

19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 736 312**

51 Int. Cl.:

H04N 19/58 (2014.01)

H04N 19/463 (2014.01)

H04N 19/70 (2014.01)

H04N 19/573 (2014.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

86 Fecha de presentación y número de la solicitud internacional: **24.06.2013 PCT/US2013/047367**

87 Fecha y número de publicación internacional: **03.01.2014 WO14004391**

96 Fecha de presentación y número de la solicitud europea: **24.06.2013 E 13734279 (6)**

97 Fecha y número de publicación de la concesión europea: **17.04.2019 EP 2868081**

54 Título: **Señalización de imágenes de referencia a largo plazo para codificación de vídeo**

30 Prioridad:

28.06.2012 US 201261665784 P
21.06.2013 US 201313924016

45 Fecha de publicación y mención en BOPI de la traducción de la patente:
27.12.2019

73 Titular/es:

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

72 Inventor/es:

RAMASUBRAMONIAN, ADARSH KRISHNAN;
WANG, YE-KUI;
JOSHI, RAJAN LAXMAN y
CHEN, YING

74 Agente/Representante:

FORTEA LAGUNA, Juan José

ES 2 736 312 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

Señalización de imágenes de referencia a largo plazo para codificación de vídeo

5 **CAMPO TÉCNICO**

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

10 **ANTECEDENTES**

15 [0002] Las capacidades del vídeo digital pueden incorporarse a una amplia gama de dispositivos, incluyendo televisores digitales, sistemas de radiodifusión digital directa, sistemas de radiodifusión inalámbrica, asistentes digitales personales (PDA), ordenadores portátiles o de escritorio, ordenadores de tableta, lectores de libros electrónicos, cámaras digitales, dispositivos de grabación digital, reproductores de medios digitales, dispositivos de videojuegos, consolas de videojuegos, teléfonos celulares o de radio por satélite, los denominados "teléfonos inteligentes", dispositivos de videoconferencia, dispositivos de transmisión continua de vídeo y similares. Los dispositivos de vídeo digital implementan técnicas de codificación de vídeo, tales como las descritas en las normas definidas por MPEG-2, MPEG-4, ITU-T H.263, ITU-T H.264/MPEG-4, Parte 10, Codificación Avanzada de Vídeo (AVC), la norma de Codificación de Vídeo de Alta Eficacia (HEVC), actualmente en desarrollo, y las ampliaciones de tales normas. Un borrador reciente de la inminente norma HEVC, denominado "HEVC Working Draft 7 [Borrador 7 de trabajo de la HEVC]" o "WD7", se describe en el documento JCTVC-I1003, de Bross et al., titulado "High Efficiency video Coding (HEVC) Text Specification Draft 7 [Memoria descriptiva textual de la Codificación de Vídeo de Alta Eficacia (HEVC), Borrador 7]", Equipo de Colaboración Conjunta en Codificación de Vídeo (JCT-VC) de ITU-T SG16 WP3 e ISO/IEC JTC1/SC29/WG11, novena reunión: San José, California, EE. UU., del 27 de abril al 7 de mayo de 2012, que a partir del 13 de mayo de 2013, puede descargarse desde http://phenix.intevry.fr/jct/doc_end_user/documents/8_San%20Jose/wg11/JCTVC-H1003-v22.zip. Los dispositivos de codificación de vídeo pueden transmitir, recibir, codificar, descodificar y/o almacenar información de vídeo digital más eficientemente, implementando tales técnicas de codificación de vídeo.

20 [0003] Las técnicas de codificación de vídeo incluyen la predicción espacial (intra-imagen) y/o la predicción temporal (entre imágenes) para reducir o eliminar la redundancia intrínseca en las secuencias de vídeo. Para la codificación de vídeo basada en bloques, un fragmento de vídeo (por ejemplo, una trama de vídeo o una parte de una trama de vídeo) puede dividirse en bloques de vídeo, que también pueden denominarse bloques arbolados, unidades de codificación (CU) y/o nodos de codificación. Los bloques de vídeo en un fragmento intracodificado (I) de una imagen se codifican usando predicción espacial con respecto a muestras de referencia en bloques contiguos en la misma imagen. Los bloques de vídeo en un fragmento intercodificado (P o B) de una imagen pueden usar la predicción espacial con respecto a muestras de referencia en bloques contiguos en la misma imagen o la predicción temporal con respecto a muestras de referencia en otras imágenes de referencia. Las imágenes pueden denominarse tramas y las imágenes de referencia pueden denominarse tramas de referencia.

25 [0004] La predicción espacial o temporal da como resultado un bloque predictivo para un bloque que se va a codificar. Los datos residuales representan diferencias de píxeles entre el bloque original que a codificar y el bloque predictivo. Un bloque intercodificado se codifica de acuerdo con un vector de movimiento que apunta a un bloque de muestras de referencia que forman el bloque predictivo, y los datos residuales que indican la diferencia entre el bloque codificado y el bloque predictivo. Un bloque intracodificado se codifica de acuerdo con un modo de intracodificación y a los datos residuales. Para una mayor compresión, los datos residuales pueden transformarse desde el dominio del píxel a un dominio de transformada, dando como resultado coeficientes de transformada residuales, que a continuación se pueden cuantificar. Los coeficientes de transformada cuantificados, dispuestos inicialmente en una formación bidimensional, pueden explorarse con el fin de generar un vector unidimensional de coeficientes de transformada, y puede aplicarse codificación por entropía para lograr aún más compresión.

30 **SUMARIO**

35 [0005] En general, esta divulgación describe técnicas para señalar imágenes de referencia a largo plazo para codificación de vídeo. Una imagen de referencia en general corresponde a una imagen que puede usarse como referencia para la codificación temporal predictiva de datos de otras imágenes. En general, las imágenes de referencia a largo plazo se almacenan en una memoria intermedia de imágenes descodificadas durante más tiempo que otras imágenes de referencia, por ejemplo, imágenes de referencia a corto plazo. Esta divulgación describe varias técnicas relacionadas con la señalización relacionada con imágenes de referencia a largo plazo, por ejemplo, señalización de imágenes de referencia a largo plazo en cabeceras de fragmento o fragmentos de imágenes. Las técnicas de esta divulgación pueden proporcionar una mejora para los procedimientos de señalización de imágenes de referencia a largo plazo en la cabecera de fragmento, por ejemplo, en términos de eficiencia y/o aplicabilidad.

40 [0006] En un ejemplo, un procedimiento incluye descodificar un primer valor representativo de una diferencia entre un valor de bits más significativos (MSB) de un valor de recuento de orden de imágenes (POC) de una imagen actual de datos de vídeo y un primer valor de MSB de un primer valor de POC de una primera imagen de referencia a largo

plazo de los datos de vídeo, descodificar un segundo valor representativo de una diferencia entre un segundo valor de MSB de un segundo valor de POC de una segunda imagen de referencia a largo plazo de los datos de vídeo y el primer valor de MSB, en el que el primer valor de POC y el segundo valor de POC tienen diferentes valores de bits menos significativos (LSB), y descodificar al menos una parte de una imagen actual de los datos de vídeo en relación con al menos una de las primeras imágenes de referencia a largo plazo utilizando el primer valor y la segunda imagen de referencia a largo plazo utilizando el primer valor y el segundo valor.

[0007] En otro ejemplo, un dispositivo para la descodificación de datos de vídeo incluye un descodificador de vídeo configurado para descodificar un primer valor representativo de una diferencia entre un valor de bits más significativos (MSB) de un valor de recuento de orden de imágenes (POC) de una imagen actual de datos de vídeo y un primer valor de MSB de un primer valor de POC de una primera imagen de referencia a largo plazo de los datos de vídeo, descodificar un segundo valor representativo de una diferencia entre un segundo valor de MSB de un segundo valor de POC de una segunda imagen de referencia a largo plazo de los datos de vídeo y el primer valor de MSB, en el que el primer valor de POC y el segundo valor de POC tienen diferentes valores de bits menos significativos (LSB), y descodificar al menos una parte de una imagen actual de los datos de vídeo en relación con al menos una de la primera imagen de referencia a largo plazo utilizando el primer valor y la segunda imagen de referencia a largo plazo utilizando el primer valor y el segundo valor.

[0008] En otro ejemplo, un dispositivo incluye medios para la descodificación de un primer valor representativo de una diferencia entre un valor de bits más significativos (MSB) de un valor de recuento de orden de imágenes (POC) de una imagen actual de datos de vídeo de base y un primer valor de MSB de un primer valor de POC de una primera imagen de referencia a largo plazo de los datos de vídeo, medios para descodificar un segundo valor representativo de una diferencia entre un segundo valor de MSB de un segundo valor de POC de una segunda imagen de referencia a largo plazo de los datos de vídeo y el primer valor de MSB, en el que el primer valor de POC y el segundo valor de POC tienen diferentes valores de bits menos significativos (LSB), y medios para descodificar al menos una parte de una imagen actual de los datos de vídeo en relación con al menos una de la primera imagen de referencia a largo plazo utilizando el primer valor y la segunda imagen de referencia a largo plazo utilizando el primer valor y el segundo valor.

[0009] En otro ejemplo, un medio de almacenamiento legible por ordenador que ha almacenado en el mismo instrucciones que, al ejecutarse, hacen que un procesador descodifique un primer valor representativo de una diferencia entre un valor de bits más significativos (MSB) de un valor de recuento de orden de imágenes (POC) de una imagen actual de datos de vídeo y un primer valor de MSB de un primer valor de POC de una primera imagen de referencia a largo plazo de los datos de vídeo, descodifique un segundo valor representativo de una diferencia entre un segundo valor de MSB de un segundo valor de POC de una segunda imagen de referencia a largo plazo de los datos de vídeo y el primer valor de MSB, en el que el primer valor de POC y el segundo valor de POC tienen diferentes valores de bits menos significativos (LSB), y descodificar al menos una parte de una imagen actual de los datos de vídeo relativos a al menos una de la primera imagen de referencia a largo plazo utilizando el primer valor y la segunda imagen de referencia a largo plazo utilizando el primer valor y el segundo valor.

[0010] En otro ejemplo, un procedimiento de codificación de datos de vídeo incluye codificar al menos una primera parte de una imagen actual de datos de vídeo con respecto a una primera imagen de referencia a largo plazo y al menos una segunda parte de la imagen actual a una segunda imagen de referencia a largo plazo, codificar un primer valor representativo de una diferencia entre un valor de bits más significativos (MSB) base de un valor de recuento de orden de imágenes (POC) de una imagen actual de los datos de vídeo y un primer valor de MSB de un primer valor de POC de la primera imagen de referencia a largo plazo de los datos de vídeo, y codificar un segundo valor representativo de una diferencia entre un segundo valor de MSB de un segundo valor de POC de la segunda imagen de referencia a largo plazo de los datos de vídeo y el primer valor de MSB, en el que el primer valor de POC y el segundo valor de POC tienen diferentes valores de bits menos significativos (LSB).

[0011] En otro ejemplo, un dispositivo para la codificación de datos de vídeo incluye un codificador de vídeo configurado para codificar al menos una primera parte de una imagen actual de datos de vídeo con respecto a una primera imagen de referencia a largo plazo y al menos una segunda parte de la imagen actual a una segunda imagen de referencia a largo plazo, codificar un primer valor representativo de una diferencia entre un valor base de bits más significativos (MSB) de un valor de recuento de orden de imágenes (POC) de una imagen actual de los datos de vídeo y un primer valor de MSB de un primer valor de POC de la primera imagen de referencia a largo plazo de los datos de vídeo, y codificar un segundo valor representativo de una diferencia entre un segundo valor de MSB de un segundo valor de POC de la segunda imagen de referencia a largo plazo de los datos de vídeo y el primer valor de MSB, en el que el primer valor de POC y el segundo valor de POC tienen diferentes valores de bits menos significativos (LSB).

[0012] En otro ejemplo, un dispositivo para la codificación de datos de vídeo incluye medios para codificar al menos una primera parte de una imagen actual de datos de vídeo con respecto a una primera imagen de referencia a largo plazo y al menos una segunda parte de la imagen actual a una segunda imagen de referencia a largo plazo, medios para codificar un primer valor representativo de una diferencia entre un valor base de bits más significativos (MSB) de un valor de recuento de orden de imágenes (POC) de una imagen actual de los datos de vídeo y un primer valor de MSB de un primer valor de POC de la primera imagen de referencia a largo plazo de los datos de vídeo, y medios

para codificar un segundo valor representativo de una diferencia entre un segundo valor de MSB de un segundo valor de POC de la segunda imagen de referencia a largo plazo de los datos de vídeo y el primer valor de MSB, en el que el primer valor de POC y el segundo valor de POC tienen diferentes valores de bits menos significativos (LSB).

5 **[0013]** En otro ejemplo, un medio de almacenamiento legible por ordenador ha almacenado en el mismo instrucciones que hacen que un procesador codifique al menos una primera parte de una imagen actual de datos de vídeo con respecto a una primera imagen de referencia a largo plazo y al menos una segunda parte de la imagen actual en una
10 segunda imagen de referencia a largo plazo, codifique un primer valor representativo de una diferencia entre un valor base de bits más significativos (MSB) de un valor de recuento de orden de imágenes (POC) de una imagen actual de los datos de vídeo y un el primer valor de MSB de un primer valor de POC de la primera imagen de referencia a largo
15 plazo de los datos de vídeo, y codifique un segundo valor representativo de una diferencia entre un segundo valor de MSB de un segundo valor de POC de la segunda imagen de referencia a largo plazo de la datos de vídeo y el primer valor de MSB, en el que el primer valor de POC y el segundo valor de POC tienen diferentes valores de bits menos significativos (LSB).

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

20 BREVE DESCRIPCIÓN DE LOS DIBUJOS

[0015]

25 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 una mejor señalización de imágenes de referencia a largo plazo en cabeceras de fragmento.

30 La FIG. 2 es un diagrama de bloques que ilustra un ejemplo de un codificador de vídeo que puede implementar técnicas para una mejor señalización de imágenes de referencia a largo plazo en cabeceras de fragmento.

La FIG. 3 es un diagrama de bloques que ilustra un ejemplo de un descodificador de vídeo que puede implementar técnicas para una mejor señalización de imágenes de referencia a largo plazo en cabeceras de fragmento.

35 La FIG. 4 es un diagrama conceptual que ilustra una secuencia de imágenes de vídeo codificadas.

La FIG. 5 es un diagrama de flujo que ilustra un procedimiento de ejemplo para codificar valores de recuento de orden de imágenes (POC) para imágenes de referencia a largo plazo de acuerdo con las técnicas de esta divulgación.

40 La FIG. 6 es un diagrama de flujo que ilustra un procedimiento de ejemplo para descodificar valores de POC para imágenes de referencia a largo plazo de acuerdo con las técnicas de esta divulgación.

DESCRIPCIÓN DETALLADA

45 **[0016]** En general, los datos de vídeo se representan por una secuencia de imágenes que son capturadas o exhibidas en rápida sucesión. Se debería entender que, en algunos ejemplos, se pueden generar imágenes o partes de las imágenes, por ejemplo, usando gráficos de ordenador, en lugar de (o además de) ser capturadas. El orden en que se han de visualizar las imágenes (que puede en general ser el mismo que el orden en que se capturan o generan las imágenes) puede diferir del orden en que las imágenes se codifican. El orden de visualización de las imágenes en
50 general se representa mediante los valores de recuento de orden de imágenes (POC), mientras que el orden de codificación de las imágenes en general puede representarse mediante valores de número de trama (frame_num).

[0017] La codificación de imágenes en general implica aprovechar los datos redundantes que aparecen en las imágenes. Por ejemplo, la codificación espacial, también denominada intra-codificación-predictiva, aprovecha las
55 redundancias que se producen entre los bloques espacialmente contiguos de valores de píxel de una imagen común. Como otro ejemplo, la codificación temporal, también conocida como inter-codificación-predictiva, aprovecha las redundancias que se producen entre diferentes imágenes que se capturan, generan o exhiben en diferentes instancias temporales. Más en particular, una imagen se puede dividir en fragmentos, que se pueden designar para predicción intra o predicción inter. Además, cada fragmento puede dividirse en bloques (por ejemplo, las unidades de codificación más grandes (LCU)), y cada uno de los bloques puede adicionalmente dividirse o codificarse basándose en si el
60 fragmento correspondiente ha sido designado para la predicción intra o la predicción inter.

[0018] Con respecto al ejemplo de la codificación con predicción inter, los bloques de una imagen que se está codificando actualmente (también denominada "imagen actual") pueden predecirse a partir de una imagen de
65 referencia. El documento WD7 define una "imagen de referencia" como una imagen con un nal_ref_flag igual a 1. El nal_ref_flag es un elemento sintáctico de una unidad de capa de abstracción de red (NAL) que indica si los datos

incluidos en la unidad de NAL deben tratarse como una imagen de referencia. El documento WD7 también provee que una imagen de referencia contiene muestras (es decir, valores de píxel) que pueden usarse para la predicción inter en el proceso de descodificación de imágenes posteriores en orden de descodificación.

5 **[0019]** Por otra parte, el documento WD7 distingue imágenes de referencia a largo plazo de imágenes de referencia a corto plazo. Por ejemplo, el documento WD7 define una imagen de referencia a largo plazo como una imagen de referencia que está marcada como "utilizada para referencia a largo plazo". El documento WD7 provee que un indicador en un conjunto de parámetros de secuencia (SPS) indica si las imágenes de referencia a largo plazo se señalan en absoluto para una secuencia de vídeo codificada. De acuerdo con el documento WD7, los codificadores de vídeo señalan los bits menos significativos (LSB) de los valores de POC de las imágenes de referencia a largo plazo en las cabeceras de fragmento de los fragmentos. La señalización de los LSB puede dar como resultado ahorros de bits, en relación con la señalización del valor de POC completo. Los codificadores de vídeo, de acuerdo con el documento WD7, pueden señalar los bits más significativos (MSB) de los valores de POC de una imagen de referencia a largo plazo si hay más de una imagen de referencia en una memoria intermedia de imágenes descodificadas (DPB) que tiene los mismos bits LSB que la imagen de referencia a largo plazo. Además, de acuerdo con el documento WD7, los codificadores de vídeo pueden usar un indicador para indicar si las imágenes de referencia a largo plazo señaladas pueden ser utilizadas por una imagen actual como referencia.

20 **[0020]** Los codificadores de vídeo puede mantener los MSB de la imagen actual, a la que se hace referencia en el presente documento como un "valor de MSB base," y señalar el MSB de imagen de referencia a largo plazo i usando valores de desviación relativos a los valores MSB base, por ejemplo, utilizando el elemento sintáctico $\text{delta_poc_msb_cycle_lt}[i]$ para la imagen i . Este elemento sintáctico se menciona a continuación como el valor del "ciclo de MSB". En otras palabras, para calcular el valor de POC para una imagen en particular, un codificador de vídeo puede calcular: (ciclo de MSB + MSB base para la imagen) concatenado con (LSB para la imagen).

25 **[0021]** En WD7, el valor del ciclo de MSB se codificó como una diferencia entre el valor base MSB y los MSB de una imagen para la que estaba siendo señalado el ciclo de MSB, es decir, la imagen i . Solo cuando se señalaran dos o más imágenes de referencia a largo plazo con los mismos valores LSB, el ciclo de MSB de la primera imagen de referencia a largo plazo se señalarían como tal y los valores del ciclo de MSB para el resto de las imágenes de referencia a largo plazo se señalarían como desviaciones relativas entre sí. Sin embargo, esta señalización de imágenes de referencia a largo plazo, de acuerdo con WD7, puede experimentar una o más deficiencias. Por ejemplo, si hay dos imágenes de referencia a largo plazo con diferentes valores de LSB, pero ambas tienen señalización $\text{delta_poc_msb_cycle_lt}[i]$, la sintaxis actual señalaría la $\text{delta_poc_msb_cycle_lt}[i]$ como se comentó anteriormente, es decir, como una desviación relativa al valor de MSB base. Sin embargo, esta divulgación reconoce que la señalización de la diferencia entre los valores de ciclo de MSB para las imágenes puede ser más eficiente.

35 **[0022]** La señalización de los datos de identificación de imágenes de referencia a largo plazo en la cabecera de fragmento, de acuerdo con WD7, puede experimentar una o más deficiencias. Por ejemplo, si una imagen actual tiene un valor de POC igual a $\text{MaxPicOrderCntLsb} * N - 1$, y la imagen que tiene un valor de POC igual a $\text{MaxPicOrderCntLsb} * N$ es la primera imagen de referencia a largo plazo para la cual se señala el $\text{delta_poc_msb_cycle_lt}[i]$ entonces WD7 no proporciona técnicas para señalar $\text{delta_poc_msb_cycle_lt}[i]$. El valor de $\text{delta_poc_msb_cycle_lt}[i]$, en este caso, debe ser, pero no puede ser, -1 (uno negativo).

45 **[0023]** Como otro ejemplo, si hay dos imágenes de referencia a largo plazo con diferentes valores LSB, pero ambas tienen $\text{delta_poc_msb_cycle_lt}[i]$ señalado, la sintaxis actual señalaría la $\text{delta_poc_msb_cycle_lt}[i]$ como tal. Sin embargo, esta divulgación reconoce que señalar la diferencia para una de las imágenes sería más eficiente.

