimiento 122 puede usar una variedad de métricas para determinar cuán estrechamente un conjunto de valores de píxel en una imagen de referencia se corresponde con los valores de píxel de una PU. Por ejemplo, el módulo de estimación de movimiento 122 puede determinar cuán estrechamente un conjunto de valores de píxel en una imagen de

referencia se corresponde con los valores de píxel de una PU mediante la suma de la diferencia absoluta (SAD), la suma de la diferencia cuadrada (SSD) u otra diferencia métrica.

**[0095]** Después de identificar una muestra de referencia de una PU de una CU en un fragmento P, el módulo de estimación de movimiento 122 puede generar un índice de referencia que indica la imagen de referencia de la lista 0 que contiene la muestra de referencia y un vector de movimiento que indica un desplazamiento espacial entre la PU y la muestra de referencia. En varios ejemplos, el módulo de estimación de movimiento 122 puede generar vectores de movimiento con diferentes grados de precisión. Por ejemplo, el módulo de estimación de movimiento 122 puede generar vectores de movimiento con un cuarto de precisión de píxel, un octavo de precisión de píxel u otra fracción de precisión de píxel. El módulo de estimación de movimiento 122 puede enviar información de movimiento para la PU al módulo de codificación por entropía 56 y al módulo de compensación de movimiento 124. La información de movimiento para la PU puede incluir el índice de referencia y el vector de movimiento de la PU. El módulo de compensación de movimiento 124 puede usar la información de movimiento de la PU de la CU para identificar y recuperar las muestras de referencia de las PU. El módulo de compensación de movimiento 124 puede entonces utilizar valores de píxel de muestras de referencia de PU para generar los datos de predicción para la CU.

**[0096]** Si la CU está en un fragmento B, la imagen que contiene la CU puede estar asociada con dos listas de imágenes de referencia, denominadas "lista 0" y "lista 1." Cada una de las imágenes de referencia en la lista 0 contiene muestras que pueden usarse para la predicción inter de imágenes posteriores en orden de descodificación. Las imágenes de referencia de la lista 1 se producen antes de la imagen en orden de descodificación pero después de la imagen en orden de presentación. 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.

**[0097]** Además, si la PU está en un fragmento B, el módulo de estimación de movimiento 122 puede llevar a cabo la predicción unidireccional o la predicción bidireccional para las PU de la CU. Cuando el módulo de estimación de movimiento 122 realiza la predicción unidireccional para la PU, el módulo de estimación de movimiento 122 puede buscar en las imágenes de referencia de la lista 1 una muestra de referencia para la PU. El módulo de estimación de movimiento 122 puede entonces generar un índice de referencia que indica la imagen de referencia en la lista 1 que contiene la muestra de referencia y un vector de movimiento que indica un desplazamiento espacial entre la PU y la muestra de referencia. El módulo de estimación de movimiento 122 puede enviar información de movimiento para las PU de la CU al módulo de codificación por entropía 56 y al módulo de compensación de movimiento 124. La información de movimiento para la PU puede incluir el índice de referencia, un indicador de dirección de predicción y el vector de movimiento del 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. El módulo de compensación de movimiento 124 puede usar la información de movimiento de la PU de la CU para identificar y recuperar las muestras de referencia de las PU. El módulo de compensación de movimiento 124 puede entonces utilizar valores de píxel de muestras de referencia de PU para generar los datos de predicción para la CU.

**[0098]** Cuando el módulo de estimación de movimiento 122 realiza la predicción bidireccional para una PU, el módulo de estimación de movimiento 122 puede buscar las imágenes de referencia de la lista 0 para una muestra de referencia para la PU y también puede buscar las imágenes de referencia de la lista 1 para otra muestra de referencia para la PU. El módulo de estimación de movimiento 122 puede entonces generar índices de referencia que indican las muestras de referencia y los vectores de movimiento que indican desplazamientos espaciales entre las muestras de referencia y la PU. El módulo de estimación de movimiento 122 puede enviar información de movimiento de la PU al módulo de codificación por entropía 116 y al módulo de compensación de movimiento 124. La información de movimiento para la PU puede incluir los índices de referencia y los vectores de movimiento de la PU. El módulo de compensación de movimiento 124 puede utilizar la información de movimiento para identificar y recuperar las muestras de referencia de las PU. El módulo de compensación de movimiento 124 puede entonces interpolar los valores de píxel de los datos de predicción de la CU a partir de los valores de píxel en las muestras de referencia de las PU de la CU.

**[0099]** En general, el módulo de predicción 100 puede implementar las técnicas descritas en esta divulgación para formar el RPS señalado anteriormente y generar los elementos sintácticos de nivel de fragmento y/o SPS que identifican imágenes de referencia en el RPS. El módulo de predicción 100 puede representar una unidad de hardware y combinación de hardware y software (que puede incluirse dentro de una unidad de hardware o combinación de hardware y software más grande) que implementa las técnicas descritas con más detalle anteriormente.

**[0100]** Como parte de la realización de una operación de codificación en una CU, el módulo de predicción intra 126 puede llevar a cabo la predicción intra en la CU. En otras palabras, el módulo de predicción intra 126 puede generar datos de predicción para la CU basándose en los valores de píxel descodificados de otras CU. La predicción intra puede proporcionar compresión espacial.



**[0101]** Para realizar la predicción intra en una CU, el módulo de predicción intra 126 puede usar múltiples modos de predicción intra para generar múltiples conjuntos de datos de predicción para la CU. Cuando el módulo de predicción intra 126 usa un modo de predicción intra para generar un conjunto de datos de predicción para una CU, el módulo de predicción intra 126 puede dividir la CU en una o más PU. El módulo de predicción intra 126 puede entonces, para cada una de las PU, extender los valores de píxel de las PU contiguas a través de la PU en una dirección y/o gradiente asociados 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. El módulo de predicción intra 46 puede usar varios números de modos de predicción intra, por ejemplo, 33 modos de predicción intra direccional, dependiendo del tamaño de la CU.

**[0102]** El módulo de predicción intra 126 puede seleccionar uno de los conjuntos de datos de predicción para la CU. En varios ejemplos, el módulo de predicción intra 126 puede seleccionar el conjunto de datos de predicción para la CU de varias maneras. Por ejemplo, el módulo de predicción intra 126 puede seleccionar el conjunto de datos de predicción para la CU calculando las tasas de distorsión para los conjuntos de datos de predicción y seleccionando el conjunto de datos de predicción que tiene la tasa de distorsión más baja.

**[0103]** El módulo de predicción 100 puede seleccionar los datos de predicción para una CU entre los datos de predicción generados por el módulo de compensación de movimiento 124 para la CU o los datos de predicción generados por el módulo de predicción intra 126 para la CU. En algunos ejemplos, el módulo de predicción 100 selecciona los datos de predicción para la CU basándose en el error (es decir, la distorsión) de los conjuntos de datos de predicción.

**[0104]** Después de que el módulo de predicción 100 seleccione los datos de predicción para una CU, el módulo de generación residual 102 puede generar datos residuales para la CU restando los datos de predicción seleccionados de la CU de los valores de píxel de la CU. Los datos residuales de una CU pueden incluir bloques residuales en 2D que corresponden a diferentes componentes de píxel de los píxeles en la CU. Por ejemplo, los datos residuales pueden incluir un bloque de residual que corresponde a diferencias entre componentes de luminancia de píxeles en los datos de predicción de la CU y componentes de luminancia de píxeles en los píxeles originales de la CU. Además, los datos residuales de la CU pueden incluir bloques residuales que corresponden a las diferencias entre componentes de crominancia de píxeles en los datos de predicción de la CU y los componentes de crominancia de los píxeles originales de la CU.

**[0105]** Una CU puede tener una o más unidades de transformada (TU). Cada TU de una CU puede corresponder a una parte diferente de los datos residuales de la CU. Los tamaños de las TU de una CU pueden o no estar basados en los tamaños de las PU de la CU. En algunos ejemplos, una CU puede subdividirse en unidades más pequeñas mediante una estructura de árbol cuaternario conocida como "árbol cuaternario residual" (RQT). Las TU pueden corresponder a nodos de hoja del RQT.

**[0106]** El módulo de transformada 104 puede generar uno o más bloques de coeficientes para cada TU de una CU aplicando una transformada a los datos residuales correspondientes a la TU. Cada uno de los bloques de coeficientes puede ser una matriz en 2D de coeficientes. En varios ejemplos, el módulo de transformada 104 puede aplicar diversas transformadas a los datos residuales correspondientes a una TU. Por ejemplo, el módulo de transformada puede aplicar una transformada de coseno discreta (DCT), una transformada direccional o una transformada conceptualmente similar.

**[0107]** Después de que el módulo de transformada 104 genere un bloque de coeficientes para un TU, el módulo de cuantificación 106 puede cuantificar los coeficientes en el bloque de coeficientes. La cuantificación en general se refiere a un proceso en el que los coeficientes en un bloque de coeficientes se cuantifican para reducir posiblemente la cantidad de datos usados para representar los coeficientes, proporcionando compresión adicional. La cuantificación puede reducir la profundidad de bits asociada con algunos o todos los coeficientes. Por ejemplo, un valor de  $n$  bits puede redondearse hacia abajo hasta un valor de  $m$  bits durante la cuantificación, donde  $n$  es mayor que  $m$ .

**[0108]** El módulo de cuantificación inversa 108 y el módulo de transformada inversa 110 pueden aplicar la cuantificación inversa y transformación inversa al bloque de coeficientes, respectivamente, para reconstruir datos residuales a partir del bloque de coeficientes. El módulo de reconstrucción 112 puede agregar los datos residuales reconstruidos a los datos de predicción generados por el módulo de compensación de movimiento 124 o el módulo de predicción intra 126 para producir un bloque de vídeo reconstruido para su almacenamiento en la memoria intermedia de imágenes descodificadas 114. El módulo de estimación de movimiento 122 y el módulo de compensación de movimiento 124 pueden utilizar una imagen de referencia que contiene el bloque de vídeo reconstruido para realizar predicción inter en las CU de las imágenes posteriores. Además, el módulo de predicción intra 126 puede usar valores de píxel reconstruidos de las CU de la imagen actual para realizar la predicción intra.

**[0109]** El módulo de codificación por entropía 116 puede recibir datos desde otros componentes funcionales del codificador de vídeo 20. Por ejemplo, el módulo de codificación por entropía 116 puede recibir bloques de

coeficientes del módulo de cuantificación 106 y puede recibir elementos sintácticos del módulo de predicción 100. Cuando el módulo de codificación por entropía 116 recibe los datos, el módulo 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 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.

**[0110]** El módulo de codificación por entropía 116 genera un flujo de bits que incluye una serie de unidades NAL. Como se ha analizado anteriormente, 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. Cada unidad NAL de fragmento codificado en el flujo de bits contiene un fragmento codificado. Un fragmento codificado incluye cabecera de fragmento y datos de fragmento codificados. Los datos del fragmento pueden incluir bloques de árbol codificados. Los bloques de árbol codificados pueden incluir una o más CU codificadas. Cada CU codificada puede incluir uno o más bloques de coeficientes codificados por entropía. El módulo de codificación por entropía 116 genera el flujo de bits para la descodificación en tiempo real o casi en tiempo real o para el almacenamiento y la descodificación posterior mediante el descodificador 30.

**[0111]** Como se describió anteriormente, el codificador de vídeo 20 puede codificar datos de vídeo de acuerdo con los primeros aspectos de las técnicas de esta divulgación. Para codificar datos de vídeo de acuerdo con el primer aspecto de las técnicas, el módulo de estimación de movimiento 122 del codificador de vídeo 20 puede determinar, para una imagen actual de los datos de vídeo, una imagen de referencia a largo plazo que se utilizará cuando se codifique al menos una parte de una imagen actual de los datos de vídeo de la forma descrita anteriormente. En algunos casos, esta imagen de referencia a largo plazo es una imagen de referencia a largo plazo que está presente en la memoria intermedia de imágenes descodificadas 114, pero que no se especifica como una o más imágenes de referencia a largo plazo en un conjunto de parámetros de secuencia asociado a la imagen actual.

**[0112]** El módulo de estimación de movimiento 122 puede entonces determinar un número de bits que se usarán para representar uno o más LSB de un valor POC que identifica la imagen de referencia a largo plazo en un flujo de bits representativo de una versión codificada de los datos de vídeo. En otras palabras, el módulo de estimación de movimiento 122 puede determinar la "v" en la codificación binaria sin signo denotada como "u(v)" con la "u", denotando nuevamente la codificación binaria sin signo. A menudo, este número de LSB se define basándose en el número de LSB necesarios para desambiguar la LTRP de otras LTRP o cualquier otra imagen de referencia almacenada en la DPB. En otras palabras, este número puede variar en un fragmento, una imagen o un GOP. En algunos casos, el número de LSB puede definirse para un flujo de bits completo, o partes del mismo. En algunos casos, el número de LSB utilizados para identificar esta diferencia es estático o predefinido, como en la norma HEVC. En algunos casos, el número de LSB utilizados para identificar la LTRP puede obtenerse basándose en los elementos sintácticos determinados para la imagen actual y/o las imágenes codificadas previamente de los datos de vídeo, donde el módulo de estimación de movimiento 122 tal vez no necesite señalar el número de bits utilizados para representar este LSB de POC dado que el módulo de estimación de movimiento 122 puede realizar un proceso de obtención similar para obtener el número "v" de bits utilizados para representar el LSB de POC.

**[0113]** En cualquier caso, el módulo 122 de estimación de movimiento puede pasar estos uno o más LSB del valor POC al módulo de codificación por entropía 116, que puede especificar el uno o más LSB del valor POC que identifica la LTRP en el flujo de bits utilizando el número determinado de bits utilizados para representar uno o más bits menos significativos del valor de recuento del orden de imágenes que identifica la imagen de referencia a largo plazo. En otras palabras, el módulo de codificación por entropía 116 puede especificar los LSB de POC anotados anteriormente en el flujo de bits utilizando el número "v" de bits como un número binario sin signo. Típicamente, el módulo de codificación por entropía 116 especifica uno o más bits menos significativos del valor de recuento del orden de imágenes que identifica la imagen de referencia a largo plazo en una cabecera de fragmento asociada con al menos la parte de la imagen actual.

**[0114]** En algunos casos, el módulo de estimación de movimiento 122 determina el número de bits que se usarán para representar uno o más bits menos significativos del valor de recuento del orden de imágenes que identifica la imagen de referencia a largo plazo en el flujo de bits basándose en los elementos sintácticos previamente determinados para uno o más de la imagen actual y las imágenes previamente codificadas de los datos de vídeo para evitar especificar en el flujo de bits el número determinado de bits que se utilizarán para representar el uno o más bits menos significativos del valor de recuento del orden de imágenes que identifica la imagen de referencia a largo plazo. En algunos casos, como cuando el número de bits "v" no se obtiene de otros elementos sintácticos (a menudo como una opción de implementación), el módulo de codificación por entropía 116 también especifica el número determinado de bits utilizados para representar el uno o más bits menos significativos del valor de recuento del orden de imágenes que identifica la imagen de referencia a largo plazo en el flujo de bits. Los diversos módulos del codificador de vídeo 20, que incluyen el módulo de estimación de movimiento 122, la unidad de compensación de movimiento 124, el sumador 102, el módulo de transformada 104, el módulo de cuantificación 106 y el módulo de codificación por entropía 116, pueden codificar al menos la parte de la imagen actual utilizando la LTRP.

**[0115]** El codificador de vídeo 20 también puede implementar el segundo aspecto de las técnicas descritas en esta divulgación, potencialmente en combinación con el primer aspecto de las técnicas descritas en esta divulgación. De acuerdo con el segundo aspecto de las técnicas, el módulo de estimación de movimiento 122 puede determinar, para la imagen actual de los datos de vídeo, uno o más bits menos significativos de un valor de recuento del orden de imágenes que identifica una imagen de referencia a largo plazo para ser utilizada al codificar al menos una parte de la imagen actual. El módulo de estimación de movimiento 122 puede entonces determinar si estos bits menos significativos son suficientes para identificar de manera inequívoca la imagen de referencia a largo plazo de cualquier otra imagen de referencia almacenada en la memoria intermedia de imágenes descodificadas 114 en el que la imagen de referencia a largo plazo también se almacena (o se almacenará dependiendo del orden en que se almacena la imagen de referencia a largo plazo en comparación con cuando se realiza esta determinación).

**[0116]** El módulo de estimación de movimiento 122 puede determinar si estos bits menos significativos son suficientes para identificar de manera inequívoca la imagen de referencia a largo plazo de cualquier otra imagen de referencia almacenada en la memoria intermedia de imágenes descodificadas 114, al menos en parte, determinando los bits menos significativos para al menos un valor de recuento del orden de imágenes que identifica una respectiva de las otras imágenes almacenadas en la memoria intermedia de imágenes descodificadas 114. El módulo de estimación de movimiento 122 puede entonces determinar si los bits menos significativos para este valor de recuento del orden de imágenes son los mismos que los bits menos significativos determinados del valor de recuento del orden de imágenes que identifica la imagen de referencia a largo plazo. Si estos dos bits menos significativos del valor de recuento del orden de imágenes son iguales, el módulo de estimación de movimiento 122 puede entonces determinar que los bits menos significativos del valor de recuento del orden de imágenes que identifican la imagen de referencia a largo plazo que se utilizará al codificar y descodificar al menos la parte de la imagen actual no es suficiente para identificar de manera única la imagen de referencia a largo plazo apropiada. Sin embargo, si estos dos bits menos significativos del valor de recuento del orden de imágenes no son iguales, el módulo de estimación de movimiento 122 puede continuar determinando los bits menos significativos de los valores de recuento del orden de imágenes que identifican otra imagen de referencia a largo plazo almacenada en la memoria intermedia de imágenes descodificadas 114, comparando los dos bits menos significativos del recuento del orden de imágenes hasta que se encuentra una coincidencia o todos los valores de recuento del orden de imágenes que identifican las imágenes almacenadas en la memoria intermedia de imágenes descodificadas 114 han pasado por esta comparación.

**[0117]** Cuando uno o más bits menos significativos de cualquier otro valor de recuento del orden de imágenes que identifican cualquier otra imagen almacenada en la memoria intermedia de imágenes descodificadas 114 en la que también se almacena la imagen de referencia a largo plazo son los mismos que el uno o más bits menos significativos determinados del valor de recuento del orden de imágenes que identifica la imagen de referencia a largo plazo, el módulo de estimación de movimiento 122 puede determinar uno o más bits más significativos del valor de recuento del orden de imágenes que identifica la imagen de referencia a largo plazo que se utilizará al codificar al menos la parte de imagen actual. Para determinar el uno o más bits más significativos del valor de recuento del orden de imágenes, el módulo de estimación de movimiento 122 puede determinar los bits más significativos del valor de recuento del orden de imágenes que identifica la imagen de referencia a largo plazo que se utilizará al codificar al menos la parte de la imagen actual, de modo que los bits más significativos en combinación con los bits menos significativos determinados sean suficientes para identificar de forma única el valor de recuento del orden de imágenes que identifica la imagen de referencia a largo plazo de cualquier otro valor de recuento del orden de imágenes que identifica cualquier otra imagen en la memoria intermedia de imágenes descodificadas.

**[0118]** El módulo de estimación de movimiento 122 puede a continuación reenviar estos bits menos significativos y los bits más significativos al módulo de codificación por entropía 116. El módulo de codificación por entropía 116 puede especificar uno o más de los bits menos significativos determinados y uno o más de los bits más significativos determinados del valor de recuento del orden de imágenes que identifica la imagen de referencia a largo plazo que se utilizará al codificar al menos la parte de la imagen actual en un flujo de bits representativo de una versión codificada de los datos de vídeo. Los diversos módulos del codificador de vídeo 20, que incluyen el módulo de estimación de movimiento 122, la unidad de compensación de movimiento 124, el sumador 102, el módulo de transformada 104, el módulo de cuantificación 106 y el módulo de codificación por entropía 116, pueden codificar al menos la parte de la imagen actual utilizando la LTRP.

