

19



OFICINA ESPAÑOLA DE  
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 776 147**

51 Int. Cl.:

**H04N 19/70** (2014.01)

**H04N 19/50** (2014.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

86 Fecha de presentación y número de la solicitud internacional: **20.09.2012 PCT/US2012/056361**

87 Fecha y número de publicación internacional: **28.03.2013 WO13043886**

96 Fecha de presentación y número de la solicitud europea: **20.09.2012 E 12777970 (0)**

97 Fecha y número de publicación de la concesión europea: **18.12.2019 EP 2759132**

54 Título: **Construcción de lista de imágenes de referencia para codificación de video**

30 Prioridad:

**23.09.2011 US 201161538787 P**

**26.09.2011 US 201161539433 P**

**30.09.2011 US 201161542034 P**

**19.09.2012 US 201213622929**

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

**29.07.2020**

73 Titular/es:

**VELOS MEDIA INTERNATIONAL LIMITED**

**(100.0%)**

**Unit 32, the Hyde Building The Park,**

**Carrickmines**

**Dublin 18, IE**

72 Inventor/es:

**WANG, YE-KUI y**

**CHEN, YING**

74 Agente/Representante:

**MILTENYI , Peter**

ES 2 776 147 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

Construcción de lista de imágenes de referencia para codificación de video

## 5 CAMPO TÉCNICO

Esta divulgación se refiere a la codificación de video (video coding) y, más particularmente, a técnicas para codificar datos de video (coding video data).

## ANTECEDENTES

- 10 Se pueden incorporar capacidades de video digital en una amplia gama de dispositivos, que incluyen televisores digitales, sistemas de transmisión directa digital, sistemas de transmisión inalámbricos, asistentes digitales personales (PDA), ordenadores portátiles o de escritorio, ordenadores tableta, lectores de libros electrónicos, cámaras digitales, dispositivos de grabación digital, reproductores de medios digitales, dispositivos de videojuegos, consolas de videojuegos, teléfonos celulares o satelitales, los denominados "teléfonos inteligentes", dispositivos de video telefonía,
- 15 dispositivos de videoconferencia, dispositivos de streaming de video, y similares. Los dispositivos de video digital aplican técnicas de compresión de video, tales como las descritas en los estándares MPEG-2, MPEG-4, ITU-T H.263, ITU-T H.264/MPEG-4, Parte 10, Codificación avanzada de video (AVC), el estándar de codificación de video de alta eficiencia (HEVC) actualmente en desarrollo, y ampliaciones de dichos estándares. Los dispositivos de video pueden transmitir, recibir, codificar, decodificar y/o almacenar información de video digital de forma más eficiente
- 20 implementando dichas técnicas de compresión de video.

- Las técnicas de compresión de video realizan predicciones espaciales (intra imagen) y/o temporales (inter imagen) para reducir o eliminar la redundancia inherente a las secuencias de video. Para la codificación de video basada en bloques, un corte (slice) de video (es decir, una imagen de video o una parte de una imagen de video) se puede
- 25 particionar en bloques de video, que también se pueden denominar bloques de árbol, bloques de árbol de codificación (CTB: Coding Tree Blocks), unidades de árbol de codificación (CTU: Coding Tree Units), unidades de codificación (CU: Coding Units) y/o nodos de codificación. Los bloques de video en un corte (slice) intra codificado (I) de una imagen se codifican utilizando predicción espacial con respecto a muestras de referencia en bloques vecinos en la misma imagen. Los bloques de video en un corte inter codificada (P o B) de una imagen pueden utilizar la predicción
- 30 espacial con respecto a muestras de referencia en bloques vecinos en la misma imagen o la predicción temporal con respecto a muestras de referencia en otras imágenes de referencia. Las imágenes se pueden denominar fotogramas, y las imágenes de referencia se pueden denominar fotogramas de referencia. Por ejemplo, sjöberg ET AL ("Absolute signaling of reference pictures", documento de entrada a JCT-VC, documento: JCTVC-F493, WG11 número: m20923) propone procesos de imagen de referencia en el estándar HEVC describiendo imágenes de referencia en un modo
- 35 absoluto en la cabecera de corte (slice) de una imagen.

- La predicción espacial o temporal resulta en un bloque predictivo para un bloque a convertir según una codificación. Los datos residuales representan diferencias de píxeles entre el bloque original a codificar y el bloque predictivo. Un
- 40 bloque inter codificado es codificado según 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 convertido según codificación y el bloque predictivo. Un bloque intra codificado es codificado según un modo de intra codificación y los datos residuales. Para una mayor compresión, los datos residuales pueden ser transformados del dominio de píxeles a un dominio de transformación, resultando en coeficientes de transformación residuales, que entonces pueden ser
- 45 cuantificados. Los coeficientes de transformación cuantificados, dispuestos inicialmente en una matriz bidimensional, pueden ser escaneados para producir un vector unidimensional de coeficientes de transformación, y se puede aplicar codificación de entropía para conseguir aún más compresión.

## RESUMEN

- En general, esta divulgación describe técnicas relacionadas con derivar un conjunto de imágenes de referencia para
- 50 uso en la codificación de video (video coding), según definen las reivindicaciones adjuntas. Por ejemplo, el conjunto de imágenes de referencia puede constituir una combinación de una pluralidad de subconjuntos de imágenes de referencia. Cada uno de los subconjuntos de imágenes de referencia puede identificar una pluralidad de imágenes de referencia potenciales, pero menos que todas las imágenes de referencia potenciales. En técnicas de ejemplo que se describen en esta divulgación, un codificador de video (video coder), que puede ser un codificador de video (video
- 55 encoder) o un decodificador de video (video decoder), puede construir múltiples listas cada una de las cuales incluye identificadores de un subconjunto de las imágenes de referencia potenciales. A partir de estas múltiples listas, el codificador de video (video coder) puede construir la pluralidad de subconjuntos de imágenes de referencia, lo que produce como resultado que el codificador de video (video coder) derive el conjunto de imágenes de referencia.
- 60 Además de las técnicas relacionadas con derivar el conjunto de imágenes de referencia, esta divulgación describe técnicas simplificadas de inicialización de una lista de imágenes de referencia. Esta inicialización de la lista de imágenes de referencia puede eliminar la necesidad de reordenar las imágenes de referencia. Por ejemplo, si no es necesario modificar la lista de imágenes de referencia, entonces las listas iniciales de imágenes de referencia pueden

formar las listas finales de imágenes de referencia, y puede que no sean necesarias más reordenaciones. Las técnicas también pueden estar enfocadas a construir la lista de imágenes de referencia de tal manera que el codificador de video (video coder) añada repetidamente imágenes de referencia a la lista de imágenes de referencia hasta que las entradas de la lista de imágenes de referencia sean iguales al número máximo de entradas permisibles.

5

En algunos ejemplos, las técnicas están enfocadas a una modificación de la lista de imágenes de referencia. Por ejemplo, el codificador de video (video coder) puede modificar la lista inicial de imágenes de referencia referenciando a uno o más de los subconjuntos de imágenes de referencia, y añadiendo una o más imágenes de los subconjuntos de imágenes de referencia a la lista de imágenes de referencia después de la construcción de la lista inicial de

10

imágenes de referencia. En algunos ejemplos, el codificador de video (video coder) puede realizar una gestión de un búfer de imágenes decodificadas (DPB: decoded picture buffer). En estos ejemplos, el codificador de video (video coder) puede eliminar imágenes decodificadas del búfer DPB si la imagen decodificada no pertenece al conjunto de imágenes de referencia. En algunos casos, el codificador de video (video coder) puede eliminar la imagen decodificada antes de la conversión según codificación de la imagen actual.

15

En un ejemplo, la divulgación describe un procedimiento para codificar datos de video que incluye codificar información indicativa de imágenes de referencia que pertenecen a un conjunto de imágenes de referencia. En este ejemplo, el conjunto de imágenes de referencia identifica las imágenes de referencia que se pueden utilizar potencialmente para la inter predicción de una imagen actual y que se pueden utilizar potencialmente para la inter predicción de una o más imágenes posteriores a la imagen actual en un orden de decodificación. El procedimiento también incluye construir una pluralidad de subconjuntos de imágenes de referencia cada uno de los cuales identifica cero o más de las imágenes de referencia del conjunto de imágenes de referencia, y codificar la imagen actual en base a la pluralidad de subconjuntos de imágenes de referencia.

20

25

En un ejemplo, la divulgación describe un dispositivo para codificar datos de video. El dispositivo incluye un codificador de video (video coder) que está configurado para codificar información indicativa de imágenes de referencia que pertenecen a un conjunto de imágenes de referencia. En este ejemplo, el conjunto de imágenes de referencia identifica las imágenes de referencia que se pueden utilizar potencialmente para la inter predicción de una imagen actual y que se pueden utilizar potencialmente para la inter predicción de una o más imágenes posteriores a la imagen actual en un orden de decodificación. El codificador de video (video coder) también está configurado para construir una pluralidad de subconjuntos de imágenes de referencia cada uno de los cuales identifica cero o más de las imágenes de referencia del conjunto de imágenes de referencia, y codificar la imagen actual en base a la pluralidad de subconjuntos de imágenes de referencia.

30

35

En un ejemplo, la divulgación describe un medio de almacenamiento legible por ordenador que tiene almacenadas en el mismo unas instrucciones que, cuando son ejecutadas, hacen que un procesador de un dispositivo para codificar datos de video convierta según codificación información indicativa de imágenes de referencia que pertenecen a un conjunto de imágenes de referencia. En este ejemplo, el conjunto de imágenes de referencia identifica las imágenes de referencia que se pueden utilizar potencialmente para la inter predicción de una imagen actual y que se pueden utilizar potencialmente para la inter predicción de una o más imágenes posteriores a la imagen actual en un orden de decodificación. Las instrucciones también hacen que el procesador construya una pluralidad de subconjuntos de imágenes de referencia cada uno de los cuales identifica cero o más de las imágenes de referencia del conjunto de imágenes de referencia, y codificar la imagen actual en base a la pluralidad de subconjuntos de imágenes de referencia.

40

45

En un ejemplo, la divulgación describe un dispositivo para codificar datos de video. El dispositivo incluye un medio para codificar información indicativa de imágenes de referencia que pertenecen a un conjunto de imágenes de referencia. En este ejemplo, el conjunto de imágenes de referencia identifica las imágenes de referencia que se pueden utilizar potencialmente para la inter predicción de una imagen actual y que se pueden utilizar potencialmente para la inter predicción de una o más imágenes posteriores a la imagen actual en un orden de decodificación. El dispositivo también incluye un medio para construir una pluralidad de subconjuntos de imágenes de referencia cada uno de los cuales identifica cero o más de las imágenes de referencia del conjunto de imágenes de referencia, y un medio para codificar la imagen actual en base a la pluralidad de subconjuntos de imágenes de referencia.

50

55

En un ejemplo, la divulgación describe un procedimiento para codificar datos de video, el procedimiento incluye codificar información indicativa de imágenes de referencia que pertenecen a un conjunto de imágenes de referencia. En este ejemplo, el conjunto de imágenes de referencia identifica las imágenes de referencia que se pueden utilizar potencialmente para la inter predicción de una imagen actual y que se pueden utilizar potencialmente para la inter predicción de una o más imágenes posteriores a la imagen actual en un orden de decodificación. El procedimiento también incluye construir una pluralidad de subconjuntos de imágenes de referencia cada uno de los cuales identifica cero o más de las imágenes de referencia del conjunto de imágenes de referencia, añadir imágenes de referencia de

60

un primer subconjunto de la pluralidad de subconjuntos de imágenes de referencia, seguido de imágenes de referencia de un segundo subconjunto de la pluralidad de subconjuntos de imágenes de referencia, y seguido de imágenes de referencia de un tercer subconjunto de la pluralidad de subconjuntos de imágenes de referencia a una lista de imágenes de referencia, mientras el número de entradas en la lista de imágenes de referencia no sea superior a un número máximo de entradas permisibles en la lista de imágenes de referencia, y codificar la imagen actual en base a la lista de imágenes de referencia.

En un ejemplo, la divulgación describe un dispositivo para codificar datos de video. El dispositivo incluye un codificador de video (video coder) configurado para codificar información indicativa de imágenes de referencia que pertenecen a un conjunto de imágenes de referencia. En este ejemplo, el conjunto de imágenes de referencia identifica las imágenes de referencia que se pueden utilizar potencialmente para la inter predicción de una imagen actual y que se pueden utilizar potencialmente para la inter predicción de una o más imágenes posteriores a la imagen actual en un orden de decodificación. El codificador de video (video coder) también está configurado para construir una pluralidad de subconjuntos de imágenes de referencia cada uno de los cuales identifica cero o más de las imágenes de referencia del conjunto de imágenes de referencia, añadir imágenes de referencia de un primer subconjunto de la pluralidad de subconjuntos de imágenes de referencia, seguido de imágenes de referencia de un segundo subconjunto de la pluralidad de subconjuntos de imágenes de referencia, y seguido de imágenes de referencia de un tercer subconjunto de la pluralidad de subconjuntos de imágenes de referencia a una lista de imágenes de referencia, siempre y cuando un número de entradas en la lista de imágenes de referencia no sea superior a un número máximo de entradas permisibles en la lista de imágenes de referencia, y codificar la imagen de referencia actual en base a la lista de imágenes de referencia.

En un ejemplo, la divulgación describe un medio de almacenamiento legible por ordenador que tiene almacenadas en el mismo unas instrucciones que, cuando son ejecutadas, hacen que un procesador de un dispositivo para codificar datos de video convierta según codificación información indicativa de imágenes de referencia que pertenecen a un conjunto de imágenes de referencia. En este ejemplo, el conjunto de imágenes de referencia identifica las imágenes de referencia que se pueden utilizar potencialmente para la inter predicción de una imagen actual y que se pueden utilizar potencialmente para la inter predicción de una o más imágenes posteriores a la imagen actual en un orden de decodificación. Las instrucciones también hacen que el procesador construya una pluralidad de subconjuntos de imágenes de referencia cada uno de los cuales identifica cero o más de las imágenes de referencia del conjunto de imágenes de referencia, añade imágenes de referencia de un primer subconjunto de la pluralidad de subconjuntos de imágenes de referencia, seguido de imágenes de referencia de un segundo subconjunto de la pluralidad de subconjuntos de imágenes de referencia, y seguido de imágenes de referencia de un tercer subconjunto de la pluralidad de subconjuntos de imágenes de referencia a una lista de imágenes de referencia, mientras el número de entradas en la lista de imágenes de referencia no sea superior a un máximo de entradas permisibles en la lista de imágenes de referencia, y codificar la imagen actual en base a la lista de imágenes de referencia.

En un ejemplo, la divulgación describe un dispositivo para codificar datos de video. El dispositivo incluye un medio para codificar información indicativa de imágenes de referencia que pertenecen a un conjunto de imágenes de referencia. En este ejemplo, el conjunto de imágenes de referencia identifica las imágenes de referencia que se pueden utilizar potencialmente para la inter predicción de una imagen actual y que se pueden utilizar potencialmente para la inter predicción de una o más imágenes posteriores a la imagen actual en un orden de decodificación. El dispositivo también incluye un medio para construir una pluralidad de subconjuntos de imágenes de referencia cada uno de los cuales identifica cero o más de las imágenes de referencia del conjunto de imágenes de referencia, un medio para añadir imágenes de referencia de un primer subconjunto de la pluralidad de subconjuntos de imágenes de referencia, seguido de imágenes de referencia de un segundo subconjunto de la pluralidad de subconjuntos de imágenes de referencia, y seguido de imágenes de referencia de un tercer subconjunto de la pluralidad de subconjuntos de imágenes de referencia a una lista de imágenes de referencia, mientras el número de entradas en la lista de imágenes de referencia no sea superior a un máximo de entradas permisibles en la lista de imágenes de referencia, y un medio para codificar la imagen de referencia actual en base a la lista de imágenes de referencia.

En un ejemplo, la divulgación describe un procedimiento para codificar datos de video, el procedimiento incluye codificar información indicativa de imágenes de referencia que pertenecen a un conjunto de imágenes de referencia. En este ejemplo, el conjunto de imágenes de referencia identifica las imágenes de referencia que se pueden utilizar potencialmente para la inter predicción de una imagen actual y que se pueden utilizar potencialmente para la inter predicción de una o más imágenes posteriores a la imagen actual en un orden de decodificación. El procedimiento también incluye construir una pluralidad de subconjuntos de imágenes de referencia cada uno de los cuales identifica cero o más de las imágenes de referencia del conjunto de imágenes de referencia, añadir imágenes de referencia de la pluralidad de subconjuntos de imágenes de referencia a un primer conjunto de entradas de una lista de imágenes de referencia, determinar si un número de entradas en la lista de imágenes de referencia es igual a un número máximo de entradas permisibles en la lista de imágenes de referencia, cuando el número de entradas en la lista de imágenes de referencia no es igual al número máximo de entradas permisibles en la lista de imágenes de referencia, volver a añadir repetidamente una o más imágenes de referencia de al menos uno de entre la pluralidad de subconjuntos de

imágenes de referencia en entradas de la lista de imágenes de referencia que son subsiguientes respecto al primer conjunto de entradas hasta que el número de entradas en la lista de imágenes de referencia sea igual al número máximo de entradas permisibles en la lista de imágenes de referencia, y codificar la imagen actual en base a la lista de imágenes de referencia.

5

En un ejemplo, la divulgación describe un dispositivo para codificar datos de video. El dispositivo incluye un codificador de video (video coder) configurado para codificar información indicativa de imágenes de referencia que pertenecen a un conjunto de imágenes de referencia. En este ejemplo, el conjunto de imágenes de referencia identifica las imágenes de referencia que se pueden utilizar potencialmente para la inter predicción de una imagen actual y que se pueden utilizar potencialmente para la inter predicción de una o más imágenes posteriores a la imagen actual en un orden de decodificación. El codificador de video (video coder) también está configurado para construir una pluralidad de subconjuntos de imágenes de referencia cada uno de los cuales identifica cero o más de las imágenes de referencia del conjunto de imágenes de referencia, añadir imágenes de referencia de la pluralidad de subconjuntos de imágenes de referencia a un primer conjunto de entradas de una lista de imágenes de referencia, determinar si un número de entradas en la lista de imágenes de referencia es igual a un número máximo de entradas permisibles en la lista de imágenes de referencia, cuando el número de entradas en la lista de imágenes de referencia no es igual al número máximo de entradas permisibles en la lista de imágenes de referencia, volver a añadir repetidamente una o más imágenes de referencia de al menos uno de entre la pluralidad de subconjuntos de imágenes de referencia en entradas de la lista de imágenes de referencia que son subsiguientes respecto al primer conjunto de entradas hasta que el número de entradas en la lista de imágenes de referencia sea igual al número máximo de entradas permisibles en la lista de imágenes de referencia, y codificar la imagen actual en base a la lista de imágenes de referencia.

En un ejemplo, la divulgación describe un medio de almacenamiento legible por ordenador que tiene almacenadas en el mismo unas instrucciones que, cuando son ejecutadas, hacen que un procesador de un dispositivo para codificar datos de video convierta según codificación información indicativa de imágenes de referencia que pertenecen a un conjunto de imágenes de referencia. En este ejemplo, el conjunto de imágenes de referencia identifica las imágenes de referencia que se pueden utilizar potencialmente para la inter predicción de una imagen actual y que se pueden utilizar potencialmente para la inter predicción de una o más imágenes posteriores a la imagen actual en un orden de decodificación. Las instrucciones también hacen que el procesador construya una pluralidad de subconjuntos de imágenes de referencia cada uno de los cuales identifica cero o más de las imágenes de referencia del conjunto de imágenes de referencia, añada imágenes de referencia de la pluralidad de subconjuntos de imágenes de referencia a un primer conjunto de entradas de una lista de imágenes de referencia, determine si un número de entradas en la lista de imágenes de referencia es igual a un número máximo de entradas permisibles en la lista de imágenes de referencia, cuando el número de entradas en la lista de imágenes de referencia no es igual al número máximo de entradas permisibles en la lista de imágenes de referencia, volver a añadir repetidamente una o más imágenes de referencia de al menos uno de entre la pluralidad de subconjuntos de imágenes de referencia en entradas de la lista de imágenes de referencia que son subsiguientes respecto al primer conjunto de entradas hasta que el número de entradas en la lista de imágenes de referencia sea igual al número máximo de entradas permisibles en la lista de imágenes de referencia, y codificar la imagen actual en base a la lista de imágenes de referencia.

40

En un ejemplo, la divulgación describe un dispositivo para codificar datos de video. El dispositivo incluye un medio para codificar información indicativa de imágenes de referencia que pertenecen a un conjunto de imágenes de referencia. En este ejemplo, el conjunto de imágenes de referencia identifica las imágenes de referencia que se pueden utilizar potencialmente para la inter predicción de una imagen actual y que se pueden utilizar potencialmente para la inter predicción de una o más imágenes posteriores a la imagen actual en un orden de decodificación. El dispositivo también incluye un medio para construir una pluralidad de subconjuntos de imágenes de referencia cada uno de los cuales identifica cero o más de las imágenes de referencia del conjunto de imágenes de referencia, un medio para añadir imágenes de referencia de la pluralidad de subconjuntos de imágenes de referencia a un primer conjunto de entradas de una lista de imágenes de referencia, un medio para determinar si un número de entradas en la lista de imágenes de referencia es igual a un número máximo de entradas permisibles en la lista de imágenes de referencia, cuando el número de entradas en la lista de imágenes de referencia no es igual al número máximo de entradas permisibles en la lista de imágenes de referencia, un medio para volver a añadir repetidamente una o más imágenes de referencia de al menos uno de entre la pluralidad de subconjuntos de imágenes de referencia en entradas de la lista de imágenes de referencia que son subsiguientes respecto al primer conjunto de entradas hasta que el número de entradas en la lista de imágenes de referencia sea igual al número máximo de entradas permisibles en la lista de imágenes de referencia, y un medio para codificar la imagen actual en base a la lista de imágenes de referencia.

En un ejemplo, la divulgación describe un procedimiento para codificar datos de video, el procedimiento incluye codificar información indicativa de imágenes de referencia que pertenecen a un conjunto de imágenes de referencia.

En este ejemplo, el conjunto de imágenes de referencia identifica las imágenes de referencia que se pueden utilizar potencialmente para la inter predicción de una imagen actual y que se pueden utilizar potencialmente para la inter predicción de una o más imágenes posteriores a la imagen actual en un orden de decodificación. El procedimiento también incluye construir una pluralidad de subconjuntos de imágenes de referencia cada uno de los cuales identifica

cero o más de las imágenes de referencia del conjunto de imágenes de referencia, construir una lista inicial de imágenes de referencia en base a los subconjuntos de imágenes de referencia construidos y, cuando sea necesario modificar una imagen de referencia, identificar una imagen de referencia en al menos uno de entre la pluralidad de subconjuntos de imágenes de referencia construidos, y añadir la imagen de referencia identificada en una entrada  
5 actual de la imagen de referencia inicial para construir una lista de imágenes de referencia modificada. El procedimiento incluye además codificar la imagen actual en base a la lista de imágenes de referencia modificada.

En un ejemplo, la divulgación describe un dispositivo para codificar datos de video. El dispositivo incluye un codificador de video (video coder) configurado para codificar información indicativa de imágenes de referencia que pertenecen a  
10 un conjunto de imágenes de referencia. En este ejemplo, el conjunto de imágenes de referencia identifica las imágenes de referencia que se pueden utilizar potencialmente para la inter predicción de una imagen actual y que se pueden utilizar potencialmente para la inter predicción de una o más imágenes posteriores a la imagen actual en un orden de decodificación. El codificador de video (video coder) también está configurado para construir una pluralidad de subconjuntos de imágenes de referencia cada uno de los cuales identifica cero o más de las imágenes de referencia  
15 del conjunto de imágenes de referencia, construir una lista inicial de imágenes de referencia en base a los subconjuntos de imágenes de referencia construidos y, cuando sea necesario modificar una imagen de referencia, identificar una imagen de referencia en al menos uno de entre la pluralidad de subconjuntos de imágenes de referencia construidos y añadir la imagen de referencia identificada en una entrada actual de la imagen de referencia inicial para construir una lista de imágenes de referencia modificada. El codificador de video (video coder) también está configurado para  
20 codificar la imagen actual en base a la lista de imágenes de referencia modificada.

En un ejemplo, la divulgación describe un medio de almacenamiento legible por ordenador que tiene almacenadas en el mismo unas instrucciones que, cuando son ejecutadas, hacen que un procesador de un dispositivo para codificar datos de video convierta según codificación información indicativa de imágenes de referencia que pertenecen a un  
25 conjunto de imágenes de referencia. En este ejemplo, el conjunto de imágenes de referencia identifica las imágenes de referencia que se pueden utilizar potencialmente para la inter predicción de una imagen actual y que se pueden utilizar potencialmente para la inter predicción de una o más imágenes posteriores a la imagen actual en un orden de decodificación. Las instrucciones también hacen que el procesador construya una pluralidad de subconjuntos de imágenes de referencia cada uno de los cuales identifica cero o más de las imágenes de referencia del conjunto de  
30 imágenes de referencia, construya una lista inicial de imágenes de referencia en base a los subconjuntos de imágenes de referencia construidos y, cuando sea necesario modificar una imagen de referencia, identifique una imagen de referencia en al menos uno de entre la pluralidad de subconjuntos de imágenes de referencia construidos y añada la imagen de referencia identificada en una entrada actual de la imagen de referencia inicial para construir una lista de imágenes de referencia modificada. Las instrucciones también hacen que el procesador convierta según codificación  
35 la imagen actual en base a la lista de imágenes de referencia modificada.

En un ejemplo, la divulgación describe un dispositivo para codificar datos de video. El dispositivo incluye un medio para codificar información indicativa de imágenes de referencia que pertenecen a un conjunto de imágenes de referencia. En este ejemplo, el conjunto de imágenes de referencia identifica las imágenes de referencia que se pueden  
40 utilizar potencialmente para la inter predicción de una imagen actual y que se pueden utilizar potencialmente para la inter predicción de una o más imágenes posteriores a la imagen actual en un orden de decodificación. El dispositivo también incluye un medio para construir una pluralidad de subconjuntos de imágenes de referencia cada uno de los cuales identifica cero o más de las imágenes de referencia del conjunto de imágenes de referencia, un medio para construir una lista inicial de imágenes de referencia en base a los subconjuntos de imágenes de referencia construidos  
45 y, cuando sea necesario modificar una imagen de referencia, un medio para identificar una imagen de referencia en al menos uno de entre la pluralidad de subconjuntos de imágenes de referencia construidos, y un medio para añadir la imagen de referencia identificada en una entrada actual de la imagen de referencia inicial para construir una lista de imágenes de referencia modificada. El dispositivo también incluye un medio para codificar la imagen actual en base a la lista de imágenes de referencia modificada.  
50

En un ejemplo, la divulgación describe un procedimiento para codificar datos de video, el procedimiento incluye codificar información indicativa de imágenes de referencia que pertenecen a un conjunto de imágenes de referencia. En este ejemplo, el conjunto de imágenes de referencia identifica las imágenes de referencia que se pueden utilizar potencialmente para la inter predicción de una imagen actual y que se pueden utilizar potencialmente para la inter  
55 predicción de una o más imágenes posteriores a la imagen actual en un orden de decodificación. El procedimiento incluye derivar el conjunto de imágenes de referencia en base a la información convertida según codificación, determinar si una imagen decodificada que está almacenada en un búfer de imágenes decodificadas (DPB: decoded picture buffer) no es necesaria para el output y no está identificada en el conjunto de imágenes de referencia, cuando la imagen decodificada no es necesaria para el output y no está identificada en el conjunto de imágenes de referencia,  
60 eliminar la imagen decodificada del búfer DPB y, una vez eliminada la imagen decodificada, codificar la imagen actual.

En un ejemplo, la divulgación describe un dispositivo para codificar datos de video. El dispositivo incluye un codificador de video (video coder) configurado para codificar información indicativa de imágenes de referencia que pertenecen a

un conjunto de imágenes de referencia. En este ejemplo, el conjunto de imágenes de referencia identifica las imágenes de referencia que se pueden utilizar potencialmente para la inter predicción de una imagen actual y que se pueden utilizar potencialmente para la inter predicción de una o más imágenes posteriores a la imagen actual en un orden de decodificación. El codificador de video (video coder) también está configurado para derivar el conjunto de imágenes de referencia en base a la información convertida según codificación, determinar si una imagen decodificada que está almacenada en un búfer de imágenes decodificadas (DPB) no es necesaria para el output y no está identificada en el conjunto de imágenes de referencia, cuando la imagen decodificada no es necesaria para el output y no está identificada en el conjunto de imágenes de referencia, eliminar la imagen decodificada del búfer DPB y, una vez eliminada la imagen decodificada, codificar la imagen actual.

10

En un ejemplo, la divulgación describe un medio de almacenamiento legible por ordenador que tiene almacenadas en el mismo unas instrucciones que, cuando son ejecutadas, hacen que un procesador de un dispositivo para codificar datos de video convierta según codificación información indicativa de imágenes de referencia que pertenecen a un conjunto de imágenes de referencia. En este ejemplo, el conjunto de imágenes de referencia identifica las imágenes de referencia que se pueden utilizar potencialmente para la inter predicción de una imagen actual y que se pueden utilizar potencialmente para la inter predicción de una o más imágenes posteriores a la imagen actual en un orden de decodificación. Las instrucciones también hacen que el procesador derive el conjunto de imágenes de referencia en base a la información convertida según codificación, determine si una imagen decodificada que está almacenada en un búfer de imágenes decodificadas (DPB) no es necesaria para el output y no está identificada en el conjunto de imágenes de referencia, cuando la imagen decodificada no es necesaria para el output y no está identificada en el conjunto de imágenes de referencia, elimine la imagen decodificada del búfer DPB y, una vez eliminada la imagen decodificada, convierta según codificación la imagen actual.

15

20

En un ejemplo, la divulgación describe un dispositivo para codificar datos de video. El dispositivo incluye un medio para codificar información indicativa de imágenes de referencia que pertenecen a un conjunto de imágenes de referencia. En este ejemplo, el conjunto de imágenes de referencia identifica las imágenes de referencia que se pueden utilizar potencialmente para la inter predicción de una imagen actual y que se pueden utilizar potencialmente para la inter predicción de una o más imágenes posteriores a la imagen actual en un orden de decodificación. El dispositivo también incluye un medio para derivar el conjunto de imágenes de referencia en base a la información convertida según codificación, un medio para determinar si una imagen decodificada que está almacenada en un búfer de imágenes decodificadas (DPB) no es necesaria para el output y no está identificada en el conjunto de imágenes de referencia, cuando la imagen decodificada no es necesaria para el output y no está identificada en el conjunto de imágenes de referencia, un medio para eliminar la imagen decodificada del búfer DPB y, una vez eliminada la imagen decodificada, un medio para codificar la imagen actual.

35

En un ejemplo, la divulgación describe un procedimiento de conversión según codificación de datos de video, el procedimiento incluye codificar elementos de sintaxis que indican imágenes de referencia de largo plazo candidatas identificadas en un conjunto de parámetros. En este ejemplo, una o más de las imágenes de referencia de largo plazo candidatas pertenecen a un conjunto de imágenes de referencia de una imagen actual. Además, en este ejemplo, el conjunto de imágenes de referencia identifica las imágenes de referencia que se pueden utilizar potencialmente para la inter predicción de la imagen actual y que se pueden utilizar potencialmente para la inter predicción de una o más imágenes posteriores a la imagen actual en un orden de decodificación. El procedimiento también incluye codificar elementos de sintaxis que indican qué imágenes de referencia de largo plazo candidatas, identificadas en el conjunto de parámetros, pertenecen al conjunto de imágenes de referencia de la imagen actual, y construir al menos uno de una pluralidad de subconjuntos de imágenes de referencia en base a la indicación de qué imágenes de referencia de largo plazo candidatas pertenecen al conjunto de imágenes de referencia de la imagen actual. En este ejemplo, la pluralidad de subconjuntos de imágenes de referencia forma el conjunto de imágenes de referencia.

40

45

En un ejemplo, la divulgación describe un dispositivo para codificar datos de video. El dispositivo incluye un codificador de video (video coder) configurado para codificar elementos de sintaxis que indican imágenes de referencia de largo plazo candidatas identificadas en un conjunto de parámetros. En este ejemplo, una o más de las imágenes de referencia de largo plazo candidatas pertenecen a un conjunto de imágenes de referencia de una imagen actual. Además, en este ejemplo, el conjunto de imágenes de referencia identifica imágenes de referencia que se pueden utilizar potencialmente para la inter predicción de la imagen actual y que se pueden utilizar potencialmente para la inter predicción de una o más imágenes posteriores a la imagen actual en un orden de decodificación. El codificador de video (video coder) también está configurado para codificar elementos de sintaxis que indican qué imágenes de referencia de largo plazo candidatas, identificadas en el conjunto de parámetros, pertenecen al conjunto de imágenes de referencia de la imagen actual, y construir al menos uno de una pluralidad de subconjuntos de imágenes de referencia en base a la indicación de qué imágenes de referencia de largo plazo candidatas pertenecen al conjunto de imágenes de referencia de la imagen actual. En este ejemplo, la pluralidad de subconjuntos de imágenes de referencia forma el conjunto de imágenes de referencia.

50

55

60

En un ejemplo, la divulgación describe un medio de almacenamiento legible por ordenador que tiene almacenadas en el mismo unas instrucciones que, cuando son ejecutadas, hacen que un procesador de un dispositivo para codificar datos de video convierta según codificación elementos de sintaxis que indican imágenes de referencia de largo plazo candidatas identificadas en un conjunto de parámetros. En este ejemplo, una o más de las imágenes de referencia de largo plazo candidatas pertenecen a un conjunto de imágenes de referencia de una imagen actual. Además, en este ejemplo, el conjunto de imágenes de referencia identifica las imágenes de referencia que se pueden utilizar potencialmente para la inter predicción de la imagen actual y que se pueden utilizar potencialmente para la inter predicción de una o más imágenes posteriores a la imagen actual en un orden de decodificación. Las instrucciones también hacen que el procesador convierta según codificación elementos de sintaxis que indican qué imágenes de referencia de largo plazo candidatas, identificadas en el conjunto de parámetros, pertenecen al conjunto de imágenes de referencia de la imagen actual, y construya al menos uno de una pluralidad de subconjuntos de imágenes de referencia en base a la indicación de qué imágenes de referencia de largo plazo candidatas pertenecen al conjunto de imágenes de referencia de la imagen actual. En este ejemplo, la pluralidad de subconjuntos de imágenes de referencia forma el conjunto de imágenes de referencia.

En un ejemplo, la divulgación describe un dispositivo para codificar datos de video. El dispositivo incluye un medio para codificar elementos de sintaxis que indican imágenes de referencia de largo plazo candidatas identificadas en un conjunto de parámetros. En este ejemplo, una o más de las imágenes de referencia de largo plazo candidatas pertenecen a un conjunto de imágenes de referencia de una imagen actual. Además, en este ejemplo, el conjunto de imágenes de referencia identifica las imágenes de referencia que se pueden utilizar potencialmente para la inter predicción de la imagen actual y que se pueden utilizar potencialmente para la inter predicción de una o más imágenes posteriores a la imagen actual en un orden de decodificación. El dispositivo también incluye un medio para codificar elementos de sintaxis que indican qué imágenes de referencia de largo plazo candidatas, identificadas en el conjunto de parámetros, pertenecen al conjunto de imágenes de referencia de la imagen actual, y un medio para construir al menos uno de una pluralidad de subconjuntos de imágenes de referencia en base a la indicación de qué imágenes de referencia de largo plazo candidatas pertenecen al conjunto de imágenes de referencia de la imagen actual. En este ejemplo, la pluralidad de subconjuntos de imágenes de referencia forma el conjunto de imágenes de referencia.

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

#### BREVE DESCRIPCIÓN DE LOS DIBUJOS

La figura 1 es un diagrama de bloques que ilustra un ejemplo de sistema de codificación y decodificación de video que puede utilizar las técnicas que se describen en esta divulgación.

La figura 2 es un diagrama conceptual que ilustra una secuencia de video de ejemplo que incluye una pluralidad de imágenes que son codificadas y transmitidas.

La figura 3 es un diagrama de bloques que ilustra un codificador de video (video encoder) de ejemplo que puede implementar las técnicas que se describen en esta divulgación.

La figura 4 es un diagrama de bloques que ilustra un ejemplo de decodificador de video que puede implementar las técnicas que se describen en esta divulgación.

La figura 5 es un diagrama de flujo que ilustra un ejemplo de operación de derivar un conjunto de imágenes de referencia.

La figura 6 es un diagrama de flujo que ilustra un ejemplo de operación de construcción de una lista de imágenes de referencia.

La figura 7 es un diagrama de flujo que ilustra otro ejemplo de operación de construcción de una lista de imágenes de referencia.

La figura 8 es un diagrama de flujo que ilustra un ejemplo de operación de modificación de una lista inicial de imágenes de referencia.

La figura 9 es un diagrama de flujo que ilustra un ejemplo de operación de eliminar imágenes decodificadas.

La figura 10 es un diagrama de flujo que ilustra una operación de ejemplo para determinar qué imágenes de referencia de largo plazo pertenecen al conjunto de imágenes de referencia de una imagen actual.

#### DESCRIPCIÓN DETALLADA



Las técnicas de esta divulgación están generalmente enfocadas a la gestión de imágenes de referencia que se utilizan para la inter predicción. Por ejemplo, un codificador de video (video coder) (por ejemplo, un codificador de video (video encoder) o un decodificador de video) incluye un búfer de imágenes decodificadas (DPB). El búfer DPB almacena imágenes decodificadas, que incluyen imágenes de referencia. Las imágenes de referencia son imágenes que se pueden utilizar para la inter predicción de una imagen. En otras palabras, el codificador de video (video coder) puede pronosticar una imagen, durante la conversión según codificación (codificación o decodificación) de esa imagen, en base a una o más imágenes de referencia almacenadas en el búfer DPB.

Para utilizar eficientemente el búfer DPB, se puede especificar el proceso de gestión del búfer DPB, tal como el proceso de almacenamiento de imágenes decodificadas en el búfer DPB, el proceso de marcado de las imágenes de referencia, el proceso de output y eliminada de las imágenes decodificadas del búfer DPB, etc. En general, en algunos estándares de conversión según codificación de video actuales y en desarrollo, la gestión del búfer DPB puede incluir uno o más de los siguientes aspectos: identificación de imagen e identificación de imagen de referencia, construcción de la lista de imágenes de referencia, marcado de imágenes de referencia, output de imagen del búfer DPB, inserción de imagen en el búfer DPB y eliminada de imagen del búfer DPB.

Para ayudar a la comprensión, a continuación se ofrece una breve descripción de cómo se puede realizar la construcción de la lista de imágenes de referencia y el marcado de imágenes de referencia de acuerdo con algunos estándares de codificación de video. Algunas de las técnicas que se describen en esta divulgación abordan cuestiones que pueden estar presentes en el marcado de imágenes de referencia, la construcción de una lista de imágenes de referencia y la eliminada y output de imágenes del búfer DPB para mejorar la eficiencia de la utilización del búfer DPB.

Para el marcado de imágenes de referencia, el número máximo, denominado M (num\_ref\_frames), de imágenes de referencia utilizadas para la inter predicción se indica en el conjunto de parámetros de secuencia activo. Cuando se decodifica una imagen de referencia, se marca como "utilizada como referencia". Si la decodificación de la imagen de referencia ha causado más de M imágenes marcadas como "utilizadas como referencia", al menos una imagen se debe marcar como "no utilizada como referencia". El proceso de eliminación del búfer DPB eliminaría entonces imágenes marcadas como "no utilizadas como referencia" del búfer DPB si tampoco son necesarias para el output.

Cuando se decodifica una imagen, puede ser una imagen de no referencia o una imagen de referencia. Una imagen de referencia puede ser una imagen de referencia de largo plazo o una imagen de referencia de corto plazo, y cuando se marca como "no utilizada como referencia", puede dejar de ser necesaria como referencia. En algunos estándares de conversión según codificación de video, puede haber operaciones de marcado de imágenes de referencia que cambian el estado de las imágenes de referencia.

Puede haber dos tipos de operaciones para el marcado de imágenes de referencia: ventana corrediza (sliding window) y control de memoria adaptativo. El modo de operación para el marcado de imágenes de referencia se puede seleccionar en función de las imágenes; mientras que la operación de la ventana corrediza puede funcionar como una cola de primera en entrar es la primera en salir con un número fijo de imágenes de referencia de corto plazo. En otras palabras, las imágenes de referencia de corto plazo con el tiempo de decodificación más antiguo pueden ser las primeras en ser eliminadas (marcadas como imágenes no utilizadas como referencia), de forma implícita.

Sin embargo, el control de memoria adaptativo elimina explícitamente imágenes de corto plazo o de largo plazo. También permite cambiar el estado de las imágenes de corto plazo y de largo plazo, etc. Por ejemplo, en el control de memoria adaptativo, un codificador de video (video encoder) puede señalar elementos de sintaxis que especifican qué imágenes se deben marcar como utilizadas como referencia. El decodificador de video puede recibir los elementos de sintaxis y marcar las imágenes según se especifica. En la ventana corrediza, es posible que el codificador de video (video encoder) no tenga que señalar qué imágenes se deben marcar como utilizadas como referencia. Más bien, el decodificador de video puede determinar implícitamente (es decir, sin recibir elementos de sintaxis) qué imágenes se deben marcar como utilizadas como referencia, en base a qué imágenes están dentro de la ventana corrediza.

El codificador de video (video coder) también puede tener la tarea de construir listas de imágenes de referencia que indican qué imágenes de referencia se pueden utilizar para propósitos de inter predicción. Dos de estas listas de imágenes de referencia se denominan Lista 0 y Lista 1, respectivamente. El codificador de video (video coder) utiliza en primer lugar técnicas de construcción por defecto para construir la Lista 0 y la Lista 1 (por ejemplo, esquemas de construcción preconfigurados para construir la Lista 0 y la Lista 1). Opcionalmente, después de construir la Lista 0 y la Lista 1 iniciales, el decodificador de video puede decodificar elementos de sintaxis, cuando estén presentes, que ordenen al decodificador de video que modifique la Lista 0 y la Lista 1 iniciales.

El codificador de video (video encoder) puede señalar elementos de sintaxis que son indicativos de un identificador o identificadores de imágenes de referencia en el búfer DPB, y el codificador de video (video encoder) también puede señalar elementos de sintaxis que incluyen índices, dentro de la Lista 0, la Lista 1, o ambas, la Lista 0 y la Lista 1, que indican qué imagen o imágenes de referencia se deben utilizar para decodificar un bloque codificado de una

imagen actual. El decodificador de video, a su vez, utiliza el identificador recibido para identificar el valor o valores de índice para una imagen de referencia o imágenes de referencia listadas en la Lista 0, la Lista 1, o en ambas, Lista 0 y Lista 1. A partir del valor o valores de índice y del identificador o identificadores de la imagen o imágenes de referencia, el decodificador de video recupera la imagen o imágenes de referencia, o una parte o partes de las mismas, del búfer 5 DPB, y decodifica el bloque codificado de la imagen actual en base a la imagen o imágenes de referencia recuperadas y uno o más vectores de movimiento que identifican bloques dentro de la imagen o imágenes de referencia que se utilizan para decodificar el bloque codificado.

Por ejemplo, construir una lista de imágenes de referencia para la primera o segunda lista de imágenes de referencia 10 de una imagen pronosticada dos veces incluye dos etapas: inicialización de la lista de imágenes de referencia y modificación de la lista de imágenes de referencia (también conocida como reordenación de la lista de imágenes de referencia). La inicialización de la lista de imágenes de referencia puede ser un mecanismo implícito que coloca las imágenes de referencia en la memoria de imágenes de referencia (también conocida como búfer de imágenes decodificadas) en una lista basada en el orden de valores de POC (Picture Order Count – contador de orden de las 15 imágenes, alineado con el orden de visualización de una imagen). El mecanismo de reordenación de la lista de imágenes de referencia puede modificar la posición de una imagen que se incluyó en la lista durante la inicialización de la lista de imágenes de referencia a cualquier nueva posición, o poner cualquier imagen de referencia de la memoria de imágenes de referencia en cualquier posición, incluso si la imagen no pertenece a la lista inicializada. Algunas imágenes después de la reordenación (modificación) de la lista de imágenes de referencia, se pueden colocar en una 20 posición muy avanzada de la lista. Sin embargo, si la posición de una imagen supera el número de imágenes de referencia activas de la lista, la imagen no se considera como una entrada de la lista final de imágenes de referencia. El número de imágenes de referencia activas puede ser señalado en la cabecera del corte (slice) para cada lista.

Las técnicas que se describen en esta divulgación pueden ser aplicables a diversos estándares de conversión según 25 codificación de video. Algunos ejemplos de los estándares de conversión según codificación de video son ITU-T H.261, ISO/IEC MPEG-1 Visual, ITU-T H.262 o ISO/IEC MPEG-2 Visual, ITU-T H.263, ISO/IEC MPEG-4 Visual y ITU-T H.264 (también conocido como ISO/IEC MPEG-4 AVC), incluidas sus extensiones SVC (Scalable Video Coding – codificación de video escalable) y MVC (Multiview Video Coding – codificación de video multivisión). Además, existe un nuevo estándar de codificación de video, en concreto el estándar HEVC (High Efficiency Video Coding – codificación 30 de video de alta eficiencia), que está siendo desarrollado por el equipo de colaboración conjunta sobre codificación de video (JCT-VC: Joint Collaboration Team on Video Coding) del grupo de expertos en codificación de video (VCEG: Video Coding Experts Group) del estándar ITU-T y el grupo de expertos de imágenes en movimiento (MPEG: Motion Picture Experts Group) del estándar ISO/IEC.

35 Sólo para fines de ilustración, las técnicas se describen en el contexto del estándar HEVC. Un Borrador de Trabajo (WD: Working Draft) reciente del estándar HEVC, y denominado WD8 del estándar HEVC en lo sucesivo, está disponible, desde el 20 de julio de 2012, en [http://phenix.int-evry.fr/jct/doc\\_end\\_user/documents/10\\_Stockholm/wg11/JCTVC-J1003-v8.zip](http://phenix.int-evry.fr/jct/doc_end_user/documents/10_Stockholm/wg11/JCTVC-J1003-v8.zip).

40 Según se ha descrito anteriormente, las técnicas que se describen en esta divulgación pueden abordar cuestiones que pueden estar presentes en soluciones existentes para la gestión del búfer de imágenes decodificadas (DPB). Como ejemplo, en algunos ejemplos de técnicas que se describen en esta divulgación, puede no ser necesario el marcado de las imágenes de referencia como "no utilizadas como referencia". Por ejemplo, las técnicas que se describen en esta divulgación pueden abordar cuestiones relacionadas con técnicas de gestión del búfer DPB que 45 pueden no ser adecuadas para una escalabilidad temporal, cuestiones relacionadas con la sobrecarga de la señalización de imágenes de referencia de largo plazo, cuestiones relacionadas con la eficiencia y la complejidad con la inicialización y modificación de la lista de imágenes de referencia. Las técnicas que se describen en esta divulgación también pueden abordar cuestiones relacionadas con el marcado de "imagen de no referencia" para entradas no completadas de una lista de imágenes de referencia durante la inicialización de la lista de imágenes de referencia, 50 cuestiones relacionadas con el output de imagen decodificada, la inserción en el búfer DPB y eliminación del mismo, así como cuestiones relacionadas con posibles valores de para valores de contador de orden de las imágenes (POC).

De acuerdo con las técnicas que se describen en esta divulgación, las listas de imágenes de referencia se construyen a partir de un conjunto de imágenes de referencia. Un conjunto de imágenes de referencia se define como un conjunto 55 de imágenes de referencia asociadas a una imagen, formado por todas las imágenes de referencia que son anteriores a la imagen asociada en un orden de decodificación, que se pueden utilizar para la inter predicción de bloques en la imagen asociada o en cualquier imagen que sigue a la imagen asociada en un orden de decodificación, por ejemplo, hasta la siguiente imagen de refresco de decodificación instantánea (IDR: instantaneous decoding refresh) o imagen de acceso de enlace roto (BLA: broken link access). En otras palabras, las imágenes de referencia en el conjunto de 60 imágenes de referencia pueden requerir las siguientes características: (1) todas ellas son anteriores a la imagen actual en un orden de decodificación, y (2) se pueden utilizar para la inter predicción de la imagen actual y/o inter predicción de cualquier imagen que sigue a la imagen actual en un orden de decodificación y, en algunos ejemplos, hasta la

siguiente imagen de IDR o imagen de BLA. Puede haber otras definiciones alternativas del conjunto de imágenes de referencia, que se proporcionan a continuación.

En las técnicas de ejemplo que se describen en esta divulgación, el codificador de video (video coder) puede derivar el conjunto de imágenes de referencia, y después de dicha derivación, el codificador de video (video coder) puede construir las listas de imágenes de referencia. Por ejemplo, solamente imágenes de referencia en el conjunto de imágenes de referencia pueden ser imágenes de referencia candidatas que se utilizan para construir las listas de imágenes de referencia.

Para construir el conjunto de imágenes de referencia, el codificador de video (video coder) puede construir una pluralidad de subconjuntos de imágenes de referencia. La combinación de los subconjuntos de imágenes de referencia puede formar el conjunto de imágenes de referencia. Por ejemplo, un codificador de video (video encoder) puede señalar explícitamente, en un flujo de bits codificado, valores que permiten que un decodificador de video determine identificadores para las imágenes de referencia que están incluidas en el conjunto de imágenes de referencia. Por ejemplo, los identificadores de las imágenes de referencia pueden ser los contadores según el orden de las imágenes. Cada imagen está asociada a un contador del orden de las imágenes, denominado PicOrderCnt. Este PicOrderCnt indica el orden de output u orden de visualización de la correspondiente imagen en relación con la imagen de IDR anterior en orden de decodificación y, en otras alternativas, indica la posición de la imagen asociada en el orden de output en relación con las posiciones del orden de output de las otras imágenes en la misma secuencia de video codificada.