50 **[0024]** Como otro ejemplo, cuando hay más de una imagen de referencia en el DPB que tiene los mismos LSB que los de la imagen de referencia a largo plazo señalada, WD7 exige que el ciclo de MSB del LTRP también se señale utilizando el elemento sintáctico $\text{delta_poc_msb_cycle_lt}[i]$. Sin embargo, esta divulgación reconoce que, si hay una imagen en la DPB que se marcará como imagen de referencia a corto plazo en la imagen actual, entonces incluir la imagen de referencia a corto plazo en el conjunto candidato de imágenes de referencia a largo plazo puede acabar señalando más bits de MSB de los necesarios.

55 **[0025]** Como otro ejemplo, cuando hay n imágenes de referencia en el DPB que tienen POC LSB igual a $\text{poc_lsb_lt}[i]$, y donde k imágenes de referencia a largo plazo, donde k es inferior o igual a n , se señalan en la imagen actual que tiene LSB igual a $\text{poc_lsb_lt}[i]$, la sintaxis definida en WD7 exige que todas las k imágenes tengan $\text{delta_poc_msb_present_flag}[i]$ establecida en 1. Sin embargo, si $k = n$, esta divulgación reconoce que un codificador de vídeo no necesita establecer $\text{delta_poc_msb_present_flag}[i]$ como 1 para todas las k de las imágenes de referencia a largo plazo, sino que sería suficiente establecer $\text{delta_poc_msb_present_flag}[i]$ como 1 para $n - 1$ imágenes. Las técnicas de esta divulgación, como se describe con mayor detalle con respecto a las figuras a continuación, pueden superar cualquiera o todas estas deficiencias. En general, esta divulgación describe varias técnicas que pueden mejorar la señalización para imágenes de referencia a largo plazo en una cabecera de fragmento.

65 **[0026]** La FIG. 1 es un diagrama de bloques que ilustra un sistema de codificación y descodificación de vídeo de ejemplo 10 que puede utilizar técnicas para la señalización mejorada de imágenes de referencia a largo plazo en

cabeceras de fragmento. Como se muestra en la figura 1, el sistema 10 incluye un dispositivo de origen 12 que proporciona datos de vídeo codificados, a descodificar más adelante por un dispositivo de destino 14. En particular, el dispositivo de origen 12 proporciona los datos de vídeo al dispositivo de destino 14 por medio de un medio legible por ordenador 16. El dispositivo de origen 12 y el dispositivo de destino 14 pueden comprender cualquiera entre una amplia gama de dispositivos, incluidos ordenadores de escritorio, ordenadores plegables (es decir, portátiles), ordenadores de tableta, descodificadores, equipos telefónicos de mano tales como los denominados teléfonos «inteligentes», los denominados paneles «inteligentes», televisores, cámaras, dispositivos de visualización, reproductores de medios digitales, consolas de videojuegos, dispositivos de transmisión de vídeo en tiempo real o similares. En algunos casos, el dispositivo de origen 12 y el dispositivo de destino 14 pueden estar equipados para la comunicación inalámbrica.

[0027] El dispositivo de destino 14 puede recibir los datos de vídeo codificados que se van a descodificar por medio del medio legible por ordenador 16. El medio legible por ordenador 16 puede comprender cualquier tipo de medio o dispositivo capaz de desplazar los datos de vídeo codificados desde el dispositivo de origen 12 al dispositivo de destino 14. En un ejemplo, el medio legible por ordenador 16 puede comprender un medio de comunicación para habilitar el dispositivo de origen 12 para transmitir datos de vídeo codificados directamente al dispositivo de destino 14 en tiempo real. Los datos de vídeo codificados se pueden modular de acuerdo con una norma de comunicación, tal como un protocolo de comunicación inalámbrica, y transmitir al dispositivo de destino 14. El medio de comunicación puede comprender cualquier medio de comunicación, inalámbrica o cableada, tal como un espectro de radiofrecuencia (RF) o una o más líneas físicas de transmisión. El medio de comunicación puede formar parte de una red basada en paquetes, tal como una red de área local, una red de área amplia o una red global tal como Internet. El medio de comunicación puede incluir encaminadores, conmutadores, estaciones base o cualquier otro equipo que pueda ser útil para facilitar la comunicación desde el dispositivo de origen 12 hasta el dispositivo de destino 14.

[0028] En algunos ejemplos, pueden emitirse datos codificados desde la interfaz de salida 22 hasta un dispositivo de almacenamiento. De forma similar, se puede acceder a los datos codificados del dispositivo de almacenamiento mediante una interfaz de entrada. El dispositivo de almacenamiento puede incluir cualquiera entre varios medios de almacenamiento de datos, distribuidos o de acceso local, tales como un disco duro, discos Blu-ray, DVD, CD-ROM, memoria flash, memoria volátil o no volátil, o cualquier otro medio de almacenamiento digital adecuado para almacenar datos de vídeo codificados. En un ejemplo adicional, el dispositivo de almacenamiento puede corresponder a un servidor de ficheros o a otro dispositivo de almacenamiento intermedio que pueda almacenar el vídeo codificado generado por el dispositivo de origen 12.

[0029] El dispositivo de destino 14 puede acceder a datos de vídeo almacenados desde el dispositivo de almacenamiento por medio de transmisión continua o descarga. El servidor de ficheros puede ser cualquier tipo de servidor capaz de almacenar datos de vídeo codificados y transmitir esos datos de vídeo codificados al dispositivo de destino 14. Los servidores de ficheros de ejemplo incluyen un servidor de la Red (por ejemplo, para una sede de la Red), un servidor del FTP, dispositivos de almacenamiento conectados en red (NAS) o una unidad de disco local. El dispositivo de destino 14 puede acceder a los datos de vídeo codificados a través de cualquier conexión de datos estándar, incluida una conexión a Internet. Esto puede incluir un canal inalámbrico (por ejemplo, una conexión de Wi-Fi), una conexión cableada (por ejemplo, DSL, módem de cable, etc.) o una combinación de ambas que sea adecuada para acceder a datos de vídeo codificados almacenados en un servidor de ficheros. La transmisión de datos de vídeo codificados desde el dispositivo de almacenamiento puede ser una transmisión continua, una transmisión de descarga o una combinación de ambas.

[0030] Las técnicas de esta divulgación no están limitadas a aplicaciones o configuraciones inalámbricas. Las técnicas se pueden aplicar a la codificación de vídeo como soporte de cualquiera entre varias aplicaciones de multimedia, tales como difusiones de televisión por el aire, transmisiones de televisión por cable, transmisiones de televisión por satélite, transmisiones continuas de vídeo por Internet, tales como la transmisión continua adaptativa dinámica por HTTP (DASH), el vídeo digital que se codifica en un medio de almacenamiento de datos, la descodificación de vídeo digital almacenado en un medio de almacenamiento de datos u otras aplicaciones. En algunos ejemplos, el sistema 10 puede estar configurado para prestar soporte a una transmisión de vídeo unidireccional o bidireccional, a fin de prestar soporte a aplicaciones tales como la transmisión de vídeo en tiempo real, la reproducción de vídeo, la difusión de vídeo y/o la videotelefonía.

[0031] En el ejemplo de la FIG. 1, el dispositivo de origen 12 incluye un origen de vídeo 18, un codificador de vídeo 20 y una interfaz de salida 22. El dispositivo de destino 14 incluye una interfaz de entrada 28, un descodificador de vídeo 30 y un dispositivo de visualización 32. De acuerdo con esta divulgación, el codificador de vídeo 20 del dispositivo de origen 12 puede estar configurado para aplicar las técnicas para la señalización mejorada de imágenes de referencia a largo plazo en cabeceras de fragmento. 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 un origen de vídeo externo 18, 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.

[0032] El sistema 10 ilustrado de la FIG. 1 es simplemente un ejemplo. Las técnicas para la señalización mejorada de imágenes de referencia a largo plazo en cabeceras de fragmento pueden ser realizadas por cualquier dispositivo de codificación y/o decodificación de vídeo digital. Aunque, en general, las técnicas de la presente divulgación son llevadas a cabo por un dispositivo de codificación de vídeo, las técnicas también pueden ser llevadas a cabo por un codificador/decodificador de vídeo, denominado típicamente "CÓDEC". Además, las técnicas de la presente divulgación también se pueden llevar a cabo mediante un preprocesador de vídeo. El dispositivo de origen 12 y el dispositivo de destino 14 son simplemente ejemplos de dichos dispositivos de codificación en los que el dispositivo de origen 12 genera datos de vídeo codificados para su transmisión al dispositivo de destino 14. 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 decodificación de vídeo. De ahí que el sistema 10 pueda prestar soporte a la transmisión de vídeo unidireccional o bidireccional entre los dispositivos de vídeo 12, 14, por ejemplo, para la transmisión continua de vídeo, la reproducción de vídeo, la difusión de vídeo o la videotelefonía.

[0033] El origen de vídeo 18 del dispositivo de origen 12 puede incluir un dispositivo de captación de vídeo, tal como una videocámara, un archivo de vídeo que contiene vídeo captado previamente y/o una interfaz de alimentación de vídeo para recibir vídeo desde un proveedor de contenidos de vídeo. Como otra alternativa, el origen 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 el origen 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. Sin embargo, como se ha mencionado anteriormente, las técnicas descritas en la presente divulgación pueden ser aplicables a la codificación de vídeo en general, y se pueden aplicar a aplicaciones inalámbricas y/o cableadas. En cada caso, el codificador de vídeo 20 puede codificar el vídeo capturado, precaptado o generado por ordenador. La información de vídeo codificada se puede emitir, a continuación, por la interfaz de salida 22 hacia un medio legible por ordenador 16.

[0034] El medio legible por ordenador 16 puede incluir medios transitorios, tales como una radiodifusión inalámbrica o transmisión de red cableada, o medios de almacenamiento (es decir, medios de almacenamiento no transitorio), tales como un disco duro, una unidad de memoria flash, un disco compacto, un disco de vídeo digital, un disco Blu-ray u otros medios legibles por ordenador. En algunos ejemplos, un servidor de red (no mostrado) puede recibir datos de vídeo codificados desde el dispositivo de origen 12 y proporcionar los datos de vídeo codificados al dispositivo de destino 14, por ejemplo, mediante transmisión por red. De forma similar, un dispositivo informático de una utilidad de producción de un medio, tal como una utilidad de grabación de discos, puede recibir datos de vídeo codificados desde el dispositivo de origen 12 y producir un disco que contiene los datos de vídeo codificados. Por lo tanto, se puede entender que el medio legible por ordenador 16 incluye uno o más medios legibles por ordenador de diversas formas, en diversos ejemplos.

[0035] La interfaz de entrada 28 del dispositivo de destino 14 recibe información desde el medio legible por ordenador 16. La información del medio legible por ordenador 16 puede incluir información sintáctica definida por el codificador de vídeo 20, que también usa el decodificador de vídeo 30, que incluye elementos sintácticos que describen características y/o procesamiento de bloques y otras unidades codificadas, por ejemplo, los GOP. El dispositivo de visualización 32 muestra los datos de vídeo descodificados a un usuario, y puede comprender cualquiera entre varios 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.

[0036] El codificador de vídeo 20 y el decodificador de vídeo 30 pueden funcionar de acuerdo con una norma de codificación de vídeo, tal como la norma de codificación de vídeo de alta eficacia (HEVC), actualmente en fase de desarrollo, y se pueden ajustar al modelo de prueba de la HEVC (HM). De forma alternativa, el codificador de vídeo 20 y el decodificador de vídeo 30 pueden funcionar de acuerdo con otras normas de propiedad o industriales, tales como la norma ITU-T H.264, denominada de forma alternativa MPEG-4, parte 10, codificación avanzada de vídeo (AVC), o ampliaciones de dichas normas. Sin embargo, las técnicas de esta divulgación no están limitadas a ninguna norma de codificación particular. Otros ejemplos de normas de codificación de vídeo incluyen MPEG-2 e ITU-T H.263. Aunque no se muestra en la figura 1, en algunos aspectos, tanto el codificador de vídeo 20 como el decodificador de vídeo 30 pueden estar integrados con un codificador y decodificador 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 se pueden ajustar al protocolo multiplexador H.223 de la ITU o a otros protocolos tales como el protocolo de datagramas de usuario (UDP).

[0037] La norma ITU-T H.264/MPEG-4 (AVC) se formuló por el grupo de expertos en codificación de vídeo (VCEG) de ITU-T, conjuntamente con el grupo de expertos en imágenes en movimiento (MPEG) de ISO/IEC, como el producto de una asociación colectiva conocida como el equipo de vídeo conjunto (JVT). En algunos aspectos, las técnicas descritas en la presente divulgación se pueden aplicar a dispositivos que se ajustan en general a la norma H.264. La norma H.264 está descrita en la recomendación H.264 de ITU-T, codificación avanzada de vídeo para servicios audiovisuales genéricos, por el grupo de estudio de ITU-T, y con fecha de marzo de 2005, que se puede denominar

en el presente documento la norma H.264 o la especificación H.264, o la norma o especificación H.264/AVC. El equipo de vídeo conjunto (JVT) continúa trabajando en extensiones para la H.264/MPEG-4 AVC.

[0038] El codificador de vídeo 20 y el descodificador de vídeo 30 se pueden implementar, cada uno, como cualquiera entre una variedad de circuitos codificadores adecuados, tales como uno o más microprocesadores, procesadores de señales digitales (DSP), circuitos integrados específicos de la aplicación (ASIC), formaciones de puertas programables *in situ* (FPGA), lógica discreta, software, hardware, firmware o cualquier combinación de los mismos. Cuando las técnicas se implementan parcialmente en software, un dispositivo puede almacenar instrucciones para el software en un medio adecuado no transitorio legible por ordenador, y ejecutar las instrucciones en hardware usando uno o más procesadores para realizar las técnicas de esta divulgación. Cada uno entre el codificador de vídeo 20 y el descodificador de vídeo 30 se puede incluir en uno o más codificadores o descodificadores, cualquiera de los cuales se puede integrar como parte de un codificador/descodificador (CÓDEC) combinado en un dispositivo respectivo.

[0039] El equipo JCT-VC está trabajando en el desarrollo del estándar HEVC. La iniciativa de normalización HEVC se basa en un modelo en evolución de un dispositivo de codificación de vídeo denominado modelo de prueba HEVC (HM). El HM supone varias capacidades adicionales de los dispositivos de codificación de vídeo respecto a los dispositivos existentes de acuerdo con, por ejemplo, la norma ITU-T H.264/AVC. Por ejemplo, mientras que la norma H.264 proporciona nueve modos de codificación mediante predicción intra, el HM puede proporcionar hasta treinta y tres modos de codificación mediante predicción intra.

[0040] En general, el modelo de funcionamiento del HM describe que una trama o imagen de vídeo puede dividirse en una secuencia de bloques arbolados o unidades de codificación de máximo tamaño (LCU), que incluyen muestras tanto de luma como de croma. Los datos sintácticos dentro de un flujo de bits pueden definir un tamaño para la LCU, que es una máxima unidad de codificación en lo que respecta al número de píxeles. Un fragmento incluye un número de bloques de árbol consecutivos en orden de codificación. Una trama o imagen de vídeo puede dividirse en uno o más fragmentos. Cada bloque arbolado puede dividirse en unidades de codificación (CU) de acuerdo con un árbol cuádruple. En general, una estructura de datos de árbol cuádruple incluye un nodo por CU, con un nodo raíz correspondiente al bloque arbolado. Si una CU se divide en cuatro sub-CU, el nodo correspondiente a la CU incluye cuatro nodos hoja, cada uno de los cuales corresponde a una de las sub-CU.

[0041] Cada nodo de la estructura de datos de árbol cuádruple puede proporcionar datos sintácticos para la CU correspondiente. Por ejemplo, un nodo en el árbol cuádruple puede incluir un indicador de división, que indica si la CU correspondiente al nodo está dividida en varias 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. Si una CU no está dividida adicionalmente, se denomina CU hoja. En esta divulgación, cuatro sub-CU de una CU hoja también se denominarán CU hojas incluso aunque no haya ninguna división explícita de la CU hoja original. Por ejemplo, si una CU con un tamaño de 16x16 no se divide adicionalmente, las cuatro sub-CU de tamaño 8x8 también se denominarán CU hojas aunque la CU de tamaño 16x16 no se haya dividido nunca.

[0042] Una CU tiene un propósito similar a un macrobloque del estándar H.264, excepto que una CU no tiene una distinción de tamaño. Por ejemplo, un bloque arbolado puede dividirse en cuatro nodos secundarios (también denominados sub-CU) y cada nodo secundario puede a su vez ser un nodo principal y dividirse en otros cuatro nodos secundarios. Un nodo secundario final, no dividido, denominado nodo hoja del árbol cuádruple, comprende un nodo de codificación, también denominado CU hoja. Los datos sintácticos asociados a un flujo de bits codificado pueden definir un número máximo de veces en que puede dividirse un bloque arbolado, denominado profundidad de CU máxima, y también pueden definir un tamaño mínimo de los nodos de codificación. Por consiguiente, un flujo de bits también puede definir una unidad de codificación de tamaño mínimo (SCU). Esta divulgación utiliza el término "bloque" para referirse a cualquiera entre una CU, PU o TU, en el contexto de la HEVC, o a estructuras de datos similares en el contexto de otras normas (por ejemplo, macrobloques y subbloques de los mismos en la norma H.264/AVC).

[0043] Una CU incluye un nodo de codificación y unidades de predicción (PU) y unidades de transformada (TU) asociadas al nodo de codificación. Un tamaño de la CU corresponde a un tamaño del nodo de codificación y debe ser de forma cuadrada. El tamaño de la CU puede variar desde 8x8 píxeles hasta el tamaño del bloque arbolado, con un máximo de 64x64 píxeles o más. Cada CU puede contener una o más PU y una o más TU. Los datos sintácticos asociados a una CU pueden describir, por ejemplo, la división de la CU en una o más PU. Los modos de división pueden diferir entre si la CU está codificada en modo de salto o directa, codificada en modo de predicción intra o codificada en modo de predicción inter. Las PU pueden dividirse para tener forma no cuadrada. Los datos sintácticos asociados a una CU también pueden describir, por ejemplo, la división de la CU en una o más TU de acuerdo con un árbol cuádruple. Una TU puede tener forma cuadrada o no cuadrada (por ejemplo, rectangular).

[0044] El estándar HEVC soporta transformaciones de acuerdo con las TU, que pueden ser diferentes para diferentes CU. El tamaño de las TU se basa típicamente en el tamaño de las PU dentro de una CU definida dada para una LCU dividida, aunque puede que no sea siempre así. Las TU son típicamente del mismo tamaño o de un tamaño más pequeño que las PU. En algunos ejemplos, las muestras residuales correspondientes a una CU pueden subdividirse en unidades más pequeñas utilizando una estructura de árbol cuádruple conocida como «árbol cuádruple residual» (RQT). Los nodos hoja del RQT pueden denominarse unidades de transformada (TU). Los valores de diferencias de

píxeles asociados a las TU pueden transformarse para generar coeficientes de transformada, que pueden cuantificarse.

[0045] Una CU hoja puede incluir una o más unidades de predicción (PU). En general, una PU representa una zona espacial correspondiente a la totalidad, o a una parte, de la CU correspondiente, y puede incluir datos para recuperar una muestra de referencia para la PU. Además, una PU incluye datos relacionados con la predicción. Por ejemplo, cuando la PU está codificada en modo intra, pueden incluirse datos para la PU en un árbol cuádruple residual (RQT), que puede incluir datos que describen un modo de predicción intra para una TU correspondiente a la PU. Como otro ejemplo, cuando la PU está codificada en modo inter, la PU puede incluir datos que definen uno o más vectores de movimiento para la PU. Los datos que definen el vector de movimiento para una PU pueden describir, por ejemplo, una componente horizontal del vector de movimiento, una componente vertical del vector de movimiento, una resolución para el vector de movimiento (por ejemplo, precisión de un cuarto de píxel o precisión de un octavo de píxel), una imagen de referencia a la que apunta el vector de movimiento y/o una lista de imágenes de referencia (por ejemplo, lista 0, lista 1 o lista C) para el vector de movimiento.

[0046] Una CU hoja que tiene una o más PU también puede incluir una o más unidades de transformada (TU). Las unidades de transformada pueden especificarse usando un RQT (también denominado estructura de árbol cuaternario de TU), como se ha expuesto anteriormente. Por ejemplo, un indicador de división puede indicar si una CU hoja está dividida en cuatro unidades de transformada. A continuación, cada unidad de transformada puede dividirse adicionalmente en sub-TU adicionales. Cuando una TU no se divide adicionalmente, puede denominarse una TU hoja. En general, en lo que respecta a la intra-codificación, todas las TU hojas que pertenecen a una CU hoja comparten el mismo modo de predicción intra. Es decir, el mismo modo de predicción intra se aplica en general para calcular valores predichos para todas las TU de una CU hoja. En lo que respecta a la intra-codificación, un codificador de vídeo puede calcular un valor residual para cada TU hoja usando el modo de predicción intra, como una diferencia entre la parte de la CU correspondiente a la TU y el bloque original. Una TU no está necesariamente limitada al tamaño de una PU. De este modo, las TU pueden ser mayores o menores que una PU. En lo que respecta a la intra-codificación, una PU puede estar co-situada con una TU hoja correspondiente para la misma CU. En algunos ejemplos, el tamaño máximo de una TU hoja puede corresponder al tamaño de la CU hoja correspondiente.

