

19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 613 003**

51 Int. Cl.:

H04N 19/85 (2014.01)

H04N 19/70 (2014.01)

H04N 19/149 (2014.01)

H04N 19/44 (2014.01)

H04N 19/46 (2014.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

86 Fecha de presentación y número de la solicitud internacional: **20.12.2013 PCT/US2013/077283**

87 Fecha y número de publicación internacional: **10.07.2014 WO14107362**

96 Fecha de presentación y número de la solicitud europea: **20.12.2013 E 13821601 (5)**

97 Fecha y número de publicación de la concesión europea: **02.11.2016 EP 2941886**

54 Título: **Señalización de información de obtención de pulso de reloj para la temporización de vídeo en la codificación de vídeo**

30 Prioridad:

07.01.2013 US 201361749866 P

23.10.2013 US 201314061215

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

19.05.2017

73 Titular/es:

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

72 Inventor/es:

WANG, YE-KUI

74 Agente/Representante:

FORTEA LAGUNA, Juan José

ES 2 613 003 T3

Aviso: En el plazo de nueve meses a contar desde la fecha de publicación en el Boletín Europeo de Patentes, de la mención de concesión de la patente europea, cualquier persona podrá oponerse ante la Oficina Europea de Patentes a la patente concedida. La oposición deberá formularse por escrito y estar motivada; sólo se considerará como formulada una vez que se haya realizado el pago de la tasa de oposición (art. 99.1 del Convenio sobre Concesión de Patentes Europeas).

DESCRIPCIÓN

Señalización de información de obtención de pulso de reloj para la temporización de vídeo en la codificación de vídeo

5 Esta solicitud reivindica el beneficio de la solicitud provisional estadounidense Nº 61/749.866, presentada el 7 de enero de 2013.

Campo técnico

10 Esta divulgación se refiere a la codificación de vídeo y al procesamiento de vídeo y, más en particular, a técnicas para la señalización de información de temporización en información de vídeo.

Antecedentes

15 Las capacidades del vídeo digital pueden incorporarse en una amplia gama de dispositivos, incluyendo televisores digitales, sistemas de difusión directa digital, sistemas de difusión inalámbrica, asistentes digitales personales (PDA), ordenadores portátiles o de escritorio, ordenadores de tableta, lectores de libros electrónicos, cámaras digitales, dispositivos de grabación digital, reproductores de medios digitales, dispositivos de videojuegos, consolas de videojuegos, teléfonos celulares o de radio por satélite, los denominados teléfonos "inteligentes", dispositivos de videoconferencia, dispositivos de transmisión por flujo de vídeo, y similares. Los dispositivos de vídeo digital implementan técnicas de compresión de vídeo, tales como las descritas en las normas definidas por MPEG-2, MPEG-4, ITU-T H.263, ITU-T H.264/MPEG-4, Parte 10, Codificación Avanzada de Vídeo (AVC) (H.264/AVC), la norma de codificación de vídeo de alta eficacia (HEVC) y las extensiones de tales normas. Los dispositivos de vídeo pueden transmitir, recibir, codificar, decodificar y/o almacenar información de vídeo digital más eficazmente, implementando tales técnicas de compresión de vídeo.

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

30 La predicción espacial o temporal utiliza un bloque predictivo. Los datos residuales representan diferencias de píxeles entre el bloque original a codificar y el bloque predictivo. Un bloque inter-codificado se codifica 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 codificado y el bloque predictivo. Un bloque intra-codificado se codifica según una modalidad de intra-codificación y los datos residuales. Para una mayor compresión, los datos residuales pueden transformarse desde el dominio de píxeles a un dominio de transformación, dando como resultado coeficientes de transformación residuales, los cuales pueden cuantizarse posteriormente. Los coeficientes de transformación cuantizados, inicialmente dispuestos en una formación bidimensional, pueden recorrerse con el fin de producir un vector unidimensional de coeficientes de transformación, y puede aplicarse codificación por entropía para lograr aún más compresión.

40 Una secuencia dada de vídeo codificado para un flujo de bits incluye una secuencia ordenada de imágenes codificadas. En las normas H.264 / AVC y HEVC, el orden de decodificación de las imágenes codificadas para un flujo de bits es equivalente a la secuencia ordenada. Sin embargo, las normas también dan soporte a un orden de salida de imágenes decodificadas, que difiere del orden de decodificación, y en tales casos cada una de las imágenes codificadas se asocia a un valor del recuento de orden de imágenes (POC) que especifica el orden de salida de la imagen en la secuencia de vídeo.

50 La información de temporización de vídeo para una secuencia de vídeo se puede señalar en los elementos sintácticos de una o más estructuras sintácticas (denominadas alternativamente "estructuras de conjuntos de parámetros" o simplemente "conjuntos de parámetros"). Las estructuras sintácticas pueden incluir un conjunto de parámetros de secuencia (SPS) que incluye información de codificación que se aplica a todos los fragmentos de una secuencia de vídeo codificada. El SPS puede por sí mismo incluir parámetros denominados información de facilidad de uso de vídeo (VUI), que incluyen información de decodificador de referencia hipotético (HRD), así como información para mejorar el uso de la secuencia de vídeo correspondiente para diversos fines. La misma información de HRD puede señalizarse usando una estructura sintáctica de HRD incluíble dentro de otras estructuras sintácticas como la estructura sintáctica de VUI. Las estructuras sintácticas también pueden incluir un conjunto de parámetros de vídeo (VPS) que describe las características de una secuencia de vídeo correspondiente, tales como elementos

sintácticos comunes compartidos por múltiples capas o puntos de operación, así como otra información del punto de operación, que pueden ser comunes a varios conjuntos de parámetros de secuencia, tal como información de HRD para diversas capas o subcapas.

5 Se hace referencia a las siguientes propuestas al JCTVC:

Documento N°	Autor	Título	Fecha
JCTVC-K1003	Bross B et al	"Codificación de vídeo de alta eficacia (HEVC) - borrador de memoria descriptiva textual 9 (SoDIS) - versión v1"	21/10/2012
JCTVC-J0272	Haque M et al	"Simplificaciones de los parámetros de HRD para la ajustabilidad temporal a escala"	03/07/2012
JCTVC-J0548	Haque M et al	"AHG10: diseños sintácticos de VUI y HRD acordados por el BoG sobre VPS y NUH"	11/07/2012
JCTVC-J0570	Wang Y-K	"AHG9: Sobre el indicador_de_velocidad_de_imagen_fija"	15/07/2012
JCTVC-K0125	Wang Y-K et al	"AHG9: Sobre el conjunto de parámetros de vídeo"	01/10/2012

También se hace referencia a los artículos siguientes:

10 D Sachin et al "Un decodificador de referencia hipotético mejorado para HEVC", Actas de SPIE, vol 8666, el 21 de febrero de 2013, página 866608;

J. Webb "Conformidad con HRD para la codificación de vídeo en tiempo real de la norma H.264", Conferencia Internacional del IEEE sobre Procesamiento de Imágenes, 1 de septiembre de 2007;

15 P Chou et al, "Un decodificador de referencia hipotético generalizado para la norma H.264/AVC", Transacciones del IEEE sobre Circuitos y Sistemas para Tecnología de Vídeo, vol. 13, N° 7, páginas 674 a 687, 1 de julio de 2008.