El PicOrderCnt se puede denominar valor de contador del orden de las imágenes (POC: picture order count). Un valor de POC puede indicar el orden de output o de visualización de una imagen y se puede utilizar para identificar una imagen. Por ejemplo, dentro de una secuencia de video codificado, se suministra como output (is outputted) o se visualiza (is displayed) una imagen con un menor valor de POC antes que una imagen con un mayor valor de POC.

El decodificador de video puede determinar los identificadores para las imágenes de referencia, y a partir de estos identificadores construir la pluralidad de subconjuntos de imágenes de referencia. A partir de estos subconjuntos de imágenes de referencia, el decodificador de video puede derivar el conjunto de imágenes de referencia, según se describe con más detalle a continuación. En algunos ejemplos, cada uno de los subconjuntos de imágenes de referencia incluye diferentes imágenes de referencia, en el sentido de que no hay solapamiento de imágenes de referencia en los subconjuntos de imágenes de referencia. De esta manera, cada una de las imágenes de referencia puede estar en sólo uno de los subconjuntos de imágenes de referencia, y en ningún otro subconjunto de imágenes de referencia. Sin embargo, aspectos de esta divulgación no se deben considerar tan limitados.

Después de determinar los identificadores (por ejemplo, valores de POC) de las imágenes de referencia en el conjunto de imágenes de referencia o sus subconjuntos, el decodificador de video puede construir los subconjuntos de imágenes de referencia. Según se describe con más detalle a continuación, el decodificador de video puede construir seis subconjuntos de imágenes de referencia, aunque puede ser posible que el decodificador de video construya más o menos subconjuntos de imágenes de referencia.

Estos seis subconjuntos de imágenes de referencia se denominan de la siguiente manera: RefPicSetStCurr0, RefPicSetStCurr1, RefPicSetStFoll0, RefPicSetStFoll1, RefPicSetLtCurr y RefPicSetLtFoll. El subconjunto de imágenes de referencia RefPicSetStCurr0 se puede denominar subconjunto de imágenes de referencia RefPicSetStCurrBefore, y el subconjunto de imágenes de referencia RefPicSetStCurr1 se puede denominar subconjunto de imágenes de referencia RefPicSetStCurrAfter.

Los subconjuntos de imágenes de referencia RefPicSetStCurr0, RefPicSetStCurr1, RefPicSetStFoll0 y RefPicSetStFoll1 pueden identificar imágenes de referencia de corto plazo. En algunos ejemplos, estos subconjuntos de imágenes de referencia pueden identificar imágenes de referencia de corto plazo en base a si las imágenes de referencia de corto plazo son anteriores en el orden de visualización o posteriores en el orden de visualización con respecto a la imagen actual que se está convirtiendo según codificación, así como si las imágenes de referencia de corto plazo se pueden utilizar potencialmente para la inter predicción de la imagen actual y las imágenes posteriores a la imagen actual en el orden de decodificación, o se pueden utilizar potencialmente para la inter predicción de las imágenes posteriores a la imagen actual en el orden de decodificación.

Por ejemplo, el subconjunto de imágenes de referencia RefPicSetStCurr0 puede incluir, y puede incluir sólo, información de identificación, tal como valores de POC, de todas las imágenes de referencia de corto plazo que tienen un orden de output o de visualización anterior a la imagen actual, y que se pueden utilizar potencialmente como referencia en la inter predicción de la imagen actual, y que se pueden utilizar potencialmente como referencia en la inter predicción de una o más imágenes posteriores a la imagen actual en el orden de decodificación. El subconjunto de imágenes de referencia RefPicSetStCurr1 puede incluir, y puede incluir sólo, información de identificación de todas las imágenes de referencia de corto plazo que tienen un orden de output o de visualización posterior al de la imagen

actual y se puede utilizar potencialmente como referencia en la inter predicción de la imagen actual, y que se pueden utilizar potencialmente como referencia en la inter predicción de una o más imágenes posteriores a la imagen actual en el orden de decodificación.

- 5 El subconjunto de imágenes de referencia RefPicSetStFoll0 puede incluir, y puede incluir sólo, información de identificación de todas las imágenes de referencia de corto plazo que tienen un orden de output o de visualización anterior al de la imagen actual, que se pueden utilizar potencialmente como referencia en la inter predicción de una o más imágenes posteriores a la imagen actual en un orden de decodificación, y que no se pueden utilizar como referencia en la inter predicción de la imagen actual. El subconjunto de imágenes de referencia RefPicSetStFoll1 puede
- 10 incluir, y puede incluir sólo, información de identificación de todas las imágenes de referencia de corto plazo que tienen un orden de output o de visualización posterior al de la imagen actual, que se pueden utilizar potencialmente como referencia en la inter predicción de una o más imágenes posteriores a la imagen actual en un orden de decodificación y que no se pueden utilizar como referencia en la inter predicción de la imagen actual.
- 15 Los subconjuntos de imágenes de referencia RefPicSetLtCurr y RefPicSetLtFoll pueden identificar imágenes de referencia de largo plazo. En algunos ejemplos, estos subconjuntos de imágenes de referencia pueden identificar imágenes de referencia de largo plazo en base a si las imágenes de referencia de largo plazo son anteriores en el orden de visualización o posteriores en el orden de visualización con respecto a la imagen actual que se está convirtiendo según codificación.
- 20 Por ejemplo, el subconjunto de imágenes de referencia RefPicSetLtCurr puede incluir, y puede incluir sólo, la información de identificación de todas las imágenes de referencia de largo plazo que se pueden utilizar potencialmente como referencia en la inter predicción de la imagen actual, y que se pueden utilizar potencialmente como referencia en la inter predicción de una o más imágenes posteriores a la imagen actual en un orden de decodificación. El
- 25 subconjunto de imágenes de referencia RefPicSetLtFoll puede incluir, y puede incluir sólo, la información de identificación de todas las imágenes de referencia de largo plazo que se pueden utilizar potencialmente como referencia en la inter predicción de una o más imágenes posteriores a la imagen actual en un orden de decodificación, y que no se pueden utilizar como referencia en la inter predicción de la imagen actual.
- 30 Después de construir los subconjuntos de imágenes de referencia, el decodificador de video puede ordenar los subconjuntos de imágenes de referencia en un orden diferente para derivar el conjunto de imágenes de referencia. Por ejemplo, el orden del conjunto de imágenes de referencia puede ser RefPicSetStCurr0, RefPicSetStCurr1, RefPicSetFoll0, RefPicSetFoll1, RefPicSetLtCurr, y RefPicSetLtFoll. Sin embargo, otro orden de los subconjuntos puede ser posible para derivar el conjunto de imágenes de referencia. Por ejemplo, En otro ejemplo, el orden del
- 35 conjunto de imágenes de referencia puede ser el subconjunto de imágenes de referencia RefPicSetStCurr0, seguido del subconjunto de imágenes de referencia RefPicSetStCurr1, seguido del subconjunto de imágenes de referencia RefPicSetLtCurr, seguido del subconjunto de imágenes de referencia RefPicSetStFoll0, seguido del subconjunto de imágenes de referencia RefPicSetLtFoll1, y seguido del subconjunto de imágenes de referencia RefPicSetLtFoll.
- 40 De acuerdo con las técnicas que se describen en esta divulgación, los subconjuntos de imágenes de referencia RefPicSetStCurr0, RefPicSetStCurr1 y RefPicSetLtCurr incluyen todas las imágenes de referencia que se pueden utilizar para la inter predicción de un bloque en la imagen actual y que se pueden utilizar para la inter predicción de una o más imágenes posteriores a la imagen actual en un orden de decodificación. Los subconjuntos de imágenes de referencia RefPicSetStFoll0, RefPicSetStFoll1 y RefPicSetLtFoll incluyen todas las imágenes de referencia que no se
- 45 utilizan para la inter predicción del bloque en la imagen actual, pero que se pueden utilizar para la inter predicción de una o más imágenes posteriores a la imagen actual en el orden de decodificación.

Se debe entender que los seis subconjuntos de imágenes de referencia se describen con fines de ilustración solamente, y no se deben considerar limitantes. En ejemplos alternativos, puede haber más o menos subconjuntos de

50 imágenes de referencia. Estos subconjuntos de imágenes de referencia, en estos ejemplos alternativos, se describen con más detalle a continuación.

En algunas técnicas descritas en estas divulgaciones, el decodificador de video puede no necesitar marcar las imágenes decodificadas como "utilizadas como referencia", "no utilizadas como referencia", "utilizadas como

55 referencia de corto plazo" o "utilizadas como referencia de largo plazo". Más bien, si una imagen decodificada que está almacenada en el búfer DPB es necesaria para la inter predicción, se indica en función de si está incluida en el conjunto de imágenes de referencia de la imagen actual. En ejemplos alternativos, puede ser posible que el decodificador de video marque imágenes decodificadas como "utilizadas como referencia", "no utilizadas como referencia", "utilizadas como referencia de corto plazo" o "utilizadas como referencia de largo plazo". En estos

60 ejemplos, después de que el decodificador de video haya decodificado una imagen, es una imagen de referencia y es marcada como "utilizada como referencia". Entonces, después de la invocación del proceso de derivación del conjunto de imágenes de referencia, todas las imágenes de referencia almacenadas en el búfer DPB, pero no incluidas en el conjunto de imágenes de referencia de la imagen actual, son marcadas como "no utilizadas como referencia", antes

de la posible eliminación de imágenes decodificadas del búfer DPB. Por lo tanto, si una imagen decodificada que está almacenada en el búfer DPB es necesaria para la inter predicción, se puede indicar en función de si está marcada como "utilizada como referencia".

- 5 Una vez que el decodificador de video deriva el conjunto de imágenes de referencia a partir de la pluralidad de subconjuntos de imágenes de referencia, el decodificador de video puede construir las listas de imágenes de referencia (por ejemplo, la Lista 0 y la Lista 1) a partir del conjunto de imágenes de referencia. Por ejemplo, la construcción de las listas de imágenes de referencia puede incluir una etapa de inicialización y posiblemente una etapa de modificación. Derivando el conjunto de imágenes de referencia de la manera que se ha descrito anteriormente, el decodificador de  
10 video puede ser capaz de mejorar la eficiencia y reducir la complejidad de la inicialización de la lista de imágenes de referencia y la modificación de la lista de imágenes de referencia.

- El decodificador de video puede construir las listas de imágenes de referencia de diversas maneras. Las técnicas que se describen en esta divulgación proporcionan un mecanismo a través del cual el decodificador de video puede  
15 construir listas de imágenes de referencia sin necesidad de reordenar las imágenes de referencia a incluir en la lista (inicial) de imágenes de referencia. Por ejemplo, el decodificador de video puede estar configurado para implementar una técnica de construcción de lista de referencia por defecto en la que el decodificador de video utiliza los subconjuntos de imágenes de referencia para construir listas iniciales de imágenes de referencia. Entonces, si la modificación de la lista de imágenes de referencia no es necesaria, las listas finales de imágenes de referencia pueden  
20 ser las mismas que las listas iniciales de imágenes de referencia, sin necesidad de una reordenación adicional de las listas de imágenes de referencia.

- En algunos ejemplos, las técnicas que se describen en esta divulgación pueden estar relacionadas con la construcción de las listas de imágenes de referencia de tal manera que no haya entradas no completadas. Por ejemplo, las técnicas  
25 pueden añadir repetidamente, a una lista de imágenes de referencia, imágenes de referencia de uno o más de los subconjuntos de imágenes de referencia. Por ejemplo, después de que el decodificador de video haya añadido imágenes de referencia de uno o más de los subconjuntos de imágenes de referencia para construir una lista inicial de imágenes de referencia, el decodificador de video puede determinar si el número de entradas en la lista de imágenes de referencia es inferior al número máximo de entradas permitidas. Si el número de entradas en la lista de  
30 imágenes de referencia es inferior al número máximo de entradas permitidas, el decodificador de video puede volver a añadir, a la lista de imágenes de referencia, al menos una de las imágenes de referencia de uno de los subconjuntos de imágenes de referencia utilizados para construir la lista de imágenes de referencia. Esta re-agregación o re-añadido (también denominado re-listado) de la imagen de referencia se puede producir en una ubicación diferente dentro de las listas de imágenes de referencia, en comparación con la ubicación en la que la imagen de referencia fue  
35 añadida por primera vez por el decodificador de video.

- Según se usa en esta divulgación, el re-listado o re-añadido se refiere a añadir de nuevo (por ejemplo, identificar de nuevo) una imagen de referencia que fue añadida (por ejemplo, identificada) previamente a la lista inicial de imágenes de referencia. Sin embargo, cuando se re-añade una imagen de referencia, la imagen de referencia puede estar  
40 ubicada en dos entradas diferentes de la lista inicial de imágenes de referencia. En otras palabras, cuando se re-añade una imagen de referencia, puede haber dos valores de índice en la lista inicial de imágenes de referencia que identifican la misma imagen de referencia.

- En algunos ejemplos, las técnicas que se describen en esta divulgación pueden estar relacionadas con la modificación de una lista inicial de imágenes de referencia. Por ejemplo, el decodificador de video puede construir una lista inicial de imágenes de referencia. El decodificador de video puede determinar que es necesario modificar la lista de imágenes de referencia en base a elementos de sintaxis señalizados por el codificador de video en el flujo de bits codificado. Cuando es necesario modificar la lista de imágenes de referencia, el decodificador de video puede identificar una imagen de referencia en al menos uno de entre la pluralidad de subconjuntos de imágenes de referencia construidos.  
50 El decodificador de video puede listar (por ejemplo, añadir) la imagen de referencia identificada a una entrada actual de la lista inicial de imágenes de referencia para construir una lista de imágenes de referencia modificada. El decodificador de video puede entonces decodificar la imagen actual en base a la lista de imágenes de referencia modificada.

- 55 En algunos ejemplos, las técnicas que se describen en esta divulgación pueden estar relacionadas con el output y la eliminación de imágenes decodificadas de un búfer de imágenes decodificadas (DPB). Las técnicas de ejemplo pueden eliminar una imagen decodificada del búfer DPB antes de codificar una imagen actual. Por ejemplo, las técnicas de ejemplo pueden eliminar la imagen decodificada si esa imagen decodificada no está identificada en el conjunto de imágenes de referencia de la imagen actual y si esa imagen decodificada no es necesaria para el output (es decir, o bien no estaba destinada para el output o bien estaba destinada para el output, pero ya ha sido suministrada como output).  
60

La figura 1 es un diagrama de bloques que ilustra un ejemplo de sistema de codificación y decodificación de video 10 que puede utilizar las técnicas que se describen en esta divulgación. En general, un conjunto de imágenes de referencia se define como un conjunto de imágenes de referencia asociadas a una imagen, formado por todas las imágenes de referencia que son anteriores a la imagen asociada en un orden de decodificación, que se pueden utilizar 5 para la inter predicción de la imagen asociada o de cualquier imagen que sigue a la imagen asociada en un orden de decodificación. En algunos ejemplos, las imágenes de referencia que son anteriores a la imagen asociada pueden ser imágenes de referencia hasta la siguiente imagen de refresco de decodificación instantánea (IDR) o imagen de acceso de enlace roto (BLA). En otras palabras, las imágenes de referencia del conjunto de imágenes de referencia pueden ser todas anteriores a la imagen actual en un orden de decodificación. Además, las imágenes de referencia del 10 conjunto de imágenes de referencia se pueden utilizar para la inter predicción de la imagen actual y/o para la inter predicción de cualquier imagen que sigue a la imagen actual en un orden de decodificación hasta la siguiente imagen de IDR o imagen de BLA.

Puede haber otras definiciones alternativas de conjunto de imágenes de referencia. Por ejemplo, el conjunto de 15 imágenes de referencia puede ser un conjunto de imágenes de referencia asociadas a una imagen, formado por todas las imágenes de referencia, excluyendo la propia imagen asociada, que se pueden utilizar para la inter predicción de la imagen asociada o de cualquier imagen que sigue a la imagen asociada en un orden de decodificación, y que tienen un valor de temporal\_id inferior o igual al de la imagen asociada. El temporal\_id puede ser un valor de identificación temporal. El valor de identificación temporal puede ser un valor jerárquico que indica qué imágenes se pueden utilizar 20 para codificar la imagen actual. En general, una imagen con un determinado valor de temporal\_id puede ser una imagen de referencia para imágenes con valores de temporal\_id iguales o superiores, pero no vice versa. Por ejemplo, una imagen con un valor de temporal\_id igual a 1 puede ser una imagen de referencia para imágenes con valores de temporal\_id iguales a 1, 2, 3,..... pero no para una imagen con un valor de temporal\_id igual a 0.

El valor de temporal\_id más bajo también puede indicar la velocidad de visualización más baja. Por ejemplo, si un 25 decodificador de video sólo decodifica imágenes con valores de temporal\_id iguales a 0, la velocidad o tasa de visualización puede ser de 7,5 imágenes por segundo. Si un decodificador de video sólo decodifica imágenes con valores de temporal\_id iguales a 0 y 1, la velocidad o tasa de visualización puede ser de 15 imágenes por segundo, y así sucesivamente. 30

En otro ejemplo, el conjunto de imágenes de referencia puede ser un conjunto de imágenes de referencia asociadas con una imagen, formado por todas las imágenes de referencia, excluyendo la propia imagen asociada, que se pueden utilizar para la inter predicción de la imagen asociada o de cualquier imagen que sigue a la imagen asociada en un orden de decodificación. Como otro ejemplo más, el conjunto de imágenes de referencia se puede definir como un 35 conjunto de imágenes de referencia asociadas a una imagen, formado por todas las imágenes de referencia, posiblemente incluyendo la propia imagen asociada, que se pueden utilizar para la inter predicción de la imagen asociada o de cualquier imagen que sigue a la imagen asociada en un orden de decodificación. Como otro ejemplo adicional, el conjunto de imágenes de referencia se puede definir como un conjunto de imágenes de referencia asociadas a una imagen, formado por todas las imágenes de referencia, incluyendo posiblemente la propia imagen 40 asociada, que se pueden utilizar para la inter predicción de la imagen asociada o de cualquier imagen que sigue a la imagen asociada en un orden de decodificación, y que tienen temporal\_id inferior o igual al de la imagen asociada.

En otro ejemplo, en las definiciones anteriores de un conjunto de imágenes de referencia, la frase "se pueden utilizar para la inter predicción de" se sustituye por "se utilizan para la inter predicción de". Aunque puede haber definiciones 45 alternativas del conjunto de imágenes de referencia, en esta divulgación, los ejemplos se describen con la definición del conjunto de imágenes de referencia como un conjunto de imágenes de referencia asociadas con una imagen, formado por todas las imágenes de referencia que son anteriores a la imagen asociada en un orden de decodificación, que se pueden utilizar para la inter predicción de la imagen asociada o de cualquier imagen que sigue a la imagen asociada en el orden de decodificación. 50

Por ejemplo, algunas de las imágenes de referencia del conjunto de imágenes de referencia son imágenes de referencia que se pueden utilizar para la inter predicción de un bloque de la imagen actual, y no imágenes posteriores a la imagen actual en un orden de decodificación. Algunas de las imágenes de referencia del conjunto de imágenes de referencia son imágenes de referencia que se pueden utilizar potencialmente para la inter predicción de un bloque 55 de la imagen actual, y bloques en una o más imágenes posteriores a la imagen actual en orden de decodificación. Algunas de las imágenes de referencia del conjunto de imágenes de referencia son imágenes de referencia que se pueden utilizar para la inter predicción de bloques en una o más imágenes posteriores a la imagen actual en un orden de decodificación, y no se pueden utilizar para la inter predicción de un bloque en la imagen actual.

Según se utiliza en esta divulgación, las imágenes de referencia que se pueden utilizar potencialmente para la inter 60 predicción se refieren a imágenes de referencia que se pueden utilizar para la inter predicción, pero no necesariamente tienen que utilizarse para la inter predicción. Por ejemplo, el conjunto de imágenes de referencia puede identificar imágenes de referencia que se pueden utilizar potencialmente para la inter predicción. Sin embargo, esto no significa

que todas las imágenes de referencia identificadas deban utilizarse para la inter predicción. Más bien, una o más de estas imágenes de referencia identificadas se podrían utilizar para la inter predicción, pero no todas tienen que utilizarse necesariamente para la inter predicción.

5 Según se muestra en la figura 1, el sistema 10 incluye un dispositivo de origen 12 que genera video codificado para su decodificación por parte del dispositivo de destino 14. El dispositivo de origen 12 y el dispositivo de destino 14 pueden ser cada uno un ejemplo de un dispositivo de conversión según codificación de video. El dispositivo de origen 12 puede transmitir el video codificado al dispositivo de destino 14 a través del canal de comunicación 16 o puede almacenar el video codificado en un medio de almacenamiento o en un servidor de archivos 19, de forma que el  
10 dispositivo de destino 14 pueda acceder al video codificado como desee.

El dispositivo de origen 12 y el dispositivo de destino 14 pueden comprender cualquiera de entre una amplia gama de dispositivos, que incluye teléfonos inalámbricos, tales como los denominados teléfonos "inteligentes", los denominados pad "inteligentes" u otros dispositivos inalámbricos similares equipados para comunicación inalámbrica. Otros  
15 ejemplos del dispositivo de origen 12 y del dispositivo de destino 14 incluyen, entre otros, una televisión digital, un dispositivo en un sistema de radiodifusión digital directa, un dispositivo en un sistema de radiodifusión inalámbrica, asistentes digitales personales (PDA), un ordenador portátil, un ordenador de sobremesa, un ordenador de tableta, un lector de libros electrónicos, una cámara digital, un dispositivo de grabación digital, un reproductor de medios digitales, un dispositivo de videojuegos, una consola de videojuegos, un radioteléfono celular, un teléfono por satélite, un  
20 dispositivo de video telefonía, y un dispositivo de transmisión de flujos de video, un dispositivo de comunicación inalámbrica, o similares.

Según se ha indicado anteriormente, en muchos casos, el dispositivo de origen 12 y/o el dispositivo de destino 14 pueden estar equipados para la comunicación inalámbrica. Por lo tanto, el canal de comunicación 16 puede  
25 comprender un canal inalámbrico, un canal alámbrico o una combinación de canales inalámbricos y alámbricos adecuados para la transmisión de datos de video codificados. De modo similar, el dispositivo de destino 14 puede acceder al servidor de archivos a través de cualquier conexión de datos estándar, incluida una conexión a Internet. Esto puede incluir un canal inalámbrico (por ejemplo, una conexión WiFi), una conexión por cable o alámbrica (por ejemplo, DSL, módem por cable, etc.), o una combinación de ambas que sea adecuada para acceder a datos de video  
30 codificados que están almacenados en un servidor de archivos.

Las técnicas de esta divulgación, sin embargo, se pueden aplicar a la conversión según codificación de video en soporte de una diversidad de aplicaciones multimedia, tales como radiodifusión de televisión por aire, transmisiones de televisión por cable, transmisiones de televisión por satélite, transmisiones de flujos de video, por ejemplo, a través  
35 de Internet, codificación de video digital para su almacenamiento en un medio de almacenamiento de datos, decodificación de video digital que está almacenado en un medio de almacenamiento de datos u otras aplicaciones. En algunos ejemplos, el sistema 10 puede estar configurado para soportar la transmisión de video unidireccional o bidireccional para soportar aplicaciones tales como la transmisión de flujos de video, la reproducción de video, la radiodifusión de video y/o la video telefonía.

40 En el ejemplo de la figura 1, el dispositivo de origen 12 incluye una fuente de video 18, un codificador de video (video encoder) 20, un modulador/demodulador (MODEM) 22 y una interfaz de output 24. En el dispositivo de origen 12, la fuente de video 18 puede incluir una fuente tal como un dispositivo de captura de video, tal como una cámara de video, un archivo de video que contiene video capturado previamente, una interfaz de recepción de video para recibir video  
45 procedente de un proveedor de contenidos de video y/o un sistema de gráficos por ordenador para generar datos de gráficos por ordenador como fuente de video, o una combinación de dichas fuentes. Por ejemplo, si la fuente de video 18 es una cámara de video, el dispositivo de origen 12 y el dispositivo de destino 14 pueden formar los denominados teléfonos con cámara o teléfonos de video. Sin embargo, las técnicas que se describen en esta divulgación pueden ser aplicables a la conversión según codificación de video en general, y pueden ser aplicadas a aplicaciones  
50 inalámbricas y/o alámbricas.

El video capturado, pre capturado o generado por ordenador se puede codificar mediante el codificador de video (video encoder) 20. La información de video codificada puede ser modulada por el módem 22 de acuerdo con un estándar de comunicación, tal como un protocolo de comunicación inalámbrica, y transmitida al dispositivo de destino 14 a  
55 través de la interfaz de output 24. El módem 22 puede incluir diversos mezcladores, filtros, amplificadores u otros componentes diseñados para la modulación de la señal. La interfaz de output 24 puede incluir circuitos diseñados para la transmisión de datos, que incluyen amplificadores, filtros y una o más antenas.

El video capturado, pre capturado o generado por ordenador y que es codificado por el codificador de video (video encoder) 20 también se puede almacenar en un medio de almacenamiento 17 o en un servidor de archivos 19 para su posterior consumo. El medio de almacenamiento 17 puede incluir discos Blu-ray, DVD, CD-ROM, memoria flash o cualquier otro medio de almacenamiento digital adecuado para almacenar video codificado. El video codificado que  
60

está almacenado en el medio de almacenamiento 17 puede ser accedido por el dispositivo de destino 14 para su decodificación y reproducción.

El servidor de archivos 19 puede ser cualquier tipo de servidor capaz de almacenar video codificado y de transmitirlo al dispositivo de destino 14. Servidores de archivos de ejemplo incluyen un servidor web (por ejemplo, para un sitio web), un servidor FTP, dispositivos de almacenamiento en red (NAS. Network Attached Storage), una unidad de disco local o cualquier otro tipo de dispositivo capaz de almacenar datos de video codificados y de transmitirlos a un dispositivo de destino. La transmisión de datos de video codificados por parte del servidor de archivos puede ser una transmisión de flujos de video, una transmisión de descarga o una combinación de ambas. El servidor de archivos 19 puede ser accedido por el dispositivo de destino 14 a través de cualquier conexión de datos estándar, incluida una conexión a Internet. Esto puede incluir un canal inalámbrico (por ejemplo, una conexión WiFi), una conexión por cable o alámbrica (por ejemplo, DSL, módem de cable, Ethernet, USB, etc.), o una combinación de ambas que sea adecuada para acceder a datos de video codificados que están almacenados en un servidor de archivos.

El dispositivo de destino 14, en el ejemplo de la figura 1, incluye una interfaz de input 26, un módem 28, un decodificador de video 30 y un dispositivo de visualización 32. La interfaz de input 26 del dispositivo de destino 14 recibe información a través del canal 16, por ejemplo, o procedente del medio de almacenamiento 17 o del servidor de archivos 17, como ejemplos alternativos, y el módem 28 demodula la información para producir un flujo de bits demodulado para el decodificador de video 30. El flujo de bits demodulado puede incluir una diversidad de información de sintaxis generada por el codificador de video (video encoder) 20 para su uso por el decodificador de video 30 en la decodificación de datos de video. Esta sintaxis también puede ser incluida con los datos de video codificados que están almacenados en un medio de almacenamiento 17 o en un servidor de archivos 19. Por ejemplo, la sintaxis puede estar embebida con los datos de video codificados, aunque aspectos de esta divulgación no se deben considerar limitados a dicho requisito. La información de sintaxis definida por el codificador de video (video encoder) 20, que también es utilizada por el decodificador de video 30, puede incluir elementos de sintaxis que describen características y/o procesamiento de bloques de video, tales como unidades de árbol de codificación (CTU), bloques de árbol de codificación (CTB), unidades de predicción (PU), unidades de codificación (CU) u otras unidades de video codificado, por ejemplo, cortes de video, imágenes de video y secuencias o grupos de imágenes (GOP) de video. Cada uno del codificador de video (video encoder) 20 y decodificador de video 30 puede formar parte de un respectivo codificador-decodificador (CODEC) que es capaz de codificar o decodificar datos de video.

El dispositivo de visualización 32 puede estar integrado con el dispositivo de destino o ser externo al mismo 14. En algunos ejemplos, el dispositivo de destino 14 puede incluir un dispositivo de visualización integrado y también estar configurado para actuar de interfaz con un dispositivo de visualización externo. En otros ejemplos, el dispositivo de destino 14 puede ser un dispositivo de visualización. En general, el dispositivo de visualización 32 muestra o visualiza los datos de video decodificados a un usuario, y puede comprender una cualquiera de entre una diversidad de dispositivos de visualización tales como una pantalla de cristal líquido (LCD), una pantalla de plasma, una pantalla de diodos emisores de luz orgánicos (OLED) u otro tipo de dispositivo de visualización.

En el ejemplo de la figura 1, el canal de comunicación 16 puede comprender cualquier medio de comunicación inalámbrico o alámbrico, tal como un espectro de radiofrecuencia (RF) o una o más líneas de transmisión físicas, o cualquier combinación de medios inalámbricos y alámbricos. El canal de comunicación 16 puede formar parte de una red basada en paquetes, tal como una red de área local, una red de área amplia o una red global tal como Internet. El canal de comunicación 16 representa en general cualquier medio de comunicación adecuado, o colección de diferentes medios de comunicación, para transmitir datos de video desde el dispositivo de origen 12 al dispositivo de destino 14, incluyendo cualquier combinación adecuada de medios alámbricos o inalámbricos. El canal de comunicación 16 puede incluir enrutadores, conmutadores, estaciones base o cualquier otro equipo que pueda ser útil para facilitar la comunicación entre el dispositivo de origen 12 y el dispositivo de destino 14.

El codificador de video (video encoder) 20 y el decodificador de video 30 pueden funcionar de acuerdo con una estándar de compresión de video, tal como las extensiones ITU-T H.261, ISO/IEC MPEG-1 Visual, ITU-T H.262 o ISO/IEC MPEG-2 Visual, ITU-T H.263, ISO/IEC MPEG-4 Visual y ITU-T H.264 (también conocida como ISO/IEC MPEG-4 AVC), incluidas sus extensiones de codificación de video escalable (SVC) y codificación de video multi vista (MVC). Además, existe un nuevo estándar de codificación de video, en concreto, el estándar de codificación de video de alta eficiencia (HEVC), que actualmente está en fase de desarrollo por el Equipo Conjunto de Colaboración en Codificación de Video (JCT-VC) del Grupo de Expertos en Codificación de Video según ITU-T (VCEG) y del Grupo de Expertos de imágenes en movimiento según ISO/IEC (MPEG). Un Borrador de Trabajo (WD) reciente del estándar HEVC, y denominado WD8 del estándar HEVC en lo sucesivo, está disponible, a partir del 20 de julio de 2012, en [http://phenix.int-evey.fr/jct/doc\\_end\\_user/documents/10\\_Stockholm/wg11/JCTVC-J1003-v8.zip](http://phenix.int-evey.fr/jct/doc_end_user/documents/10_Stockholm/wg11/JCTVC-J1003-v8.zip).

Las técnicas de esta divulgación, sin embargo, no se limitan a ningún estándar de codificación en particular. Las técnicas se describen de acuerdo con el estándar HEVC sólo a modo ilustrativo.



Aunque no se muestra en la figura 1, en algunos aspectos, el codificador de video (video encoder) 20 y el decodificador de video 30 pueden estar integrados cada uno de ellos con un codificador y decodificador de audio, y pueden incluir unidades MUX–DEMUX apropiadas, u otro hardware y software, para realizar codificación de audio y video en una secuencia o flujo de datos común o en flujos de datos separados. Si procede, las unidades MUX–DEMUX pueden ajustarse al protocolo multiplexor H.223 de ITU o a otros protocolos tales como el protocolo de datagramas de usuario (UDP).

El codificador de video (video encoder) 20 y el decodificador de video 30 pueden ser implementados cada uno de ellos como cualquiera de entre una diversidad de circuitos codificadores adecuados, tales como uno o más procesadores que incluyen microprocesadores, procesadores de señal digital (DSP), circuitos integrados específicos de aplicación (ASIC), matrices de puertas programables de campo (FPGA), lógica discreta, software, hardware, firmware o cualquier combinación de los mismos. Cuando las técnicas son implementadas parcialmente en software, un dispositivo puede almacenar instrucciones para el software en un medio no transitorio legible por ordenador adecuado y ejecutar las instrucciones en el hardware utilizando uno o más procesadores para realizar las técnicas de esta divulgación.

Cada uno del codificador de video (video encoder) 20 y decodificador de video 30 puede estar incluido en uno o más codificadores o decodificadores, cualquiera de los cuales puede estar integrado como parte de un codificador/decodificador (CODEC) combinado en un respectivo dispositivo. En algunos casos, el codificador de video (video encoder) 20 y el decodificador de video 30 se pueden denominar comúnmente como convertidor de codificación de video que convierte información (por ejemplo, imágenes y elementos de sintaxis) según una codificación. La conversión según codificación de información se puede referir a la codificación cuando el codificador de video (video coder) corresponde al codificador de video (video encoder) 20. La conversión según codificación de información se puede referir a la decodificación cuando el codificador de video (video coder) corresponde al decodificador de video 30.

Además, las técnicas que se describen en esta divulgación se pueden referir a información de señalización por parte del codificador de video (video encoder) 20. Cuando el codificador de video (video encoder) 20 señala información, las técnicas de esta divulgación se refieren en general a cualquier manera en la que el codificador de video (video encoder) 20 proporciona la información. Por ejemplo, cuando el codificador de video (video encoder) 20 señala elementos de sintaxis al decodificador de video 30, puede significar que el codificador de video (video encoder) 20 ha transmitido los elementos de sintaxis al decodificador de video 30 a través de la interfaz de output 24 y del canal de comunicación 16, o que el codificador de video (video encoder) 20 ha almacenado los elementos de sintaxis a través de la interfaz de output 24 en el medio de almacenamiento 17 y/o en el servidor de archivos 19 para su recepción final por parte del decodificador de video 30. De este modo, la señalización por parte del codificador de video (video encoder) 20 al decodificador de video 30 no se debe interpretar como que requiere una transmisión por parte del codificador de video (video encoder) 20 que es recibida inmediatamente por el decodificador de video 30, aunque esto puede ser posible. Más bien, la señalización por parte del codificador de video (video encoder) 20 al decodificador de video 30 se debe interpretar como cualquier técnica mediante la cual el codificador de video (video encoder) 20 proporciona información para su recepción final por parte del decodificador de video 30, ya sea directamente o a través de un almacenamiento intermedio (por ejemplo, en un medio de almacenamiento 17 y/o en un servidor de archivos 19).

El codificador de video (video encoder) 20 y el decodificador de video 30 pueden estar configurados para implementar las técnicas de ejemplo que se describen en esta divulgación para derivar un conjunto de imágenes de referencia. Por ejemplo, el decodificador de video 30 puede invocar el proceso para derivar el conjunto de imágenes de referencia una vez por cada imagen. El decodificador de video 30 puede invocar el proceso para derivar el conjunto de imágenes de referencia después de la decodificación de una cabecera de corte (slice), pero antes de la decodificación de cualquier unidad de codificación y antes del proceso de decodificación para la construcción de la lista de imágenes de referencia del corte (slice).

Según se ha descrito anteriormente, el conjunto de imágenes de referencia es una descripción absoluta de las imágenes de referencia utilizadas en el proceso de decodificación de la imagen actual y de futuras imágenes codificadas en un orden de decodificación hasta la siguiente imagen de refresco de decodificación instantánea (IDR: Instant Decoding Refresh) o imagen de acceso de enlace roto (BLA). En ejemplos descritos en esta divulgación, el codificador de video (video encoder) 20 puede señalar explícitamente valores a partir de los cuales el decodificador de video 30 puede determinar identificadores para las imágenes de referencia que pertenecen al conjunto de imágenes de referencia. La señalización del conjunto de imágenes de referencia es explícita en el sentido de que se enumeran o listan explícitamente todas las imágenes de referencia incluidas en el conjunto de imágenes de referencia, excepto en el caso de determinadas imágenes, por ejemplo, imágenes de IDR, no se incluyen elementos de sintaxis del conjunto de imágenes de referencia en la cabecera de corte (slice) y se deriva el conjunto de imágenes de referencia como conjunto vacío.

Puede haber diversas maneras en las que el codificador de video (video encoder) 20 puede señalar elementos de sintaxis en un flujo de bits codificado que el decodificador de video 30 puede utilizar para derivar el conjunto de imágenes de referencia. Por ejemplo, el codificador de video (video encoder) 20 puede señalar los elementos de sintaxis en el conjunto de parámetros de imagen (PPS), el conjunto de parámetros de secuencia (SPS), la cabecera de la imagen (si la hay), la cabecera del corte (slice) o cualquier combinación de los mismos. Sólo con fines ilustrativos, el codificador de video (video encoder) 20 puede señalar los elementos de sintaxis utilizando el SPS, el PPS y la cabecera de corte (slice), según se describe con más detalle.

Para derivar el conjunto de imágenes de referencia, el decodificador de video 30 puede implementar un proceso de decodificación para determinar los identificadores de imágenes que pertenecen al conjunto de imágenes de referencia. El decodificador de video 30 puede entonces construir una pluralidad de subconjuntos de imágenes de referencia, en el que cada uno de los subconjuntos identifica cero o más de las imágenes de referencia que pertenecen al conjunto de imágenes de referencia. El decodificador de video 30 puede derivar el conjunto de imágenes de referencia a partir de los subconjuntos de imágenes de referencia analizados. Por ejemplo, el decodificador de video 30 puede listar o enumerar la pluralidad de subconjuntos de imágenes de referencia en un orden particular para derivar el conjunto de imágenes de referencia.

Puede haber diversas maneras en las que el decodificador de video 30 puede determinar los identificadores de las imágenes que pertenecen al conjunto de imágenes de referencia. En general, el codificador de video (video encoder) 20 puede señalar valores a partir de los cuales el decodificador de video 30 puede determinar los identificadores de las imágenes, incluidas las imágenes que pertenecen al conjunto de imágenes de referencia. Los identificadores de las imágenes pueden ser el PicOrderCnt (es decir, valores de contador del orden de las imágenes (POC)). Según se ha descrito anteriormente, el valor de POC puede indicar el orden de visualización o de output de una imagen, en el que las imágenes con valores de POC más pequeños se visualizan antes que las imágenes con valores de POC más grandes. El valor de POC de una imagen determinada puede ser relativo con respecto a la imagen anterior de refresco de decodificación instantánea (IDR). Por ejemplo, el PicOrderCnt (es decir, el valor de POC) para una imagen de IDR puede ser igual a 0, el valor de POC para la imagen posterior a la imagen de IDR en orden de visualización o de output puede ser igual a 1, el valor de POC para la imagen posterior a la imagen con valor de POC igual a 1 en orden de visualización o de output puede ser igual a 2, y así sucesivamente.

De acuerdo con las técnicas que se describen en esta divulgación, cuando la imagen actual no es una imagen de IDR, se puede aplicar lo siguiente para derivar el valor de POC para la imagen actual. Lo siguiente pretende ayudar a la comprensión, y no se debe considerar como una limitación.

Por ejemplo, considérese la variable de lista listD que incluye como elementos los valores de PicOrderCnt (valores de POC) asociados con la lista de imágenes que incluye todo lo siguiente: (1) la primera imagen de la lista es la imagen de IDR anterior en orden de decodificación, y (2) todas las demás imágenes que siguen en orden de decodificación después de la primera imagen de la lista y preceden a la imagen actual en orden de decodificación o son la imagen actual. En este ejemplo, se incluye la imagen actual en la listD antes de invocar el proceso de derivación para el conjunto de imágenes de referencia. Considérese también la variable listO, que incluye los elementos de la listD ordenados en orden ascendente de los valores de POC. En este ejemplo, es posible que la listO no contenga un valor de POC que tenga un valor igual al valor de POC de otra imagen.

En algunos ejemplos, los valores de POC se pueden limitar al rango de  $-2^{\text{pocLen}-1}$  hasta  $2^{\text{pocLen}-1} - 1$ , inclusive. En este ejemplo, pocLen puede ser igual a  $\text{long\_term\_ref\_pic\_id\_len\_delta} + \text{long\_term\_ref\_pic\_id\_delta\_len\_minus4} + 4$ . El  $\text{long\_term\_ref\_pic\_id\_len\_delta}$ , y el  $\text{long\_term\_ref\_pic\_id\_delta\_len\_minus4}$  pueden ser elementos de sintaxis que el decodificador de video 30 recibe en el flujo de bits codificado como parte de la sintaxis del conjunto de parámetros de imagen, según se describe con más detalle a continuación. En otro ejemplo, los valores de POC se pueden limitar al rango de  $-2^{31}$  hasta  $2^{31} - 1$ , inclusive.

Por ejemplo, el decodificador de video 30 puede recibir en el flujo de bits codificado (es decir, el flujo de bits señalado por el codificador de video (video encoder) 20), el elemento de sintaxis pic\_order\_cnt\_lsb. El elemento de sintaxis pic\_order\_cnt\_lsb puede especificar el módulo contador de orden de las imágenes MaxPicOrderCntLsb para la imagen codificada. La longitud del elemento de sintaxis pic\_order\_cnt\_lsb puede ser igual a  $\log_2 \text{max\_pic\_order\_cnt\_lsb\_minus4} + 4$  bits. El valor de pic\_order\_cnt\_lsb puede estar en el rango de 0 hasta  $\text{MaxPicOrderCntLsb} - 1$ , inclusive. El decodificador de video 30 puede recibir el elemento de sintaxis pic\_order\_cnt\_lsb en la sintaxis de la cabecera de corte (slice) para la imagen a decodificar.

El decodificador de video 30 también puede recibir el elemento de sintaxis log2\_max\_pic\_order\_cnt\_lsb\_minus4 en el flujo de bits codificado señalado por el codificador de video (video encoder) 20. El decodificador de video 30 puede recibir el elemento de sintaxis log2\_max\_pic\_order\_cnt\_lsb\_minus4 en el conjunto de parámetros de secuencia. El valor de log2\_max\_pic\_order\_cnt\_lsb\_minus4 puede estar en el rango de 0 hasta 12, inclusive. El elemento de sintaxis

`log2_max_pic_order_cnt_lsb_minus4` puede especificar el valor de la variable `MaxPicOrderCntLsb` que el decodificador de video 30 utiliza en el proceso de decodificación para determinar los valores de POC. Por ejemplo:

$$\text{MaxPicOrderCntLsb} = 2^{(\text{log2\_max\_pic\_order\_cnt\_lsb\_minus4} + 4)}$$

5

A partir de estos elementos de sintaxis recibidos, el decodificador de video 30 puede determinar el valor de POC para la imagen actual de la siguiente manera. Por ejemplo, el decodificador de video 30 puede determinar el `PicOrderCntMsb` para la imagen actual. El valor de POC para la imagen actual puede ser el `PicOrderCntMsb` determinado para la imagen actual más el `pic_order_cnt_lsb` recibido para la imagen actual.

10

A continuación, la función `PicOrderCnt(picX)` es igual al valor de POC para la imagen X. La función `DiffPicOrderCnt(picA, picB)` es igual a `PicOrderCnt(picA)` menos `PicOrderCnt(picB)`. En algunos ejemplos, es posible que el flujo de bits codificado no incluya datos que produzcan como resultado los valores de `DiffPicOrderCnt(picA, picB)` utilizados en el proceso de decodificación que excedan el rango de  $-2^{15}$  hasta  $2^{15} - 1$ ; inclusive. Además, considérese que X es la imagen actual e Y y Z son otras dos imágenes en la misma secuencia, en la que Y y Z se consideran la misma dirección de orden de output con respecto a X cuando tanto `DiffPicOrderCnt(X, Y)` como `DiffPicOrderCnt(X, Z)` son positivas o ambas son negativas. Además, en algunos ejemplos, el codificador de video (video encoder) 20 puede asignar `PicOrderCnt` proporcionalmente al tiempo de muestreo de la correspondiente imagen relativa con respecto al tiempo de muestreo de la imagen de IDR anterior.

20

Como parte del proceso de determinar el valor de POC para la imagen actual, el decodificador de video 30 puede determinar las variables `prevPicOrderCntMsb` y `prevPicOrderCntLsb`. Por ejemplo, si la imagen actual es una imagen de IDR, el decodificador de video 30 puede establecer `prevPicOrderCntMsb` igual a 0 y `prevPicOrderCntLsb` igual a 0. De lo contrario (es decir, si la imagen actual no es una imagen de IDR), el decodificador de video 30 puede establecer `prevPicOrderCntMsb` igual a `PicOrderCntMsb` de la imagen de referencia anterior en un orden de decodificación con un valor de `temporal_id` inferior o igual al de la imagen actual, y establecer `prevPicOrderCntLsb` igual al valor de `pic_order_cnt_lsb` de la imagen de referencia anterior en orden de decodificación con un valor de `temporal_id` inferior o igual al de la imagen actual.

25

30 Con estos valores de variables y los valores de los elementos de sintaxis (por ejemplo, los valores de `prevPicOrderCntMsb`, `prevPicOrderCntLsb`, `pic_order_cnt_lsb` y `MaxPicOrderCntLsb`), el decodificador de video 30 puede determinar el valor de `PicOrderCntMsb` en base a las etapas indicadas en el siguiente pseudo código. Se debe entender que el decodificador de video 30 puede implementar las etapas indicadas en el siguiente pseudo código para determinar el `PicOrderCntMsb` para cada imagen actual, que se utiliza para derivar el valor de POC para la imagen actual.

35

```
si( ( pic_order_cnt_lsb < prevPicOrderCntLsb ) && ( ( prevPicOrderCntLsb - pic_order_cnt_lsb ) >= ( MaxPicOrderCntLsb / 2 ) ) )
```

40

```
    PicOrderCntMsb = prevPicOrderCntMsb + MaxPicOrderCntLsb
```

```
sino si( ( pic_order_cnt_lsb > prevPicOrderCntLsb ) && ( ( pic_order_cnt_lsb - prevPicOrderCntLsb ) > ( MaxPicOrderCntLsb / 2 ) ) )
```

```
    PicOrderCntMsb = prevPicOrderCntMsb - MaxPicOrderCntLsb
```

45 sino

```
    PicOrderCntMsb = prevPicOrderCntMsb
```

Después de determinar el `PicOrderCntMsb` para la imagen actual, el decodificador de video 30 puede determinar el valor de POC para la imagen actual en base al `PicOrderCntMsb` para la imagen actual y el `pic_order_cnt_lsb` para la imagen actual. El decodificador de video 30 puede determinar el valor de POC para la imagen actual de la siguiente manera:

50

```
PicOrderCnt = PicOrderCntMsb + pic_order_cnt_lsb.
```

55 Después de decodificar una imagen, el decodificador de video 30 puede almacenar el valor de `PicOrderCntMsb`, el valor de `pic_order_cnt_lsb` y el valor de POC para esa imagen, que incluyen cada una de las imágenes de referencia que pertenecen al conjunto de imágenes de referencia, en un búfer de imágenes decodificadas (DPB) del decodificador de video 30. De este modo, cada imagen en el búfer DPB se asocia con un valor de POC, un valor de `PicOrderCntMsb` y un valor de `pic_order_cnt_lsb`.

60

A continuación se describen con más detalle procedimientos para determinar los valores de POC de las imágenes de referencia incluidas en el conjunto de imágenes de referencia de una imagen actual. A partir de los valores de POC determinados, el decodificador de video 30 puede implementar el proceso de derivación para el conjunto de imágenes

de referencia. Sin embargo, antes de describir la manera en la que el decodificador de video 30 implementa el proceso de derivación para el conjunto de imágenes de referencia, a continuación se proporcionan tablas de elementos de sintaxis que el decodificador de video 30 puede recibir en el flujo de bits codificado señalizado por el codificador de video (video encoder) 20. Por ejemplo, el codificador de video (video encoder) 20 puede señalar los elementos de sintaxis de las siguientes tablas en el flujo de bits codificado que recibe el decodificador de video 30. Algunos de estos elementos de sintaxis se han descrito anteriormente. A partir de los elementos de sintaxis, el decodificador de video 30 puede determinar los valores de POC de las imágenes de referencia incluidas en el conjunto de imágenes de referencia y derivar el conjunto de imágenes de referencia.

- 10 Por ejemplo, en las técnicas que se describen en esta divulgación, se modifican las siguientes estructuras de sintaxis en relación con los anteriores estándares de codificación de video: la sintaxis de carga útil de secuencia de bytes en bruto o sin procesar (Rbsp: raw byte sequence payload) del conjunto de parámetros de secuencia (SPS: sequence parameter set), seq\_parameter\_set\_rbsp( ), la sintaxis de Rbsp del conjunto de parámetros de imagen (PPS), pic\_parameter\_set\_rbsp( ), la sintaxis de cabecera de corte (slice), slice\_header( ), y la sintaxis de modificación de la lista de imágenes de referencia, ref\_pic\_list\_modification( ).
- 15 La modificación de la lista de imágenes de referencia se describe con más detalle siguiendo la descripción de derivar el conjunto de imágenes de referencia, e inicializar una o más listas de imágenes de referencia.

Además, de acuerdo con las técnicas que se describen en esta divulgación, se añaden las siguientes estructuras de sintaxis al flujo de bits codificado: la sintaxis del conjunto de imágenes de referencia de corto plazo, short\_term\_ref\_pic\_set( ), y la sintaxis del conjunto de imágenes de referencia de largo plazo, long\_term\_ref\_pic\_set( ). El decodificador de video 30 puede utilizar la sintaxis del conjunto de imágenes de referencia de corto plazo y la sintaxis del conjunto de imágenes de referencia de largo plazo para construir subconjuntos de imágenes de referencia, a partir de los cuales el decodificador de video 30 deriva el conjunto de imágenes de referencia.