[0047] Además, las TU de las CU hojas también pueden asociarse a las respectivas estructuras de datos de árbol cuádruple, denominadas árboles cuádruples residuales (RQT). Es decir, una CU hoja puede incluir un árbol cuaternario que indica cómo la CU hoja está dividida en las TU. El nodo raíz de un árbol cuaternario de TU corresponde en general a una CU hoja, mientras que el nodo raíz de un árbol cuaternario de CU corresponde en general a un bloque arbolado (o LCU). Las TU del RQT que no están divididas se denominan TU hojas. En general, esta divulgación usa los términos CU y TU para hacer referencia a una CU hoja y a una TU hoja, respectivamente, a no ser que se indique lo contrario.

[0048] Una secuencia de vídeo incluye típicamente una serie de tramas o imágenes de vídeo. Un grupo de imágenes (GOP) comprende, en general, una serie de una o más de las imágenes de vídeo. Un GOP puede incluir datos sintácticos en una cabecera del GOP, una cabecera de una o más de las imágenes, o en otras ubicaciones, que describen una serie de imágenes incluidas en el GOP. Cada fragmento de una imagen puede incluir datos sintácticos de fragmento que describen un modo de codificación para el fragmento respectivo. El codificador de vídeo 20 actúa típicamente sobre bloques de vídeo dentro de fragmentos de vídeo individuales con el fin de codificar los datos de vídeo. Un bloque de vídeo puede corresponder a un nodo de codificación dentro de una CU. Los bloques de vídeo pueden tener tamaños fijos o variables y pueden diferir en tamaño de acuerdo con una norma de codificación especificada.

[0049] Como ejemplo, el HM presta soporte a la predicción en diversos tamaños de PU. Si se supone que el tamaño de una CU particular es $2N \times 2N$, el HM admite predicción intra en tamaños de PU de $2N \times 2N$ o $N \times N$ e predicción inter en tamaños de PU simétricas de $2N \times 2N$, $2N \times N$, $N \times 2N$ o $N \times N$. El HM también admite la división asimétrica para predicción inter en tamaños de PU de $2N \times nU$, $2N \times nD$, $nL \times 2N$ y $nR \times 2N$. En la división asimétrica, una dirección de una CU no está dividida, mientras que la otra dirección está dividida en un 25 % y un 75 %. La parte de la CU correspondiente a la división del 25 % está indicada por una «n» seguida por una indicación de «arriba», «abajo», «izquierda» o «derecha». Así pues, por ejemplo, « $2N \times nU$ » se refiere a una CU de tamaño $2N \times 2N$ que está dividida horizontalmente, con una PU de tamaño $2N \times 0,5N$ encima y una PU de tamaño $2N \times 1,5N$ debajo.

[0050] En esta 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 tamaño 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 que en la dirección vertical. Por ejemplo, los bloques pueden comprender NxM píxeles, donde M no es necesariamente igual a N.

[0051] Tras la codificación de predicción intra o predicción inter mediante las PU de una CU, el codificador de vídeo 20 puede calcular los datos residuales para las TU de la CU. Las PU pueden comprender datos sintácticos que

describen un procedimiento o modo de generación de datos de píxeles predictivos en el dominio espacial (también denominado el dominio de píxeles) y las TU pueden comprender coeficientes en el dominio de transformada, tras la aplicación de una transformada, por ejemplo, una transformada de coseno discreta (DCT), una transformada entera, una transformada de ondículas o una transformada conceptualmente similar, a los datos de vídeo residuales. Los datos residuales pueden corresponder a diferencias de píxeles entre píxeles de la imagen no codificada y valores de predicción correspondientes a las PU. El codificador de vídeo 20 puede formar las TU incluyendo los datos residuales para la CU y, a continuación, transformar las TU para generar coeficientes de transformada para la CU.

[0052] Tras cualquier transformada para generar coeficientes de transformada, el codificador de vídeo 20 puede realizar la cuantificación de los coeficientes de transformada. La cuantificación se refiere, en general, a un proceso en el que unos coeficientes de transformada se cuantifican para reducir posiblemente la cantidad de datos usados para representar los coeficientes, proporcionando compresión adicional. El proceso de cuantificación puede reducir la profundidad de bits asociada a algunos, o a la totalidad, de los coeficientes. Por ejemplo, un valor de n bits puede redondearse a la baja hasta un valor de m bits durante la cuantificación, donde n es mayor que m .

[0053] Después de la cuantificación, el codificador de vídeo puede recorrer los coeficientes de transformada, produciendo un vector unidimensional a partir de la matriz bidimensional que incluye los coeficientes de transformada cuantificados. El recorrido puede estar diseñado para colocar los coeficientes de energía más alta (y por lo tanto de menor frecuencia) al frente de la formación y para colocar los coeficientes de energía más baja (y por lo tanto de mayor frecuencia) al final de la formación. En algunos ejemplos, el codificador de vídeo 20 puede utilizar un orden de exploración predefinido para explorar los coeficientes de transformada cuantificados, para producir un vector serializado que se pueda codificar por entropía. En otros ejemplos, el codificador de vídeo 20 puede realizar una exploración adaptativa. Después de recorrer los coeficientes de transformada cuantificados para formar un vector unidimensional, el codificador de vídeo 20 puede realizar la codificación por entropía del vector unidimensional, por ejemplo, de acuerdo con la codificación de longitud variable adaptable al contexto (CAVLC), la codificación aritmética binaria adaptable al contexto (CABAC), la codificación aritmética binaria adaptable al contexto basada en la sintaxis (SBAC), la codificación por entropía por división de intervalos de probabilidad (PIPE) o con otra metodología de codificación por entropía. El codificador de vídeo 20 también puede realizar la codificación por entropía de los elementos sintácticos asociados a los datos de vídeo codificados, para su uso por el descodificador de vídeo 30 en la descodificación de los datos de vídeo.

[0054] Para realizar la CABAC, el codificador de vídeo 20 puede asignar un contexto dentro de un modelo contextual a un símbolo que se va a transmitir. El contexto se puede referir, por ejemplo, a si los valores contiguos del símbolo son distintos de cero o no. Para realizar la CAVLC, el codificador de vídeo 20 puede seleccionar un código de longitud variable para un símbolo a transmitir. Las palabras de código en la VLC se pueden construir de modo que los códigos relativamente más cortos correspondan a símbolos más probables, mientras que los códigos más largos correspondan a símbolos menos probables. De esta forma, el uso de la VLC puede lograr un ahorro en bits con respecto, por ejemplo, al uso de palabras de código de igual longitud para cada símbolo que se va a transmitir. La determinación de la probabilidad se puede basar en un contexto asignado al símbolo.

[0055] De acuerdo con las técnicas de esta divulgación, el codificador de vídeo 20 y el descodificador de vídeo 30 pueden estar configurados para utilizar técnicas mejoradas para datos de señalización para la identificación de imágenes de referencia a largo plazo, por ejemplo, en las cabeceras de fragmento, conjuntos de parámetros (incluyendo conjuntos de parámetros de imagen (PPS) y conjuntos de parámetros de secuencia (SPS)), o similares. En general, el codificador de vídeo 20 y el descodificador de vídeo 30 pueden configurarse para realizar una cualquiera o todas las técnicas de esta divulgación, solas o en cualquier combinación. En algunos ejemplos, el codificador de vídeo 20 y el descodificador de vídeo 30 pueden configurarse para realizar procedimientos para señalar valores negativos de ciclo delta MSB de POC para imágenes de referencia a largo plazo. En algunos ejemplos, el codificador de vídeo 20 y el descodificador de vídeo 30 pueden configurarse para realizar procedimientos para una señalización más eficiente de los valores de ciclo delta MSB de POC para imágenes de referencia a largo plazo con diferentes valores de LSB usando codificación diferencial.

[0056] En algunos ejemplos, las imágenes de referencia que se señalan como imágenes de referencia a corto plazo para una imagen actual (o que se incluyen en el conjunto de imágenes de referencia a corto plazo de la imagen actual) pueden excluirse de las imágenes de referencia utilizadas para decidir si para enviar valores de ciclo delta MSB de POC para imágenes de referencia a largo plazo, para evitar el envío innecesario de valores de ciclo delta MSB de POC. En algunos ejemplos, cuando hay n imágenes de referencia (incluidas o excluidas las imágenes de referencia a corto plazo) en el DPB que tienen un POC LSB igual a $poc_lsb_lt[i]$, y cuando en la imagen actual se señalan n imágenes de referencia a largo plazo que tienen los valores LSB iguales a $poc_lsb_lt[i]$, el codificador de vídeo 20 solo puede solo enviar y el descodificador de vídeo 30 puede solo recibir los valores de ciclo delta MSB de POC para $n - 1$ imágenes de referencia a largo plazo, aunque es posible que estos datos se envíen para todas las n de las imágenes de referencia a largo plazo.

[0057] Cualquiera de las diversas técnicas anteriores puede implementarse de acuerdo con el siguiente ejemplo de implementación. Un elemento sintáctico de WD7 que describe el ciclo de MSB de una imagen de referencia a largo plazo puede modificarse, como se muestra a continuación. Esta divulgación también describe un procedimiento de

ejemplo para obtener la matriz DeltaPocMSBCycleLt []. Esta divulgación describe además un ejemplo de proceso de descodificación para un conjunto de imágenes de referencia. Además, esta divulgación describe la semántica para los elementos sintácticos delta_poc_msb_present_flag[i] y poc_lsb_lt[i]. La sintaxis y la semántica de otros elementos sintácticos de WD7 pueden permanecer sin cambios.

5 [0058] La tabla 1 representa un conjunto de ejemplo de sintaxis para una cabecera de fragmento. El texto subrayado representa cambios en la sintaxis de la cabecera de fragmento en relación con WD7. La semántica para los elementos sintácticos cambiados, así como la semántica cambiada para otros elementos sintácticos, se analizan a continuación.

10 **TABLA 1**

	Descriptor
slice_header() {	
...	
short_term_ref_pic_set_idx	u(v)
if(long_term_ref_pics_present_flag) {	
num_long_term_pics	ue(v)
for (i = 0; i < num_long_term_pics; i++) {	
poc_lsb_lt[i]	u(v)
delta_poc_msb_present_flag[i]	u(1)
if(delta_poc_msb_present_flag[i])	
<u>delta_poc_msb_cycle_lt_plus1[i]</u>	<u>ue(v)</u>
used_by_curr_pic_lt_flag[i]	u(1)
}	
}	
...	
byte_alignment ()	
}	

15 [0059] En el ejemplo de la Tabla 1, la sintaxis de cabecera de fragmento incluye el elemento adicional delta_poc_msb_cycle_lt_plus1[i]. La semántica para este elemento sintáctico se describe a continuación. En algunos ejemplos, poc_lsb_lt[i] puede especificar el valor de 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 del elemento sintáctico poc_lsb_lt[i] puede ser log2_max_pic_order_cnt_lsb_minus4 + 4 bits.

20 [0060] En algunos ejemplos, delta_poc_msb_present_flag[i] igual a 1 pueden especificar que delta_poc_msb_cycle_lt_plus1[i] está presente. delta_poc_msb_present_flag[i] igual a 0 puede especificar que delta_poc_msb_cycle_lt_plus1[i] no está presente. delta_poc_msb_present_flag[i] puede establecerse en 1 cuando hay j imágenes de referencia, donde j es un valor entero mayor que 1, en la memoria intermedia de imágenes descodificadas, excluyendo aquellas imágenes que están marcadas como "usadas para referencia a corto plazo" por la imagen actual, con los bits menos significativos del valor de recuento del orden de imágenes igual a poc_lsb_lt[i], y cuando la i-ésima imagen de referencia a largo plazo no es la j-ésima imagen de referencia a largo plazo con los bits menos significativos del valor de recuento del orden de imágenes igual a poc_lsb_lt[i] que está señalado por la imagen actual.

30 [0061] En algunos ejemplos, delta_poc_msb_cycle_lt_plus1[i] se puede usar para determinar el valor de los bits más significativos del valor de recuento de 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. Cuando no se indica, delta_poc_msb_cycle_lt_plus1[i] puede tomar el valor 1.

35 [0062] La variable DeltaPocMSBCycleLt[i] puede obtenerse de acuerdo con el siguiente algoritmo, donde (7-37) se refiere a una sección de HEVC, es decir, una sección cambiada en relación con WD7. Tenga en cuenta que la siguiente obtención del ejemplo implica que los ciclos de MSB se envían en orden creciente de sus valores.

```

40 DeltaPocMSBCycleLt[i] = delta_poc_msb_present_flag[i] *
    (delta_poc_msb_cycle_lt_plus1[i] - 1);
    if( i != 0) (7-37)
        DeltaPocMSBCycleLt[i] += DeltaPocMSBCycleLt[i - 1];

```

[0063] El valor de $\Delta\text{PocMSBCycleLt}[i] * \text{MaxPicOrderCntLsb} + \text{pic_order_cnt_lsb} - \text{poc_lsb_lt}[i]$ puede estar en el rango de 1 a $2^{24} - 1$, inclusive. De forma alternativa, el ciclo de MSB se puede enviar mediante la señalización de $\text{delta_poc_msb_cycle_lt_plusN}[i]$ como $ue(v)$ (es decir, como un elemento sintáctico con codificación Exp-Golomb entero sin signo con el bit izquierdo primero), en lugar de $\text{delta_poc_msb_cycle_lt_plus1}[i]$, donde N es un valor mayor que 1. De forma alternativa, el ciclo de MSB puede enviarse señalando $\text{delta_poc_msb_cycle_lt}[i]$ como $se(v)$ (es decir, como un elemento sintáctico con codificación Exp-Golomb entero con signo con el bit izquierdo primero), en lugar de $\text{delta_poc_msb_cycle_lt_plus1}[i]$.

[0064] En otras palabras, de acuerdo con las técnicas de esta divulgación, un valor representativo del ciclo de MSB para la imagen i (por ejemplo, $\text{delta_poc_msb_cycle_lt_plus1}[i]$, que en realidad se ha señalado en el flujo de bits) puede calcularse en relación con el ciclo de MSB de una imagen de referencia a largo plazo previamente codificada, por ejemplo, la imagen i-1, en lugar de en relación con el valor de MSB base. Esto es matemáticamente equivalente al cálculo del valor representativo del ciclo de MSB para la imagen i como la diferencia entre los MSB del valor de POC para la imagen i y los MSB del valor de POC para la imagen i-1, suponiendo que este valor se agrega finalmente al valor del ciclo de MSB para la imagen i-1 para representar el verdadero valor del ciclo de MSB para la imagen i (que es la diferencia entre los MSB para la imagen i y los MSB base). Esto puede lograr ahorros de bits en un flujo de bits, porque los valores que representan las diferencias entre los valores de ciclo de MSB secuenciales pueden ser más pequeños que las diferencias entre los MSB de imágenes de referencia a largo plazo y los MSB base.

[0065] En algunos ejemplos, el codificador de vídeo 20 y el decodificador de vídeo 30 pueden realizar un proceso de obtención para el conjunto de imágenes de referencia y el marcado de imágenes de acuerdo con los siguientes pasos ordenados, donde DPB se refiere a la memoria intermedia de imágenes decodificadas, como se describe en el Anexo C del documento WD7. Tenga en cuenta que el orden de imágenes de referencia a corto plazo y las obtenciones del subconjunto de imagen de referencia a largo plazo en el siguiente ejemplo se intercambian en relación con el orden en WD7. Las referencias parentéticas cerca del margen derecho están destinadas a referirse a las secciones de WD7. Este y otros cambios relacionados con el algoritmo de WD7 se anotan usando el texto subrayado en el siguiente ejemplo:

1. Se aplica lo siguiente:

```

for(i = 0; i < NumPocStCurrBefore; i++)
    if(there is a short-term reference picture picX in the DPB
        with PicOrderCntVal equal to PocStCurrBefore[i])
35         RefPicSetStCurrBefore[i] = picX
    else
        RefPicSetStCurrBefore[i] = "no reference picture"
for(i = 0; i < NumPocStCurrAfter; i++)
    if(there is a short-term reference picture picX in the DPB
        with PicOrderCntVal equal to PocStCurrAfter[i])
40         RefPicSetStCurrAfter[i] = picX
    else
        RefPicSetStCurrAfter[i] = "no reference picture" (8-7)
for(i = 0; i < NumPocStFoll; i++)
45         if(there is a short-term reference picture picX in the DPB
            with PicOrderCntVal equal to PocStFoll[i])
                RefPicSetStFoll[i] = picX
        else
            RefPicSetStFoll[i] = "no reference picture"
50

```

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(!delta_poc_msb_present_flag[i]) {
        if(there is a long-term reference picture picX in the DPB
            with pic_order_nt_lsb equal to PocLtCurr[i]
            and is not included in any of RefPicSetStCurrBefore,
60             RefPicSetStCurrAfter and RefPicSetStFoll)
                RefPicSetLtCurr[i] = picX
        else if(there is a short-term reference picture picY in the DPB
            with pic_order_cnt_lsb equal to PocLtCurr[i]
            and is not included in any of RefPicSetStCurrBefore,
65             RefPicSetStCurrAfter and RefPicSetStFoll)

```

```

    RefPicSetLtCurr[i] = picY
  else
    RefPicSetLtCurr[i] = "no reference picture"
} else {
5   if(there is a long-term reference picture picX in the DPB
      with PicOrderCntVal equal to PoclCurr[i]
      and is not included in any of RefPicSetStCurrBefore,
      RefPicSetStCurrAfter and RefPicSetStFoll)
      RefPicSetLtCurr[i] = picX
10  else if(there is a short-term reference picture picY in the DPB
      with PicOrderCntVal equal to PoclCurr[i]
      and is not included in any of RefPicSetStCurrBefore,
      RefPicSetStCurrAfter and RefPicSetStFoll)
      RefPicSetLtCurr[i] = picY
15  else
      RefPicSetLtCurr[i] = "no reference picture"
}
}(8-6)
for(i = 0; i < NumPoclFoll; i++) {
20  if(!delta_poc_msb_present_flag[i]){
      if(there is a long-term reference picture picX in the DPB
      with pic_order_cnt_lsb equal to PoclFoll[i]
      and is not included in any of RefPicSetStCurrBefore,
      RefPicSetStCurrAfter and RefPicSetStFoll)
25  RefPicSetLtFoll[i] = picX
      else if(there is a short-term reference picture picY in the DPB
      with pic_order_cnt_lsb equal to PoclFoll[i]
      and is not included in any of RefPicSetStCurrBefore,
      RefPicSetStCurrAfter and RefPicSetStFoll)
30  RefPicSetLtFoll[i] = picY
      else
      RefPicSetLtFoll[i] = "no reference picture"
} else {
35  if(there is a long-term reference picture picX in the DPB
      with PicOrderCntVal to PoclFoll[i]
      and is not included in any of RefPicSetStCurrBefore,
      RefPicSetStCurrAfter and RefPicSetStFoll)
      RefPicSetLtFoll[i] = picX
40  else if(there is a short-term reference picture picY in the DPB
      with PicOrderCntVal equal to PoclFoll[i]
      and is not included in any of RefPicSetStCurrBefore,
      RefPicSetStCurrAfter and RefPicSetStFoll)
      RefPicSetLtFoll[i] = picY
45  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. 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".

[0066] 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. Las entradas en RefPicSetStFoll o RefPicSetLtFoll que son iguales a "sin imagen de referencia" deberían ignorarse. A menos que cualquiera de las dos condiciones siguientes sea verdadera, se debe deducir una pérdida de imagen involuntaria para cada entrada en RefPicSetStCurrBefore, RefPicSetStCurrAfter y RefPicSetLtCurr que sea igual a "sin imagen de referencia": a) la primera imagen codificada en el flujo de bits es una imagen de CRA y la imagen codificada actual es una imagen TFD asociada a la primera imagen codificada en el flujo de bits; b) la imagen de RAP anterior que precede a la imagen codificada actual en orden de descodificación es una imagen de BLA y la imagen codificada actual es una imagen TFD asociada a la imagen de BLA.

65

[0067] Así, de acuerdo con las técnicas de la presente divulgación, un codificador de vídeo (por ejemplo, el codificador de vídeo 20 y/o el descodificador de vídeo 30) puede calcular los valores de ciclo de MSB para las imágenes que tienen diferentes valores LSB utilizando codificación diferencial. Además, el codificador de vídeo puede codificar una imagen actual en relación con una de las imágenes de referencia a largo plazo, por ejemplo, la imagen $i-1$ o la imagen i , usando los valores representativos de los ciclos MSB. Por ejemplo, el codificador de vídeo puede determinar los valores de POC para imágenes de referencia usando la fórmula descrita anteriormente (fórmula 7-37) y/o el algoritmo descrito anteriormente.