También se hace referencia a la siguiente propuesta para JCT3V:

20

Documento N°	Autor	Título	Fecha
JCT3V-A0099	Chen Y et al	"3D-HLS: Conjunto de parámetros de vídeo para 3D-HEVC "	11/07/2012

Sumario

25 La invención se define en las reivindicaciones a las que se hace ahora referencia. En general, esta divulgación describe técnicas para la codificación de vídeo y, más en particular, técnicas para la señalización de información de temporización, por ejemplo, para especificar la temporización de la emisión de imágenes y/o para definir un modelo de memoria intermedia, tal como un decodificador de referencia hipotético (HRD). En algunos ejemplos, las técnicas pueden incluir la generación, para una secuencia de vídeo codificado, de un flujo de bits codificado para señalar la escala temporal y el número de unidades en un pulso de reloj, como máximo, una vez, tanto en la estructura sintáctica del conjunto de parámetros de vídeo (VPS) como en la parte de información de facilidad de uso de vídeo (VUI) de la estructura sintáctica del conjunto de parámetros de secuencia (SPS) para la secuencia de vídeo codificado. Es decir, en algunos ejemplos, en una estructura sintáctica dada del VPS para una secuencia de vídeo codificado, la estructura sintáctica del VPS puede incluir los elementos sintácticos de la escala temporal y el número de unidades en un pulso de reloj, como máximo, una vez. Análogamente, en algunos ejemplos, en una estructura sintáctica dada de la VUI (por ejemplo, la parte de VUI de una estructura sintáctica del SPS) para una secuencia de vídeo codificado, la estructura sintáctica de la VUI puede incluir los elementos sintácticos de la escala temporal y del número de unidades en un pulso de reloj, como máximo, una vez.

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

Breve descripción de los dibujos

45 La FIG. 1 es un diagrama de bloques que ilustra un sistema ejemplar de codificación y decodificación de vídeo que

puede utilizar las técnicas descritas en esta divulgación.

La FIG. 2 es un diagrama de bloques que ilustra un codificador ejemplar de vídeo que puede implementar las técnicas descritas en esta divulgación.

La FIG. 3 es un diagrama de bloques que ilustra un decodificador ejemplar de vídeo que puede implementar las técnicas descritas en esta divulgación.

La FIG. 4 es un diagrama de bloques que ilustra la información de temporización para una estructura ejemplar de codificación para un conjunto de imágenes de referencia, según las técnicas descritas en este documento.

La FIG. 5 es un diagrama de flujo que muestra un procedimiento ejemplar de funcionamiento según las técnicas descritas en esta divulgación.

Las FIGs. 6A a 6B son diagramas de flujo que ilustran procedimientos ejemplares de funcionamiento según las técnicas descritas en esta divulgación.

La FIG. 7 es un diagrama de flujo que ilustra un procedimiento ejemplar de funcionamiento según las técnicas descritas en esta divulgación.

La FIG. 8 es un diagrama de flujo que ilustra un procedimiento ejemplar de funcionamiento según las técnicas descritas en esta divulgación.

Las FIGs. 9A a 9B son diagramas de flujo que ilustran procedimientos ejemplares de funcionamiento según las técnicas descritas en esta divulgación.

La FIG. 10 es un diagrama de flujo que ilustra un procedimiento ejemplar de funcionamiento según las técnicas descritas en esta divulgación.

Descripción detallada

Esta divulgación describe varias técnicas para la codificación de vídeo y, más en particular, para las técnicas para la señalización de información de temporización, por ejemplo, para especificar la temporización de emisión de imágenes y/o para definir un modelo de memoria intermedia o de decodificación, tal como un decodificador de referencia hipotético (HRD). En general, el término "señalización" se utiliza en esta divulgación para hacer referencia a la señalización que tiene lugar dentro de un flujo de bits codificado. Un codificador puede generar elementos sintácticos para señalar la información en un flujo de bits como parte de un proceso de codificación de vídeo. Un dispositivo de decodificación, u otro dispositivo de procesamiento de vídeo, pueden recibir un flujo de bits codificado, e interpretar los elementos sintácticos en el flujo de bits codificado como parte de un proceso de decodificación de vídeo u otro procesamiento de vídeo. Por ejemplo, para indicar la temporización de emisión para la conmutación de una imagen dada a la siguiente imagen en una secuencia de vídeo codificado según el orden de salida, la información de temporización para la secuencia de vídeo codificado puede, en algunos casos, señalar un número de pulsos de reloj que corresponde a una diferencia de valores del recuento de orden de imágenes (POC) igual a uno. Una diferencia de valores del POC igual a uno puede representar la diferencia entre un valor del POC para una imagen dada y el valor del POC para la próxima imagen según el orden de salida, por ejemplo, el valor del POC para la 2ª imagen y el valor del POC para la 3ª imagen, según el orden de salida. La información de temporización de vídeo puede incluir también una condición que especifica si la información de temporización de vídeo señala o no el número de pulsos de reloj que corresponden a una diferencia de valores del recuento de orden de imágenes igual a uno. En otras palabras, solo si se cumple la condición, la información de temporización de vídeo señala el número de pulsos de reloj que corresponden a una diferencia de valores del recuento de orden de imágenes igual a uno. En algunos casos, la condición no se cumple y la información de temporización de vídeo no señala el número de pulsos de reloj que corresponden a una diferencia de valores del recuento de orden de imágenes igual a uno. El número de pulsos de reloj puede depender de la escala temporal (correspondiente, por ejemplo, a una frecuencia de oscilador - tal como 27 MHz - que define un sistema de coordenadas de tiempo para la información señalizada) y del número de unidades de tiempo de un reloj funcionando en la escala temporal que corresponde a un incremento de un contador de pulsos de reloj, lo que se conoce como un "pulso de reloj".