- 25 Por ejemplo, para que el decodificador de video 30 determine los valores de POC para las imágenes de referencia que pertenecen al conjunto de imágenes de referencia, el codificador de video (video encoder) 20 puede señalar información de identificación de imágenes de referencia que el decodificador de video 30 utiliza para determinar los valores de POC, en el conjunto de parámetros de imagen y se puede referenciar el índice de la lista en la cabecera de corte (slice). Sin embargo, éste es un ejemplo de cómo el codificador de video (video encoder) 20 puede señalar dicha información de identificación de imágenes de referencia.

En un ejemplo alternativo, el codificador de video (video encoder) 20 puede señalar información de imágenes de referencia en el conjunto de parámetros de secuencia y se puede referenciar el índice de la lista en la cabecera de corte (slice), lo que puede reducir la sobrecarga de señalización. En otro ejemplo alternativo, el codificador de video (video coder) puede señalar la información de imágenes de referencia en un nuevo tipo de conjunto de parámetros (por ejemplo, conjunto de parámetros del conjunto de imágenes de referencia (RPSPS: reference picture set parameter set)), y el identificador del RPSPS, así como el índice de la lista de información de identificación de imágenes de referencia, pueden ser referenciados en la cabecera de corte (slice). Esto puede reducir la sobrecarga de señalización así como no aumentar la necesidad del número de conjuntos de parámetros de imagen o conjuntos de parámetros de secuencia. En otros ejemplos, el codificador de video (video encoder) 20 puede utilizar cualquier combinación de estas técnicas de ejemplo para señalar la información de identificación de imágenes de referencia.

45 **Tabla 1. Sintaxis de carga útil de secuencia de bytes en bruto o sin procesar (Rbsp) del conjunto de parámetros de secuencia**

	<b>Descriptor</b>
seq_parameter_set_rbsp( ) {	
<b>profile_idc</b>	u(8)
<b>reserved_zero_8bits</b> /* igual a 0 */	u(8)
<b>level_idc</b>	u(8)
<b>seq_parameter_set_id</b>	ue(v)
<b>max_temporal_layers_minus1</b>	u(3)
<b>pic_width_in_luma_samples</b>	u(16)
<b>pic_height_in_luma_samples</b>	u(16)
<b>bit_depth_luma_minus8</b>	ue(v)
<b>bit_depth_chroma_minus8</b>	ue(v)
<b>pcm_bit_depth_luma_minus1</b>	u(4)
<b>pcm_bit_depth_chroma_minus1</b>	u(4)
<b>log2_max_pic_order_cnt_lsb_minus4</b>	ue(v)
<b>max_num_ref_frames</b>	ue(v)
<b>log2_min_coding_block_size_minus3</b>	ue(v)
<b>log2_diff_max_min_coding_block_size</b>	ue(v)

<b>log2_min_transform_block_size_minus2</b>	ue(v)
<b>log2_diff_max_min_transform_block_size</b>	ue(v)
<b>log2_min_pcm_coding_block_size_minus3</b>	ue(v)
<b>max_transform_hierarchy_depth_inter</b>	ue(v)
<b>max_transform_hierarchy_depth_intra</b>	ue(v)
<b>chroma_pred_from_luma_enabled_flag</b>	u(1)
<b>loop_filter_across_slice_flag</b>	u(1)
<b>sample_adaptive_offset_enabled_flag</b>	u(1)
<b>adaptive_loop_filter_enabled_flag</b>	u(1)
<b>pcm_loop_filter_disable_flag</b>	u(1)
<b>cu_qp_delta_enabled_flag</b>	u(1)
<b>temporal_id_nesting_flag</b>	u(1)
<b>inter_4 × 4_enabled_flag</b>	u(1)
<b>rsbsp_trailing_bits()</b>	
}	

**pic\_width\_in\_luma\_samples** puede especificar el ancho de cada imagen decodificada en muestras de luma. El valor de **pic\_width\_in\_luma\_samples** puede estar en el rango de 0 hasta  $2^{16} - 1$ , inclusive.

5

**pic\_height\_in\_luma\_samples** puede especificar la altura de cada imagen decodificada en muestras de luma. El valor de **pic\_height\_in\_luma\_samples** puede estar en el rango de 0 hasta  $2^{16} - 1$ , inclusive.

Según se indica en la Tabla 1, el decodificador de video 30 puede recibir en el conjunto de parámetros de secuencia (SPS) el elemento de sintaxis **log2\_max\_pic\_order\_cnt\_lsb\_minus4**. Según se ha descrito anteriormente, el valor de **log2\_max\_pic\_order\_cnt\_lsb\_minus4** puede especificar el valor de la variable **MaxPicOrderCntLsb** que el decodificador de video 30 utiliza en el proceso de decodificación para determinar los valores de POC, en el que  $MaxPicOrderCntLsb = 2^{(log2\_max\_pic\_order\_cnt\_lsb\_minus4 + 4)}$ .

15 **Tabla 2. Sintaxis de carga útil de secuencia de bytes en bruto o sin procesar (Rbsp) del conjunto de parámetros de imagen**

	<b>Descriptor</b>
<b>pic_parameter_set_rbsp()</b> {	
<b>pic_parameter_set_id</b>	ue(v)
<b>seq_parameter_set_id</b>	ue(v)
<b>entropy_coding_mode_flag</b>	u(1)
<b>num_short_term_ref_pic_sets_pps</b>	ue(v)
para(i = 0; i < num_short_term_ref_pic_sets_pps; i++)	
short_term_ref_pic_set()	
<b>long_term_ref_pics_present_flag</b>	u(1)
si(long_term_ref_pics_present_flag) {	
<b>long_term_ref_pic_id_delta_len_minus4</b>	ue(v)
<b>long_term_ref_pic_id_len_delta</b>	ue(v)
<b>num_long_term_ref_pics_pps</b>	ue(v)
para(i = 0; i < num_long_term_ref_pics_pps; i++)	
<b>long_term_ref_pic_id_pps[i]</b>	i(v)
}	
<b>num_temporal_layer_switching_point_flags</b>	ue(v)
para(i = 0; i < num_temporal_layer_switching_point_flags; i++)	
<b>temporal_layer_switching_point_flag[i]</b>	u(1)
<b>num_ref_idx_10_default_active_minus1</b>	ue(v)
<b>num_ref_idx_11_default_active_minus1</b>	ue(v)
<b>pic_init_qp_minus26</b> /* en relación a 26 */	se(v)
<b>constrained_intra_pred_flag</b>	u(1)
<b>slice_granularity</b>	u(2)
<b>shared_pps_info_enabled_flag</b>	u(1)
si(shared_pps_info_enabled_flag)	
si(adaptive_loop_filter_enabled_flag)	
alf_param()	
si(cu_qp_delta_enabled_flag)	
<b>max_cu_qp_delta_depth</b>	u(4)

rbsp_trailing_bits( )	
}	

**num\_short\_term\_ref\_pic\_sets\_pps** especifica el número de estructuras de sintaxis short\_term\_ref\_pic\_set( ) incluidas en el conjunto de parámetros de imagen. El valor de num\_short\_term\_ref\_pic\_sets\_pps estará en el rango de 0 hasta 32, ambos inclusive.

**long\_term\_ref\_pics\_present\_flag** igual a 0 especifica que no se utiliza ninguna imagen de referencia de largo plazo para la inter predicción de ninguna imagen codificada que haga referencia al conjunto de parámetros de imagen y no están presentes los elementos de sintaxis long\_term\_ref\_pic\_id\_delta\_len\_minus4, long\_term\_ref\_pic\_id\_len\_delta y num\_long\_term\_ref\_pics\_pps. El indicador long\_term\_ref\_pics\_present\_flag igual a 1 especifica que se pueden utilizar imágenes de referencia de largo plazo para la inter predicción de una o más imágenes codificadas que hacen referencia al conjunto de parámetros de imagen y están presentes los elementos de sintaxis long\_term\_ref\_pic\_id\_delta\_len\_minus4, long\_term\_ref\_pic\_id\_len\_delta y num\_long\_term\_ref\_pics\_pps.

**long\_term\_ref\_pic\_id\_delta\_len\_minus4** más 4 especifica la longitud en bits de los elementos de sintaxis long\_term\_ref\_pic\_id\_delta\_add\_foll[i]. El valor de long\_term\_ref\_pic\_id\_delta\_len\_minus4 estará en el rango de 0 hasta 12, ambos inclusive.

**long\_term\_ref\_pic\_id\_len\_delta** más long\_term\_ref\_pic\_id\_delta\_len\_minus4 más 4 especifica la longitud en bits del elemento de sintaxis long\_term\_ref\_pic\_id\_pps[i]. El valor de long\_term\_ref\_pic\_id\_len\_delta puede estar en el rango de 0 hasta 28 – long\_term\_ref\_pic\_id\_delta\_len\_minus4, inclusive. El valor de long\_term\_ref\_pic\_id\_len\_delta + long\_term\_ref\_pic\_id\_delta\_len\_minus4 + 4 en todos los conjuntos de parámetros de imagen que hacen referencia a un conjunto de parámetros de secuencia en particular puede ser idéntico.

**num\_long\_term\_ref\_pics\_pps** especifica el número de identificaciones de imágenes de referencia de largo plazo incluidas en el conjunto de parámetros de imagen. El valor de num\_long\_term\_ref\_pics\_pps puede estar en el rango de 0 hasta 32, inclusive.

**long\_term\_ref\_pic\_id\_pps[i]** especifica información de identificación de la i-ésima imagen de referencia de largo plazo incluida en el conjunto de parámetros de imagen. El número de bits utilizados para representar long\_term\_ref\_pic\_id\_pps[i] puede ser igual a long\_term\_ref\_pic\_id\_len\_delta + long\_term\_ref\_pic\_id\_len\_minus4 + 4.

**Tabla 3. Sintaxis de conjunto de imágenes de referencia de corto plazo**

35

short term ref pic set( ) {	Descriptor
<b>num_short_term_curr0</b>	ue(v)
<b>num_short_term_curr1</b>	ue(v)
<b>num_short_term_foll0</b>	ue(v)
<b>num_short_term_foll1</b>	ue(v)
NumShortTerm = num_short_term_curr0 + num_short_term_curr1 + num_short_term_foll0 + num_short_term_foll1	
para( i = 0; i < NumShortTerm; i++)	
<b>short_term_ref_pic_id_delta_minus1[i]</b>	ue(v)
}	

La sintaxis del conjunto de imágenes de referencia de corto plazo puede ser para imágenes de corto plazo. Una imagen de corto plazo se puede definir como una imagen de referencia cuya información de identificación se incluye en la estructura de sintaxis short\_term\_ref\_pic\_set( ) de una imagen codificada, bien incluida en la cabecera o cabeceras de corte (slice) o bien incluida en el conjunto de parámetros de imagen y referencia por medio del elemento de sintaxis short\_term\_ref\_pic\_set\_idx en la cabecera o cabeceras de corte (slice). Los elementos de sintaxis de la cabecera de corte (slice) se presentan más adelante en la Tabla 4.

**num\_short\_term\_curr0** especifica el número de imágenes de referencia de corto plazo en RefPicSetStCurr0 cuando se utiliza la estructura de sintaxis short\_term\_ref\_pic\_set( ) para derivar el conjunto de imágenes de referencia de una imagen codificada, según se describe a continuación. El valor de num\_short\_term\_curr0 puede estar en el rango de 0 hasta max\_num\_ref\_frames, inclusive.

**num\_short\_term\_curr1** especifica el número de imágenes de referencia de corto plazo en RefPicSetStCurr1 cuando se utiliza la estructura de sintaxis short\_term\_ref\_pic\_set( ) para derivar el conjunto de imágenes de referencia de una

imagen codificada, según se describe a continuación. El valor de num\_short\_term\_curr1 puede estar en el rango de 0 hasta max\_num\_ref\_frames – num\_short\_term\_curr0, inclusive.

5 **num\_short\_term\_foll0** especifica el número de imágenes de referencia de corto plazo en RefPicSetStFoll0 cuando se utiliza la estructura de sintaxis short\_term\_ref\_pic\_set( ) para derivar el conjunto de imágenes de referencia de una imagen codificada, según se describe a continuación. El valor de num\_short\_term\_foll0 puede estar en el rango de 0 hasta max\_num\_ref\_frames – num\_short\_term\_curr0 – num\_short\_term\_curr1, inclusive.

10 **num\_short\_term\_foll1** especifica el número de imágenes de referencia de corto plazo en RefPicSetStFoll1 cuando se utiliza la estructura de sintaxis short\_term\_ref\_pic\_set( ) para derivar el conjunto de imágenes de referencia de una imagen codificada, según se describe a continuación. El valor de num\_short\_term\_foll1 estará en el rango de 0 hasta max\_num\_ref\_frames – num\_short\_term\_curr0 – num\_short\_term\_curr1 – num\_short\_term\_foll0, inclusive.

15 **short\_term\_ref\_pic\_id\_delta\_minus1[i]** especifica la información de identificación de la i-ésima imagen de referencia de corto plazo incluida en la estructura de sintaxis short\_term\_ref\_pic\_set( ).

**Tabla 4. Sintaxis de cabecera de corte (slice)**

	<b>Descriptor</b>
slice_header() {	
<b>lightweight slice flag</b>	u(1)
si( !lightweight slice flag ) {	
<b>slice type</b>	ue(v)
<b>pic parameter set id</b>	ue(v)
si( !IdrPicFlag ) {	
<b>idr pic id</b>	ue(v)
<b>no output of prior pics flag</b>	u(1)
}	
<b>pic order cnt lsb</b>	u(v)
si( !IdrPicFlag ) {	
<b>short term ref pic set pps flag</b>	u(1)
si( short term ref pic set pps flag )	
<b>short term ref pic set idx</b>	ue(v)
Sino	
short_term_ref_pic_set()	
si( long_term_ref_pics_present_flag )	
long_term_ref_pic_set()	
}	
si( slice_type == P    slice_type == B ) {	
<b>num ref idx active override flag</b>	u(1)
si( num_ref_idx_active_override_flag ) {	
<b>num ref idx l0 active minus1</b>	ue(v)
si( slice_type == B )	
<b>num ref idx l1 active minus1</b>	ue(v)
}	
}	
ref_pic_list_modification()	
ref_pic_list_combination()	
}	
si(entropy_coding_mode_flag && slice_type != I)	
<b>cabac init idc</b>	ue(v)
<b>first slice in pic flag</b>	u(1)
si( first slice in pic flag == 0 )	
<b>slice address</b>	u(v)
si( !lightweight_slice_flag ) { (en inglés)	
<b>slice qp delta</b>	se(v)
si( sample_adaptive_offset_enabled_flag )	
sao_param()	
si( deblocking_filter_control_present_flag ) {	
<b>disable deblocking filter idc</b>	
si( disable_deblocking_filter_idc != 1 ) {	
<b>slice_alpha_c0_offset_div2</b>	

<b>slice_beta_offset_div2</b>	
}	
}	
si( slice_type == B )	
<b>collocated_from_I0_flag</b>	u(1)
si( adaptive_loop_filter_enabled_flag ) {	
si( !shared_pps_info_enabled_flag )	
alf_param( )	
alf_cu_control_param( )	
}	
}	
}	

**no\_output\_of\_prior\_pics\_flag** especifica cómo se tratan las imágenes decodificadas previamente en el búfer de imágenes decodificadas después de decodificar una imagen de IDR. Cuando la imagen de IDR es la primera imagen de IDR en el flujo de bits, es posible que el valor de **no\_output\_of\_prior\_pics\_flag** no tenga ningún efecto en el proceso de decodificación. Cuando la imagen de IDR no es la primera imagen de IDR en el flujo de bits y el valor de **pic\_width\_in\_luma\_samples** o **pic\_height\_in\_luma\_samples** o **max\_dec\_frame\_buffering** derivado a partir del conjunto de parámetros de secuencia activo puede ser diferente del valor de **pic\_width\_in\_luma\_samples** o **pic\_height\_in\_luma\_samples** o **max\_dec\_frame\_buffering** derivado a partir del conjunto de parámetros de secuencia activo para la imagen anterior, **no\_output\_of\_prior\_pics\_flag** igual a 1 puede ser, pero no necesariamente, inferido por el decodificador, con independencia del valor real de **no\_output\_of\_prior\_pics\_flag**.

**short\_term\_ref\_pic\_set\_pps\_flag** igual a 1 especifica que la información de identificación del conjunto de imágenes de referencia de corto plazo incluida en el conjunto de imágenes de referencia para la imagen actual se encuentra presente en el conjunto de parámetros de imagen referenciado. **short\_term\_ref\_pic\_set\_pps\_flag** igual a 0 especifica que la información de identificación del conjunto de imágenes de referencia de corto plazo incluida en el conjunto de imágenes de referencia para la imagen actual no se encuentra presente en el conjunto de parámetros de imagen referenciado.

**short\_term\_ref\_pic\_set\_idx** especifica el índice de la estructura de sintaxis **short\_term\_ref\_pic\_set( )**, que se incluye en el conjunto de parámetros de imagen referenciado, que incluye la información de identificación del conjunto de imágenes de referencia de corto plazo en el conjunto de imágenes de referencia para la imagen actual.

Las variables **NumShortTermCurr0** y **NumShortTermCurr1** se especifican de la siguiente manera:

**NumShortTermCurr0** = num\_short\_term\_curr0

**NumShortTermCurr1** = num\_short\_term\_curr1

En el que **num\_short\_term\_curr0** y **num\_short\_term\_curr1** son los elementos de sintaxis de los mismos nombres, respectivamente, en la estructura de sintaxis **short\_term\_ref\_pic\_set( )**, bien presente en el conjunto de parámetros de imagen referenciado y que está referenciado por el índice **short\_term\_ref\_pic\_set\_idx**, o bien presente directamente en la cabecera de corte (slice).

**num\_ref\_idx\_l0\_active\_minus1** especifica el índice de referencia máximo para la lista de imágenes de referencia 0 a utilizar para decodificar el corte (slice).

Cuando el corte (slice) actual es un corte P o B y no se encuentra presente **num\_ref\_idx\_l0\_active\_minus1**, se puede inferir que **num\_ref\_idx\_l0\_active\_minus1** es igual a **num\_ref\_idx\_l0\_default\_active\_minus1**.

El valor de **num\_ref\_idx\_l0\_active\_minus1** puede estar en el rango de 0 hasta 15, inclusive.

**num\_ref\_idx\_l1\_active\_minus1** especifica el índice de referencia máximo para la Lista 1 de imágenes de referencia a utilizar para decodificar el corte (slice).

Cuando el corte (slice) actual es un corte P o B y no se encuentra presente **num\_ref\_idx\_l1\_active\_minus1**, se puede inferir que **num\_ref\_idx\_l1\_active\_minus1** es igual a **num\_ref\_idx\_l1\_default\_active\_minus1**.

El valor de **num\_ref\_idx\_l1\_active\_minus1** puede estar en el rango de 0 hasta 15, inclusive.

50

**Tabla 5. Sintaxis de conjunto de imágenes de referencia de largo plazo**



	Descriptor
long_term_ref_pic_set() {	
<b>num_long_term_pps_curr</b>	ue(v)
<b>num_long_term_add_curr</b>	ue(v)
<b>num_long_term_pps_foll</b>	ue(v)
<b>num_long_term_add_foll</b>	ue(v)
para( i = 0; i < num_long_term_pps_curr + num_long_term_pps_foll; i++ )	
<b>long_term_ref_pic_set_idx_pps[i]</b>	ue(v)
para( i = 0; i < num_long_term_add_curr + num_long_term_add_foll; i++ )	
<b>long_term_ref_pic_id_delta_add[i]</b>	i(v)
}	

La sintaxis del conjunto de imágenes de referencia de largo plazo puede ser para imágenes de largo plazo. Una imagen de largo plazo se puede definir como una imagen de referencia cuya información de identificación se incluye en la estructura de sintaxis long\_term\_ref\_pic\_set() para una imagen codificada.

**num\_long\_term\_pps\_curr** especifica el número de todas las imágenes de referencia de largo plazo que la información de identificación está incluida en el conjunto de parámetros de imagen referenciado y que se pueden utilizar para la inter predicción de la imagen actual. Si num\_long\_term\_pps\_curr no se encuentra presente, el valor se puede derivar igual a 0. El valor de num\_long\_term\_pps\_curr puede estar en el rango de 0 hasta max\_num\_ref\_frames, inclusive.

**num\_long\_term\_add\_curr** especifica el número de todas las imágenes de referencia de largo plazo que la información de identificación no está incluida en el conjunto de parámetros de imagen referenciado y que se puede utilizar para la inter predicción de la imagen actual. Si num\_long\_term\_add\_curr no se encuentra presente, el valor se puede derivar igual a 0. El valor de num\_long\_term\_add\_curr puede estar en el rango de 0 hasta max\_num\_ref\_frames – num\_long\_term\_pps\_curr, inclusive.

La variable NumLongTermCurr se especifica de la siguiente manera:

$$\text{NumLongTermCurr} = \text{num\_long\_term\_pps\_curr} + \text{num\_long\_term\_add\_curr}$$

**num\_long\_term\_pps\_foll** especifica el número de todas las imágenes de referencia de largo plazo que la información de identificación está incluida en el conjunto de parámetros de imagen referenciado, que no se utilizan para la inter predicción de la imagen actual y que se pueden utilizar para la inter predicción de cualquiera de las imágenes posteriores a la imagen actual en orden de decodificación. Si num\_long\_term\_pps\_foll no se encuentra presente, el valor se puede derivar igual a 0. El valor de num\_long\_term\_pps\_foll puede estar en el rango de 0 hasta max\_num\_ref\_frames, inclusive.

**num\_long\_term\_add\_foll** especifica el número de todas las imágenes de referencia de largo plazo que la información de identificación no está incluida en el conjunto de parámetros de imagen referenciado, que no se utilizan para la inter predicción de la imagen actual y que se pueden utilizar para la inter predicción de cualquiera de las imágenes subsiguientes en orden de decodificación. Si num\_long\_term\_add\_foll no se encuentra presente, el valor se puede derivar igual a 0. El valor de num\_long\_term\_add\_foll puede estar en el rango de 0 hasta max\_num\_ref\_frames – num\_long\_term\_pps\_foll, inclusive.

**long\_term\_ref\_pic\_set\_idx\_pps[i]** especifica el índice, a la lista de información de identificación de imágenes de referencia de largo plazo incluida en el conjunto de parámetros de imagen de referencia, de la i-ésima imagen de referencia de largo plazo que se hereda del conjunto de parámetros de imagen referenciado al conjunto de imágenes de referencia de la imagen actual. El valor de long\_term\_ref\_pic\_set\_idx\_pps[i] puede estar en el rango de 0 hasta 31, inclusive.

**long\_term\_ref\_pic\_id\_delta\_add[i]** especifica la información de identificación de imágenes de referencia de largo plazo de la i-ésima imagen de referencia de largo plazo que no se hereda del conjunto de parámetros de imagen referenciado sino que está incluida en el conjunto de imágenes de referencia de la imagen actual. El número de bits utilizados para representar long\_term\_ref\_pic\_id\_add\_curr[i] puede ser igual a long\_term\_pic\_id\_len\_minus4 + 4.

Con los valores arriba señalizados o derivados (es decir, los valores de las Tablas 1 – 5), el decodificador de video puede derivar el conjunto de imágenes de referencia. Según se ha descrito anteriormente, el conjunto de imágenes de referencia derivado puede identificar imágenes de referencia que se pueden utilizar potencialmente para codificar / pronosticar la imagen actual (es decir, la imagen que se está decodificando actualmente) e imágenes posteriores a la imagen actual en un orden de decodificación. De acuerdo con las técnicas que se describen en esta divulgación, el

orden de decodificación de todas las imágenes de referencia en el conjunto de imágenes de referencia derivado es anterior al orden de decodificación de la imagen actual.

El proceso de derivación puede incluir la construcción del conjunto de imágenes de referencia a partir de una pluralidad de subconjuntos de imágenes de referencia. Este proceso puede ser invocado una vez por cada imagen, después de la decodificación de una cabecera de corte (slice) pero antes de la decodificación de cualquier unidad de codificación y antes del proceso de decodificación para la construcción de la lista de imágenes de referencia del corte (slice). Por ejemplo, a partir del valor o valores derivados y los elementos de sintaxis señalizados, el decodificador de video 30 puede determinar los valores de POC para las imágenes de referencia que pertenecen al conjunto de imágenes de referencia. A partir del valor o valores de POC determinados, el decodificador de video 30 puede construir subconjuntos de imágenes de referencia que juntos forman el conjunto de imágenes de referencia. De esta manera, construyendo los subconjuntos de imágenes de referencia, el decodificador de video 30 puede construir el conjunto de imágenes de referencia. Por ejemplo, el decodificador de video 30 puede ordenar los subconjuntos de imágenes de referencia de una manera particular para derivar el conjunto de imágenes de referencia. La ordenación se puede referir a la manera en la que el decodificador de video 30 lista los subconjuntos de imágenes de referencia para derivar el conjunto de imágenes de referencia.

Según se ha descrito anteriormente, para derivar el conjunto de imágenes de referencia, el decodificador de video 30 puede construir una pluralidad de subconjuntos de imágenes de referencia. En algunos ejemplos, el decodificador de video 30 puede construir seis subconjuntos de imágenes de referencia. Los seis subconjuntos de imágenes de referencia se pueden denominar: RefPicSetStCurr0, RefPicSetStCurr1, RefPicSetStFoll0, RefPicSetStFoll1, RefPicSetLtCurr y RefPicSetLtFoll. RefPicSetStCurr0 se puede denominar RefPicSetStCurrBefore, y RefPicSetStCurr1 se puede denominar RefPicSetStCurrAfter.

Se debe entender que los seis subconjuntos de imágenes de referencia se describen a efectos de ilustración y no se deben interpretar como limitantes. Por ejemplo, el decodificador de video 30 puede construir menos de seis subconjuntos de imágenes de referencia, por ejemplo, combinando algunos de los subconjuntos. A continuación se describen algunos de estos ejemplos en los que el decodificador de video 30 construye menos de seis subconjuntos de imágenes de referencia. Sin embargo, a efectos de ilustración, las técnicas se describen con ejemplos en los que el decodificador de video 30 construye seis subconjuntos de imágenes de referencia.

Los subconjuntos de imágenes de referencia RefPicSetStCurr0, RefPicSetStCurr1, RefPicSetStFoll0 y RefPicSetStFoll1 pueden identificar imágenes de referencia de corto plazo. En algunos ejemplos, estos subconjuntos de imágenes de referencia pueden identificar imágenes de referencia de corto plazo en base a si las imágenes de referencia de corto plazo son anteriores en el orden de visualización o posteriores en el orden de visualización con respecto a la imagen actual que se está codificando, así como si las imágenes de referencia de corto plazo se pueden utilizar potencialmente para la inter predicción de la imagen actual y de imágenes posteriores a la imagen actual en orden de decodificación, o se pueden utilizar potencialmente para la inter predicción de sólo las imágenes posteriores a la imagen actual en el orden de decodificación.

Por ejemplo, el subconjunto de imágenes de referencia RefPicSetStCurr0 puede incluir, y puede incluir sólo, información de identificación, tal como valores de POC, de todas las imágenes de referencia de corto plazo que tienen un orden de output o de visualización anterior a la imagen actual, y que se pueden utilizar potencialmente como referencia en la inter predicción de la imagen actual, y que se pueden utilizar potencialmente como referencia en la inter predicción de una o más imágenes posteriores a la imagen actual en orden de decodificación. El subconjunto de imágenes de referencia RefPicSetStCurr1 puede incluir, y puede incluir sólo, información de identificación de todas las imágenes de referencia de corto plazo que tienen un orden de output o de visualización posterior al de la imagen actual y se puede utilizar potencialmente como referencia en la inter predicción de la imagen actual, y que se pueden utilizar potencialmente como referencia en la inter predicción de una o más imágenes posteriores a la imagen actual en orden de decodificación.

El subconjunto de imágenes de referencia RefPicSetStFoll0 puede incluir, y puede incluir sólo, información de identificación de todas las imágenes de referencia de corto plazo que tienen un orden de output o de visualización anterior al de la imagen actual, que se pueden utilizar potencialmente como referencia en la inter predicción de una o más imágenes posteriores a la imagen actual en un orden de decodificación, y que no se pueden utilizar como referencia en la inter predicción de la imagen actual. El subconjunto de imágenes de referencia RefPicSetStFoll1 puede incluir, y puede incluir sólo, información de identificación de todas las imágenes de referencia de corto plazo que tienen un orden de output o de visualización posterior al de la imagen actual, que se pueden utilizar potencialmente como referencia en la inter predicción de una o más imágenes posteriores a la imagen actual en un orden de decodificación y que no se pueden utilizar como referencia en la inter predicción de la imagen actual.

Los subconjuntos de imágenes de referencia RefPicSetLtCurr y RefPicSetLtFoll pueden identificar imágenes de referencia de largo plazo. En algunos ejemplos, estos subconjuntos de imágenes de referencia pueden identificar

imágenes de referencia de largo plazo en base a si las imágenes de referencia de largo plazo son anteriores en el orden de visualización o posteriores en el orden de visualización con respecto a la imagen actual que se está convirtiendo según codificación.

- 5 Por ejemplo, el subconjunto de imágenes de referencia RefPicSetLtCurr puede incluir, y puede incluir sólo, la información de identificación de todas las imágenes de referencia de largo plazo que se pueden utilizar potencialmente como referencia en la inter predicción de la imagen actual, y que se pueden utilizar potencialmente como referencia en la inter predicción de una o más imágenes posteriores a la imagen actual en orden de decodificación. El subconjunto de imágenes de referencia RefPicSetLtFoll puede incluir, y puede incluir sólo, la información de identificación de todas las imágenes de referencia de largo plazo que se pueden utilizar potencialmente como referencia en la inter predicción de una o más imágenes posteriores a la imagen actual en orden de decodificación, y que no se pueden utilizar como referencia en la inter predicción de la imagen actual.

- 15 Si la imagen actual a decodificar es una imagen de IDR, el decodificador de video 30 puede vaciar los subconjuntos de imágenes de referencia RefPicSetStCurr0, RefPicSetStCurr1, RefPicSetStFoll0, RefPicSetStFoll1, RefPicSetLtCurr y RefPicSetLtFoll. Esto se puede deber a que la imagen de IDR puede no ser inter pronosticada y a que ninguna imagen posterior a la imagen de IDR en un orden de decodificación puede utilizar cualquier imagen anterior a la imagen de IDR en la decodificación como referencia. De lo contrario (por ejemplo, cuando la imagen actual no es una imagen de IDR), el decodificador de video 30 puede construir los subconjuntos de imágenes de referencia de corto plazo y los subconjuntos de imágenes de referencia de largo plazo implementando el siguiente pseudo código.

- 25 Por ejemplo, cuando el decodificador de video 30 decodifica una instancia de la estructura de sintaxis short\_term\_ref\_pic\_set( ), ya sea en la cabecera de corte (slice) o por referencia al conjunto de parámetros de imagen referenciado, el decodificador de video 30 puede implementar el siguiente pseudo código para construir los subconjuntos de imágenes de referencia RefPicSetStCurr0, RefPicSetStCurr1, RefPicSetStFoll0 y RefPicSetStFoll1.

```

cIdx = 0
para( i = 0, pocPred = PicOrderCnt( CurrPic ); i < num_short_term_curr0; pocPred =
RefPicSetStCurr0[i], i++, cIdx++ )
30 RefPicSetStCurr0[i] = pocPred - short_term_ref_pic_id_delta_minus1[ cIdx ] - 1

para( i = 0, pocPred = PicOrderCnt( CurrPic ); i < num_short_term_curr1; pocPred =
RefPicSetStCurr1[i], i++, cIdx++ )
35 RefPicSetStCurr1[i] = short_term_ref_pic_id_delta_minus1[ cIdx ] + 1 - pocPred

para( i = 0, pocPred = PicOrderCnt( CurrPic ); i < num_short_term_foll0; pocPred =
RefPicSetStFoll0[i], i++, cIdx++ )
40 RefPicSetStFoll0[i] = pocPred - short_term_ref_pic_id_delta_minus1[ cIdx ] - 1

para( i = 0, pocPred = PicOrderCnt( CurrPic ); i < num_short_term_foll1; pocPred =
RefPicSetStFoll1[i], i++, cIdx++ )
45 RefPicSetStFoll1[i] = short_term_ref_pic_id_delta_minus1[ cIdx ] + 1 - pocPred

```

- 45 Si el decodificador de video 30 determina que el indicador long\_term\_ref\_pics\_present\_flag es igual a 0, el decodificador de video 30 puede establecer RefPicSetLtCurr y RefPicSetLtFoll como vacías porque, en este caso, no hay imágenes de referencia de largo plazo. De lo contrario, si el decodificador de video 30 decodifica una instancia de la estructura de sintaxis long\_term\_ref\_pic\_set( ) en la cabecera de corte (slice), el decodificador de video 30 puede implementar el siguiente pseudo código para construir los subconjuntos de imágenes de referencia RefPicSetLtCurr y RefPicSetLtFoll.

```

cIdx = 0
para( i = 0; i < num_long_term_pps_curr; i++, cIdx++ )
{
55 pIdx = long_term_ref_pic_idx_pps[i]
RefPicSetLtCurr[ cIdx ] = long_term_ref_pic_id_pps[ pIdx ]
}

para( i = 0; i < num_long_term_add_curr; i++, cIdx++ )
60 {
picIdDelta = long_term_ref_pic_id_delta_add[i]
RefPicSetLtCurr[ cIdx ] = PicOrderCnt( CurrPic ) - picIdDelta
}

```

```

cIdx = 0
para( i = 0; i < num_long_term_pps_foll; i++, cIdx++ )
{
5   pIdx = long_term_ref_pic_idx_pps[ i + num_long_term_pps_curr ]
   RefPicSetLtFoll[ cIdx ] = long_term_ref_pic_id_pps[ pIdx ]
}

para( i = 0; i < num_long_term_add_foll; i++, cIdx++ )
10 {
   picIdDelta = long_term_ref_pic_id_delta_add[ i + num_long_term_add_curr ]
   RefPicSetLtFoll[ cIdx ] = PicOrderCnt( CurrPic ) - picIdDelta
}

```

15

De acuerdo con las técnicas que se describen en esta divulgación, una imagen de referencia con un valor particular de PicOrderCnt (valor de POC) se puede denominar incluida en el conjunto de imágenes de referencia de una imagen codificada si la imagen de referencia está incluida en cualquiera de los seis subconjuntos del conjunto de imágenes de referencia de esa imagen codificada. Una imagen de referencia con un valor particular de PicOrderCnt es referenciada como incluida en un subconjunto particular del conjunto de imágenes de referencia si el valor particular de PicOrderCnt (valor de POC) es igual a uno de los valores de PicOrderCnt incluidos en dicho subconjunto.

Después de construir los subconjuntos de imágenes de referencia, el decodificador de video 30 puede derivar el conjunto de imágenes de referencia. Por ejemplo, el decodificador de video 30 puede ordenar los subconjuntos de imágenes de referencia para derivar el conjunto de imágenes de referencia. Por ejemplo, el decodificador de video 30 puede listar el subconjunto de imágenes de referencia RefPicSetStCurr0, seguido del subconjunto de imágenes de referencia RefPicSetStCurr1, seguido del subconjunto de imágenes de referencia RefPicSetStFoll0, seguido del subconjunto de imágenes de referencia RefPicSetStFoll1, seguido del subconjunto de imágenes de referencia RefPicSetLtCurr y, a continuación, del subconjunto de imágenes de referencia RefPicSetLtFoll. En otro ejemplo, el decodificador de video 30 puede listar el subconjunto de imágenes de referencia RefPicSetStCurr0, seguido del subconjunto de imágenes de referencia RefPicSetStCurr1, seguido del subconjunto de imágenes de referencia RefPicSetLtCurr, seguido del subconjunto de imágenes de referencia RefPicSetStFoll0, seguido del subconjunto de imágenes de referencia RefPicSetStFoll1 y, a continuación, del subconjunto de imágenes de referencia RefPicSetLtFoll.

35

Otras permutaciones de la manera en la que el decodificador de video 30 ordena los subconjuntos de imágenes de referencia pueden ser posibles para derivar el conjunto de imágenes de referencia. En algunos ejemplos, se pueden combinar la construcción de los subconjuntos de imágenes de referencia y la derivación del conjunto de imágenes de referencia. Por ejemplo, la construcción de los subconjuntos de imágenes de referencia puede dar lugar automáticamente a que el decodificador de video 30 derive el conjunto de imágenes de referencia. En otras palabras, el decodificador de video 30 puede no necesitar realizar diferentes etapas para construir los subconjuntos de imágenes de referencia y derivar el conjunto de imágenes de referencia, aunque puede ser posible que el decodificador de video 30 primero construya los subconjuntos de imágenes de referencia y luego derive el conjunto de imágenes de referencia.

45

Además, de acuerdo con las técnicas que se describen en esta divulgación, la construcción del conjunto de imágenes de referencia de la manera que se ha descrito anteriormente puede dar lugar a que el decodificador de video 30 cumpla con las siguientes restricciones. Por ejemplo, una imagen de referencia particular con un valor particular de PicOrderCnt puede no incluirse en más de uno de los subconjuntos de imágenes de referencia del conjunto de imágenes de referencia de la imagen actual. En otras palabras, una imagen de referencia identificada en uno de los subconjuntos de imágenes de referencia puede no identificarse en ninguno de los otros subconjuntos de imágenes de referencia. En otro ejemplo, en el conjunto de imágenes de referencia derivado, puede que no haya ninguna imagen de referencia con temporal\_id superior a la imagen actual que esté incluida en cualquiera de los subconjuntos de imágenes de referencia que forman el conjunto de imágenes de referencia.

55

Según se ha descrito anteriormente, el valor de identificación temporal (temporal\_id) puede ser un valor jerárquico que indica qué imágenes se pueden utilizar para codificar / pronosticar la imagen actual. En general, una imagen con un determinado valor de temporal\_id puede ser una imagen de referencia para imágenes con valores de temporal\_id iguales o superiores, pero no vice versa. Por ejemplo, una imagen con un valor de temporal\_id igual a 1 puede ser una imagen de referencia para imágenes con valores de temporal\_id iguales a 1, 2, 3,..... pero no para una imagen con un valor de temporal\_id igual a 0.

60

El valor de `temporal_id` más bajo también puede indicar la velocidad de visualización más baja. Por ejemplo, si el decodificador de video 30 sólo decodifica imágenes con valores de `temporal_id` iguales a 0, la velocidad o tasa de visualización puede ser de 7,5 imágenes por segundo. Si el decodificador de video 30 sólo decodifica imágenes con valores de `temporal_id` iguales a 0 y 1, la velocidad o tasa de visualización puede ser de 15 imágenes por segundo, y así sucesivamente.

En algunos ejemplos, sólo se pueden incluir en el conjunto de imágenes de referencia de la imagen actual imágenes con valores de `temporal_id` inferiores o iguales al `temporal_id` de la imagen actual. Según se ha descrito anteriormente, sólo se pueden utilizar como imágenes de referencia imágenes con valores de `temporal_id` inferiores o iguales al valor de `temporal_id` de la imagen actual. Por lo tanto, todas las imágenes de referencia con valores de `temporal_id` inferiores o iguales pueden ser utilizadas por la imagen actual para la inter predicción y pueden ser incluidas en el conjunto de imágenes de referencia. Además, se excluyen del conjunto de imágenes de referencia algunas imágenes de referencia que tienen valores de `temporal_id` mayores que la imagen actual y que van a ser utilizadas por imágenes posteriores a la imagen actual en orden de decodificación y que tienen valores de `temporal_id` mayores que la imagen actual.

Con estas técnicas, no es necesaria la señalización del `temporal_id` adicionalmente a la identificación de imágenes para la derivación del conjunto de imágenes de referencia; por lo tanto, la señalización del conjunto de imágenes de referencia resulta más eficiente. Por ejemplo, es posible que el codificador de video (video encoder) 20 no señalice valores del `temporal_id` de las imágenes de referencia que pertenecen al conjunto de imágenes de referencia, y que el decodificador de video 30 no necesite recibir valores del `temporal_id` de las imágenes de referencia que pertenecen al conjunto de imágenes de referencia con el propósito de derivar el conjunto de imágenes de referencia.

Además, de esta manera, los subconjuntos de imágenes de referencia construidos no pueden identificar imágenes de referencia con valores de `temporal_id` superiores a los de la imagen actual. Por ejemplo, el decodificador de video 30 puede ser capaz de construir los subconjuntos de imágenes de referencia y garantizar que ninguna imagen de referencia identificada en alguno de los subconjuntos de imágenes de referencia tiene un valor de `temporal_id` superior al de la imagen actual, ya que la conformidad con el flujo de bits puede requerir que los valores de `temporal_id` no se incluyan en el flujo de bits señalizado por el codificador de video (video encoder) 20 y recibido por el decodificador de video 30. De esta manera, el decodificador de video 30 puede derivar el conjunto de imágenes de referencia sin recibir valores de identificación temporal para las imágenes de referencia que pertenecen al conjunto de imágenes de referencia.

En los ejemplos anteriores, el decodificador de video 30 puede construir seis subconjuntos de imágenes de referencia, cuatro para imágenes de referencia de corto plazo (es decir, `RefPicSetStCurr0`, `RefPicSetStCurr1`, `RefPicSetStFoll0` y `RefPicSetStFoll1`), y dos para imágenes de referencia de largo plazo (es decir, `RefPicSetLtCurr` y `RefPicSetLtFoll`). Sin embargo, aspectos de esta divulgación no son tan limitados. En otros ejemplos, dos o más de estos subconjuntos de imágenes de referencia se pueden combinar en un subconjunto de imágenes de referencia, lo que resulta en menos subconjuntos de imágenes de referencia que construye el decodificador de video 30. A continuación se describen algunos ejemplos no limitativos en los que el decodificador de video 30 puede construir menos subconjuntos de imágenes de referencia. Puede haber otras maneras en las que el decodificador de video 30 puede construir menos subconjuntos de imágenes de referencia.

Por ejemplo, en algunos ejemplos, puede que no haya separación entre el subconjunto para la imagen actual y el subconjunto para las imágenes subsiguientes en orden de decodificación. Por lo tanto, puede haber dos subconjuntos para imágenes de referencia de corto plazo, denominados `RefPicSetSt0` y `RefPicSetSt1`, y puede haber sólo un subconjunto para imágenes de referencia de largo plazo, denominado `RefPicSetLt`. En este ejemplo, el subconjunto de imágenes de referencia `RefPicSetSt0` puede ser la concatenación de `RefPicSetStCurr0` y `RefPicSetStFoll0`, con `RefPicSetStCurr0` al principio del resultado de la concatenación. En este ejemplo, el subconjunto de imágenes de referencia `RefPicSetSt1` puede ser la concatenación de `RefPicSetStCurr1` y `RefPicSetStFoll1`, con `RefPicSetStCurr1` al principio del resultado de la concatenación. El subconjunto de imágenes de referencia `RefPicSetLt` puede ser la concatenación de `RefPicSetLtCurr` y `RefPicSetLtFoll`, con `RefPicSetLtCurr` al principio del resultado de la concatenación.

En otro ejemplo, puede que no haya separación del subconjunto con un orden de output anterior o posterior a la imagen actual. Esto se puede aplicar sólo a imágenes de referencia de corto plazo. Por lo tanto, puede haber dos subconjuntos para imágenes de referencia de corto plazo, denominados `RefPicSetStCurr` y `RefPicSetStFoll`. El subconjunto de imágenes de referencia `RefPicSetStCurr` puede ser la concatenación de `RefPicSetStCurr0` y `RefPicSetStCurr1`, con `RefPicSetStCurr0` al principio del resultado de la concatenación. El subconjunto de imágenes de referencia `RefPicSetStFoll` puede ser la concatenación de `RefPicSetStFoll0` y `RefPicSetStFoll1`, con `RefPicSetStFoll0` al principio del resultado de la concatenación.

En otro ejemplo, pueden no aplicarse ambos tipos de separaciones mencionadas anteriormente. Por lo tanto, puede haber sólo un subconjunto para imágenes de referencia de corto plazo, denominado `RefPicSetSt`, y sólo un

subconjunto para imágenes de referencia de largo plazo, denominado RefPicSetLt. El subconjunto de imágenes de referencia RefPicSetSt puede ser una concatenación de RefPicSetStCurr0, RefPicSetStCurr1, RefPicSetStFoll0 y RefPicSetStFoll1, en el orden listado (o en cualquier otro orden), y RefPicSetLt puede ser el mismo que antes.

- 5 Las técnicas anteriores describen un ejemplo de cómo el decodificador de video 30 puede derivar el conjunto de imágenes de referencia. Durante el proceso de codificación, el codificador de video (video encoder) 20 también puede necesitar decodificar imágenes codificadas con el fin de codificar imágenes subsiguientes, en lo que se denomina proceso de reconstrucción. Por consiguiente, en algunos ejemplos, el codificador de video (video encoder) 20 también puede estar configurado para derivar el conjunto de imágenes de referencia. En algunos ejemplos, el codificador de video (video encoder) 20 puede implementar las mismas técnicas que el decodificador de video 30 ha implementado para derivar el conjunto de imágenes de referencia. Sin embargo, la derivación del conjunto de imágenes de referencia por parte del codificador de video (video encoder) 20 puede no ser necesaria en todos los ejemplos, y el decodificador de video 30 puede ser el único convertidor según codificación que deriva el conjunto de imágenes de referencia.
- 10
- 15 En consecuencia, en algunos ejemplos, un codificador de video (video coder) (por ejemplo, el codificador de video (video encoder) 20 o el decodificador de video 30) puede convertir según una codificación (por ejemplo, codificar o decodificar, respectivamente) información indicativa de imágenes de referencia que pertenecen a un conjunto de imágenes de referencia. Por ejemplo, el codificador de video (video encoder) 20 puede señalar un flujo de bits codificado que incluye valores para determinar qué imágenes de referencia pertenecen al conjunto de imágenes de referencia. De modo similar, el decodificador de video 30 puede decodificar el flujo de bits para determinar qué imágenes de referencia pertenecen al conjunto de imágenes de referencia.
- 20

El codificador de video (video coder) puede construir una pluralidad de subconjuntos de imágenes de referencia cada uno de los cuales identifica cero o más de las imágenes de referencia. Por ejemplo, el codificador de video (video coder) puede construir seis subconjuntos de imágenes de referencia (es decir, los subconjuntos de imágenes de referencia RefPicSetStCurr0, RefPicSetStCurr1, RefPicSetStFoll0, RefPicSetStFoll1, RefPicSetLtCurr y RefPicSetLtFoll), en los que cada uno de los subconjuntos identifica cero o más imágenes de referencia. En algunos ejemplos, el codificador de video (video coder) puede codificar la imagen actual en base a la pluralidad de subconjuntos de imágenes de referencia.

25

30 Por ejemplo, el codificador de video (video coder) puede derivar el conjunto de imágenes de referencia a partir de la pluralidad de subconjuntos de imágenes de referencia construida. Por ejemplo, el codificador de video (video coder) puede ordenar los subconjuntos de imágenes de referencia en cualquier orden para derivar el conjunto de imágenes de referencia, o puede derivar el conjunto de imágenes de referencia como parte de la construcción de los subconjuntos de imágenes de referencia. En algunos ejemplos, el codificador de video (video coder) puede codificar la imagen actual a partir del conjunto de imágenes de referencia derivado. A causa de que el conjunto de imágenes de referencia se deriva a partir de la pluralidad de subconjuntos de imágenes de referencia, se puede considerar que el codificador de video (video coder) convierte según codificación la imagen actual en base a la pluralidad de subconjuntos de imágenes de referencia.

35

40 En algunos ejemplos, para ordenar los subconjuntos de imágenes de referencia, el codificador de video (video coder) puede construir los subconjuntos de imágenes de referencia en el orden en el que se deben listar los subconjuntos de imágenes de referencia en el conjunto de imágenes de referencia. Por ejemplo, el codificador de video (video coder) puede construir primero el subconjunto de imágenes de referencia RefPicSetLtCurr, luego construir el subconjunto de imágenes de referencia RefPicSetLtFoll, luego construir el subconjunto de imágenes de referencia RefPicSetStCurr0, luego construir el subconjunto de imágenes de referencia RefPicSetStCurr1, luego construir el subconjunto de imágenes de referencia RefPicSetStFoll0 y luego construir el subconjunto de imágenes de referencia RefPicSetStFoll1. En este ejemplo ilustrativo, el orden de los subconjuntos de imágenes de referencia en el conjunto de imágenes de referencia puede ser RefPicSetLtCurr, RefPicSetLtFoll, RefPicSetStCurr0, RefPicSetStCurr1, RefPicSetStFoll0 y RefPicSetStFoll1, en ese orden, aunque otros órdenes son posibles.

45

50

De acuerdo con las técnicas de ejemplo que se describen en esta divulgación, después de derivar el conjunto de imágenes de referencia, el decodificador de video 30 puede decodificar cortes o cortes (slices) dentro de la imagen actual. Parte del proceso de decodificación implica la construcción de una o dos listas de imágenes de referencia. Una lista de imágenes de referencia es una lista de imágenes de referencia que se utiliza para pronosticar un corte P o B. Para el proceso de decodificación de un corte P, hay una lista de imágenes de referencia (Lista 0). Para el proceso de decodificación de un corte B, existen dos listas de imágenes de referencia (Lista 0 y Lista 1). La Lista 0, a veces denominada Lista 0 de imágenes de referencia o RefPicList0, es una lista de imágenes de referencia utilizada para la inter predicción de un corte P o B. Toda inter predicción utilizada para cortes P utiliza la Lista 0. La Lista 0 de imágenes de referencia es una de las dos listas de imágenes de referencia utilizadas para la bi-predicción de un corte B, siendo la otra lista la Lista 1 de imágenes de referencia. La Lista 1, a veces denominada Lista 1 de imágenes de referencia o RefPicList1, es una lista de imágenes de referencia utilizada para la predicción de un corte B. La Lista 1 de imágenes de referencia es una de las dos listas de imágenes de referencia utilizadas para la predicción de un corte B, siendo la

55

60

otra lista la Lista 0 de imágenes de referencia. Algunos bloques de un corte B pueden ser bi-pronosticados, utilizando tanto la Lista 0 como la Lista 1, y algunos bloques de un corte B pueden ser uni-pronosticados, utilizando la Lista 0 o la Lista 1.