[0068] El codificador de vídeo 20 y el descodificador de vídeo 30, por lo tanto, representan ejemplos de un codificador de vídeo configurado para codificar un primer valor representativo de una diferencia entre un valor base de bits más significativos (MSB) de un valor de recuento de orden de imágenes (POC) de una imagen actual de datos de vídeo y un primer valor de MSB de un primer valor de POC de una primera imagen de referencia a largo plazo de los datos de vídeo, codificar un segundo valor representativo de una diferencia entre un segundo valor de MSB de un segundo valor de POC de una segunda imagen de referencia a largo plazo de los datos de vídeo y el primer valor de MSB, en el que el primer valor de POC y el segundo valor de POC tienen diferentes valores de bits menos significativos (LSB), y codifican al menos una parte de una imagen actual de los datos de vídeo en relación con al menos una de la primera imagen de referencia a largo plazo utilizando el primer valor y la segunda imagen de referencia a largo plazo utilizando el primer valor y el segundo valor.

[0069] El codificador de vídeo 20, por ejemplo, puede determinar si hay dos o más imágenes de referencia a largo plazo, para una imagen actual, para las cuales se deben señalar los valores de ciclo de MSB (también denominados valores de MSB), incluso cuando las imágenes de referencia a largo plazo tienen diferentes valores de LSB. En el caso de que haya dos (o más) imágenes de referencia a largo plazo con diferentes valores de LSB, pero para las cuales se señalarán los valores de ciclo de MSB, el codificador de vídeo 20 puede codificar datos representativos del valor de ciclo de MSB para la segunda imagen de referencia como una diferencia (o delta) relativa al valor del ciclo de MSB para la primera imagen de referencia a largo plazo. De manera similar, el codificador de vídeo 20 puede codificar datos representativos del valor del ciclo de MSB para la primera imagen de referencia a largo plazo como una diferencia con respecto al valor de MSB base de un valor de POC para la imagen actual. El codificador de vídeo 20 puede codificar además al menos una parte de la imagen actual (por ejemplo, uno o más bloques) en relación con al menos una de la primera imagen de referencia a largo plazo y la segunda imagen de referencia a largo plazo.

[0070] Por lo tanto, cuando el descodificador de vídeo 30 recibe los datos indicativos de los valores de ciclo de MSB para imágenes de referencia a largo plazo, el descodificador de vídeo 30 puede descodificar un valor representativo de la diferencia entre los MSB base de un valor de POC para la primera imagen de referencia a largo plazo y calcular los MSB del valor de POC para la primera imagen de referencia a largo plazo agregando el valor a los MSB base. El descodificador de vídeo 30 también puede descodificar un valor representativo de la diferencia entre los MSB de un valor de POC para la segunda imagen de referencia a largo plazo y los MSB del valor de POC para la primera imagen de referencia a largo plazo, y calcular los MSB del valor de POC para la segunda imagen de referencia a largo plazo agregando el valor a los MSB del valor de POC para la primera imagen de referencia a largo plazo. Del mismo modo, el descodificador de vídeo 30 puede descodificar al menos una parte de la imagen actual en relación con al menos una de la primera imagen de referencia a largo plazo y la segunda imagen de referencia a largo plazo, utilizando los MSB de los valores de POC para las imágenes de referencia a largo plazo.

[0071] El codificador de vídeo 20 y el descodificador de vídeo 30 se pueden implementar, cada uno, como cualquiera de una variedad de circuitos de codificadores o descodificadores adecuados, según corresponda, tales como uno o más microprocesadores, procesadores de señales digitales (DSP), circuitos integrados específicos de la aplicación (ASIC), matrices de puertas programables *in situ* (FPGA), circuitos de lógica discreta, software, hardware, firmware o cualquier combinación de los mismos. 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 dispositivo que incluye el codificador de vídeo 20 y/o el descodificador de vídeo 30 puede comprender un circuito integrado, un microprocesador y/o un dispositivo de comunicación inalámbrica, tal como un teléfono celular.

[0072] La FIG. 2 es un diagrama de bloques que ilustra un ejemplo de codificador de vídeo 20 que puede implementar técnicas para señalización mejorada de imágenes de referencia a largo plazo en cabeceras de fragmento. El codificador de vídeo 20 puede realizar la intracodificación y la intercodificación de bloques de vídeo dentro de fragmentos de vídeo. La intracodificación se basa en la predicción espacial para reducir o eliminar la redundancia espacial en el vídeo dentro de una trama o imagen de vídeo dada. La intercodificación se basa en la predicción temporal para reducir o eliminar la redundancia temporal en el vídeo dentro de tramas o imágenes adyacentes de una secuencia de vídeo. El modo intra (modo I) puede referirse a cualquiera de varios modos de codificación de base espacial. Los modos inter, tales como la predicción unidireccional (modo P) o la bipredicción (modo B), pueden referirse a cualquiera de varios modos de codificación de base temporal.

[0073] Como se muestra en la FIG. 2, el codificador de vídeo 20 recibe un bloque de vídeo actual dentro de una trama de vídeo para ser codificada. En el ejemplo de la figura 2, el codificador de vídeo 20 incluye una unidad de selección de modo 40, una memoria de imágenes de referencia 64, un sumador 50, una unidad de procesamiento de

transformada 52, una unidad de cuantificación 54 y una unidad de codificación por entropía 56. A su vez, la unidad de selección de modo 40 incluye la unidad de compensación de movimiento 44, la unidad de estimación de movimiento 42, la unidad de predicción intra 46 y la unidad de división 48. Para la reconstrucción de bloques de vídeo, el codificador de vídeo 20 incluye también la unidad de cuantificación inversa 58, la unidad de transformada inversa 60 y el sumador 62. También se puede incluir un filtro de eliminación de bloques (no mostrado en la figura 2) para filtrar fronteras de bloques, para eliminar distorsiones de efecto pixelado del vídeo reconstruido. Si se desea, el filtro de eliminación de bloques filtrará típicamente la salida del sumador 62. También pueden usarse filtros adicionales (en el bucle o tras el bucle), además del filtro de eliminación de bloques. Dichos filtros no se muestran por razones de brevedad pero, si se desea, pueden filtrar la salida del sumador 50 (como un filtro en el bucle).

[0074] Durante el proceso de codificación, el codificador de vídeo 20 recibe una trama o un fragmento de vídeo que va a codificarse. La trama o el fragmento pueden dividirse en múltiples bloques de vídeo. La unidad 42 de estimación de movimiento y la unidad 44 de compensación de movimiento llevan a cabo la codificación de predicción intra del bloque de vídeo recibido con respecto a uno o más bloques en una o más tramas de referencia, para proporcionar la predicción temporal. La unidad 46 de predicción intra, de forma alternativa, puede llevar a cabo la codificación con predicción intra del bloque de vídeo recibido, con respecto a uno o más bloques contiguos en la misma trama o fragmento que el bloque a codificar, para proporcionar predicción espacial. El codificador de vídeo 20 puede llevar a cabo múltiples pases de codificación, por ejemplo, para seleccionar un modo de codificación adecuada para cada bloque de datos de vídeo.

[0075] Además, la unidad de división 48 puede dividir bloques de datos de vídeo en sub-bloques, basándose en la evaluación de los anteriores esquemas de división en las pasadas de codificación anteriores. Por ejemplo, la unidad de partición 48 puede dividir inicialmente una trama o un fragmento en varias LCU, y dividir cada una de las LCU en varias subCU, basándose en un análisis de velocidad-distorsión (por ejemplo, una optimización de velocidad-distorsión). La unidad de selección de modo 40 puede producir además una estructura de datos de árbol cuádruple, indicativa de la división de una LCU en las sub-CU. Las CU de nodos hoja del árbol cuádruple pueden incluir una o más PU y una o más TU.

[0076] La unidad de selección de modo 40 puede seleccionar uno de los modos de codificación, intra o inter, por ejemplo, basándose en los resultados de errores, y proporciona el bloque intra-codificado o inter-codificado resultante al sumador 50 para generar datos de bloque residuales y al sumador 62 para reconstruir el bloque codificado para su uso como una trama de referencia. La unidad de selección de modo 40 proporciona también elementos sintácticos, tales como vectores de movimiento, indicadores de modo intra, información de división y otra información sintáctica de este tipo, a la unidad de codificación por entropía 56. De acuerdo con las técnicas de esta divulgación, la unidad de selección de modo 40 puede proporcionar información de identificación de imagen de referencia, por ejemplo, si la imagen de referencia debe tratarse como imágenes de referencia a largo plazo o imágenes de referencia a corto plazo, así como los bits menos significativos (LSB) de los valores de recuento de orden de imágenes (POC) para imágenes de referencia a largo plazo, a la unidad de codificación por entropía 56. La unidad de codificación por entropía 56, u otra unidad de codificador de vídeo 20, puede incluir dicha información de identificación de imagen de referencia en una cabecera de fragmento de un fragmento, de acuerdo con las técnicas de esta divulgación. De forma alternativa, la unidad de codificación por entropía 56 puede incluir la información de identificación de la imagen de referencia (por ejemplo, para indicar qué imágenes de referencia son imágenes de referencia a largo plazo) en un conjunto de parámetros, como un conjunto de parámetros de imagen (PPS) o un conjunto de parámetros de secuencia (SPS).

[0077] Como se explicó anteriormente, las imágenes de referencia a largo plazo pueden señalarse usando LSB de valores de POC para las imágenes de referencia a largo plazo. Por ejemplo, la unidad de codificación por entropía 56 puede indicar que una imagen de referencia es una imagen de referencia a largo plazo al señalar los LSB del valor de POC para la imagen de referencia en una cabecera de fragmento, un PPS o un SPS. Otras imágenes de referencia que no se señalan de esta manera se pueden tratar como imágenes de referencia a corto plazo.

[0078] Al señalar información de identificación para imágenes de referencia a largo plazo, la unidad de codificación por entropía 56 proporciona datos que indican qué imágenes de referencia deben almacenarse en una memoria intermedia de imágenes descodificadas (DPB) de un descodificador de vídeo (por ejemplo, descodificador de vídeo 30) durante un período de tiempo relativamente largo, y qué imágenes de referencia se pueden eliminar de la memoria intermedia de imágenes descodificadas anteriormente. En general, las imágenes de referencia a largo plazo pueden usarse como imágenes de referencia para codificar imágenes que tienen tiempos de orden de descodificación relativamente tardíos. Por lo tanto, la señalización de la información de identificación para las imágenes de referencia a largo plazo proporciona información para un descodificador de vídeo, por ejemplo, el descodificador de vídeo 30, de manera que el descodificador de vídeo puede determinar si una imagen de referencia puede ser descartada de su propio DPB.

[0079] La unidad de estimación de movimiento 42 y la unidad de compensación de movimiento 44 pueden estar altamente integradas, pero se ilustran por separado con fines conceptuales. La estimación del movimiento, realizada por la unidad de estimación de movimiento 42, es el proceso de generación de vectores de movimiento, que estiman el movimiento para los bloques de vídeo. Un vector de movimiento, por ejemplo, puede indicar el desplazamiento de una PU de un bloque de vídeo dentro de una trama o imagen de vídeo actual, con respecto a un bloque predictivo

dentro de una trama de referencia (u otra unidad codificada), con respecto al bloque actual que se está codificando dentro de la trama actual (u otra unidad codificada). Un bloque predictivo es un bloque que se revela como estrechamente coincidente con el bloque a codificar, en lo que respecta a la diferencia de píxeles, lo cual puede determinarse mediante la suma de diferencias absolutas (SAD), suma de las diferencias al cuadrado (SSD) u otras métricas de diferencia. En algunos ejemplos, el codificador de vídeo 20 puede calcular valores para posiciones fraccionarias de píxeles de imágenes de referencia almacenadas en la memoria de imágenes de referencia 64. Por ejemplo, el codificador de vídeo 20 puede interpolar valores de posiciones de un cuarto de píxel, posiciones de un octavo de píxel u otras posiciones fraccionarias de píxel de la imagen de referencia. Por lo tanto, la unidad de estimación de movimiento 42 puede realizar una búsqueda de movimiento relativa a las posiciones de píxel completo y las posiciones de píxel fraccionario, y emitir un vector de movimiento con una precisión de píxel fraccionario.

[0080] La unidad de estimación de movimiento 42 calcula un vector de movimiento para una PU de un bloque de vídeo en un fragmento inter-codificado, comparando la posición de la PU con la posición de un bloque predictivo de una imagen de referencia. La imagen de referencia puede seleccionarse a partir de una primera lista (lista 0) de imágenes de referencia o una segunda lista (lista 1) de imágenes de referencia, cada una de las cuales identifica una o más imágenes de referencia almacenadas en una memoria de imágenes de referencia 64. La memoria de imagen de referencia 64 puede almacenar imágenes de referencia tanto a largo como a corto plazo. La unidad de estimación de movimiento 42 envía el vector de movimiento calculado a la unidad de codificación por entropía 56 y a la unidad de compensación de movimiento 44.

[0081] La compensación de movimiento, llevada a cabo por la unidad de compensación de movimiento 44, puede implicar capturar o generar el bloque predictivo basándose en el vector de movimiento determinado por la unidad de estimación de movimiento 42. De nuevo, la unidad de estimación de movimiento 42 y la unidad de compensación de movimiento 44 pueden integrarse funcionalmente, en algunos ejemplos. Tras recibir el vector de movimiento para la PU del bloque de vídeo actual, la unidad de compensación de movimiento 44 puede localizar el bloque predictivo al que apunta el vector de movimiento en una de las listas de imágenes de referencia. El sumador 50 forma un bloque de vídeo residual restando los valores de píxel del bloque predictivo a los valores de píxel del bloque de vídeo actual que se está codificando, generando valores de diferencias de píxel, como se analiza posteriormente. En general, la unidad de estimación de movimiento 42 lleva a cabo la estimación de movimiento con respecto a los componentes de luma, y la unidad de compensación de movimiento 44 utiliza los vectores de movimiento calculados basándose en los componentes de luma, tanto para los componentes de croma como para los componentes de luma.

[0082] La unidad de selección de modo 40 también puede generar elementos sintácticos asociados a los bloques de vídeo y el fragmento de vídeo para su uso mediante el descodificador de vídeo 30 en la descodificación de los bloques de vídeo del fragmento de vídeo. Por ejemplo, después de calcular un vector de movimiento para un bloque de datos de vídeo (donde el vector de movimiento apunta a un bloque de referencia en una imagen de referencia), la unidad de selección de modo 40 puede generar valores para los parámetros de movimiento, como una lista de imágenes de referencia en la que la imagen de referencia se almacena y un índice en la lista de imágenes de referencia que corresponde a la posición de la imagen de referencia en la lista de imágenes de referencia. Estos y otros parámetros de movimiento pueden codificarse basándose en un modo de codificación de vectores de movimiento. Por ejemplo, al usar la predicción avanzada de vectores de movimiento (AMVP), la unidad de codificación por entropía 56 puede identificar un predictor de vectores de movimiento (correspondiente a un bloque contiguo espacial o temporal al bloque actual) y señalar explícitamente la lista de imágenes de referencia, el índice de imágenes de referencia y los valores de diferencia de vectores de movimiento horizontales y verticales. Como otro ejemplo, usando el modo de combinación, la unidad de codificación por entropía 56 puede identificar un predictor de vector de movimiento usando un índice de combinación, y la lista de imágenes de referencia y el índice de referencia pueden heredarse del predictor de vector de movimiento (por ejemplo, la información de movimiento de un bloque contiguo, ya sea un bloque contiguo espacial o un bloque contiguo temporal).

[0083] El codificador de vídeo 20 señala información indicativa de qué imágenes de referencia son imágenes de referencia a largo plazo, tanto para indicar que estas imágenes no se deben descartar del DPB como para proporcionar información que se pueden usar durante un proceso de construcción de la lista de imágenes de referencia. De esta manera, un descodificador de vídeo, tal como el descodificador de vídeo 30, puede ser capaz de reproducir con precisión la lista de imágenes de referencia, de modo que el índice de referencia utilizado como información de movimiento para un bloque con predicción inter se refiera con precisión a la imagen de referencia correcta.

[0084] Como se señaló anteriormente, las imágenes de referencia a largo plazo pueden señalarse usando valores de POC de las imágenes de referencia a largo plazo. Los valores de POC en general corresponden a datos indicativos del orden de visualización (también conocido como orden de salida) de las imágenes, que no es necesariamente el mismo orden de descodificación (también conocido como orden de flujo de bits). Los valores de POC para imágenes de referencia a largo plazo pueden señalarse utilizando solo los bits menos significativos (LSB). Los datos representativos de MSB pueden señalarse ocasionalmente, por ejemplo, cuando dos o más imágenes de referencia a largo plazo tienen los mismos LSB. En algunos casos, el codificador de vídeo 20 puede determinar que los datos representativos de los MSB para dos (o más) imágenes deben ser señalados, incluso cuando al menos dos de las imágenes para las que se señalan los datos de los MSB tienen diferentes LSB.

[0085] De acuerdo con las técnicas de esta divulgación, el codificador de vídeo 20 puede ser señalizar los datos representativos de los MSB como una diferencia, por ejemplo, la diferencia entre un conjunto de MSB y otro conjunto de MSB. Por ejemplo, como se analizó anteriormente con respecto a la Tabla 1, el codificador de vídeo 20 puede calcular la diferencia entre un valor de MSB base y el valor de MSB para una imagen de referencia a largo plazo. La unidad de codificación por entropía 56 puede señalar esta diferencia como un valor representativo de los MSB para la imagen de referencia a largo plazo. El valor de MSB base puede corresponder a los MSB del valor de POC de una imagen actual. De forma alternativa, el valor de MSB base puede corresponder a los MSB del valor de POC de otra imagen, como una imagen de actualización de descodificador instantánea (IDR).

[0086] Después de codificar el valor de diferencia para una imagen de referencia a largo plazo, la unidad de codificación por entropía 56 puede calcular la diferencia entre MSB de valor de POC para otra imagen de referencia a largo plazo y el MSB del valor de POC para la imagen de referencia a largo plazo anterior. Del mismo modo, la unidad de codificación por entropía 56 puede señalar este segundo valor de diferencia para la imagen de referencia a largo plazo subsiguiente, de modo que un descodificador de vídeo, como el descodificador de vídeo 30, puede reconstruir los MSB del valor de POC para la imagen de referencia a largo plazo subsiguiente. De esta manera, el codificador de vídeo 20 puede configurarse para codificar datos representativos de MSB de valores de POC para imágenes de referencia a largo plazo como valores de diferencia, por ejemplo, relativos a un valor de MSB base o relativos a MSB codificados previamente para valores de POC de otras imágenes de referencia a largo plazo. Además, el codificador de vídeo 20 puede señalar estos valores de diferencia incluso cuando los LSB de los valores de POC son diferentes.

[0087] Como se analizó anteriormente, en algunos ejemplos, el codificador de vídeo 20 puede estar configurado para codificar valores enteros negativos para valores de diferencia que representan diferencias entre los MSB de valores de POC para una o más imágenes de referencia a largo plazo y un valor de MSB base (u otros MSB de referencia). Por ejemplo, cuando la unidad de codificación por entropía 56 (u otra unidad del codificador de vídeo 20) determina que una imagen actual tiene un valor de POC igual a $\text{MaxPicOrderCntLsb} * N - 1$, y que una imagen que tiene un valor de POC igual a $\text{MaxPicOrderCntLsb} * N$ es el primer LTRP para el cual se señala el `delta_poc_msb_cycle_it[i]`, la unidad de codificación por entropía 56 puede codificar un valor de -1 para los MSB del valor de POC del LTRP.

[0088] Además, o de forma alternativa, el codificador de vídeo 20 puede estar configurado para codificar datos representativos de valores de POC para imágenes de referencia a largo plazo basadas al menos en parte en un número total de imágenes de referencia almacenadas actualmente en una imagen de referencia establecida (o una memoria intermedia de imágenes descodificadas) y/o un número de imágenes de referencia a corto plazo actualmente almacenadas en el conjunto de imágenes de referencia (o la memoria intermedia de imágenes descodificadas). Por ejemplo, el codificador de vídeo 20 puede configurarse para determinar si debe señalar un valor de ciclo de MSB de POC para una imagen de referencia a largo plazo basándose en si hay otras imágenes de referencia a largo plazo en el conjunto de imágenes de referencia (o memoria intermedia de imágenes descodificadas) que tienen el mismo valor de POC LSBs que la imagen de referencia a largo plazo. Por lo tanto, cuando solo hay una o más imágenes de referencia a corto plazo en el conjunto de imágenes de referencia que tienen los mismos POC LSB que la imagen de referencia a largo plazo, el codificador de vídeo 20 puede evitar la codificación de un valor de ciclo de MSB de POC para la imagen de referencia a largo plazo.

[0089] Además, o de forma alternativa, el codificador de vídeo 20 puede determinar no codificar un valor del ciclo de MSB de POC delta para una imagen de referencia a largo plazo cuando hay N imágenes de referencia a largo plazo en la memoria intermedia de imágenes descodificadas que tienen los mismos LSB que el valor de POC para la imagen de referencia a largo plazo y cuando el codificador de vídeo 20 ya ha codificado los valores del ciclo de MSB de POC delta para N-1 de las imágenes de referencia a largo plazo.