En algunos ejemplos, las técnicas de esta divulgación pueden incluir señalar directamente, en la estructura sintáctica del VPS o en la parte de VUI de la estructura sintáctica del SPS para una secuencia de vídeo codificado, todos los elementos sintácticos que definen la condición para señalar un número de pulsos de reloj correspondientes a una diferencia de valores del POC igual a 1. Es decir, en algunos casos, la condición puede tener una o más dependencias de información adicional (es decir, los elementos sintácticos) señalizada en la estructura sintáctica del VPS o en la parte de VUI de la estructura sintáctica del SPS. Estos elementos sintácticos pueden incluir un indicador de información de temporización presente, según se describe más adelante.

Las normas de codificación de vídeo incluyen 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 e ITU-T H.264 (también conocida como ISO/IEC MPEG-4 AVC), incluyendo sus ampliaciones de codificación de vídeo ajustable a escala (SVC) y de codificación de vídeo de múltiples vistas (MVC).

5 Además, hay una nueva norma de codificación de vídeo, a saber, la codificación de vídeo de alta eficacia (HEVC), que está siendo desarrollada por el Equipo de Colaboración Conjunta en Codificación de Vídeo (JCT-VC) del Grupo de Expertos en Codificación de Vídeo (VCEG) de ITU-T y del Grupo de Expertos en Imágenes en Movimiento (MPEG) de ISO/IEC. El último Borrador de Trabajo (WD) de la HEVC, denominado en adelante HEVC WD9, o simplemente WD9, es el artículo de Bross et al., "Memoria Descriptiva Textual de las Mejoras en la redacción
10 propuestas para la Codificación de Vídeo de Alta Eficacia (HEVC), borrador 9 (SoDIS)", Equipo de Colaboración Conjunta en Codificación de Vídeo (JCT-VC) de la UIT-T SG 16 WP 3 e ISO/IEC JTC 1/SC 29/WG 11, 12ª conferencia: Ginebra, Suiza, 14 a 23 de enero de 2013, disponible en http://phenix.int-evry.fr/jct/doc_end_user/documents/12_Geneva/wg11/JCTVC-L0030-vl.zip, a partir del 7 de enero de 2013.

15 Un borrador reciente de la norma HEVC, denominado "Borrador 10 de trabajo de la HEVC" o "WD10", se describe en el documento JCTVC-L1003v34, de Bross et al., titulado "Memoria Descriptiva Textual de la Codificación de Vídeo de Alta Eficacia (HEVC), Borrador 10 (para Norma Internacional de Borrador Final y Última Llamada)", Equipo de Colaboración Conjunta en Codificación de Vídeo (JCT-VC) de ITU-T SG16 WP3 e ISO/IEC JTC1/SC29/WG11, 12ª Conferencia: Ginebra, Suiza, 14 a 23 de enero de 2013, que puede descargarse en: http://phenix.int-evry.fr/jct/doc_end_user/documents/12_Geneva/wg11/JCTVC-L1003-v34.zip.

Otro borrador de la norma HEVC, denominado en el presente documento "revisiones de WD10", se describe en el artículo de Bross et al., "Correcciones propuestas de los Redactores para la HEVC, versión 1", Equipo Conjunto de Colaboración en Codificación de Vídeo (JCT-VC) de la UIT-T SGI6 WP3 e ISO/IEC JTC1/SC29/WG11, 13ª Conferencia, Incheon, Corea, abril de 2013, que a partir del 7 de junio de 2013, está disponible en: http://phenix.int-evry.fr/jct/doc_end_user/documents/13_Incheon/wg11/JCTVC-M0432-v3.zip

Los esfuerzos de la normalización de la HEVC se basan en un modelo de un dispositivo de codificación de vídeo denominado modelo de prueba de la HEVC (HM). El HM prevé mejoras en las capacidades de los dispositivos de codificación de vídeo actuales con respecto a los dispositivos de codificación de vídeo disponibles durante el desarrollo de otras normas de codificación de vídeo anteriores, por ejemplo, la ITU-T H.264/AVC. Por ejemplo, mientras que la norma H.264 proporciona nueve modalidades de codificación de intra-predicción, la norma HEVC proporciona hasta treinta y cinco modalidades de codificación de intra-predicción.

35 Las normas de codificación de vídeo incluyen habitualmente una memoria descriptiva de un modelo de memoria intermedia de vídeo. En AVC y HEVC, el modelo de memoria intermedia se denomina decodificador de referencia hipotético (HRD), que incluye un modelo de memoria intermedia tanto de memoria intermedia de imágenes codificadas (CPB) como de memoria intermedia de imágenes decodificadas (DPB). Según se define en el HEVC WD9, un HRD es un modelo de decodificador hipotético que especifica restricciones sobre la variabilidad de los flujos de unidades de Capa de Abstracción de Red (NAL) o los flujos de octetos compatibles que un proceso de codificación puede producir. Los comportamientos de la CPB y la DPB se especifican matemáticamente. El HRD impone directamente restricciones sobre diferente temporización, tamaños de memoria intermedia y velocidades de bits, e impone indirectamente restricciones sobre características y estadísticas de flujos de bits. Un conjunto completo de parámetros de HRD incluye cinco parámetros básicos, el retardo inicial de eliminación de la CPB, el tamaño de la CPB, la velocidad de bits, el retardo inicial de salida del DPB y el tamaño de la DPB.

En la AVC y la HEVC, la conformidad del flujo de bits y la conformidad del decodificador se especifican como partes de la memoria descriptiva del HRD. Aunque un "Decodificador de referencia hipotético" incluye el término "decodificador", normalmente se necesita el HRD en el lado del codificador para garantizar la conformidad de flujo de bits y normalmente no se necesita en el lado del decodificador. Se especifican dos tipos de conformidad de flujo de bits o de HRD, concretamente Tipo I y Tipo II. Además, se especifican dos tipos de conformidad de decodificador, la conformidad de decodificador de la temporización de salida y la conformidad de decodificador del orden de salida.

En el HEVC WD9, las operaciones de HRD requieren parámetros señalizados en las estructuras sintácticas parámetros_hrd(), mensajes de información de mejora suplementaria (SEI) del período de almacenamiento intermedio, mensajes de SEI de temporización de imágenes y, a veces, también en mensajes de SEI de información de la unidad de decodificación. Las estructuras sintácticas parámetros_hrd() pueden señalizarse en el conjunto de parámetros de vídeo (VPS), el conjunto de parámetros de secuencia (SPS), o cualquier combinación de los mismos.