- 5 Para construir las listas de imágenes de referencia, el decodificador de video 30 puede implementar una técnica de construcción por defecto para construir una Lista 0 inicial y, para las cortes B, una Lista 1 inicial. La construcción de la Lista 0 inicial y la Lista 1 inicial se puede denominar proceso de inicialización. En algunos ejemplos, el flujo de bits codificado puede indicar que el decodificador de video 30 debe modificar la Lista 0 inicial y/o la Lista 1 inicial para generar la Lista 0 final y la Lista 1 final. La modificación de la Lista 0 inicial y/o la Lista 1 inicial se puede denominar proceso de modificación. El proceso de modificación puede no ser necesario en todos los ejemplos, y la manera en la que el decodificador de video 30 puede implementar el proceso de modificación se describe con más detalle a continuación. De acuerdo con las técnicas que se describen en esta divulgación, cuando la modificación de la Lista 0 inicial o de la Lista 1 inicial no es necesaria, la Lista 0 final o la Lista 1 final (es decir, la Lista 0 o 1 de imágenes de referencia que se utiliza para decodificar el corte de la imagen actual) puede ser igual a la Lista 0 inicial o a la Lista 1 inicial. De esta manera, puede que no sea necesario reordenar la lista de imágenes de referencia.

En las técnicas que se describen en esta divulgación, el decodificador de video 30 puede construir la Lista 0 inicial o la Lista 1 inicial de tal manera que el decodificador de video 30 puede no tener que realizar el reordenamiento de las imágenes de referencia a incluir en la Lista 0 inicial o la Lista 1 inicial, con independencia de si se necesita el proceso de modificación, porque las imágenes de referencia en cada uno de los subconjuntos de imágenes de referencia ya están en un orden adecuado. Por ejemplo, en algunas otras técnicas, con independencia de si se necesita el proceso de modificación, es necesario reordenar las imágenes de referencia a incluir en la Lista 0 inicial o la Lista 1 inicial según sus valores de POC cuando se añaden las imágenes de referencia a la Lista 0 inicial o la Lista 1 inicial.

- 25 En el proceso de inicialización, el decodificador de video 30 puede implementar una técnica de construcción por defecto para construir la Lista 0 inicial y la Lista 1 inicial. La técnica de construcción por defecto puede significar que el decodificador de video 30 construye las listas iniciales de imágenes de referencia sin recibir elementos de sintaxis procedentes del codificador de video (video encoder) 20 con respecto a la manera en la que el decodificador de video 30 debe construir las listas iniciales de imágenes de referencia, o qué imágenes de referencia deben identificarse en las listas iniciales de imágenes de referencia.

El decodificador de video 30 puede invocar el proceso de construcción de la lista de imágenes de referencia cuando decodifica una cabecera de corte (slice) P o B. Por ejemplo, cuando se decodifica un corte P, el decodificador de video 30 puede invocar el proceso de construcción de la Lista 0 inicial, pero no puede invocar el proceso de construcción de la Lista 1 inicial porque un bloque en un corte P sólo es uni-pronosticado con respecto a una imagen de referencia identificada en la Lista 0. Cuando se decodifica un corte (slice) B, el decodificador de video 30 puede invocar el proceso de construcción de la Lista 0 inicial y de construcción de la Lista 1 inicial porque un bloque en un corte B puede ser bi-pronosticado con respecto a imágenes de referencia identificadas en cada una de las dos listas, Lista 0 y Lista 1.

- 40 De acuerdo con las técnicas de ejemplo que se describen en esta divulgación, el decodificador de video 30 puede utilizar subconjuntos de imágenes de referencia para construir la Lista 0 inicial y la Lista 1 inicial. Por ejemplo, la Lista 0 inicial y la Lista 1 inicial pueden listar o enumerar cero o más imágenes de referencia identificadas en RefPicSetStCurr0, RefPicSetStCurr1 o RefPicSetLtCurr. En este ejemplo, cuando se invoca el proceso de construcción de la lista de imágenes de referencia, puede haber al menos una imagen de referencia en RefPicSetStCurr0, RefPicSetStCurr1 y RefPicSetLtCurr. Aunque la Lista 0 inicial y la Lista 1 inicial pueden identificar una o más imágenes de referencia de los mismos subconjuntos de imágenes de referencia, el orden en el que el decodificador de video 30 añade las imágenes de referencia a la Lista 0 inicial puede ser diferente del orden en el que el decodificador de video 30 añade las imágenes de referencia a la Lista 1 inicial.

- 50 En esta divulgación, cuando el decodificador de video 30 añade (por ejemplo, listas) imágenes de referencia procedentes de uno o más de los subconjuntos de imágenes de referencia de la Lista 0 inicial o de la Lista 1 inicial, esta divulgación se refiere a que el decodificador de video 30 identifica las imágenes de referencia de la Lista 0 inicial o de la Lista 1 inicial. Por ejemplo, la pluralidad de subconjuntos de imágenes de referencia puede identificar cero o más imágenes de referencia. Para construir la Lista 0 inicial y la Lista 1 inicial, el decodificador de video 30 puede identificar en la Lista 0 inicial o en la Lista 1 inicial una o más de las imágenes de referencia que están identificadas en los subconjuntos de imágenes de referencia.

Para evitar confusiones y para mayor claridad, esta divulgación se puede referir a que el decodificador de video 30 lista o añade en la Lista 0 inicial y la Lista 1 inicial cero o más de las imágenes de referencia que se identifican en los subconjuntos de imágenes de referencia para construir la Lista 0 inicial y la Lista 1 inicial. De esta manera, añadir o listar, por parte del decodificador de video 30, imágenes de referencia significa que el decodificador de video 30 añade o lista un identificador de la imagen de referencia identificada en el subconjunto de imágenes de referencia. Por consiguiente, la Lista 0 inicial y la Lista 1 inicial resultantes incluyen una pluralidad de identificadores para las imágenes

de referencia que se pueden utilizar potencialmente para codificar un bloque o corte de una imagen actual. Estas imágenes de referencia se almacenan en respectivos búferes de imágenes decodificadas del decodificador de video 30 y del codificador de video (video encoder) 20.

5 Por ejemplo, para construir la Lista 0 inicial, el decodificador de video 30 puede listar primero (por ejemplo, añadir) las imágenes de referencia identificadas en RefPicSetStCurr0 en la Lista 0 inicial, seguido de las imágenes de referencia identificadas en RefPicSetStCurr1 en la Lista 0 inicial y, a continuación, las imágenes de referencia identificadas en RefPicSetLtCurr en la Lista 0 inicial. Para construir la Lista 1 inicial, el decodificador de video 30 puede listar primero (por ejemplo, añadir) las imágenes de referencia identificadas en RefPicSetStCurr1 en la Lista 1 inicial, seguido de las  
10 imágenes de referencia identificadas en RefPicSetStCurr0 en la Lista 1 inicial, y luego las imágenes de referencia identificadas en RefPicSetLtCurr en la Lista 1 inicial.

Además, adicionalmente a añadir en un orden diferente las imágenes de referencia en los subconjuntos de imágenes de referencia, el decodificador de video 30 puede utilizar un número diferente de imágenes de referencia de cada uno de los subconjuntos de imágenes de referencia cuando construye la Lista 0 y la Lista 1. Por ejemplo, la Lista 0 y la  
15 Lista 1 no necesitan incluir todas las imágenes de referencia del subconjunto RefPicSetStCurr0, RefPicSetStCurr1 y RefPicSetLtCurr. Más bien, el número de imágenes de referencia que se listan o enumeran procedentes de estos subconjuntos de imágenes de referencia de ejemplo para construir la Lista 0 inicial y la Lista 1 inicial se puede basar en los elementos de sintaxis que indican el número máximo de imágenes de referencia dentro de cada una de las  
20 listas, la Lista 0 inicial y la Lista 1 inicial.

Por ejemplo, para la Lista 0 inicial, el codificador de video (video encoder) 20 puede señalar el elemento de sintaxis num\_ref\_idx\_l0\_active\_minus1 para cortes o cortes P y B en la cabecera de corte (slice), y el elemento de sintaxis num\_ref\_idx\_l1\_active\_minus1 para cortes o cortes B que son bi-pronosticadas. Según se ha descrito anteriormente,  
25 el num\_ref\_idx\_l0\_active\_minus1 puede definir el número máximo de imágenes de referencia que pueden estar en la Lista 0 inicial, y el num\_ref\_idx\_l1\_active\_minus1 puede definir el número máximo de imágenes de referencia que pueden estar en la Lista 1 inicial. En algunos ejemplos, puede ser posible que el valor de num\_ref\_idx\_l0\_active\_minus1 sea diferente del valor de num\_ref\_idx\_l1\_active\_minus1, aunque no sea necesario en todos los ejemplos. En algunos ejemplos, el valor de num\_ref\_idx\_l0\_active\_minus1 puede ser el mismo que el  
30 valor de num\_ref\_idx\_l1\_active\_minus1.

Según se ha descrito anteriormente, el decodificador de video 30 puede recibir en el flujo de bits codificado los valores de num\_short\_term\_curr0 y num\_short\_term\_curr1. El decodificador de video 30 puede definir la variable NumShortTermCurr0 igual a num\_short\_term\_curr0 y definir la variable NumShortTermCurr1 igual a  
35 num\_short\_term\_curr1. NumShortTermCurr0 puede indicar el número de imágenes de referencia en el subconjunto de imágenes de referencia RefPicSetStCurr0, y NumShortTermCurr1 puede indicar el número de imágenes de referencia en el subconjunto de imágenes de referencia RefPicSetStCurr1.

El decodificador de video 30 también puede recibir en el flujo de bits codificado los valores de num\_long\_term\_pps\_curr  
40 y num\_long\_term\_add\_curr. El decodificador de video 30 puede definir la variable NumLongTermCurr igual a num\_long\_term\_pps\_curr más num\_long\_term\_add\_curr. NumLongTermCurr puede indicar el número de imágenes de referencia en RefPicSetLtCurr.

Para construir la Lista 0 inicial, el decodificador de video 30 primero puede añadir imágenes de referencia en  
45 RefPicSetStCurr0 a la Lista 0 inicial hasta que el decodificador de video 30 haya añadido todas las imágenes de referencia en RefPicSetStCurr0 a la Lista 0 inicial, y mientras el número de entradas en la lista 0 inicial (por ejemplo, el número de imágenes de referencia identificadas en la Lista 0) sea inferior o igual a num\_ref\_idx\_l0\_active\_minus1. Por ejemplo, NumShortTermCurr0 puede indicar el número de imágenes de referencia en el subconjunto de imágenes de referencia RefPicSetStCurr0. En este ejemplo, el decodificador de video 30 puede listar (por ejemplo, añadir)  
50 imágenes de referencia del subconjunto de imágenes de referencia RefPicSetStCurr0 hasta que el número de imágenes de referencia listadas procedentes de RefPicSetStCurr0 sea igual a NumShortTermCurr0. Sin embargo, mientras se listan las imágenes de referencia del subconjunto RefPicSetStCurr0 en la Lista 0 inicial, si el número total de entradas en la Lista 0 inicial se iguala a num\_ref\_idx\_l0\_active\_minus1, el decodificador de video 30 puede dejar de añadir imágenes de referencia al subconjunto de imágenes de referencia RefPicSetStCurr0 incluso si hay imágenes  
55 adicionales en RefPicSetStCurr0 que no se han listado en la Lista 0 inicial. En este caso, el decodificador de video 30 puede haber completado la construcción de la Lista 0 inicial.

Después de que el decodificador de video 30 haya listado todas las imágenes de referencia en el subconjunto de imágenes de referencia RefPicSetStCurr0 y el número total de entradas en la Lista 0 inicial sea inferior a  
60 num\_ref\_idx\_l0\_active\_minus1, el decodificador de video 30 puede entonces añadir imágenes de referencia en RefPicSetStCurr1 hasta que el decodificador de video 30 haya identificado todas las imágenes de referencia del subconjunto RefPicSetStCurr1, y mientras el número de entradas en la Lista 0 inicial (por ejemplo el número de imágenes de referencia identificadas en la Lista 0) sea inferior o igual a num\_ref\_idx\_l0\_active\_minus1. Por ejemplo,



de forma similar a lo anterior, NumShortTermCurr1 puede indicar el número de imágenes de referencia en el subconjunto de imágenes de referencia RefPicSetStCurr1. En este ejemplo, el decodificador de video 30 puede listar imágenes de referencia procedentes del subconjunto de imágenes de referencia RefPicSetStCurr1 hasta que el número de imágenes de referencia listadas procedentes de RefPicSetStCurr1 sea igual a NumShortTermCurr1. Sin embargo, mientras se listan las imágenes de referencia del subconjunto RefPicSetStCurr1, si el número total de entradas en la Lista 0 inicial se iguala a num\_ref\_idx\_l0\_active\_minus1, entonces el decodificador de video 30 puede dejar de añadir imágenes de referencia procedentes del subconjunto de imágenes de referencia RefPicSetStCurr1 incluso si hay imágenes adicionales en RefPicSetStCurr1 que no se han listado en la Lista 0 inicial. En este caso, el decodificador de video 30 puede haber completado la construcción de la Lista 0 inicial.

Después de que el decodificador de video 30 ha listado todas las imágenes de referencia en el subconjunto de imágenes de referencia RefPicSetStCurr1 y el número total de entradas en la Lista 0 inicial es inferior a num\_ref\_idx\_l0\_active\_minus1, el decodificador de video 30 puede entonces listar imágenes de referencia en RefPicSetLtCurr hasta que el decodificador de video 30 haya listado todas las imágenes de referencia en RefPicSetLtCurr, y mientras el número de entradas en la lista 0 inicial (por ejemplo, el número de imágenes de referencia identificadas en la Lista 0) sea inferior o igual a num\_ref\_idx\_l0\_active\_minus1. Por ejemplo, de forma similar a lo anterior, NumLongTermCurr puede indicar el número de imágenes de referencia en el subconjunto de imágenes de referencia RefPicSetLtCurr. En este ejemplo, el decodificador de video 30 puede listar imágenes de referencia del subconjunto de imágenes de referencia RefPicSetLtCurr hasta que el número de imágenes de referencia listadas de RefPicSetLtCurr sea igual a NumLongTermCurr. Sin embargo, mientras se listan las imágenes de referencia del subconjunto RefPicSetLtCurr en la Lista 0 inicial, si el número total de entradas en la Lista 0 inicial se iguala a num\_ref\_idx\_l0\_active\_minus1, entonces el decodificador de video 30 puede dejar de añadir imágenes de referencia al subconjunto de imágenes de referencia RefPicSetLtCurr incluso si hay imágenes adicionales en RefPicSetLtCurr que no se han listado en la Lista 0 inicial. En este caso, es posible que el decodificador de video 30 pueda haber completado la construcción de la Lista 0 inicial.

El siguiente pseudo código ilustra la manera en la que el decodificador de video 30 puede construir la Lista 0 inicial.

```

cIdx = 0
30 mientras( cIdx <= num_ref_idx_l0_active_minus1 )
{
    para( i=0; i < NumShortTermCurr0 && cIdx <= num_ref_idx_l0_active_minus1; cIdx++, i++)
        RefPicList0[ cIdx ] = RefPicSetStCurr0[i]
35  para( i=0; i < NumShortTermCurr1 && cIdx <= num_ref_idx_l0_active_minus1; cIdx++, i++)
        RefPicList0[ cIdx ] = RefPicSetStCurr1[i]

    para( i=0; i < NumLongTermCurr && cIdx <= num_ref_idx_l0_active_minus1; cIdx++, i++)
        RefPicList0[ cIdx ] = RefPicSetLtCurr[i]
40 }

```

En el pseudo código anterior, RefPicList0 puede ser la Lista 0 inicial. En ejemplos en los que no es necesario modificar la lista 0, la Lista 0 final puede ser igual a la Lista 0 inicial. Por lo tanto, en ejemplos en los que no es necesario modificar la lista 0, RefPicList0, en el pseudo código anterior, puede ser la Lista 0 final.

El decodificador de video 30 puede construir de forma similar la Lista 1. Sin embargo, para construir la Lista 1 inicial, el decodificador de video 30 puede primero añadir las imágenes de referencia del subconjunto de imágenes de referencia RefPicSetStCurr1 a la Lista 1 inicial, seguido del subconjunto de imágenes de referencia RefPicSetStCurr0 a la Lista 1 inicial, y seguido del subconjunto de imágenes de referencia RefPicSetLtCurr0 a la Lista 1 inicial. Además, de forma similar a lo anterior, si, mientras se listan imágenes de referencia de cualquiera de los subconjuntos de imágenes de referencia RefPicSetStCurr1, RefPicSetStCurr0 y RefPicSetLtCurr, el número total de entradas en la Lista 1 inicial se iguala a num\_ref\_idx\_l1\_active\_minus1, el decodificador de video 30 puede dejar de añadir imágenes de referencia, incluso si hay imágenes de referencia adicionales en estos subconjuntos de imágenes de referencia.

Por ejemplo, para construir la Lista 1 inicial, el decodificador de video 30 puede primero listar las imágenes de referencia del subconjunto RefPicSetStCurr1 hasta que el decodificador de video 30 haya identificado todas las imágenes de referencia en RefPicSetStCurr1, y mientras el número de entradas en la lista 1 inicial (por ejemplo, el número de imágenes de referencia identificadas en la Lista 1) sea inferior o igual a num\_ref\_idx\_l1\_active\_minus1. Por ejemplo, el valor de NumShortTermCurr1 puede indicar cuándo el decodificador de video 30 ha completado la lista de todas las imágenes de referencia en el subconjunto de imágenes de referencia RefPicSetStCurr1. Sin embargo, mientras se listan imágenes de referencia del subconjunto RefPicSetStCurr1, si el número total de entradas en la Lista 1 inicial se iguala a num\_ref\_idx\_l1\_active\_minus1, entonces el decodificador de video 30 puede dejar de añadir imágenes de referencia al subconjunto de imágenes de referencia RefPicSetStCurr1 incluso si hay imágenes

adicionales en RefPicSetStCurr1 que no se han listado en la Lista 1 inicial. En este caso, el decodificador de video 30 puede haber completado la construcción de la Lista 1 inicial.

Después de que el decodificador de video 30 haya listado todas las imágenes de referencia en el subconjunto de imágenes de referencia RefPicSetStCurr1 y el número total de entradas en la Lista 1 inicial sea inferior a num\_ref\_idx\_l1\_active\_minus1, el decodificador de video 30 puede entonces listar imágenes de referencia del subconjunto RefPicSetStCurr0 hasta que el decodificador de video 30 haya listado todas las imágenes de referencia en RefPicSetStCurr0, y mientras el número de entradas en la lista 1 inicial (por ejemplo, el número de imágenes de referencia identificadas en la Lista 1) es inferior o igual a num\_ref\_idx\_l1\_active\_minus1. Por ejemplo, de forma similar a lo anterior, el valor de NumShortTermCurr0 puede indicar cuándo el decodificador de video 30 ha completado la lista de todas las imágenes de referencia en el subconjunto de imágenes de referencia RefPicSetStCurr0. Sin embargo, mientras se listan imágenes de referencia del subconjunto RefPicSetStCurr0 en la Lista 1 inicial, si el número total de entradas en la Lista 1 inicial se iguala a num\_ref\_idx\_l1\_active\_minus1, el decodificador de video 30 puede dejar de añadir imágenes de referencia al subconjunto de imágenes de referencia RefPicSetStCurr0, incluso si hay imágenes adicionales en RefPicSetStCurr0 que no se han listado en la Lista 1 inicial. En este caso, el decodificador de video 30 puede haber completado la construcción de la Lista 1 inicial.

Después de que el decodificador de video 30 haya listado todas las imágenes de referencia en el subconjunto de imágenes de referencia RefPicSetStCurr0 y el número total de entradas en la Lista 1 inicial sea inferior a num\_ref\_idx\_l1\_active\_minus1, el decodificador de video 30 puede entonces listar imágenes de referencia en RefPicSetLtCurr hasta que el decodificador de video 30 haya listado todas las imágenes de referencia en RefPicSetLtCurr, y mientras el número de entradas en la lista 1 inicial (por ejemplo, el número de imágenes de referencia identificadas en la Lista 1) sea inferior o igual a num\_ref\_idx\_l1\_active\_minus1. Por ejemplo, de forma similar a lo anterior, el valor de NumLongTermCurr puede indicar cuándo el decodificador de video 30 ha completado la lista de todas las imágenes de referencia en el subconjunto de imágenes de referencia RefPicSetLtCurr. Sin embargo, mientras se listan imágenes de referencia del subconjunto RefPicSetLtCurr, si el número total de entradas en la Lista 1 inicial se iguala a num\_ref\_idx\_l1\_active\_minus1, entonces el decodificador de video 30 puede dejar de añadir imágenes de referencia al subconjunto de imágenes de referencia RefPicSetLtCurr incluso si hay imágenes adicionales en RefPicSetLtCurr que no se han listado en la Lista 1 inicial. En este caso, el decodificador de video 30 puede haber completado la construcción de la Lista 1 inicial.

El siguiente pseudo código ilustra la manera en la que el decodificador de video 30 puede construir la Lista 1 inicial.

```

cIdx = 0
35 mientras( cIdx <= num_ref_idx_l1_active_minus1 )
    {
        para( i=0; i < NumShortTermCurr1 && cIdx <= num_ref_idx_l1_active_minus1; cIdx++, i++ )
            RefPicList1[ cIdx ] = RefPicSetStCurr1 [i]
40     para( i=0; i < NumShortTermCurr0 && cIdx <= num_ref_idx_l1_active_minus1; cIdx++, i++ )
            RefPicList1[ cIdx ] = RefPicSetStCurr0[i]

        para( i=0; i < NumLongTermCurr && cIdx <= num_ref_idx_l1_active_minus1; cIdx++, i++ )
            RefPicList1[ cIdx ] = RefPicSetLtCurr[i]
45     }

```

En el pseudo código anterior, RefPicList1 puede ser la Lista 1 inicial. En ejemplos en los que no es necesario modificar la Lista 1, la Lista 1 final puede ser igual a la Lista 1 inicial. Por lo tanto, en ejemplos en los que no es necesario modificar la Lista 1, RefPicList1, en el pseudo código anterior, puede ser la Lista 1 final.

Lo anterior es un ejemplo de cómo el decodificador de video 30 puede construir la Lista 0 final y la Lista 1 final, cuando no es necesario modificar la lista de imágenes de referencia. En otros ejemplos, el decodificador de video 30 puede añadir los subconjuntos de imágenes de referencia en un orden diferente a los que se han descrito anteriormente. En algunos otros ejemplos, el decodificador de video 30 puede añadir subconjuntos de imágenes de referencia distintos de los que se han descrito anteriormente.

Aunque los ejemplos anteriores han descrito las técnicas para la construcción de listas de imágenes de referencia realizadas por el decodificador de video 30, aspectos de esta divulgación no son tan limitativos, y el codificador de video (video encoder) 20 puede implementar técnicas similares para construir las listas de imágenes de referencia. Sin embargo, puede no ser necesario que el codificador de video (video encoder) 20 construya las listas de imágenes de referencia de la misma manera en que el decodificador de video 30 construye las listas de imágenes de referencia.

En consecuencia, un codificador de video (video coder) (por ejemplo, el codificador de video (video encoder) 20 o el decodificador de video 30) puede estar configurado para convertir según una codificación (por ejemplo, codificar o decodificar) información indicativa de imágenes de referencia que pertenecen a un conjunto de imágenes de referencia. Según se ha descrito anteriormente, el conjunto de imágenes de referencia identifica las imágenes de referencia que se pueden utilizar potencialmente para la inter predicción de una imagen actual y que se pueden utilizar potencialmente para la inter predicción de una o más imágenes posteriores a la imagen actual en un orden de decodificación.

El codificador de video (video coder) también puede estar configurado para construir una pluralidad de subconjuntos de imágenes de referencia cada uno de los cuales identifica cero o más imágenes de referencia. Por ejemplo, el codificador de video (video coder) puede construir al menos los subconjuntos de imágenes de referencia RefPicSetStCurr0, RefPicSetStCurr1 y RefPicSetLtCurr. El codificador de video (video coder) puede construir subconjuntos de imágenes de referencia adicionales, tales como los que se han descrito anteriormente.

El codificador de video (video coder) puede entonces añadir imágenes de referencia de un primer subconjunto de imágenes de referencia, seguido de imágenes de referencia de un segundo subconjunto de imágenes de referencia, y seguido de imágenes de referencia de un tercer subconjunto de imágenes de referencia a una lista inicial de imágenes de referencia, mientras el número de entradas en la lista inicial de imágenes de referencia no sea superior a un número máximo de entradas permisibles en la lista de imágenes de referencia. Por ejemplo, el codificador de video (video coder) puede listar imágenes de referencia del subconjunto de imágenes de referencia RefPicSetStCurr0, seguido del subconjunto de imágenes de referencia RefPicSetStCurr1, y seguido del subconjunto de imágenes de referencia RefPicSetLtCurr en la Lista 0 inicial mientras el número de entradas en la lista 0 inicial no sea superior a num\_ref\_idx\_l0\_active\_minus1. Una vez más, el valor de num\_ref\_idx\_l0\_active\_minus1 puede indicar el número máximo de entradas permisibles en la lista de imágenes de referencia para la Lista 0.

En algunos ejemplos, el codificador de video (video coder) puede añadir las imágenes de referencia del primer subconjunto de imágenes de referencia a la lista inicial de imágenes de referencia hasta que todas las imágenes de referencia del primer subconjunto de imágenes de referencia estén listadas o incluidas en la lista inicial de imágenes de referencia o hasta que el número de entradas en la lista inicial de imágenes de referencia sea igual al número máximo de entradas permisibles en la lista de imágenes de referencia. Cuando el número de entradas de la lista inicial de imágenes de referencia es inferior al número máximo de entradas permisibles en la lista de imágenes de referencia, y después de añadir las imágenes de referencia del primer subconjunto de imágenes de referencia, el codificador de video (video coder) puede añadir las imágenes de referencia del segundo subconjunto de imágenes de referencia a la lista inicial de imágenes de referencia hasta que todas las imágenes de referencia del segundo subconjunto de imágenes de referencia estén listadas en la lista inicial de imágenes de referencia o hasta que el número de entradas en la lista inicial de imágenes de referencia sea igual al número máximo de entradas permisibles en la lista de imágenes de referencia. Cuando el número de entradas de la lista inicial de imágenes de referencia es inferior al número máximo de entradas permisibles en la lista de imágenes de referencia, y después de añadir las imágenes de referencia del segundo subconjunto de imágenes de referencia, el codificador de video (video coder) puede añadir las imágenes de referencia del tercer subconjunto de imágenes de referencia a la lista inicial de imágenes de referencia hasta que todas las imágenes de referencia del tercer subconjunto de imágenes de referencia estén listadas en la lista inicial de imágenes de referencia o hasta que el número de entradas en la lista inicial de imágenes de referencia sea igual al número máximo de entradas permisibles en la lista de imágenes de referencia.

El codificador de video (video coder) puede construir de forma similar la Lista 1 inicial. Por ejemplo, el codificador de video (video coder) puede añadir imágenes de referencia del segundo subconjunto de imágenes de referencia, seguido de imágenes de referencia del primer subconjunto de imágenes de referencia, y seguido de imágenes de referencia del tercer subconjunto de imágenes de referencia a la Lista 1 inicial, mientras el número de entradas en la lista inicial de imágenes de referencia de la Lista 1 inicial no sea superior a num\_ref\_idx\_l1\_active\_minus1. El elemento de sintaxis num\_ref\_idx\_l1\_active\_minus1 puede definir el número máximo de entradas permisibles en la lista 1.

En algunos ejemplos, tales como cuando no se necesita ninguna modificación, la Lista 0 inicial y la Lista 1 inicial pueden ser iguales a la Lista 0 final y a la Lista 1 final. En otras palabras, el codificador de video (video coder) puede construir la Lista 0 final y la Lista 1 final sin modificar la Lista 0 inicial y la Lista 1 inicial cuando no es necesaria su modificación. En estos casos, después de construir la Lista 0 inicial y la Lista 1 inicial, el codificador de video (video coder) puede no necesitar la realización de etapas adicionales para construir la Lista 0 final y la Lista 1 final (es decir, las listas de imágenes de referencia que el codificador de video (video coder) utiliza para codificar un bloque de la imagen actual).

Según se indica en el pseudo código anterior, el decodificador de video 30 puede construir la Lista 0 inicial, mientras cldx es inferior o igual a num\_ref\_idx\_l0\_active\_minus1, y puede construir la Lista 1 inicial, mientras cldx es inferior o igual a num\_ref\_idx\_l1\_active\_minus1. Esto puede dar lugar a que el decodificador de video 30 construya la Lista 0 inicial y la Lista 1 inicial sin ninguna entrada no completada en las listas de imágenes de referencia. Por ejemplo, en

algunas otras técnicas de codificación de video, un decodificador de video para estas otras técnicas de video construiría la Lista 0 y la Lista 1 iniciales, utilizando técnicas distintas a las que se describen en esta divulgación. Para estas otras técnicas de codificación de video, si el número de entradas en la Lista 0 inicial y en la Lista 1 inicial fuera inferior al número máximo de entradas permisibles, el decodificador de video, para estas otras técnicas de codificación de video, rellenaría las entradas restantes de la Lista 0 y de la Lista 1 con "sin imagen de referencia" para las entradas no completadas. Las entradas no completadas referencian a entradas en la Lista 0 y en la Lista 1 después de la última entrada que identifica una imagen de referencia y hasta la última entrada posible.

Como ejemplo ilustrativo para ayudar a la comprensión, un decodificador de video para estas otras técnicas de codificación de video puede construir una Lista 0 con cinco entradas, en el que el número máximo de entradas permisibles es igual a diez entradas. En este ejemplo, el decodificador de video para estas otras técnicas de codificación de video rellenaría las entradas de la sexta hasta la décima con "sin imagen de referencia". En este ejemplo, las entradas no completadas serían la sexta entrada (por ejemplo, la entrada después de la última entrada que identifica una imagen de referencia) hasta la décima entrada (por ejemplo, la última entrada posible según define el número máximo de entradas permisibles).

De acuerdo con las técnicas de esta divulgación, el decodificador de video 30 puede construir la Lista 0 inicial y la Lista 1 inicial de tal manera que no haya entradas no completadas. Además, en ejemplos en los que no se requiere ninguna modificación de la lista de imágenes de referencia, la Lista 0 final y la Lista 1 final pueden ser iguales a la Lista 0 inicial y a la Lista 1 inicial. Por lo tanto, en ejemplos en los que no se requiere ninguna modificación de la lista de imágenes de referencia, el decodificador de video 30 puede construir la Lista 0 final y la Lista 1 final de forma que no haya entradas no completadas. Incluso cuando se requiera alguna modificación de la lista de imágenes de referencia, la modificación puede no dar lugar a alguna entrada no completada. Por lo tanto, incluso en ejemplos en los que se requiere alguna modificación de la lista de imágenes de referencia, el decodificador de video 30 puede construir la Lista 0 final y la Lista 1 final de forma que no haya entradas no completadas.

Por ejemplo, la Lista 0 y la Lista 1 se pueden considerar como listas con entradas, y cada entrada puede identificar una imagen de referencia (por ejemplo, a través de su valor de POC). En otras palabras, el decodificador de video 30 puede identificar una imagen de referencia a través de su valor de POC en cada entrada de la Lista 0 y la Lista 1. El número de entradas en la Lista 0 y en la Lista 1 puede estar definido por los elementos de sintaxis `num_ref_idx_l0_active_minus1` y `num_ref_idx_l1_active_minus1`, respectivamente.

Para asegurar que no haya entradas no completadas, el decodificador de video 30 puede listar repetidamente (por ejemplo, añadir o identificar) imágenes de referencia de los subconjuntos de imágenes de referencia en la Lista 0 inicial y en la Lista 1 inicial hasta que el decodificador de video 30 determine qué imagen de referencia se debe identificar en cada posible entrada de la Lista 0 inicial y de la Lista 1 inicial. Por ejemplo, según se ha descrito anteriormente, para construir la Lista 0 inicial, después de añadir imágenes de referencia de los subconjuntos de imágenes de referencia `RefPicSetStCurr0` y `RefPicSetStCurr1` a la Lista 0 inicial, el decodificador de video 30 puede añadir imágenes de referencia del subconjunto de imágenes de referencia `RefPicSetLtCurr` a la Lista 0 inicial.

En algunos ejemplos, es posible que el número total de entradas en la Lista 0 inicial sea inferior al número máximo de entradas permisibles en la Lista 0 después de que el decodificador de video 30 haya añadido imágenes de referencia del subconjunto de imágenes de referencia `RefPicSetLtCurr` a la Lista 0 inicial. Por ejemplo, en el pseudo código para construir la Lista 0 inicial, `cldx` puede indicar el número de entradas en la Lista 0. En algunos ejemplos, después de que el decodificador de video 30 haya identificado las imágenes de referencia del subconjunto `RefPicSetLtCurr` en la Lista 0 inicial, el valor de `cldx` puede ser inferior a `num_ref_idx_l0_active_minus1`, en el que `num_ref_idx_l0_active_minus1` especifica el número máximo de entradas permisibles en la Lista 0.

De acuerdo con las técnicas que se describen en esta divulgación, después de listar las imágenes de referencia de los tres subconjuntos de la pluralidad de subconjuntos de imágenes de referencia, si el número de entradas en la Lista 0 inicial es inferior al número máximo de entradas permisibles, el decodificador de video 30 puede añadir repetidamente imágenes de referencia de los tres subconjuntos de imágenes de referencia hasta que todas las entradas en la Lista 0 estén completas. Por ejemplo, después de que el decodificador de video 30 haya añadido imágenes de referencia del conjunto de imágenes de referencia `RefPicSetLtCurr` y el número de entradas en la Lista 0 inicial sea inferior al número máximo de entradas permisibles, el decodificador de video 30 puede volver a listar (por ejemplo, volver a añadir o volver a identificar) imágenes de referencia del subconjunto de imágenes de referencia `RefPicSetStCurr0`.

En aspectos descritos en esta divulgación, cuando el decodificador de video 30 lista las imágenes de referencia de los subconjuntos de imágenes de referencia `RefPicSetStCurr0`, `RefPicSetStCurr1` y `RefPicSetLtCurr`, se puede considerar que el decodificador de video 30 añade imágenes de referencia de esta pluralidad de subconjuntos de imágenes de referencia a un primer conjunto de entradas en la lista de imágenes de referencia (por ejemplo, la Lista 0). Por ejemplo, el primer conjunto de entradas pueden ser las entradas en la lista de imágenes de referencia en las que el decodificador

de video 30 ha identificado las imágenes de referencia de los subconjuntos de imágenes de referencia RefPicSetStCurr0, RefPicSetStCurr1 y RefPicSetLtCurr. Entonces, si el número de entradas en la lista de imágenes de referencia es inferior al número máximo de entradas permitidas, el decodificador de video 30 puede volver a listar (por ejemplo, volver a añadir o volver a identificar) una o más imágenes de referencia de al menos uno de entre la pluralidad de subconjuntos de imágenes de referencia RefPicSetStCurr0, RefPicSetStCurr1 y RefPicSetLtCurr en entradas de la lista de imágenes de referencia que son subsiguientes respecto al primer conjunto de entradas. Las entradas subsiguientes respecto al primer grupo de entradas pueden ser las entradas que siguen al primer grupo de entradas, en las que el decodificador de video 30 añade imágenes de referencia ya listadas de uno o más de los subconjuntos de imágenes de referencia RefPicSetStCurr0, RefPicSetStCurr1 y RefPicSetLtCurr, según se describe a continuación.

Si, mientras se vuelven a añadir imágenes de referencia del subconjunto de imágenes de referencia RefPicSetStCurr0, el número total de entradas en la Lista 0 inicial se iguala a num\_ref\_idx\_I0\_active\_minus1, el decodificador de video 30 puede dejar de volver a añadir imágenes de referencia a la Lista 0 inicial. En este caso, el decodificador de video 30 puede haber completado la construcción de la Lista 0 inicial, y puede que no haya entradas no completadas. De lo contrario, el decodificador de video 30 puede volver a añadir imágenes de referencia del subconjunto de imágenes de referencia RefPicSetStCurr0 hasta que se hayan listado de nuevo todas las imágenes de referencia del subconjunto de imágenes de referencia RefPicSetStCurr0.

Si después de volver a añadir todas las imágenes de referencia del subconjunto de imágenes de referencia RefPicSetStCurr0, el número de entradas en la Lista 0 inicial es inferior a num\_ref\_idx\_I0\_active\_minus1, el decodificador de video 30 puede entonces volver a añadir imágenes de referencia del subconjunto de imágenes de referencia RefPicSetStCurr1, de forma similar a como el decodificador de video 30 ha listado de nuevo las imágenes de referencia del subconjunto de imágenes de referencia RefPicSetStCurr0. Si después de volver a añadir todas las imágenes de referencia del subconjunto de imágenes de referencia RefPicSetStCurr1, el número de entradas en la Lista 0 inicial es inferior a num\_ref\_idx\_I0\_active\_minus1, el decodificador de video 30 puede volver a añadir imágenes de referencia del subconjunto de imágenes de referencia RefPicSetLtCurr, de forma similar a como el decodificador de video 30 ha listado de nuevo las imágenes de referencia de los subconjuntos de imágenes de referencia RefPicSetStCurr0 y RefPicSetStCurr1. El decodificador de video 30 añade repetidamente imágenes de referencia de los subconjuntos de imágenes de referencia hasta que el número de entradas en la Lista 0 inicial sea igual al número máximo de entradas permitidas para la Lista 0 (es decir, igual a num\_ref\_idx\_I0\_active\_minus1).

Por ejemplo, supóngase que hay una imagen de referencia en RefPicSetStCurr0, una imagen de referencia en RefPicSetStCurr1 y una imagen de referencia en RefPicSetLtCurr. También, supóngase que num\_ref\_idx\_I0\_active\_minus1 es igual a cinco. En este ejemplo, el decodificador de video 30 puede identificar la imagen de referencia del subconjunto RefPicSetStCurr0 en dos entradas de la Lista 0 inicial. Por ejemplo, el decodificador de video 30 puede identificar la imagen de referencia del subconjunto RefPicSetStCurr0 en la primera entrada de la Lista 0 inicial, y volver a identificar la imagen de referencia del subconjunto RefPicSetStCurr0 en la cuarta entrada de la Lista 0 inicial. En este ejemplo, el valor del índice para la imagen de referencia en RefPicSetStCurr0 puede ser index[0] para la primera entrada de la Lista 0 inicial, e index[3] para la cuarta entrada de la Lista 0 inicial. En consecuencia, en algunos ejemplos, se puede listar (por ejemplo, identificar) una imagen de referencia de uno de los subconjuntos de imágenes de referencia más de una vez en las listas iniciales de imágenes de referencia.

El decodificador de video 30 puede construir de forma similar la Lista 1 inicial de tal manera que no haya entradas no completadas en la Lista 1 inicial. Por ejemplo, el decodificador de video 30 puede añadir repetidamente imágenes de referencia de los subconjuntos de imágenes de referencia RefPicSetStCurr1, RefPicSetStCurr0 y RefPicSetLtCurr, en ese orden, hasta que el número de entradas en la Lista 1 inicial sea igual al número máximo de entradas permitidas en la Lista 1 (es decir, igual a num\_ref\_idx\_I1\_active\_minus1).

De esta manera, a causa de que los bucles "for" están anidados dentro del bucle "while", en el pseudo código anterior para construir la Lista 0 inicial y la Lista 1 inicial, el decodificador de video 30 puede construir la Lista 0 inicial y la Lista 1 inicial de tal manera que no haya entradas no completadas en la Lista 0 inicial y en la Lista 1 inicial (es decir, que no haya entradas no completadas después del proceso de inicialización). En algunos ejemplos, cada una de las entradas en la Lista 0 inicial y en la Lista 1 inicial puede identificar una imagen de referencia de uno de los subconjuntos de imágenes de referencia. En algunos ejemplos, puede ser posible que una o más imágenes de referencia de uno de los subconjuntos de imágenes de referencia sea identificada más de una vez en las listas finales de imágenes de referencia, pero en diferentes entradas con diferentes valores de índice.

Aunque los ejemplos anteriores han descrito las técnicas para la construcción de listas de imágenes de referencia sin entradas no completadas como realizadas por el decodificador de video 30, aspectos de esta divulgación no son tan limitados, y el codificador de video (video encoder) 20 puede implementar técnicas similares para construir las listas de imágenes de referencia sin entradas no completadas. Sin embargo, puede no ser necesario que el codificador de

video (video encoder) 20 construya las listas de imágenes de referencia de la misma manera en la que el decodificador de video 30 construye las listas de imágenes de referencia.

En consecuencia, un codificador de video (video coder) (por ejemplo, el codificador de video (video encoder) 20 o el decodificador de video 30) puede estar configurado para convertir según una codificación (por ejemplo, codificar o decodificar) información indicativa de imágenes de referencia que pertenecen a un conjunto de imágenes de referencia. Según se ha descrito anteriormente, el conjunto de imágenes de referencia identifica las imágenes de referencia que se pueden utilizar potencialmente para la inter predicción de una imagen actual y que se pueden utilizar potencialmente para la inter predicción de una o más imágenes posteriores a la imagen actual en un orden de decodificación.

El codificador de video (video coder) puede construir una pluralidad de subconjuntos de imágenes de referencia cada uno de los cuales identifica cero o más de las imágenes de referencia (por ejemplo, los subconjuntos de imágenes de referencia RefPicSetStCurr0, RefPicSetStCurr1 y RefPicSetLtCurr). El codificador de video (video coder) puede listar (por ejemplo, identificar o añadir) imágenes de referencia de la pluralidad de subconjuntos de imágenes de referencia en un primer conjunto de entradas de una lista de imágenes de referencia. El codificador de video (video coder) puede determinar si un número de entradas en la lista de imágenes de referencia es igual a un número máximo de entradas permisibles en la lista de imágenes de referencia.

Cuando el número de entradas en la lista de imágenes de referencia no es igual al número máximo de entradas permisibles en la lista de imágenes de referencia, el codificador de video (video coder) puede volver a añadir repetidamente (por ejemplo, volver a identificar) una o más imágenes de referencia de al menos una de las imágenes de referencia en entradas de la lista de imágenes de referencia que son subsiguientes respecto al primer grupo de entradas hasta que el número de entradas en la lista de imágenes de referencia sea igual al número máximo de entradas permisibles en la lista de imágenes de referencia. El codificador de video (video coder) puede entonces codificar la imagen actual en base a la lista de imágenes de referencia construida.

Según se ha descrito anteriormente, en algunos ejemplos, el codificador de video (video encoder) 20 puede señalar elementos de sintaxis que indican al decodificador de video 30 que modifique una lista o listas iniciales de imágenes de referencia. Por ejemplo, el decodificador de video 30 puede construir la lista o listas iniciales de imágenes de referencia de la manera que se ha descrito anteriormente. A continuación, en algunos casos, el decodificador de video 30 puede decodificar elementos de sintaxis del flujo de bits codificado que le indican al decodificador de video 30 que modifique la lista o listas iniciales de imágenes de referencia para construir la lista o listas finales de imágenes de referencia. En general, en la modificación, el decodificador de video 30 puede mapear una o más de las imágenes identificadas en uno o más de la pluralidad de subconjuntos de imágenes de referencia en una entrada de una de las listas de imágenes de referencia después de la inicialización de la lista de imágenes de referencia.

Por ejemplo, después de que el decodificador de video 30 haya construido la lista o listas iniciales de imágenes de referencia de la manera que se ha descrito anteriormente, el decodificador de video 30 puede modificar al menos una de las entradas de una de las listas iniciales de imágenes de referencia de la manera indicada por el flujo de bits codificado. Por ejemplo, el codificador de video (video encoder) 20 puede indicar como parte de los elementos de sintaxis de modificación qué imagen de uno de entre la pluralidad de subconjuntos de imágenes de referencia se debe identificar en una entrada de la lista de imágenes de referencia, aunque esa entrada de la lista de imágenes de referencia ya pueda identificar una imagen de referencia como parte del proceso de inicialización. En algunos ejemplos, las técnicas de modificación de listas de imágenes de referencia que se describen en esta divulgación pueden permitir una modificación de manera flexible. Por ejemplo, puede ser posible que el decodificador de video 30 identifique en una o ambas de las listas de imágenes de referencia una imagen de referencia que no está en las listas iniciales de imágenes de referencia.

Según se utiliza en esta divulgación, la frase "lista de imágenes de referencia modificada" se refiere a la lista de imágenes de referencia después de la modificación de la lista inicial de imágenes de referencia. La lista de imágenes de referencia modificada puede ser la lista final de imágenes de referencia. El número de entradas en la lista de imágenes de referencia modificada es  $\text{num\_ref\_idx\_I0\_active\_minus1} + 1$  para la Lista 0, y  $\text{num\_ref\_idx\_I1\_active\_minus1} + 1$  para la Lista 1. Una imagen de referencia puede aparecer en más de un índice (por ejemplo, entrada) de las listas de referencia modificadas para la Lista 0 y la Lista 1.

Para la modificación de una lista de imágenes de referencia, el codificador de video (video encoder) 20 puede señalar los elementos de sintaxis de la Tabla 6.

**Tabla 6. Sintaxis de modificación de lista de imágenes de referencia**

ref_pic_list_modification() { si( slice_type != 2 ) {	<b>Descriptor</b>
----------------------------------------------------------	-------------------

<b>ref_pic_list_modification_flag_I0</b>	u(1)
si( ref_pic_list_modification_flag_I0 )	
hacer {	
<b>modification_of_ref_pic_idc</b>	ue(v)
si( modification_of_ref_pic_idc != 3 )	
<b>ref_pic_set_idx</b>	
} mientras( modification_of_ref_pic_idc != 3 )	
}	
si( slice_type == 1 ) {	
<b>ref_pic_list_modification_flag_I1</b>	u(1)
si( ref_pic_list_modification_flag_I1 )	
hacer {	
<b>modification_of_ref_pic_idc</b>	ue(v)
si( modification_of_ref_pic_idc != 3 )	
<b>ref_pic_set_idx</b>	
} mientras( modification_of_ref_pic_idc != 3 )	
}	
}	

Los elementos sintácticos `modification_of_ref_pic_idc`, `short_term_ref_pic_set_idx` y `long_term_ref_pic_set_idx` pueden especificar el cambio de las listas iniciales de imágenes de referencia a las listas de imágenes de referencia a utilizar para decodificar el corte (slice).

**ref\_pic\_list\_modification\_flag\_I0** igual a 1 puede especificar que el elemento de sintaxis `modification_of_ref_pic_idc` se encuentra presente para especificar la Lista 0 de imágenes de referencia. `ref_pic_list_modification_flag_I0` igual a 0 especifica que este elemento de sintaxis no se encuentra presente.

10 Cuando `ref_pic_list_modification_flag_I0` es igual a 1, el número de veces que `modification_of_ref_pic_idc` no es igual a 3 después de `ref_pic_list_modification_flag_I0` no puede exceder `num_ref_idx_I0_active_minus1 + 1`.

15 **ref\_pic\_list\_modification\_flag\_I1** igual a 1 puede especificar que la sintaxis `modification_of_ref_pic_idc` se encuentra presente para especificar la Lista 1 de imágenes de referencia. `ref_pic_list_modification_flag_I1` igual a 0 puede especificar que este elemento de sintaxis no se encuentra presente.

20 Cuando `ref_pic_list_modification_flag_I1` es igual a 1, el número de veces que `modification_of_ref_pic_idc` no es igual a 3 después de `ref_pic_list_modification_flag_I1` no puede exceder `num_ref_idx_I1_active_minus1 + 1`.

**modification\_of\_ref\_pic\_idc** junto con `short_term_ref_pic_set_idx` o `long_term_ref_pic_set_idx` puede especificar cuáles de las imágenes de referencia son mapeadas de nuevo. Los valores de `modification_of_ref_pic_idc` se especifican en la Tabla 7. El valor del primer `modification_of_ref_pic_idc` que sigue inmediatamente después de `ref_pic_list_modification_flag_I0` o `ref_pic_list_modification_flag_I1` puede no ser igual a 3.

25 **ref\_pic\_set\_idx** especifica el índice, a `RefPicSetStCurr0`, `RefPicSetStCurr1` o `RefPicSetLtCurr`, de la imagen de referencia que se mueve al índice actual en la lista de imágenes de referencia. El valor de `ref_pic_set_idx` puede estar en el rango de 0 hasta `max_num_ref_frames`, inclusive.