[0090] La unidad de predicción intra 46 puede realizar predicción intra un bloque actual, de forma alternativa a la predicción inter realizada por la unidad de estimación del movimiento 42 y la unidad de compensación del movimiento 44, como se describe anteriormente. En particular, la unidad de predicción intra 46 puede determinar un modo de predicción intra a usar para codificar un bloque actual. En algunos ejemplos, la unidad de predicción intra 46 puede codificar un bloque actual usando diversos modos de predicción intra, por ejemplo, durante pases de codificación independientes, y la unidad de predicción intra 46 (o la unidad de selección de modo 40, en algunos ejemplos) puede seleccionar un modo de predicción intra adecuada para usar, a partir de los modos probados.

[0091] Por ejemplo, la unidad de predicción intra 46 puede calcular valores de distorsión de velocidad usando un análisis de distorsión de velocidad para los diversos modos de predicción intra probados, y seleccionar el modo de predicción intra que tenga las mejores características de distorsión de velocidad entre los modos probados. El análisis de velocidad-distorsión determina, en general, una magnitud de distorsión (o de errores) entre un bloque codificado y un bloque original, no codificado, que se codificó para producir el bloque codificado, así como una velocidad de transmisión de bits (es decir, un número de bits) usada para producir el bloque codificado. La unidad de predicción intra 46 puede calcular proporciones a partir de las distorsiones y velocidades para los diversos bloques codificados, para determinar qué modo de predicción intra presenta el mejor valor de velocidad-distorsión para el bloque.

[0092] Después de seleccionar un modo de predicción intra para un bloque, la unidad de predicción intra 46 puede proporcionar información indicativa del modo de predicción intra seleccionado para el bloque a la unidad de

codificación por entropía 56. La unidad de codificación por entropía 56 puede codificar la información que indica el modo de predicción intra seleccionada. El codificador de vídeo 20 puede incluir, en el flujo de bits transmitido, datos de configuración, que pueden incluir una pluralidad de tablas de índices de modo de predicción intra y una pluralidad de tablas modificadas de índices de modo de predicción intra (también denominadas tablas de correlación de palabras de código), definiciones de contextos de codificación para diversos bloques e indicaciones de un modo de predicción intra más probable, una tabla de índices de modo de predicción intra y una tabla modificada de índices de modo de predicción intra, a usar para cada uno de los contextos.

[0093] El codificador de vídeo 20 forma un bloque de vídeo residual restando los datos de predicción de la unidad de selección de modo 40 del bloque de vídeo original que se está codificado. El sumador 50 representa el componente o los componentes que realizan esta operación de resta. La unidad de procesamiento de transformada 52 aplica una transformada, tal como una transformada discreta de coseno (DCT) o una transformada conceptualmente similar, al bloque residual, produciendo un bloque de vídeo que comprende valores residuales de coeficientes de transformada. La unidad de procesamiento de transformada 52 puede llevar a cabo otras transformadas que son conceptualmente similares a la DCT. También se podrían usar transformadas de ondículas, transformadas de números enteros, transformadas de subbandas u otros tipos de transformadas. En cualquier caso, la unidad de procesamiento de transformada 52 aplica la transformada al bloque residual, produciendo un bloque de coeficientes de transformada residuales. La transformada puede convertir la información residual, desde un dominio de valores de píxel a un dominio de transformada, tal como un dominio de frecuencia. La unidad de procesamiento de transformada 52 puede enviar los coeficientes de transformada resultantes a la unidad de cuantificación 54. La unidad de cuantificación 54 cuantifica los coeficientes de transformada para reducir más la velocidad de bits. El proceso de cuantificación puede reducir la profundidad de bits asociada a algunos, o a la totalidad, de los coeficientes. El grado de cuantificación se puede modificar ajustando un parámetro de cuantificación. En algunos ejemplos, la unidad 54 de cuantificación puede realizar, a continuación, una exploración de la matriz que incluye los coeficientes de transformada cuantificados. De forma alternativa, la unidad de codificación por entropía 56 puede realizar la exploración.

[0094] Tras la cuantificación, la unidad de codificación por entropía 56 codifica por entropía los coeficientes de transformada cuantificados. Por ejemplo, la unidad de codificación por entropía 56 puede llevar a cabo la codificación de longitud variable adaptativa según el contexto (CAVLC), la codificación aritmética binaria adaptativa según el contexto (CABAC), la codificación aritmética binaria adaptativa según el contexto y basada en sintaxis (SBAC), la codificación por entropía mediante la división en intervalos de probabilidades (PIPE) u otra técnica de codificación por entropía. En el caso de la codificación por entropía basada en el contexto, el contexto se puede basar en bloques contiguos. Tras la codificación por entropía por parte de la unidad de codificación por entropía 56, el flujo de bits codificado se puede transmitir a otro dispositivo (por ejemplo, el descodificador de vídeo 30) o archivar para su posterior transmisión o recuperación.

[0095] La unidad de cuantificación inversa 58 y la unidad de transformada inversa 60 aplican la cuantificación inversa y la transformación inversa, respectivamente, para reconstruir el bloque residual en el dominio de píxeles, por ejemplo, para su uso posterior como un bloque de referencia. La unidad 44 de compensación de movimiento puede calcular un bloque de referencia añadiendo el bloque residual a un bloque predictivo de una de las tramas de la memoria 64 de imágenes de referencia. La unidad de compensación de movimiento 44 también puede aplicar uno o más filtros de interpolación al bloque residual reconstruido para calcular valores fraccionarios de píxel, para su uso en la estimación de movimiento. El sumador 62 añade el bloque residual reconstruido al bloque de predicción compensado por movimiento, producido por la unidad de compensación de movimiento 44, para producir un bloque de vídeo reconstruido para su almacenamiento en la memoria de imágenes de referencia 64. El bloque de vídeo reconstruido puede ser utilizado por la unidad de estimación de movimiento 42 y la unidad de compensación de movimiento 44, como bloque de referencia para intercodificar un bloque en una trama de vídeo posterior.

[0096] El codificador de vídeo 20 de la FIG. 2 representa un ejemplo de un codificador de vídeo que se puede configurar para codificar datos para una imagen actual de datos de vídeo, en el que los datos codificados representan un valor entero negativo para un ciclo de bits más significativos (MSB) del recuento de orden de imágenes (POC) durante una imagen de referencia a largo plazo para la imagen actual, y codifica al menos una parte de la imagen actual en relación con la imagen de referencia a largo plazo basada, al menos en parte, en el valor entero negativo para el ciclo de MSB.

[0097] El codificador de vídeo 20 de la FIG. 2 también representa un ejemplo de un codificador de vídeo que puede configurarse, de forma adicional o alternativa, para codificar al menos una primera parte de una imagen actual de datos de vídeo en relación con una primera imagen de referencia a largo plazo y al menos una segunda parte de la imagen actual en una segunda imagen de referencia a largo plazo, codificar un primer valor representativo de una diferencia entre un valor base de bits más significativos (MSB) de un valor de recuento de orden de imágenes (POC) de una imagen actual de los datos de vídeo y un primer valor de MSB de un primer valor de POC de la primera imagen de referencia a largo plazo de los datos de vídeo, y codificar un segundo valor representativo de una diferencia entre un segundo valor de MSB de un segundo valor de POC de la segunda imagen de referencia a largo plazo de los datos de vídeo y el primer valor de MSB, en el que el primer valor de POC y el segundo valor de POC tienen diferentes valores de bits menos significativos (LSB).

[0098] El codificador de vídeo 20 de la FIG. 2 también representa un ejemplo de un codificador de vídeo que puede configurarse, de forma adicional o alternativa, para determinar un número de imágenes de referencia marcadas como imágenes de referencia a corto plazo, de un número total de imágenes de referencia disponibles, para una imagen actual, codificar un valor de bits más significativos (MSB) para imágenes de referencia a largo plazo para la imagen actual basándose en el número total de imágenes de referencia y el número de imágenes de referencia marcadas como imágenes de referencia a corto plazo, y codificar al menos una parte de la imagen actual en referencia al menos a una de las imágenes de referencia disponibles.

[0099] El codificador de vídeo 20 de la FIG. 2 también representa un ejemplo de un codificador de vídeo que puede configurarse, de forma adicional o alternativa, para determinar un número N de imágenes de referencia para una imagen actual que tenga valores de recuento de orden de imágenes (POC) con valores de bit (LSB) menos significativos iguales, determinar que el conjunto de imágenes de referencia de la imagen actual puede incluir N imágenes de referencia a largo plazo, y codificar delta valores de ciclo de bits más significativos (MSB) de POC para, como máximo, N-1 de las imágenes de referencia a largo plazo basadas en las determinaciones.

[0100] La FIG. 3 es un diagrama de bloques que ilustra un ejemplo de descodificador de vídeo 30 que puede implementar técnicas para señalización mejorada de imágenes de referencia a largo plazo en cabeceras de fragmento. En el ejemplo de la figura 3, el descodificador de vídeo 30 incluye una unidad de descodificación por entropía 70, una unidad de compensación de movimiento 72, una unidad de predicción intra 74, una unidad de cuantificación inversa 76, una unidad de transformación inversa 78, una memoria de imágenes de referencia 82 y un sumador 80. En algunos ejemplos, el descodificador de vídeo 30 puede realizar un pase de descodificación, en general recíproco al pase de codificación descrito con respecto al codificador de vídeo 20 (figura 2). La unidad de compensación de movimiento 72 puede generar datos de predicción basándose en vectores de movimiento recibidos desde la unidad de descodificación por entropía 70, mientras que la unidad de predicción intra 74 puede generar datos de predicción basándose en indicadores de modo de predicción intra, recibidos desde la unidad de descodificación por entropía 70.

[0101] Durante el proceso de descodificación, el descodificador de vídeo 30 recibe un flujo de bits de vídeo codificado, que representa los bloques de vídeo de un fragmento de vídeo codificado y elementos sintácticos asociados, desde el codificador de vídeo 20. La unidad de descodificación por entropía 70 del descodificador de vídeo 30 descodifica por entropía el flujo de bits para generar coeficientes cuantificados, vectores de movimiento o indicadores de modo de predicción intra y otros elementos sintácticos. La unidad de descodificación por entropía 70 remite los vectores de movimiento y otros elementos sintácticos a la unidad de compensación de movimiento 72. El descodificador de vídeo 30 puede recibir los elementos sintácticos en el nivel de fragmento de vídeo y/o el nivel de bloque de vídeo.

[0102] Cuando el fragmento de vídeo se codifica como un fragmento intra-codificado (I), la unidad de predicción intra 74 puede generar datos de predicción para un bloque de vídeo del fragmento de vídeo actual, basándose en un modo de predicción intra señalado, y datos de bloques previamente descodificados de la trama o imagen actual. Cuando la trama de vídeo se codifica como un fragmento inter-codificado (es decir, B, P o GPB), la unidad de compensación de movimiento 72 genera bloques predictivos para un bloque de vídeo del fragmento de vídeo actual basándose en los vectores de movimiento y en otros elementos sintácticos recibidos desde la unidad de descodificación por entropía 70. Los bloques predictivos se pueden generar a partir de una de las imágenes de referencia dentro de una de las listas de imágenes de referencia. El descodificador de vídeo 30 puede construir las listas de imágenes de referencia, la Lista 0 y la Lista 1, usando técnicas de construcción por omisión, basándose en las imágenes de referencia almacenadas en la memoria de imágenes de referencia 82.

[0103] El descodificador de vídeo 30 puede determinar un conjunto de imágenes de referencia, a partir del cual el descodificador de vídeo 30 construye las listas de imágenes de referencia. En algunos ejemplos, por ejemplo, cuando el descodificador de vídeo 30 se ajusta a HEVC, el descodificador de vídeo 30 puede construir las listas de imágenes de referencia de acuerdo con la Sección 8.3.3 de HEVC WD7. El conjunto de imágenes de referencia en general corresponde a las imágenes de referencia almacenadas actualmente en una memoria intermedia de imágenes descodificadas, por ejemplo, dentro de la memoria de imágenes de referencia 82. Ciertas imágenes de referencia del conjunto de imágenes de referencia pueden marcarse como imágenes de referencia a corto plazo, mientras que otras imágenes de referencia del conjunto de imágenes de referencia pueden marcarse como imágenes de referencia a largo plazo. En general, una imagen de referencia a largo plazo se marca como "utilizada para referencia a largo plazo", por ejemplo, en una cabecera de unidad de capa de abstracción de red (NAL) de una unidad NAL que encapsula datos de vídeo codificados para la imagen de referencia.

[0104] De acuerdo con las técnicas de esta divulgación, el descodificador de vídeo 30 puede configurarse para descodificar identificadores para imágenes de referencia a largo plazo utilizando codificación diferencial. Por ejemplo, para una primera imagen de referencia a largo plazo (por ejemplo, una imagen de referencia a largo plazo que tiene un valor de índice de 0 en una secuencia de imágenes), el descodificador de vídeo 30 puede descodificar un valor de diferencia indicativo de una diferencia entre MSB de POC para la imagen de referencia a largo plazo y MSB de POC para una imagen actual. Los MSB de POC para la imagen actual pueden denominarse "MSB base". Para otras imágenes de referencia a largo plazo (por ejemplo, imágenes de referencia a largo plazo que tienen un valor de índice mayor que 0 en la secuencia de imágenes), el descodificador de vídeo 30 puede descodificar un valor de diferencia indicativo de una diferencia entre MSB de POC para la imagen de referencia a largo plazo y MSB de POC para una

imagen de referencia a largo plazo anterior. Suponiendo que la imagen de referencia a largo plazo para la que se están determinando los MSB de POC actualmente tiene un valor de índice de i , el descodificador de vídeo 30 puede referirse a los MSB de POC de la imagen de referencia a largo plazo que tiene un valor de índice de cualquier valor entre 0 e $i-1$, inclusive, para determinar los MSB de POC. Es decir, el descodificador de vídeo 30 puede referirse a los MSB de POC de una imagen de referencia a largo plazo que tiene un valor de índice de j , donde j está entre 0 y $i-1$, inclusive. El valor de la diferencia puede, en algunos casos, tener un valor entero negativo.

[0105] Después de recibir un valor de diferencia que describe la diferencia entre los MSB de POC para una imagen de referencia a largo plazo, el descodificador de vídeo 30 puede determinar un valor de MSB de POC al que se aplica el valor de diferencia (es decir, al que tiene que agregarse el valor de diferencia para reconstruir los MSB de POC para la imagen de referencia a largo plazo). Por ejemplo, si la imagen de referencia a largo plazo tiene un valor de índice de 0 (lo cual indica que la imagen de referencia a largo plazo es la primera imagen de referencia a largo plazo ordinal en una secuencia de imágenes), el descodificador de vídeo 30 puede usar el valor de MSB de POC de la imagen actual como un valor de MSB base, y agregar el valor de diferencia señalado para la imagen de referencia a largo plazo al valor de MSB base. Como otro ejemplo, si la imagen de referencia a largo plazo tiene un valor de índice mayor que 0 (lo cual indica que la imagen de referencia a largo plazo no es la primera imagen de referencia a largo plazo ordinal en la secuencia de imágenes), suponiendo que el valor del índice es i , el descodificador de vídeo 30 puede agregar el valor de diferencia señalado a los MSB de POC para la imagen de referencia a largo plazo que tiene un valor de índice $i-1$.

[0106] De esta manera, el descodificador de vídeo 30 puede descodificar un primer valor representativo de una diferencia entre un valor de bits más significativos (MSB) de un valor de recuento de orden de imágenes (POC) de una imagen actual de datos de vídeo y un primer valor de MSB de un primer valor de POC de una primera imagen de referencia a largo plazo de los datos de vídeo, y descodificar un segundo valor representativo de una diferencia entre un segundo valor de MSB de un segundo valor de POC de una segunda imagen de referencia a largo plazo de los datos de vídeo y el primer valor de MSB, en el que el primer valor de POC y el segundo valor de POC tienen diferentes valores de bits menos significativos (LSB). Del mismo modo, el descodificador de vídeo 30 puede calcular un primer valor de ciclo de MSB para la primera imagen de referencia a largo plazo como $\Delta\text{PocMSBCycleLt}[i - 1]$ utilizando el primer valor (es decir, el primer valor de diferencia), y calcular un segundo valor de ciclo de MSB para la segunda imagen de referencia a largo plazo como $\Delta\text{PocMSBCycleLt}[i]$, en el que el cálculo del segundo valor de ciclo de MSB comprende el cálculo de $\Delta\text{PocMSBCycleLt}[i]$ basado en $\Delta\text{PocMSBCycleLt}[i - 1]$ y el segundo valor (es decir, el segundo valor de diferencia).

[0107] Además, el descodificador de vídeo 30 puede descodificar LSB para los valores de POC de las imágenes de referencia a largo plazo. Los LSB pueden señalarse en una cabecera de fragmento de un fragmento para la imagen actual, un conjunto de parámetros de secuencia (SPS) para una secuencia que incluye la imagen actual, un conjunto de parámetros de imagen (PPS) para la imagen actual, o en cualquier otro lugar. El descodificador de vídeo 30 puede concatenar los MSB y los LSB para el valor de POC de una imagen de referencia a largo plazo para determinar el valor de POC completo de la imagen de referencia a largo plazo. A continuación, el descodificador de vídeo 30 puede ensamblar una lista de imágenes de referencia que incluya las imágenes de referencia a largo plazo identificadas por el valor de POC. De esta manera, el descodificador de vídeo 30 puede usar información de movimiento (por ejemplo, datos que identifican una lista de imágenes de referencia y un índice de referencia) para realizar la predicción inter. Por ejemplo, como se explica a continuación, la unidad de compensación de movimiento 72 puede determinar un bloque de referencia para un bloque de una imagen actual basándose al menos en parte en información de movimiento que identifica una imagen de referencia, correspondiente a una de las imágenes en una de las listas de imágenes de referencia. La imagen de referencia puede corresponder a una de las imágenes de referencia a largo plazo, o a una imagen de referencia a corto plazo.

[0108] La unidad de compensación de movimiento 72 determina la información de predicción para un bloque de vídeo del fragmento de vídeo actual analizando los vectores de movimiento y otros elementos sintácticos, y usa la información de predicción para producir los bloques predictivos del bloque de vídeo actual que se está descodificando. Por ejemplo, la unidad de compensación de movimiento 72 usa algunos de los elementos sintácticos recibidos para determinar un modo de predicción (por ejemplo, predicción intra o predicción inter) usada para codificar los bloques de vídeo del fragmento de vídeo, un tipo de fragmento de predicción inter (por ejemplo, fragmento B, fragmento P o fragmento GPB), información de construcción para una o más de las listas de imágenes de referencia para el fragmento, vectores de movimiento para cada bloque de vídeo intercodificado del fragmento, el estado de predicción inter para cada bloque de vídeo intercodificado del fragmento y otra información para descodificar los bloques de vídeo en el fragmento de vídeo actual.

[0109] La unidad de compensación de movimiento 72 también puede realizar la interpolación basándose en filtros de interpolación. La unidad de compensación de movimiento 72 puede usar filtros de interpolación como los usados por el codificador de vídeo 20 durante la codificación de los bloques de vídeo, para calcular valores interpolados para píxeles fraccionarios de los bloques de referencia. En este caso, la unidad de compensación de movimiento 72 puede determinar los filtros de interpolación usados por el codificador de vídeo 20 a partir de los elementos sintácticos recibidos y usar los filtros de interpolación para generar bloques predictivos.