60 En el HEVC WD9, la estructura sintáctica parámetros_hrd() incluye elementos sintácticos para la señalización de información de temporización de vídeo, incluyendo la escala temporal y el número de unidades en un pulso de reloj. La parte de información de la facilidad de uso de vídeo (VUI) del SPS incluye un indicador que indica si el valor del recuento de orden de imágenes (POC), para cada imagen en la secuencia de vídeo codificado, que no sea la primera imagen en la secuencia de vídeo codificado, en el orden de decodificación, es o no proporcional al tiempo de salida de la imagen en relación con el tiempo de salida de la primera imagen en la secuencia de vídeo codificado; si es así, entonces el número de pulsos de reloj correspondientes a una diferencia de valores del recuento de orden

de imágenes es igual a 1.

La sintaxis y la semántica relacionadas en el HEVC WD9 son las siguientes. La tabla 1 muestra una estructura sintáctica ejemplar de carga útil de secuencia de octetos en bruto (RBSP) del conjunto de parámetros de vídeo, según el WD9.

5

rbsp_conjunto_de_parámetros_de_vídeo () {	Descriptor
...	
vps_núm_parámetros_hrd	ue(v)
para(i = 0; i < vps_núm_parámetros_hrd; i++) {	
índice_conjunto_po_hrd[i]	ue(v)
si (i > 0)	
indicador_cprms_presente[i]	u(1)
parámetros_hrd(indicador_cprms_presente[i], vps_máx_sub_capas_menos_1)	
}	
...	
}	

Tabla 1: Estructura sintáctica ejemplar de RBSP del conjunto de parámetros de vídeo

10 En la tabla 1 anterior, el elemento sintáctico `vps_núm_parámetros_hrd` especifica el número de estructuras
sintácticas `parámetros_hrd()` presentes en la carga útil de secuencia de octetos en bruto (RBSP) del conjunto de
parámetros de vídeo. En los flujos de bits conforme a esta versión de esta memoria descriptiva, el valor de
`vps_núm_parámetros_hrd` deberá ser menor o igual a 1. Aunque se requiere que el valor de
15 `vps_núm_parámetros_hrd` sea menor o igual a 1 en el HEVC WD9, los decodificadores permitirán que otros valores
de `vps_núm_parámetros_hrd`, en el rango de 0 hasta 1.024, inclusive, aparezcan en la sintaxis.

El elemento sintáctico `índice_conjunto_po_hrd[i]` especifica el índice, en la lista de conjuntos de puntos de operación,
especificada por el conjunto de parámetros de vídeo, del conjunto de puntos de operación al que se aplica la *i*-ésima
estructura sintáctica `parámetros_hrd()` en el conjunto de parámetros de vídeo (VPS). En los flujos de bits conforme a
20 esta versión de esta memoria descriptiva, el valor del `índice_conjunto_po_hrd[i]` será igual a 0. Aunque se requiere
que el valor del `índice_conjunto_po_hrd[i]` sea menor o igual a 1 en el HEVC WD9, los decodificadores deberán
permitir la aparición de otros valores del `índice_conjunto_po_hrd[i]`, en el rango de 0 a 1.023, en la sintaxis.

El elemento sintáctico `indicador_cprms_presente[i]` igual a 1 especifica que los parámetros de HRD que son
25 comunes para todas las sub-capas están presentes en la *i*-ésima estructura sintáctica `parámetros_hrd()` en el
conjunto de parámetros de vídeo. El `indicador_cprms_presente[i]` igual a 0 especifica que los parámetros de HRD
que son comunes para todas las sub-capas no están presentes en la *i*-ésima estructura sintáctica `parámetros_hrd()`
en el conjunto de parámetros de vídeo y que se obtienen como los mismos que los de la (*i*-1)-ésima estructura
30 sintáctica `parámetros_hrd()` en el conjunto de parámetros de vídeo. Se deduce que el `indicador_cprms_presente[0]`
es igual a 1.

La tabla 2 siguiente muestra una estructura sintáctica de parámetros de VUI según el WD9.

<code>parámetros_vui()</code> {	Descriptor
...	
indicador_parámetros_hrd_presentes	u(1)
si(indicador_parámetros_hrd_presentes)	
parámetros_hrd(1, sps_máx_sub_capas_menos1)	
indicador_poc_proporcional_a_temporización	u(1)
si(indicador_poc_proporcional_a_temporización && indicador_inf_temporización_presente)	
núm_pulsos_dif_poc_uno_menos1	ue(v)
...	
}	

Tabla 2: Estructura sintáctica de parámetros de VUI

En la Tabla 2 anterior, el elemento sintáctico indicador_parámetros_hrd_presentes igual a 1 especifica que la estructura sintáctica parámetros_hrd() está presente en la estructura sintáctica parámetros_vui(). El indicador_parámetros_hrd_presentes igual a 0 especifica que la estructura sintáctica parámetros_hrd() no está presente en la estructura sintáctica parámetros_vui().

El elemento sintáctico indicador_poc_proporcional_a_temporización igual a 1 indica que el valor del recuento de orden de imágenes para cada imagen en la secuencia de vídeo codificado, que no sea la primera imagen en la secuencia de vídeo codificado, en orden de decodificación, es proporcional al tiempo de salida de la imagen con respecto al tiempo de salida de la primera imagen en la secuencia de vídeo codificado. El indicador_poc_proporcional_a_temporización igual a 0 indica que el valor del recuento de orden de imágenes, para cada imagen en la secuencia de vídeo codificado que no sea la primera imagen en la secuencia de vídeo codificado, en orden de decodificación, puede o no ser proporcional al tiempo de salida de la imagen con respecto al tiempo de salida de la primera imagen en la secuencia de vídeo codificado.

El elemento sintáctico núm_pulsos_dif_poc_uno_menos1 más 1 especifica el número de pulsos de reloj que corresponden a una diferencia de valores del recuento de orden de imágenes igual a 1.

La Tabla 3 a continuación muestra un ejemplo de estructura sintáctica de parámetros de HRD según el WD9.

parámetros_hrd(indicadorInfComúnPresente, máxNúmSubCapasMenos1) {	Descriptor
si(indicadorInfComúnPresente) {	
indicador_inf_temporización_presente	u(1)
si(indicador_inf_temporización_presente) {	
núm_unidades_en_pulso	
escala_temporal	u(32)
}	u(32)
...	
}	

Tabla 3: Estructura sintáctica ejemplar de parámetros de HRD

En la Tabla 3 anterior, el elemento sintáctico indicador_inf_temporización_presente igual a 1 indica que el núm_unidades_en_pulso y la escala_temporal están presentes en la estructura sintáctica parámetros_hrd(). Si el indicador_inf_temporización_presente es igual a 0, el núm_unidades_en_pulso y la escala_temporal no están presentes en la estructura sintáctica parámetros_hrd(). Si no está presente, se deduce que el valor de indicador_inf_temporización_presente es 0.