30 **Tabla 7. Operaciones `modification_of_ref_pic_idc` para modificar listas de imágenes de referencia**

<b>modification_of_ref_pic_idc</b>	<b>Modificación especificada</b>
0	Para la lista 0: <code>ref_pic_set_idx</code> está presente y corresponde a un índice de <code>RefPicSetStCurr0</code> ; Para la lista 1: <code>ref_pic_set_idx</code> está presente y corresponde a un índice de <code>RefPicSetStCurr1</code>
1	Para la lista 0: <code>ref_pic_set_idx</code> está presente y corresponde a un índice de <code>RefPicSetStCurr1</code> ; Para la lista 1: <code>ref_pic_set_idx</code> está presente y corresponde a un índice de <code>RefPicSetStCurr0</code>
2	<code>ref_pic_set_idx</code> está presente y corresponde a un índice de <code>RefPicSetLtCurr</code>
3	<code>ref_pic_idx</code> no está presente y el bucle termina para modificar la lista inicial de imágenes de referencia

- Para la modificación de la lista de imágenes de referencia, cuando `ref_pic_list_modification_flag_I0` es igual a 1, el decodificador de video 30 puede modificar la lista 0 inicial de imágenes de referencia (es decir, la Lista 0 inicial), y cuando `ref_pic_list_modification_flag_I1` es igual a 1, el decodificador de video 30 puede modificar la lista 1 inicial de imágenes de referencia (es decir, la Lista 1 inicial). Para ayudar a comprender la modificación de listas de imágenes de referencia, supóngase que la variable `refIdxL0` es un índice de la Lista 0 inicial, y la variable `refIdxL1` es un índice de la Lista 1 inicial. En otras palabras, `refIdxL0` puede identificar la entrada de la Lista 0 inicial (es decir, un índice de la Lista 0 inicial identifica una entrada de la Lista 0 inicial), y `refIdxL1` puede identificar la entrada de la Lista 1 inicial. Las variables `refIdxL0` y `refIdxL1` se pueden establecer iguales a 0 inicialmente.
- El decodificador de video 30 puede procesar los elementos de sintaxis para `modification_of_ref_pic_idc` en el proceso que se producen los elementos de sintaxis en el flujo de bits. Por ejemplo, si el codificador de video (video encoder) 20 señala que es necesario modificar la lista de imágenes de referencia para la Lista 0 inicial, entonces el decodificador de video 30 puede procesar el orden en el que el codificador de video (video encoder) 20 ha señalado los elementos de sintaxis `modification_of_ref_pic_idc` para modificar la Lista 0 inicial. De modo similar, si el codificador de video (video encoder) 20 señala que la modificación de la lista de imágenes de referencia es necesaria para la Lista 1 inicial, el decodificador de video 30 puede procesar el orden en el que el codificador de video (video encoder) 20 señala los elementos de sintaxis `modification_of_ref_pic_idc` para modificar la Lista 1 inicial.
- El valor del elemento de sintaxis `modification_of_ref_pic_idc` puede ser igual a 0, 1, 2 o 3, según se indica en la Tabla 7. Si el valor del elemento de sintaxis `modification_of_ref_pic_idc` es igual a 3, el decodificador de video 30 puede detener la modificación de la lista inicial de imágenes de referencia. De lo contrario, el decodificador de video 30 puede seguir modificando la lista inicial de imágenes de referencia hasta que el valor del elemento de sintaxis `modification_of_ref_pic_idc` sea igual a 3. Por ejemplo, el codificador de video (video encoder) 20 puede señalar una pluralidad de valores para el elemento de sintaxis `modification_of_ref_pic_idc`, y el decodificador de video 30 puede procesar cada uno de los valores en el orden en el que están presentes en el flujo de bits codificado. Cuando el decodificador de video 30 procesa el valor del elemento de sintaxis `modification_of_ref_pic_idc` para igualarlo a 3, el decodificador de video 30 puede determinar que no es necesaria ninguna otra modificación.
- El valor del elemento de sintaxis `modification_of_ref_pic_idc` que no es igual a 3 (es decir, 0, 1 o 2) puede indicar de qué subconjunto de imágenes de referencia el decodificador de video 30 debe identificar una imagen de referencia que se debe listar en (por ejemplo, añadir a) una entrada actual de la lista de imágenes de referencia. Según se ha descrito anteriormente, la entrada actual de la lista de imágenes de referencia puede identificarse por el valor de `refIdxLX`, en el que `LX` es la Lista 0 o la Lista 1. Por ejemplo, si el decodificador de video 30 está modificando la Lista 0 inicial, y `modification_of_ref_pic_idc` es igual a 0, entonces de acuerdo con la Tabla 7, el decodificador de video 30 puede determinar qué imagen de referencia del subconjunto `RefPicSetStCurr0` se debe identificar en una entrada actual de la lista de imágenes de referencia en base al valor de `ref_pic_set_idx`. Si el decodificador de video 30 está modificando la Lista 1 inicial y `modification_of_ref_pic_idc` es igual a 0, entonces, de acuerdo con la Tabla 7, el decodificador de video 30 puede determinar qué imagen de referencia del subconjunto `RefPicSetStCurr1` se debe identificar en una entrada actual de la lista de imágenes de referencia en base al valor de `ref_pic_set_idx`. Por ejemplo, la variable `curRefPicSet` puede definir qué subconjunto de imágenes de referencia debe utilizar el decodificador de video 30 para modificar la Lista 0 inicial o la Lista 1 inicial.
- Por ejemplo, si `modification_of_ref_pic_idc` es igual a 0, y el decodificador de video 30 está modificando la Lista 0 inicial, entonces `curRefPicSet` es igual al subconjunto de imágenes de referencia `RefPicSetStCurr0`. Si `modification_of_ref_pic_idc` es igual a 0, y el decodificador de video 30 está modificando la Lista 1 inicial, entonces `curRefPicSet` es igual al subconjunto de imágenes de referencia `RefPicSetStCurr1`.
- Si el decodificador de video 30 está modificando la Lista 0 inicial, y `modification_of_ref_pic_idc` es igual a 1, entonces de acuerdo con la Tabla 7, el decodificador de video 30 puede determinar qué imagen de referencia del subconjunto `RefPicSetStCurr1` se debe identificar en una entrada actual de la lista de imágenes de referencia en base al valor de `ref_pic_set_idx`. Si el decodificador de video 30 está modificando la Lista 1 inicial, y `modification_of_ref_pic_idc` es igual a 1, entonces de acuerdo con la Tabla 7, el decodificador de video 30 puede determinar qué imagen de referencia del subconjunto `RefPicSetStCurr0` se debe identificar en una entrada actual de la lista de imágenes de referencia en base al valor del `ref_pic_set_idx`.
- En este caso, si `modification_of_ref_pic_idc` es igual a 1, y el decodificador de video 30 está modificando la Lista 0 inicial, entonces `curRefPicSet` es igual al subconjunto de imágenes de referencia `RefPicSetStCurr1`. Si `modification_of_ref_pic_idc` es igual a 1, y el decodificador de video 30 está modificando la Lista 1 inicial, entonces `curRefPicSet` es igual al subconjunto de imágenes de referencia `RefPicSetStCurr0`.
- Si el decodificador de video 30 está modificando la Lista 0 inicial o la Lista 1 inicial, y `modification_of_ref_pic_idc` es igual a 2, entonces de acuerdo con la Tabla 7, el decodificador de video 30 puede determinar qué imagen de referencia



del subconjunto RefPicSetLtCurr se debe identificar en una entrada actual de la lista de imágenes de referencia en base al valor de ref\_pic\_set\_idx. En este ejemplo, si modification\_of\_ref\_pic\_idc es igual a 2, y el decodificador de video 30 está modificando la Lista 0 inicial o la Lista 1 inicial, curRefPicSet es igual al subconjunto de imágenes de referencia RefPicSetLtCurr.

5 Según se ha descrito anteriormente, el elemento de sintaxis ref\_pic\_set\_idx puede indicar el índice de uno de los subconjuntos de imágenes de referencia. En otras palabras, el elemento de sintaxis ref\_pic\_set\_idx puede indicar al decodificador de video 30 la entrada de uno de los subconjuntos de imágenes de referencia. El decodificador de video 30 puede determinar la imagen de referencia identificada en la entrada de uno de los subconjuntos de la pluralidad de  
10 subconjuntos de imágenes de referencia como la imagen de referencia que se debe identificar en un índice actual de la Lista 0 inicial o de la Lista 1 inicial.

La variable curRefPicPoc puede ser igual a PicOrderCnt(curRefPicSet[ref\_pic\_set\_idx]). De este modo, el valor de curRefPicPoc puede ser el valor de POC de la imagen de referencia identificada en la entrada ref\_pic\_set\_idx de  
15 curRefPicSet. Según se ha descrito anteriormente, curRefPicSet puede ser igual a RefPicSetStCurr0, RefPicSetStCurr1 o RefPicSetLtCurr en base al valor del elemento de sintaxis modification\_of\_ref\_pic\_idc y en base a si el decodificador de video 30 está modificando la Lista 0 inicial o la Lista 1 inicial.

El decodificador de video 30 puede implementar el siguiente pseudo código para la modificación de la lista de imágenes de referencia. Por ejemplo, en el siguiente pseudo código, el decodificador de video 30 puede identificar la imagen con el valor de POC igual a curRefPicPoc en una entrada de la lista inicial de imágenes de referencia. La variable  
20 reldxLX indica la posición de índice para la entrada en la lista inicial de imágenes de referencia. Por ejemplo, cuando el decodificador de video 30 está modificando la Lista 0 inicial, reldxLX puede ser reldxL0, y cuando el decodificador de video 30 está modificando la Lista 1 inicial, reldxLX puede ser reldxL1.

25 Después de que el decodificador de video 30 haya identificado la imagen de referencia con un valor de POC igual a curRefPicPOC en la lista inicial de imágenes de referencia (es decir, la Lista 0 inicial o la Lista 1 inicial), el decodificador de video 30 puede cambiar la posición de las otras imágenes restantes a una posición posterior en la lista. Por ejemplo, el decodificador de video 30 puede mover las imágenes de referencia identificadas en la lista inicial de imágenes de referencia a entradas subsiguientes a la entrada actual a una entrada siguiente para construir la lista de imágenes de referencia modificada. Como ejemplo ilustrativo, supóngase que la entrada actual en la lista inicial de imágenes de referencia es la tercera entrada con index[2]. El decodificador de video 30 puede mover la imagen de referencia actualmente identificada en la tercera entrada con index[2] a una siguiente entrada (por ejemplo, la cuarta entrada con index[3]). El decodificador de video 30 puede mover la imagen de referencia actualmente identificada en la cuarta  
30 entrada con index[3] a la quinta entrada con index[4]. En algunos ejemplos, el decodificador de video 30 puede comenzar desde la última entrada de la lista inicial de imágenes de referencia y mover la imagen de referencia identificada en esa entrada a una nueva entrada temporal. A continuación, mover la imagen de referencia identificada en la penúltima entrada a la última entrada, y así sucesivamente, hasta que el decodificador de video 30 alcance la entrada actual.

40 El decodificador de video 30 puede entonces incrementar el valor de la variable reldxLX. En este pseudo código, la longitud de RefPicListX (es decir, RefPicList0 o RefPicList1) se establece temporalmente un elemento más larga que la longitud necesaria para la lista final de imágenes de referencia. Por ejemplo, según se ha descrito anteriormente, el decodificador de video 30 puede comenzar desde la última entrada de la lista inicial de imágenes de referencia, mover  
45 esa última entrada a la entrada temporal, mover la penúltima entrada a la última entrada, y así sucesivamente para modificar la lista inicial de imágenes de referencia. Tras la ejecución del pseudo código, el decodificador de video 30 sólo podrá conservar las entradas del índice 0 hasta num\_ref\_idx\_IX\_active\_minus1, en el que num\_ref\_idx\_IX\_active\_minus1 es num\_ref\_idx\_I0\_active\_minus1 para la Lista 0, y num\_ref\_idx\_I1\_active\_minus1 para la Lista 1.

50 para( cIdx = num\_ref\_idx\_IX\_active\_minus1 + 1; cIdx > reldxLX; cIdx-- )  
    RefPicListX[ cIdx ] = RefPicListX[ cIdx - 1 ]  
RefPicListX[ reldxLX++ ] = imagen de referencia con PicOrderCnt igual a currRefPicPoc  
nIdx = reldxLX

55 para( cIdx = reldxLX; cIdx <= num\_ref\_idx\_IX\_active\_minus1 + 1; cIdx++ )  
    si( PicOrderCnt( RefPicListX[ cIdx ] ) != currentRefPicPoc )  
        RefPicListX[ nIdx++ ] = RefPicListX[ cIdx ]

60 En el pseudo código anterior, RefPicListX se refiere a RefPicList0 (es decir, la Lista 0 final) o a RefPicList1 (es decir, la Lista 1 final) en base a si el decodificador de video 30 está modificando la Lista 0 inicial o la Lista 1 inicial.

num\_ref\_idx\_IX\_active\_minus1 se refiere a num\_ref\_idx\_I0\_active\_minus1 o ref\_idx\_11\_active\_minus1 en base a si el decodificador de video 30 está modificando la Lista 0 inicial o la Lista 1 inicial.

Las técnicas anteriores describen un ejemplo de cómo el decodificador de video 30 puede modificar una lista inicial de imágenes de referencia. Durante el proceso de codificación, el codificador de video (video encoder) 20 también puede necesitar decodificar la imagen codificada para fines de codificar imágenes subsiguientes. Por consiguiente, en algunos ejemplos, el codificador de video (video encoder) 20 también puede estar configurado para construir las listas iniciales de imágenes de referencia y modificar las listas iniciales de imágenes de referencia de la manera que se ha descrito anteriormente. Sin embargo, es posible que el codificador de video (video encoder) 20 no necesite modificar la lista o listas iniciales de imágenes de referencia en todos los ejemplos. En algunos ejemplos, el decodificador de video 30 puede ser el único convertidor de codificación de video que modifica la imagen de referencia inicial utilizando las técnicas que se han descrito anteriormente.

Por consiguiente, en algunos ejemplos, un codificador de video (video coder) (por ejemplo, el codificador de video (video encoder) 20 o el decodificador de video 30) puede construir una lista inicial de imágenes de referencia (por ejemplo, la Lista 0 inicial o la Lista 1 inicial) utilizando las técnicas que se han descrito anteriormente. El codificador de video (video coder) puede determinar si es necesario modificar la lista de imágenes de referencia en base a los elementos de sintaxis codificados en el flujo de bits codificado. Cuando sea necesario modificar la lista de imágenes de referencia, el codificador de video (video coder) puede modificar la lista inicial de imágenes de referencia.

Por ejemplo, cuando se necesita modificar la lista de imágenes de referencia, el codificador de video (video coder) puede identificar una imagen de referencia en al menos uno de entre la pluralidad de subconjuntos de imágenes de referencia construidos. El codificador de video (video coder) puede listar (por ejemplo, añadir) el subconjunto de imágenes de referencia identificado en una entrada actual de la lista inicial de imágenes de referencia para construir una lista de imágenes de referencia modificada. El codificador de video (video coder) puede convertir según una codificación (por ejemplo, codificar o decodificar) la imagen actual en base a la lista de imágenes de referencia modificada.

Como ejemplo, el codificador de video (video coder) puede construir los subconjuntos de imágenes de referencia RefPicSetStCurr0, RefPicSetStCurr1 y RefPicSetLtCurr. Para identificar una imagen de referencia en al menos uno de estos subconjuntos de imágenes de referencia, el codificador de video (video coder) puede determinar un índice en al menos uno de estos subconjuntos de imágenes de referencia. El codificador de video (video coder) puede entonces determinar la imagen de referencia identificada en una entrada de al menos uno de estos subconjuntos de imágenes de referencia en base al índice determinado.

Por ejemplo, el codificador de video (video coder) puede convertir según una codificación (por ejemplo, codificar o decodificar) un primer elemento de sintaxis, tal como el elemento de sintaxis modification\_of\_ref\_pic\_idc, con el que el codificador de video (video coder) identifica uno de los subconjuntos de imágenes de referencia (por ejemplo, uno de los subconjuntos de imágenes de referencia RefPicSetStCurr0, RefPicSetStCurr1, y RefPicSetLtCurr). El codificador de video (video coder) puede codificar un segundo elemento de sintaxis, tal como el elemento de sintaxis ref\_pic\_set\_idx, que indica el índice en el subconjunto de imágenes de referencia identificado (por ejemplo, uno de los subconjuntos de imágenes de referencia RefPicSetStCurr0, RefPicSetStCurr1 y RefPicSetLtCurr).

En algunos ejemplos, el codificador de video (video coder) puede estar configurado para mover las imágenes de referencia en la Lista inicial de imágenes de referencia. Por ejemplo, el codificador de video (video coder) puede mover imágenes de referencia identificadas en la lista inicial de imágenes de referencia en entradas posteriores a la entrada actual a una siguiente entrada de la lista de imágenes de referencia modificada.

Los ejemplos anteriores han descrito la manera en la que el codificador de video (video encoder) 20 y el decodificador de video 30 pueden derivar un conjunto de imágenes de referencia, así como técnicas de ejemplo para construir listas de imágenes de referencia cuando no se necesita ninguna modificación y cuando se necesita alguna modificación. Sin embargo, las técnicas que se describen en esta divulgación no son tan limitadas. En algunos ejemplos, las técnicas que se describen en esta divulgación pueden estar enfocadas a la gestión de un búfer de imágenes decodificadas (DPB). El búfer DPB puede ser un búfer que almacena imágenes decodificadas.

Cada uno del codificador de video (video encoder) 20 y decodificador de video 30 puede incluir sus respectivos búferes DPB. Por ejemplo, como parte del proceso de codificación, el codificador de video (video encoder) 20 puede decodificar una imagen actual, almacenar la imagen decodificada en el búfer DPB del codificador de video (video encoder) 20 y utilizar la imagen decodificada almacenada en el búfer DPB para la inter predicción de una imagen posterior. De modo similar, como parte del proceso de decodificación, el decodificador de video 30 puede decodificar una imagen actual y almacenar la imagen decodificada en el búfer DPB del decodificador de video 30. El decodificador de video 30 puede utilizar entonces la imagen decodificada para la inter predicción de una imagen posterior.

En algunos ejemplos, el búfer DPB del codificador de video (video encoder) 20 o del decodificador de video 30 puede almacenar imágenes decodificadas para reordenación de outputs o retardo de outputs. Por ejemplo, el decodificador de video 30 puede determinar que las imágenes decodificadas deben reordenarse para el output o que se debe retardar el output de una imagen decodificada. En estos ejemplos, el búfer DPB del decodificador de video 30 puede almacenar las imágenes decodificadas para reordenación de outputs o retardo de outputs.

Las técnicas de gestión del búfer DPB que se describen en esta divulgación pueden estar enfocadas a la manera en la que el búfer DPB produce y elimina las imágenes decodificadas. El elemento de sintaxis `output_flag` puede afectar al output de imágenes decodificadas y al proceso de eliminación, y se puede definir como parte de la semántica de la unidad de capa de abstracción de red (NAL: network abstraction layer). Una unidad de NAL se puede definir como una estructura de sintaxis que incluye una indicación del tipo de datos a seguir y bytes que incluyen esos datos en forma de una carga útil de secuencia de bytes en bruto (RBSP: raw byte sequence payload) intercalada, según sea necesario, con bytes de prevención de emulación. La carga útil RBSP puede ser una estructura de sintaxis que incluye un número entero de bytes encapsulados en una unidad de NAL. Una carga útil RBSP puede estar vacía o tener la forma de una cadena de bits de datos que incluye elementos de sintaxis seguidos de un bit de parada de carga útil RBSP y seguidos de cero o más bits subsiguientes iguales a 0. La Tabla 8 define la sintaxis de unidades de NAL.

**Tabla 8. Sintaxis de unidades de NAL**

nal_unit( NumBytesInNALunit ) {	Descriptor
<b>forbidden zero bit</b>	f(1)
<b>nal_ref_idc</b>	u(2)
<b>nal_unit_type</b>	u(5)
NumBytesInRBSP = 0	
nalUnitHeaderBytes = 1	
si( nal_unit_type == 1    nal_unit_type == 4    nal_unit_type == 5 ) {	
<b>temporal_id</b>	u(3)
<b>output_flag</b>	u(1)
<b>reserved_one_4bits</b>	u(4)
nalUnitHeaderBytes += 1	
}	
para( i = nalUnitHeaderBytes; i < NumBytesInNALunit; i++ ) {	
si( I + 2 < NumBytesInNALunit && next_bits(24) == 0x000003 ) {	
<b>rbsp_byte</b> [ NumBytesInRBSP++ ]	b(8)
<b>rbsp_byte</b> [ NumBytesInRBSP++ ]	b(8)
i += 2	
<b>emulation_prevention_three_byte</b> /* igual a 0x03 */	f(8)
} sino	
<b>rbsp_byte</b> [ NumBytesInRBSP++ ]	b(8)
}	
}	

20

En la Tabla 8, el indicador `output_flag` puede afectar al output de imágenes decodificadas y al proceso de eliminación según se describe con más detalle a continuación. Para cualquier imagen, si `output_flag` es igual a 1, la imagen está destinada para el output. De lo contrario, la imagen nunca se incluye en el output. En las técnicas que se describen en esta divulgación, la variable `OutputFlag` es igual al elemento de sintaxis `output_flag`.

En algunos ejemplos, cualquier unidad de NAL de corte (slice) codificada de la imagen codificada de la unidad de acceso actual puede ser diferente de cualquier unidad de NAL de corte codificada de la imagen codificada de la unidad de acceso anterior de una o más de las siguientes maneras. Por ejemplo, los valores de `pic_parameter_set_id` pueden ser diferentes, los valores de `nal_ref_idc` pueden ser diferentes, siendo uno de los valores de `nal_ref_idc` igual a 0. Los valores de `pic_order_cnt_lsb` pueden ser diferentes. Los valores de `IdrPicFlag` pueden ser diferentes. El `IdrPicFlag` puede ser igual a 1 en ambos casos, y los valores de `idr_pic_id` pueden ser diferentes.

En las técnicas que se describen en esta divulgación, la unidad de acceso se puede definir como un conjunto de unidades de NAL que son consecutivas en un orden de decodificación y contienen una imagen codificada. Además de la imagen codificada, una imagen codificada auxiliar, u otras unidades de NAL pueden no contener cortes (slices) de una imagen codificada. En algunos ejemplos, la decodificación de una unidad de acceso puede dar lugar a una imagen decodificada. La imagen codificada puede ser la representación codificada de una imagen a utilizar en el proceso de decodificación.

40

Según se indica en la Tabla 4, la sintaxis de la cabecera de corte (slice) puede incluir el elemento de sintaxis pic\_parameter\_set\_id, el elemento de sintaxis pic\_order\_cnt\_lsb, el elemento de sintaxis IdrPicFlag, y el elemento de sintaxis idr\_pic\_id. Según se indica en la Tabla 8, la sintaxis de unidad de NAL puede incluir el elemento de sintaxis nal\_ref\_idc.

5

A efectos de ilustración, las técnicas de gestión del búfer DPB se describen desde la perspectiva de un hipotético decodificador de referencia (HRD: hypothetical reference decoder). El HRD se puede definir como un modelo hipotético de decodificador que especifica restricciones a la variabilidad de flujos de unidades de NAL de conformación o de flujos de bytes de conformación que puede producir un proceso de codificación. Sin embargo, de acuerdo con las técnicas que se describen en esta divulgación, el decodificador de video 30 puede implementar las técnicas de gestión del búfer DPB, y en algunos ejemplos, puede ser posible que el codificador de video (video encoder) 20 también implemente las técnicas de gestión del búfer DPB.

El modelo HDR puede definir un búfer de imágenes codificadas (CPB: coded picture buffer), un proceso de decodificación instantánea y un búfer de imágenes decodificadas (DPB: decoded picture buffer). El búfer CPB puede ser similar al búfer CPB de los modelos HDR definidos en otros estándares anteriores (es decir, el búfer CPB puede almacenar imágenes codificadas). Las técnicas que se describen en esta divulgación están enfocadas a operaciones del búfer DPB que son diferentes a las operaciones en otros estándares. Una vez más, se debe entender que el decodificador de video 30 y posiblemente el codificador de video (video encoder) 20 pueden implementar las operaciones del búfer DPB según se describe a continuación.

En general, en las técnicas que se describen en esta divulgación están relacionadas con el output y eliminación de imágenes decodificadas en el búfer DPB. El output de una imagen decodificada, en este contexto, significa el output de la imagen decodificada para su visualización, almacenamiento u otros fines. Sin embargo, una imagen decodificada que se suministra como output no tiene por qué ser necesariamente eliminada del búfer DPB. Por ejemplo, es posible que el decodificador de video 30 no elimine del búfer DPB una imagen decodificada que se suministra como output porque es posible que el decodificador de video 30 necesite utilizar esa imagen decodificada como imagen de referencia para la inter predicción de una imagen posterior. La eliminación de una imagen decodificada, en este contexto, significa la eliminación de la imagen decodificada del búfer DPB.

30

Por ejemplo, el decodificador de video 30 puede almacenar imágenes decodificadas en el búfer DPB del decodificador de video 30 en el orden en el que se decodifican las imágenes. Sin embargo, el orden de decodificación de las imágenes puede no ser el mismo que el orden de output de las imágenes. Por ejemplo, puede haber imágenes subsiguientes respecto a una imagen actual en un orden de decodificación que se deben suministrar como output antes que la imagen actual. Por consiguiente, en algunos ejemplos, el decodificador de video 30 puede realizar un reordenamiento mediante el cual el decodificador de video 30 reordena en un orden de output las imágenes en el búfer DPB que están ordenadas en orden de decodificación. El decodificador de video 30 puede entonces suministrar como output las imágenes decodificadas en su orden de output. El decodificador de video 30 también puede eliminar imágenes de la imagen decodificada si la imagen no es necesaria para el output (es decir, si ha sido suministrada como output o no está destinada para el output) y no es necesaria para la inter predicción (es decir, no es necesario utilizarla como imagen de referencia para la inter predicción).

En las técnicas que se describen en esta divulgación, el decodificador de video 30 puede eliminar una imagen decodificada del búfer DPB si la imagen decodificada ha sido suministrada como output o no está destinada a ser suministrada como output, y si la imagen decodificada no está identificada en el conjunto de imágenes de referencia derivado, lo que equivale a que ya no se necesita para la referencia de inter predicción (es decir, ya no es necesario utilizarla como imagen de referencia para la inter predicción). Una vez más, según se ha descrito anteriormente, el conjunto de imágenes de referencia puede identificar imágenes de referencia que se pueden utilizar potencialmente para la inter predicción de la imagen actual, y que se pueden utilizar potencialmente para la inter predicción de una o más imágenes posteriores a la imagen actual en un orden de decodificación. De acuerdo con las técnicas que se describen en esta divulgación, si una imagen decodificada no está identificada en el conjunto de imágenes de referencia derivado, entonces esa imagen decodificada puede no ser necesaria como imagen de referencia para la inter predicción (por ejemplo, la decodificación) de la imagen actual y una o más imágenes posteriores a la imagen actual en un orden de decodificación. Por lo tanto, dicha imagen decodificada se puede eliminar del búfer DPB si no es necesaria para el output, ya que puede no ser necesario mantenerla en el búfer DPB, ya que la imagen decodificada no se utilizará para la inter predicción.

Además, en las técnicas que se describen en esta divulgación, el decodificador de video 30 puede eliminar una imagen decodificada antes de decodificar una imagen actual. Por ejemplo, según se ha descrito anteriormente, el decodificador de video 30 puede derivar un conjunto de imágenes de referencia y construir una o más listas de imágenes de referencia antes de decodificar la imagen actual. A causa de que el decodificador de video 30 puede derivar el conjunto de imágenes de referencia antes de decodificar la imagen actual, el decodificador de video 30 puede estar configurado para determinar si una imagen decodificada que no es necesaria para el output debe eliminarse antes de decodificar

60

la imagen actual. Por ejemplo, después de derivar el conjunto de imágenes de referencia y antes de decodificar la imagen actual, el decodificador de video 30 puede determinar si una imagen decodificada suministrada como output o una imagen decodificada no destinada para el output no está identificada en el conjunto de imágenes de referencia. Entonces, antes de decodificar la imagen actual, el decodificador de video 30 puede eliminar la imagen decodificada que no se necesita para el output (es decir, que ya ha sido suministrada como output o que no está destinada para el output) si la imagen decodificada no está identificada en el conjunto de imágenes de referencia.

En algunos ejemplos, el decodificador de video 30 puede eliminar una imagen decodificada antes de decodificar la imagen actual. Sin embargo, el decodificador de video 30 puede eliminar la imagen decodificada después de construir la lista o listas de imágenes de referencia. Por ejemplo, el decodificador de video 30 puede derivar el conjunto de imágenes de referencia y puede construir las listas de imágenes de referencia en base al conjunto de imágenes de referencia. A continuación, antes de decodificar la imagen actual, el decodificador de video 30 puede eliminar la imagen decodificada. En algunos ejemplos, el decodificador de video 30 también puede suministrar como output la imagen decodificada después de construir la lista o listas de imágenes de referencia.

Esta divulgación describe las técnicas de eliminar imágenes decodificadas en el búfer DPB desde al menos dos perspectivas. En la primera perspectiva, el decodificador de video 30 puede eliminar imágenes decodificadas en base a un momento de output si las imágenes están destinadas a ser suministradas como output. En la segunda perspectiva, el decodificador de video 30 puede eliminar imágenes decodificadas en base a los valores de POC si las imágenes están destinadas a ser suministradas como output. En cualquiera de las dos perspectivas, el decodificador de video 30 puede eliminar las imágenes decodificadas que no se necesitan para el output (es decir, que ya se han suministrado como output o que no están destinadas a ser suministradas como output) cuando la imagen decodificada no está en el conjunto de imágenes de referencia, y antes de decodificar la imagen actual.

El búfer DPB puede incluir una pluralidad de búferes, y cada búfer puede almacenar una imagen decodificada a utilizar como imagen de referencia o que se guarda para ser suministrada como output en el futuro. Inicialmente, el búfer DPB está vacío (es decir, la compleción del búfer DPB está establecida igual a cero). En las técnicas de ejemplo que se describen, la eliminación de las imágenes decodificadas del búfer DPB se puede producir antes de la decodificación de la imagen actual, pero después de que el decodificador de video 30 analice la cabecera de corte (slice) del primer corte de la imagen actual.

En la primera perspectiva, las siguientes técnicas se pueden producir instantáneamente en el tiempo  $t_r(n)$  en la siguiente secuencia. En este ejemplo,  $t_r(n)$  es el momento de eliminación del búfer CPB (es decir, el tiempo de decodificación) de la unidad de acceso  $n$  que contiene la imagen actual. Según se describe en esta divulgación, las técnicas que se producen instantáneamente pueden significar que en el modelo HDR, se asume que la decodificación de una imagen es instantánea (es decir, rápida e ilimitada) con un período de tiempo para decodificar una imagen igual a cero.

Si la imagen actual es una imagen de IDR, y cuando la imagen de IDR no es la primera imagen de IDR y el valor de `pic_width_in_luma_samples` o `pic_height_in_luma_samples` o `max_dec_frame_buffering` derivado a partir del conjunto de parámetros de secuencia activo es diferente del valor de `pic_width_in_luma_samples` o `pic_height_in_luma_samples` o `max_dec_frame_buffering` derivado a partir del conjunto de parámetros de secuencia que estaba activo para la imagen anterior, respectivamente, el decodificador de video 30 puede inferir que el elemento de sintaxis `no_output_of_prior_pics_flag` es igual a 1, con independencia del valor real del indicador `no_output_of_prior_pics_flag`. Si la imagen actual es una imagen de IDR, y cuando el indicador `no_output_of_prior_pics_flag` es igual a 1 o se infiere que es igual a 1, el decodificador de video 30 puede vaciar todos los búferes del búfer DPB sin suministrar como output las imágenes que hay en el búfer DPB, y puede establecer la compleción del búfer DPB igual a 0.

Según se ha indicado anteriormente en la Tabla 1, el conjunto de parámetros de secuencia puede incluir los elementos de sintaxis `pic_width_in_luma_samples` y `pic_height_in_luma_samples`. El conjunto de parámetros de secuencia también puede incluir el elemento de sintaxis `max_dec_frame_buffering`. Según se ha indicado en la Tabla 4, la sintaxis de la cabecera de corte (slice) puede incluir el elemento de sintaxis `no_output_of_prior_pics_flag`.

Cuando la imagen actual no es una imagen de IDR, el decodificador de video 30 puede eliminar todas las imágenes (m) del búfer DPB para las cuales las siguientes condiciones son ciertas. La primera condición puede ser que la imagen no esté incluida en el conjunto de imágenes de referencia de la imagen actual. La segunda condición puede ser que la imagen tenga un `OutputFlag` igual a 0 o que su momento de output del búfer DPB sea inferior o igual al momento de eliminación del búfer CPB de la imagen actual. En este ejemplo, el momento de eliminación del búfer CPB es  $t_r(n)$ , que es la instancia en la que se produce el proceso de eliminación (por ejemplo, un momento anterior a la decodificación de la imagen actual). El momento de output del búfer DPB de una imagen decodificada  $m$  puede ser definido por la variable  $t_{o,dpb}(m)$ . Por lo tanto, el momento de output del búfer DPB que es inferior o igual al momento

de eliminación del búfer CPB se puede representar como  $t_{o,dpb}(m) \leq t_r(n)$ . La derivación del momento de output del búfer DPB ( $t_{o,dpb}$ ) se define con más detalle a continuación.

De esta manera, el decodificador de video 30 puede eliminar imágenes decodificadas del búfer DPB, antes de  
5 decodificar una imagen, en base al momento de output de la imagen decodificada y cuando la imagen decodificada no está identificada en el conjunto de imágenes de referencia. Cuando el decodificador de video 30 elimina una imagen decodificada del búfer DPB, el decodificador de video 30 puede decrementar la compleción del búfer DPB en uno.

A continuación se describe la manera en la que el decodificador de video 30 puede determinar el momento de output  
10 de la imagen decodificada (por ejemplo, el momento de output del búfer DPB de una imagen decodificada), y también se describe cuándo el decodificador de video 30 puede almacenar la imagen decodificada en el búfer DPB. Según se ha descrito anteriormente, el momento de output del búfer DPB de la imagen puede ser un factor a tener en cuenta para determinar si esa imagen se elimina o no del búfer DPB.

15 Cuando el decodificador de video 30 decodifica una imagen, el decodificador de video 30 almacena la imagen en el búfer DPB e incrementa la compleción del búfer DPB en uno. Cuando la imagen tiene un OutputFlag igual a 1, el decodificador de video 30 puede derivar el momento de output del búfer DPB para la imagen en base a la siguiente ecuación.

$$20 \quad t_{o,dpb}(n) = t_r(n) + t_c * dpb\_output\_delay(n)$$

En la ecuación, se puede especificar el  $dpb\_output\_delay(n)$  en el mensaje SEI de temporización de imagen asociado con la unidad de acceso que incluye la imagen. El mensaje SEI puede estar definido en algunos estándares tales como el estándar H.264/AVC.

25 El valor  $t_{o,dpb}(n)$  puede definir cuándo se debe suministrar como output una imagen. Por ejemplo, si OutputFlag es igual a 1 y  $t_{o,dpb}(n)$  se iguala con  $t_r(n)$ , el decodificador de video 30 puede suministrar como output la imagen. De lo contrario, si OutputFlag es igual a 0, es posible que el decodificador de video 30 no suministre la imagen como output. En los casos en que OutputFlag es igual a 1 y  $t_{o,dpb}(n)$  es superior a  $t_r(n)$ , el decodificador de video 30 puede suministrar  
30 como output la imagen en un momento posterior (por ejemplo, en el momento  $t_{o,dpb}(n)$ ).

En algunos ejemplos, cuando el decodificador de video 30 suministra una imagen como output, el decodificador de video 30 puede recortar la imagen. Por ejemplo, el decodificador de video 30 puede utilizar el rectángulo de recorte especificado en el conjunto de parámetros de secuencia activo para la imagen. Las técnicas para recortar la imagen  
35 están generalmente bien establecidas y descritas en estándares tales como el estándar H.264/AVC.

En algunos ejemplos, el decodificador de video 30 puede determinar una diferencia entre el momento de output del búfer DPB para una imagen y el momento de output del búfer DPB para una imagen que sigue a la imagen en orden de output. Por ejemplo, cuando la imagen (n) es una imagen que el decodificador de video 30 suministra como output,  
40 y no es la última imagen del flujo de bits que se suministra como output, el decodificador de video 30 puede determinar el valor de  $\Delta t_{o,dpb}(n)$  se define como:

$$\Delta t_{o,dpb}(n) = t_{o,dpb}(n_n) - t_{o,dpb}(n)$$

45 En la ecuación anterior,  $n_n$  indica la imagen que sigue después de la imagen (n) en orden de output y tiene OutputFlag igual a 1. Además, en la ecuación anterior,  $\Delta t_{o,dpb}(n)$  representa la diferencia en los momentos de output del búfer DPB entre una imagen y la imagen siguiente en orden de output.

En la segunda perspectiva para la eliminación de imágenes decodificadas, el HDR puede implementar las técnicas  
50 instantáneamente cuando se elimina una unidad de acceso del búfer CPB. Una vez más, el decodificador de video 30 puede implementar la eliminación de imágenes decodificadas del búfer DPB, y el decodificador de video 30 puede no incluir necesariamente el búfer CPB. En general, en esta divulgación, la eliminación de imágenes decodificadas es realizada por el decodificador de video 30, y también puede realizarla el codificador de video (video encoder) 20. En estos ejemplos, es posible que el decodificador de video 30 y el codificador de video (video encoder) 20 no necesiten  
55 el búfer CPB. Más bien, el búfer CPB se describe como parte del modelo HDR sólo a efectos de ilustración.

Según lo anterior, en la segunda perspectiva para eliminar imágenes decodificadas, el decodificador de video 30 puede eliminar las imágenes del búfer DPB antes de la decodificación de la imagen actual, pero después de analizar la  
60 cabecera de corte (slice) de la primera corte de la imagen actual. También, de forma similar a la primera perspectiva para eliminar imágenes decodificadas, en la segunda perspectiva, el decodificador de video 30 puede realizar funciones similares a las que se han descrito anteriormente con respecto a la primera perspectiva cuando la imagen actual es una imagen de IDR.

De lo contrario, si la imagen actual no es una imagen de IDR, el decodificador de video 30 puede vaciar, sin suministrar como output, los búferes del búfer DPB que almacenan una imagen que está marcada como "no necesaria para el output" y que almacenan imágenes no incluidas en el conjunto de imágenes de referencia de la imagen actual. El decodificador de video 30 también puede decrementar la compleción del búfer DPB en el número de búferes que el  
 5 decodificador de video 30 ha vaciado. Cuando no hay ningún búfer vacío (es decir, la compleción del búfer DPB es igual al tamaño del búfer DPB), el decodificador de video 30 puede implementar un proceso de "choque" ("bumping") que se describe a continuación. En algunos ejemplos, cuando no hay ningún búfer vacío, el decodificador de video 30 puede implementar el proceso de choque repetidamente hasta que haya un búfer vacío en el que el decodificador de video 30 puede almacenar la imagen decodificada actual.

10

Cuando la imagen actual es una imagen de IDR para la cual el indicador no\_output\_of\_prior\_pics\_flag no es igual a 1, y no se infiere que es igual a 1, el decodificador de video 30 puede realizar lo siguiente. El decodificador de video 30 puede vaciar, sin suministrar como output, búferes del búfer DPB que almacenan una imagen que está marcada como "no necesaria para el output" y que no está incluida en el conjunto de imágenes de referencia de la imagen  
 15 actual. El decodificador de video 30 puede vaciar todos los búferes no vacíos del búfer DPB invocando repetidamente el proceso de "choque" o "bumping", y puede establecer la compleción del búfer DPB igual a 0.

En otras palabras, cuando la imagen actual es una imagen de IDR, el decodificador de video 30 puede implementar técnicas para vaciar todos los búferes del búfer DPB. Cuando la imagen actual no es una imagen de IDR, el  
 20 decodificador de video 30 puede implementar técnicas para eliminar imágenes decodificadas para liberar búferes para almacenar la imagen decodificada actual.

Por ejemplo, después de que el decodificador de video 30 haya decodificado la imagen actual, el decodificador de video 30 puede almacenar la imagen actual en el búfer DPB, e incrementar la compleción del búfer DPB en uno. En  
 25 algunos ejemplos, si el indicador OutputFlag de la imagen actual es igual a 1, el decodificador de video 30 puede marcar la imagen actual como "necesaria para el output". De lo contrario, si el indicador OutputFlag de la imagen actual es igual a 0, el decodificador de video 30 puede marcar la imagen actual como "no necesaria para el output".

Según se ha descrito anteriormente, en algunos ejemplos, el decodificador de video 30 puede implementar un proceso  
 30 de choque. En general, el proceso de choque implica suministrar como output imágenes decodificadas. Por ejemplo, el decodificador de video 30 puede implementar el proceso de choque cuando la imagen actual es una imagen de IDR y el indicador no\_output\_of\_prior\_pics\_flag no es igual a 1, y no se infiere que sea igual a 1. El decodificador de video 30 también puede implementar el proceso de choque si no hay un búfer vacío en el búfer DPB (es decir, la compleción del búfer DPB es igual al tamaño del búfer DPB), y se necesita un búfer vacío para almacenar una imagen decodificada  
 35 (no IDR).

En general, el decodificador de video 30 puede implementar las siguientes etapas para implementar el proceso de choque (bumping). El decodificador de video 30 puede determinar primero la imagen a suministrar como output. Por  
 40 ejemplo, el decodificador de video 30 puede seleccionar la imagen que tiene el valor de PicOrderCnt (POC) más pequeño de todas las imágenes en el búfer DPB que están marcadas como "necesarias para el output". El decodificador de video 30 puede recortar la imagen seleccionada utilizando el rectángulo de recorte especificado en el conjunto de parámetros de secuencia activo para la imagen. El decodificador de video 30 puede suministrar como output la imagen recortada y puede marcarla como "no necesaria para el output". El decodificador de video 30 puede comprobar el búfer del búfer DPB que ha almacenado la imagen recortada y suministrada como output. Si la imagen  
 45 no está incluida en el conjunto de imágenes de referencia, el decodificador de video 30 puede vaciar ese búfer y puede decrementar la compleción del búfer DPB en uno.

Aunque las técnicas anteriores para la gestión del búfer DPB se han descrito desde el contexto del decodificador de video 30, en algunos ejemplos, el codificador de video (video encoder) 20 puede implementar técnicas similares. Sin  
 50 embargo, el codificador de video (video encoder) 20 que implementa técnicas similares no es necesario en todos los ejemplos. En algunos ejemplos, el decodificador de video 30 puede implementar estas técnicas, y el codificador de video (video encoder) 20 puede no implementar estas técnicas.

De esta manera, un codificador de video (video coder) (por ejemplo, el codificador de video (video encoder) 20 o el  
 55 decodificador de video 30) puede convertir según una codificación información indicativa de imágenes de referencia que pertenecen a un conjunto de imágenes de referencia. Una vez más, el conjunto de imágenes de referencia puede identificar las imágenes de referencia que se pueden utilizar potencialmente para la inter predicción de la imagen actual y que se pueden utilizar potencialmente para la inter predicción de una o más imágenes posteriores a la imagen actual en un orden de decodificación.

60

El codificador de video (video coder) puede derivar el conjunto de imágenes de referencia de cualquier manera, incluyendo las técnicas de ejemplo que se han descrito anteriormente. El codificador de video (video coder) puede determinar si una imagen decodificada que está almacenada en el búfer de imágenes decodificadas no es necesaria

para el output y no está identificada en el conjunto de imágenes de referencia. Cuando se ha suministrado la imagen decodificada como output y no está identificada en el conjunto de imágenes de referencia, el codificador de video (video coder) puede eliminar la imagen decodificada del búfer de imágenes decodificadas. Después de eliminar la imagen decodificada, el codificador de video (video coder) puede codificar la imagen actual. Por ejemplo, el codificador de video (video coder) puede construir la lista o listas de imágenes de referencia según se ha descrito anteriormente, y codificar la imagen actual en base a la lista o listas de imágenes de referencia.

Los ejemplos anteriores han descrito las técnicas que el codificador de video (video encoder) 20 y el decodificador de video 30 pueden utilizar para derivar el conjunto de imágenes de referencia, construyendo listas de imágenes de referencia a partir de las listas de imágenes de referencia cuando no es necesario modificarlas y cuando es necesario hacerlo, así como técnicas para la gestión del búfer de imágenes decodificadas (DPB). Sin embargo, los aspectos de esta divulgación no son tan limitados. En algunos ejemplos, las técnicas que se describen en esta divulgación pueden estar relacionadas con la manera en la que el codificador de video (video encoder) señala qué imágenes pertenecen al conjunto de imágenes de referencia y son imágenes de referencia de largo plazo (o en otras palabras, qué imagen pertenece al conjunto de imágenes de referencia de largo plazo), y la manera en la que el decodificador de video determina qué imagen pertenece al conjunto de imágenes de referencia de largo plazo.

Por ejemplo, la Tabla 2 incluye los elementos de sintaxis `num_long_term_ref_pics_pps` y `long_term_ref_pic_id_pps[i]` como parte del conjunto de parámetros de imagen. Sin embargo, aspectos de esta divulgación no son tan limitados. En otros ejemplos, el conjunto de parámetros de secuencia (por ejemplo, la Tabla 1) puede incluir los elementos de sintaxis `num_long_term_ref_pics_pps` y `long_term_ref_pic_id_pps[i]`. En ejemplos en los que el conjunto de parámetros de secuencia incluye estos elementos de sintaxis, esta divulgación se puede referir a los elementos de sintaxis como `num_long_term_ref_pics_sps` y `long_term_ref_pic_id_sps[i]` para evitar confusiones. Con fines de ilustración, las técnicas se describen con ejemplos en los que el conjunto de parámetros de secuencia incluye estos elementos de sintaxis.

De modo similar a la definición de `num_long_term_ref_pics_pps`, el elemento de sintaxis `num_long_term_ref_pics_sps` puede especificar el número de imágenes de referencia de largo plazo candidatas que se incluyen en el conjunto de parámetros de secuencia. El valor de `num_long_term_ref_pics_sps` puede estar en el rango de 0 hasta 32, inclusive. De forma similar a la definición del elemento de sintaxis `long_term_ref_pic_id_pps[i]`, el elemento de sintaxis `long_term_ref_pic_id_sps[i]` puede especificar la información de identificación de la *i*-ésima imagen de referencia de largo plazo incluida en el conjunto de parámetros de secuencia.

En algunos ejemplos, el elemento de sintaxis `long_term_ref_pic_id_sps[i]` puede indicar imágenes de referencia de largo plazo candidatas que pertenecen a un conjunto de imágenes de referencia de una imagen actual. Las imágenes de referencia de largo plazo candidatas son una o más de las imágenes de referencia de largo plazo candidatas que pueden ser imágenes de referencia de largo plazo que el decodificador de video 30 puede utilizar para la inter predicción de la imagen actual o una o más imágenes posteriores a la imagen actual en un orden de decodificación. En otras palabras, las imágenes de referencia de largo plazo candidatas pueden indicar imágenes que son imágenes de referencia de largo plazo y que se pueden utilizar posiblemente para la inter predicción de la imagen actual y para la inter predicción de una o más imágenes posteriores a la imagen actual en un orden de decodificación. En algunos ejemplos, el elemento de sintaxis `long_term_ref_pic_id_sps[i]` puede incluir los valores de POC para las imágenes de referencia de largo plazo candidatas.

Sin embargo, no todas las imágenes de referencia de largo plazo candidatas se utilizan necesariamente para la inter predicción. Por ejemplo, no todas las imágenes de referencia de largo plazo candidatas pertenecen al conjunto de imágenes de referencia de una imagen actual. Más bien, cero o más de las imágenes de referencia de largo plazo candidatas pertenecen al conjunto de imágenes de referencia.

En las técnicas que se describen en esta divulgación, el codificador de video (video encoder) 20 puede señalar el elemento de sintaxis `long_term_ref_pic_id` en un conjunto de parámetros (por ejemplo, el elemento de sintaxis `long_term_ref_pic_id_sps` en el conjunto de parámetros de secuencia, o el elemento de sintaxis `long_term_ref_pic_id_pps` en el conjunto de parámetros de imagen). El decodificador de video 30 puede recibir el elemento de sintaxis `long_term_ref_pic_id` e identificar las imágenes de referencia de largo plazo candidatas. De acuerdo con las técnicas que se describen en esta divulgación, el decodificador de video 30 puede determinar además cuáles de las imágenes de referencia de largo plazo candidatas pertenecen al conjunto de imágenes de referencia. Por ejemplo, el decodificador de video 30 puede estar configurado para realizar esta determinación en base a elementos de sintaxis adicionales señalizados por el codificador de video (video encoder) 20 en el flujo de bits codificado.