**[0119]** En otras palabras, en lugar de solo determinar que los LSB de POC que identifican la imagen de referencia a largo plazo son suficientes para identificar de manera única la imagen de referencia a largo plazo de cualquier otra imagen de referencia a largo plazo especificada en la cabecera de SPS, PPS y/o fragmento, el módulo de estimación de movimiento 122 puede determinar que los LSB de POC que identifican la imagen de referencia a largo plazo son suficientes para identificar de manera única la imagen de referencia a largo plazo de cualquier otra imagen de referencia a largo plazo almacenada en la memoria intermedia de imágenes descodificadas 114. A este respecto, el codificador de vídeo 20 puede evitar aquellos casos que den lugar a problemas señalados anteriormente en los que el descodificador de vídeo no puede identificar correctamente una imagen de referencia

a largo plazo mediante los LSB de POC señalados cuando dos o más imágenes de referencia a largo plazo que tienen el mismo LSB de POC se almacenan en la DPB pero solo una de estas imágenes de referencia a largo plazo se especificó en la cabecera de SPS, PPS y/o fragmento. Como resultado, al implementar estas técnicas, el codificador de vídeo 20 puede codificar de manera más robusta los datos de vídeo en comparación con los codificadores de vídeo convencionales que solo identifican cuándo codificar de manera más robusta la imagen de referencia a largo plazo señalada con respecto a las imágenes de referencia a largo plazo señaladas en la cabecera de SPS, PPS y/o fragmento.

**[0120]** La FIG. 3 es un diagrama de bloques que ilustra una configuración de ejemplo de un descodificador de vídeo 30. La FIG. 3 se proporciona para propósitos de explicación y no se limita a las técnicas como las ampliamente ejemplificadas y descritas en esta divulgación. Para los propósitos de explicación, esta divulgación describe el descodificador de vídeo 30 en el contexto de la codificación HEVC. Sin embargo, las técnicas de esta divulgación pueden ser aplicables a otras normas o procedimientos de codificación.

**[0121]** 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 un módulo de descodificación por entropía 150, un módulo de predicción 152, un módulo de cuantificación inversa 154, un módulo de transformada inversa 156, un módulo de reconstrucción 158 y una memoria intermedia de imágenes descodificadas 160. El módulo de predicción 152 incluye un módulo de compensación de movimiento 162 y un módulo 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. Por ejemplo, el descodificador de vídeo 30 puede incluir un filtro de desbloqueo para filtrar la salida del módulo de reconstrucción 158 para eliminar las distorsiones de bloqueo del vídeo reconstruido.

**[0122]** El descodificador de vídeo 30 puede recibir un flujo de bits que comprende datos de vídeo codificados. Cuando el descodificador de vídeo 30 recibe el flujo de bits, el módulo de descodificación por entropía 150 realiza una operación de análisis en el flujo de bits. Como resultado de realizar la operación de análisis en el flujo de bits, el módulo de descodificación por entropía 150 puede generar elementos sintácticos descodificados por entropía. Los elementos sintácticos descodificados por entropía pueden incluir bloques de coeficientes descodificados por entropía. El módulo de predicción 152, el módulo de cuantificación inversa 154, el módulo de transformada inversa 156 y el módulo de reconstrucción 158 pueden realizar una operación de descodificación que utiliza los elementos sintácticos para generar datos de vídeo descodificados.

**[0123]** En algunos ejemplos, el módulo de descodificación por entropía 150 puede analizar los elementos sintácticos relacionados con los bits de un recuento del orden de imágenes utilizado para identificar imágenes de referencia y formar conjuntos de imágenes de referencia de acuerdo con diversos aspectos de las técnicas descritas con más detalle anteriormente. El módulo de descodificación por entropía 150 puede proporcionar este elemento sintáctico al módulo de predicción 152, que puede implementar las técnicas descritas en esta divulgación para determinar el recuento del orden de imágenes a partir de los bits, formar un conjunto de imágenes de referencia que incluya una imagen de referencia asociada o identificada por el recuento del orden de imágenes y descodificar una o más imágenes codificadas de los datos de vídeo usando el conjunto de imágenes de referencia.

**[0124]** 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 en el flujo de bits, el módulo de descodificación por entropía 150 puede realizar operaciones de análisis 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 SEI NAL, etc. Un conjunto de parámetros de secuencia es una estructura sintáctica que contiene elementos sintácticos que se aplican a cero o más secuencias de vídeo codificadas completas. Un conjunto de parámetros de imagen es una estructura sintáctica que contiene elementos sintácticos que se aplican a cero o más imágenes codificadas. Un conjunto de parámetros de imagen asociado con una imagen dada puede incluir un elemento sintáctico que identifica un conjunto de parámetros de secuencia asociado con la imagen dada.

**[0125]** Además, las unidades NAL del flujo de bits pueden incluir unidades NAL de fragmentos codificados. Como parte de realizar la operación de análisis en el flujo de bits, el módulo de descodificación por entropía 150 puede realizar operaciones de análisis que extraen y descodifican por entropía fragmentos codificados a partir de las unidades de NAL de fragmentos codificado. 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 de fragmento pueden incluir un elemento sintáctico que identifica un conjunto de parámetros de imagen asociado con una imagen que contiene el fragmento. El módulo de descodificación por entropía 150 puede realizar una operación de descodificación por entropía, tal como una operación de descodificación de CAVLC, en la cabecera de fragmento codificada para recuperar la cabecera de fragmento.

**[0126]** Después de extraer los datos de fragmentos de unidades NAL de fragmentos codificados, el módulo de descodificación por entropía 150 puede extraer bloques de árbol codificados a partir de los datos de fragmento. El módulo de descodificación por entropía 150 puede entonces extraer CU codificadas de los bloques de árbol codificados. El módulo de descodificación por entropía 150 puede realizar operaciones de análisis que extraen elementos sintácticos de las CU codificadas. Los elementos sintácticos extraídos pueden incluir bloques de coeficientes codificados por entropía. El módulo de descodificación por entropía 150 puede entonces realizar operaciones de descodificación por entropía en los elementos sintácticos. Por ejemplo, el módulo de descodificación por entropía 150 puede realizar operaciones CABAC en los bloques de coeficientes.

**[0127]** Cuando el módulo de descodificación por entropía 150 realiza una operación de descodificación por entropía en un conjunto de datos, el módulo de descodificación por entropía 150 puede seleccionar un modelo de contexto. En los ejemplos en los que el módulo de descodificación por entropía 150 utiliza CABAC, el modelo de contexto puede indicar probabilidades de bins particulares. En los ejemplos en los que el módulo de descodificación por entropía 150 utiliza CAVLC, el modelo de contexto puede indicar una asignación entre las palabras de código y los datos correspondientes. El módulo de descodificación por entropía 150 puede utilizar a continuación el modelo de contexto seleccionado para realizar la operación de descodificación por entropía en el conjunto de datos.

Después de que el módulo de descodificación por entropía 150 realice una operación de análisis en una CU no dividida, el descodificador de vídeo 30 puede realizar una operación de descodificación en la CU no dividida. Para realizar la operación de descodificación en una CU no dividida, el descodificador de vídeo 30 puede, en cada nivel del árbol cuaternario residual de la CU, realizar una operación de descodificación en cada TU de la CU. Realizando la operación de descodificación para cada TU de la CU, el descodificador de vídeo 30 puede reconstruir los datos residuales de la CU.

**[0128]** Como parte de la realización de una operación de descodificación en una TU no dividida, el módulo de cuantificación inversa 154 puede cuantificar de forma inversa, es decir, descuantificar, el bloque de coeficientes asociado con la TU. El módulo de cuantificación inversa 154 puede cuantificar de forma inversa los bloques de coeficientes de una manera similar a los procesos de cuantificación inversa propuestos para HEVC o definidos por la norma de descodificación H.264. El módulo de cuantificación inversa 154 puede utilizar un parámetro de cuantificación QPY calculado por el codificador de vídeo 20 para una CU del bloque de coeficientes para determinar un grado de cuantificación y, asimismo, un grado de cuantificación inversa para el módulo de cuantificación inversa 154 a aplicar.

**[0129]** Después de que el módulo de cuantificación inversa 154 cuantifique de forma inversa un bloque de coeficientes, el módulo de transformada inversa 156 puede generar datos residuales para la TU asociada con el bloque de coeficientes. El módulo de transformada inversa 156 puede generar los datos residuales para la TU, al menos en parte, aplicando una transformada inversa al bloque de coeficiente. Por ejemplo, el módulo 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. En algunos ejemplos, el módulo de transformada inversa 156 puede determinar una transformada inversa para aplicar al bloque de coeficientes basándose en la señalización del codificador de vídeo 20. En tales ejemplos, el módulo de transformada inversa 156 puede determinar la transformada inversa basada en una transformada señalada en el nodo raíz de un árbol cuaternario para un bloque de árbol asociado con el bloque de coeficientes. En otros ejemplos, el módulo 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, el módulo de transformada inversa 156 puede aplicar una transformada inversa en cascada.

**[0130]** Si una CU se codificó usando predicción inter, el módulo de compensación de movimiento 162 puede llevar a cabo la compensación de movimiento para generar datos de predicción para la CU. El módulo de compensación de movimiento 162 puede utilizar información de movimiento para que la PU de la CU identifique muestras de referencia para las PU. La información de movimiento para una PU puede incluir un vector de movimiento, un índice de imagen de referencia y una dirección de predicción. El módulo de compensación de movimiento 162 puede a continuación utilizar las muestras de referencia para que las PU generen los datos de predicción para la CU.

**[0131]** En algunos ejemplos, el módulo de compensación de movimiento 162 puede refinar los datos de predicción para una CU realizando interpolación basándose en filtros de interpolación. Los identificadores para filtros de interpolación que se van a usar para la compensación de movimiento con una precisión de subpíxel pueden incluirse en los elementos sintácticos. El módulo de compensación de movimiento 162 puede usar los mismos filtros de interpolación usados por el codificador de vídeo 20 durante la generación de los datos de predicción de la CU para calcular valores interpolados para píxeles de sub-enteros de un bloque de referencia. El módulo 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 generar datos de predicción.

**[0132]** Si se codifica una CU usando predicción intra, el módulo de predicción intra 164 puede realizar predicción intra para generar datos de predicción para la CU. Por ejemplo, el módulo de predicción intra 164 puede determinar un modo de predicción intra para la CU basándose en elementos sintácticos en el flujo de bits. El módulo de predicción intra 164 puede entonces utilizar el modo de predicción intra para generar datos de predicción (por ejemplo, valores de píxel predichos) para la CU basándose en los valores de píxel de las CU contiguas.

**[0133]** El módulo de reconstrucción 158 puede usar los datos residuales de una CU y los datos de predicción para que la CU reconstruya los valores de píxel para la CU. En algunos ejemplos, el descodificador de vídeo 30 puede aplicar un filtro de desbloqueo para eliminar las distorsiones de bloqueo del filtro de valores de píxel reconstruidos de un fragmento o imagen. La memoria intermedia de imágenes descodificadas 160 puede almacenar los valores de píxel descodificados para imágenes de los datos de vídeo. 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.

**[0134]** Como se describió anteriormente, el descodificador de vídeo 30 puede implementarse o configurarse para realizar el primer aspecto de las técnicas descritas anteriormente. El módulo de descodificación por entropía 150 del descodificador de vídeo 30 puede recibir el flujo de bits representativo de los datos de vídeo codificados que se han codificado de acuerdo con el primer aspecto de las técnicas descritas en esta divulgación. El módulo de descodificación por entropía 150 puede determinar primero, para una parte de una imagen actual de los datos de vídeo codificados, el número de bits utilizados para especificar uno o más LSB del valor POC que identifica una LTRP para usar en la descodificación de la imagen actual. En algunos ejemplos, el número de bits puede señalarse explícitamente en el flujo de bits como un elemento sintáctico, por ejemplo, en la cabecera de fragmento, como se describe anteriormente. En otros ejemplos, el módulo de descodificación por entropía 150 puede obtener el número de bits en función de otros elementos sintácticos especificados en uno o más de los SPS, un conjunto de parámetros de imagen (PPS) y la cabecera de fragmento.

**[0135]** El módulo de descodificación por entropía 150 puede luego analizar el número determinado de bits del flujo de bits, donde el número determinado de bits analizados representa el uno o más LSB del POC que identifican la LTRP. A menudo, el valor de recuento del orden de imágenes que identifica la imagen de referencia a largo plazo es un valor de recuento del orden de imágenes que identifica una imagen de referencia a largo plazo que está presente en la memoria intermedia de imágenes descodificadas 160 pero que no se especifica como una o más imágenes de referencia a largo plazo en un conjunto de parámetros de secuencia asociados con la imagen actual. El número determinado de bits puede especificar lo que se ha denominado anteriormente como los LSB de POC. El módulo de descodificación por entropía 150 puede pasar el valor LSB de POC al módulo de compensación de movimiento 162. El módulo de compensación de movimiento 162 luego puede recuperar, de la memoria intermedia de imágenes descodificadas 160, la LTRP que se usará cuando se descodifique al menos la parte de la imagen actual basándose en los LSB de POC.

**[0136]** Para recuperar esta LTRP de la memoria intermedia de imágenes descodificadas, el módulo de compensación de movimiento 162 puede comparar los bits menos significativos del valor POC que identifica la LTRP con un número igual de bits menos significativos de un valor POC asociado con la al menos una imagen almacenada en la memoria intermedia de imágenes descodificadas 160. El módulo de compensación de movimiento 162 recupera una de las al menos una imagen almacenada en la memoria intermedia de imágenes descodificadas 160 para usar como imagen de referencia a largo plazo en la descodificación de la imagen actual que tiene los mismos LSB para el valor POC asociado con el uno de al menos una imagen almacenada en la memoria intermedia de imágenes descodificadas 160 como los bits menos significativos determinados de la LTRP que se utilizarán para descodificar la imagen actual. En otras palabras, el módulo de compensación de movimiento 162 puede hacer coincidir los LSB de POC con los bits de valores de POC menos significativos para las imágenes almacenadas en la memoria intermedia de imágenes descodificadas 160 e identificar la imagen con los LSB de POC correspondientes como la imagen de referencia a largo plazo presente en la memoria intermedia de imágenes descodificadas 160 pero no especificada en el SPS asociado con la imagen actual.

**[0137]** El módulo de compensación de movimiento 162 luego puede usar la imagen de referencia a largo plazo identificada por el valor de recuento del orden de imágenes obtenido para descodificar la parte de la imagen actual. En otras palabras, uno o más vectores de movimiento especificados para una unidad de codificación del fragmento actual de la imagen actual pueden hacer referencia a la LTRP identificada por los LSB de POC. El módulo de compensación de movimiento 162 puede recuperar varios bloques de la LTRP identificada por estos vectores de movimiento, utilizando estos diversos bloques como bloques de referencia al realizar la compensación de movimiento. El descodificador de vídeo 30 puede agregar datos residuales codificados para bloques del fragmento actual de la imagen actual a estos bloques de referencia para generar bloques descodificados de datos de vídeo y así reconstruir los datos de vídeo originales, como se describe anteriormente. El descodificador de vídeo 30 puede almacenar los bloques descodificados de datos de vídeo en la memoria intermedia de imágenes descodificadas 160 para su uso posterior en la descodificación de imágenes posteriores. El descodificador de vídeo

luego puede descodificar otros fragmentos, si los hay, de la imagen actual, de manera similar a como se describió anteriormente.

**[0138]** El descodificador de vídeo 30 también puede implementar los segundos aspectos de las técnicas descritas en esta divulgación, potencialmente en combinación con el primer aspecto de las técnicas descritas en esta divulgación. De acuerdo con los segundos aspectos de las técnicas, el módulo de descodificación por entropía 150 puede determinar, para una imagen actual de estos datos de vídeo codificados, los bits menos significativos de un valor de recuento del orden de imágenes que identifica una imagen de referencia a largo plazo que se utilizará al descodificar al menos una parte de la imagen actual. Estos uno o más bits menos significativos pueden no identificar de manera única el valor de recuento del orden de imágenes de la imagen de referencia a largo plazo con respecto al uno o más bits menos significativos de cualquier otro valor de recuento del orden de imágenes que identifique cualquier otra imagen presente en la memoria intermedia de imágenes descodificadas 160 en la que también se almacena la imagen de referencia a largo plazo. Como se señaló anteriormente, los bits menos significativos del valor de recuento del orden de imágenes pueden identificar de forma única el valor de recuento del orden de imágenes de la imagen de referencia a largo plazo con respecto a los bits menos significativos de cualquier otro valor de recuento del orden de imágenes determinado con respecto a un valor de recuento del orden de imágenes que identifica cualquier otra imagen de referencia a largo plazo que se utilizará como candidata para descodificar el fragmento de la imagen actual.

**[0139]** A este respecto, incluso cuando los LSB de POC delta son únicos para un subconjunto dado de las imágenes almacenadas en la memoria intermedia de imágenes descodificadas 160, el módulo de descodificación por entropía 150 puede determinar los bits más significativos del valor de recuento del orden de imágenes que identifica la imagen de referencia a largo plazo. Los bits más significativos en combinación con los bits menos significativos determinados son potencialmente suficientes para distinguir el valor de recuento del orden de imágenes que identifica la imagen de referencia a largo plazo de cualquier otro valor de recuento del orden de imágenes que identifica cualquier otra imagen en la memoria intermedia de imágenes descodificadas 160. El módulo de descodificación por entropía 150 puede pasar los bits menos significativos y los bits más significativos al módulo de compensación de movimiento 162. El módulo de compensación de movimiento 162 luego puede recuperar la imagen de referencia a largo plazo de la memoria intermedia de imágenes descodificadas 160 basándose en los bits menos significativos determinados del valor de recuento del orden de imágenes y los bits más significativos determinados del valor de recuento del orden de imágenes. El módulo de compensación de movimiento 162 (en combinación uno o más de los módulos 154, 156 y 158) puede usar la imagen de referencia a largo plazo identificada por el valor de recuento del orden de imágenes determinado para descodificar la parte de la imagen actual de una manera similar a la descrito anteriormente.

**[0140]** La FIG. 4 es un diagrama de flujo que ilustra la operación a modo de ejemplo de un codificador de vídeo al realizar el primer aspecto de las técnicas descritas en esta divulgación. Como ejemplo, el codificador de vídeo 20 mostrado en el ejemplo de la FIG. 2 puede codificar datos de vídeo de acuerdo con los primeros aspectos de las técnicas de esta divulgación. Para codificar los datos de vídeo de acuerdo con el primer aspecto de las técnicas, el módulo de estimación de movimiento 122 del codificador de vídeo 20 puede determinar, para una imagen actual de los datos de vídeo, una imagen de referencia a largo plazo que se usará cuando se realice una predicción inter para una imagen actual de los datos de vídeo de la manera descrita anteriormente (170). En algunos casos, esta imagen de referencia a largo plazo es una imagen de referencia a largo plazo que está presente en la memoria intermedia de imágenes descodificadas 114, pero que no se especifica como una o más imágenes de referencia a largo plazo en un conjunto de parámetros de secuencia asociado a la imagen actual.