- 5 **[0110]** La unidad de cuantificación inversa 76 cuantifica de manera inversa, es decir, descuantifica, los coeficientes de transformada cuantificados proporcionados en el flujo de bits y descodificados por la unidad de descodificación por entropía 70. El proceso de cuantificación inversa puede incluir el uso de un parámetro de cuantificación QPY, calculado por el descodificador 30 de vídeo de cada bloque de vídeo en el fragmento de vídeo para determinar un grado de cuantificación y, asimismo, un grado de cuantificación inversa que debería aplicarse. La unidad de transformada inversa 78 aplica una transformada inversa, por ejemplo, una DCT inversa, una transformada entera inversa, o un proceso de transformada inversa conceptualmente similar, a los coeficientes de transformada, con el fin de generar bloques residuales en el dominio del píxel.
- 10 **[0111]** Después de que la unidad de compensación de movimiento 72 genera el bloque predictivo para el bloque de vídeo actual, basándose en los vectores de movimiento y otros elementos sintácticos, el descodificador de vídeo 30 forma un bloque de vídeo descodificado sumando los bloques residuales procedentes de la unidad de transformada inversa 78 a los correspondientes bloques predictivos generados por la unidad de compensación de movimiento 72. El sumador 80 representa el componente o los componentes que realizan esta operación de suma. Si se desea, también se puede aplicar un filtro de eliminación de bloques para filtrar los bloques descodificados, a fin de eliminar distorsiones de efecto pixelado. También se pueden usar otros filtros de bucle (en el bucle de codificación o bien después del bucle de codificación) para allanar las transiciones de píxeles o mejorar de otro modo la calidad del vídeo. Los bloques de vídeo descodificados de una trama o imagen determinada se almacenan a continuación en la memoria de imágenes de referencia 82, que almacena imágenes de referencia usadas para una subsiguiente compensación de movimiento. La memoria de imágenes de referencia 82 almacena también vídeo descodificado para su presentación posterior en un dispositivo de visualización, tal como el dispositivo de visualización 32 de la figura 1.
- 15 **[0112]** De esta manera, el descodificador de vídeo 30 de la FIG. 3 representa un ejemplo de un descodificador de vídeo que puede configurarse para descodificar un primer valor representativo de una diferencia entre un valor de bits más significativos (MSB) base de un valor de recuento de orden de imágenes (POC) de una imagen actual de datos de vídeo y un primer valor de MSB de un primer valor de POC de una primera imagen de referencia a largo plazo de los datos de vídeo, descodificar un segundo valor representativo de una diferencia entre un segundo valor de MSB de un segundo valor de POC de una segunda imagen de referencia a largo plazo de los datos de vídeo y el primer valor de MSB, en el que el primer valor de POC y el segundo valor de POC tienen diferentes valores de bits menos significativos (LSB), y descodificar al menos una parte de una imagen actual de los datos de vídeo en relación con al menos una de las primeras imágenes de referencia a largo plazo utilizando el primer valor y la segunda imagen de referencia a largo plazo utilizando el primer valor y el segundo valor.
- 20 **[0113]** El descodificador de vídeo 30 de la FIG. 3 también representa un ejemplo de un descodificador de vídeo que se puede configurar, de forma adicional o alternativa, para descodificar datos para una imagen actual de datos de vídeo, en el que los datos descodificados representan un valor entero negativo para un recuento de orden de imágenes (POC) del ciclo de bits más significativos (MSB) para obtener una imagen de referencia a largo plazo relativa a la imagen actual, y descodificar al menos una parte de la imagen actual en relación con la imagen de referencia a largo plazo basándose, al menos en parte, en el valor entero negativo para el ciclo de MSB.
- 25 **[0114]** El descodificador de vídeo 30 de la FIG. 3 también representa un ejemplo de un descodificador de vídeo que puede configurarse, de forma adicional o alternativa, para determinar un número de imágenes de referencia marcadas como imágenes de referencia a corto plazo, de un número total de imágenes de referencia disponibles, para una imagen actual, descodificar un valor de bits más significativos (MSB) para imágenes de referencia a largo plazo para la imagen actual basándose en el número total de imágenes de referencia y el número de imágenes de referencia marcadas como imágenes de referencia a corto plazo, y descodificar al menos una parte de la imagen actual relativa al menos a una de las imágenes de referencia disponibles.
- 30 **[0115]** El descodificador de vídeo 30 de la FIG. 3 también representa un ejemplo de un descodificador de vídeo que puede configurarse, de forma adicional o alternativa, para determinar un número N de imágenes de referencia para una imagen actual que tenga valores de recuento de orden de imágenes (POC) con valores de bits menos significativos (LSB) iguales, determinar que el conjunto de imágenes de referencia puede incluir N imágenes de referencia a largo plazo, y descodificar delta valores del ciclo del bit más significativo (MSB) de POC para como máximo N-1 de las imágenes de referencia a largo plazo basándose en las determinaciones.
- 35 **[0116]** La FIG. 4 es un diagrama conceptual que ilustra una secuencia de imágenes de vídeo codificadas 100-132. Las imágenes están sombreadas de manera diferente para indicar las posiciones dentro de una estructura de predicción jerárquica. Por ejemplo, las imágenes 100, 116 y 132 están sombreadas en negro para representar que las imágenes 100, 116, 132 están en la parte superior de la estructura de predicción jerárquica. Las imágenes 100, 116, 132 pueden comprender, por ejemplo, imágenes intra-codificadas o imágenes inter-codificadas que se predicen a partir de otras imágenes en una única dirección (por ejemplo, imágenes P). Cuando están intra-codificadas, las imágenes 100, 116, 132 se predicen únicamente a partir de datos dentro de la misma imagen. Cuando se inter-codifican, la imagen 116, por ejemplo, puede codificarse con relación a los datos de la imagen 100, como se indica mediante la flecha discontinua desde la imagen 116 a la imagen 100. Las imágenes 116, 132 forman imágenes clave de los grupos de imágenes (GOP) 134, 136, respectivamente.
- 40 **[0117]** El descodificador de vídeo 30 de la FIG. 3 también representa un ejemplo de un descodificador de vídeo que puede configurarse, de forma adicional o alternativa, para determinar un número de imágenes de referencia marcadas como imágenes de referencia a corto plazo, de un número total de imágenes de referencia disponibles, para una imagen actual, descodificar un valor de bits más significativos (MSB) para imágenes de referencia a largo plazo para la imagen actual basándose en el número total de imágenes de referencia y el número de imágenes de referencia marcadas como imágenes de referencia a corto plazo, y descodificar al menos una parte de la imagen actual relativa al menos a una de las imágenes de referencia disponibles.
- 45 **[0118]** El descodificador de vídeo 30 de la FIG. 3 también representa un ejemplo de un descodificador de vídeo que puede configurarse, de forma adicional o alternativa, para determinar un número N de imágenes de referencia para una imagen actual que tenga valores de recuento de orden de imágenes (POC) con valores de bits menos significativos (LSB) iguales, determinar que el conjunto de imágenes de referencia puede incluir N imágenes de referencia a largo plazo, y descodificar delta valores del ciclo del bit más significativo (MSB) de POC para como máximo N-1 de las imágenes de referencia a largo plazo basándose en las determinaciones.
- 50 **[0119]** El descodificador de vídeo 30 de la FIG. 3 también representa un ejemplo de un descodificador de vídeo que puede configurarse, de forma adicional o alternativa, para determinar un número N de imágenes de referencia para una imagen actual que tenga valores de recuento de orden de imágenes (POC) con valores de bits menos significativos (LSB) iguales, determinar que el conjunto de imágenes de referencia puede incluir N imágenes de referencia a largo plazo, y descodificar delta valores del ciclo del bit más significativo (MSB) de POC para como máximo N-1 de las imágenes de referencia a largo plazo basándose en las determinaciones.
- 55 **[0120]** El descodificador de vídeo 30 de la FIG. 3 también representa un ejemplo de un descodificador de vídeo que puede configurarse, de forma adicional o alternativa, para determinar un número N de imágenes de referencia para una imagen actual que tenga valores de recuento de orden de imágenes (POC) con valores de bits menos significativos (LSB) iguales, determinar que el conjunto de imágenes de referencia puede incluir N imágenes de referencia a largo plazo, y descodificar delta valores del ciclo del bit más significativo (MSB) de POC para como máximo N-1 de las imágenes de referencia a largo plazo basándose en las determinaciones.
- 60 **[0121]** El descodificador de vídeo 30 de la FIG. 3 también representa un ejemplo de un descodificador de vídeo que puede configurarse, de forma adicional o alternativa, para determinar un número N de imágenes de referencia para una imagen actual que tenga valores de recuento de orden de imágenes (POC) con valores de bits menos significativos (LSB) iguales, determinar que el conjunto de imágenes de referencia puede incluir N imágenes de referencia a largo plazo, y descodificar delta valores del ciclo del bit más significativo (MSB) de POC para como máximo N-1 de las imágenes de referencia a largo plazo basándose en las determinaciones.
- 65 **[0122]** El descodificador de vídeo 30 de la FIG. 3 también representa un ejemplo de un descodificador de vídeo que puede configurarse, de forma adicional o alternativa, para determinar un número N de imágenes de referencia para una imagen actual que tenga valores de recuento de orden de imágenes (POC) con valores de bits menos significativos (LSB) iguales, determinar que el conjunto de imágenes de referencia puede incluir N imágenes de referencia a largo plazo, y descodificar delta valores del ciclo del bit más significativo (MSB) de POC para como máximo N-1 de las imágenes de referencia a largo plazo basándose en las determinaciones.

[0117] Las imágenes 108, 124 están oscurecidas para indicar que están en la jerarquía de codificación a continuación de las imágenes 100, 116 y 132. Las imágenes 108, 124 pueden comprender imágenes bidireccionales codificadas con modo de predicción inter. Por ejemplo, la imagen 108 puede predecirse a partir de los datos de las imágenes 100 y 116, mientras que la imagen 124 puede predecirse a partir de las imágenes 116 y 132. Las imágenes 104, 112, 120 y 128 están ligeramente sombreadas para indicar que están a continuación en la jerarquía de codificación que sigue a las imágenes 108 y 124. Las imágenes 104, 112, 120 y 128 también pueden comprender imágenes bidireccionales codificadas con modo de predicción inter. Por ejemplo, la imagen 104 puede predecirse a partir de las imágenes 100 y 108, la imagen 112 puede predecirse a partir de las imágenes 108 y 116, la imagen 120 puede predecirse a partir de las imágenes 116 y 124, y la imagen 128 puede predecirse a partir de la imagen 124 y 132. En general, las imágenes que están más abajo en la jerarquía pueden codificarse a partir de cualquier imagen de referencia que sea superior en la jerarquía, suponiendo que las imágenes de referencia todavía están almacenadas temporalmente en una memoria intermedia de imágenes descodificadas, y suponiendo que las imágenes de referencia fueron codificadas antes que la imagen que actualmente está siendo codificada.

[0118] Finalmente, las imágenes 102, 106, 110, 114, 118, 122, 126 y 130 están sombreadas de color blanco para indicar que estas imágenes son las últimas en la jerarquía de codificación. Las imágenes 102, 106, 110, 114, 118, 122, 126 y 130 pueden ser imágenes bidireccionales codificadas con modo de predicción inter. La imagen 102 puede predecirse a partir de las imágenes 100 y 104, la imagen 106 puede predecirse a partir de las imágenes 104 y 108, la imagen 110 puede predecirse a partir de las imágenes 108 y 112, la imagen 114 puede predecirse a partir de las imágenes 112 y 116, la imagen 118 puede predecirse a partir de las imágenes 116 y 120, la imagen 122 puede predecirse a partir de las imágenes 120 y 124, la imagen 126 puede predecirse a partir de las imágenes 124 y 128, y la imagen 130 puede predecirse a partir de las imágenes 128 y 132. Nuevamente, debe entenderse que las imágenes inferiores en la jerarquía de codificación pueden codificarse a partir de otras imágenes que están más alto en la jerarquía de codificación. Por ejemplo, cualquiera de, o todas, las imágenes 102, 106, 110 o 114 pueden predecirse en relación con cualquiera de las imágenes 100, 116 o 108, además o de forma alternativa.

[0119] Las imágenes 100-132 se ilustran en orden de visualización. Es decir, después de la descodificación, la imagen 100 se visualiza antes que la imagen 102, la imagen 102 se visualiza antes que la imagen 104, y así sucesivamente. Como se ha expuesto anteriormente, los valores de POC en general describen un orden de visualización de imágenes, que también es esencialmente el mismo que el orden en el que las imágenes en bruto fueron capturadas o generadas antes de ser codificadas. Sin embargo, debido a la jerarquía de codificación, las imágenes 100-132 pueden descodificarse en un orden diferente. Además, mientras son codificadas, las imágenes 100-132 pueden disponerse en orden de descodificación en un flujo de bits que incluye datos codificados para las imágenes 100-132. Por ejemplo, la imagen 116 se puede mostrar última entre las imágenes del GOP 134. Sin embargo, debido a la jerarquía de codificación, la imagen 116 puede descodificarse como la primera del GOP 134. Es decir, para descodificar adecuadamente la imagen 108, por ejemplo, la imagen 116 puede necesitar ser descodificada primero, para actuar como una imagen de referencia para la imagen 108. Asimismo, la imagen 108 puede actuar como una imagen de referencia para las imágenes 104, 106, 110 y 112, y por lo tanto puede necesitar ser descodificada antes que las imágenes 104, 106, 110 y 112.

[0120] Además, ciertas imágenes pueden tratarse como imágenes de referencia a largo plazo, mientras que otras imágenes pueden tratarse como imágenes de referencia a corto plazo. Supongamos, por ejemplo, que las imágenes 100 y 116 representan imágenes de referencia a largo plazo, mientras que las imágenes 108, 104 y 112 representan imágenes de referencia a corto plazo. Puede ser el caso, en este ejemplo, que las imágenes 102 y 106 puedan predecirse con respecto a cualquiera de las imágenes 100, 116, 108 o 104, pero que las imágenes 110 y 114 puedan predecirse con relación a cualquiera de las imágenes 100, 116, 108 o 112. En otras palabras, la imagen 104 puede no estar disponible para referencia cuando se codifican las imágenes 110 y 114. Como otro ejemplo, suponiendo que las imágenes 100 y 116 representan imágenes de referencia a largo plazo y las imágenes 108, 104 y 112 representan imágenes de referencia a corto plazo, las imágenes 108, 104 y 112 pueden no estar disponibles como referencia cuando se codifican las imágenes 118, 122, 126 y 130.

[0121] De acuerdo con las técnicas de esta divulgación, los datos con respecto a las imágenes de referencia a largo plazo pueden señalarse en las cabeceras de fragmento de los fragmentos para cualquiera de, o todas, las imágenes 100-132. De forma alternativa, los datos pueden señalarse en PPS, SPS u otras estructuras de datos.

[0122] Suponiendo de nuevo que las imágenes 100 y 116 representan imágenes de referencia a largo plazo, el codificador de vídeo 20 puede codificar MSB de POC para la imagen 100 como una diferencia con respecto a un valor de MSB base, por ejemplo, MSB de POC de una imagen actual, como la imagen 102. Es decir, el codificador de vídeo 20 puede calcular una diferencia entre los MSB de POC para la imagen 100 y los MSB de POC para la imagen 102, y codificar el valor de diferencia para la imagen 100. Del mismo modo, el codificador de vídeo 20 puede calcular una diferencia entre los MSB de POC para la imagen 116 y los MSB de POC para la imagen 100, y codificar el valor de diferencia para la imagen 116. El codificador de vídeo 20 también puede codificar valores de LSB para las imágenes 100 y 116.

[0123] Por lo tanto, el descodificador de vídeo 30 puede reconstruir valores de POC para imágenes 100 y 116 mediante la descodificación del valor de diferencia para la imagen 100 y añadiendo el valor de diferencia al valor de MSB base

para la imagen 102. Del mismo modo, el descodificador de vídeo 30 puede descodificar el valor de diferencia para la imagen 116 y agregar el valor de diferencia a los MSB de POC para la imagen 100. El descodificador de vídeo 30 también puede descodificar valores de POC LSBs para las imágenes 100 y 116. El descodificador de vídeo 30 puede concatenar los MSB de POC para la imagen 100 con los POC LSB para la imagen 100 para reproducir el valor de POC para la imagen 100. Del mismo modo, el descodificador de vídeo 30 puede concatenar los MSB de POC para la imagen 116 con los LSB de POC para que la imagen 116 reproduzca el valor de POC para la imagen 116.

[0124] De esta manera, tanto el codificador de vídeo 20 como el descodificador de vídeo 30 pueden configurarse para codificar un primer valor representativo de una diferencia entre un valor de bits más significativos (MSB) base de un valor de recuento de orden de imágenes (POC) de una imagen actual de datos de vídeo y un primer valor de MSB de un primer valor de POC de una primera imagen de referencia a largo plazo de los datos de vídeo, codificar un segundo valor representativo de una diferencia entre un segundo valor de MSB de un segundo valor de POC de una segunda imagen de referencia a largo plazo de los datos de vídeo y el primer valor de MSB, en el que el primer valor de POC y el segundo valor de POC tienen diferentes valores de bits menos significativos (LSB), y codificar al menos una parte de una imagen actual de los datos de vídeo en relación con al menos una de la primera imagen de referencia a largo plazo utilizando el primer valor y la segunda imagen de referencia a largo plazo utilizando el primer valor y el segundo valor.

[0125] Más particularmente, el codificador de vídeo 20 puede codificar datos representativos de valores de POC para imágenes de referencia a largo plazo que se incluirán en una lista de imágenes de referencia para una imagen actual, por ejemplo, la imagen 102. En general, el codificador de vídeo 20 puede indicar que una imagen de referencia debe incluirse en una lista de imágenes de referencia cuando al menos una parte de la imagen actual está codificada en relación con la imagen de referencia. Una parte de una imagen puede considerarse codificada en relación con una imagen de referencia cuando un bloque de la imagen está intercodificado en relación con la imagen de referencia, de modo que un vector de movimiento para el bloque apunta a un bloque de referencia de la imagen de referencia. La información de movimiento para el bloque puede incluir un identificador de lista de imágenes de referencia (por ejemplo, Lista 0 o Lista 1), así como un índice de referencia que corresponde a una imagen de referencia en la lista de imágenes de referencia identificada por el identificador de lista de imágenes de referencia.

[0126] Por lo tanto, el codificador de vídeo 20 puede codificar un bloque de imagen 102 con respecto a un bloque de referencia de la imagen 100. Es decir, el codificador de vídeo 20 puede realizar una búsqueda de movimiento del bloque y determinar que el bloque de referencia de la imagen 100 produce un valor de error aceptable, en relación con otros bloques de referencia. De manera similar, el codificador de vídeo 20 puede codificar un bloque diferente de la imagen 102 en relación con un bloque de referencia de la imagen 116.

[0127] La FIG. 5 es un diagrama de flujo que ilustra un procedimiento de ejemplo para codificar valores de POC para imágenes de referencia a largo plazo de acuerdo con las técnicas de esta divulgación. Con fines explicativos, el procedimiento de la FIG. 5 se explica con respecto al codificador de vídeo 20 de las FIGs. 1 y 2 y los componentes del mismo. Sin embargo, debe entenderse que otros dispositivos de codificación de vídeo pueden configurarse para realizar el procedimiento de la FIG. 5.

[0128] Inicialmente, el codificador de vídeo 20 puede codificar datos de una imagen actual utilizando una o más imágenes de referencia (150). El codificador de vídeo 20 puede entonces determinar las imágenes de referencia que se usaron para codificar la imagen actual (152). Más particularmente, el codificador de vídeo 20 puede determinar qué imágenes de referencia en un conjunto de imágenes de referencia disponibles se usan realmente como referencia cuando se codifica la imagen actual. En algunos ejemplos, los pasos 150 y 152 pueden integrarse y realizarse de manera sustancialmente simultánea. Entre las imágenes de referencia disponibles pueden incluirse imágenes de referencia a largo plazo (LTRP) e imágenes de referencia a corto plazo (STRPs). El codificador de vídeo 20 puede determinar si establecer una imagen de referencia como una imagen de referencia a largo plazo o una imagen de referencia a corto plazo basándose en los resultados de codificación obtenidos a lo largo de varios pases de codificación de prueba. Por ejemplo, cuando los resultados de la codificación (por ejemplo, las métricas de optimización de distorsión de la velocidad (RDO)) son en general mejores cuando una imagen de referencia en particular se trata como una imagen de referencia a largo plazo (potencialmente desviada por la cantidad de memoria consumida para almacenar la imagen de referencia a largo plazo), el codificador de vídeo 20 puede tratar esa imagen de referencia como una imagen de referencia a largo plazo.

[0129] En cualquier caso, después de determinar el conjunto de imágenes de referencia y esas imágenes de referencia que han de ser tratadas como imágenes de referencia a largo plazo, el codificador de vídeo 20 pueden codificar información que indica que las imágenes han de incluirse en listas de imágenes de referencia para la imagen actual como imágenes de referencia a largo plazo. La información codificada puede comprender datos representativos de valores de POC para las imágenes de referencia a largo plazo. El codificador de vídeo 20 puede codificar esta información en una cabecera de fragmento de un fragmento de la imagen actual, un PPS correspondiente a la imagen actual, un SPS para una secuencia que incluye la imagen actual, o en otro lugar.

[0130] El codificador de vídeo 20 puede codificar datos indicativos de MSB del valor de POC para una primera imagen de referencia a largo plazo ordinal, así como datos indicativos de LSB del valor de POC para la primera imagen de

referencia a largo plazo ordinal (154). Para codificar los MSB, el codificador de vídeo 20 puede calcular una diferencia entre los MSB del valor de POC para la primera imagen de referencia a largo plazo y los MSB del valor de POC para la imagen actual, y codificar este valor de diferencia calculado.

5 **[0131]** El codificador de vídeo 20 puede entonces determinar si los LSB de un valor de POC para un LTRP subsiguiente son iguales a los LSB de un valor de POC para otra imagen de referencia para la imagen actual (156). Cuando el
 10 codificador de vídeo 20 determina que los LSB del valor de POC para el LTRP subsiguiente son iguales a los LSB del valor de POC para otra referencia para la imagen actual (rama "Sí" de 156), el codificador de vídeo 20 puede codificar
 15 datos representativos de una diferencia entre los MSB del valor de POC para el LTRP subsiguiente y los MSB del valor de POC para un LTRP anterior (es decir, el LTRP más reciente para el que se codificó la información MSB de POC) (158). Los LSB del valor de POC para el LTRP más reciente no son necesariamente los mismos que los LSB del valor de POC para el LTRP actualmente codificado. Como se analizó anteriormente, para calcular el valor de la diferencia, suponiendo que el LTRP posterior tiene un valor de índice de i , el codificador de vídeo 20 puede calcular la diferencia entre los MSB del valor de POC para LTRP i y los MSB del valor de POC para LTRP j si LTRP j fue el LTRP anterior para el que se codificó la información MSB de POC. El codificador de vídeo 20 puede codificar el valor de diferencia. Este valor de diferencia puede corresponder a $\text{delta_poc_msb_cycle_lt}[i]$ (o $\text{delta_poc_msb_cycle_lt_plus1}[i]$).