Según se ha indicado en la Tabla 4, el codificador de video (video encoder) 20 puede señalar la estructura de sintaxis `long_term_ref_pic_set( )` en la cabecera de corte (slice) de la imagen actual. La Tabla 5 describe la estructura de sintaxis `long_term_ref_pic_set( )`. Por ejemplo, la estructura de sintaxis `long_term_ref_pic_set( )` puede incluir los



- elementos de sintaxis `num_long_term_pps_curr` y `num_long_term_pps_foll`. Una vez más, se debe tener en cuenta que, aunque los elementos de sintaxis `num_long_term_pps_curr` y `num_long_term_pps_foll` se definen como el número de imágenes de referencia de largo plazo incluidas en el conjunto de parámetros de secuencia, en ejemplos en los que las imágenes de referencia de largo plazo candidatas se incluyen en el conjunto de parámetros de secuencia, estos elementos de sintaxis pueden definir el número de imágenes de referencia de largo plazo candidatas incluidas en el conjunto de parámetros de secuencia. Por ejemplo, para evitar confusiones, el elemento de sintaxis `num_long_term_pps_curr` se puede denominar elemento de sintaxis `num_long_term_sps_curr`, y el elemento de sintaxis `num_long_term_pps_foll` se puede denominar elemento de sintaxis `num_long_term_sps_foll`.
- 10 De forma similar al elemento de sintaxis `num_long_term_pps_curr`, el elemento de sintaxis `num_long_term_sps_curr` puede definir el número de todas las imágenes de referencia de largo plazo cuya información de identificación se incluye en el conjunto de parámetros de secuencia referenciado, como las imágenes de referencia de largo plazo candidatas, y que se pueden utilizar para la inter predicción de la imagen actual y de una o más imágenes posteriores a la imagen actual en un orden de decodificación. De forma similar al elemento de sintaxis `num_long_term_pps_foll`,
- 15 el elemento de sintaxis `num_long_term_sps_foll` puede definir el número de todas las imágenes de referencia de largo plazo cuya información de identificación se incluye en el conjunto de parámetros de secuencia, como las imágenes de referencia de largo plazo candidatas, que no se utilizan para la inter predicción de la imagen actual, y que se pueden utilizar para la inter predicción de una o más imágenes posteriores a la imagen actual en un orden de decodificación.
- 20 Además, la estructura de sintaxis `long_term_ref_pic_set( )` señalizada en la cabecera de corte (slice) puede incluir el elemento de sintaxis `long_term_ref_pic_set_idx_pps[i]`. Una vez más, en ejemplos en los que las imágenes de referencia de largo plazo candidatas se señalizan en el conjunto de parámetros de secuencia, el elemento de sintaxis `long_term_ref_pic_set_idx_pps[i]` se puede considerar como el elemento de sintaxis `long_term_ref_pic_set_idx_sps[i]`. De forma similar al elemento de sintaxis `long_term_ref_pic_set_idx_pps[i]`, el elemento de sintaxis
- 25 `long_term_ref_pic_set_idx_sps[i]` puede definir el índice, en la lista de la información de identificación de imágenes de referencia de largo plazo candidatas en el conjunto de parámetros de secuencia referenciado, de la i-ésima imagen de referencia de largo plazo heredada del conjunto de parámetros de imagen de referencia al conjunto de imágenes de referencia de la imagen actual. En otras palabras, el elemento de sintaxis `long_term_ref_pic_set_idx_sps[i]` puede identificar un índice en la lista de las imágenes de referencia de largo plazo candidatas en el conjunto de parámetros de secuencia. A partir del índice, el decodificador de video 30 puede identificar una imagen de referencia de largo plazo en las imágenes de referencia de largo plazo candidatas, y puede determinar que la imagen de referencia de largo plazo identificada pertenece al conjunto de imágenes de referencia de la imagen actual.

Por ejemplo, el decodificador de video 30 puede implementar el siguiente pseudo código, similar al de la Tabla 5, para determinar cuáles de las imágenes de referencia de largo plazo candidatas pertenecen al conjunto de imágenes de referencia de la imagen actual.

```
para( i = 0; i < num_long_term_sps_curr + num_long_term_sps_foll; i++ )
    long_term_ref_pic_set_idx_sps[i]
```

- 40 De esta manera, el decodificador de video 30 puede decodificar elementos de sintaxis que indican imágenes de referencia de largo plazo que se identifican en un conjunto de parámetros. Por ejemplo, si el conjunto de parámetros es el conjunto de parámetros de secuencia, el decodificador de video 30 puede decodificar los elementos de sintaxis `long_term_ref_pic_id_sps[i]` que indican imágenes de referencia de largo plazo candidatas que están identificadas en el conjunto de parámetros de secuencia. Si el conjunto de parámetros es el conjunto de parámetros de imagen, el decodificador de video 30 puede decodificar los elementos de sintaxis `long_term_ref_pic_id_pps[i]` que indican imágenes de referencia de largo plazo candidatas que están identificadas en el conjunto de parámetros de imagen.

El decodificador de video 30 también puede decodificar elementos de sintaxis que indican qué imágenes de referencia de largo plazo, identificadas en el conjunto de parámetros, pertenecen al conjunto de imágenes de referencia de la imagen actual. Por ejemplo, si el conjunto de parámetros es un conjunto de parámetros de secuencia, el decodificador de video 30 puede decodificar los elementos de sintaxis `num_long_term_sps_curr`, `num_long_term_sps_foll` y `num_ref_pic_set_idx_sps[i]`, y si el conjunto de parámetros es el conjunto de parámetros de imagen, el decodificador de video 30 puede decodificar los elementos de sintaxis `num_long_term_pps_curr`, `num_long_term_pps_foll` y

55 `long_term_ref_pic_set_idx_pps[i]`. En cualquier ejemplo, el decodificador de video 30 puede decodificar estos elementos de sintaxis de la cabecera de corte (slice) de la imagen actual.

De acuerdo con las técnicas que se describen en esta divulgación, los elementos de sintaxis `long_term_ref_pic_id_sps[i]` y `long_term_ref_pic_id_pps[i]` se pueden considerar como una lista de valores de contador de orden de las imágenes (POC) para las imágenes de referencia de largo plazo candidatas que pertenecen al conjunto de imágenes de referencia y pueden ser convertidas según una codificación (es decir, codificadas o decodificadas) como parte de un conjunto de parámetros (por ejemplo, el conjunto de parámetros de imagen y el conjunto de parámetros de secuencia). Se puede considerar que el elemento de sintaxis `long_term_ref_pic_set_idx_sps[i]` o

long\_term\_ref\_pic\_set\_idx\_pps[i] proporciona un valor de índice en la lista de valores de POC de las imágenes de referencia de largo plazo candidatas (por ejemplo, un índice de long\_term\_ref\_pic\_id\_sps[i] o long\_term\_ref\_pic\_id\_pps[i]). En algunos ejemplos, el elemento de sintaxis long\_term\_ref\_pic\_set\_idx\_sps[i] o el elemento de sintaxis long\_term\_ref\_pic\_set\_idx\_pps[i] se puede codificar como parte de la cabecera de corte (slice) de la imagen actual.

Lo anterior describe una de las formas en las que el codificador de video (video encoder) 20 y el decodificador de video 30 pueden codificar o decodificar, respectivamente, elementos de sintaxis para indicar qué imágenes pertenecen al conjunto de imágenes de referencia de largo plazo de la imagen actual. A partir de los elementos de sintaxis, el decodificador de video 30 y el codificador de video (video encoder) 20 pueden determinar qué imágenes pertenecen al conjunto de imágenes de referencia de largo plazo de la imagen actual. Después de que el decodificador de video 30 y el codificador de video (video encoder) 20 hayan determinado qué imágenes pertenecen al conjunto de imágenes de referencia de largo plazo, el codificador de video (video encoder) 20 y el decodificador de video 30 pueden construir al menos un subconjunto de imágenes de referencia de la pluralidad de subconjuntos de imágenes de referencia y derivar el conjunto de imágenes de referencia de la manera que se ha descrito anteriormente. Por ejemplo, en base a la determinación de qué imágenes pertenecen al conjunto de imágenes de referencia de largo plazo, el codificador de video (video encoder) 20 y el decodificador de video 30 pueden construir el subconjunto de imágenes de referencia RefPicSetLtCurr que utiliza el codificador de video (video encoder) 20 y el decodificador de video 30 para derivar el conjunto de imágenes de referencia.

En algunos ejemplos, puede haber imágenes que pertenecen al conjunto de imágenes de referencia de largo plazo que no están incluidas en las imágenes de referencia de largo plazo candidatas. Por consiguiente, puede haber otras formas adicionales en las que el codificador de video (video encoder) 20 y el decodificador de video 30 pueden determinar qué imágenes pertenecen al conjunto de imágenes de referencia de largo plazo de la imagen actual.

Por ejemplo, según se indica en la Tabla 5, la estructura de sintaxis long\_term\_ref\_pic\_set( ) de la cabecera de corte (slice) incluye el elemento de sintaxis long\_term\_ref\_pic\_id\_delta\_add[i]. Este elemento de sintaxis puede especificar la identificación de imágenes de referencia de largo plazo, tal como los valores de POC, de la i-ésima imagen de referencia de largo plazo que no se hereda del conjunto de parámetros de imagen de referencia, sino que está incluida en el conjunto de imágenes de referencia de la imagen actual. Una vez más, en ejemplos en los que se identifican imágenes de referencia de largo plazo candidatas en el conjunto de parámetros de secuencia, el elemento de sintaxis long\_term\_ref\_pic\_id\_delta\_add[i] puede especificar la identificación de imágenes de referencia de largo plazo de la i-ésima imagen de referencia de largo plazo que no se hereda del conjunto de parámetros de secuencia, sino que está incluida en el conjunto de imágenes de referencia de la imagen actual.

En otras palabras, el elemento de sintaxis long\_term\_ref\_pic\_id\_pps[i] o long\_term\_ref\_pic\_id\_sps[i] puede identificar imágenes de referencia de largo plazo candidatas, pero no necesariamente identificar todas las imágenes de referencia de largo plazo del conjunto de imágenes de referencia de la imagen actual. Por ejemplo, puede haber imágenes de referencia de largo plazo a utilizar para la inter predicción de la imagen actual y una o más imágenes posteriores a la imagen actual en un orden de decodificación que no están incluidas en la lista de imágenes de referencia de largo plazo candidatas. Para dichas imágenes de referencia de largo plazo, el codificador de video (video encoder) 20 y el decodificador de video 30 pueden codificar o decodificar, respectivamente, información de identificación que identifica imágenes de referencia de largo plazo que pertenecen al conjunto de imágenes de referencia de la imagen actual.

Por ejemplo, según se indica en la Tabla 5, el codificador de video (video encoder) 20 y el decodificador de video 30 pueden codificar o decodificar, respectivamente, los elementos de sintaxis num\_long\_term\_add\_curr y num\_long\_term\_add\_foll. El elemento de sintaxis num\_long\_term\_add\_curr puede definir el número de todas las imágenes de referencia de largo plazo cuya información de identificación no está incluida en el conjunto de parámetros de imagen o secuencia referenciado (según corresponda) y que se puede utilizar en la inter predicción de la imagen actual y de una o más imágenes posteriores a la imagen actual en un orden de decodificación. El elemento de sintaxis num\_long\_term\_add\_foll puede definir el número de todas las imágenes de referencia de largo plazo cuya información de identificación no está incluida en el conjunto de parámetros de imagen o secuencia referenciado (según corresponda), que no se puede utilizar para la inter predicción de la imagen actual y que se puede utilizar para la inter predicción de una o más imágenes posteriores a la imagen actual en un orden de decodificación.

El decodificador de video 30 puede implementar el siguiente pseudo código para determinar qué imágenes de referencia de largo plazo pertenecen al conjunto de imágenes de referencia. En este ejemplo, las imágenes de referencia de largo plazo pueden no estar incluidas en las imágenes de referencia de largo plazo candidatas.

```

60 para( i = 0; i < num_long_term_add_curr + num_long_term_add_foll; i++)
    long_term_ref_pic_id_delta_add[i]

```

Según se ha descrito anteriormente, las técnicas que se describen en esta divulgación pueden ser realizadas de acuerdo con el estándar HEVC. La siguiente es una breve descripción del estándar HEVC para ayudar a la comprensión. Además, aunque las técnicas se describen en el contexto del estándar HEVC, las técnicas pueden ser ampliables a otros estándares, incluidos estándares propietarios.

5

El equipo JCT-VC está trabajando en el desarrollo del estándar HEVC. Los esfuerzos de estandarización del HEVC se basan en un modelo en evolución de un dispositivo de codificación de video denominado HEVC Test Model (HM). El modelo HM presupone diversas capacidades adicionales de los dispositivos de codificación de video en relación con dispositivos existentes de acuerdo con, por ejemplo, el estándar ITU-T H.264/AVC. Por ejemplo, si bien el estándar H.264 proporciona nueve modos de codificación de inter predicción, el modelo HM puede proporcionar hasta treinta y tres modos de codificación de inter predicción.

10

En general, el modelo de trabajo del modelo HM describe que un fotograma o imagen de video puede ser dividido en una secuencia de bloques de árbol o unidades de codificación más grandes (LCU: largest coding units) que incluyen tanto muestras de luma como de croma. Un bloque de árbol tiene un propósito similar al de un macro bloque del estándar H.264. un corte (slice) incluye un número de bloques de árbol consecutivos en orden de decodificación. Un fotograma o imagen de video se puede particionar en uno o más cortes. Cada bloque de árbol se puede particionar en unidades de codificación (CU: coding units) según un árbol cuaternario. Por ejemplo, un bloque de árbol, como nodo raíz del árbol cuaternario, se puede particionar en cuatro nodos hijo, y cada nodo hijo puede a su vez ser un

20

25

nodo padre y dividirse en otros cuatro nodos hijo. Un nodo hijo final no dividido, como nodo hoja del árbol cuaternario, comprende un nodo de codificación, es decir, un bloque de video codificado. Datos de sintaxis asociados con un flujo de bits codificado pueden definir un número máximo de veces que se puede dividir un bloque de árbol, y también pueden definir un tamaño mínimo de los nodos de codificación. Los bloques de árbol se pueden denominar LCU en algunos ejemplos.

Una unidad de codificación incluye un nodo de codificación y unidades de predicción (unidad de predicción: prediction units) y unidades de transformación (TU: transform units) asociadas con el nodo de codificación. Un tamaño de la unidad de codificación corresponde a un tamaño del nodo de codificación y debe tener forma cuadrada. El tamaño de la unidad de codificación puede variar desde  $8 \times 8$  píxeles hasta el tamaño del bloque de árbol con un máximo de  $64 \times 64$  píxeles o más. Cada unidad de codificación puede contener una o más unidades de predicción y una o más unidades de transformación. Datos de sintaxis asociados a una unidad de codificación pueden describir, por ejemplo, el particionado de la unidad de codificación en una o más unidades de predicción. Los modos de particionado pueden diferir entre si la unidad de codificación está codificada en modo de salto (skip) o directo, codificada en modo de intra predicción o codificada en modo de inter predicción. Las unidades de predicción pueden ser divididas para que no tengan forma cuadrada. Datos de sintaxis asociados a una unidad de codificación también pueden describir, por ejemplo, el particionado de la unidad de codificación en una o más unidades de transformación según un árbol cuaternario. Una unidad de transformación puede tener forma cuadrada o no cuadrada.

30

35

40

45

50

55

60

65

70

75

80

85

90

95

100

105

110

115

120

125

130

135

140

145

150

155

160

165

170

175

180

185

190

195

píxeles que pueden ser transformados en coeficientes de transformación, cuantificados y escaneados usando las unidades de transformación para producir coeficientes de transformación serializados para la codificación de entropía. Esta divulgación utiliza normalmente el término "bloque de video" para referirse a un nodo de codificación de una unidad de codificación. En algunos casos específicos, esta divulgación también puede utilizar el término "bloque de video" para referirse a un bloque de árbol, es decir, una unidad LCU, o una unidad de codificación CU, que incluye un nodo de codificación y unidades de predicción PU y unidades de transformación TU.

Una secuencia de video suele incluir una serie de fotogramas o imágenes de video. Un grupo de imágenes (GOP) generalmente comprende una serie de una o más de las imágenes de video. Un GOP puede incluir datos de sintaxis en una cabecera del GOP, una cabecera de una o más de las imágenes, u otra ubicación, que describe un número de imágenes incluidas en el GOP. Cada corte de una imagen puede incluir datos de sintaxis de corte que describen un modo de codificación para el respectivo corte. El codificador de video (video encoder) funciona normalmente en bloques de video dentro de cortes de video individuales para codificar los datos de video. Un bloque de video puede corresponder a un nodo de codificación dentro de una unidad de codificación. Los bloques de video pueden tener tamaños fijos o variables, y pueden diferir en tamaño según un estándar de codificación específico.

Como ejemplo, el modelo HM soporta la predicción en diversos tamaños de unidad de predicción. Asumiendo que el tamaño de una unidad de codificación en particular es igual a  $2N \times 2N$ , el modelo HM soporta la intra predicción en tamaños de unidad de predicción de  $2N \times 2N$  o  $N \times N$ , y la inter predicción en tamaños de unidad de predicción simétricos de  $2N \times 2N$ ,  $2N \times N$ ,  $N \times 2N$ , o  $N \times N$ . El modelo HM también soporta particionados asimétricos para la inter predicción en tamaños de unidad de predicción de  $2N \times nU$ ,  $2N \times nD$ ,  $nL \times 2N$ , y  $nR \times 2N$ . En el particionado asimétrico, una dirección de una unidad de codificación no es particionada, si bien la otra dirección es particionada en un 25% y 75%. La parte de la unidad de codificación que corresponde a la partición del 25% se indica con una "n" seguida de una indicación de "Up" (superior), "Down" (inferior), "Left" (izquierda) o "Right" (derecha). Así, por ejemplo, " $2N \times nU$ " se refiere a una unidad de codificación de  $2N \times 2N$  que es particionada horizontalmente con una unidad de predicción de  $2N \times 0,5N$  en la parte superior y una unidad de predicción de  $2N \times 1,5N$  en la parte inferior.

En esta divulgación, " $N \times N$ " y " $N$  por  $N$ " se pueden utilizar indistintamente para referirse a las dimensiones en píxeles de un bloque de video en términos de dimensiones verticales y horizontales, por ejemplo,  $16 \times 16$  píxeles o 16 por 16 píxeles. En general, un bloque de  $16 \times 16$  tendrá 16 píxeles en dirección vertical ( $y = 16$ ) y 16 píxeles en dirección horizontal ( $x = 16$ ). Asimismo, un bloque de  $N \times N$  generalmente tiene  $N$  píxeles en dirección vertical y  $N$  píxeles en dirección horizontal, en el que  $N$  representa un valor entero no negativo. Los píxeles de un bloque pueden estar dispuestos en filas y columnas. Además, los bloques no tienen por qué tener el mismo número de píxeles en la dirección horizontal que en la vertical. Por ejemplo, los bloques pueden comprender  $N \times M$  píxeles, en el que  $M$  no es necesariamente igual a  $N$ .

Tras la codificación intra predictiva o inter predictiva utilizando las unidades de predicción de una unidad de codificación, el codificador de video (video encoder) puede calcular los datos residuales para las unidades de transformación de la unidad de codificación. Las unidades de predicción pueden incluir datos de píxeles en el dominio espacial (también denominado dominio de píxeles) y las unidades de transformación pueden incluir coeficientes en el dominio de transformación tras la aplicación de una transformación, por ejemplo, una transformación de coseno discreta (DCT), una transformación entera, una transformación de ondículas o una transformación conceptualmente similar a los datos de video residuales. Los datos residuales pueden corresponder a diferencias de píxeles entre píxeles de la imagen sin codificar y valores de predicción que corresponden a las unidades de predicción. El codificador de video (video encoder) puede formar las unidades de transformación incluyendo los datos residuales para la unidad de codificación, y luego transformar las unidades de transformación para producir coeficientes de transformación para la unidad de codificación.

Después de cualesquiera transformaciones para producir coeficientes de transformación, el codificador de video (video encoder) puede realizar la cuantificación de los coeficientes de transformación. La cuantificación se refiere generalmente a un proceso en el que se cuantifican los coeficientes de transformación para reducir posiblemente la cantidad de datos utilizados para representar los coeficientes, lo que proporciona una mayor compresión. El proceso de cuantificación puede reducir la profundidad de bits asociada con algunos o todos los coeficientes. Por ejemplo, un valor de  $n$ -bits puede ser redondeado a un valor de  $m$ -bits durante la cuantificación, en el que  $n$  es superior a  $m$ .

En algunos ejemplos, el codificador de video (video encoder) puede utilizar un orden de escaneo predefinido para escanear los coeficientes de transformación cuantificados y producir un vector serializado que puede ser codificado por entropía. En otros ejemplos, el codificador de video (video encoder) puede realizar una exploración o escaneo adaptativo. Después de escanear los coeficientes de transformación cuantificados para formar un vector unidimensional, el codificador de video (video encoder) puede codificar por entropía el vector unidimensional, por ejemplo, de acuerdo con una codificación de longitud variable adaptable al contexto (CAVLC: context adaptive variable length coding), una codificación aritmética binaria adaptable al contexto (CABAC: context adaptive binary arithmetic coding), una codificación aritmética binaria adaptable al contexto basada en sintaxis (SBAC: syntax-based context-

adaptive binary arithmetic coding), una codificación de entropía de particionado a intervalos de probabilidad (PIPE: Probability Interval Partitioning Entropy) u otra metodología de codificación de entropía. El codificador de video (video encoder) 20 también puede codificar elementos de sintaxis asociados con datos de video codificados para su uso por el decodificador de video 30 en la decodificación de los datos de video.

5

Para realizar una codificación CABAC, el codificador de video (video encoder) 20 puede asignar un contexto dentro de un modelo de contexto a un símbolo a transmitir. El contexto puede estar relacionado, por ejemplo, con si los valores vecinos del símbolo son distintos de cero o no. Para realizar una codificación CAVLC, el codificador de video (video encoder) 20 puede seleccionar un código de longitud variable para un símbolo a transmitir. En una codificación

10

VLC, se pueden construir palabras clave de tal manera que códigos relativamente más cortos corresponden a símbolos más probables, mientras que códigos más largos corresponden a símbolos menos probables. De esta manera, el uso de la codificación VLC puede conseguir un pequeño ahorro de bits en comparación, por ejemplo, con el uso de palabras clave de igual longitud para cada símbolo a transmitir. La determinación de probabilidad se puede basar en un contexto asignado al símbolo.

15

La figura 2 es un diagrama conceptual que ilustra una secuencia de video de ejemplo 33 que incluye una pluralidad de imágenes que son codificadas y transmitidas. En algunos casos, la secuencia de video 33 se puede denominar grupo de imágenes (GOP). La secuencia de video 33, según se ilustra, incluye las imágenes 35A, 36A, 38A, 35B, 36B, 38B y 35C, y la imagen final 39 en orden de visualización. La imagen 34 es una imagen final en el orden de

20

visualización para una secuencia que se produce antes de la secuencia 33. La figura 2 representa en general una estructura de predicción de ejemplo para una secuencia de video y tiene por objeto únicamente ilustrar las referencias de imagen utilizadas para pronosticar bloques de video de diferentes tipos de corte o imagen (por ejemplo, imagen o corte P, o imagen o corte B). Una secuencia de video real puede contener más o menos imágenes de video de diferentes tipos de imagen y en un orden de visualización diferente. La secuencia de video 33 puede incluir más o

25

menos imágenes que las ilustradas en la figura 2, y las imágenes ilustradas en la secuencia de video 33 se ilustran a efectos de comprensión y como ejemplos.

Para la codificación de video por bloques, cada una de las imágenes de video incluidas en la secuencia 33 se puede particionar en bloques de video, tales como unidades de codificación (CU) o unidades de predicción (PU). Por ejemplo,

30

cada unidad de codificación de una imagen de video puede incluir una o más unidades de predicción. Se pronostican bloques de video en una imagen intra codificada (I) utilizando la predicción espacial con respecto a bloques vecinos en la misma imagen. Los bloques de video en una imagen inter codificada (P o B) pueden utilizar la predicción espacial con respecto a bloques vecinos en la misma imagen o predicción temporal con respecto a otras imágenes de referencia.

35

Los bloques de video en una imagen B pueden pronosticarse utilizando una predicción bidireccional para calcular dos vectores de movimiento a partir de dos listas de imágenes de referencia diferentes (por ejemplo, las listas 0 y 1 de imágenes de referencia, denominadas Lista 0 y Lista 1). En algunos casos, los bloques de video en una imagen B pueden pronosticarse utilizando una predicción unidireccional a partir de una de las dos listas de imágenes de

40

referencia diferentes (por ejemplo, codificación B unidireccional). Los bloques de video en una imagen P pueden pronosticarse utilizando la predicción unidireccional para calcular un solo vector de movimiento a partir de una sola lista de imágenes de referencia. De acuerdo con el nuevo estándar HEVC, los bloques de video pueden codificarse utilizando la predicción unidireccional para calcular un solo vector de movimiento a partir de una de las dos listas de imágenes de referencia o la predicción bidireccional para calcular dos vectores de movimiento a partir de las dos listas

45

de imágenes de referencia. Las dos listas de imágenes de referencia pueden contener imágenes de referencia pasadas o futuras, o imágenes de referencia pasadas y futuras en orden de visualización u orden de output, y siempre imágenes de referencia pasadas en orden de decodificación, por ejemplo.

En el ejemplo de la figura 2, se designa la imagen final 39 para una codificación intra modal como una imagen I. En

50

otros ejemplos, la imagen final 39 se puede codificar con una codificación inter modal (por ejemplo, como imagen P) con referencia a la imagen final 34 de la secuencia anterior, que puede ser una imagen I. Las imágenes de video 35A – 35C (colectivamente "imágenes de video 35") se designan para codificar como imágenes B utilizando la predicción bidireccional con referencia a una imagen pasada y una imagen futura. En el ejemplo ilustrado, la imagen 35A se codifica como una imagen B con referencia a la imagen final 34 y a la imagen 36A, según indican las flechas desde

55

las imágenes 34 y 36A hacia la imagen de video 35A. Las imágenes 35B y 35C son codificadas de forma similar.

Las imágenes de video 36A – 36B (colectivamente "imágenes de video 36") pueden designarse para su codificación como imágenes utilizando la predicción unidireccional con referencia a una imagen pasada. En el ejemplo ilustrado, la imagen 36A se codifica como una imagen P con referencia a la imagen final 34, según indica la flecha desde la

60

imagen 34 hacia la imagen de video 36A. La imagen 36B es codificada de forma similar.

Las imágenes de video 38A – 38B (colectivamente "imágenes de video 38") pueden ser designadas para codificación utilizando predicción bidireccional con referencia a la misma imagen pasada. En otros ejemplos, las imágenes de video

pueden codificarse utilizando la predicción bidireccional con referencia a imágenes pasadas sustancialmente similares incluidas en las listas de imágenes de referencia. En el ejemplo ilustrado, la imagen 38A se codifica con dos referencias a la imagen 36A, según indican las dos flechas desde la imagen 36A hacia la imagen de video 38A. La imagen 38B es codificada de forma similar.

5

De acuerdo con las técnicas que se describen en esta divulgación, el codificador de video (video encoder) 20 puede señalar un conjunto de imágenes de referencia para cada una de las imágenes de la secuencia 33. Por ejemplo, para la imagen 35A, este conjunto de imágenes de referencia puede identificar todas las imágenes de referencia que se pueden utilizar para la inter predicción de la imagen 35A, así como todas las imágenes de referencia que se pueden utilizar potencialmente para la inter predicción de imágenes posteriores a la imagen 35A en un orden de decodificación. Por ejemplo, el conjunto de imágenes de referencia para la imagen 35A puede incluir el valor de POC para la imagen 34 y la imagen 36A, así como los valores de POC para imágenes de referencia adicionales tales como las que se pueden utilizar potencialmente para la inter predicción de imágenes subsiguientes respecto a la imagen 35A en un orden de decodificación. Las imágenes posteriores a la imagen 35A pueden ser aquellas que siguen a la imagen 35A en un orden de decodificación, y que están dentro de la secuencia de video 33, en este ejemplo.

El decodificador de video 30 puede derivar entonces el conjunto de imágenes de referencia para la imagen 35A de la manera que se ha descrito anteriormente. Por ejemplo, el decodificador de video 30 puede determinar los valores de POC para las imágenes de referencia que pertenecen al conjunto de imágenes de referencia, según se ha descrito anteriormente. El decodificador de video 30 puede construir al menos cuatro o al menos cinco subconjuntos de imágenes de referencia y, en algunos ejemplos, hasta los seis subconjuntos de imágenes de referencia descritos anteriormente. El decodificador de video 30 puede disponer los seis conjuntos de imágenes de referencia en un orden particular para derivar el conjunto de imágenes de referencia para la imagen 35A.

El decodificador de video 30 puede construir además listas iniciales de imágenes de referencia de la manera que se ha descrito anteriormente, en la que no es necesario reordenar las imágenes a incluir en las listas iniciales de imágenes de referencia. Cuando la modificación de listas de imágenes de referencia está desactivada, el decodificador de video 30 puede establecer las listas finales de imágenes de referencia iguales a las listas iniciales de imágenes de referencia. Además, el decodificador de video 30 puede construir las listas de imágenes de referencia de tal manera que no haya entradas no completadas en las listas de imágenes de referencia. Por ejemplo, el decodificador de video 30 puede listar repetidamente imágenes de referencia de los subconjuntos de imágenes de referencia hasta que el número de entradas en las listas de imágenes de referencia sea igual al número máximo de entradas permisibles de las listas de imágenes de referencia. En algunos ejemplos, el decodificador de video 30 puede modificar las listas iniciales de imágenes de referencia de la manera que se ha descrito anteriormente (por ejemplo, en base a imágenes de referencia en al menos uno de entre la pluralidad de subconjuntos de imágenes de referencia construidos). Además, en algunos ejemplos, el decodificador de video 30 puede eliminar imágenes decodificadas del búfer DPB del decodificador de video 30 utilizando las técnicas de ejemplo que se describen en esta divulgación, tales como la eliminación de imágenes decodificadas que no están identificadas en el conjunto de imágenes de referencia de la imagen actual a decodificar y que no son necesarias para el output. También, en algunos ejemplos, el decodificador de video 30 puede determinar qué imágenes de referencia de largo plazo pertenecen al conjunto de imágenes de referencia de la manera que se ha descrito anteriormente, en la que la información de identificación de una lista de imágenes de referencia de largo plazo candidatas puede estar incluida en un conjunto de parámetros.

La figura 3 es un diagrama de bloques que ilustra un ejemplo de codificador de video (video encoder) 20 que puede implementar las técnicas que se describen en esta divulgación. El codificador de video (video encoder) 20 puede realizar una intra codificación e inter codificación de bloques de video dentro de cortes de video. La intra codificación se basa en la predicción espacial para reducir o eliminar la redundancia espacial en el video dentro de un fotograma o imagen de video determinado. La inter codificación se basa en la predicción temporal para reducir o eliminar la redundancia temporal en el video dentro de fotogramas o imágenes adyacentes de una secuencia de video. Intra-modos (modo I) se puede referir a cualquiera de los diversos modos de compresión basados en el espacio. Los inter modos, tales como la predicción unidireccional (modo P) o la bi-predicción (modo B), se pueden referir a cualquiera de los diversos modos de compresión basados en el tiempo.

En el ejemplo de figura 3, el codificador de video (video encoder) 20 incluye una unidad de particionado 35, un módulo de predicción 41, un búfer de imágenes decodificadas (DPB) 64, un sumador 50, un módulo de transformación 52, una unidad de cuantificación 54 y una unidad de codificación de entropía 56. El módulo de predicción 41 incluye una unidad de estimación de movimiento 42, una unidad de compensación de movimiento 44 y un módulo de inter predicción 46. Para la reconstrucción de bloques de video, el codificador de video (video encoder) 20 también incluye una unidad de cuantificación inversa 58, un módulo de transformación inversa 60 y un sumador 62. También se puede incluir un filtro de desbloqueo (que no se muestra en la figura 3) para filtrar los límites de los bloques y eliminar los artefactos de bloqueo del video reconstruido. Si se desea, el filtro de desbloqueo filtrará normalmente el output del sumador 62. También se pueden utilizar filtros en bucle adicionales (en bucle o en bucle posterior) además del filtro de desbloqueo.

Según se muestra en la figura 3, el codificador de video (video encoder) 20 recibe datos de video y la unidad de particionado 35 los particiona en bloques de video. Este particionado también puede incluir un particionado en cortes o cortes, piezas de mosaico u otras unidades más grandes, así como un particionado en bloques de video, por ejemplo, de acuerdo con una estructura de árbol cuaternario de unidades LCU y unidades de codificación. El codificador de video (video encoder) 20 generalmente ilustra los componentes que codifican bloques de video dentro de un corte de video a codificar. el corte se puede particionar en diversos bloques de video (y posiblemente en conjuntos de bloques de video denominados piezas de mosaico). El módulo de predicción 41 puede seleccionar uno de una pluralidad de modos de codificación posibles, tal como uno de una pluralidad de modos de intra codificación o uno de una pluralidad de modos de inter codificación, para el bloque de video actual en base a resultados de error (por ejemplo, la velocidad de codificación y el nivel de distorsión). El módulo de predicción 41 puede proporcionar el bloque intra codificado o inter codificado resultante al sumador 50 para generar datos de bloques residuales y al sumador 62 para reconstruir el bloque codificado y utilizarlo como imagen de referencia.

El módulo de intra predicción 46 dentro del módulo de predicción 41 puede realizar una codificación intra predictiva del bloque de video actual en relación con uno o más bloques vecinos en la misma imagen o corte que el bloque actual a codificar para proporcionar una compresión espacial. La unidad de estimación de movimiento 42 y la unidad de compensación de movimiento 44 dentro del módulo de predicción 41 realizan una codificación inter predictiva del bloque de video actual en relación con uno o más bloques predictivos en una o más imágenes de referencia para proporcionar una compresión temporal.

La unidad de estimación de movimiento 42 puede estar configurada para determinar el modo de inter predicción para un corte (slice) de video de acuerdo con un patrón predeterminado para una secuencia de video. El patrón predeterminado puede designar cortes de video en la secuencia como cortes P o cortes B. La unidad de estimación de movimiento 42 y la unidad de compensación de movimiento 44 pueden estar altamente integradas, pero se ilustran por separado para fines conceptuales. La estimación de movimiento, realizada por la unidad de estimación de movimiento 42, es el proceso de generación de vectores de movimiento, que estiman el movimiento de bloques de video. Un vector de movimiento, por ejemplo, puede indicar el desplazamiento de una unidad de predicción de un bloque de video dentro de una imagen de video actual en relación con un bloque predictivo dentro de una imagen de referencia.

Un bloque predictivo es un bloque que se determina que coincide estrechamente con la unidad de predicción del bloque de video a codificar en términos de diferencias de píxeles, lo cual se puede determinar mediante la suma de la diferencia absoluta (SAD: sum of absolute difference), suma de la diferencia cuadrática (SSD: sum of square difference) u otras métricas de diferencia. En algunos ejemplos, el codificador de video (video encoder) 20 puede calcular unos valores para posiciones de píxeles sub enteros de imágenes de referencia almacenadas en el búfer de imágenes decodificadas 64. Por ejemplo, el codificador de video (video encoder) 20 puede interpolar valores de posiciones de un cuarto de píxel, un octavo de píxel u otras posiciones de píxel fraccionadas de la imagen de referencia. Por lo tanto, la unidad de estimación de movimiento 42 puede realizar una búsqueda de movimiento relativo a posiciones de píxeles completas y a posiciones de píxeles fraccionados y generar como output un vector de movimiento con precisión de píxeles fraccionados.

La unidad de estimación de movimiento 42 calcula un vector de movimiento para una unidad de predicción de un bloque de video en un corte (slice) inter-codificado comparando la posición de la unidad de predicción con la posición de un bloque predictivo de una imagen de referencia. La imagen de referencia se puede seleccionar de una primera lista de imágenes de referencia (Lista 0) o de una segunda lista de imágenes de referencia (Lista 1), cada una de las cuales identifica una o más imágenes de referencia almacenadas en el búfer de imágenes decodificadas 64. La unidad de estimación de movimiento 42 suministra el vector de movimiento calculado a la unidad de codificación de entropía 56 y a la unidad de compensación de movimiento 44.

La compensación de movimiento, realizada por la unidad de compensación de movimiento 44, puede implicar la obtención o generación del bloque predictivo en base al vector de movimiento determinado por la estimación de movimiento, realizando posiblemente interpolaciones con una precisión de subpíxel. Al recibir el vector de movimiento para la unidad de predicción del bloque de video actual, la unidad de compensación de movimiento 44 puede localizar el bloque predictivo al que apunta el vector de movimiento en una de las listas de imágenes de referencia. El codificador de video (video encoder) 20 forma un bloque de video residual sustrayendo valores de píxel del bloque predictivo de valores de píxel del bloque de video actual que se está codificando, formando valores de diferencias de píxel. Los valores de diferencias de píxeles forman datos residuales para el bloque, y pueden incluir componentes de diferencias de luminancia y de croma. El sumador 50 representa el componente o componentes que realizan esta operación de sustracción. La unidad de compensación de movimiento 44 también puede generar elementos de sintaxis asociados con los bloques de video y el corte de video para que los utilice el decodificador de video 30 para decodificar los bloques de video del corte (slice) de video.

El módulo de intra predicción 46 puede realizar la intra predicción de un bloque actual, como alternativa a la inter predicción realizada por la unidad de estimación de movimiento 42 y la unidad de compensación de movimiento 44, según se ha descrito anteriormente. En particular, el módulo de intra predicción 46 puede determinar un modo de intra predicción a utilizar para codificar un bloque actual. En algunos ejemplos, el módulo de intra predicción 46 puede

5 codificar un bloque actual utilizando diversos modos de intra predicción, por ejemplo, durante pases de codificación separados, y el módulo de intra predicción 46 (o la unidad de selección de modo 40, en algunos ejemplos) puede seleccionar un modo de intra predicción apropiado a utilizar de entre los modos probados. Por ejemplo, el módulo de intra predicción 46 puede calcular valores de tasa-distorsión o velocidad-distorsión utilizando un análisis de tasa-distorsión para los distintos modos de intra predicción probados, y seleccionar el modo de intra predicción que tenga

10 las mejores características de tasa-distorsión entre los modos probados. El análisis de tasa-distorsión generalmente determina una cantidad de distorsión (o error) entre un bloque codificado y un bloque original no codificado que ha sido codificado para producir el bloque codificado, así como una velocidad o tasa de bits (es decir, un número de bits) utilizada para producir el bloque codificado. El módulo de intra predicción 46 puede calcular ratios a partir de las distorsiones y tasas de los distintos bloques codificados para determinar qué modo de intra predicción presenta el

15 mejor valor de tasa-distorsión para el bloque.

Después de seleccionar un modo de intra predicción para un bloque, el módulo de intra predicción 46 puede proporcionar información indicativa del modo de intra predicción seleccionado para el bloque a la unidad de codificación de entropía 56. La unidad de codificación de entropía 56 puede codificar la información que indica el modo

20 de intra predicción seleccionado de acuerdo con las técnicas de esta divulgación. El codificador de video (video encoder) 20 puede incluir en los datos de configuración de flujo de bits transmitidos, cuáles pueden incluir una pluralidad de tablas de índices de modo de intra predicción y una pluralidad de tablas de índices de modo de intra predicción modificadas (también denominadas tablas de mapeo de palabras clave), definiciones de contextos de codificación para diversos bloques e indicaciones de un modo de intra predicción más probable, una tabla de índices

25 de modo de intra predicción y una tabla de índices de modo de intra predicción modificada a utilizar para cada uno de los contextos.

Después de que el módulo de predicción 41 haya generado el bloque predictivo para el bloque de video actual mediante inter predicción o intra predicción, el codificador de video (video encoder) 20 forma un bloque de video

30 residual sustrayendo el bloque predictivo del bloque de video actual. Los datos de video residuales en el bloque residual pueden ser incluidos en una o más unidades de transformación y aplicados al módulo de transformación 52. El módulo de transformación 52 transforma los datos de video residuales en coeficientes de transformación residuales utilizando una transformación, tal como una transformación de coseno discreta (DCT) o una transformación conceptualmente similar. El módulo de transformación 52 puede convertir los datos de video residuales de un dominio

35 de píxeles a un dominio de transformación, tal como un dominio de frecuencia.

El módulo de transformación 52 puede enviar los coeficientes de transformación resultantes a la unidad de cuantificación 54. La unidad de cuantificación 54 cuantifica los coeficientes de transformación para reducir aún más la velocidad o tasa de bits. El proceso de cuantificación puede reducir la profundidad de bits asociada con algunos o

40 todos los coeficientes. El grado de cuantificación se puede modificar ajustando un parámetro de cuantificación. En algunos ejemplos, la unidad de cuantificación 54 puede entonces realizar un escaneo o exploración de la matriz que incluye los coeficientes de transformación cuantificados. Alternativamente, la unidad de codificación de entropía 56 puede realizar el escaneo.

Después de la cuantificación, la unidad de codificación de entropía 56 codifica por entropía los coeficientes de transformación cuantificados. Por ejemplo, la unidad de codificación de entropía 56 puede realizar una codificación de longitud variable adaptable al contexto (CAVLC), una codificación aritmética binaria adaptable al contexto (CABAC), una codificación aritmética binaria adaptable al contexto basada en sintaxis (SBAC), codificación de entropía de particionado por intervalo de probabilidad (PIPE) u otra metodología o técnica de codificación de entropía. Tras la

50 codificación de entropía realizada por la unidad de codificación de entropía 56, se puede transmitir el flujo de bits codificado al decodificador de video 30, o archivarlo para su posterior transmisión o recuperación por parte del decodificador de video 30. La unidad de codificación de entropía 56 también puede codificar los vectores de movimiento y los otros elementos de sintaxis para el corte de video que se está codificando.

La unidad de cuantificación inversa 58 y el módulo de transformación inversa 60 aplican la cuantificación inversa y la transformación inversa, respectivamente, para reconstruir el bloque residual en el dominio de píxeles para su uso posterior como bloque de referencia de una imagen de referencia. La unidad de compensación de movimiento 44 puede calcular un bloque de referencia añadiendo el bloque residual a un bloque predictivo de una de las imágenes de referencia dentro de una de las listas de imágenes de referencia. La unidad de compensación de movimiento 44

60 también puede aplicar uno o más filtros de interpolación al bloque residual reconstruido para calcular valores de píxeles sub enteros para utilizarlos en la estimación de movimiento. El sumador 62 añade el bloque residual reconstruido al bloque de predicción compensado de movimiento producido por la unidad de compensación de movimiento 44 para producir un bloque de referencia para almacenarlo en el búfer de imágenes decodificadas 64. El bloque de referencia



puede ser utilizado por la unidad de estimación de movimiento 42 y la unidad de compensación de movimiento 44 como un bloque de referencia para inter pronosticar un bloque en una imagen o fotograma de video posterior.

De acuerdo con esta divulgación, el módulo de predicción 41 representa una unidad de ejemplo para realizar las funciones de ejemplo que se han descrito anteriormente. Por ejemplo, el módulo de predicción 41 puede determinar qué imágenes de referencia pertenecen a un conjunto de imágenes de referencia y hacer que el codificador de video (video encoder) 20 codifique información indicativa de las imágenes de referencia que pertenecen al conjunto de imágenes de referencia. Además, durante el proceso de reconstrucción (por ejemplo, el proceso utilizado para reconstruir una imagen para utilizarla como imagen de referencia y almacenarla en el búfer de imágenes decodificadas 64), el módulo de predicción 41 puede construir la pluralidad de subconjuntos de imágenes de referencia cada uno de los cuales identifica una o más de las imágenes de referencia. El módulo de predicción 41 también puede derivar el conjunto de imágenes de referencia a partir de la pluralidad de subconjuntos de imágenes de referencia construida. Además, el módulo de predicción 41 puede implementar cualquiera del uno o más de los conjuntos de pseudo código de ejemplo que se han descrito anteriormente para implementar una o más técnicas de ejemplo que se describen en esta divulgación.

En algunos ejemplos, el módulo de predicción 41 puede construir las listas iniciales de imágenes de referencia de la manera que se ha descrito anteriormente. En algunos ejemplos, no es necesario reordenar las imágenes a incluir en las listas iniciales de imágenes de referencia. Además, el módulo de predicción 41 puede construir las listas de imágenes de referencia de tal manera que no haya entradas no completadas en las listas de imágenes de referencia. En algunos ejemplos, el módulo de predicción 41 también puede modificar las listas iniciales de imágenes de referencia de la manera que se ha descrito anteriormente para construir la lista de imágenes de referencia modificada. Además, en algunos ejemplos, el módulo de predicción 41 puede implementar la eliminación de imágenes decodificadas del búfer DPB 64 de la manera que se ha descrito anteriormente. Además, en algunos ejemplos, el módulo de predicción 41 puede estar configurado para determinar qué imágenes de referencia de largo plazo pertenecen al conjunto de imágenes de referencia para la imagen actual de la manera que se ha descrito anteriormente.

En otros ejemplos, una unidad distinta del módulo de predicción 41 puede implementar los ejemplos que se han descrito anteriormente. En algunos otros ejemplos, el módulo de predicción 41 junto con una o más unidades de codificador de video (video encoder) 20 puede implementar los ejemplos que se han descrito anteriormente. En algunos otros ejemplos, un procesador o unidad de codificador de video (video encoder) 20 (que no se muestra en la figura 3) puede, en solitario o junto con otras unidades de codificador de video (video encoder) 20, implementar los ejemplos que se han descrito anteriormente.

La figura 4 es un diagrama de bloques que ilustra un ejemplo de decodificador de video 30 que puede implementar las técnicas que se describen en esta divulgación. En el ejemplo de la figura 4, el decodificador de video 30 incluye una unidad de decodificación de entropía 80, un módulo de predicción 81, una unidad de cuantificación inversa 86, una unidad de transformación inversa 88, un sumador 90 y un búfer de imágenes decodificadas (DPB) 92. El módulo de predicción 81 incluye una unidad de compensación de movimiento 82 y un módulo de intra predicción 84. El decodificador de video 30 puede, en algunos ejemplos, realizar una etapa de decodificación generalmente recíproca a la etapa de codificación descrita con respecto al codificador de video (video encoder) 20 de la figura 3.

Durante el proceso de decodificación, el decodificador de video 30 recibe un flujo de bits de video codificado que representa bloques de video de un corte de video codificado y elementos de sintaxis asociados del codificador de video (video encoder) 20. La unidad de decodificación de entropía 80 del decodificador de video 30 decodifica por entropía el flujo de bits para generar coeficientes cuantificados, vectores de movimiento y otros elementos de sintaxis. La unidad de decodificación de entropía 80 suministra los vectores de movimiento y otros elementos de sintaxis al módulo de predicción 81. El decodificador de video 30 puede recibir los elementos de sintaxis a nivel de corte (slice) de video y/o a nivel de bloque de video.

Cuando el corte de video está codificado como un corte intra codificado (I), el módulo de intra predicción 84 del módulo de predicción 81 puede generar datos de predicción para un bloque de video del corte de video actual en base a un modo de intra predicción señalizada y datos de bloques previamente decodificados de la imagen actual. Cuando la imagen de video está codificada como un corte inter-codificado (es decir, B o P), la unidad de compensación de movimiento 82 del módulo de predicción 81 produce bloques predictivos para un bloque de video del corte (slice) de video actual en base a los vectores de movimiento y otros elementos de sintaxis recibidos procedentes de la unidad de decodificación de entropía 80. Los bloques predictivos pueden producirse a partir de una de las imágenes de referencia dentro de una de las listas de imágenes de referencia. El decodificador de video 30 puede construir las listas de fotogramas de referencia, la Lista 0 y la Lista 1, utilizando técnicas de construcción por defecto en base a imágenes de referencia almacenadas en el búfer de imágenes decodificadas 92. En algunos ejemplos, el decodificador de video 30 puede construir la Lista 0 y la Lista 1 a partir de las imágenes de referencia identificadas en el conjunto de imágenes de referencia derivado.

La unidad de compensación de movimiento 82 determina información de predicción para un bloque de video del corte (slice) de video actual analizando los vectores de movimiento y otros elementos de sintaxis, y utiliza la información de predicción para producir los bloques predictivos para el bloque de video actual que se está decodificando. Por ejemplo, la unidad de compensación de movimiento 82 utiliza algunos de los elementos de sintaxis recibidos para determinar

5 un modo de predicción (por ejemplo, intra predicción o inter predicción) que se utiliza para codificar los bloques de video del corte de video, un tipo de corte de inter predicción (por ejemplo, un tipo de corte de B o P), información de construcción para una o más de las listas de imágenes de referencia para el corte, vectores de movimiento para cada bloque de video inter-codificado del corte, estado de inter predicción para cada bloque de video inter-codificado del corte, y otra información para decodificar los bloques de video en el corte de video actual.

10

La unidad de compensación de movimiento 82 también puede realizar interpolación en base a filtros de interpolación. La unidad de compensación de movimiento 82 puede utilizar filtros de interpolación como los que utiliza el codificador de video (video encoder) 20 durante la codificación de los bloques de video para calcular valores interpolados de píxeles sub enteros de los bloques de referencia. En este caso, la unidad de compensación de movimiento 82 puede

15 determinar los filtros de interpolación utilizados por el codificador de video (video encoder) 20 a partir de los elementos de sintaxis recibidos y utilizar los filtros de interpolación para producir bloques predictivos.

La unidad de cuantificación inversa 86 cuantifica inversamente, es decir, de-cuantifica, los coeficientes de transformación cuantificados proporcionados en el flujo de bits y decodificados por la unidad de decodificación de

20 entropía 80. El proceso de cuantificación inversa puede incluir el uso de un parámetro de cuantificación calculado por el codificador de video (video encoder) 20 para cada bloque de video en el corte de video para determinar un grado de cuantificación y, de modo similar, un grado de cuantificación inversa que se debe aplicar. El módulo de transformación inversa 88 aplica una transformación inversa, por ejemplo, una DCT inversa, una transformación entera inversa, o un proceso de transformación inversa conceptualmente similar, a los coeficientes de transformación para

25 producir bloques residuales en el dominio de píxeles.

Después de que el módulo de predicción 81 haya generado el bloque predictivo para el bloque de video actual en base a la inter o intra predicción, el decodificador de video 30 forma un bloque de video decodificado sumando los bloques residuales procedentes del módulo de transformación inversa 88 con los bloques predictivos correspondientes

30 generados por el módulo de predicción 81. El sumador 90 representa el componente o componentes que realizan esta operación de totalización. Si se desea, también se puede aplicar un filtro de desbloqueo para filtrar los bloques decodificados con el fin de eliminar artefactos de bloqueo. Otros filtros de bucle (ya sea en el bucle de codificación o después del bucle de codificación) también se pueden utilizar para suavizar transiciones de píxeles o para mejorar la calidad del video. Los bloques de video decodificados en una imagen determinada se almacenan en el búfer de

35 imágenes decodificadas 92, que almacena imágenes de referencia utilizadas para la posterior compensación de movimiento. El búfer de imágenes decodificadas 92 también almacena video decodificado para su posterior presentación en un dispositivo de visualización, tal como por ejemplo el dispositivo de visualización 32 de la figura 1.