**[0141]** El módulo de estimación de movimiento 122 puede identificar esta imagen de referencia a largo plazo seleccionada por el valor de recuento del orden de imágenes descrito anteriormente. El módulo de estimación de movimiento 122 puede acceder a un conjunto de parámetros de secuencia para determinar si el valor de recuento del orden de imágenes que identifica la imagen de referencia a largo plazo seleccionada está presente en el conjunto de parámetros de secuencia, que como se indicó anteriormente puede denotarse como "SPS" (172). En otras palabras, el módulo de estimación de movimiento 122 puede determinar si este SPS ya está señalando esta imagen de referencia a largo plazo al determinar si el SPS incluye, como un ejemplo, el valor de recuento del orden de imágenes que identifica esta imagen de referencia a largo plazo seleccionada. Cuando el SPS incluye el valor de recuento del orden de imágenes ("SÍ" 172), el módulo de estimación de movimiento 122 determina un índice de esta imagen de referencia a largo plazo identificada por el valor de recuento del orden de imágenes en, como un ejemplo, una lista de valores de recuento del orden de imágenes que se señalan en el SPS para su uso en la codificación de la imagen actual (y posiblemente otras imágenes) (174). El módulo de estimación de movimiento 122 puede a continuación pasar este índice al módulo de codificación por entropía 116. El módulo de codificación por entropía 116 puede entonces especificar este índice de la imagen de referencia a largo plazo en un flujo de bits representativo de una versión codificada de los datos de vídeo (176).

**[0142]** Sin embargo, si el valor de recuento del orden de imágenes que identifica la imagen de referencia a largo plazo seleccionada no se especifica en el SPS ("NO" 172), el módulo de estimación de movimiento 122 puede luego determinar un número de bits que se usarán para representar el uno o más LSB de un valor POC que identifica la imagen de referencia a largo plazo, como se describe anteriormente (178). El módulo de estimación

de movimiento 122 puede luego pasar estos uno o más LSB del valor POC al módulo de codificación por entropía 116. El módulo de codificación por entropía 116 puede especificar uno o más LSB del valor POC que identifica la LTRP en el flujo de bits utilizando el número determinado de bits utilizados para representar el uno o más bits menos significativos del valor de recuento del orden de imágenes que identifica la imagen de referencia a largo plazo (180).

**[0143]** El codificador de vídeo 20 puede usar la imagen de referencia a largo plazo identificada por el valor de recuento del orden de imágenes para codificar la parte de la imagen actual. En otras palabras, el módulo de compensación de movimiento 124 también puede configurarse para determinar un bloque predicho mediante la predicción inter utilizando la imagen de referencia a largo plazo de la manera descrita anteriormente (182). El módulo de compensación de movimiento 124 luego determina un bloque de vídeo residual basado en el bloque predicho (184). Es decir, el módulo de compensación de movimiento 124 luego determina un bloque de vídeo residual restando los valores de píxel del bloque predictivo de los valores de píxel del bloque de vídeo actual que se está codificando, formando datos residuales como valores de diferencia de píxel. El sumador 50 representa el componente o los componentes que realizan esta operación de resta.

**[0144]** El módulo de transformada 104 luego transforma el bloque residual para determinar coeficientes de transformada utilizando una transformada, tal como una transformada de coseno discreta (DCT) o una transformada conceptualmente similar (186). El módulo de transformada 104 puede convertir los datos de vídeo residuales de un dominio de píxel (espacial) a un dominio de transformada, tal como un dominio de frecuencia. El módulo 104 de transformada puede enviar los coeficientes de transformada resultantes a la unidad 54 de cuantificación. El módulo de cuantificación 106 cuantifica los coeficientes de transformada para reducir más la velocidad de bits (188). El proceso de cuantificación puede reducir la profundidad de bits asociada a algunos, o a la totalidad, de los coeficientes. Tras la cuantificación, el módulo de codificación por entropía 116 codifica por entropía los coeficientes de transformada cuantificados (190), insertando los coeficientes de transformada cuantificados por entropía en el flujo de bits, asociados con el índice y/o los bits menos significativos del valor de recuento del orden de imágenes (que se especifican comúnmente en una cabecera de fragmento asociada con el bloque codificado de datos de vídeo).

**[0145]** La FIG. 5 es un diagrama de flujo que ilustra el funcionamiento a modo de ejemplo de un descodificador de vídeo al realizar el primer aspecto de las técnicas descritas en esta divulgación. Como ejemplo, el descodificador de vídeo 30 implementa o está configurado para realizar el primer aspecto de las técnicas descritas anteriormente. El módulo de descodificación por entropía 150 puede descodificar por entropía un flujo de bits representativo de los datos de vídeo codificados para determinar los coeficientes de transformada cuantificados correspondientes a un bloque (que puede representar una parte) de los datos de vídeo codificados (200). El módulo de descodificación por entropía 150 puede luego pasar estos coeficientes de transformada cuantificados al módulo de cuantificación inversa 154. El módulo de cuantificación inversa 154 puede realizar una cuantificación inversa con respecto a los coeficientes de transformada cuantificados para determinar los coeficientes de transformada (202). El módulo de transformada inversa 156 puede luego realizar una transformada inversa con respecto a los coeficientes de transformada para generar datos residuales para la imagen actual (204).

**[0146]** El módulo de descodificación por entropía 150 también puede analizar el flujo de bits para determinar, para el bloque de la imagen actual, una imagen de referencia a largo plazo utilizada cuando se realiza una predicción para el bloque de la imagen actual (206). El módulo de descodificación por entropía 150 puede determinar si la imagen de referencia a largo plazo ya se ha señalado en el SPS asociado con la imagen actual (208). Si esta imagen de referencia a largo plazo ya se ha señalado en el SPS ("Sí" 208), el módulo de descodificación por entropía 150 puede determinar un índice de esta imagen de referencia a largo plazo como se especifica en una lista de imágenes de referencia a largo plazo en el SPS (210). El módulo de descodificación por entropía 150 puede entonces determinar el valor de recuento del orden de imágenes (POC) que identifica la imagen de referencia a largo plazo del SPS basándose en el índice (212).

**[0147]** Sin embargo, si la imagen de referencia a largo plazo no se especifica en el SPS ("NO" 208), el módulo de descodificación por entropía 150 determina el número de bits utilizados para señalar uno o más LSB del valor POC que identifica la imagen de referencia a largo plazo para usar en la descodificación de la imagen actual (214). En algunos ejemplos, el número de bits se puede señalar explícitamente en el flujo de bits como un elemento sintáctico, por ejemplo, en la cabecera de fragmento, como se describe anteriormente. En otros ejemplos, el módulo de descodificación por entropía 150 puede obtener el número de bits en función de otros elementos sintácticos especificados en uno o más de los SPS, un conjunto de parámetros de imagen (PPS) y la cabecera de fragmento.

**[0148]** El módulo de descodificación por entropía 150 puede luego analizar el número determinado de bits del flujo de bits, donde el número determinado de bits analizados representa el uno o más LSB del POC que identifica la LTRP como un número binario sin signo (216). A menudo, el valor de recuento del orden de imágenes que identifica la imagen de referencia a largo plazo es un valor de recuento del orden de imágenes que identifica una imagen de referencia a largo plazo que está presente en la memoria intermedia de imágenes descodificadas 160 pero que no se especifica como una o más imágenes de referencia a largo plazo en un conjunto de parámetros de



secuencia asociados con la imagen actual. Nuevamente, el número determinado de bits puede especificar lo que se ha denominado como los LSB de POC anteriores. El módulo de descodificación por entropía 150 puede pasar el valor LSB de POC al módulo de compensación de movimiento 162. El módulo de compensación de movimiento 162 luego puede recuperar, de la memoria intermedia de imágenes descodificadas 160, la LTRP que se usará cuando se descodifique al menos la parte de la imagen actual basada en los LSB de POC de la manera descrita anteriormente (218).

**[0149]** El módulo de compensación de movimiento 162 recupera una de las al menos una imagen almacenada en la memoria intermedia de imágenes descodificadas 160 para usar como imagen de referencia a largo plazo en la descodificación de la imagen actual que tiene los mismos LSB para el valor POC asociado con el uno de al menos una imagen almacenada en la memoria intermedia de imágenes descodificadas 160 como los bits menos significativos determinados de la LTRP que se utilizarán para descodificar la imagen actual. En otras palabras, el módulo de compensación de movimiento 162 puede hacer coincidir los LSB de POC con los bits de valores de POC menos significativos para las imágenes almacenadas en la memoria intermedia de imágenes descodificadas 160 e identificar la imagen con los LSB de POC correspondientes como la imagen de referencia a largo plazo presente en la memoria intermedia de imágenes descodificadas 160 pero no especificada en el SPS asociado con la imagen actual.

**[0150]** El módulo de compensación de movimiento 162 luego puede usar la imagen de referencia a largo plazo identificada por el valor de recuento del orden de imágenes obtenido para descodificar la parte de la imagen actual. En otras palabras, uno o más vectores de movimiento especificados para una unidad de codificación del fragmento actual de la imagen actual pueden hacer referencia a la LTRP identificada por los LSB de POC. El módulo de compensación de movimiento 162 puede recuperar varios bloques de la LTRP identificada por estos vectores de movimiento, utilizando estos diversos bloques como bloques de referencia al realizar la compensación de movimiento. El descodificador de vídeo 30 puede agregar datos residuales codificados para bloques del fragmento actual de la imagen actual a estos bloques de referencia para generar bloques descodificados de datos de vídeo y así reconstruir los datos de vídeo originales, como se describe anteriormente. En este sentido, el descodificador de vídeo 30 puede reconstruir el bloque de la imagen actual basándose en la imagen de referencia a largo plazo y los datos residuales (220). El descodificador de vídeo 30 puede almacenar los bloques descodificados de datos de vídeo en la memoria intermedia de imágenes descodificadas 160 para su uso posterior en la descodificación de imágenes posteriores. El descodificador de vídeo luego puede descodificar otros fragmentos, si los hay, de la imagen actual, de manera similar a como se describió anteriormente.

**[0151]** La FIG. 6 es un diagrama de flujo que ilustra el funcionamiento a modo de ejemplo de un codificador de vídeo al realizar el segundo aspecto de las técnicas descritas en esta divulgación. Como ejemplo, el codificador de vídeo 20 está configurado para implementar los segundos aspectos de las técnicas descritas en esta divulgación, potencialmente en combinación con el primer aspecto de las técnicas. De acuerdo con los segundos aspectos de las técnicas, el módulo de estimación de movimiento 122 del codificador de vídeo 20 puede determinar, para la imagen actual de los datos de vídeo, una imagen de referencia a largo plazo para usar cuando se realiza una predicción inter para la imagen actual (230). El módulo de estimación de movimiento 122 puede determinar uno o más bits menos significativos de un valor de recuento del orden de imágenes que identifica la imagen de referencia a largo plazo (232). El módulo de estimación de movimiento 122 puede entonces determinar si estos bits menos significativos son suficientes para identificar de manera inequívoca la imagen de referencia a largo plazo de cualquier otra imagen de referencia almacenada en la memoria intermedia de imágenes descodificadas 114 en la que también se almacena la imagen de referencia a largo plazo (o se almacenará dependiendo del orden en que se almacena la imagen de referencia a largo plazo en comparación con cuando se realiza esta determinación) de la manera descrita anteriormente.

**[0152]** El módulo de estimación de movimiento 122 puede determinar si estos bits menos significativos son suficientes para identificar de manera inequívoca la imagen de referencia a largo plazo de cualquier otra imagen de referencia almacenada en la memoria intermedia de imágenes descodificadas 114, al menos en parte, determinando los bits menos significativos para al menos una valor de recuento del orden de imágenes que identifica una respectiva de las otras imágenes almacenadas en la memoria intermedia de imágenes descodificadas 114. El módulo de estimación de movimiento 122 puede entonces determinar si los bits menos significativos para este valor de recuento del orden de imágenes son los mismos que los bits menos significativos determinados del valor de recuento del orden de imágenes que identifica la imagen de referencia a largo plazo. Si estos dos bits menos significativos del valor de recuento del orden de imágenes son iguales, el módulo de estimación de movimiento 122 puede entonces determinar que los bits menos significativos del valor de recuento del orden de imágenes que identifican la imagen de referencia a largo plazo que se utilizará al codificar y descodificar al menos el parte de la imagen actual no es suficiente para identificar de manera única la imagen de referencia a largo plazo apropiada. Sin embargo, si estos dos bits menos significativos del valor de recuento del orden de imágenes no son iguales, el módulo de estimación de movimiento 122 puede continuar determinando los bits menos significativos de los valores de recuento del orden de imágenes que identifican otra imagen de referencia a largo plazo almacenada en la memoria intermedia de imágenes descodificadas 114, comparando los dos bits menos significativos del recuento del orden de imágenes hasta que se encuentra una coincidencia o todos

los valores de recuento del orden de imágenes que identifican las imágenes almacenadas en la memoria intermedia de imágenes descodificadas 114 han pasado por esta comparación.

**[0153]** En otras palabras, la unidad de estimación de movimiento 122 puede comparar los bits menos significativos determinados con los bits menos significativos de cada uno de los valores de recuento del orden de imágenes que identifican cada una de las otras imágenes almacenadas en la memoria intermedia de imágenes descodificadas 114 (234). Cuando uno o más bits menos significativos de cualquier otro valor de recuento del orden de imágenes que identifica cualquier otra imagen almacenada en la memoria intermedia de imágenes descodificadas 114 en la que también se almacena la imagen de referencia a largo plazo son los mismos que el uno o más bits determinados menos significativos del valor de recuento del orden de imágenes que identifica la imagen de referencia a largo plazo ("SI" 236), el módulo de estimación de movimiento 122 puede determinar uno o más bits más significativos del valor de recuento del orden de imágenes determinado que identifica la imagen de referencia a largo plazo que se utilizará al codificar al menos la parte de la imagen actual de la manera descrita anteriormente (238).

**[0154]** El módulo de estimación de movimiento 122 puede a continuación reenviar estos bits menos significativos y los bits más significativos al módulo de codificación por entropía 116. El módulo de codificación por entropía 116 puede especificar los bits más significativos del valor de recuento del orden de imágenes que identifica la imagen de referencia a largo plazo que se utilizará cuando se codifique al menos la parte de la imagen actual en un flujo de bits representativo de una versión codificada de los datos de vídeo (240). El módulo de codificación por entropía 116 también puede especificar uno o más bits menos significativos determinados del valor de recuento del orden de imágenes que identifica la imagen de referencia a largo plazo que se utilizará cuando se codifique al menos la parte de la imagen actual en el flujo de bits (242).

**[0155]** Sin embargo, cuando uno o más bits menos significativos de cualquier otro valor de recuento del orden de imágenes que identifican cualquier otra imagen almacenada en la memoria intermedia de imágenes descodificadas 114 en la que también se almacena la imagen de referencia a largo plazo no son los mismos que el uno o más bits menos significativos determinados del valor de recuento del orden de imágenes que identifica la imagen de referencia a largo plazo ("NO" 236), la unidad de estimación de movimiento 122 puede reenviar solo los bits menos significativos determinados del valor de recuento del orden de imágenes que identifica la imagen de referencia a largo plazo sin determinar de otro modo los bits más significativos porque los bits menos significativos identifican de manera única la imagen de referencia a largo plazo con respecto a las imágenes almacenadas en la memoria intermedia de imágenes descodificadas 114. El módulo de codificación por entropía 122 puede luego especificar uno o más bits menos significativos determinados del valor de recuento del orden de imágenes que identifica la imagen de referencia a largo plazo que se utilizará cuando se codifique al menos la parte de la imagen actual en el flujo de bits sin especificar los bits más significativos de nuevo, porque los bits menos significativos identifican de manera única la imagen de referencia a largo plazo con respecto a las imágenes almacenadas en la memoria intermedia de imágenes descodificadas 114 (242).

**[0156]** El codificador de vídeo 20 puede usar la imagen de referencia a largo plazo identificada por el valor de recuento del orden de imágenes para codificar la parte de la imagen actual. En otras palabras, el módulo de compensación de movimiento 124 también puede configurarse para determinar un bloque predicho mediante la predicción inter utilizando la imagen de referencia a largo plazo de la manera descrita anteriormente (244). El módulo de compensación de movimiento 124 luego determina un bloque de vídeo residual basado en el bloque predicho (246). Es decir, el módulo de compensación de movimiento 124 luego determina un bloque de vídeo residual restando los valores de píxel del bloque predictivo de los valores de píxel del bloque de vídeo actual que se está codificando, formando datos residuales como valores de diferencia de píxel. El sumador 50 representa el componente o los componentes que realizan esta operación de resta.

**[0157]** El módulo de transformada 104 luego transforma el bloque residual para determinar coeficientes de transformada utilizando una transformada, tal como una transformada de coseno discreta (DCT) o una transformada conceptualmente similar (248). El módulo de transformada 104 puede convertir los datos de vídeo residuales de un dominio de píxel (espacial) a un dominio de transformada, tal como un dominio de frecuencia. El módulo 104 de transformada puede enviar los coeficientes de transformada resultantes a la unidad 54 de cuantificación. El módulo de cuantificación 106 cuantifica los coeficientes de transformada para reducir más la velocidad de bits (250). El proceso de cuantificación puede reducir la profundidad de bits asociada a algunos, o a la totalidad, de los coeficientes. Tras la cuantificación, el módulo de codificación por entropía 116 codifica por entropía los coeficientes de transformada cuantificados (252), insertando los coeficientes de transformada cuantificados por entropía en el flujo de bits, asociados con el índice y/o los bits menos significativos del valor de recuento del orden de imágenes (que se especifican comúnmente en una cabecera de fragmento asociada con el bloque codificado de datos de vídeo).

**[0158]** La FIG. 7 es un diagrama de flujo que ilustra el funcionamiento a modo de ejemplo de un descodificador de vídeo al realizar el segundo aspecto de las técnicas descritas en esta divulgación. Como ejemplo, el descodificador de vídeo 30 puede implementar los segundos aspectos de las técnicas descritas en esta divulgación, potencialmente en combinación con el primer aspecto de las técnicas descritas en esta divulgación.

El módulo de descodificación por entropía 150 puede descodificar por entropía un flujo de bits representativo de los datos de vídeo codificados para determinar los coeficientes de transformada cuantificados correspondientes a un bloque (que puede representar una parte) de los datos de vídeo codificados (260). El módulo de descodificación por entropía 150 puede luego pasar estos coeficientes de transformada cuantificados al módulo de cuantificación inversa 154. El módulo de cuantificación inversa 154 puede realizar una cuantificación inversa con respecto a los coeficientes de transformada cuantificados para determinar los coeficientes de transformada (262). El módulo de transformada inversa 156 puede entonces realizar una transformada inversa con respecto a los coeficientes de transformada para generar datos residuales para la imagen actual (264).

**[0159]** El módulo de descodificación por entropía 150 también puede analizar el flujo de bits para determinar, para el bloque de la imagen actual, una imagen de referencia a largo plazo utilizada cuando se realiza una predicción inter para el bloque de la imagen actual. De acuerdo con los segundos aspectos de las técnicas, el módulo de descodificación por entropía 150 puede descodificar por entropía los bits menos significativos del flujo de bits de un recuento del orden de imágenes que identifica una imagen de referencia a largo plazo que se utilizará cuando se descodifique, es decir, mediante predicción inter, al menos una parte de una imagen actual (266). El módulo de descodificación por entropía 150 puede determinar si los bits más significativos (MSB) para el valor de recuento del orden de imágenes que identifica la imagen de referencia a largo plazo están presentes en el flujo de bits (a menudo, analizando un elemento sintáctico en forma de un indicador del flujo de bits que indica si los bits más significativos están presentes) (268).

**[0160]** Si estos bits más significativos están presentes ("Sí" 268), el módulo de descodificación por entropía 150 puede descodificar por entropía el flujo de bits para determinar los bits más significativos del valor de recuento del orden de imágenes que identifica la imagen de referencia a largo plazo que se utilizará cuando se realice la predicción inter para la imagen actual (270). El módulo de descodificación por entropía 150 puede pasar los bits menos significativos y los bits más significativos al módulo de compensación de movimiento 162. El módulo de compensación de movimiento 162 luego puede recuperar la imagen de referencia a largo plazo de la memoria intermedia de imágenes descodificadas 160 basándose en los bits menos significativos determinados del valor de recuento del orden de imágenes y los bits más significativos determinados del valor de recuento del orden de imágenes (272). Si estos bits más significativos no están presentes ("NO" 268), el módulo de descodificación por entropía 150 puede pasar los bits menos significativos al módulo de compensación de movimiento 162. El módulo de compensación de movimiento 162 luego puede recuperar la imagen de referencia a largo plazo de la memoria intermedia de imágenes descodificadas 160 basándose en los bits menos significativos determinados del valor de recuento del orden de imágenes (274).