El elemento sintáctico núm_unidades_en_pulso es el número de unidades de tiempo de un reloj que funciona en la frecuencia de escala_temporal Hz que corresponde a un incremento (llamado un pulso de reloj) de un contador de pulsos de reloj. El valor del elemento sintáctico núm_unidades_en_pulso será mayor que 0. Un pulso de reloj es el intervalo de tiempo mínimo que se puede representar en los datos codificados cuando el indicador_sub_imagen_cpb_parámetros_presentes es igual a 0. Por ejemplo, cuando la frecuencia de imágenes de una señal de vídeo es de 25 Hz, la escala temporal puede ser igual a 27.000.000 y el núm_unidades_en_pulso puede ser igual a 1.080.000.

El elemento sintáctico escala_temporal es el número de unidades de tiempo que pasan en un segundo. Por ejemplo, un sistema de coordenadas de tiempo que mide el tiempo usando un reloj de 27 MHz tiene una escala_temporal de 27.000.000. El valor del elemento sintáctico escala_temporal será mayor que 0.

La señalización de temporización, como se especifica en el HEVC WD9 y como se ha descrito anteriormente, puede presentar una serie de problemas. En primer lugar, la condición para la señalización del elemento sintáctico núm_pulsos_dif_poc_uno_menos1 es "si(indicador_poc_proporcional_a_temporización && indicador_inf_temporización_presente)". Esta condición implica dependencias de dos elementos sintácticos señalizados: indicador_poc_proporcional_a_temporización e indicador_inf_temporización_presente. Sin embargo, no está claro, a partir de la memoria descriptiva HEVC WD9, si el indicador_inf_temporización_presente para la condición hace referencia al elemento sintáctico indicador_inf_temporización_presente de una estructura sintáctica parámetros_hrd() (si está presente) en la parte de VUI del SPS o si hace referencia al elemento sintáctico indicador_inf_temporización_presente de una estructura sintáctica parámetros_hrd() en el VPS.

Además, múltiples capas o múltiples posibles subconjuntos de flujos de bits de un flujo de bits de vídeo ajustable a escala pueden compartir los valores comunes de la escala temporal y el número de unidades en un pulso de reloj, que se especifican en el HEVC WD9 en los elementos sintácticos `escala_temporal` y `núm_unidades_en_pulso` de una estructura sintáctica `parámetros_hrd()`, que puede señalizarse repetidamente en la parte de VUI del SPS y en el VPS, por ejemplo. Tal replicación, si está presente en el flujo de bits, puede dar lugar al desperdicio de bits.

Además, los valores del recuento de orden de imágenes (POC) son normalmente proporcionales a los tiempos de salida para todas las capas de un flujo de bits de vídeo ajustable a escala, si los valores del POC son proporcionales a los tiempos de salida para cualquiera de las capas de un flujo de bits de vídeo ajustable a escala. Sin embargo, la memoria descriptiva HEVC WD9 no prevé la señalización, en el flujo de bits de vídeo ajustable a escala, de una indicación de que los valores del POC son proporcionales a los tiempos de salida para todas las capas o todos los posibles subconjuntos de flujos de bits del flujo de bits de vídeo ajustable a escala. La referencia a un "capa" de un flujo de bits de vídeo ajustable a escala puede referirse a una capa ajustable a escala, una vista de la textura y/o una vista de la profundidad, por ejemplo. Además, aunque el HEVC WD9 especifica que el indicador `indicador_poc_proporcional_a_temporización` siempre se señala en la estructura sintáctica de la VUI del SPS, el indicador `indicador_poc_proporcional_a_temporización` no tiene ninguna utilidad si los elementos sintácticos `escala_temporal` y `núm_unidades_en_pulso` no se han señalado también en el flujo de bits.

Las técnicas de esta divulgación pueden resolver uno o más de los problemas anteriores, así como proporcionar otras mejoras, para permitir la señalización eficaz de los parámetros para las operaciones de HRD. Varios ejemplos de las técnicas se describen en el presente documento con referencia al HEVC WD9 y a posibles mejoras del mismo. Las soluciones se aplican a cualquier norma de codificación de vídeo, incluyendo AVC y HEVC, por ejemplo, que incluya una memoria descriptiva para un modelo de memoria intermedia de vídeo, aunque con fines de ilustración la descripción es específica para la señalización de parámetros de HRD definida en el HEVC WD9 y modificada según las técnicas de esta divulgación.

La FIG. 1 es un diagrama de bloques que ilustra un sistema ejemplar de codificación y decodificación de vídeo 10 que puede utilizar las técnicas descritas en esta divulgación. Como se muestra en la FIG. 1, el sistema 10 incluye un dispositivo de origen 12 que genera datos de vídeo codificado, a decodificar en un momento posterior por parte de un dispositivo de destino 14. El dispositivo de origen 12 y el dispositivo de destino 14 pueden comprender cualquiera entre una amplia gama de dispositivos, incluyendo ordenadores de sobremesa, ordenadores plegables (es decir, portátiles), ordenadores de tableta, decodificadores, equipos telefónicos de mano tales como los denominados teléfonos "inteligentes", los denominados paneles "inteligentes", televisores, cámaras, dispositivos de visualización, reproductores de medios digitales, consolas de videojuegos, un dispositivo de flujos de transmisión de vídeo o similares. En algunos casos, el dispositivo de origen 12 y el dispositivo de destino 14 pueden estar equipados para la comunicación inalámbrica.

El dispositivo de destino 14 puede recibir los datos de vídeo codificado que se van a decodificar, mediante un enlace 16. El enlace 16 puede comprender cualquier tipo de medio o dispositivo capaz de desplazar los datos de vídeo codificado desde el dispositivo de origen 12 al dispositivo de destino 14. En un ejemplo, el enlace 16 puede comprender un medio de comunicación para permitir al dispositivo de origen 12 transmitir datos de vídeo codificado directamente al dispositivo de destino 14 en tiempo real. Los datos de vídeo codificado pueden ser modulados según una norma de comunicación, tal como un protocolo de comunicación inalámbrica, y transmitidos al dispositivo de destino 14. El medio de comunicación puede comprender cualquier medio de comunicación inalámbrico o cableado, tal como un espectro de radiofrecuencia (RF) o una o más líneas de transmisión física. El medio de comunicación puede formar parte de una red basada en paquetes, tal como una red de área local, una red de área extensa o una red global tal como Internet. El medio de comunicación puede incluir encaminadores, conmutadores, estaciones base o cualquier otro equipo que pueda ser útil para facilitar la comunicación desde el dispositivo de origen 12 al dispositivo de destino 14.