20 **[0132]** Después de codificar el valor de diferencia, o cuando los LSB del valor de POC para el próximo LTRP no son iguales a los LSB de otra imagen de referencia (rama "NO" de 156), el codificador de vídeo 20 puede codificar los LSB del valor de POC para el LTRP (160). El codificador de vídeo 20 puede entonces determinar si hay más LTRP para la imagen actual (162). Cuando hay más LTRP para la imagen actual (rama "Sí" de 162), el codificador de vídeo 20 puede proceder a codificar los LSB de los valores de POC para los LTRP restantes y, cuando sea necesario, los MSB utilizan la codificación diferencial como se analizó anteriormente. Sin embargo, cuando no hay más LTRP para la
 25 imagen actual (rama "NO" de 162), el codificador de vídeo 20 puede generar los datos codificados (164), que pueden incluir bloques codificados para la imagen actual, valores de POC codificados para imágenes de referencia a largo plazo para la imagen actual (LSB incluidos y, cuando los LSB no son únicos, valores de diferencia para MSB), y similares.

30 **[0133]** De esta manera, el procedimiento de la FIG. 5 representa un ejemplo de un procedimiento que incluye codificar al menos una primera parte de una imagen actual de datos de vídeo en relación con una primera imagen de referencia a largo plazo y al menos una segunda parte de la imagen actual a una segunda imagen de referencia a largo plazo, codificar un primer valor representativo de una diferencia entre un valor de bits más significativos (MSB) base de un valor de recuento de orden de imágenes (POC) de una imagen actual de los datos de vídeo y un primer valor de MSB
 35 de un primer valor de POC de la primera imagen de referencia a largo plazo de los datos de vídeo, y codificación de un segundo valor representativo de una diferencia entre un segundo valor de MSB de un segundo valor de POC de la segunda imagen de referencia a largo plazo de los datos de vídeo y el primer valor de MSB, en el que el primer valor de POC y el segundo valor de POC tiene diferentes valores de bits menos significativos (LSB).

40 **[0134]** La FIG. 6 es un diagrama de flujo que ilustra un procedimiento de ejemplo para descodificar valores de POC para imágenes de referencia a largo plazo de acuerdo con las técnicas de esta divulgación. Con fines explicativos, el procedimiento de la FIG. 6 se explica con respecto al descodificador de vídeo 30 de las FIGs. 1 y 3 y los componentes del mismo. Sin embargo, debe entenderse que otros dispositivos de descodificación de vídeo pueden configurarse para realizar el procedimiento de la FIG. 6.
 45

[0135] Inicialmente, el descodificador de vídeo 30 puede descodificar MSB y LSB de un valor de POC para una primera imagen de referencia a largo plazo (LTRP) ordinal de una imagen actual (200). Por ejemplo, el descodificador de vídeo 30 puede descodificar un valor de diferencia representativo de una diferencia entre los MSB del valor de POC para el primer LTRP y los MSB del valor de POC para la imagen actual. A continuación, el descodificador de vídeo 30 puede
 50 agregar el valor de diferencia a los MSB del valor de POC para la imagen actual, para obtener los MSB del primer LTRP. El descodificador de vídeo 30 también puede descodificar los LSB del valor de POC para el primer LTRP y concatenar los MSB y los LSB del valor de POC del primer LTRP para reproducir el valor de POC para el LTRP. El descodificador de vídeo 30 también puede descodificar los LSB del valor de POC para el primer LTRP solo, cuando no se envían los MSB, y puede usar los valores de LSB solo para identificar si no hay otras imágenes de referencia con el mismo valor de POC LSB.
 55

[0136] A continuación, el descodificador de vídeo 30 puede descodificar los LSB de un valor de POC para un LTRP siguiente (202). En el ejemplo de la FIG. 6, el descodificador de vídeo 30 determina si los LSB del valor de POC para el próximo LTRP son iguales a los LSB de otro LTRP para la imagen actual (204). En otros ejemplos, por ejemplo,
 60 como se explica con respecto a la Tabla 1 anterior, el descodificador de vídeo 30 puede determinar si un valor de indicador indica que los MSB están señalados para el valor de POC del siguiente LTRP. En cualquier caso, el descodificador de vídeo 30 puede determinar si los MSB se señalan para el valor de POC del siguiente LTRP. Cuando los MSB se señalan para el valor de POC del siguiente LTRP (por ejemplo, cuando los LSB del valor de POC para el siguiente LTRP son iguales a los LSB de otra imagen de referencia (rama "YES" de 204)), el descodificador de vídeo 30 puede descodificar un valor de diferencia de MSB relativo a MSB de un valor de POC para un LTRP anterior (206). Es decir, el descodificador de vídeo 30 puede descodificar el valor de diferencia señalado para LTRP i , y agregar el
 65

valor de diferencia a los MSB del valor de POC para LTRP j donde LTRP j era el LTRP anterior para el cual se descodificó la información MSB. El descodificador de vídeo 30 puede volver a ensamblar el valor de POC del siguiente LTRP concatenando los MSB (ya sea calculados o determinados implícitamente) con los LSB para el siguiente LTRP.

5 **[0137]** Después de determinar que los LSB del valor de POC para el próximo LTRP no son iguales a los LSB de otra imagen de referencia (rama "NO" de 204), o después de descodificar el valor de diferencia para los MSB del siguiente LTRP, el descodificador de vídeo 30 puede determinar si hay más LTRP para la imagen actual (210). Si hay más LTRP para la imagen actual (rama "Sí" de 210), el descodificador de vídeo 30 puede proceder a descodificar los valores de POC para un LTRP posterior, por ejemplo, como se analizó anteriormente. Una vez que no haya más LTRP para la imagen actual (rama "NO" de 210), el descodificador de vídeo 30 puede ensamblar una o más listas de imágenes de referencia que incluyen los LTRP señalados (212). A continuación, el descodificador de vídeo 30 puede descodificar la imagen actual usando la(s) lista(s) de imágenes de referencia (214).

15 **[0138]** Por ejemplo, cuando se realiza predicción inter de un bloque, el bloque puede incluir una indicación de una lista de imágenes de referencia y un índice de referencia en la lista de imágenes de referencia, que indica una imagen de referencia para el bloque. El bloque puede incluir además datos para reconstruir un vector de movimiento para el bloque. Por lo tanto, el descodificador de vídeo 30 puede recuperar un bloque de referencia de la imagen de referencia usando el vector de movimiento. El descodificador de vídeo 30 puede descodificar además un valor de error para el bloque, por ejemplo, mediante descodificación, cuantificación inversa y transformación inversa de coeficientes de transformada para el bloque. A continuación, el descodificador de vídeo 30 puede combinar el bloque de referencia con un bloque residual reconstruido para descodificar el bloque.

25 **[0139]** De esta manera, el procedimiento de la FIG. 6 representa un ejemplo de un procedimiento que incluye descodificar un primer valor representativo de una diferencia entre un valor de bits de base más significativos (MSB) de un valor de recuento de orden de imágenes (POC) de una imagen actual de datos de vídeo y un primer valor de MSB de un primer valor de POC de una primera imagen de referencia a largo plazo de los datos de vídeo, descodificar un segundo valor representativo de una diferencia entre un segundo valor de MSB de un segundo valor de POC de una segunda imagen de referencia a largo plazo de los datos de vídeo y el primer valor de MSB, en el que el primer valor de POC y el segundo valor de POC tienen diferentes valores de bits menos significativos (LSB), y descodificar al menos una parte de una imagen actual de los datos de vídeo en relación con al menos una de la primera imagen de referencia a largo plazo utilizando primer valor y la segunda imagen de referencia a largo plazo utilizando el primer valor y el segundo valor.

35 **[0140]** 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 agregarse, fusionarse 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 determinados ejemplos, los actos o sucesos se pueden llevar a cabo simultáneamente, por ejemplo, mediante procesamiento de múltiples hilos, procesamiento de interrupciones o múltiples procesadores, en lugar de secuencialmente.

40 **[0141]** En uno o más ejemplos, las funciones descritas se pueden implementar 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.

55 **[0142]** A modo de ejemplo, y no de limitación, dichos 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 se pueda usar para almacenar un código de programa deseado en forma de instrucciones o estructuras de datos y al que se pueda acceder mediante un ordenador. Además, cualquier conexión recibe adecuadamente la denominación de medio legible por ordenador. Por ejemplo, si las instrucciones se transmiten desde una sede de la Red, un servidor u otro origen remoto usando un cable coaxial, un cable de fibra óptica, un par trenzado, una línea de abonado digital (DSL) o tecnologías inalámbricas tales como infrarrojos, radio y microondas, entonces el cable coaxial, el cable de fibra óptica, el par trenzado, la DSL o las tecnologías inalámbricas tales como infrarrojos, radio y microondas se incluyen en la definición de medio. Sin embargo, debería entenderse que los medios de almacenamiento legibles por ordenador y los medios de almacenamiento de datos no incluyen conexiones, ondas portadoras, señales u otros medios transitorios, sino que, en cambio, están orientados a medios de almacenamiento tangibles no transitorios. El término

disco, como se usa en el presente documento, incluye el disco compacto (CD), el disco láser, el disco óptico, el disco versátil digital (DVD), el disco flexible y el disco Blu-ray, donde algunos discos normalmente emiten datos magnéticamente, mientras que otros discos emiten datos ópticamente con láseres. Las combinaciones de lo anterior también deberían incluirse dentro del alcance de los medios legibles por ordenador.

5 [0143] Las instrucciones se pueden ejecutar por uno o más procesadores, tales como uno o más procesadores de señales digitales (DSP), microprocesadores de propósito general, circuitos integrados específicos de la aplicación (ASIC), matrices lógicas 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
10 estructuras anteriores o a cualquier otra estructura adecuada para la implementación de las técnicas descritas en el presente documento. Además, en algunos aspectos, la funcionalidad descrita en el presente documento se puede proporcionar dentro de módulos de hardware y/o software dedicados, configurados para la codificación y la decodificación, o incorporados en un códec combinado. Además, las técnicas se podrían implementar totalmente en uno o más circuitos o elementos lógicos.

15 [0144] Las técnicas de esta divulgación se pueden implementar en una amplia variedad de dispositivos o aparatos, incluyendo un teléfono 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
20 mediante diferentes unidades de hardware. En cambio, como se ha descrito anteriormente, diversas unidades se pueden combinar en una unidad de hardware de códec, o ser proporcionadas por un grupo de unidades de hardware interoperativas, incluyendo uno o más procesadores, como se ha descrito anteriormente, conjuntamente con software y/o firmware adecuados.

REIVINDICACIONES

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

5 descodificar un elemento sintáctico poc_lsb_lt para una primera imagen de referencia a largo plazo, LTRP, de los datos de vídeo, en el que el elemento sintáctico especifica el valor de los bits menos significativos, LSB, del valor de recuento de orden de imágenes, POC, del primer LTRP;

10 descodificar (200) un primer valor representativo de una diferencia entre un valor de bits más significativos, MSB, base de un valor de POC de una imagen actual de datos de vídeo y un primer valor de MSB de un primer valor de POC del primer LTRP de datos de vídeo;

15 descodificar un elemento sintáctico poc_lsb_lt para un segundo LTRP, en el que el elemento sintáctico especifica el valor de los LSB del valor de POC del segundo LTRP;

20 descodificar (206) un segundo valor representativo de una diferencia entre un segundo valor de MSB de un segundo valor de POC del segundo LTRP de los datos de vídeo y el primer valor de MSB, en el que el primer valor de POC y el segundo valor de POC tienen diferentes valores de bits menos significativos, LSB, en los que los LTRP se señalan en orden creciente de sus valores de ciclo de MSB con respecto a la imagen actual; y

25 descodificar (214) al menos una primera parte de la imagen actual de los datos de vídeo en relación con la primera imagen de referencia a largo plazo utilizando el primer valor y descodificar al menos una segunda parte de la imagen actual en relación con la segunda imagen de referencia a largo plazo utilizando el primer valor y el segundo valor.

2. El procedimiento de acuerdo con la reivindicación 1, que comprende además:

30 calcular un primer valor de ciclo de MSB para la primera imagen de referencia a largo plazo, DeltaPocMSBCycleLt [i - 1], utilizando el primer valor; y

35 calcular un segundo valor de ciclo de MSB para la segunda imagen de referencia a largo plazo, DeltaPocMSBCycleLt[i], en el que calcular el segundo valor de ciclo de MSB comprende calcular DeltaPocMSBCycleLt[i] basándose en DeltaPocMSBCycleLt[i - 1] y el segundo valor.

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

40 codificar (150) al menos una primera parte de una imagen actual de datos de vídeo en relación con una primera imagen de referencia a largo plazo, LTRP, y al menos una segunda parte de la imagen actual en relación con una segunda LTRP;

45 codificar un elemento sintáctico poc_lsb_lt para el primer LTRP, en el que el elemento sintáctico para el primer LTRP especifica el valor de los bits menos significativos, LSB, del valor de recuento de orden de imágenes, POC, del primer LTRP;

50 codificar (154) un primer valor representativo de una diferencia entre un valor de bits más significativos, MSB, de base de un valor de POC de la imagen actual de los datos de vídeo y un primer valor de MSB de un primer valor de POC del primer LTRP de datos de vídeo;

55 codificar un elemento sintáctico poc_lsb_lt para el segundo LTRP, en el que el elemento sintáctico para el segundo LTRP especifica el valor de los LSB del valor de POC del segundo LTRP; y

60 codificar (158) un segundo valor representativo de una diferencia entre un segundo valor de MSB de un segundo valor de POC de la segunda imagen de referencia a largo plazo de los datos de vídeo y el primer valor de MSB, en el que el primer valor de POC y el segundo valor de POC tienen diferentes valores de bits menos significativos, LSB, y en el que los LTRP se señalan en orden creciente de sus valores de ciclo de MSB en relación con la imagen actual.

4. El procedimiento de acuerdo con la reivindicación 3, que comprende además:

65 calcular el primer valor como una diferencia entre un primer valor de ciclo de MSB para la primera imagen de referencia a largo plazo, DeltaPocMSBCycleLt[i - 1], y el valor de MSB base; y

 calcular el segundo valor como una diferencia entre un segundo valor de ciclo de MSB para la segunda imagen de referencia a largo plazo, DeltaPocMSBCycleLt[i], y el primer valor de MSB.

5. El procedimiento según la reivindicación 1 o 3, en el que la primera imagen de referencia a largo plazo corresponde a un índice de j , en el que la segunda imagen de referencia a largo plazo corresponde a un índice de i , en el que el primer valor comprende $\text{delta_poc_msb_cycle_lt_plus1}[j]$, en el que el segundo valor comprende $\text{delta_poc_msb_cycle_lt_plus1}[i]$, y en la que j está en el rango de 0 a $i-1$, inclusive.
6. El procedimiento según la reivindicación 1 o 3, en el que al menos uno del primer valor y el segundo valor comprende un valor entero negativo.
7. Un dispositivo para descodificar datos de vídeo, comprendiendo el dispositivo:
- medios para descodificar un elemento sintáctico poc_lsb_lt para una primera imagen de referencia a largo plazo, LTRP, de los datos de vídeo, en el que el elemento sintáctico especifica el valor de los bits menos significativos, LSB, del valor de recuento de orden de imágenes, POC, del primer LTRP de datos de vídeo;
- medios para descodificar un primer valor representativo de una diferencia entre un valor de bits más significativos, MSB, de base de un valor de POC de una imagen actual de datos de vídeo y un primer valor de MSB de un primer valor de POC de la primera imagen de referencia a largo plazo de los datos de vídeo;
- medios para descodificar un elemento sintáctico poc_lsb_lt para un segundo LTRP, en el que el elemento sintáctico especifica el valor de los LSB del valor de POC del segundo LTRP;
- medios para descodificar un segundo valor representativo de una diferencia entre un segundo valor de MSB de un segundo valor de POC del segundo LTRP de los datos de vídeo y el primer valor de MSB, en el que el primer valor de POC y el segundo valor de POC tienen diferentes valores de bits menos significativos (LSB), en el que los LTRP se señalan en orden creciente de sus valores de ciclo de MSB con respecto a la imagen actual; y
- medios para descodificar al menos una primera parte de la imagen actual de los datos de vídeo en relación con el primer LTRP usando el primer valor y descodificar al menos una segunda parte de una imagen actual relacionada con el segundo LTRP usando el primer valor y el segundo valor.
8. El dispositivo de acuerdo con la reivindicación 7, que comprende, además:
- medios para calcular un primer valor de ciclo de MSB para la primera imagen de referencia a largo plazo, $\text{DeltaPocMSBCycleLt}[i - 1]$, utilizando el primer valor; y
- medios para calcular un segundo valor de ciclo de MSB para la segunda imagen de referencia a largo plazo, $\text{DeltaPocMSBCycleLt}[i]$, en el que el cálculo del segundo valor de ciclo de MSB comprende calcular $\text{DeltaPocMSBCycleLt}[i]$ basado en $\text{DeltaPocMSBCycleLt}[i - 1]$ y el segundo valor.
9. Un dispositivo para codificar datos de vídeo, comprendiendo el dispositivo:
- medios para codificar al menos una primera parte de una imagen actual de datos de vídeo en relación con una primera imagen de referencia a largo plazo, LTRP, y al menos una segunda parte de la imagen actual a una segunda LTRP;
- medios para codificar un elemento sintáctico poc_lsb_lt para el primer LTRP, en el que el elemento sintáctico especifica el valor de los bits menos significativos, LSB, del valor de POC del primer LTRP;
- medios para codificar un primer valor representativo de una diferencia entre un valor de MSB de bits más significativos base, de un valor de POC de la imagen actual de los datos de vídeo y un primer valor de MSB de un primer valor de POC de la primera imagen de referencia a largo plazo de los datos de vídeo y
- medios para codificar un elemento sintáctico poc_lsb_lt para el segundo LTRP, en el que el elemento sintáctico especifica el valor de los LSB de la POC del segundo LTRP;
- medios para codificar un segundo valor representativo de una diferencia entre un segundo valor de MSB de un segundo valor de POC de la segunda imagen de referencia a largo plazo de los datos de vídeo y el primer valor de MSB, en el que el primer valor de POC y el segundo valor de POC tienen diferentes valores de bits menos significativos, LSB, y, en el que los LTRP se señalan en orden creciente de sus valores de ciclo de MSB en relación con la imagen actual.
10. El dispositivo de la reivindicación 9, que comprende además medios para calcular el primer valor como una diferencia entre un primer valor de ciclo de MSB para la primera imagen de referencia a largo plazo, $\text{DeltaPocMSBCycleLt}[i - 1]$, y el valor de MSB base; y

medios para calcular el segundo valor como una diferencia entre un segundo valor de ciclo de MSB para la segunda imagen de referencia a largo plazo, $\Delta_{\text{PocMSBCycleLt}}[i]$, y el primer valor de MSB.

- 5
11. El dispositivo de la reivindicación 7 o 9, en el que la primera imagen de referencia a largo plazo corresponde a un índice de j , en el que la segunda imagen de referencia a largo plazo corresponde a un índice de i , en el que el primer valor comprende $\Delta_{\text{poc_msb_cycle_lt_plus1}}[j]$, en el que el segundo valor comprende $\Delta_{\text{poc_msb_cycle_lt_plus1}}[i]$, y en el que j está en el rango de 0 a $i-1$, inclusive.
- 10
12. El dispositivo de la reivindicación 7 o 9, en el que al menos uno del primer valor y el segundo valor comprende un valor entero negativo.
- 15
13. Un medio de almacenamiento legible por ordenador que tiene instrucciones almacenadas en el mismo que, cuando son ejecutadas, hacen que un procesador realice el procedimiento según cualquiera de las reivindicaciones 1 a 6.