**[0161]** Después de recuperar esta imagen de referencia a largo plazo, el módulo de compensación de movimiento 162 (en combinación con uno o más de los módulos 154, 156 y 158) puede luego usar la imagen de referencia a largo plazo identificada por el valor de recuento del orden de imágenes determinado para descodificar la parte de la imagen actual de una manera similar a la descrita anteriormente. En otras palabras, uno o más vectores de movimiento especificados para una unidad de codificación del fragmento actual de la imagen actual pueden hacer referencia a la LTRP identificada por los LSB de POC. El módulo de compensación de movimiento 162 puede recuperar varios bloques de la LTRP identificada por estos vectores de movimiento, utilizando estos diversos bloques como bloques de referencia al realizar la compensación de movimiento. El descodificador de vídeo 30 puede agregar datos residuales codificados para bloques del fragmento actual de la imagen actual a estos bloques de referencia para generar bloques descodificados de datos de vídeo y así reconstruir los datos de vídeo originales, como se describe anteriormente. En este sentido, el descodificador de vídeo 30 puede reconstruir el bloque de la imagen actual basándose en la imagen de referencia a largo plazo y los datos residuales (276). El descodificador de vídeo 30 puede almacenar los bloques descodificados de datos de vídeo en la memoria intermedia de imágenes descodificadas 160 para su uso posterior en la descodificación de imágenes posteriores. El descodificador de vídeo luego puede descodificar otros fragmentos, si los hay, de la imagen actual, de manera similar a como se describió anteriormente.

**[0162]** Si bien se describió anteriormente con respecto a los valores de recuento del orden de imágenes completo, las técnicas también pueden implementarse con respecto a los valores de recuento del orden de imágenes delta calculados como la diferencia entre el valor de recuento del orden de imágenes que identifica la imagen actual y el valor de recuento del orden de imágenes que identifica la imagen de referencia a largo plazo que se usará al codificar (que es un término que puede usarse para referirse tanto a la codificación como a la descodificación) la imagen actual de los datos de vídeo. A este respecto, la referencia a un valor de recuento del orden de imágenes puede referirse tanto a valores de recuento del orden de imágenes completo como a valores de recuento del orden de imágenes delta. Por consiguiente, las técnicas no deberían estar limitadas en este sentido.

**[0163]** Además, se describen otras varias técnicas relacionadas con los aspectos primero y segundo de las técnicas y, de alguna manera, se describen a continuación. Como se indicó anteriormente, la última versión de HEVC WD6 utiliza un mecanismo basado en un conjunto de imágenes de referencia (RPS) para administrar las imágenes de referencia. Un RPS se refiere a un conjunto de imágenes de referencia asociadas con una imagen, que consta de todas las imágenes de referencia que son anteriores a la imagen asociada en orden de

descodificación que puede usarse para la predicción inter de la imagen asociada o cualquier imagen que siga a la imagen asociada en orden de descodificación. El orden de descodificación puede referirse al orden en que el proceso de descodificación procesa los elementos sintácticos.

**[0164]** En HEVC, un RPS para cada imagen codificada puede señalarse directamente. Los elementos sintácticos para la señalización de RPS se incluyen tanto en un conjunto de parámetros de secuencia (SPS) como en una cabecera de fragmento asociada con cada imagen codificada. Para una imagen codificada en particular, el RPS puede ser uno de esos conjuntos alternativos incluidos en el SPS, según lo indicado por un indicador en la cabecera de fragmento, o directamente señalado en la cabecera de fragmento.

**[0165]** El RPS para cada imagen consta de cinco listas diferentes de imágenes de referencia, que también pueden referirse a los cinco subconjuntos de RPS. El primer subconjunto de RPS puede referirse a RefPicSetStCurrBefore, que incluye todas las imágenes de referencia a corto plazo (STRP) que son anteriores a la imagen actual, tanto en el orden de descodificación como en el orden de salida, y que se pueden usar para predicción inter de la imagen actual. El segundo subconjunto de RPS puede referirse a RefPicSetStCurrAfter, que incluye todas las imágenes de referencia a corto plazo que son anteriores a la imagen actual en orden de descodificación, que suceden a la imagen actual en orden de salida, y que se pueden usar para predicción inter de la imagen actual. El tercer subconjunto de RPS puede referirse a RefPicSetStFoll, que incluye todas las imágenes de referencia a corto plazo que se pueden usar en la predicción inter de una o más de las imágenes que siguen a la imagen actual en orden de descodificación y que no se usan en la predicción intra de la imagen actual. El cuarto subconjunto de RPS puede referirse a RefPicSetLtCurr, que incluye todas las imágenes de referencia a largo plazo (LTRP) que pueden usarse en la predicción inter de la imagen actual. El quinto subconjunto de RPS puede referirse a RefPicSetLtFoll, que incluye todas las imágenes de referencia a largo plazo que se pueden usar en la predicción inter de una o más de las imágenes que siguen a la imagen actual en orden de descodificación y que no se usan en la predicción parcial de la imagen actual. El orden de salida puede referirse al orden en que las imágenes descodificadas se envían desde la memoria intermedia de imágenes descodificadas (DPB) en caso de que las imágenes descodificadas se envíen desde la DPB. El orden de salida de una imagen puede ser especificado por el valor de recuento del orden de imágenes (POC), a menudo sin importar si la imagen se va a emitir.

**[0166]** En cualquier caso, hay una serie de problemas potenciales con respecto a cómo HEVC WD6 indica que estos subconjuntos de RPS pueden obtenerse y señalarse. El primer problema se refiere a HEVC WD6 subcláusula 8.3.2, ecuación (Ecn.) 8-6, donde la obtención de los subconjuntos RPS de acuerdo con esta Ecn. 8-6 puede depender del estado de las imágenes de referencia en la DPB como resultado de la imagen anterior (debido a la redacción de "hay una imagen de referencia a largo plazo picX en la DPB" o similar). Sin embargo, esta redacción en la subcláusula 8.3.2 puede ser contraria a un principio fundamental del RPS de que el estado de las imágenes de referencia en la DPB esté "intra-codificado", lo cual puede significar que no depende del estado anterior. En otras palabras, determinar un RPS actual para una imagen actual puede depender del almacenamiento previo de una LTRP ("picX") en la DPB, a pesar de que un principio fundamental del RPS es que el estado de la imagen de referencia en la DPB está intracodificado y no intercodificado y, por lo tanto, debe utilizarse para la toma de decisiones de intercodificación temporal.

**[0167]** El segundo problema se refiere a la obtención de las LTRP en el RPS. Donde la Ecn. 8-6 en la subcláusula 8.3.2 de HEVC WD6, la LTRP se identifica por su bit POC menos significativo (LSB) (que se denota como "pic\_order\_cnt\_lsb"), si el delta\_poc\_msb\_present\_flag [i] es igual a 0, o su valor POC completo (indicado por "PicOrderCntVal") en caso contrario. Sin embargo, en este último caso, la parte en Ecn. 8-5 de HEVC WD6 para la obtención de un valor POC completo de una LTRP " $((\text{PicOrderCntVal} - \text{DeltaPocLt}[i] + \text{MaxPicOrderCntLsb}) \% \text{MaxPicOrderCntLsb}) - (\text{DeltaPocMSBCycleLt}[i]) * \text{MaxPicOrderCntLsb}$ " puede ser incorrecta.

**[0168]** Por ejemplo, si  $\text{MaxPicOrderCntLsb} = 256$ , y la imagen actual tiene un valor POC completo igual a  $256 * 10 + 100$  (LSB igual a 100), las imágenes con valores POC completos  $256 * 8 + 100$  y  $256 * 6 + 100$  dan como resultado dos LTRPs con el mismo valor LSB de POC de 100. Dado lo anterior,  $\text{DeltaPocMSBCycleLt}[i]$  para los dos LTRP sería igual a 2 y 4, respectivamente. De acuerdo con Ecn. 8-5, los valores POC completos de los dos LTRP serían  $-256 * 2$  y  $-256 * 4$ , respectivamente, lo cual es incorrecto.

**[0169]** Un tercer problema asociado con la obtención de RPS en HEVC WD6 puede ser que una vez que una imagen se señala como una LTRP, nunca se puede señalar como una imagen de referencia a corto plazo (STRP).

**[0170]** Un cuarto problema asociado con la obtención de RPS en HEVC WD6 puede ser que el LSB de POC de una LTRP que se incluirá en el RPS de una imagen se señala en la cabecera de fragmento como un delta del LSB de los valores POC delta, codificados usando ue(v). Esto puede ser ineficiente, ya que el delta del LSB de los valores POC delta son típicamente grandes para las LTRP, por lo que requieren muchos bits para representarse. Esto tampoco es sencillo para los codificadores para determinar qué valor señalar en la cabecera de fragmento.

**[0171]** Un quinto problema asociado con la obtención de RPS en HEVC WD6 puede ser que, cuando hay dos o más LTRP que tienen el mismo valor de LSB de POC, los valores de POC completos de todos estos LTRP deben

señalarse, en el que los valores de POC MSB se señalan como un delta de dos LTRP continuos que tienen los mismos LSB de POC, usando  $ue(v)$ . Para el primero de un conjunto de LTRP señalados continuamente que tienen los mismos LSB de POC, el POC MSB se señala directamente, utilizando  $ue(v)$ . Puede haber dos aspectos de ineficiencia aquí. El primero es que el delta de los valores de POC MSB y el valor POC MSB directamente señalado para el primer LTRP en un conjunto pueden ser grandes y, por lo tanto, pueden requerir muchos bits para representar los valores. El segundo es que, en la mayoría de los casos, agregar un poco más de LSB (o, de manera equivalente, hacer que la longitud del MSB sea un poco más larga) puede permitir la identificación única de todos las LTRP. Por consiguiente, puede que no sea necesario señalar todos los valores de MSB para las LTRP que tienen los mismos valores de LSB.

**[0172]** Para resolver potencialmente los problemas identificados anteriormente, a continuación se describen varias técnicas para la obtención del conjunto de imágenes de referencia (RPS) y la señalización de imágenes de referencia a largo plazo (LTRP) que se incluirán en el RPS de una imagen codificada, incluido el primer y el segundo aspectos mencionados anteriormente de las técnicas descritas anteriormente con respecto a las FIGs. 1-7. Nuevamente, aunque se describe a continuación con respecto a HEVC WD6, las técnicas pueden implementarse con respecto a cualquier norma de codificación de vídeo u otras técnicas de codificación que involucren la predicción inter con respecto a los datos de referencia. Además, el resaltado en gris se usa a continuación para indicar los cambios en HEVC WD6.

**[0173]** En un aspecto, se puede modificar la sintaxis y la semántica de la cabecera de fragmento como se especifica HEVC WD6, donde la sintaxis de la cabecera de fragmento se cambia de la siguiente manera: se agregan los elementos sintácticos `poc_lsb_len_delta[i]` y `poc_lsb_lt[i]` y se eliminan los elementos sintácticos `delta_poc_lsb_lt[i]`, `delta_poc_msb_present_flag[i]` y `delta_poc_msb_cycle_lt_minus[i]`, mientras que los otros elementos sintácticos no se modifican. Por lo tanto, la sintaxis de la cabecera de fragmento puede parecerse a la que se muestra a continuación en la Tabla 1, donde las partes en negrita y en cursiva que se muestran a continuación reflejan adiciones o cambios a HEVC en soporte de uno o más aspectos de las técnicas descritas en esta divulgación.

Tabla 1

<code>slice_header( ) {</code>	Descriptor
<code>...</code>	
<code>if(long_term_ref_pics_present_flag) {</code>	
<b><code>num_long_term_pics</code></b>	$ue(v)$
<code>for(i = 0; i &lt; num_long_term_pics; i++) {</code>	
<b><i><code>poc_lsb_len_delta[i]</code></i></b>	<b><i><math>ue(v)</math></i></b>
<b><i><code>poc_lsb_lt[i]</code></i></b>	<b><i><math>u(v)</math></i></b>
<b><i><code>used_by_curr_pic_lt_flag[i]</code></i></b>	$u(1)$
<code>}</code>	
<code>}</code>	
<code>...</code>	
<code>}</code>	

**[0174]** La semántica de la cabecera de fragmento también se puede cambiar de la siguiente manera: se agrega la semántica de los elementos sintácticos `poc_lsb_len_delta[i]` y `poc_lsb_lt[i]`, y se elimina la semántica de los elementos sintácticos `delta_poc_lsb_lt[i]`, `delta_poc_msb_present_flag[i]` y `delta_poc_msb_cycle_lt_minus[i]`, y la semántica para otros elementos sintácticos se mantiene sin cambios.

**[0175]** El elemento sintáctico **`poc_lsb_len_delta[i]`** puede especificar el número de bits para representar `poc_lsb_lt[i]`.

**[0176]** La variable `PocLsbLtLen[i]` se obtiene de la forma siguiente:

```

if(i == 0)
    PocLsbLtLen[i] =
log2_max_pic_order_cnt_lsb_minus4 + 4 + poc_lsb_len_delta[i]
else
    PocLsbLtLen[i] = PocLsbLtLen[i - 1] + poc_lsb_len_delta[i]

```

[0177] El elemento sintáctico **poc\_lsb\_lt[i]** puede especificar los bits menos significativos del valor de recuento del orden de imágenes de la i-ésima imagen de referencia a largo plazo que se incluye en el conjunto de imágenes de referencia a largo plazo de la imagen actual. Este elemento sintáctico **poc\_lsb\_lt[i]** puede estar en el rango de 0 a  $(1 < \text{PocLsbLtLen}[i]) - 1$ , inclusive. La longitud de **poc\_lsb\_lt[i]** puede ser **PocLsbLtLen[i]**. De forma alternativa, la variable **PocLsbLtLen[i]** se puede obtener de la siguiente manera:

$$\text{PocLsbLtLen}[i] = \log_2 \text{max\_pic\_order\_cnt\_lsb\_minus4} + 4 + \text{poc\_lsb\_len\_delta}[i]$$

[0178] El proceso de decodificación para los conjuntos de imágenes de referencia especificados en la sección 8.3.3 de HEVC WD6 se puede cambiar de la siguiente manera, donde el resaltado o sombreado en gris indica adiciones y/o cambios en la sección 8.3.2 de HEVC WD6 según lo especificado actualmente. Lo siguiente reproduce gran parte de la sección 8.3.2 de HEVC WD6.

Este proceso se invoca una vez por imagen, después de la decodificación de una cabecera de fragmento, pero antes de la decodificación de cualquier unidad de codificación y antes del proceso de decodificación para la construcción de la lista de imágenes de referencia del fragmento, como se especifica en la subcláusula 8.3.3 de HEVC WD6. El proceso puede dar como resultado marcar una o más imágenes de referencia como "no usada como referencia".

[0179] NOTA 1 - El conjunto de imágenes de referencia es una descripción absoluta de las imágenes de referencia utilizadas en el proceso de decodificación de las imágenes codificadas actuales y futuras. La señalización del conjunto de imágenes de referencia es explícita en el sentido de que todas las imágenes de referencia incluidas en el conjunto de imágenes de referencia se enumeran explícitamente y que no existe ningún proceso predeterminado de construcción del conjunto de imágenes de referencia en el decodificador que dependa del estado de la memoria intermedia de imágenes decodificadas.

Las imágenes de referencia a corto plazo pueden identificarse mediante sus valores **PicOrderCntVal**. **Las imágenes de referencia a largo plazo se identifican mediante los bits menos significativos de sus valores de PicOrderCntVal.**

[0180] Se elaboran cinco listas de valores de recuento del orden de imágenes **o de los bits menos significativos de recuento del orden de imágenes** para obtener el conjunto de imágenes de referencia: **PocStCurrBefore**, **PocStCurrAfter**, **PocStFoll**, **PocLtCurr** y **PocLtFoll**, con los números de elementos **NumPocStCurrBefore**, **NumPocStCurrAfter**, **NumPocStFoll**, **NumPocLtCurr** y **NumPocLtFoll** número de elementos, respectivamente.

- Si la imagen actual es una imagen IDR, **PocStCurrBefore**, **PocStCurrAfter**, **PocStFoll**, **PocLtCurr**, y **PocLtFoll** se establecen como vacías, y **NumPocStCurrBefore**, **NumPocStCurrAfter**, **NumPocStFoll**, **NumPocLtCurr**, y **NumPocLtFoll** se establecen a 0.

- De lo contrario, vale lo siguiente para la obtención de las cinco listas de valores de recuento del orden de imágenes y de los números de entradas.

```

for( i = 0, j = 0, k = 0; i < NumNegativePics[ StRpsIdx ]; i++ )
    if( UsedByCurrPicS0[ StRpsIdx ][ i ] )
        PocStCurrBefore[ j++ ] = PicOrderCntVal + DeltaPocS0[ StRpsIdx ][ i ]
    else
        PocStFoll[ k++ ] = PicOrderCntVal + DeltaPocS0[ StRpsIdx ][ i ]
NumPocStCurrBefore = j

for( i = 0, j = 0; i < NumPositivePics[ StRpsIdx ]; i++ )
    if( UsedByCurrPicS1[ StRpsIdx ][ i ] )
        PocStCurrAfter[ j++ ] = PicOrderCntVal + DeltaPocS1[ StRpsIdx ][ i ]
    else
        PocStFoll[ k++ ] = PicOrderCntVal + DeltaPocS1[ StRpsIdx ][ i ]
NumPocStCurrAfter = j
NumPocStFoll = k

```

(8-5)

```

for( i = 0, j = 0, k = 0; i < num_long_term_pics; i++ )
    if( used_by_curr_pic_lt_flag[ i ] ) {
        PocLtCurr[ j ] = poc_lsb_lt[ i ]

        PocLsbLenCurr[ j ] = PocLsbLtLen[ i ]
        j++
    }
    else {
        PocLtFoll[ k ] = poc_lsb_lt[ i ]
        ExtPocLsbLenFoll[ k ] = PocLsbLtLen[ i ]
        k++
    }
NumPocLtCurr = j
NumPocLtFoll = k

```

5 donde PicOrderCntVal es el recuento del orden de imágenes de la imagen actual como se especifica en la subcláusula 8.2.1 de HEVC WD6.

10 **[0181]** NOTA 2 - Un valor de StRpsIdx en el rango de 0 a num\_short\_term\_ref\_pic\_sets - 1, inclusive, indica que se está utilizando un conjunto de imágenes de referencia a corto plazo del conjunto de parámetros de secuencia activa, donde StRpsIdx es el índice del conjunto de imágenes de referencia a corto plazo para la lista de conjuntos de imágenes de referencia a corto plazo en el orden en que se señalan en el conjunto de parámetros de secuencia. StRpsIdx igual a num\_short\_term\_ref\_pic\_sets indica que se está utilizando un conjunto de imágenes de referencia a corto plazo explícitamente señalado en la cabecera de fragmento.

15 **[0182]** El conjunto de imágenes de referencia consta de cinco listas de imágenes de referencia: RefPicSetStCurrBefore, RefPicSetStCurrAfter, RefPicSetStFoll, RefPicSetLtCurr y RefPicSetLtFoll. La variable NumPocTotalCurr se fija igual a NumPocStCurrBefore + NumPocStCurrAfter + NumPocLtCurr. Al descodificar un fragmento P o B, se requiere, para la conformidad del flujo de bits, que el valor de NumPocTotalCurr no sea igual a 0.

20 **[0183]** NOTA 3 - RefPicSetStCurrBefore, RefPicSetStCurrAfter y RefPicSetLtCurr contienen todas las imágenes de referencia que pueden usarse en la predicción inter de la imagen actual y que pueden usarse en la predicción inter de una o más de las imágenes que siguen a la imagen actual en orden de descodificación. El RefPicSetStFoll y el RefPicSetLtFoll constan de todas las imágenes de referencia que no se utilizan en la predicción inter de la imagen actual, pero se pueden usar en la predicción inter de una o más de las imágenes siguientes a la imagen actual en orden de descodificación.