Como alternativa, los datos codificados pueden ser emitidos desde la interfaz de salida 22 a un dispositivo de almacenamiento 32. De forma similar, se puede acceder a los datos codificados desde el dispositivo de almacenamiento 32 mediante una interfaz de entrada. El dispositivo de almacenamiento 32 puede incluir cualquiera entre una diversidad de medios de almacenamiento de datos, de acceso distribuido o local, tales como una unidad de disco, discos Blu-ray, discos DVD, discos CD-ROM, memoria flash, memoria volátil o no volátil, u otros medios adecuados cualesquiera de almacenamiento digital, para almacenar datos de vídeo codificado. En un ejemplo adicional, el dispositivo de almacenamiento 32 puede corresponder a un servidor de ficheros o a otro dispositivo de almacenamiento intermedio que pueda retener los datos de vídeo codificado, generados por el dispositivo de origen 12. El dispositivo de destino 14 puede acceder a datos de vídeo almacenados, procedentes del dispositivo de almacenamiento 32, mediante flujos de transmisión o descarga. El servidor de ficheros puede ser cualquier tipo de servidor capaz de almacenar datos de vídeo codificados y transmitir esos datos de vídeo codificado al dispositivo de destino 14. Entre los ejemplos de servidores de ficheros se incluyen un servidor de Internet (por ejemplo, para una sede de la Red), un servidor del FTP, dispositivos de almacenamiento conectados a la red (NAS) o una unidad de disco local. El dispositivo de destino 14 puede acceder a los datos de vídeo codificado mediante cualquier conexión de datos estándar, incluyendo una conexión a Internet. Esto puede incluir un canal inalámbrico (por ejemplo, una conexión de Wi-Fi), una conexión por cable (por ejemplo, DSL, módem de cable, etc.), o una combinación de ambos

que sea adecuada para acceder a datos de vídeo codificado, almacenados en un servidor de ficheros. La transmisión de datos de vídeo codificado desde el dispositivo de almacenamiento 34 puede ser una transmisión por flujo, una transmisión de descarga o una combinación de ambas.

5 Las técnicas de esta divulgación no están limitadas necesariamente a aplicaciones o configuraciones inalámbricas. Las técnicas pueden aplicarse a la codificación de vídeo, en soporte de cualquiera entre una diversidad de aplicaciones de multimedios, tales como difusiones de televisión por el aire, transmisiones de televisión por cable, transmisiones de televisión por satélite, transmisiones de vídeo por flujo, por ejemplo, mediante Internet, codificación de vídeo digital para su almacenamiento en un medio de almacenamiento de datos, decodificación de vídeo digital
10 almacenado en un medio de almacenamiento de datos, u otras aplicaciones. En algunos ejemplos, el sistema 10 puede configurarse para dar soporte a la transmisión de vídeo unidireccional o bidireccional, para prestar soporte a aplicaciones tales como los flujos de transmisión de vídeo, la reproducción de vídeo, la difusión de vídeo y/o la videotelefonía.

15 En el ejemplo de la FIG. 1, el dispositivo de origen 12 incluye un origen de vídeo 18, un codificador de vídeo 20 y una interfaz de salida 22. En algunos casos, la interfaz de salida 22 puede incluir un modulador/demodulador (módem) y/o un transmisor. En el dispositivo de origen 12, el origen de vídeo 18 puede incluir un origen tal como un dispositivo de captura de vídeo, por ejemplo, una videocámara, un archivo de vídeo que contiene vídeo previamente capturado, una interfaz de alimentación de vídeo para recibir vídeo desde un proveedor de contenido de vídeo y/o un sistema de gráficos de ordenador para generar datos de gráficos de ordenador como el vídeo de origen, o una combinación de tales orígenes. Como un ejemplo, si el origen de vídeo 18 es una videocámara, el dispositivo de origen 12 y el dispositivo de destino 14 pueden formar los denominados teléfonos con cámara o videoteléfonos. Sin embargo, las técnicas descritas en esta divulgación pueden aplicarse a la codificación de vídeo en general, y pueden aplicarse a aplicaciones inalámbricas y/o cableadas.

25 El vídeo capturado, pre-capturado o generado por ordenador puede ser codificado por el codificador de vídeo 20. Los datos de vídeo codificado pueden ser transmitidos directamente al dispositivo de destino 14 mediante la interfaz de salida 22 del dispositivo de origen 12. Los datos de vídeo codificado también (o como alternativa) pueden almacenarse en el dispositivo de almacenamiento 32 para un acceso posterior por el dispositivo de destino 14 u
30 otros dispositivos, para decodificación y/o reproducción.

El dispositivo de destino 14 incluye una interfaz de entrada 28, un decodificador de vídeo 30 y un dispositivo de visualización 32. En algunos casos, la interfaz de entrada 28 puede incluir un receptor y/o un módem. La interfaz de entrada 28 del dispositivo de destino 14 recibe los datos de vídeo codificado por el enlace 16. Los datos de vídeo
35 codificado, comunicados por el enlace 16, o proporcionados en el dispositivo de almacenamiento 32, pueden incluir una diversidad de elementos sintácticos generados por el codificador de vídeo 20, para su uso por un decodificador de vídeo, tal como el decodificador de vídeo 30, en la decodificación de los datos de vídeo. Tales elementos sintácticos pueden incluirse con los datos de vídeo codificado, transmitidos en un medio de comunicación, almacenarse en un medio de almacenamiento o almacenarse en un servidor de ficheros.

40 El dispositivo de visualización 32 puede estar integrado con, o ser externo a, el dispositivo de destino 14. En algunos ejemplos, el dispositivo de destino 14 puede incluir un dispositivo de visualización integrado y también configurarse para mantener interfaces 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 exhibe los datos de vídeo decodificado a un usuario, y puede comprender cualquiera entre una diversidad de dispositivos de visualización, tales como una pantalla de cristal líquido (LCD), una pantalla de plasma, una pantalla de diodos orgánicos emisores de luz (OLED) u otro tipo de dispositivo de visualización.