De acuerdo con esta divulgación, el módulo de predicción 81 representa una unidad de ejemplo para realizar las

40 funciones de ejemplo que se han descrito anteriormente. Por ejemplo, el módulo de predicción 81 puede determinar qué imágenes de referencia pertenecen a un conjunto de imágenes de referencia. Además, el módulo de predicción 81 puede construir la pluralidad de subconjuntos de imágenes de referencia cada uno de los cuales identifica una o más de las imágenes de referencia. El módulo de predicción 81 también puede derivar el conjunto de imágenes de referencia a partir de la pluralidad de subconjuntos de imágenes de referencia construida. Además, el módulo de

45 predicción 81 puede implementar uno o más de los conjuntos de pseudo código de ejemplo que se han descrito anteriormente para implementar una o más técnicas de ejemplo que se describen en esta divulgación.

En algunos ejemplos, el módulo de predicción 81 puede construir las listas iniciales de imágenes de referencia de la manera que se ha descrito anteriormente. En algunos ejemplos, no es necesario reordenar las imágenes a incluir en

50 las listas iniciales de imágenes de referencia. Además, el módulo de predicción 81 puede construir las listas de imágenes de referencia de tal manera que no haya entradas no completadas en las listas de imágenes de referencia. En algunos ejemplos, el módulo de predicción 81 también puede modificar las listas iniciales de imágenes de referencia de la manera que se ha descrito anteriormente para construir la lista de imágenes de referencia modificada. Además, en algunos ejemplos, el módulo de predicción 81 puede implementar la eliminación de imágenes decodificadas del

55 búfer DPB 94 de la manera que se ha descrito anteriormente. Además, en algunos ejemplos, el módulo de predicción 81 puede estar configurado para determinar qué imágenes de referencia de largo plazo pertenecen al conjunto de imágenes de referencia para la imagen actual de la manera que se ha descrito anteriormente.

En otros ejemplos, una unidad distinta del módulo de predicción 81 puede implementar los ejemplos que se han

60 descrito anteriormente. En algunos otros ejemplos, el módulo de predicción 81 junto con una o más unidades de decodificador de video 30 puede implementar los ejemplos que se han descrito anteriormente. En algunos otros ejemplos, un procesador o unidad de decodificador de video 30 (que no se muestra en la figura 4) puede, en solitario o junto con otras unidades de decodificador de video 30, implementar los ejemplos que se han descrito anteriormente.

La figura 5 es un diagrama de flujo que ilustra un ejemplo de operación de derivar un conjunto de imágenes de referencia. Para fines de ilustración solamente, el procedimiento de la figura 5 puede ser realizado por un codificador de video (video coder) que corresponde al codificador de video (video encoder) 20 o al decodificador de video 30. Por ejemplo, un codificador de video (video coder) (por ejemplo, el codificador de video (video encoder) 20 o el decodificador de video 30) puede convertir según una codificación (por ejemplo, codificar o decodificar) información indicativa de imágenes de referencia que pertenecen a un conjunto de imágenes de referencia (94). El conjunto de imágenes de referencia puede identificar las imágenes de referencia que se pueden utilizar potencialmente para la inter predicción de una imagen actual y que se pueden utilizar potencialmente para la inter predicción de una o más imágenes posteriores a la imagen actual en un orden de decodificación.

Por ejemplo, cuando el codificador de video (video encoder) 20 realiza la etapa 94, el codificador de video (video encoder) 20 puede codificar valores que indican identificadores para las imágenes de referencia que pertenecen al conjunto de imágenes de referencia. Por ejemplo, el codificador de video (video encoder) 20 puede señalar en el flujo de bits el elemento de sintaxis `pic_order_cnt_lsb` y el elemento de sintaxis `log2_max_pic_order_cnt_lsb_minus4`. Cuando el decodificador de video 30 realiza la etapa 94, a partir del elemento de sintaxis `log2_max_pic_order_cnt_lsb_minus4`, el decodificador de video 30 puede determinar el valor de `MaxPicOrderCntLsb`. El decodificador de video 30 puede entonces determinar los identificadores (por ejemplo, los valores de POC) para las imágenes de referencia que pertenecen al conjunto de imágenes de referencia.

El codificador de video (video coder) puede construir una pluralidad de subconjuntos de imágenes de referencia cada uno de los cuales identifica cero o más de las imágenes de referencia (96). Por ejemplo, el codificador de video (video coder) puede construir los subconjuntos de imágenes de referencia `RefPicSetStCurr0`, `RefPicSetStCurr1`, `RefPicSetStFoll0`, `RefPicSetStFoll1`, `RefPicSetLtCurr` y `RefPicSetLtFoll`. Sin embargo, aspectos de esta divulgación no son tan limitados. En algunos ejemplos, el codificador de video (video coder) puede construir cinco subconjuntos de imágenes de referencia, cuatro de los cuales pueden ser cuatro de los subconjuntos de imágenes de referencia `RefPicSetStCurr0`, `RefPicSetStCurr1`, `RefPicSetStFoll0`, `RefPicSetStFoll1`, `RefPicSetLtCurr` y `RefPicSetLtFoll`, y el quinto puede ser una combinación de dos de los seis subconjuntos de imágenes de referencia restantes (por ejemplo, una combinación de los subconjuntos de imágenes de referencia `RefPicSetFoll0` y `RefPicSetFoll1`).

En algunos ejemplos, el codificador de video (video coder) puede construir al menos dos de los siguientes cuatro subconjuntos de imágenes de referencia. En otros ejemplos, el codificador de video (video coder) puede construir al menos los siguientes cuatro subconjuntos de imágenes de referencia. El primer subconjunto de imágenes de referencia puede identificar imágenes de referencia de corto plazo que son anteriores a la imagen actual en un orden de decodificación y anteriores a la imagen actual en orden de output y que se pueden utilizar potencialmente para la inter predicción de la imagen actual y una o más de las imágenes posteriores a la imagen actual en orden de decodificación. El segundo subconjunto de imágenes de referencia puede identificar imágenes de referencia de corto plazo que son anteriores a la imagen actual en orden de decodificación y posteriores a la imagen actual en orden de output y que se pueden utilizar potencialmente para la inter predicción de la imagen actual y una o más de las imágenes posteriores a la imagen actual en orden de decodificación.

El tercer subconjunto de imágenes de referencia puede identificar imágenes de referencia de largo plazo que son anteriores a la imagen actual en orden de decodificación y que se pueden utilizar potencialmente para la inter predicción de la imagen actual y una o más de las imágenes posteriores a la imagen actual en orden de decodificación. El cuarto subconjunto de imágenes de referencia puede identificar imágenes de referencia de largo plazo que son anteriores a la imagen actual en orden de decodificación y que no se pueden utilizar para la inter predicción de la imagen actual, y que se pueden utilizar potencialmente para la inter predicción de una o más de las imágenes posteriores a la imagen actual en orden de decodificación.

El codificador de video (video coder) puede derivar el conjunto de imágenes de referencia a partir de la pluralidad de subconjuntos de imágenes de referencia (98). Por ejemplo, el codificador de video (video coder) puede ordenar al menos dos de los subconjuntos de imágenes de referencia `RefPicSetStCurr0`, `RefPicSetStCurr1`, `RefPicSetStFoll0`, `RefPicSetStFoll1`, `RefPicSetLtCurr` y `RefPicSetLtFoll` en un orden determinado para obtener el conjunto de imágenes de referencia.

En algunos ejemplos, la ordenación realizada por el código de video puede significar que las imágenes en cada uno de los subconjuntos de imágenes de referencia se pueden identificar secuencialmente dentro del conjunto de imágenes de referencia. En estos ejemplos, el codificador de video (video coder) puede referenciar a las imágenes de referencia del conjunto de imágenes de referencia a través de un valor de índice en el conjunto de imágenes de referencia.

El codificador de video (video coder) puede codificar la imagen actual en base al conjunto de imágenes de referencia derivado (100). Se debe entender que a causa de que el codificador de video (video coder) deriva el conjunto de

- imágenes de referencia a partir de los subconjuntos de imágenes de referencia, se puede considerar que el codificador de video (video coder) convierte según codificación la imagen actual en base a la pluralidad de subconjuntos de imágenes de referencia. Por ejemplo, el codificador de video (video coder) puede construir al menos una de entre una primera lista de imágenes de referencia y una segunda lista de imágenes de referencia en base a la pluralidad de subconjuntos de imágenes de referencia (por ejemplo, a partir del conjunto de imágenes de referencia derivado que se deriva a partir de la pluralidad de subconjuntos de imágenes de referencia). El codificador de video (video coder) puede entonces codificar la imagen actual en base a por lo menos una de entre la primera lista de imágenes de referencia y la segunda lista de imágenes de referencia.
- 10 La figura 6 es un diagrama de flujo que ilustra un ejemplo de operación de construcción de una lista de imágenes de referencia. Para fines de ilustración solamente, el procedimiento de la figura 6 puede ser realizado por un codificador de video (video coder) que corresponde al codificador de video (video encoder) 20 o al decodificador de video 30. De forma similar a la figura 5, el codificador de video (video coder) puede codificar información indicativa de imágenes de referencia (102), y construir una pluralidad de subconjuntos de imágenes de referencia (104).
- 15 El codificador de video (video coder) puede entonces añadir imágenes de referencia de subconjuntos de imágenes de referencia, a una lista inicial de imágenes de referencia, para construir la lista inicial de imágenes de referencia (106). En algunos ejemplos, tanto el codificador de video (video encoder) 20 como el decodificador de video 30 pueden construir la lista inicial de imágenes de referencia. Por ejemplo, el codificador de video (video encoder) 20 puede construir la lista inicial de imágenes de referencia para crear los bloques de video reconstruidos para su almacenamiento en el búfer DPB 64. El decodificador de video 30 puede construir la lista inicial de imágenes de referencia como parte de su proceso de decodificación, y puede implementar una técnica de construcción por defecto en la que el decodificador de video 30 no necesita recibir información procedente del codificador de video (video encoder) 20 sobre la manera de construir la lista inicial de imágenes de referencia.
- 20 En algunos ejemplos, para construir la lista inicial de imágenes de referencia, el codificador de video (video coder) puede añadir imágenes de referencia de un primer subconjunto de la pluralidad de subconjuntos de imágenes de referencia a la lista inicial de imágenes de referencia, seguido de imágenes de referencia de un segundo subconjunto a la lista inicial de imágenes de referencia, y luego seguido de imágenes de referencia de un tercer subconjunto a la lista inicial de imágenes de referencia. El codificador de video (video coder) puede añadir imágenes de referencia de estos subconjuntos de imágenes de referencia mientras el número total de imágenes de referencia listadas o incluidas en la lista inicial de imágenes de referencia no sea superior a un número máximo de entradas permisibles en la lista inicial de imágenes de referencia. Por ejemplo, si en cualquier momento durante la agregación de las imágenes de referencia a la lista de imágenes de referencia, el número de entradas en la lista inicial de imágenes de referencia se iguala con el número máximo de entradas permisibles en la lista de referencia inicial, el codificador de video (video coder) puede dejar de añadir imágenes adicionales a la lista inicial de imágenes de referencia.
- 25 El codificador de video (video coder) puede construir de forma similar otra lista inicial de imágenes de referencia, tal como en ejemplos en los que el bloque de video de la imagen actual es bi-pronosticado. En este ejemplo, para construir esta otra lista inicial de imágenes de referencia, el codificador de video (video coder) puede añadir imágenes de referencia del segundo subconjunto a la otra lista inicial de imágenes de referencia, seguido de imágenes de referencia del primer subconjunto a la otra lista inicial de imágenes de referencia, y seguido entonces de imágenes de referencia del tercer subconjunto a la otra lista inicial de imágenes de referencia, siempre y cuando el número total de entradas en esta otra lista inicial de imágenes de referencia no sea superior al número de entradas permisibles. En estos ejemplos, el primer subconjunto puede ser el subconjunto de imágenes de referencia RefPicSetStCurr0, el segundo puede ser el subconjunto de imágenes de referencia RefPicSetStCurr1 y el tercer subconjunto puede ser el subconjunto de imágenes de referencia RefPicSetLtCurr.
- 30 En algún ejemplo, para añadir imágenes de referencia identificadas en los subconjuntos de imágenes de referencia RefPicSetStCurr0, RefPicSetStCurr1 y RefPicSetLtCurr, el codificador de video (video coder) puede codificar (por ejemplo, codificar o decodificar) elementos de sintaxis a partir de los cuales el codificador de video (video coder) puede determinar el número de imágenes de referencia en cada uno de estos subconjuntos de imágenes de referencia. Por ejemplo, el codificador de video (video coder) puede codificar un elemento de sintaxis num\_short\_term\_curr0 y un elemento de sintaxis num\_short\_term\_curr1. El elemento de sintaxis num\_short\_term\_curr0 y el elemento de sintaxis num\_short\_term\_curr1 pueden indicar el número de imágenes de referencia identificadas en el subconjunto de imágenes de referencia RefPicSetStCurr0 y en el subconjunto de imágenes de referencia RefPicSetStCurr1, respectivamente.
- 35 El codificador de video (video coder) también puede codificar un elemento de sintaxis num\_long\_term\_pps\_curr, y un elemento de sintaxis num\_long\_term\_add\_curr. El elemento de sintaxis num\_long\_term\_pps\_curr puede indicar un número de imágenes de referencia de largo plazo cuya identificación está incluida en un conjunto de parámetros de imagen (PPS), y el elemento de sintaxis num\_long\_term\_add\_curr puede indicar un número de imágenes de referencia de largo plazo cuya información de identificación no está incluida en el PPS. En este ejemplo, estas imágenes de

referencia de largo plazo se pueden utilizar potencialmente para la inter predicción de la imagen actual y se pueden utilizar potencialmente para la inter predicción de una o más imágenes posteriores a la imagen actual en orden de decodificación.

5 El codificador de video (video coder) puede determinar el número de imágenes de referencia en el subconjunto de imágenes de referencia RefPicSetLtCurr en base al elemento de sintaxis num\_long\_term\_pps\_curr y al elemento de sintaxis num\_long\_term\_add\_curr. Por ejemplo, el codificador de video (video coder) puede sumar los valores del elemento de sintaxis num\_long\_term\_pps\_curr y del elemento de sintaxis num\_long\_term\_add\_curr para determinar el número de imágenes de referencia en el subconjunto de imágenes de referencia RefPicSetLtCurr.

10

El codificador de video (video coder) puede codificar la imagen actual en base a la lista o listas de imágenes de referencia (108). Por ejemplo, el codificador de video (video coder) puede construir al menos una de entre una primera lista de imágenes de referencia y una segunda lista de imágenes de referencia en base al conjunto de imágenes de referencia derivado. El codificador de video (video coder) puede entonces codificar la imagen actual en base a por lo menos una de entre la primera lista de imágenes de referencia y la segunda lista de imágenes de referencia.

15

La figura 7 es un diagrama de flujo que ilustra otro ejemplo de operación de construcción de una lista de imágenes de referencia. Para fines de ilustración solamente, el procedimiento de la figura 7 puede ser realizado por un codificador de video (video coder) que corresponde al codificador de video (video encoder) 20 o al decodificador de video 30. De forma similar a la figura 5, el codificador de video (video coder) puede codificar información indicativa de imágenes de referencia (110), y construir una pluralidad de subconjuntos de imágenes de referencia (112). El codificador de video (video coder) puede añadir las imágenes de referencia a un primer conjunto de entradas de una lista de imágenes de referencia (114). Por ejemplo, el número de entradas en la lista de imágenes de referencia puede ser igual al número máximo de entradas permisibles definido por el elemento de sintaxis num\_ref\_idx\_l0\_active\_minus1 o el elemento de sintaxis num\_ref\_idx\_l1\_active\_minus1.

20

25

En este ejemplo, el codificador de video (video coder) puede listar en el (por ejemplo, añadir al) primer conjunto de entradas las imágenes de referencia identificadas en los subconjuntos de imágenes de referencia RefPicSetStCurr0, RefPicSetStCurr1 y RefPicSetLtCurr. Por ejemplo, para la Lista 0, el codificador de video (video coder) puede añadir al primer conjunto de entradas en la Lista 0, en orden, las imágenes de referencia identificadas en los subconjuntos de imágenes de referencia RefPicSetStCurr0, RefPicSetStCurr1 y RefPicSetLtCurr. Para la Lista 1, el codificador de video (video coder) puede añadir al primer conjunto de entradas en la Lista 1, en orden, las imágenes de referencia identificadas en los subconjuntos de imágenes de referencia RefPicSetStCurr1, RefPicSetStCurr0 y RefPicSetLtCurr.

30

35

El codificador de video (video coder) puede entonces determinar si el número de entradas en la lista de imágenes de referencia es igual a un número máximo de entradas permisibles en la lista de imágenes de referencia (116). Si el número de entradas en la lista de imágenes de referencia no es inferior al número máximo de entradas permisibles ("NO" de 116), el codificador de video (video coder) puede codificar la imagen actual en base a la lista de imágenes de referencia (118).

40

De lo contrario, si el número de entradas en la lista de imágenes de referencia es inferior al número máximo de entradas permisibles ("SÍ" de 116), el codificador de video (video coder) puede volver a listar (por ejemplo, volver a identificar o volver a añadir) una o más imágenes de referencia de al menos uno de entre la pluralidad de subconjuntos de imágenes de referencia en entradas de la lista de imágenes de referencia que son subsiguientes respecto al primer grupo de entradas (120). Por ejemplo, el codificador de video (video coder) puede añadir una o más imágenes de referencia identificadas en el subconjunto de imágenes de referencia RefPicSetStCurr0 a la Lista 0 en entradas subsiguientes respecto al primer conjunto de entradas en la Lista 0, o una o más imágenes de referencia identificadas en el subconjunto de imágenes de referencia RefPicSetStCurr1 a la Lista 1 en entradas subsiguientes respecto al primer conjunto de entradas en la Lista 1. De esta forma, el codificador de video (video coder) puede identificar al menos una imagen de referencia del primer subconjunto de imágenes de referencia en más de una entrada de la lista de imágenes de referencia.

45

50

El codificador de video (video coder) puede entonces determinar si el número de entradas en la lista de imágenes de referencia es igual a un número máximo de entradas permisibles en la lista de imágenes de referencia (122). Si el número de entradas en la lista de imágenes de referencia no es inferior al número máximo de entradas permisibles ("NO" de 122), el codificador de video (video coder) puede codificar la imagen actual en base a la lista de imágenes de referencia (124).

55

De lo contrario, si el número de entradas en la lista de imágenes de referencia es inferior al número máximo de entradas permisibles ("SÍ" de 122), el codificador de video (video coder) puede volver a listar (por ejemplo, volver a identificar o volver a añadir) una o más imágenes de referencia de al menos uno de entre la pluralidad de subconjuntos de imágenes de referencia en entradas de la lista de imágenes de referencia que son subsiguientes respecto al primer grupo de entradas (120). Por ejemplo, en esta situación, el codificador de video (video coder) puede volver a añadir

60

imágenes de referencia adicionales identificadas en el primer subconjunto de imágenes de referencia. Si el codificador de video (video coder) ya ha vuelto a añadir todas las imágenes de referencia del primer subconjunto de imágenes de referencia, el codificador de video (video coder) puede volver a añadir una o más imágenes de referencia del segundo subconjunto de imágenes de referencia (por ejemplo, el subconjunto de imágenes de referencia RefPicSetStCurr0 para la Lista 0, o el subconjunto de imágenes de referencia RefPicSetStCurr1 para la Lista 1). Este proceso se puede repetir hasta que el número de entradas en la lista de imágenes de referencia no sea inferior al número máximo de entradas permisibles ("NO" de 122).

De esta manera, cuando el número de entradas en la lista de imágenes de referencia no es igual al número máximo de entradas permisibles en la lista de imágenes de referencia, el codificador de video (video coder) puede volver a listar (por ejemplo, volver a identificar o volver a añadir) repetidamente una o más imágenes de referencia de al menos uno de entre la pluralidad de subconjuntos de imágenes de referencia en entradas de la lista de imágenes de referencia que son subsiguientes respecto al primer conjunto de entradas hasta que el número de entradas en la lista de imágenes de referencia sea igual al número máximo de entradas permisibles en la lista de imágenes de referencia. Esto puede dar lugar a que el codificador de video (video coder) liste o incluya imágenes de referencia en entradas de la lista de imágenes de referencia de forma que cada entrada de la lista de imágenes de referencia identifique una de las imágenes de referencia y que al menos dos entradas de la lista de imágenes de referencia identifiquen una misma imagen de referencia de las imágenes de referencia.

La figura 8 es un diagrama de flujo que ilustra un ejemplo de operación de modificación de una lista inicial de imágenes de referencia. Para fines de ilustración solamente, el procedimiento de la figura 8 puede ser realizado por un codificador de video (video coder) que corresponde al codificador de video (video encoder) 20 o al decodificador de video 30. De forma similar a la figura 6, el codificador de video (video coder) puede codificar información indicativa de imágenes de referencia (126), y construir una pluralidad de subconjuntos de imágenes de referencia (128). La pluralidad de subconjuntos de imágenes de referencia puede incluir los subconjuntos de imágenes de referencia RefPicSetStCurr0, RefPicSetStCurr1 y RefPicSetLtCurr. El codificador de video (video coder) puede construir una lista inicial de imágenes de referencia de la manera que se ha descrito anteriormente en base a los subconjuntos de imágenes de referencia RefPicSetStCurr0, RefPicSetStCurr1 y RefPicSetLtCurr (130).

El codificador de video (video coder) puede determinar si es necesario modificar la lista de imágenes de referencia (132). Por ejemplo, el codificador de video (video coder) puede codificar elementos de sintaxis tales como ref\_pic\_list\_modification\_flag\_I0 y ref\_pic\_list\_modification\_flag\_I1 que indican si la Lista 0 inicial o la Lista 1 inicial necesita ser modificada. Si no es necesario modificar la lista inicial de imágenes de referencia ("NO" de 132), el codificador de video (video coder) puede codificar la imagen actual en base a la lista inicial de imágenes de referencia (134).

Si se necesita alguna modificación ("SÍ" de 132), el codificador de video (video coder) puede identificar una imagen de referencia en al menos uno de entre la pluralidad de subconjuntos de imágenes de referencia construidos (136). Por ejemplo, el codificador de video (video coder) puede codificar el elemento de sintaxis modification\_of\_ref\_pic\_idc. El valor del elemento de sintaxis modification\_of\_ref\_pic\_idc puede indicar qué subconjunto de imágenes de referencia debe utilizar el codificador de video (video coder) para identificar la imagen de referencia. El codificador de video (video coder) también puede codificar el elemento de sintaxis ref\_pic\_set\_idx que indica el índice en el subconjunto de imágenes de referencia que el codificador de video (video coder) debe utilizar para identificar la imagen de referencia.

El codificador de video (video coder) puede listar (por ejemplo, añadir o identificar) la imagen de referencia identificada en la lista inicial de imágenes de referencia en una entrada actual para construir la lista de imágenes de referencia modificada (138). La entrada actual puede ser inicialmente la entrada en la lista inicial de imágenes de referencia definida por el índice 0. Para cada instancia del elemento de sintaxis modification\_of\_ref\_pic\_idc en el flujo de bits codificado en el que el valor del elemento de sintaxis modification\_of\_ref\_pic\_idc no es igual a 3, el codificador de video (video coder) puede incrementar el valor de la entrada inicial en uno (por ejemplo, el siguiente valor de la entrada estará definido por el índice 1). El codificador de video (video coder) puede codificar la imagen actual en base a la lista de imágenes de referencia modificada (140).

La figura 9 es un diagrama de flujo que ilustra un ejemplo de operación de eliminar imágenes decodificadas. Para fines de ilustración solamente, el procedimiento de la figura 8 puede ser realizado por un codificador de video (video coder) que corresponde al codificador de video (video encoder) 20 o al decodificador de video 30. De forma similar a la figura 5, un codificador de video (video coder) (por ejemplo, el codificador de video (video encoder) 20 o el decodificador de video 30) puede codificar (por ejemplo, codificar o decodificar) información indicativa de imágenes de referencia de un conjunto de imágenes de referencia (142), y puede derivar el conjunto de imágenes de referencia a partir de la información convertida según codificación (por ejemplo, codificada o decodificada) (144).

El codificador de video (video coder) puede determinar si una imagen decodificada que está almacenada en un búfer de imágenes decodificadas (DPB) no es necesaria para el output (es decir, ya ha sido suministrada como output o no

está destinada a ser suministrada como output) y no está identificada en el conjunto de imágenes de referencia (146). Si la imagen decodificada es necesaria para el output o se identifica en el conjunto de imágenes de referencia ("NO" de 146), el codificador de video (video coder) puede no eliminar la imagen decodificada del búfer DPB y puede conservar la imagen decodificada almacenada en el búfer DPB (148).

5 Cuando se ha suministrado la imagen decodificada como output y no está identificada en el conjunto de imágenes de referencia ("SI" de 146), el codificador de video (video coder) puede eliminar la imagen decodificada del búfer DPB (150). El codificador de video (video coder) puede entonces codificar la imagen actual después de eliminar la imagen decodificada (152).

10 Según se ha descrito anteriormente, el codificador de video (video coder) puede construir una lista de imágenes de referencia en base al conjunto de imágenes de referencia para codificar la imagen actual. En algunos ejemplos, el codificador de video (video coder) puede eliminar la imagen decodificada del búfer DPB después de construir la lista de imágenes de referencia. Además, para la imagen decodificada suministrada como output, el codificador de video (video coder) puede determinar el momento de output de la imagen decodificada, y puede suministrar como output la imagen decodificada en base al momento determinado y antes de la conversión según codificación de la imagen actual.

En algunos ejemplos, el codificador de video (video coder) puede almacenar la imagen convertida según codificación (por ejemplo, decodificada) en el búfer DPB. En algunos ejemplos, el codificador de video (video coder) puede determinar que el búfer DPB está lleno antes del almacenamiento. En estos ejemplos, el codificador de video (video coder) puede seleccionar una imagen decodificada que actualmente está almacenada en el búfer DPB que está marcada como "necesaria para el output" y que tiene un valor de contador del orden de las imágenes (POC) que es el menor de todas las imágenes decodificadas almacenadas en el búfer DPB. El codificador de video (video coder) puede suministrar como output la imagen seleccionada. Además, el codificador de video (video coder) puede determinar que la imagen suministrada como output no está incluida en el conjunto de imágenes de referencia de la imagen actual. En este caso, el codificador de video (video coder) puede vaciar un búfer dentro del búfer DPB que ha almacenado la imagen suministrada como output, y puede almacenar la imagen convertida según codificación en ese búfer dentro del búfer DPB.

30 La figura 10 es un diagrama de flujo que ilustra una operación de ejemplo para determinar qué imágenes de referencia de largo plazo pertenecen al conjunto de imágenes de referencia de una imagen actual. Para fines de ilustración solamente, el procedimiento de la figura 10 puede ser realizado por un codificador de video (video coder) que corresponde al codificador de video (video encoder) 20 o al decodificador de video 30.

35 El codificador de video (video coder) (por ejemplo, el codificador de video (video encoder) 20 o el decodificador de video 30) puede codificar elementos de sintaxis que indican imágenes de referencia de largo plazo candidatas identificadas en un conjunto de parámetros (154). En algunos ejemplos, una o más de las imágenes de referencia de largo plazo candidatas pertenecen a un conjunto de imágenes de referencia de la imagen actual. De acuerdo con las técnicas que se describen en esta divulgación, el conjunto de parámetros puede ser el conjunto de parámetros de secuencia o el conjunto de parámetros de imagen. En algunos ejemplos, para codificar los elementos de sintaxis que indican imágenes de referencia de largo plazo candidatas, el codificador de video (video coder) puede codificar una lista de valores de POC para las imágenes de referencia de largo plazo candidatas en el conjunto de parámetros, y codificar un valor de índice para la lista en la cabecera de corte (slice) de la imagen actual.

45 El codificador de video (video coder) puede codificar elementos de sintaxis que indican qué imágenes de referencia de largo plazo candidatas, identificadas en el conjunto de parámetros, pertenecen al conjunto de imágenes de referencia de la imagen actual (156). Por ejemplo, el codificador de video (video coder) puede codificar en la cabecera de corte (slice) de la imagen actual los elementos de sintaxis que indican qué imágenes de referencia de largo plazo candidatas pertenecen al conjunto de imágenes de referencia.

En algunos ejemplos, no todas las imágenes de referencia de largo plazo que pertenecen al conjunto de imágenes de referencia son incluidas en las imágenes de referencia de largo plazo candidatas. En estos ejemplos, el codificador de video (video coder) puede codificar más elementos de sintaxis que indican imágenes de referencia de largo plazo que pertenecen al conjunto de imágenes de referencia (158). El codificador de video (video coder) puede entonces construir al menos uno de entre la pluralidad de subconjuntos de imágenes de referencia en función de la indicación de qué imágenes de referencia de largo plazo candidatas pertenecen al conjunto de imágenes de referencia de la imagen actual (160).

60 En los ejemplos anteriores se describen diversos ejemplos, junto con posibles alternativas a los ejemplos. A continuación se describen algunas alternativas adicionales a los ejemplos, que se pueden considerar como ejemplos alternativos a uno o más de los ejemplos que se han descrito anteriormente. Además, estos ejemplos alternativos se

pueden utilizar junto con los ejemplos que se han descrito anteriormente o por separado de los ejemplos que se han descrito anteriormente.

Por ejemplo, lo que se ha descrito anteriormente son ejemplos alternativos para el concepto de conjunto de imágenes de referencia, ejemplos alternativos para la señalización de conjuntos de imágenes de referencia en los conjuntos de parámetros, ejemplos alternativos de subconjuntos del conjunto de imágenes de referencia, y ejemplos alternativos para el marcado de imágenes de referencia. A continuación se ofrecen otros ejemplos alternativos.

Por ejemplo, para la cabecera de la unidad de NAL, en los ejemplos anteriores, la sintaxis de la cabecera de la unidad de NAL puede incluir los elementos de sintaxis `nal_ref_idc`, `temporal_id` y `output_flag`. En un ejemplo alternativo, `nal_ref_idc` (2 bits) se sustituye por `nal_ref_flag` (1 bit). En este ejemplo, `nal_ref_flag` igual a 1 tiene la semántica de que `nal_ref_idc` es superior a 0, y `nal_ref_flag` igual a 0 tiene la misma semántica que `nal_ref_idc` igual a 0. En un ejemplo alternativo, se elimina `nal_ref_idc` (2 bits). La definición de imagen de referencia puede cambiarse a: Una imagen que contiene muestras que se pueden utilizar para la inter predicción en el proceso de decodificación de imágenes subsiguientes en un orden de decodificación. En un ejemplo alternativo, `temporal_id` no se encuentra presente en la sintaxis de la cabecera de la unidad de NAL, y se deriva que el valor es el mismo para todas las imágenes. En un ejemplo alternativo, `output_flag` no se encuentra presente en la sintaxis de la cabecera de la unidad de NAL, y se deriva que el valor es igual a 1 para todas las imágenes.

Para la señalización y el cálculo del contador del orden de las imágenes, las técnicas anteriores pueden utilizar una forma de señalización y cálculo del contador del orden de las imágenes que puede ser similar al contador del orden de las imágenes de tipo 0 en AVC. Se pueden aplicar dos formas alternativas para la señalización y el cálculo del contador del orden de las imágenes, las mismas que los tipos de contador del orden de las imágenes 1 y 2 en AVC, respectivamente. También se puede aplicar cualquier combinación de dos de las tres maneras, o la combinación de las tres maneras.

Para la identificación de imágenes, las técnicas anteriores pueden utilizar el valor de contador del orden de las imágenes (POC) para la identificación de imágenes. En un ejemplo alternativo, se utiliza una referencia temporal (TR: temporal reference) como identificación de imágenes. Una manera de definir la referencia temporal TR es la misma que el POC cuando el POC está restringido de tal manera que la diferencia de POC entre dos imágenes cualesquiera es proporcional al momento de presentación o a la diferencia de momentos de muestreo. En un ejemplo alternativo, se puede utilizar como identificación de imagen un `frame_num` que indica el orden de decodificación (mientras que POC indica el orden de output o de visualización), o una variable (por ejemplo, denominada `UnWrappedFrameNum`) que puede ser cualquier valor de entre unos valores enteros sin signo de 32 bits, derivado a partir de `frame_num`. Básicamente, `UnWrappedFrameNum` puede ser la versión sin reinicio de `frame_num`. Por ejemplo, si `frame_num` está representado por 8 bits, el valor máximo de `frame_num` es igual a 255. El siguiente valor posterior a 255 para `frame_num` es igual a 0. El valor de `UnWrappedFrameNum` continúa aumentando en las posiciones en las que `frame_num` vuelve al valor 0.

Para la señalización de imágenes de referencia de largo plazo en conjuntos de parámetros, en los ejemplos anteriores, el codificador de video (video encoder) puede señalar una lista de información de identificación absoluta de imágenes de referencia de largo plazo en el conjunto de parámetros de imagen, puede señalar información sobre la longitud, en bits, de la información de identificación absoluta de imágenes de referencia de largo plazo, y se puede referenciar el índice de la lista en la cabecera de corte (slice), con lo que se reduce la sobrecarga de señalización. En un ejemplo alternativo, se puede señalar una lista de información de identificación diferencial de imágenes de referencia de largo plazo en el conjunto de parámetros de secuencia y se puede referenciar el índice de la lista en la cabecera de corte (slice), para reducir con ello la sobrecarga de señalización. En un ejemplo alternativo, se puede señalar información sobre la longitud, en bits, de la información de identificación absoluta de imágenes de referencia de largo plazo, pero la longitud se puede considerar como una constante, por ejemplo, 32. En diversos ejemplos, se aplica una combinación de cualquiera de los ejemplos anteriores.

Para la señalización de imágenes de referencia de corto plazo, en los ejemplos anteriores, para imágenes de referencia de corto plazo en el conjunto de imágenes de referencia de una imagen convertida según codificación, o bien todas se señalizan en el conjunto de parámetros de imagen referenciado, o bien todas se señalizan en la cabecera de corte (slice). En un ejemplo alternativo, para imágenes de referencia de corto plazo en el conjunto de imágenes de referencia de una imagen convertida según codificación, algunas son señalizadas en el conjunto de parámetros de imagen referenciado, y otras son señalizadas en la cabecera de corte (slice).

Para la inicialización de la lista de imágenes de referencia, en un ejemplo alternativo, en el proceso de inicialización de la lista de imágenes de referencia, en primer lugar se pueden añadir imágenes de referencia de corto plazo a una lista y, en segundo lugar, opcionalmente, se pueden añadir imágenes de referencia de largo plazo a la lista. Después de esto, si el número de entradas en la lista de imágenes de referencia (`RefPicList0` o `RefPicList1`) sigue siendo inferior al valor de `num_ref_idx_lx_active_minus1 + 1` (en el que `lx` es 0 o 1), se pueden marcar las entradas restantes como



- “sin imagen de referencia”. En un ejemplo alternativo, también puede ser posible si la lista de imágenes de referencia todavía no está completa después de añadir las entradas a RefPicSetStCurr0 y RefPicSetStCurr1 y posiblemente las imágenes de referencia de largo plazo, pueden añadirse las imágenes en RefPicSetStFoll0 y/o RefPicSetStFoll1. En un ejemplo alternativo, el proceso de inicialización de la lista de imágenes de referencia puede añadir sólo imágenes
- 5 de referencia de corto plazo a una lista de imágenes de referencia. En dicho caso, una imagen de referencia de largo plazo sólo puede ser añadida a una lista de imágenes de referencia utilizando comandos de modificación de la lista de imágenes de referencia según se indica en la tabla de sintaxis de modificación de la lista de imágenes de referencia (RPLM: Reference Picture List Modification).
- 10 Para la modificación de la lista de imágenes de referencia, en un ejemplo alternativo, la modificación de la lista de imágenes de referencia también puede señalar el ref\_pic\_set\_idx de manera diferencial, en la que se utiliza el índice anterior como predictor del índice actual. En este ejemplo, diferentes valores de modification\_of\_ref\_pic\_idc pueden corresponder a diferentes categorías de indexación (RefPicSetStCurrx para RefPicListx, RefPicSetStCurrx para RefPicList(1-x) o RefPicSetLtCurr), cada una de las cuales mantiene un predictor diferente. Un predictor puede ser
- 15 actualizado una vez que un elemento de sintaxis perteneciente a la misma categoría ha sido analizado. En un ejemplo alternativo, las modificaciones de la lista de imágenes de referencia se pueden basar en la diferencia de números de imagen. En un ejemplo alternativo, las modificaciones de la lista de imágenes de referencia se pueden basar en la diferencia de valores de POC.
- 20 Para operaciones de búfer de imágenes decodificadas (DPB), en el ejemplo anterior, después de decodificar la imagen actual y antes de analizar una cabecera de corte (slice) de la siguiente imagen codificada en orden de decodificación, se almacena la imagen decodificada actual en el búfer DPB. En un ejemplo alternativo, después de decodificar la imagen actual y antes de analizar una cabecera de corte (slice) de la siguiente imagen codificada en orden de decodificación, se almacena la imagen decodificada actual en una memoria temporal (no en el búfer DPB). Se
- 25 almacena en el búfer DPB después de analizar una cabecera de corte (slice) de la siguiente imagen codificada en orden de decodificación y construcción del conjunto de imágenes de referencia de esa imagen, si todavía se necesita como referencia o como output. En este momento, si no se necesita ni como referencia ni como output, la imagen decodificada puede ser simplemente descartada (del búfer temporal).
- 30 Además, en los ejemplos anteriores, la eliminación de una imagen decodificada del búfer DPB se produce inmediatamente después de analizar una cabecera de corte (slice) de la imagen actual y antes de decodificar cualquier corte de la imagen actual. En un ejemplo alternativo, el marcado, si existe, y la eliminación de una imagen decodificada del búfer DPB se produce después de que la imagen actual se haya decodificado completamente.
- 35 En los ejemplos anteriores, los subconjuntos de imágenes de referencia RefPicSetStCurr0 y RefPicSetStCurr1 del conjunto de imágenes de referencia para la imagen actual se derivan para todas las imágenes decodificadas. Sin embargo, esto puede no ser necesario para la intra imagen. Para la derivación de conjuntos de imágenes de referencia intra imagen, en un ejemplo alternativo, para una imagen no IDR que está intra codificada (es decir, todas las cortes de la imagen codificada son cortes I), RefPicSetStCurr0 y RefPicSetStCurr1 no son derivados, ya que incluso si no
- 40 están vacíos después de haber sido derivados, no son necesarias para la decodificación de la imagen codificada. Permitiendo RefPicSetStCurr0 o RefPicSetStCurr1 no vacíos para una intra imagen no IDR puede permitir compartir una estructura de sintaxis short\_term\_ref\_pic\_set( ) de instancia de una o más imágenes inter codificadas, para las cuales RefPicSetStCurr0 y RefPicSetStCurr1 no pueden estar ambos vacíos.
- 45 Para la detección de pérdidas, pueden ser posibles los siguientes procedimientos para la detección de la pérdida de una imagen de referencia o la detección temprana de si es posible decodificar correctamente la imagen actual. En diversos ejemplos, después de la derivación del conjunto de imágenes de referencia, el decodificador de video 30 (por ejemplo, el lado del decodificador) puede comprobar la presencia de imágenes de referencia incluidas en RefPicSetStCurr0, RefPicSetStCurr1 y RefPicSetLtCurr. Si alguna de las imágenes de referencia incluidas en
- 50 RefPicSetStCurr0, RefPicSetStCurr1 y RefPicSetLtCurr no se encuentra presente en el búfer DPB, el lado del decodificador puede inferir que esa imagen de referencia se ha perdido y que es probable que la imagen actual no se decodifique correctamente, y puede tomar alguna medida para mejorar la situación, por ejemplo, notificando al lado del codificador (por ejemplo, el codificador de video (video encoder) 20) de la pérdida o pérdidas de alguna imagen, y el codificador puede retransmitir la imagen o imágenes de referencia perdidas o codificar las imágenes subsiguientes
- 55 utilizando únicamente imágenes de referencia que se sabe que son correctas en el lado del decodificador para referencia en la inter predicción.
- En diversos ejemplos, después de la derivación del conjunto de imágenes de referencia, el lado del decodificador puede comprobar la presencia de imágenes de referencia incluidas en RefPicSetStFoll0, RefPicSetStFoll1 y
- 60 RefPicSetLtFoll. Si alguna de las imágenes de referencia incluidas en RefPicSetStFoll0, RefPicSetStFoll1 y RefPicSetLtFoll no se encuentra presente en el búfer DPB, el lado del decodificador puede inferir que se ha perdido la imagen de referencia y que algunas de las imágenes subsiguientes en orden de decodificación probablemente no se decodificarán correctamente a menos que se tomen algunas medidas y se pueda tomar alguna medida para

remediar la situación, por ejemplo, notificando al lado del codificador de la pérdida o pérdidas de imagen o imágenes, y el codificador puede retransmitir la imagen o imágenes de referencia perdidas o codificar la imagen o imágenes subsiguientes utilizando sólo las imágenes de referencia que se conocen correctamente en el lado del decodificador para referencia en la inter predicción.

5

Para la composición del conjunto de imágenes de referencia en el lado del codificador (por ejemplo, el codificador de video (video encoder) 20), con los ejemplos anteriores, pueden ser posibles los siguientes procedimientos diferentes para la composición del conjunto de imágenes de referencia en el lado del codificador. Por ejemplo, en diversos ejemplos, el codificador compone las estructuras de sintaxis relacionadas con el conjunto de imágenes de referencia, de forma que después de la derivación del conjunto de imágenes de referencia en el lado del decodificador para la imagen actual: (1) RefPicSetStCurr0 incluye y sólo incluye la información de identificación de todas las imágenes de referencia de corto plazo que tienen un orden de output anterior al de la imagen actual y que se utilizan como referencia en la inter predicción de la imagen actual, (2) RefPicSetStCurr1 incluye y sólo incluye la información de identificación de todas las imágenes de referencia de corto plazo que tienen un orden de output posterior al de la imagen actual y que se utilizan como referencia en la inter predicción de la imagen actual, y (3) RefPicSetLtCurr incluye y sólo incluye la información de identificación de todas las imágenes de referencia de largo plazo que se utilizan como referencia en la inter predicción de la imagen actual.

En diversos ejemplos, el codificador (por ejemplo, el codificador de video (video encoder) 20) puede componer las estructuras de sintaxis relacionadas con el conjunto de imágenes de referencia de forma que, después de la derivación del conjunto de imágenes de referencia en el lado del decodificador para la imagen actual: (1) RefPicSetStCurr0 incluye y sólo incluye la información de identificación de 1) todas las imágenes de referencia de referencia de corto plazo que tienen un orden de output anterior al de la imagen actual y que se utilizan como referencia en la inter predicción de la imagen actual, así como 2) una o más imágenes de referencia de corto plazo que tienen un orden de output anterior al de la imagen actual y que no se utilizan como referencia en la inter predicción de la imagen actual, 2) RefPicSetStCurr1 incluye y sólo incluye la información de identificación de 1) todas las imágenes de referencia de corto plazo que tienen un orden de output posterior al de la imagen actual y que se utilizan como referencia en la inter predicción de la imagen actual, así como 2) una o más imágenes de referencia de corto plazo que tienen un orden de output posterior al de la imagen actual y que no se utilizan como referencia en la predicción de la imagen actual, y RefPicSetLtCurr incluye y sólo incluye la información de identificación de 1) todas las imágenes de referencia de largo plazo que se utilizan como referencia en la inter predicción de la imagen actual, así como 2) una o más imágenes de referencia de largo plazo que no se utilizan como referencia en la inter predicción de la imagen actual.

De esta manera, las técnicas individuales anteriores o cualquier combinación de las mismas, incluyendo cualquier combinación de los ejemplos alternativos, pueden proporcionar técnicas relacionadas con lo siguiente. Sin embargo, se proporciona la siguiente lista para facilitar la comprensión y no se debe considerar una limitación. Una o más de las técnicas anteriores se pueden implementar conjuntamente o por separado. Además, las técnicas anteriores son ejemplos, y no deben ser consideradas como limitantes de aquellas técnicas de ejemplo específicas.

Restricciones en temporal\_id para el conjunto de imágenes de referencia, de modo que los procedimientos de gestión del búfer DPB se adaptan bien a la escalabilidad temporal, se puede reducir la sobrecarga de señalización y se puede habilitar un proceso simple de extracción del flujo de bits para subconjuntos del flujo de bits extraídos de forma limpia.

Subconjuntos de imágenes de referencia de largo plazo señalizados en el conjunto de parámetros de imagen y el índice se pueden incluir en la cabecera de corte (slice). Esto puede proporcionar una señalización eficaz de imágenes de largo plazo.

La separación del conjunto de imágenes de referencia en diversos subconjuntos, que incluye la separación de la imagen actual o de las imágenes subsiguientes en orden de decodificación, la separación de las que tienen un orden de output anterior o posterior al de la imagen actual. Esto puede proporcionar una mayor eficiencia y una menor complejidad para la inicialización de la lista de referencia y la modificación de la lista de imágenes de referencia.

La codificación de doble diferencial en la señalización de identificación de imágenes de corto plazo puede proporcionar una mayor eficiencia. La identificación extendida y restringida de imágenes de largo plazo puede proporcionar una mayor eficiencia y flexibilidad. La inicialización simplificada de la lista de imágenes de referencia puede eliminar la necesidad de marcar como "sin imagen de referencia" las entradas no completadas en una lista de imágenes de referencia; sin embargo, esto puede no ser necesario en todos los ejemplos.

Procesos simplificados para el output de imágenes decodificadas, inserción en y eliminación del búfer DPB. El contador del orden de las imágenes (POC) puede ser negativo. Esto puede permitir algunos casos de uso importantes que pueden no haber sido permitidos si el POC no puede ser negativo. La señalización de si una imagen es una imagen de referencia no es necesaria en el proceso de decodificación, aunque es posible que todavía pueda estar señalizada. Las marcas en imágenes de referencia como "no utilizadas como referencia" pueden ya no ser necesarias.

En uno o más ejemplos, las funciones descritas pueden ser implementadas en hardware, software, firmware o cualquier combinación de los mismos. Si se implementan en el software, las funciones pueden almacenarse o transmitirse, como una o más instrucciones o códigos, en un medio legible por ordenador y ser ejecutadas por una  
 5 unidad de procesamiento basada en hardware. Los medios legibles por ordenador pueden incluir medios de almacenamiento legibles por ordenador, que corresponden a un medio tangible tal como medios de almacenamiento de datos, o medios de comunicación, que incluyen cualquier medio que facilite la transferencia de un programa de ordenador de un lugar a otro, por ejemplo, de acuerdo con un protocolo de comunicación. De esta manera, los medios legibles por ordenador generalmente pueden corresponder a (1) medios tangibles de almacenamiento legibles por  
 10 ordenador que no son transitorios o (2) medios de comunicación tales como una señal u onda portadora. Los medios de almacenamiento de datos pueden ser cualquier medio disponible al que puedan acceder uno o más ordenadores o uno o más procesadores para recuperar instrucciones, código y/o estructuras de datos para la implementación de las técnicas que se describen en esta divulgación. Un producto de programa de ordenador puede incluir un medio legible por ordenador.

15 A modo de ejemplo, y sin limitación, dichos medios de almacenamiento legibles por ordenador pueden incluir RAM, ROM, EEPROM, CD-ROM u otro almacenamiento en disco óptico, almacenamiento en disco magnético u otros dispositivos de almacenamiento magnético, memoria flash o cualquier otro medio que pueda utilizarse para almacenar el código de programa deseado en forma de instrucciones o estructuras de datos y al que pueda acceder un ordenador.  
 20 Además, cualquier conexión se denomina correctamente como medio legible por ordenador. Por ejemplo, si las instrucciones se transmiten desde un sitio web, servidor u otra fuente remota utilizando un cable coaxial, cable de fibra óptica, par trenzado, línea de suscriptor 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 tecnologías inalámbricas tales como infrarrojos, radio y microondas se incluyen en la definición de medio. Se debe entender, sin embargo, que los  
 25 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 están enfocados a medios de almacenamiento tangibles y no transitorios. El término disco, según se utilizan en el presente documento, incluye el disco compacto (CD), disco láser, disco óptico, disco versátil digital (DVD), el disquete y el disco Blu-ray, en el que los discos suelen reproducir datos magnéticamente, si bien los discos reproducen los datos ópticamente con láseres. Las combinaciones  
 30 de lo anterior también deben incluirse en el alcance de los medios legibles por ordenador.

Las instrucciones pueden ser ejecutadas por uno o más procesadores, tales como uno o más procesadores de señales digitales (DSP: digital signal processors), microprocesadores de propósito general, circuitos integrados específicos de aplicación (ASIC: application specific integrated circuits), matrices lógicas programables de campo (FPGA: field  
 35 programmable logic arrays), u otros circuitos lógicos integrados o discretos equivalentes. Por consiguiente, el término "procesador", según se utiliza en el presente documento, se puede referir a cualquiera de las estructuras anteriores o a cualquier otra estructura adecuada para la implementación de las técnicas descritas en el presente documento. Además, en algunos aspectos, la funcionalidad descrita en el presente documento puede ser proporcionada dentro de módulos dedicados de hardware y/o software configurados para la codificación y decodificación, o incorporados en un  
 40 códec combinado. Además, las técnicas pueden ser implementadas completamente en uno o más circuitos o elementos lógicos.

Las técnicas de esta divulgación pueden ser implementadas en una amplia diversidad de dispositivos o aparatos, que incluyen un teléfono inalámbrico, un circuito integrado (IC: integrated circuit) o un conjunto de circuitos integrados (por  
 45 ejemplo, un conjunto de chips). En esta divulgación se describen diversos componentes, módulos, o unidades para enfatizar aspectos funcionales de dispositivos configurados para realizar las técnicas divulgadas, pero no necesariamente requieren su realización por unidades de hardware diferentes. Más bien, según se ha descrito anteriormente, se pueden combinar diversas unidades en una unidad de hardware de códec o pueden ser proporcionadas por una colección de unidades de hardware inter operativas, que incluyen uno o más procesadores  
 50 según se ha descrito anteriormente, junto con un software y/o firmware adecuados.