**[0184]** El marcado de una imagen de referencia puede ser "no usada para referencia", "usada para referencia a corto plazo" o "usada para referencia a largo plazo", pero solo uno de estos tres. Cuando se hace referencia a una imagen de referencia como marcada como "utilizada como referencia", esta se refiere colectivamente a la imagen marcada como "usada para referencia a corto plazo" o "utilizada para referencia a largo plazo" (pero no ambas). Una imagen de referencia que está marcada como "utilizada para referencia a corto plazo" se denomina imagen de referencia a corto plazo. Una imagen de referencia que está marcada como "utilizada para referencia a largo plazo" se denomina imagen de referencia a largo plazo.

**[0185]** El proceso de obtención para el conjunto de imágenes de referencia y el marcado de imágenes se llevan a cabo de acuerdo con los siguientes pasos ordenados, donde DPB se refiere a la memoria intermedia de imágenes descodificadas, como se describe en el Anexo C de HEVC WD6:

**1. Se aplica lo siguiente:**

*for( i = 0; i < NumPocStCurrBefore; i++ )*

If (hay una imagen de referencia picX en la DPB con PicOrderCntVal igual

a *PocStCurrBefore[ i ]* )

*RefPicSetStCurrBefore[ i ] = picX*

*else*

*RefPicSetStCurrBefore[ i ] = "no reference picture"*

*for( i = 0; i < NumPocStCurrAfter; i++ )*

If (hay una imagen de referencia picX en la DPB con PicOrderCntVal igual

a *PocStCurrAfter[ i ]* )

*RefPicSetStCurrAfter[ i ] = picX*

*else*

*RefPicSetStCurrAfter[ i ] = "no reference picture"* (8-6)

*for( i = 0; i < NumPocStFoll; i++ )*

If (hay una imagen de referencia picX en la DPB con PicOrderCntVal igual

a *PocStFoll[ i ]* )

*RefPicSetStFoll[ i ] = picX*

*else*

*RefPicSetStFoll[ i ] = "no reference picture"*

**2. Todas las imágenes de referencia incluidas en RefPicSetStCurrBefore, RefPicSetStCurrAfter y RefPicSetStFoll están marcadas como "utilizadas para referencia a corto plazo".**

**3. Se aplica lo siguiente:**



*for( i = 0; i < NumPocLtCurr; i++ ) {*

If (hay una imagen de referencia picX en la DPB, no incluida en

*RefPicSetStCurrBefore,*

*RefPicSetStCurrAfter* o *RefPicSetStFoll*, con *PicOrderCntVal*,

para la que

*Abs( PicOrderCntVal ) % ( 2<sup>PocLsbLenCurr[i]</sup> )* es igual a

*PocLtCurr[ i ] )*

*RefPicSetLtCurr[ i ] = picX*

*else*

*RefPicSetLtCurr[ i ] = "no reference picture"*

*}*

*for( i = 0; i < NumPocLtFoll; i++ ) {*

If (hay una imagen de referencia picX en la DPB, no incluida en

*RefPicSetStCurrBefore,*

*RefPicSetStCurrAfter* o *RefPicSetStFoll*, con *PicOrderCntVal*

para la que

*Abs( PicOrderCntVal ) % ( 2<sup>PocLsbLenFoll[i]</sup> )* es igual a

*PocLtFoll[ i ] )*

*RefPicSetLtFoll[ i ] = picX*

*else*

*RefPicSetLtFoll[ i ] = "no reference picture"*

*}*

**4. Todas las imágenes de referencia incluidas en el RefPicSetLtCurr y RefPicSetLtFoll están marcadas como "utilizadas para referencia a largo plazo".**

5

5. Todas las imágenes de referencia en la memoria intermedia de imágenes descodificadas que no están incluidas en RefPicSetLtCurr, RefPicSetLtFoll, RefPicSetStCurrBefore, RefPicSetStCurrAfter o RefPicSetStFoll están marcadas como "no utilizada como referencia".

10

**[0186]** NOTA 4 - Puede haber una o más imágenes de referencia que están incluidas en el conjunto de imágenes de referencia pero que no están presentes en la memoria intermedia de imágenes descodificadas. Cuando la primera imagen codificada en el flujo de bits es una imagen IDR o la imagen codificada actual no es una imagen principal de la primera imagen codificada en el flujo de bits, las entradas en RefPicSetStFoll o RefPicSetLtFoll que son iguales a "ninguna imagen de referencia" deben ignorarse. Cuando la primera imagen codificada en el flujo de bits no es una imagen de CRA o la imagen codificada actual no es una imagen principal de la primera imagen codificada en el flujo de bits, debe inferirse una pérdida de imagen involuntaria para cada entrada en RefPicSetStCurrBefore, RefPicSetStCurrAfter y RefPicSetLtCurr que es igual a "sin imagen de referencia".

15

20

**[0187]** De forma alternativa, en los pasos anteriores, las palabras "hay una imagen de referencia picX" se cambian a "hay una imagen picX".

**[0188]** Es un requisito de conformidad con el flujo de bits que el conjunto de imágenes de referencia esté restringido de la siguiente manera:

25

- **No habrá ninguna imagen de referencia con temporal\_ID mayor que el de la imagen actual incluida en RefPicSetStCurrBefore, RefPicSetStCurrAfter o RefPicSetLtCurr.**

- Para cada valor de  $i$  en el rango de 0 a  $\text{NumPocLtCurr} - 1$ , inclusive, no habrá más de una imagen de referencia en la DPB que no esté incluida en  $\text{RefPicSetStCurrBefore}$ ,  $\text{RefPicSetStCurrAfter}$  o  $\text{RefPicSetStFoll}$ , y que tenga un  $\text{PicOrderCntVal}$  para el que  $\text{Abs}(\text{PicOrderCntVal}) \% (2^{\text{PocLsbLenCurr}[i]})$  sea igual a  $\text{PocLtCurr}[i]$ . Para cada valor de  $i$  en el rango de 0 a  $\text{NumPocLtFoll} - 1$ , inclusive, no habrá más de una imagen de referencia en el DPB que no esté incluida en  $\text{RefPicSetStCurrBefore}$ ,  $\text{RefPicSetStCurrAfter}$  o  $\text{RefPicSetStFoll}$ , y que tenga  $\text{PicOrderCntVal}$  para el cual  $\text{Abs}(\text{PicOrderCntVal}) \% (2^{\text{PocLsbLenFoll}[i]})$  sea igual a  $\text{PocLtFoll}[i]$ .
- No se incluirá ninguna imagen de referencia en el conjunto de imágenes de referencia que preceda, en orden de salida, a cualquier imagen de CRA que preceda a la imagen actual, tanto en orden de descodificación como en orden de salida.
- Cuando la primera imagen codificada en el flujo de bits es una imagen IDR, o si la imagen codificada actual no es una imagen principal de la primera imagen codificada en el flujo de bits, no habrá ninguna entrada en  $\text{RefPicSetStCurrBefore}$ ,  $\text{RefPicSetStCurrAfter}$  o  $\text{RefPicSetLtCurr}$  que sea igual a "sin imagen de referencia".

**[0189]** NOTA 5 - Una imagen de referencia no se puede incluir en más de una de las cinco listas de conjuntos de imágenes de referencia.

**[0190]** En otro aspecto, las técnicas pueden modificar la sintaxis y la semántica de la cabecera de fragmento para agregar los elementos sintácticos  $\text{poc\_lsb\_lt}[i]$ ,  $\text{add\_poc\_lsb\_len}[i]$  y  $\text{add\_poc\_lsb}[i]$ , y eliminar los elementos sintácticos  $\text{delta\_poc\_lsb\_lt}[i]$ ,  $\text{delta\_poc\_msb\_present\_flag}[i]$  y  $\text{delta\_poc\_msb\_cycle\_lt\_minusl}[i]$ , dejando los otros elementos sintácticos sin cambios. La siguiente Tabla 2 ilustra estos cambios, donde las adiciones y/o cambios se denotan utilizando resaltado en gris o sombreado:

Tabla 2

<code>slice_header() {</code>	<b>Descript or</b>
<code>...</code>	
<code>if(long_term_ref_pics_present_flag) {</code>	
<b><code>num_long_term_pics</code></b>	<b><code>ue(v)</code></b>
<code>for(i = 0; i &lt; num_long_term_pics; i++) {</code>	
<b><code>poc_lsb_lt[i]</code></b>	<b><code>u(v)</code></b>
<b><code>add_poc_lsb_len[i]</code></b>	<b><code>ue(v)</code></b>
<b><code>if(add_poc_lsb_len[i] != 0)</code></b>	
<b><code>add_poc_lsb[i]</code></b>	<b><code>u(v)</code></b>
<b><code>used_by_curr_pic_lt_flag[i]</code></b>	<b><code>u(1)</code></b>
<code>}</code>	
<code>}</code>	
<code>...</code>	
<code>}</code>	

**[0191]** La semántica de cabecera de fragmento se puede cambiar para agregar semántica de elementos sintácticos  $\text{poc\_lsb\_lt}[i]$ ,  $\text{add\_poc\_lsb\_len}[i]$  y  $\text{add\_poc\_lsb}[i]$ , y eliminar la semántica de los elementos sintácticos  $\text{delta\_poc\_lsb\_lt}[i]$ ,  $\text{delta\_poc\_msb\_present\_flag}[i]$  y  $\text{delta\_poc\_msb\_cycle\_lt\_minusl}[i]$ , dejando la semántica sin cambios para otros elementos sintácticos.

**[0192]** El elemento sintáctico **`poc_lsb_lt[i]`** puede especificar los bits menos significativos del valor de recuento del orden de imágenes de la  $i$ -ésima imagen de referencia a largo plazo que se incluye en el conjunto de imágenes de referencia a largo plazo de la imagen actual. La longitud de  $\text{poc\_lsb\_lt}[i]$  puede ser  $\text{log2\_max\_pic\_order\_cnt\_lsb\_minus4} + 4$  bits.

**[0193]** El elemento sintáctico **`add_poc_lsb_len[i]`** puede especificar el número de bits menos significativos adicionales del valor de recuento del orden de imágenes de la  $i$ -ésima imagen de referencia a largo plazo que se incluye en el conjunto de imágenes de referencia a largo plazo de la imagen actual.  $\text{add\_poc\_lsb\_len}[i]$  estará en el rango de 0 a  $32 - \text{log2\_max\_pic\_order\_cnt\_lsb\_minus4} - 4$ , inclusive.

[0194] El elemento sintáctico **add\_poc\_lsb[ i]** puede especificar el valor de los bits adicionales menos significativos del valor de recuento del orden de imágenes de la *i*-ésima imagen de referencia a largo plazo que se incluye en el conjunto de imágenes de referencia a largo plazo de la imagen actual. **poc\_lsb\_lt[ i]** estará en el rango de 1 a  $(1 \ll \text{add\_poc\_lsb\_len}[i]) - 1$ , inclusive. La longitud de **add\_poc\_lsb[i]** es **add\_poc\_lsb\_len[i]** bits.

[0195] El proceso de decodificación para el conjunto de imágenes de referencia especificado en la sección 8.3.3 de HEVC WD6 se puede cambiar de la siguiente manera, donde nuevamente el sombreado o resaltado en gris indica adiciones y/o cambios en la sección 8.3.2 de HEVC WD6 según lo especificado actualmente. Lo siguiente reproduce gran parte de la sección 8.3.2 de HEVC WD6.

[0196] Este proceso se invoca una vez por imagen, después de la decodificación de una cabecera de fragmento, pero antes de la decodificación de cualquier unidad de codificación y antes del proceso de decodificación para la construcción de la lista de imágenes de referencia del fragmento, como se especifica en la subcláusula 8.3.3. El proceso puede dar como resultado marcar una o más imágenes de referencia como "no usada como referencia".

[0197] NOTA 1 - El conjunto de imágenes de referencia es una descripción absoluta de las imágenes de referencia utilizadas en el proceso de decodificación de las imágenes codificadas actuales y futuras. La señalización del conjunto de imágenes de referencia es explícita en el sentido de que todas las imágenes de referencia incluidas en el conjunto de imágenes de referencia se enumeran explícitamente y que no existe ningún proceso predeterminado de construcción del conjunto de imágenes de referencia en el decodificador que dependa del estado de la memoria intermedia de imágenes decodificadas.

[0198] Las imágenes de referencia a corto plazo pueden identificarse mediante sus valores **PicOrderCntVal**. **Las imágenes de referencia a largo plazo se identifican mediante los bits menos significativos de sus valores de PicOrderCntVal.**

[0199] Se elaboran cinco listas de valores de recuento del orden de imágenes **o de los bits menos significativos de recuento del orden de imágenes** para obtener el conjunto de imágenes de referencia: **PocStCurrBefore**, **PocStCurrAfter**, **PocStFoll**, **PocLtCurr** y **PocLtFoll**, con los números de elementos **NumPocStCurrBefore**, **NumPocStCurrAfter**, **NumPocStFoll**, **NumPocLtCurr** y **NumPocLtFoll** número de elementos, respectivamente.

- Si la imagen actual es una imagen IDR, **PocStCurrBefore**, **PocStCurrAfter**, **PocStFoll**, **PocLtCurr**, y **PocLtFoll** se establecen como vacías, y **NumPocStCurrBefore**, **NumPocStCurrAfter**, **NumPocStFoll**, **NumPocLtCurr**, y **NumPocLtFoll** se establecen a 0.

- De lo contrario, vale lo siguiente para la obtención de las cinco listas de valores de recuento del orden de imágenes y de los números de entradas.

```

for( i = 0, j = 0, k = 0; i < NumNegativePics[ StRpsIdx ]; i++ )
    if( UsedByCurrPicS0[ StRpsIdx ][ i ] )
        PocStCurrBefore[ j++ ] = PicOrderCntVal + DeltaPocS0[ StRpsIdx ][ i ]
    else
        PocStFoll[ k++ ] = PicOrderCntVal + DeltaPocS0[ StRpsIdx ][ i ]
NumPocStCurrBefore = j

for( i = 0, j = 0; i < NumPositivePics[ StRpsIdx ]; i++ )
    if( UsedByCurrPicS1[ StRpsIdx ][ i ] )
        PocStCurrAfter[ j++ ] = PicOrderCntVal + DeltaPocS1[ StRpsIdx ][ i ]
    else
        PocStFoll[ k++ ] = PicOrderCntVal + DeltaPocS1[ StRpsIdx ][ i ]
NumPocStCurrAfter = j
NumPocStFoll = k
(8-5)

for( i = 0, j = 0, k = 0; i < num_long_term_pics; i++ )
    if( used_by_curr_pic_lt_flag[ i ] ) {
        PocLtCurr[ j ] = add_poc_lsb[ i ] * MaxPicOrderCntLsb + poc_lsb_lt[ i ]
        ExtPocLsbLenCurr[ j ] =
            log2_max_pic_order_cnt_lsb_minus4 + 4 + add_poc_lsb_len[ i ]
        j++
    }
    else {
        PocLtFoll[ k ] = add_poc_lsb[ i ] * MaxPicOrderCntLsb + poc_lsb_lt[ i ]
        ExtPocLsbLenFoll[ k ] =
            log2_max_pic_order_cnt_lsb_minus4 + 4 + add_poc_lsb_len[ i ]
        k++
    }
NumPocLtCurr = j
NumPocLtFoll = k

```

donde PicOrderCntVal es el recuento del orden de imágenes de la imagen actual como se especifica en la subcláusula 8.2.1.

5

**[0200]** NOTA 2 - Un valor de StRpsIdx en el rango de 0 a num\_short\_term\_ref\_pic\_sets - 1, inclusive, indica que se está utilizando un conjunto de imágenes de referencia a corto plazo del conjunto de parámetros de secuencia activa, donde StRpsIdx es el índice del conjunto de imágenes de referencia a corto plazo para la lista de conjuntos de imágenes de referencia en el orden en que se señalan en el conjunto de parámetros de secuencia. StRpsIdx igual a num\_short\_term\_ref\_pic\_sets indica que se está utilizando un conjunto de imágenes de referencia a corto plazo explícitamente señalado en la cabecera de fragmento.

10

**[0201]** El conjunto de imágenes de referencia consta de cinco listas de imágenes de referencia: RefPicSetStCurrBefore, RefPicSetStCurrAfter, RefPicSetStFoll, RefPicSetLtCurr y RefPicSetLtFoll. La variable NumPocTotalCurr se fija igual a NumPocStCurrBefore + NumPocStCurrAfter + NumPocLtCurr. Al decodificar un fragmento P o B, se requiere, para la conformidad del flujo de bits, que el valor de NumPocTotalCurr no sea igual a 0.

15

**[0202]** NOTA 3 - RefPicSetStCurrBefore, RefPicSetStCurrAfter y RefPicSetLtCurr contienen todas las imágenes de referencia que pueden usarse en la predicción inter de la imagen actual y que pueden usarse en la predicción inter de una o más de las imágenes que siguen a la imagen actual en orden de descodificación. El RefPicSetStFoll y el RefPicSetLtFoll constan de todas las imágenes de referencia que no se utilizan en la predicción inter de la imagen actual, pero se pueden usar en la predicción inter de una o más de las imágenes siguientes a la imagen actual en orden de descodificación.

**[0203]** El marcado de una imagen de referencia puede ser "no usada para referencia", "usada para referencia a corto plazo" o "usada para referencia a largo plazo", pero solo uno de estos tres. Cuando se hace referencia a una imagen de referencia marcada como "utilizada como referencia", esto se refiere colectivamente a la imagen marcada como "usada para referencia a corto plazo" o "utilizada para referencia a largo plazo" (pero no ambas). Una imagen de referencia que está marcada como "utilizada para referencia a corto plazo" se denomina imagen de referencia a corto plazo. Una imagen de referencia que está marcada como "utilizada para referencia a largo plazo" se denomina imagen de referencia a largo plazo.

**[0204]** El proceso de obtención para el conjunto de imágenes de referencia y el marcado de imágenes se llevan a cabo de acuerdo con los siguientes pasos ordenados, donde DPB se refiere a la memoria intermedia de imágenes descodificadas, como se describe en el Anexo C de HEVC WD6:

**1. Se aplica lo siguiente:**

*for( i = 0; i < NumPocStCurrBefore; i++ )*

If (hay una imagen de referencia picX en la DPB con PicOrderCntVal igual

a *PocStCurrBefore[ i ]* )

*RefPicSetStCurrBefore[ i ] = picX*

*else*

*RefPicSetStCurrBefore[ i ] = "no reference picture"*

*for( i = 0; i < NumPocStCurrAfter; i++ )*

If (hay una imagen de referencia picX en la DPB con PicOrderCntVal igual

a *PocStCurrAfter[ i ]* )

*RefPicSetStCurrAfter[ i ] = picX*

*else*

*RefPicSetStCurrAfter[ i ] = "no reference picture"* (8-67)

*for( i = 0; i < NumPocStFoll; i++ )*

If (hay una imagen de referencia picX en la DPB con PicOrderCntVal igual

a *PocStFoll[ i ]* )

*RefPicSetStFoll[ i ] = picX*

*else*

*RefPicSetStFoll[ i ] = "no reference picture"*

**2. Todas las imágenes de referencia incluidas en RefPicSetStCurrBefore, RefPicSetStCurrAfter y RefPicSetStFoll están marcadas como "utilizadas para referencia a corto plazo".**

**3. Se aplica lo siguiente:**

*for( i = 0; i < NumPocLtCurr; i++ ) {*

If (hay una imagen de referencia picX en la DPB, no incluida en

*RefPicSetStCurrBefore,*

*RefPicSetStCurrAfter or RefPicSetStFoll, with PicOrderCntVal,*

para la que

*Abs( PicOrderCntVal ) % ( 2<sup>ExtPocLsbLenCurr[i]</sup> )* es igual a *PocLtCurr[i]*

*RefPicSetLtCurr[ i ] = picX*

*else*

*RefPicSetLtCurr[ i ] = "no reference picture"*

*}*

*for( i = 0; i < NumPocLtFoll; i++ ) {*

If (hay una imagen de referencia picX en la DPB, no incluida en

*RefPicSetStCurrBefore,*

*RefPicSetStCurrAfter or RefPicSetStFoll, with PicOrderCntVal*

para la que

*Abs( PicOrderCntVal ) % ( 2<sup>ExtPocLsbLenFoll[i]</sup> )* es igual a

*PocLtFoll[ i ] )*

*RefPicSetLtFoll[ i ] = picX*

*else*

*RefPicSetLtFoll[ i ] = "no reference picture"*

*}*

5 **4. Todas las imágenes de referencia incluidas en el RefPicSetLtCurr y el RefPicSetLtFoll están marcadas como "utilizada para referencia a largo plazo"**

5. Todas las imágenes de referencia en la memoria intermedia de imágenes descodificadas que no están incluidas en RefPicSetLtCurr, RefPicSetLtFoll, RefPicSetStCurrBefore, RefPicSetStCurrAfter o RefPicSetStFoll están marcadas como "no utilizada como referencia".