45 El codificador de vídeo 20 y el decodificador de vídeo 30 pueden funcionar según una norma de compresión de vídeo, tal como la norma de Codificación de Vídeo de Alta Eficacia (HEVC), actualmente en desarrollo, y pueden ser conformes al Modelo de Prueba de la HEVC (HM). Como alternativa, el codificador de vídeo 20 y el decodificador de vídeo 30 pueden funcionar según otras normas patentadas o industriales, tales como la norma ITU-T H.264, también denominada MPEG-4, Parte 10, Codificación Avanzada de Vídeo (AVC), o extensiones de tales normas. Sin embargo, las técnicas de esta divulgación no están limitadas a ninguna norma de codificación particular. Otros
50 ejemplos de normas de compresión de vídeo incluyen la MPEG-2 y la ITU-T H.263.

Aunque no se muestra en la FIG. 1, en algunos aspectos, el codificador de vídeo 20 y el decodificador de vídeo 30 pueden estar integrados con un codificador y decodificador de audio, y pueden incluir unidades adecuadas de multiplexado y demultiplexado, u otro hardware y software, para llevar a cabo la codificación, tanto de audio como de
60 vídeo, en un flujo de datos común o en flujos de datos diferentes. Si procede, en algunos ejemplos, las unidades de multiplexado y demultiplexado pueden ajustarse al protocolo de multiplexado ITU H.223 o a otros protocolos, tales como el protocolo de datagramas de usuario (UDP).

65 El codificador de vídeo 20 y el decodificador de vídeo 30 pueden implementarse como cualquiera entre una diversidad de circuitos codificadores adecuados, tales como uno o más microprocesadores, procesadores de señales digitales (DSP), circuitos integrados específicos de la aplicación (ASIC), formaciones de compuertas

programables en el terreno (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 adecuado, legible por ordenador, y ejecutar las instrucciones en hardware usando uno o más procesadores para realizar las técnicas de esta divulgación. Tanto el codificador de vídeo 20 como el decodificador de vídeo 30 pueden incluirse en uno o más codificadores o decodificadores, donde cualquiera de los mismos puede estar integrado como parte de un codificador/decodificador combinado (códec) en un dispositivo respectivo.

El JCT-VC está trabajando en el desarrollo de la norma HEVC. Los esfuerzos de normalización de la HEVC se basan en un modelo en evolución de un dispositivo de codificación de vídeo denominado modelo de prueba de la HEVC (HM). El HM supone varias capacidades adicionales de los dispositivos de codificación de vídeo, con respecto a dispositivos existentes, según, por ejemplo, la ITU-T H.264/AVC. Por ejemplo, mientras que la norma H.264 proporciona nueve modalidades de codificación de intra-predicción, el HM puede proporcionar hasta treinta y tres modalidades de codificación de intra-predicción.

En general, el modelo de funcionamiento del HM describe que una trama o imagen de vídeo puede dividirse en una secuencia de bloques arbolados, o unidades máximas de codificación (LCU), que incluyen muestras tanto de luma como de croma. Un bloque arbolado tiene un fin similar al de un macro-bloque de la norma H.264. Un fragmento incluye un cierto número de bloques arbolados consecutivos en orden de codificación. Una trama o imagen de vídeo puede dividirse en uno o más fragmentos. Cada bloque arbolado puede dividirse en unidades de codificación (CU) según un árbol cuádruple. Por ejemplo, un bloque arbolado, como un nodo raíz del árbol cuádruple, puede dividirse en cuatro nodos secundarios, y cada nodo secundario puede a su vez ser un nodo principal y dividirse en otros cuatro nodos secundarios. Un nodo secundario final, no dividido, como un nodo de hoja del árbol cuádruple, comprende un nodo de codificación, es decir, un bloque de vídeo codificado. Los datos sintácticos asociados a un flujo de bits codificado pueden definir un número máximo de veces que puede dividirse un bloque arbolado, y también pueden definir un tamaño mínimo de los nodos de codificación.

Una CU incluye un nodo de codificación y unidades de predicción (PU) y unidades de transformación (TU) asociadas al nodo de codificación. Un tamaño de la CU corresponde, en general, a un tamaño del nodo de codificación y normalmente debe tener una forma cuadrada. El tamaño de la CU puede variar desde 8 x 8 píxeles hasta el tamaño del bloque arbolado, con un máximo de 64 x 64 píxeles, o más. Cada CU puede contener una o más PU y una o más TU. Los datos sintácticos asociados a una CU pueden describir, por ejemplo, la división de la CU en una o más PU. Las modalidades de división pueden diferir entre si la CU está codificada en modalidad de omisión o directa, codificada en modo de intra-predicción o codificada en modo de inter-predicción. Las PU pueden dividirse para que no tengan forma cuadrada. Los datos sintácticos asociados a una CU también pueden describir, por ejemplo, la división de la CU en una o más TU según un árbol cuádruple. Una TU puede tener forma cuadrada o no cuadrada.

La norma de HEVC admite transformaciones según las TU, que pueden ser distintas para distintas CU. Las TU son habitualmente dimensionadas en base al tamaño de las PU dentro de una CU dada, definida para una LCU dividida, aunque esto puede no ser siempre el caso. Las TU tienen habitualmente el mismo tamaño, o un tamaño más pequeño, que las PU. En algunos ejemplos, las muestras residuales correspondientes a una CU pueden subdividirse en unidades más pequeñas usando una estructura de árbol cuádruple conocida como "árbol cuádruple residual" (RQT). Los nodos de hoja del RQT pueden denominarse unidades de transformación (TU). Los valores de diferencias de píxeles asociados a las TU pueden transformarse para producir coeficientes de transformación, que pueden cuantizarse.

En general, una PU incluye datos relacionados con el proceso de predicción. Por ejemplo, cuando la PU está codificada en intra-modalidad, la PU puede incluir datos que describen una modalidad de intra-predicción para la PU. Como otro ejemplo, cuando la PU está codificada en inter-modalidad, la PU puede incluir datos que definen un vector de movimiento para la PU. Los datos que definen el vector de movimiento para una PU pueden describir, por ejemplo, un componente horizontal del vector de movimiento, un componente vertical del vector de movimiento, una resolución para el vector de movimiento (por ejemplo, precisión de un cuarto de píxel o precisión de un octavo de píxel), una imagen de referencia a la que apunta el vector de movimiento y/o una lista de imágenes de referencia (por ejemplo, la Lista 0, la Lista 1 o la Lista C) para el vector de movimiento.