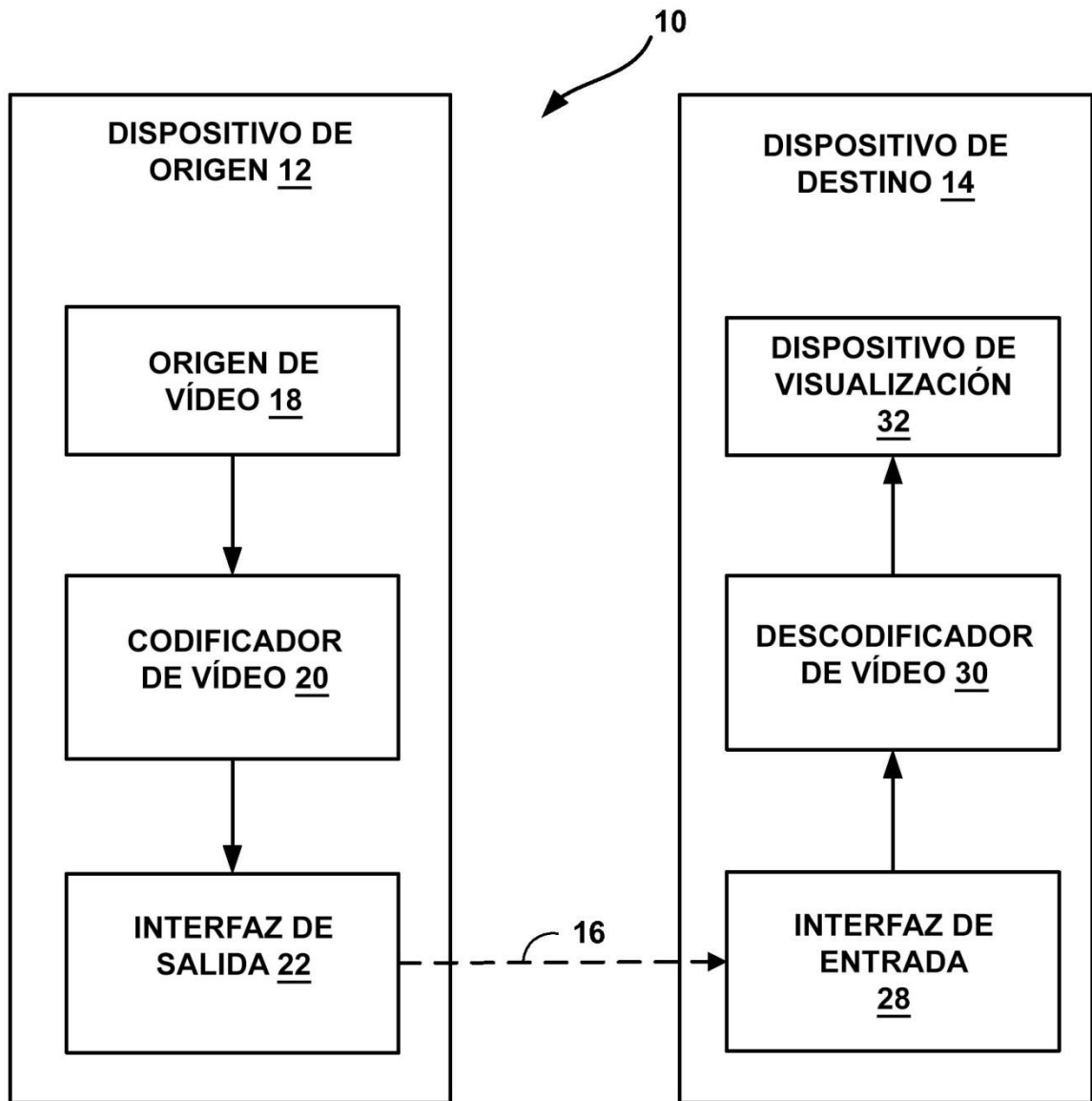


FIG. 1

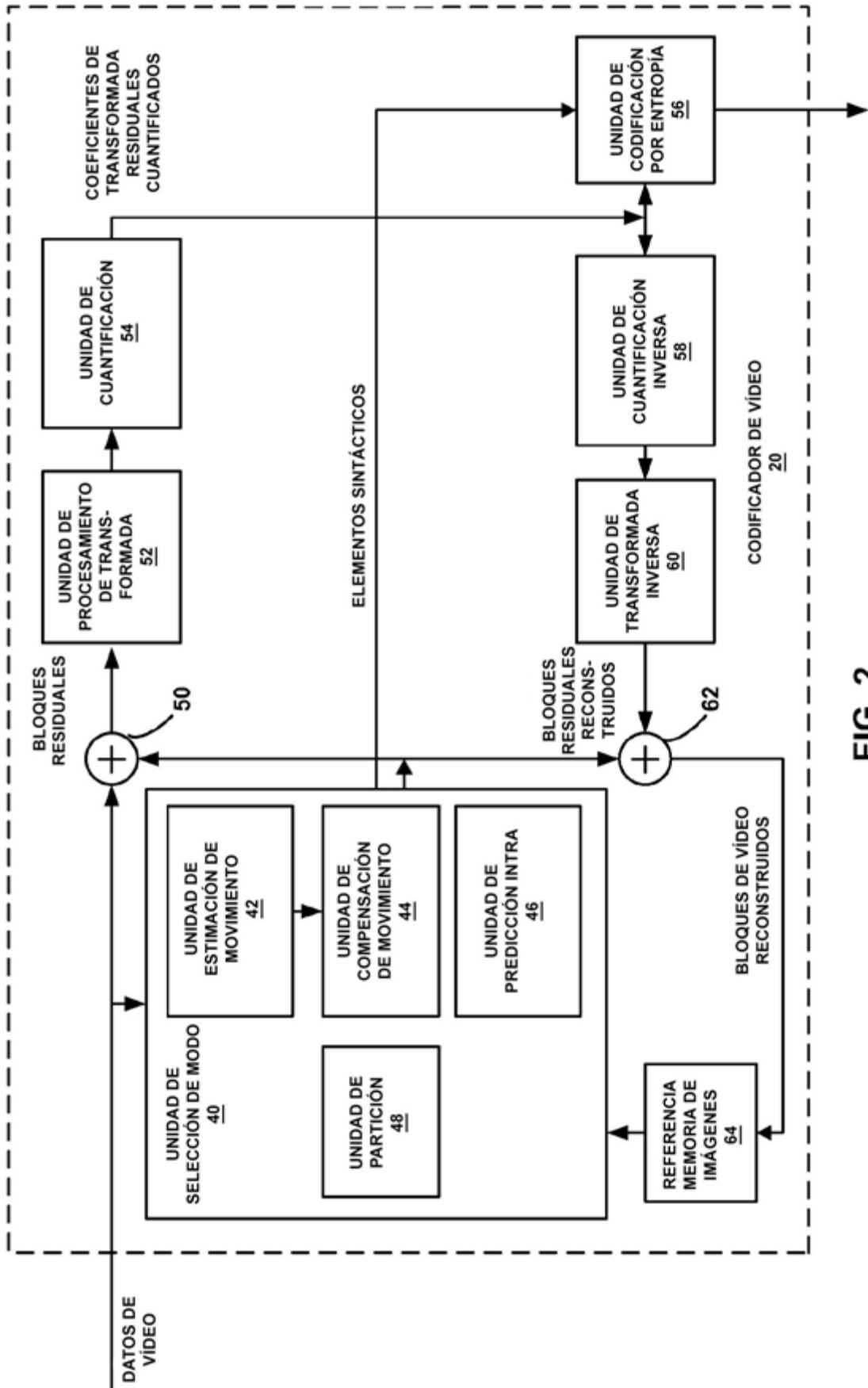


FIG. 2

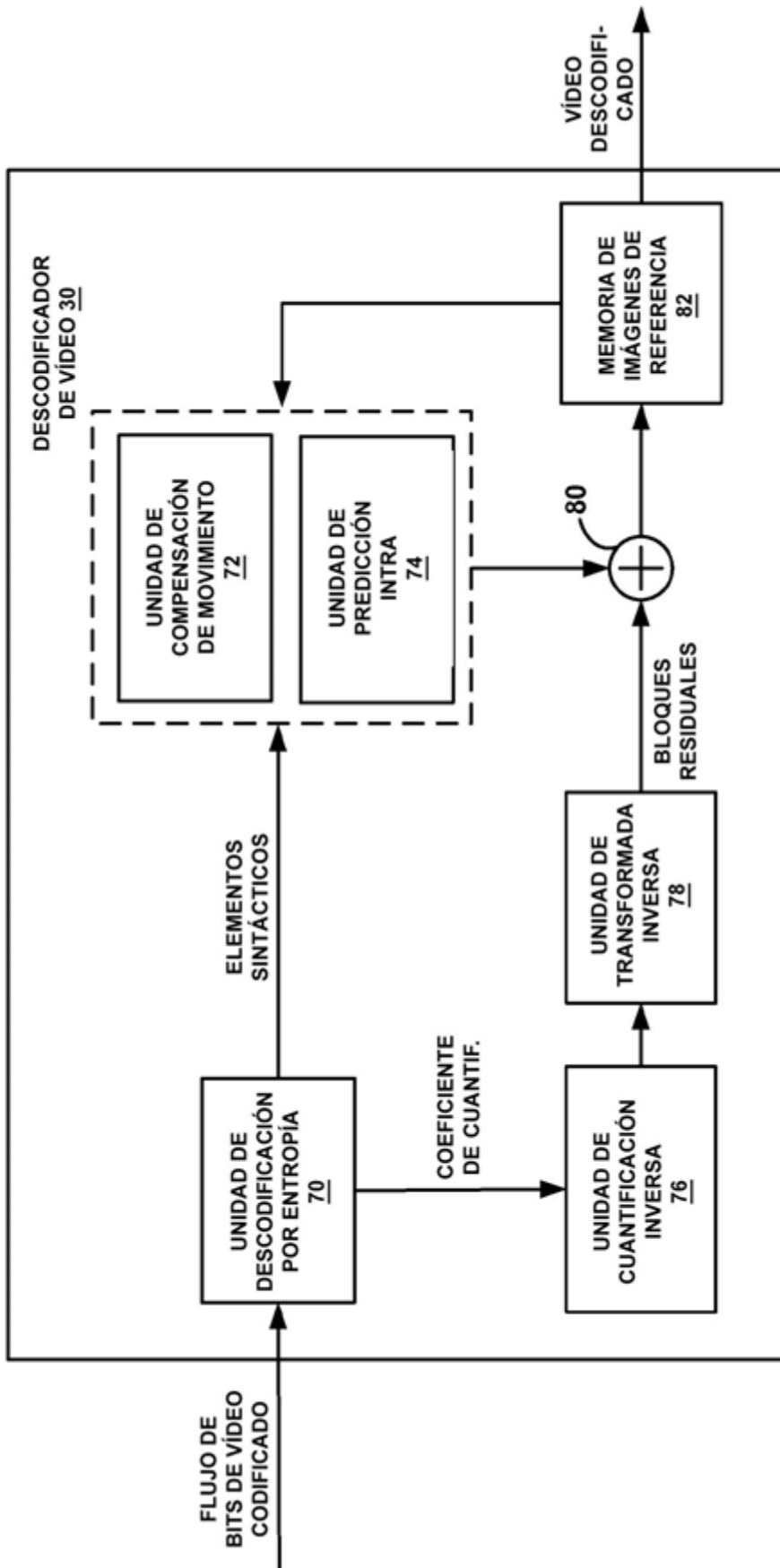


FIG. 3

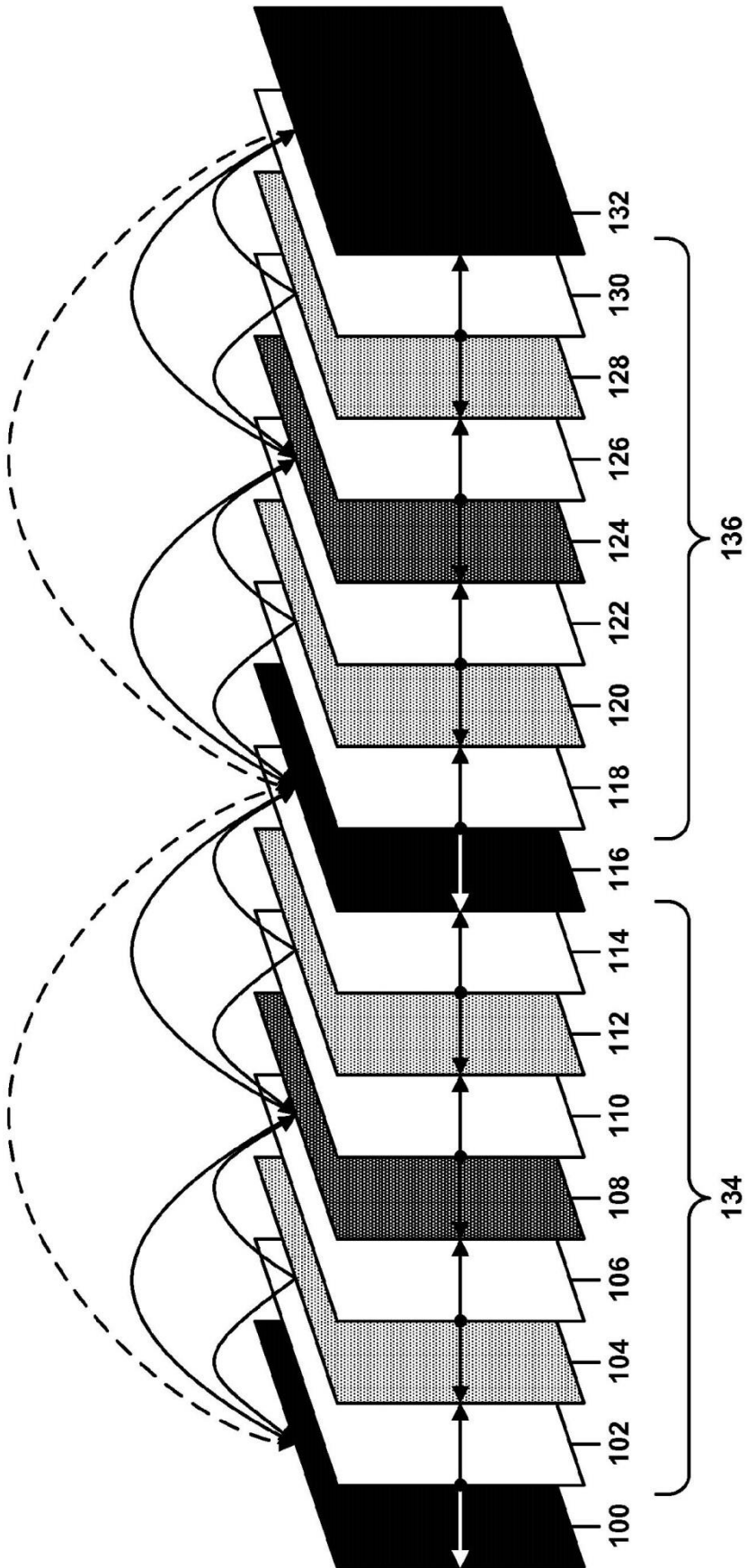


FIG. 4

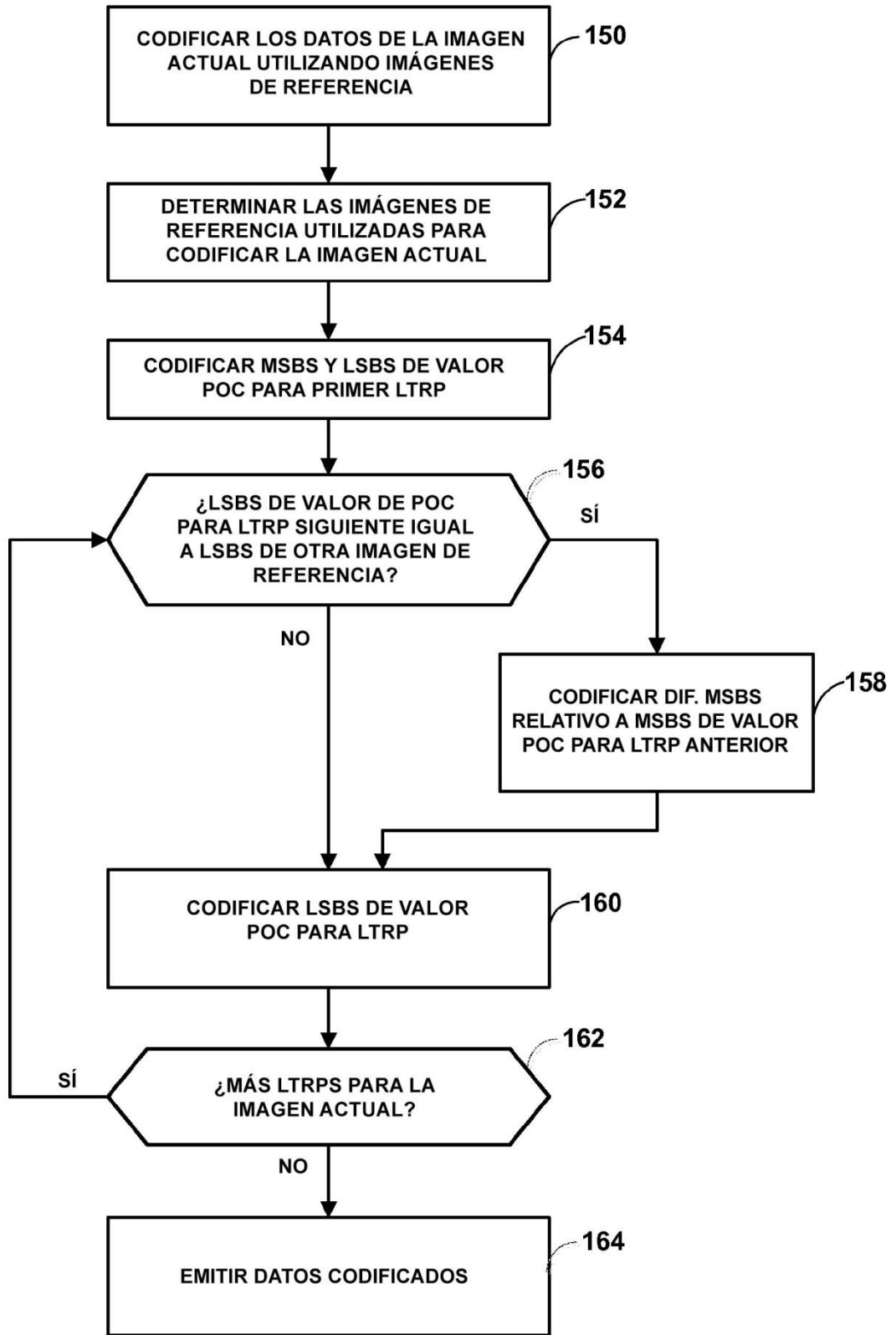


FIG. 5

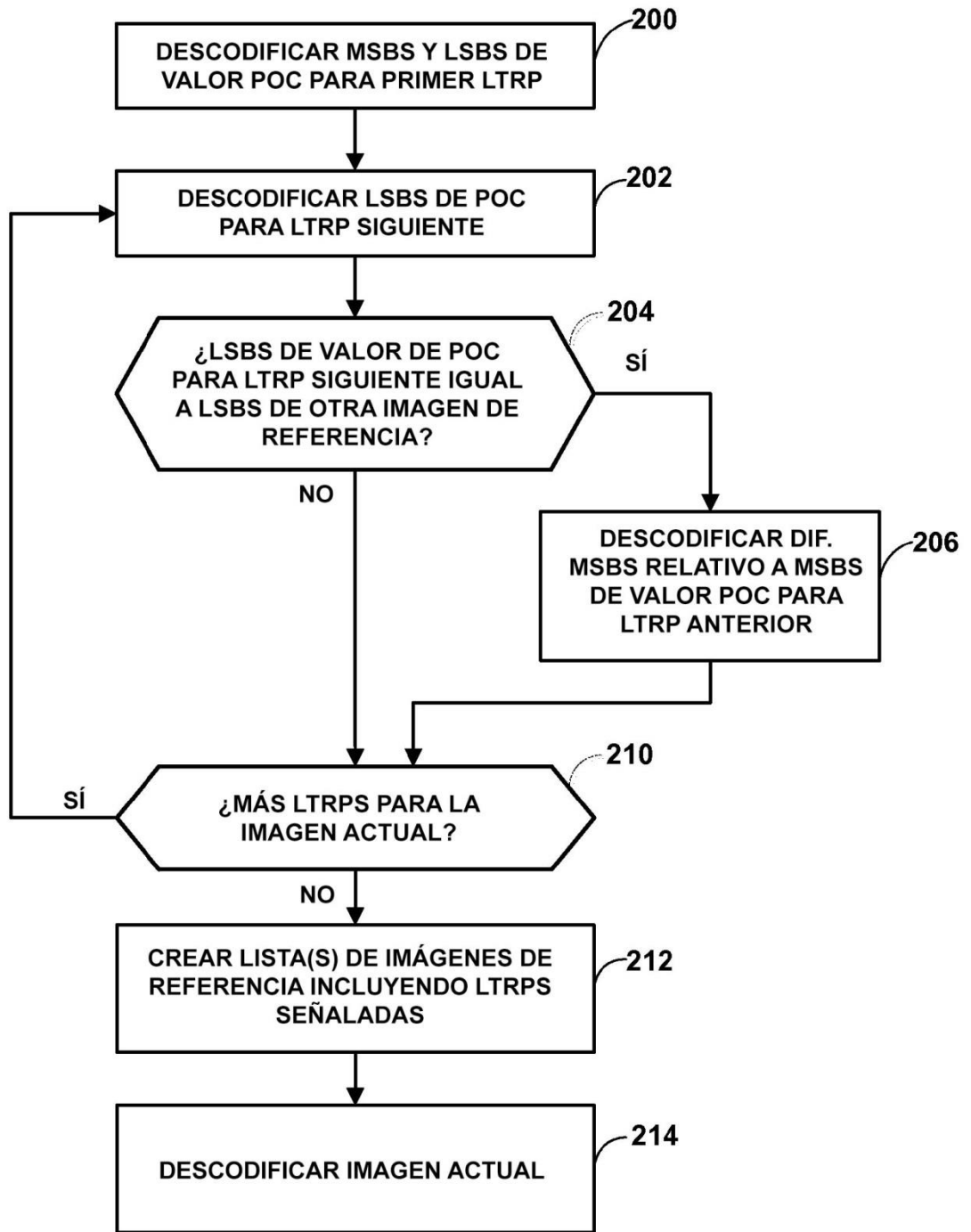


FIG. 6