[0205] NOTA 4 - Puede haber una o más imágenes de referencia que están incluidas en el conjunto de imágenes de referencia pero que no están presentes en la memoria intermedia de imágenes descodificadas. Cuando la primera imagen codificada en el flujo de bits es una imagen IDR o la imagen codificada actual no es una imagen principal de la primera imagen codificada en el flujo de bits, las entradas en RefPicSetStFoll o RefPicSetLtFoll que son iguales a "ninguna imagen de referencia" deben ignorarse. Cuando la primera imagen codificada en el flujo de bits no es una imagen de CRA o la imagen codificada actual no es una imagen principal de la primera imagen codificada en el flujo de bits, debe inferirse una pérdida de imagen involuntaria para cada entrada en RefPicSetStCurrBefore, RefPicSetStCurrAfter y RefPicSetLtCurr que es igual a "sin imagen de referencia".

[0206] De forma alternativa, en los pasos anteriores, las palabras "hay una imagen de referencia picX" se cambian a "hay una imagen picX".

[0207] Es un requisito de conformidad con el flujo de bits que el conjunto de imágenes de referencia esté restringido de la siguiente manera:

- **No habrá ninguna imagen de referencia con temporal\_id mayor que el de la imagen actual incluida en RefPicSetStCurrBefore, RefPicSetStCurrAfter o RefPicSetLtCurr.**

- **Para cada valor de i en el rango de 0 a NumPocLtCurr - 1, inclusive, no habrá más de una imagen de referencia en la DPB que no esté incluida en RefPicSetStCurrBefore, RefPicSetStCurrAfter o RefPicSetStFoll, y que tenga PicOrderCntVal para el que  $Abs(PicOrderCntVal) \% (2^{ExtPocLsbLenCurr[i]})$  sea**

*igual a  $PocLtCurr[i]$ . Para cada valor de  $i$  en el rango de 0 a  $NumPocLtFoll-1$ , inclusive, no habrá más de una imagen de referencia en la DPB que no esté incluida en  $RefPicSetStCurrBefore$ ,  $RefPicSetStCurrAfter$  o  $RefPicSetStFoll$ , y que tenga  $PicOrderCntVal$  para el cual  $Abs(PicOrderCntVal) \% (2^{ExtPocLsbLenFoll[i]})$  sea igual a  $PocLtFoll[i]$ .*

- No se incluirá ninguna imagen de referencia en el conjunto de imágenes de referencia que preceda, en orden de salida, a cualquier imagen de CRA que preceda a la imagen actual, tanto en orden de decodificación como en orden de salida.
- Cuando la primera imagen codificada en el flujo de bits es una imagen IDR, o si la imagen codificada actual no es una imagen principal de la primera imagen codificada en el flujo de bits, no habrá ninguna entrada en  $RefPicSetStCurrBefore$ ,  $RefPicSetStCurrAfter$  o  $RefPicSetStFoll$  que sea igual a "sin imagen de referencia".

**[0208]** NOTA 5 - Una imagen de referencia no se puede incluir en más de una de las cinco listas de conjuntos de imágenes de referencia.

**[0209]** Un tercer aspecto de las técnicas puede basarse en el primer aspecto de las técnicas descritas anteriormente. En este tercer aspecto, en la sintaxis de la cabecera de fragmento, se puede agregar un elemento sintáctico más para cada LTRP que se incluirá en el RPS para señalar el bit de inicio del POC de la LTRP, con el valor 0 que indica, como ejemplo, comenzar en el bit menos significativo (bit 0), el valor 1 indica comenzar desde el segundo bit menos significativo (bit 1), y así sucesivamente.  $PocLsbLtLen[i]$  se puede obtener entonces de la misma manera que en el primer aspecto de las técnicas descritas anteriormente, donde  $PocLsbLtLen[i]$  se obtiene de  $poc\_lsb\_len\_delta[i]$ . Sin embargo, en este tercer aspecto de las técnicas,  $PocLsbLtLen[i]$  puede representar el número de bits del POC de la LTRP comenzando desde el bit de inicio como se identificó anteriormente mediante el elemento sintáctico adicional, donde  $poc\_lsb\_lt[i]$  especifica el valor de los  $PocLsbLtLen[i]$  bits del POC del LTRP que comienzan desde el bit de inicio como se identificó anteriormente mediante el elemento sintáctico adicional.

**[0210]** A continuación puede ser necesario que en la DPB haya solo una imagen de referencia (o, de forma alternativa, solo una imagen) que tenga el mismo conjunto de bits del valor POC igual al indicado para la LTRP, excluyendo los incluidos en  $RefPicSetStCurrBefore$ ,  $RefPicSetStCurrAfter$  o  $RefPicSetStFoll$ . Por lo tanto, la LTRP puede entonces identificarse de forma única mediante el conjunto de bits, y en la Ecn. 8-7 de HEVC WD6, esta condición (en lugar de la condición de que el conjunto de bits menos significativos del valor POC igual al indicado para la LTRP) se puede usar para identificar una imagen de referencia (o, de forma alternativa, una imagen) para ser incluida en  $RefPicSetLtCurr[i]$  o  $RefPicSetLtFoll[i]$ . Las otras partes del tercer aspecto de las técnicas pueden ser similares a las descritas anteriormente con respecto al primer aspecto de las técnicas.

**[0211]** Un cuarto aspecto de las técnicas descritas en esta divulgación puede representar una variación de los segundos aspectos de las técnicas descritas anteriormente. En este cuarto aspecto, la sintaxis de la cabecera de fragmento puede incluir  $add\_poc\_lsb\_len[i]$ , que representa un delta de la longitud del LSB adicional señalado por  $add\_poc\_lsb[i]$ . Si  $i$  es igual a 0,  $add\_poc\_lsb\_len[i]$  puede especificar la longitud del LSB adicional señalada por  $add\_poc\_lsb[i]$ ; de lo contrario,  $add\_poc\_lsb\_len[i]$  puede especificar la diferencia entre la longitud  $add\_poc\_lsb[i]$  y la longitud de  $add\_poc\_lsb[i - 1]$ . Las otras partes del cuarto aspecto de las técnicas pueden ser similares a las descritas anteriormente con respecto al segundo aspecto de las técnicas.

**[0212]** Un quinto aspecto de las técnicas puede basarse en el primer aspecto de las técnicas descritas anteriormente. En este quinto aspecto, para indicar una LTRP, en lugar de señalar el LSB de POC a partir del bit menos significativo, se puede señalar un LSB de POC truncado.

**[0213]** La sintaxis de la cabecera de fragmento se puede cambiar para agregar los elementos sintácticos  $poc\_lsb\_len\_delta[i]$  y  $poc\_lsb\_lt[i]$  y eliminar los elementos sintácticos  $delta\_poc\_lsb\_lt[i]$ ,  $delta\_poc\_msb\_present\_flag[i]$  y  $delta\_poc\_msb\_cycle\_lt\_minus[i]$ , dejando los otros elementos sintácticos sin cambiar. La siguiente **Tabla 3** ilustra cómo estos cambios pueden afectar a la cabecera de fragmento, con resaltado en gris o sombreado que denota cambios y/o adiciones.

Tabla 3

slice_header() {	Descript or
...	
if(long_term_ref_pics_present_flag) {	
<b>num_long_term_pics</b>	ue(v)
for(i = 0; i < num_long_term_pics; i++) {	
<b>poc_lsb_len_truncated_delta[i]</b>	<b>ue(v)</b>
<b>poc_lsb_lt_truncated[i]</b>	<b>u(v)</b>
<b>used_by_curr_pic_lt_flag[i]</b>	u(1)
}	
}	
...	
}	

**[0214]** **poc\_lsb\_len\_truncated\_delta[i]** se puede usar para especificar el número de bits que representan poc\_lsb\_truncated\_lt[i]. La variable PocLsbLtTrunLen[i] puede obtenerse de la forma siguiente:

```

if(i == 0)
    PocLsbLtTrunLen[i] =
        log2_max_pic_order_cnt_lsb_minus4 + 4 + poc_lsb_len_truncated_delta[i] -
        NumTruncatedBits
else
    PocLsbLtTrunLen[i] =
        PocLsbLtTrunLen[i - 1] + poc_lsb_len_truncated_delta[i]
    NumTruncatedBits puede establecerse igual a
        log2_max_pic_order_cnt_lsb_minus4 + 4, y en esta PocLsbLtTrunLen[0] es
        igual a poc_lsb_len_truncated_delta[0]. Este valor también se puede señalar en PPS o SPS.

```

**[0215]** **poc\_lsb\_truncated\_lt[i]** puede especificar los bits menos significativos, después de trancar los bits NumTruncatedBits, del valor de recuento del orden de imágenes de la i-ésima imagen de referencia a largo plazo que se incluye en el conjunto de imágenes de referencia a largo plazo de la imagen actual, en otros palabras, del bit NumTruncatedBits al bit NumTruncatedBits + PocLsbLtTrunLen[i] - 1, incluido el valor de recuento del orden de imágenes de la i-ésima imagen de referencia a largo plazo que se incluye en el conjunto de imágenes de referencia a largo plazo de la imagen actual. Estos bits se conocen como el conjunto de bits truncados para la i-ésima LTRP que se incluye en el conjunto de imágenes de referencia a largo plazo de la imagen actual. El elemento sintáctico poc\_lsb\_truncated\_lt[i] puede estar en el rango de 0 a (1 < PocLsbLtTrunLen[i]-1, inclusive. La longitud de poc\_lsb\_lt[i] puede, como ejemplo, ser PocLsbLtTrunLen[i].

**[0216]** Las imágenes de referencia a largo plazo se identifican típicamente mediante el conjunto de bits truncados de sus valores de PicOrderCntVal. No hay dos imágenes de referencia (o, de forma alternativa, no hay dos imágenes) en la DPB (excluyendo las incluidas en RefPicSetStCurrBefore, RefPicSetStCurrAfter o RefPicSetStFoll) que puedan tener el mismo valor para el mismo conjunto de bits para sus valores PicOrderCntVal. Por lo tanto, la LTRP puede ser identificada de forma única por el conjunto de bits truncados, y en la Ecn. 8-7 de HEVC WD6, esta condición (en lugar de la condición de que el conjunto de bits menos significativos del valor POC igual al indicado para la LTRP) se usa para identificar una imagen de referencia (o, de forma alternativa, una imagen) para ser incluida en RefPicSetLtCurr[i] o RefPicSetLtFoll[i].

**[0217]** De forma alternativa, los NumTruncatedBits pueden ser señalados en cada fragmento, cuando se señalan las imágenes LTRP. La Tabla 4 ilustra esta alternativa, con resaltado en gris o sombreado que denota cambios y/o adiciones:



Tabla 4

slice_header() {	Descript or
...	
if(long_term_ref_pics_present_flag) {	
<b>num_long_term_pics</b>	ue(v)
<b>num_truncated_bits</b>	u(v)
for(i = 0; i < num_long_term_pics; i++) {	
<b>poc_lsb_len_truncated_delta[i]</b>	ue(v)
<b>poc_lsb_lt_truncated[i]</b>	u(v)
<b>used_by_curr_pic_lt_flag[i]</b>	u(1)
}	
}	
...	
}	

En la Tabla 4, **num\_truncated\_bits** puede especificar los bits menos significativos de un valor POC que se truncarán al identificar una LTRP desde el LSB de POC. Este elemento sintáctico puede estar en el rango de 0 a  $\log_2_{\max\_pic\_order\_cnt\_lsb\_minus4+4}$ , inclusive, y la longitud puede ser  $\text{ceil}(\log_2(\log_2_{\max\_pic\_order\_cnt\_lsb\_minus4+4+1}))$ . Como otra alternativa más, num\_truncated\_bits puede señalarse en ue(v) y puede ser más grande que  $\log_2_{\max\_pic\_order\_cnt\_lsb\_minus4+4}$ .

**[0218]** Debe reconocerse que, dependiendo del ejemplo, ciertos actos o sucesos de cualquiera de las técnicas descritas en el presente documento pueden realizarse en una secuencia distinta, pueden añadirse, combinarse u omitirse por completo (por ejemplo, no todos los actos o sucesos descritos son necesarios para la puesta en práctica de las técnicas). Además, en ciertos ejemplos, los actos o sucesos pueden realizarse simultáneamente, por ejemplo, mediante el procesamiento de múltiples hilos, el procesamiento de interrupciones o múltiples procesadores, en lugar de secuencialmente.

**[0219]** En uno o más ejemplos, las funciones descritas pueden implementarse en hardware, software, firmware o cualquier combinación de los mismos. Si se implementan en software, las funciones pueden almacenarse en, y transmitirse por, un medio legible por ordenador, como una o más instrucciones o código, y ejecutarse mediante una unidad de procesamiento basada en hardware. Los medios legibles por ordenador pueden incluir medios de almacenamiento legibles por ordenador, que correspondan a un medio tangible tal como medios de almacenamiento de datos o medios de comunicación que incluyan cualquier medio que facilite la transferencia de un programa informático desde un lugar a otro, por ejemplo, de acuerdo con un protocolo de comunicación. De esta manera, los medios legibles por ordenador pueden corresponder en general a (1) medios de almacenamiento tangibles legibles por ordenador que sean no transitorios o (2) un medio de comunicación tal como una señal o una onda portadora. Los medios de almacenamiento de datos pueden ser medios disponibles cualesquiera a los que se pueda acceder desde uno o más ordenadores o uno o más procesadores para recuperar instrucciones, código y/o estructuras de datos para la implementación de las técnicas descritas en esta divulgación. Un producto de programa informático puede incluir un medio legible por ordenador.

**[0220]** A modo de ejemplo, y no de manera limitativa, tales medios de almacenamiento legibles por ordenador pueden comprender RAM, ROM, EEPROM, CD-ROM u otro almacenamiento de disco óptico, almacenamiento de disco magnético u otros dispositivos de almacenamiento magnético, memoria flash o cualquier otro medio que pueda usarse para almacenar código de programa deseado en forma de instrucciones o estructuras de datos y al que pueda accederse mediante un ordenador. Además, cualquier conexión recibe debidamente la denominación de medio legible por ordenador. Por ejemplo, si las instrucciones se transmiten desde una sede de la Red, un servidor u otro origen remoto usando un cable coaxial, un cable de fibra óptica, un par trenzado, una línea de abonado digital (DSL) o tecnologías inalámbricas tales como infrarrojos, radio y microondas, entonces el cable coaxial, el cable de fibra óptica, el par trenzado, la DSL o las tecnologías inalámbricas tales como infrarrojos, radio y microondas se incluyen en la definición de medio. Sin embargo, debería entenderse que los medios de almacenamiento legibles por ordenador y los medios de almacenamiento de datos no incluyen conexiones, ondas portadoras, señales u otros medios transitorios, sino que, en cambio, están orientados a medios de almacenamiento tangibles no transitorios. Los discos, como se usan en el presente documento, incluyen un disco compacto (CD), un disco láser, un disco óptico, un disco versátil digital (DVD), un disco flexible y un disco Blu-ray, donde algunos discos reproducen usualmente los datos magnéticamente, mientras que otros discos reproducen

los datos ópticamente con láseres. Las combinaciones de lo anterior también deberían incluirse dentro del alcance de los medios legibles por ordenador.

- 5 **[0221]** Las instrucciones pueden ser ejecutadas por uno o más procesadores, tales como uno o más procesadores de señales digitales (DSP), microprocesadores de propósito general, circuitos integrados específicos de la aplicación (ASIC), formaciones de puertas programables in situ (FPGA) u otros circuitos lógicos, integrados o discretos, equivalentes. En consecuencia, el término "procesador", como se usa en el presente documento, se puede referir a cualquiera de las estructuras anteriores o a cualquier otra estructura adecuada para la implementación de las técnicas descritas en el presente documento. Además, en algunos aspectos, la
- 10 funcionalidad descrita en el presente documento se puede proporcionar dentro de módulos de hardware y/o software dedicados, configurados para la codificación y la decodificación, o incorporados en un códec combinado. Además, las técnicas se podrían implementar totalmente en uno o más circuitos o elementos lógicos.
- 15 **[0222]** 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 (IC) o un conjunto de IC (por ejemplo, un conjunto de chips). Diversos componentes, módulos o unidades se describen en esta divulgación para enfatizar aspectos funcionales de dispositivos configurados para realizar las técnicas divulgadas, pero no requieren necesariamente su realización mediante diferentes unidades de hardware. En cambio, como se ha descrito anteriormente, diversas unidades se pueden combinar en una unidad de hardware de códec, o ser proporcionadas
- 20 por un grupo de unidades de hardware interoperativas, incluyendo uno o más procesadores, como se ha descrito anteriormente, en combinación con software y/o firmware adecuados.

## REIVINDICACIONES

1. Un procedimiento de codificación de datos de vídeo, el procedimiento que comprende:

determinar, para una imagen actual de los datos de vídeo, uno o más bits menos significativos de un valor de recuento del orden de imágenes que identifica una imagen de referencia a largo plazo que se utilizará al codificar al menos una parte de la imagen actual;

codificar al menos la parte de la imagen actual utilizando la imagen de referencia a largo plazo;

cuando uno o más bits menos significativos de cualquier otro valor de recuento del orden de imágenes que identifican cualquier otra imagen almacenada en la memoria intermedia de imágenes descodificadas en la que también se almacena la imagen de referencia a largo plazo no son los mismos que el uno o más bits menos significativos determinados del valor de recuento del orden de imágenes, especificar solo uno o más de los bits menos significativos determinados del valor de recuento del orden de imágenes en un flujo de bits representativo de una versión codificada de los datos de vídeo; y

cuando el uno o más bits menos significativos de cualquier otro valor de recuento del orden de imágenes que identifique cualquier otra imagen almacenada en la memoria intermedia de imágenes descodificadas en la que también se almacena la imagen de referencia a largo plazo son los mismos que el uno o más bits menos significativos determinados del valor de recuento del orden de imágenes, especificar el uno o más bits menos significativos determinados y uno o más bits más significativos del valor de recuento del orden de imágenes en el flujo de bits representativo de una versión codificada de los datos de vídeo,

**caracterizado por que** el número de uno o más bits más significativos que se especifica en el flujo de bits se selecciona basándose en un análisis de todas las imágenes almacenadas en la memoria intermedia de imágenes descodificadas de tal manera que el uno o más bits más significativos permiten desambiguar el valor de recuento del orden de imágenes que identifica la imagen de referencia a largo plazo en la memoria intermedia de imágenes descodificadas, y de tal manera que el procedimiento soporta que el número del uno o más bits menos significativos cuando se suma al número seleccionado de uno o más bits más significativos da como resultado un total que es menor que un número total de bits para especificar el valor de recuento del orden de imágenes.