Se han descrito diversos ejemplos. Estos y otros ejemplos están dentro del alcance de las siguientes reivindicaciones.

## REIVINDICACIONES

1. Un procedimiento de codificar datos de video, comprendiendo el procedimiento:
- codificar información indicativa de imágenes de referencia que pertenecen a un conjunto de imágenes de referencia,
  - 5 en el que el conjunto de imágenes de referencia identifica las imágenes de referencia que se pueden utilizar potencialmente para una inter predicción de una imagen actual y que se pueden utilizar potencialmente para una inter predicción de una o más imágenes posteriores a la imagen actual en orden de decodificación;
  - construir una pluralidad de subconjuntos de imágenes de referencia cada uno de los cuales identifica imágenes de referencia del conjunto de imágenes de referencia, en el que la pluralidad de subconjuntos de imágenes de referencia
  - 10 incluye:
    - un primer subconjunto de imágenes de referencia que identifica imágenes de referencia de corto plazo que son anteriores a la imagen actual en orden de decodificación y anteriores a la imagen actual en orden de output y que se pueden utilizar potencialmente para la inter predicción de la imagen actual y de una o más de la una o más imágenes posteriores a la imagen actual en orden de decodificación;
    - 15 un segundo subconjunto de imágenes de referencia que identifica imágenes de referencia de corto plazo que son anteriores a la imagen actual en orden de decodificación y posteriores a la imagen actual en orden de output y que se pueden utilizar potencialmente para la inter predicción de la imagen actual y de una o más de la una o más imágenes posteriores a la imagen actual en orden de decodificación; y
    - un tercer subconjunto de imágenes de referencia que identifica imágenes de referencia de largo plazo que son
    - 20 anteriores a la imagen actual en orden de decodificación y que se pueden utilizar potencialmente para la inter predicción de la imagen actual y de una o más de la una o más imágenes posteriores a la imagen actual en orden de decodificación;
    - añadir imágenes de referencia del primer subconjunto de la pluralidad de subconjuntos de imágenes de referencia, seguido de imágenes de referencia del segundo subconjunto de la pluralidad de subconjuntos de imágenes de referencia, y seguido de imágenes de referencia del tercer subconjunto de la pluralidad de subconjuntos de imágenes de referencia a una lista de imágenes de referencia mientras un número de entradas en la lista de imágenes de referencia no sea superior a un número máximo de entradas permisibles en la lista de imágenes de referencia; y
    - 25 codificar la imagen actual en base a la lista de imágenes de referencia.
- 30 2. El procedimiento de la reivindicación 1, en el que añadir las imágenes de referencia comprende:
- añadir las imágenes de referencia del primer subconjunto de imágenes de referencia a la lista de imágenes de referencia hasta que se hayan añadido todas las imágenes de referencia del primer subconjunto de imágenes de referencia a la lista de imágenes de referencia o el número de entradas en la lista de imágenes de referencia sea igual al número máximo de entradas permisibles en la lista de imágenes de referencia;
  - 35 cuando el número de entradas en la lista de imágenes de referencia sea inferior al número máximo de entradas permisibles en la lista de imágenes de referencia, y después de añadir las imágenes de referencia del primer subconjunto de imágenes de referencia, añadir las imágenes de referencia del segundo subconjunto de imágenes de referencia a la lista de imágenes de referencia hasta que se hayan identificado todas las imágenes de referencia del segundo subconjunto de imágenes de referencia en la lista de imágenes de referencia o el número de entradas en la
  - 40 lista de imágenes de referencia sea igual al número máximo de entradas permisibles en la lista de imágenes de referencia; y
  - cuando el número de entradas en la lista de imágenes de referencia sea inferior al número máximo de entradas permisibles en la lista de imágenes de referencia, y después de añadir las imágenes de referencia del segundo subconjunto de imágenes de referencia, añadir las imágenes de referencia del tercer subconjunto de imágenes de referencia a la lista de imágenes de referencia hasta que se hayan identificado todas las imágenes de referencia del tercer subconjunto de imágenes de referencia en la lista de imágenes de referencia o el número de entradas en la lista de imágenes de referencia sea igual al número máximo de entradas permisibles en la lista de imágenes de referencia.
  - 45
3. El procedimiento de la reivindicación 2, que comprende además:
- 50 codificar un primer elemento de sintaxis que indica un número de imágenes de referencia en el primer subconjunto de imágenes de referencia; y
  - codificar un segundo elemento de sintaxis que indica un número de imágenes de referencia en el segundo subconjunto de imágenes de referencia;
  - codificar un tercer elemento de sintaxis que indica un número de imágenes de referencia de largo plazo cuya
  - 55 información de identificación está incluida en un conjunto de parámetros de imagen (PPS: Picture Parameter Set), en el que las imágenes de referencia de largo plazo se pueden utilizar potencialmente para la inter predicción de la imagen actual y se pueden utilizar potencialmente para la inter predicción de una o más imágenes posteriores a la imagen actual en orden de decodificación;
  - codificar un cuarto elemento de sintaxis que indica un número de las imágenes de referencia de largo plazo cuya
  - 60 información de identificación no está incluida en el conjunto de parámetros de imagen (PPS); y
  - determinar un número de imágenes de referencia en el tercer subconjunto de imágenes de referencia en base al tercer elemento de sintaxis y al cuarto elemento de sintaxis,

en el que añadir las imágenes de referencia del primer subconjunto de imágenes de referencia a la lista de imágenes de referencia comprende añadir las imágenes de referencia del primer subconjunto de imágenes de referencia hasta que un número de imágenes de referencia añadidas de la primera lista de imágenes de referencia sea igual a un valor del primer elemento de sintaxis o el número de entradas en la lista de imágenes de referencia sea igual al número

5 máximo de entradas permisibles en la lista de imágenes de referencia,

en el que añadir las imágenes de referencia del segundo subconjunto de imágenes de referencia a la lista de imágenes de referencia comprende añadir las imágenes de referencia del segundo subconjunto de imágenes de referencia hasta que un número de imágenes de referencia añadidas de la segunda lista de imágenes de referencia sea igual a un valor del segundo elemento de sintaxis o el número de entradas en la lista de imágenes de referencia

10 sea igual al número máximo de entradas permisibles en la lista de imágenes de referencia, y

en el que añadir las imágenes de referencia del tercer subconjunto de imágenes de referencia a la lista de imágenes de referencia comprende añadir las imágenes de referencia del tercer subconjunto de imágenes de referencia hasta que un número de imágenes de referencia añadidas del tercer subconjunto de imágenes de referencia sea igual al número determinado de imágenes de referencia en el tercer conjunto de imágenes de referencia o el número de

15 entradas en la lista de imágenes de referencia sea igual al número máximo de entradas permisibles en la lista de imágenes de referencia.

4. El procedimiento de la reivindicación 3, en el que el primer elemento de sintaxis comprende un elemento de sintaxis num\_short\_term\_curr0, el segundo elemento de sintaxis comprende un elemento de sintaxis num\_short\_term\_curr1,

20 el tercer elemento de sintaxis comprende un elemento de sintaxis num\_long\_term\_pps\_curr, y el cuarto elemento de sintaxis comprende un elemento de sintaxis num\_long\_term\_add\_curr.

5. Procedimiento de la reivindicación 1, en el que la lista de imágenes de referencia comprende una primera lista de imágenes de referencia, el número de entradas en la lista de imágenes de referencia comprende un número de

25 entradas en la primera lista de imágenes de referencia, y el número máximo de entradas permisibles en la lista de imágenes de referencia comprende un número máximo de entradas permisibles en la primera lista de imágenes de referencia, comprendiendo el procedimiento además:

añadir las imágenes de referencia del segundo subconjunto de la pluralidad de subconjuntos de imágenes de referencia, seguido de las imágenes de referencia del primer subconjunto de la pluralidad de subconjuntos de

30 imágenes de referencia, y seguido de las imágenes de referencia del tercer subconjunto de la pluralidad de subconjuntos de imágenes de referencia a una segunda lista de imágenes de referencia mientras un número de entradas en la segunda lista de imágenes de referencia no sea superior a un número máximo de entradas permisibles en la segunda lista de imágenes de referencia,

en el que codificar la imagen actual comprende codificar la imagen actual en base a la primera lista de imágenes

35 de referencia y la segunda lista de imágenes de referencia.

6. El procedimiento de la reivindicación 1,

en el que codificar comprende decodificar información, con un decodificador de video, de imágenes de referencia que pertenecen al conjunto de imágenes de referencia,

40 en el que construir comprende construir, con el decodificador de video, la pluralidad de subconjuntos de imágenes de referencia cada uno de los cuales identifica cero o más de las imágenes de referencia del conjunto de imágenes de referencia;

en el que añadir comprende añadir, con el decodificador de video, imágenes de referencia del primer subconjunto de la pluralidad de subconjuntos de imágenes de referencia, seguido de imágenes de referencia del segundo

45 subconjunto de la pluralidad de subconjuntos de imágenes de referencia, y seguido de imágenes de referencia del tercer subconjunto de la pluralidad de subconjuntos de imágenes de referencia a la lista de imágenes de referencia mientras el número de entradas en la lista de imágenes de referencia no sea superior al número máximo de entradas permisibles en la lista de imágenes de referencia; y

en el que codificar comprende decodificar, con el decodificador de video, la imagen actual en base a la lista de

50 imágenes de referencia.

7. El procedimiento de la reivindicación 1,

en el que codificar comprende codificar información, con un codificador de video (video encoder), de imágenes de referencia que pertenecen al conjunto de imágenes de referencia,

55 en el que construir comprende construir, con el codificador de video (video encoder), la pluralidad de subconjuntos de imágenes de referencia cada uno de los cuales identifica cero o más de las imágenes de referencia del conjunto de imágenes de referencia;

en el que añadir comprende añadir, con el codificador de video (video encoder), imágenes de referencia del primer subconjunto de la pluralidad de subconjuntos de imágenes de referencia, seguido de imágenes de referencia del

60 segundo subconjunto de la pluralidad de subconjuntos de imágenes de referencia, y seguido de imágenes de referencia del tercer subconjunto de la pluralidad de subconjuntos de imágenes de referencia a la lista de imágenes de referencia mientras el número de entradas en la lista de imágenes de referencia no sea superior al número máximo de entradas permisibles en la lista de imágenes de referencia; y

en el que codificar comprende codificar, con el codificador de video (video encoder), la imagen actual en base a la lista de imágenes de referencia.

8. Un dispositivo para codificar datos de video, comprendiendo el dispositivo:

- 5 medios para codificar información indicativa de imágenes de referencia que pertenecen a un conjunto de imágenes de referencia, en el que el conjunto de imágenes de referencia identifica las imágenes de referencia que se pueden utilizar potencialmente para una inter predicción de una imagen actual y que se pueden utilizar potencialmente para una inter predicción de una o más imágenes posteriores a la imagen actual en orden de decodificación;
- 10 medios para construir una pluralidad de subconjuntos de imágenes de referencia cada uno de los cuales identifica imágenes de referencia del conjunto de imágenes de referencia, en el que la pluralidad de subconjuntos de imágenes de referencia incluye:
- 15 un primer subconjunto de imágenes de referencia que identifica imágenes de referencia de corto plazo que son anteriores a la imagen actual en orden de decodificación y anteriores a la imagen actual en orden de output y que se pueden utilizar potencialmente para la inter predicción de la imagen actual y de una o más de la una o más imágenes posteriores a la imagen actual en orden de decodificación;
- 20 un segundo subconjunto de imágenes de referencia que identifica imágenes de referencia de corto plazo que son anteriores a la imagen actual en orden de decodificación y posteriores a la imagen actual en orden de output y que se pueden utilizar potencialmente para la inter predicción de la imagen actual y de una o más de la una o más imágenes posteriores a la imagen actual en orden de decodificación; y
- 25 un tercer subconjunto de imágenes de referencia que identifica imágenes de referencia de largo plazo que son anteriores a la imagen actual en orden de decodificación y que se pueden utilizar potencialmente para la inter predicción de la imagen actual y de una o más de la una o más imágenes posteriores a la imagen actual en orden de decodificación;
- 30 medios para añadir imágenes de referencia del primer subconjunto de la pluralidad de subconjuntos de imágenes de referencia, seguido de imágenes de referencia del segundo subconjunto de la pluralidad de subconjuntos de imágenes de referencia, y seguido de imágenes de referencia del tercer subconjunto de la pluralidad de subconjuntos de imágenes de referencia a una lista de imágenes de referencia mientras un número de entradas en la lista de imágenes de referencia no sea superior a un número máximo de entradas permisibles en la lista de imágenes de referencia; y
- 30 medios para codificar la imagen actual en base a la lista de imágenes de referencia.

9. El dispositivo de la reivindicación 8, en el que los medios para añadir las imágenes de referencia comprenden:

- 35 medios para añadir las imágenes de referencia del primer subconjunto de imágenes de referencia a la lista de imágenes de referencia hasta que se hayan añadido todas las imágenes de referencia del primer subconjunto de imágenes de referencia a la lista de imágenes de referencia o el número de entradas en la lista de imágenes de referencia sea igual al número máximo de entradas permisibles en la lista de imágenes de referencia;
- 40 cuando el número de entradas en la lista de imágenes de referencia sea inferior al número máximo de entradas permisibles en la lista de imágenes de referencia, y después de añadir las imágenes de referencia del primer subconjunto de imágenes de referencia, medios para añadir las imágenes de referencia del segundo subconjunto de imágenes de referencia a la lista de imágenes de referencia hasta que se hayan identificado todas las imágenes de referencia del segundo subconjunto de imágenes de referencia en la lista de imágenes de referencia o el número de entradas en la lista de imágenes de referencia sea igual al número máximo de entradas permisibles en la lista de imágenes de referencia; y
- 45 cuando el número de entradas en la lista de imágenes de referencia sea inferior al número máximo de entradas permisibles en la lista de imágenes de referencia, y después de añadir las imágenes de referencia del segundo subconjunto de imágenes de referencia, medios para añadir las imágenes de referencia del tercer subconjunto de imágenes de referencia a la lista de imágenes de referencia hasta que se hayan identificado todas las imágenes de referencia del tercer subconjunto de imágenes de referencia en la lista de imágenes de referencia o el número de entradas en la lista de imágenes de referencia sea igual al número máximo de entradas permisibles en la lista de imágenes de referencia.
- 50 imágenes de referencia.

10. El procedimiento de la reivindicación 8, en el que el dispositivo comprende un decodificador de video, y en el que el decodificador de video está configurado para:

- 55 decodificar información de imágenes de referencia que pertenecen al conjunto de imágenes de referencia, construir la pluralidad de subconjuntos de imágenes de referencia cada uno de los cuales identifica cero o más de las imágenes de referencia del conjunto de imágenes de referencia;
- 60 añadir imágenes de referencia del primer subconjunto de la pluralidad de subconjuntos de imágenes de referencia, seguido de imágenes de referencia del segundo subconjunto de la pluralidad de subconjuntos de imágenes de referencia, y seguido de imágenes de referencia del tercer subconjunto de la pluralidad de subconjuntos de imágenes de referencia a la lista de imágenes de referencia mientras el número de entradas en la lista de imágenes de referencia no sea superior al número máximo de entradas permisibles en la lista de imágenes de referencia; y decodificar la imagen actual en base a la lista de imágenes de referencia.

11. El procedimiento de la reivindicación 8, en el que el dispositivo comprende un codificador de video (video encoder), y en el que el codificador de video (video encoder) está configurado para:
- codificar información de imágenes de referencia que pertenecen al conjunto de imágenes de referencia,
  - construir la pluralidad de subconjuntos de imágenes de referencia cada uno de los cuales identifica cero o más de
- 5 las imágenes de referencia del conjunto de imágenes de referencia;
- añadir imágenes de referencia del primer subconjunto de la pluralidad de subconjuntos de imágenes de referencia, seguido de imágenes de referencia del segundo subconjunto de la pluralidad de subconjuntos de imágenes de referencia, y seguido de imágenes de referencia del tercer subconjunto de la pluralidad de subconjuntos de imágenes de referencia a la lista de imágenes de referencia mientras el número de entradas en la lista de imágenes de referencia
- 10 no sea superior al número máximo de entradas permisibles en la lista de imágenes de referencia; y
- codificar la imagen actual en base a la lista de imágenes de referencia.
12. Un medio de almacenamiento legible por ordenador que tiene almacenadas en el mismo unas instrucciones que, cuando son ejecutadas, hacen que un procesador de un dispositivo para codificar datos de video realice el
- 15 procedimiento de cualquiera de las reivindicaciones 1 a 7.

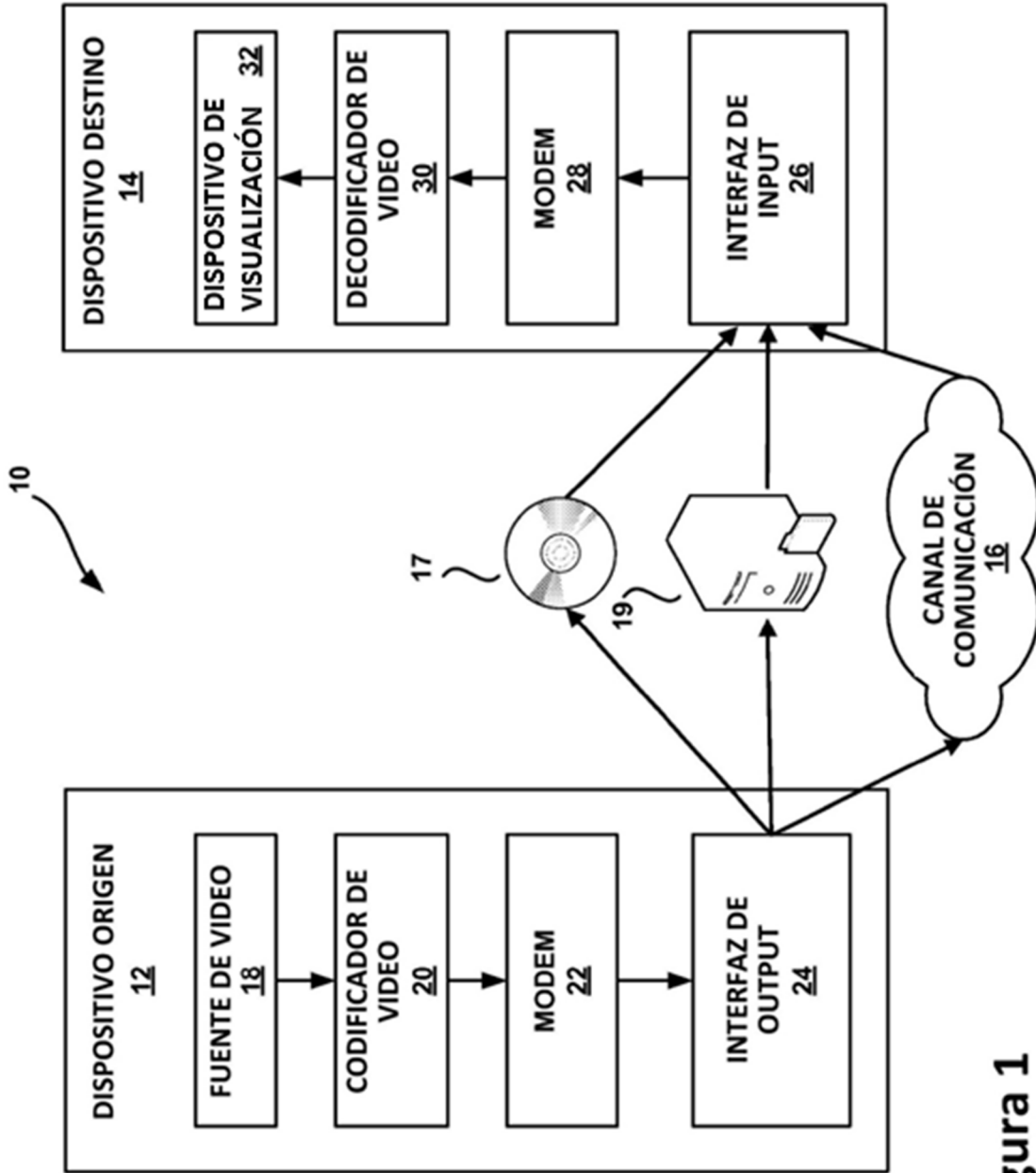


Figura 1



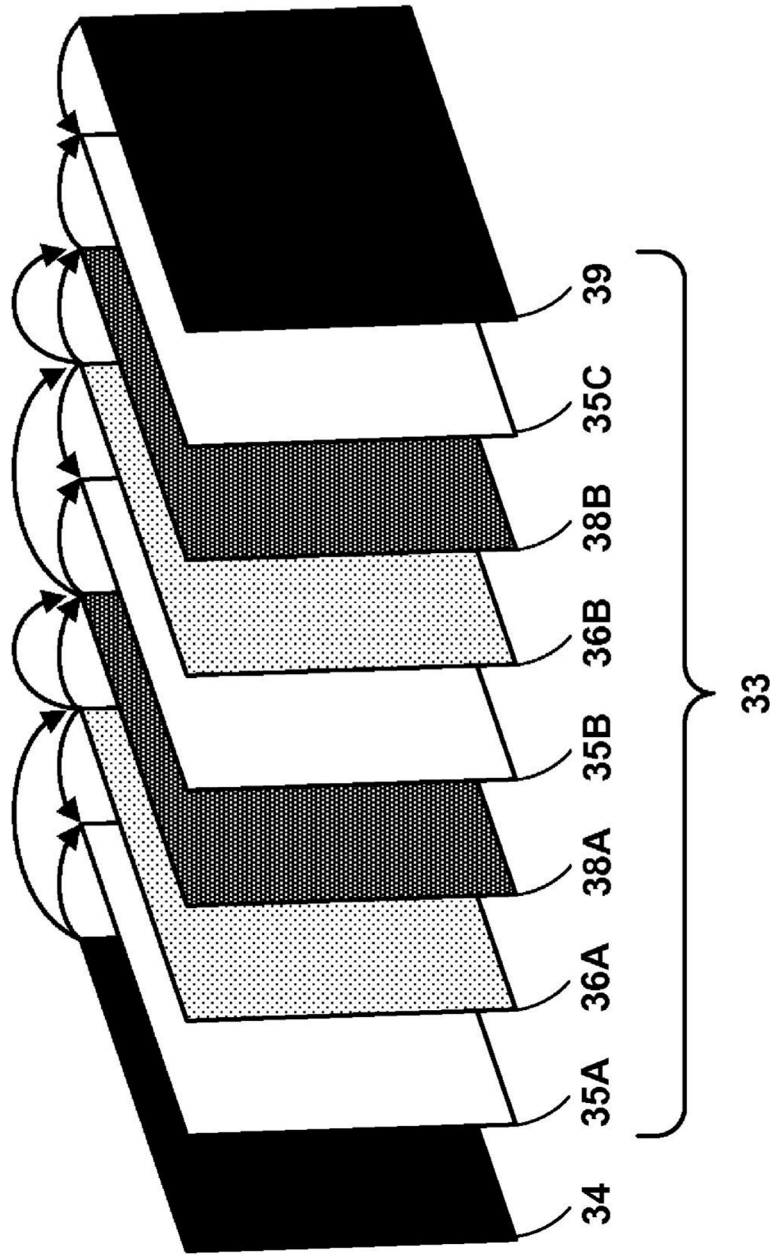


Figura 2

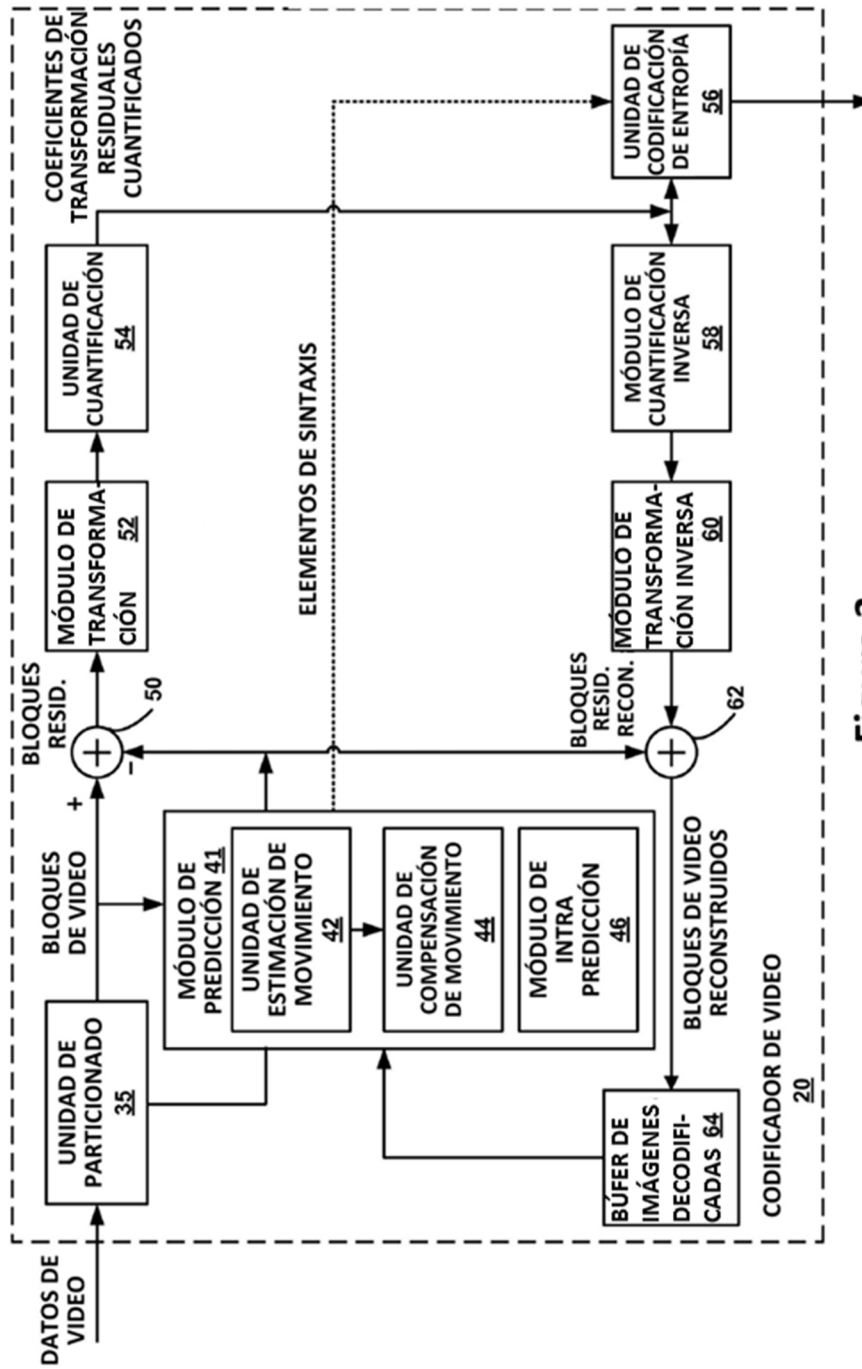


Figura 3

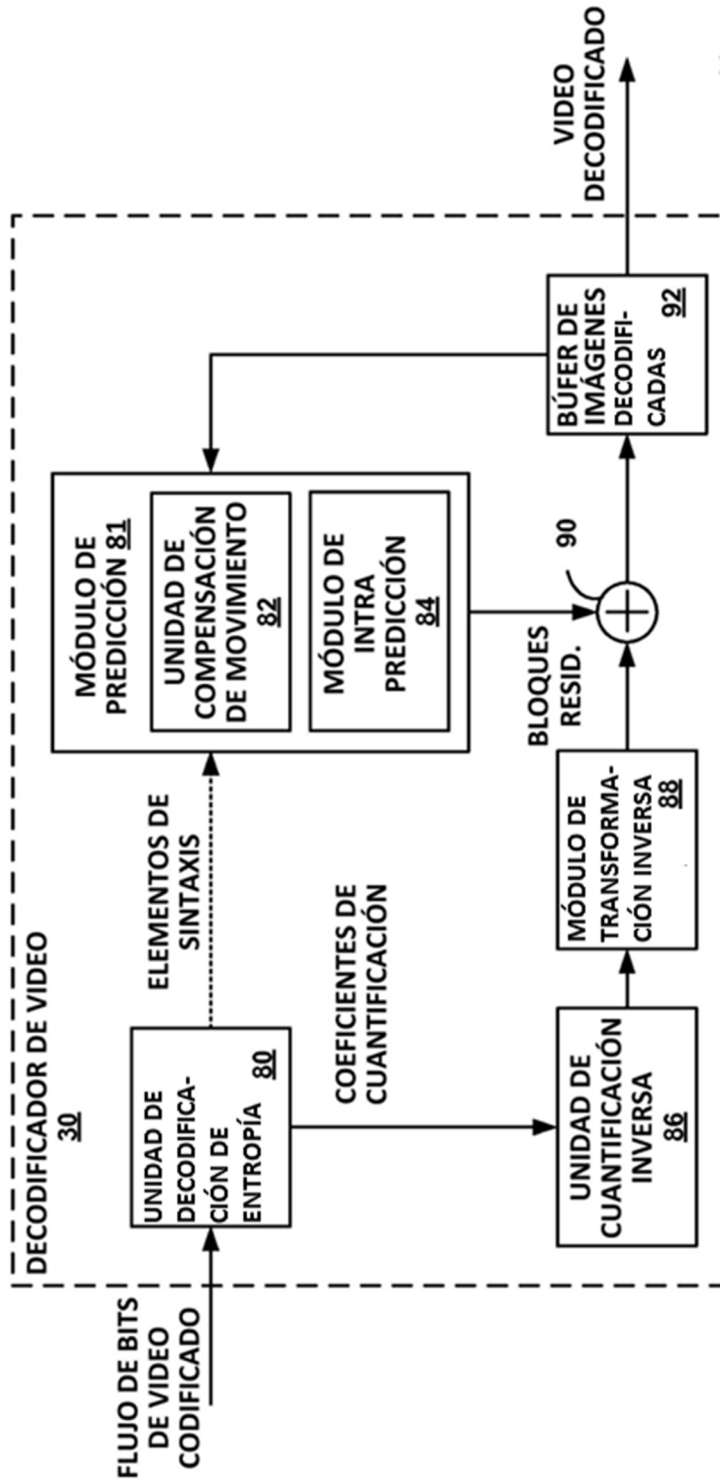
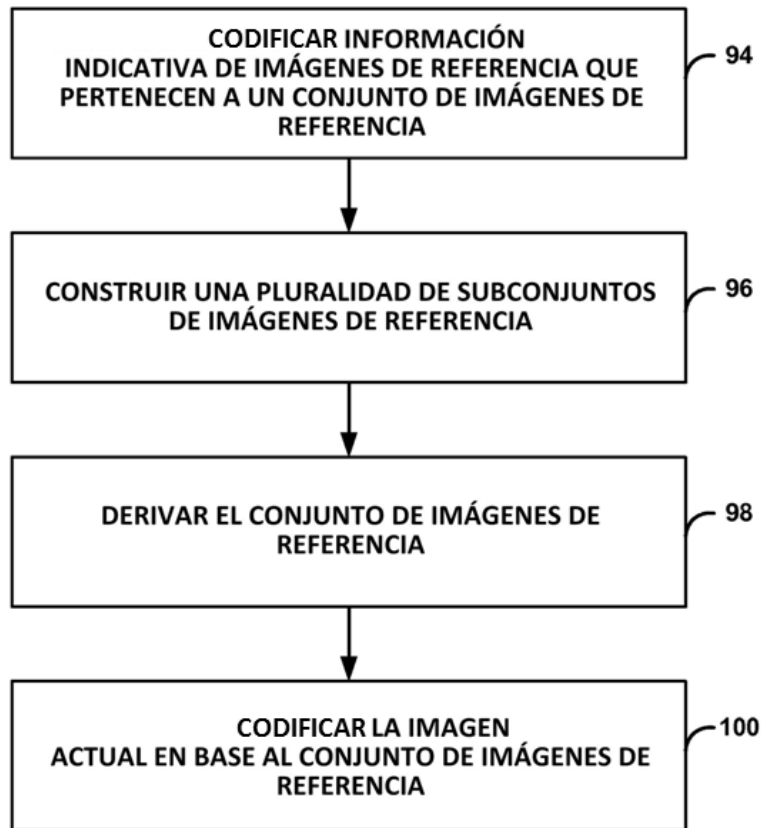
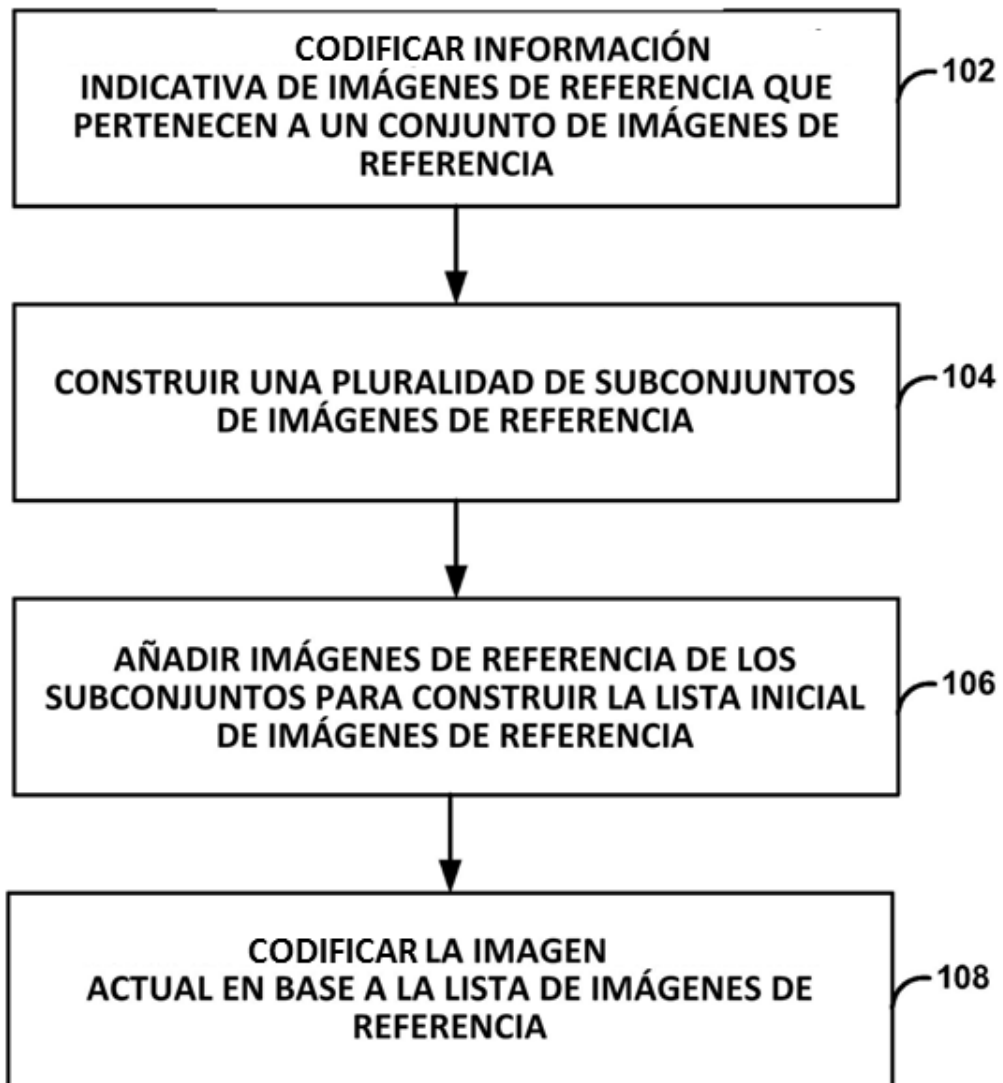


Figura 4



**Figura 5**



**Figura 6**

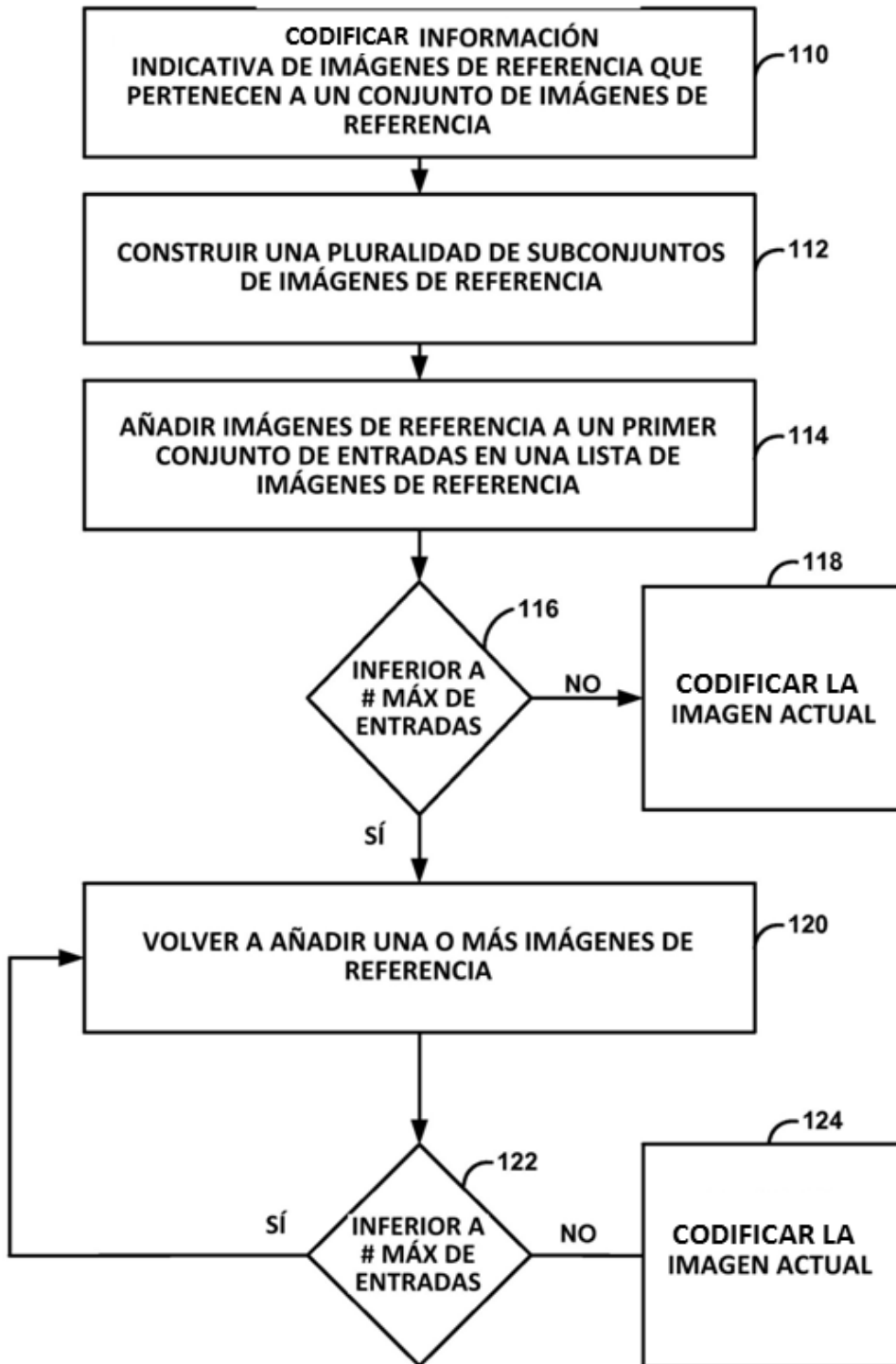


Figura 7

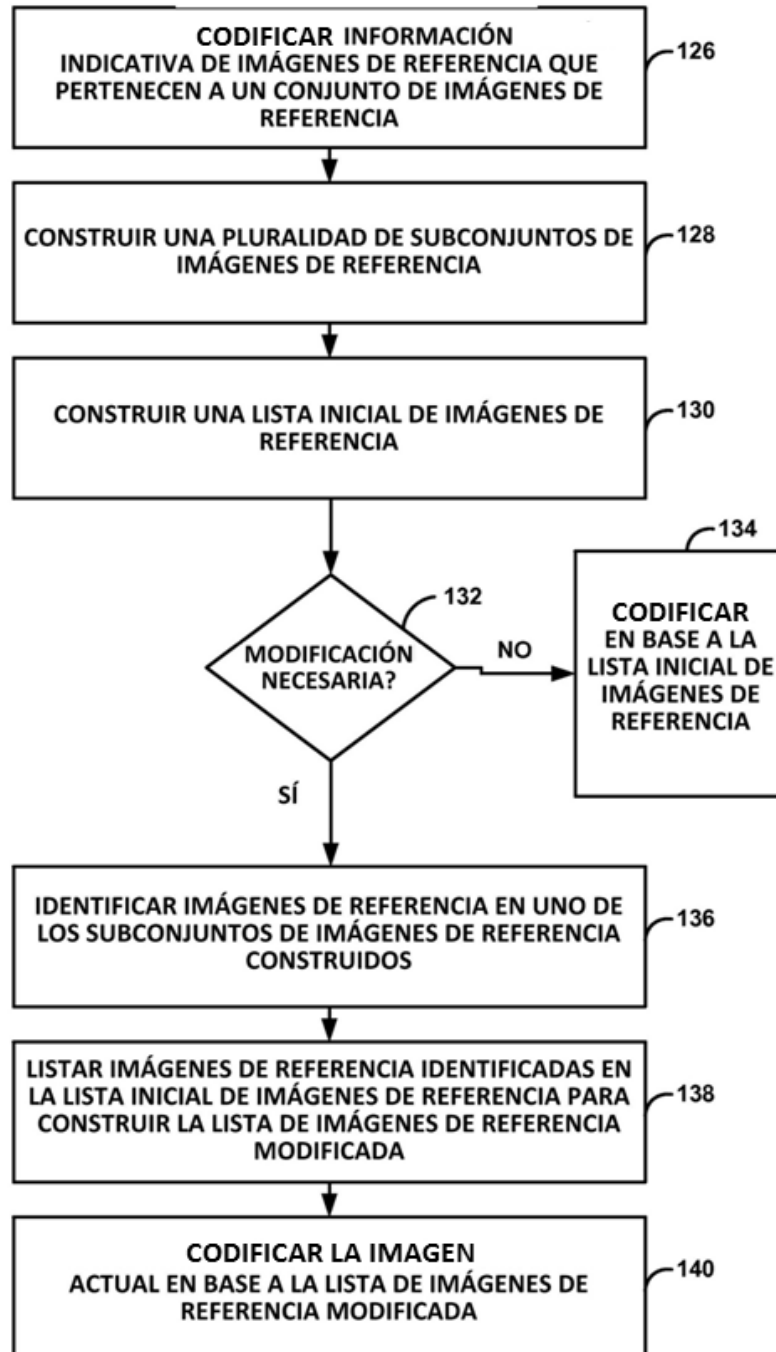


Figura 8

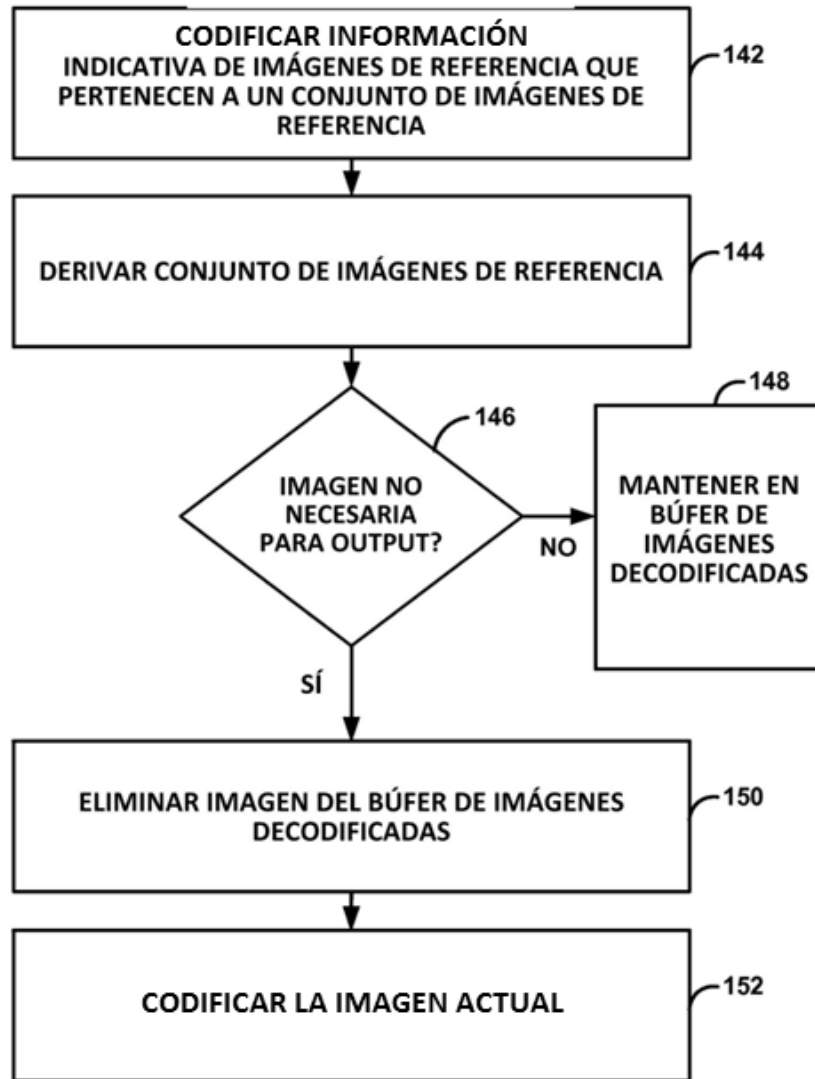
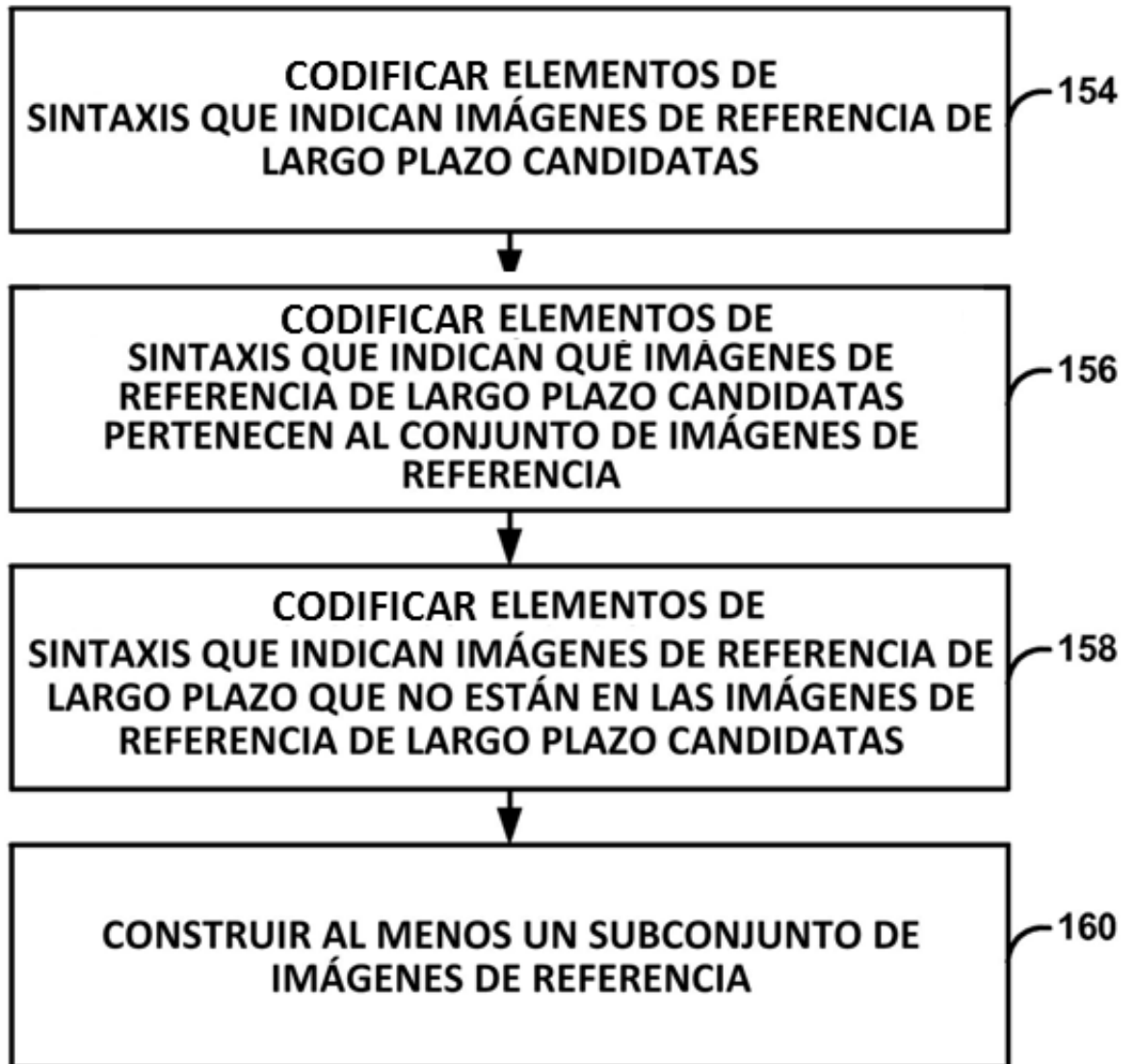


Figura 9





**Figura 10**