2. El procedimiento según la reivindicación 1, en el que la imagen de referencia a largo plazo es una imagen de referencia a largo plazo que está presente en una memoria intermedia de imágenes descodificadas pero que no se especifica como una o más imágenes de referencia a largo plazo en un conjunto de parámetros de secuencia asociado con la imagen actual.

3. El procedimiento según la reivindicación 1, en el que especificar uno o más bits menos significativos determinados y uno o más bits más significativos determinados del valor de recuento del orden de imágenes comprende especificar uno o más bits menos significativos del valor de recuento del orden de imágenes que identifica la imagen de referencia a largo plazo en una cabecera de fragmento asociada con al menos la parte de la imagen actual.

4. Un dispositivo de codificación de vídeo para codificar datos de vídeo, el dispositivo de codificación de vídeo que comprende:

medios para determinar, para una imagen actual de los datos de vídeo, uno o más bits menos significativos de un valor de recuento del orden de imágenes que identifica una imagen de referencia a largo plazo que se utilizará cuando se codifique al menos una parte de la imagen actual;

medios para codificar al menos la parte de la imagen actual utilizando la imagen de referencia a largo plazo; y

medios para especificar bits del valor de recuento del orden de imágenes para identificar la imagen de referencia a largo plazo que se utilizará al codificar al menos una parte de la imagen actual, en el que:

cuando uno o más bits menos significativos de cualquier otro valor de recuento del orden de imágenes que identifican cualquier otra imagen almacenada en la memoria intermedia de imágenes descodificadas en la que también se almacena la imagen de referencia a largo plazo no son los mismos que el uno o más bits menos significativos determinados del valor de recuento del orden de imágenes, especificar solo uno o más de los bits más significativos determinados del valor de recuento del orden de imágenes en un flujo de bits representativo de una versión codificada de los datos de vídeo; y,

cuando uno o más bits menos significativos de cualquier otro valor de recuento del orden de imágenes que identifican cualquier otra imagen almacenada en la memoria intermedia de imágenes descodificadas en el que también se almacena la imagen de referencia a largo plazo son los mismos que el uno o más bits menos significativos determinados del valor de recuento del orden de imágenes, especificar el uno o más bits menos significativos determinados y uno o más bits más significativos del valor de recuento del orden de imágenes en un flujo de bits representativo de una versión codificada de los datos de vídeo,

**caracterizado por que** el número de uno o más bits más significativos que se especifica en el flujo de bits se selecciona basándose en un análisis de todas las imágenes almacenadas en la memoria intermedia de imágenes descodificadas de tal manera que el uno o más bits más significativos permiten desambiguar el valor de recuento del orden de imágenes que identifica la imagen de referencia a largo plazo en la memoria intermedia de imágenes descodificadas, y de tal manera que el dispositivo soporta que el número de uno o más bits menos significativos cuando se suma al número de uno o más bits más significativos da como resultado un total que es menor que un número total de bits para especificar el valor de recuento del orden de imágenes.

5. Un procedimiento de descodificación de datos de vídeo codificados, el procedimiento que comprende:

determinar, para una imagen actual de los datos de vídeo codificados, una imagen de referencia a largo plazo que se utilizará cuando se descodifique al menos una parte de la imagen actual;

recuperar la imagen de referencia a largo plazo de una memoria intermedia de imágenes descodificadas; y,

descodificar al menos la parte de la imagen actual utilizando la imagen de referencia a largo plazo recuperada, el procedimiento que comprende:

determinar, a partir de un flujo de bits representativo de los datos de vídeo codificados, uno o más bits menos significativos de un valor de recuento del orden de imágenes que identifican la imagen de referencia a largo plazo que se utilizará al descodificar al menos la parte de la imagen actual;

cuando el uno o más bits menos significativos no son los mismos que el uno o más bits menos significativos de cualquier otro valor de recuento del orden de imágenes que identifique cualquier otra imagen almacenada en la memoria intermedia de imágenes descodificadas, utilizar solo uno o más de los bits menos significativos determinados del valor de recuento del orden de imágenes para identificar la imagen de referencia a largo plazo, y recuperar la imagen de referencia a largo plazo de la memoria intermedia de imágenes descodificadas basándose en uno o más bits menos significativos determinados del valor de recuento del orden de imágenes;

cuando el uno o más bits menos significativos no identifican de manera única el valor de recuento del orden de imágenes de la imagen de referencia a largo plazo con respecto al uno o más bits menos significativos de cualquier otro valor de recuento del orden de imágenes que identifique cualquier otra imagen presente en una memoria intermedia de imágenes descodificadas, determinar, a partir del flujo de bits representativo de los datos de vídeo codificados, uno o más bits más significativos del valor de recuento del orden de imágenes, y,

recuperar la imagen de referencia a largo plazo de la memoria intermedia de imágenes descodificadas basándose en uno o más bits menos significativos determinados del valor de recuento del orden de imágenes y el uno o más bits más significativos determinados del valor de recuento del orden de imágenes;

**caracterizado por que** el número de uno o más bits más significativos que se especifica en el flujo de bits se determina basándose en un análisis de todas las imágenes almacenadas en la memoria intermedia de imágenes descodificadas de tal manera que uno o más bits más significativos permiten desambiguar el valor de recuento del orden de imágenes que identifica la imagen de referencia a largo plazo en la memoria intermedia de imágenes descodificadas, y el procedimiento soporta que el número de uno o más bits menos significativos cuando se suma al número de uno o más bits más significativos da como resultado un total que es menor que un número total de bits para especificar el valor de recuento del orden de imágenes.

6. El procedimiento según la reivindicación 5,

en el que la determinación de uno o más bits menos significativos del valor de recuento del orden de imágenes que identifica la imagen de referencia a largo plazo que se utilizará al descodificar al menos la

parte de la imagen actual comprende el análisis, a partir del flujo de bits representativo de los datos de vídeo codificados, el uno o más bits menos significativos del valor de recuento del orden de imágenes que identifica la imagen de referencia a largo plazo que se utilizará al descodificar al menos la parte de la imagen actual; y

en el que la determinación de uno o más bits más significativos del valor de recuento del orden de imágenes que identifica la imagen de referencia a largo plazo comprende analizar, a partir del flujo de bits, el uno o más bits más significativos del valor de recuento del orden de imágenes que identifica la imagen de referencia a largo plazo.

7. El procedimiento según la reivindicación 5, en el que la determinación de uno o más bits menos significativos del valor de recuento del orden de imágenes que identifica la imagen de referencia a largo plazo que se utilizará al descodificar al menos la parte de la imagen actual comprende el análisis, a partir de uno o más de una cabecera de fragmento, de un conjunto de parámetros de imagen y un conjunto de parámetros de secuencia asociados con al menos la parte de la imagen actual, el uno o más bits menos significativos del valor de recuento del orden de imágenes que identifica la imagen de referencia a largo plazo que se utilizará al descodificar al menos la parte de la imagen actual.

8. El procedimiento según la reivindicación 1 o la reivindicación 5, en el que la determinación de uno o más bits más significativos del valor de recuento del orden de imágenes que identifica la imagen de referencia a largo plazo que se utilizará al descodificar al menos la parte de la imagen actual comprende el análisis, a partir de una o más de una cabecera de fragmento, de un conjunto de parámetros de imagen y un conjunto de parámetros de secuencia asociados con al menos la parte de la imagen actual, el uno o más bits más significativos del valor de recuento del orden de imágenes que identifican la imagen de referencia a largo plazo que debe usarse cuando se descodifica al menos la parte de la imagen actual.

9. El procedimiento según la reivindicación 1 o la reivindicación 5, en el que el valor de recuento del orden de imágenes comprende un delta de uno o más bits más significativos del valor de recuento del orden de imágenes calculado como la diferencia entre uno o más bits de un valor de recuento del orden de imágenes que identifica la imagen actual y el uno o más bits más significativos del valor de recuento del orden de imágenes que identifica la imagen de referencia a largo plazo que se utilizará al codificar al menos la parte de la imagen actual.

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

determinar un número de bits utilizados para representar uno o más bits menos significativos del valor de recuento del orden de imágenes que identifica la imagen de referencia a largo plazo que se utilizará al descodificar al menos la parte de la imagen actual; y

analizar el número determinado de bits del flujo de bits representativo de los datos de vídeo codificados, en el que el número determinado de bits analizados representa el uno o más bits menos significativos del valor de recuento del orden de imágenes que identifica la imagen de referencia a largo plazo.

11. El procedimiento según la reivindicación 10, en el que el valor de recuento del orden de imágenes que identifica la imagen de referencia a largo plazo comprende un valor de recuento del orden de imágenes que identifica una imagen de referencia a largo plazo que está presente en la memoria intermedia de imágenes descodificadas pero que no se especifica como una o más imágenes de referencia a largo plazo en un conjunto de parámetros de secuencia asociados con la imagen actual.

12. El procedimiento según la reivindicación 10, en el que el número determinado de bits analizados representa los bits menos significativos del valor de recuento del orden de imágenes que identifica la imagen de referencia a largo plazo como un número binario sin signo.

13. El procedimiento según la reivindicación 10, en el que determinar el número de bits utilizados para representar uno o más bits menos significativos del valor de recuento del orden de imágenes que identifica la imagen de referencia a largo plazo que se usará cuando se descodifique al menos una parte de las imágenes actuales comprende analizar el número de bits utilizados para representar uno o más bits menos significativos del valor de recuento del orden de imágenes que identifica la imagen de referencia a largo plazo que se utilizará al descodificar al menos una parte de las imágenes actuales del flujo de bits.

14. El procedimiento según la reivindicación 10, en el que determinar el número de bits comprende determinar el número de bits que se usarán para representar el uno o más bits menos significativos del valor de recuento del orden de imágenes que identifica la imagen de referencia a largo plazo basada en elementos sintácticos determinados previamente para una o más de la imagen actual y las imágenes codificadas previamente de los datos de vídeo sin tener que analizar a partir del flujo de bits el número determinado de

bits utilizados para representar el uno o más bits menos significativos del valor de recuento del orden de imágenes que identifica la imagen de referencia a largo plazo.

- 5      **15.** Un dispositivo de descodificación de vídeo para descodificar datos de vídeo codificados, el dispositivo de descodificación de vídeo que comprende:

medios para determinar, para una imagen actual de los datos de vídeo codificados, una imagen de referencia a largo plazo que se utilizará cuando se descodifique al menos una parte de la imagen actual;

10      medios para recuperar la imagen de referencia a largo plazo de una memoria intermedia de imágenes descodificadas;

medios para descodificar al menos la parte de la imagen actual utilizando la imagen de referencia a largo plazo recuperada;

15      medios para determinar, a partir de un flujo de bits representativo de los datos de vídeo codificados, uno o más bits menos significativos de un valor de recuento del orden de imágenes que identifican la imagen de referencia a largo plazo que se utilizará al descodificar al menos una parte de la imagen actual, en el que:

20      cuando el uno o más bits menos significativos no son los mismos que el uno o más bits menos significativos de cualquier otro valor de recuento del orden de imágenes que identifique cualquier otra imagen almacenada en la memoria intermedia de imágenes descodificadas en la que también se almacena la imagen de referencia a largo plazo, utilizar solo uno o más de los bits menos significativos determinados del valor de recuento del orden de imágenes para identificar la imagen de referencia a largo plazo, en el que los medios para recuperar la imagen de referencia a largo plazo recuperan la imagen de referencia a largo plazo basándose en uno o más bits menos significativos determinados del valor de recuento del orden de imágenes;

30      cuando el uno o más bits menos significativos no identifican de manera única el valor de recuento del orden de imágenes de la imagen de referencia a largo plazo con respecto al uno o más bits menos significativos de cualquier otro valor de recuento del orden de imágenes que identifique cualquier otra imagen presente en una memoria intermedia de imágenes descodificadas en la que también se almacena la imagen de referencia a largo plazo, utilizar medios para determinar, a partir del flujo de bits representativo de los datos de vídeo codificados, uno o más bits más significativos del valor de recuento del orden de imágenes, y, en el que los medios para recuperar la imagen de referencia a largo plazo de la memoria intermedia de imágenes descodificadas recuperan la imagen de referencia a largo plazo basándose en el uno o más bits menos significativos determinados del valor de recuento del orden de imágenes y el uno o más bits más significativos determinados del valor de recuento del orden de imágenes;

45      **caracterizado por que** el número de uno o más bits más significativos que se especifica en el flujo de bits se determina basándose en un análisis de todas las imágenes almacenadas en la memoria intermedia de imágenes descodificadas de tal manera que uno o más bits más significativos permiten desambiguar el valor de recuento del orden de imágenes que identifica la imagen de referencia a largo plazo en la memoria intermedia de imágenes descodificadas, y el dispositivo soporta que el número de uno o más bits menos significativos cuando se suma al número de uno o más bits más significativos da como resultado un total que es menor que un número total de bits para especificar el valor de recuento del orden de imágenes.

- 50      **16.** Un medio de almacenamiento no transitorio legible por ordenador que tiene instrucciones almacenadas en el mismo que, al ejecutarse, hacen que uno o más procesadores de dispositivo de codificación de vídeo realicen el procedimiento de acuerdo con cualquiera de las reivindicaciones 1 a 3 o 5 a 14.

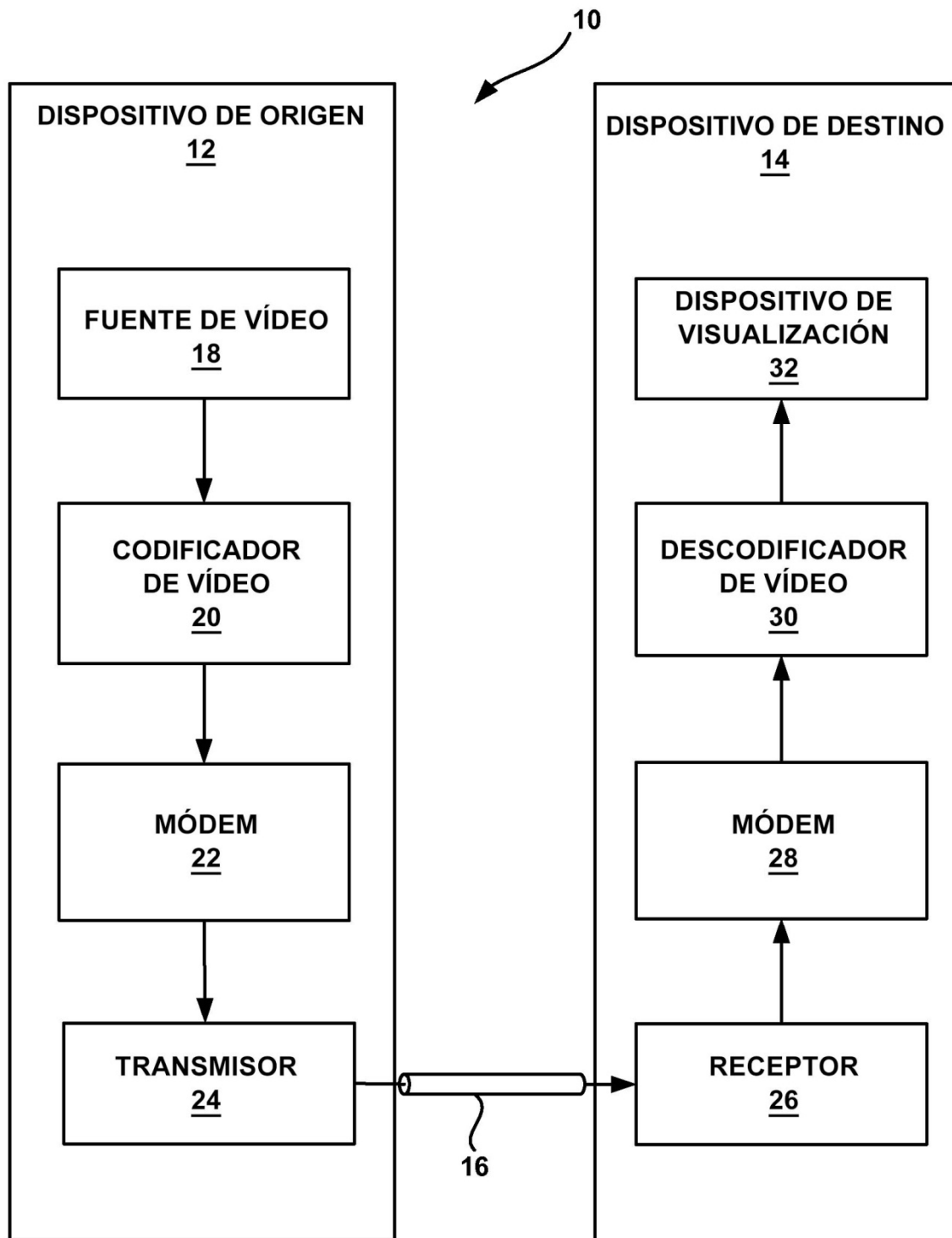
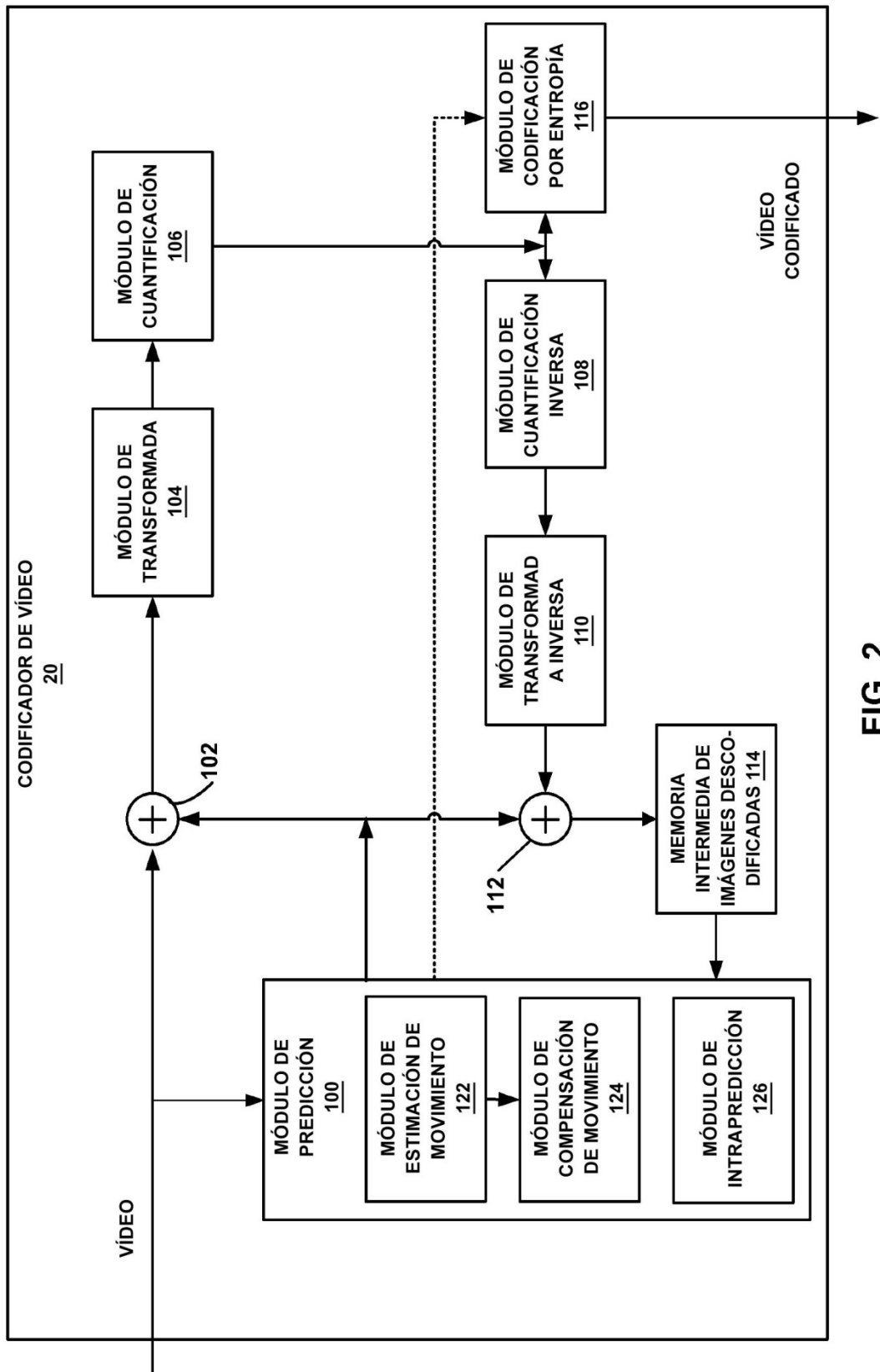


FIG. 1





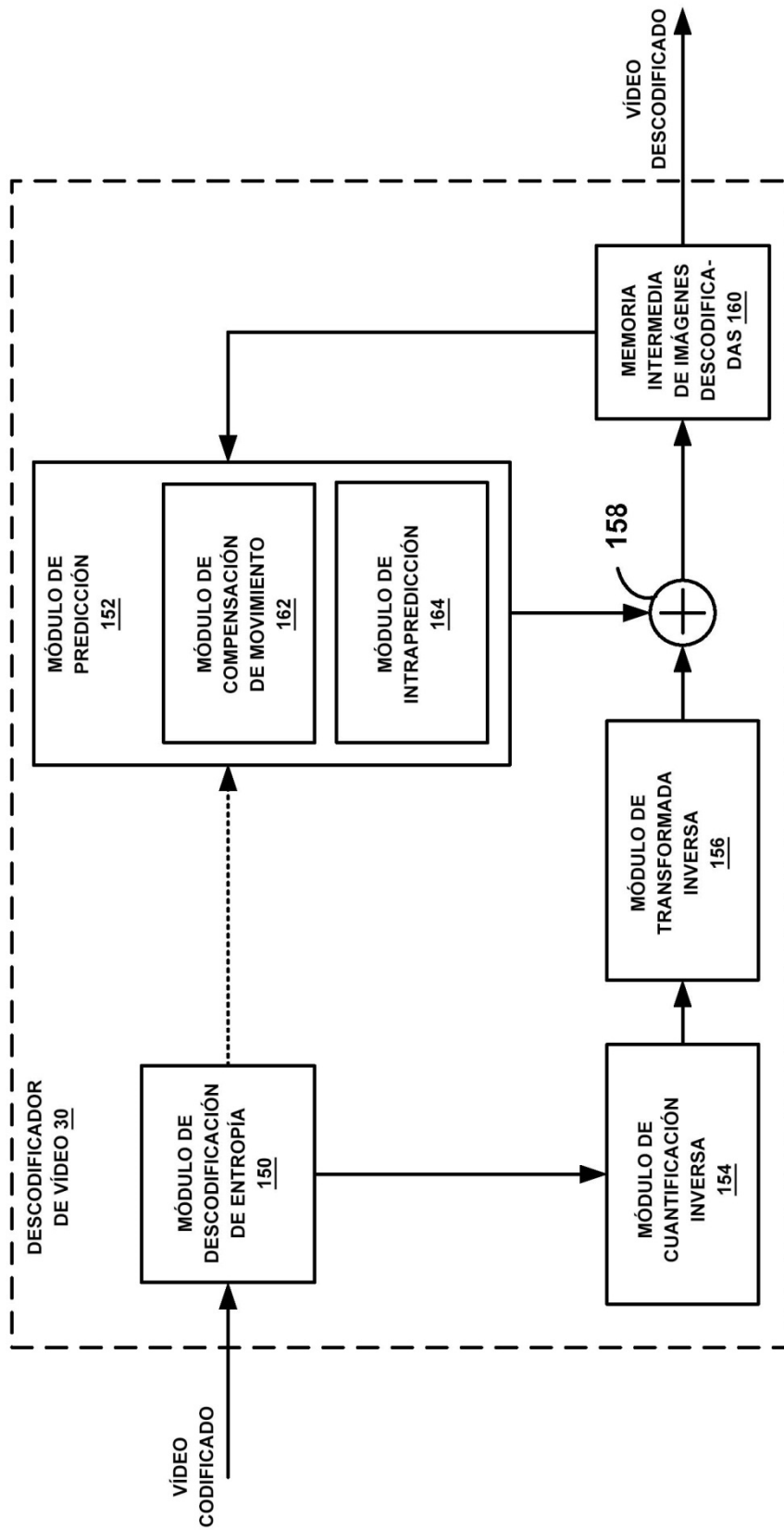


FIG. 3

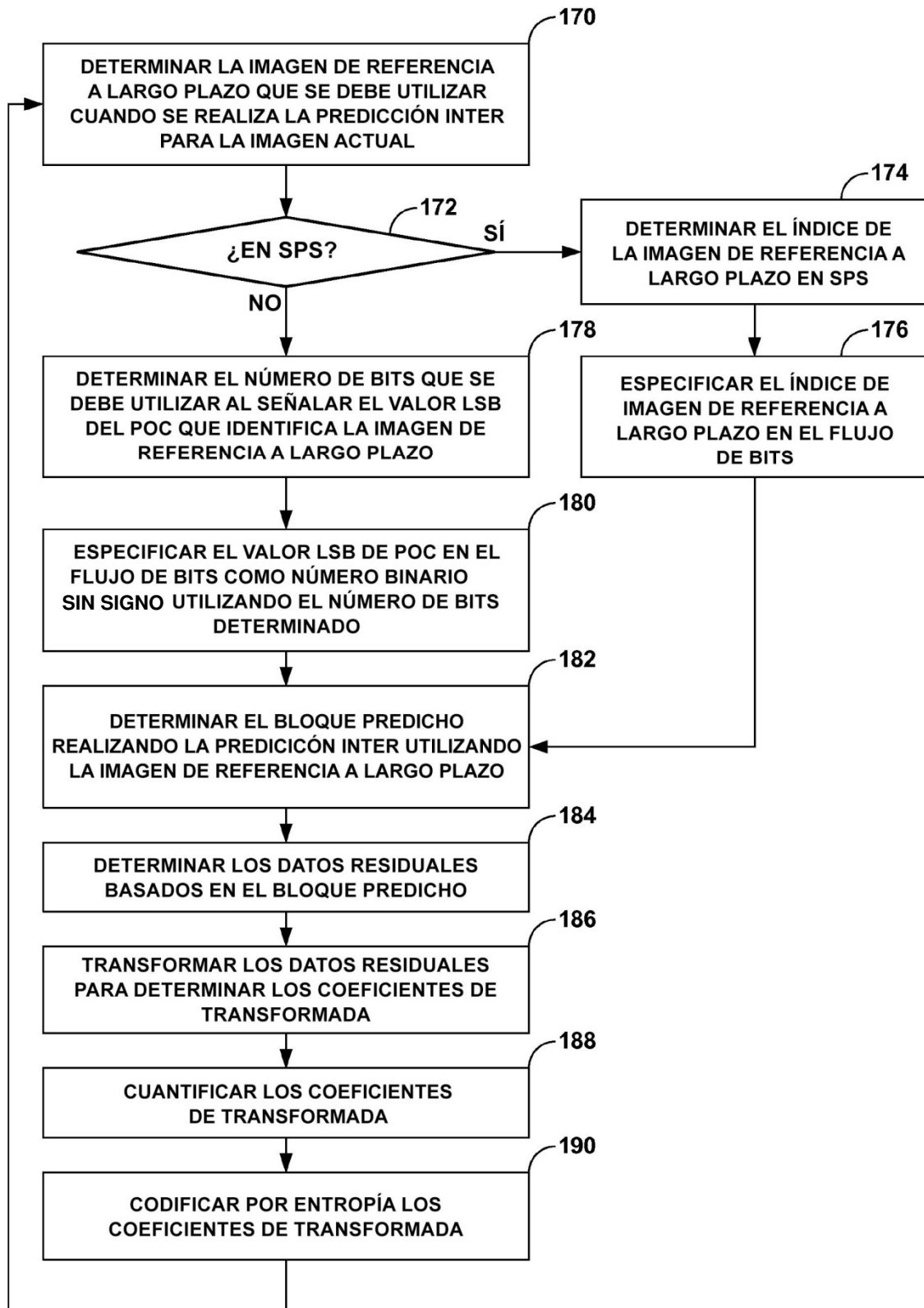


FIG. 4

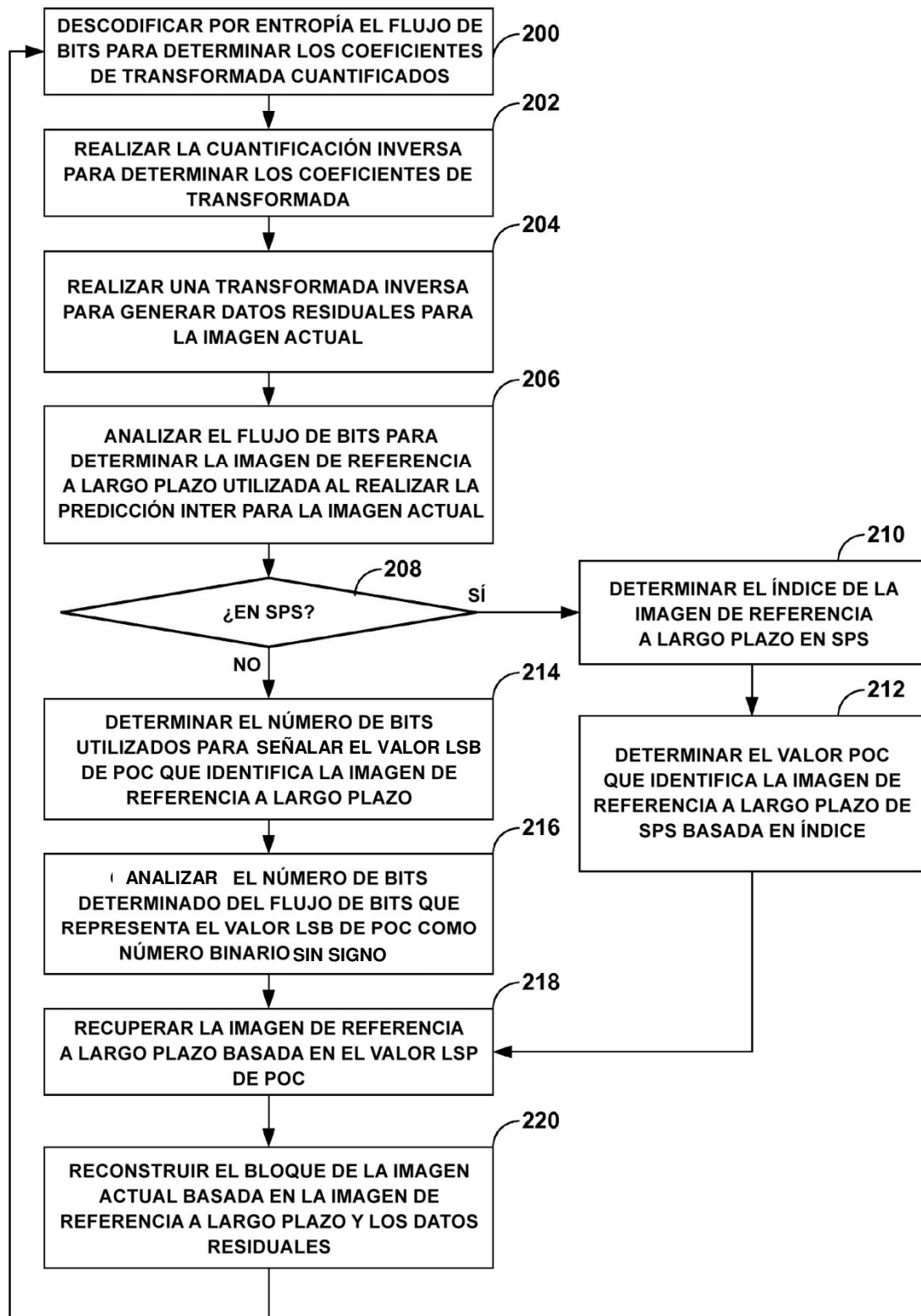


FIG. 5

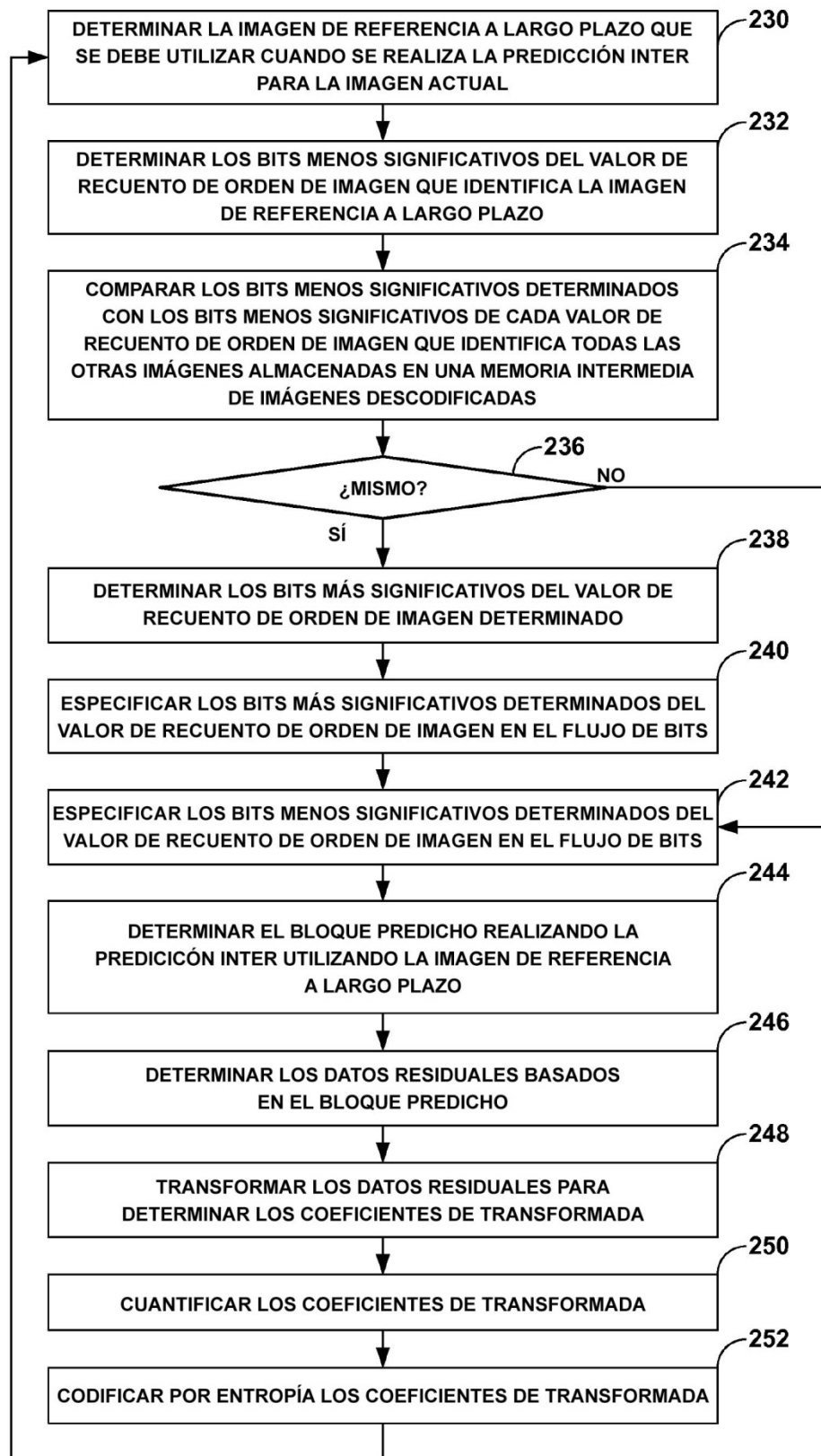


FIG. 6

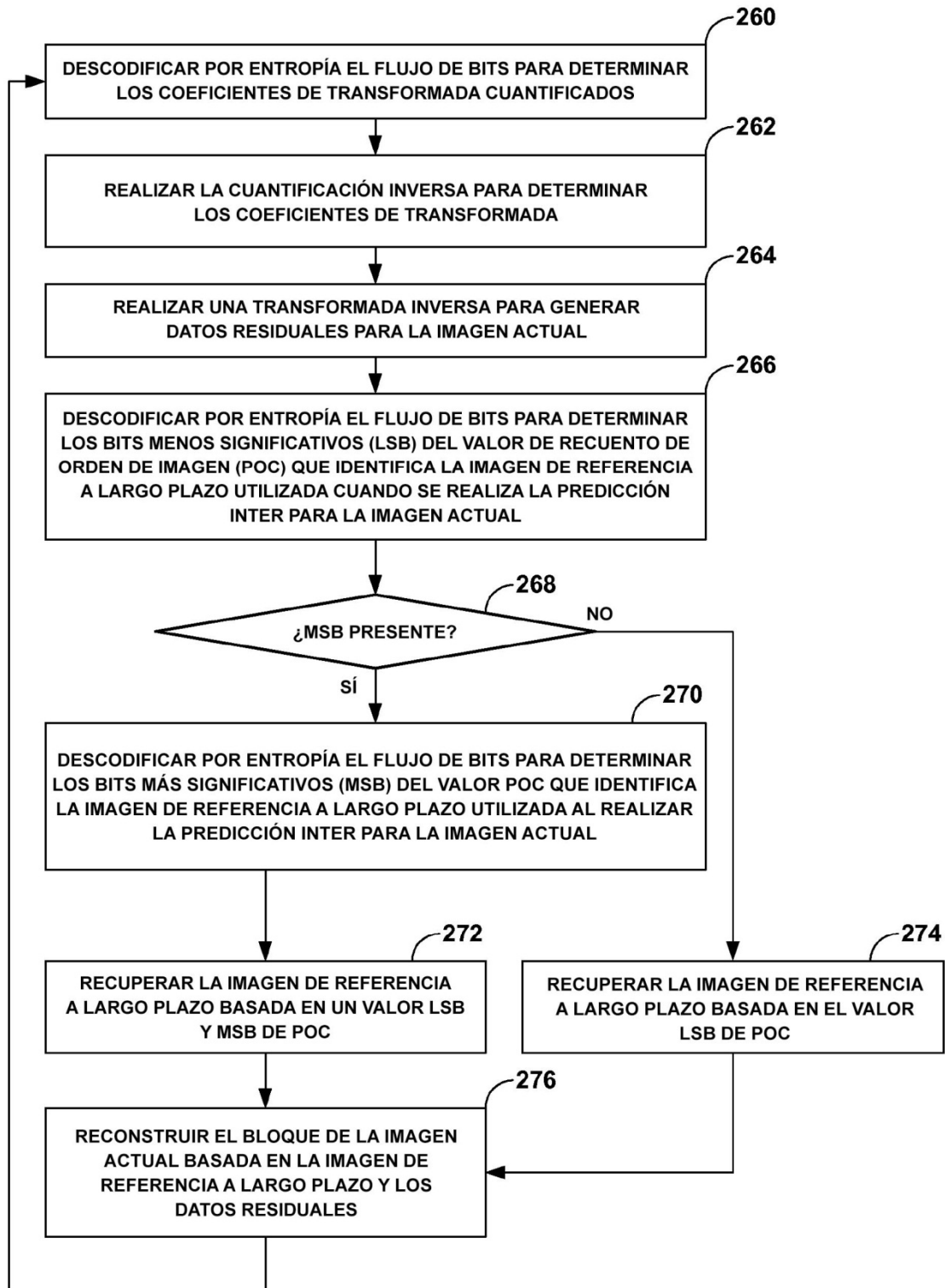


FIG. 7