En general, se usa una TU para los procesos de transformación y cuantización. Una CU dada con una o más PU también puede incluir una o más unidades de transformación (TU). Después de la predicción, el codificador de vídeo 20 puede calcular los valores residuales a partir del bloque de vídeo identificado por el nodo de codificación, según la PU. El nodo de codificación se actualiza a continuación, para hacer referencia a los valores residuales, en lugar del bloque de vídeo original. Los valores residuales comprenden valores de diferencias de píxeles que pueden transformarse en coeficientes de transformación, cuantizarse y escanearse usando las transformaciones y otra información de transformación especificada en las TU, para producir coeficientes de transformación serializados para la codificación por entropía. El nodo de codificación puede, una vez más, ser actualizado para referirse a estos coeficientes de transformación serializados. Esta divulgación usa habitualmente el término "bloque de vídeo" para referirse a un nodo de codificación de una CU. En algunos casos específicos, esta divulgación también puede usar el término "bloque de vídeo" para referirse a un bloque arbolado, es decir, una LCU, o una CU, que incluye un nodo

de codificación y las PU y las TU.

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

Como ejemplo, el HM da soporte a la predicción en diversos tamaños de PU. Suponiendo que el tamaño de una CU específica sea $2N \times 2N$, el HM presta soporte a la intra-predicción en tamaños de PU de $2N \times 2N$ o $N \times N$, y a la intra-predicción en tamaños de PU simétricos de $2N \times 2N$, $2N \times N$, $N \times 2N$ o $N \times N$. El HM también presta soporte a la división asimétrica para la inter-predicción en tamaños de PU de $2N \times nU$, $2N \times nD$, $nL \times 2N$ y $nR \times 2N$. En la división asimétrica, una dirección de una CU no está dividida, mientras que la otra dirección está dividida entre el 25 % y el 75 %. La parte de la CU correspondiente a la división del 25 % está indicada por una "n" seguida por una indicación de "Arriba", "Abajo", "Izquierda" o "Derecha". Así, por ejemplo, "2NxnU" se refiere a una CU de tamaño $2N \times 2N$ que está dividida horizontalmente, con una PU de tamaño $2N \times 0,5N$ encima y una PU de tamaño $2N \times 1,5N$ debajo.

En esta divulgación, "NxN" y "N por N" pueden usarse de manera intercambiable para hacer referencia a las dimensiones de píxeles de un bloque de vídeo, en lo que respecta a dimensiones verticales y horizontales, por ejemplo, 16×16 píxeles o 16 por 16 píxeles. En general, un bloque de tamaño 16×16 tendrá 16 píxeles en la dirección vertical ($y = 16$) y 16 píxeles en la dirección horizontal ($x = 16$). Asimismo, un bloque de tamaño NxN tiene generalmente N píxeles en la dirección vertical y N píxeles en la dirección horizontal, donde N representa un valor entero no negativo. Los píxeles en un bloque pueden estar dispuestos en filas y columnas. Además, los bloques no necesitan tener necesariamente el mismo número de píxeles en la dirección horizontal y en la dirección vertical. Por ejemplo, los bloques pueden comprender $N \times M$ píxeles, donde M no es necesariamente igual a N.

Tras la codificación intra-predictiva o inter-predictiva, usando las PU de una CU, el codificador de vídeo 20 puede calcular datos residuales a los que se aplican las transformaciones especificadas por las TU de la CU. Los datos residuales pueden corresponder a diferencias de píxeles entre píxeles de la imagen no codificada y valores de predicción correspondientes a las CU. El codificador de vídeo 20 puede formar los datos residuales para la CU, y a continuación transformar los datos residuales para producir coeficientes de transformación.

Tras cualquier transformación para producir coeficientes de transformación, el codificador de vídeo 20 puede realizar la cuantización de los coeficientes de transformación. La cuantización se refiere generalmente a un proceso en el que los coeficientes de transformación se cuantizan para reducir posiblemente la cantidad de datos usados para representar los coeficientes, proporcionando una compresión adicional. El proceso de cuantización puede reducir la profundidad de bits asociada a algunos de, o todos, los coeficientes. Por ejemplo, un valor de n bits puede redondearse por lo bajo a un valor de m bits durante la cuantización, donde n es mayor que m .

En algunos ejemplos, el codificador de vídeo 20 puede utilizar un orden de escaneo predefinido para escanear los coeficientes de transformación cuantizados, para producir un vector serializado que pueda ser codificado por entropía. En otros ejemplos, el codificador de vídeo 20 puede realizar un escaneo adaptativo. Después de escanear los coeficientes de transformación cuantizados, para formar un vector unidimensional, el codificador de vídeo 20 puede codificar por entropía el vector unidimensional, por ejemplo, según la codificación de longitud variable adaptativa al contexto (CAVLC), la codificación aritmética binaria adaptativa al contexto (CABAC), la codificación aritmética binaria adaptativa al contexto y de base sintáctica (SBAC), la codificación por Entropía por División de Intervalos de Probabilidad (PIPE) u otra metodología de codificación por entropía. El codificador de vídeo 20 también puede codificar por entropía elementos sintácticos asociados a los datos de vídeo codificado, para su uso por el decodificador de vídeo 30, en la decodificación de los datos de vídeo.

Para realizar la CABAC, el codificador de vídeo 20 puede asignar un contexto dentro de un modelo contextual a un símbolo a transmitir. El contexto puede referirse, por ejemplo, a si los valores adyacentes del símbolo son distintos a cero o no. Para realizar la CAVLC, el codificador de vídeo 20 puede seleccionar un código de longitud variable para un símbolo a transmitir. Las palabras de código en la VLC pueden ser construidas de modo que los códigos relativamente más cortos correspondan a símbolos más probables, mientras que los códigos más largos correspondan a símbolos menos probables. De esta manera, el uso de VLC puede lograr un poco de ahorro con respecto, por ejemplo, a usar palabras de código de igual longitud para cada símbolo a transmitir. La determinación de la probabilidad puede basarse en un contexto asignado al símbolo.

El dispositivo de origen 12 puede generar un flujo de bits codificado para incluir elementos sintácticos que se ajustan a una estructura sintáctica según técnicas descritas en esta divulgación. En algunos ejemplos, el codificador de vídeo 20 puede generar el flujo de bits codificado para señalar directamente, en la estructura sintáctica del

conjunto de parámetros de vídeo (VPS) o en la parte de información sobre facilidad de uso de vídeo (VUI) de la estructura sintáctica del conjunto de parámetros de secuencia (SPS) para una secuencia de vídeo codificado, todas las variables que definen la condición para señalar una serie de pulsos de reloj que corresponden a una diferencia de valores del recuento de orden de imágenes (POC) igual a 1. En otras palabras, en lugar de señalar los elementos sintácticos para la condición para señalar una serie de pulsos de reloj que corresponden a una diferencia de valores del recuento de orden de imágenes (POC) igual a 1 en otra estructura sintáctica (tal como la estructura sintáctica de parámetros de HRD) que está incorporada en una estructura sintáctica del VPS o en la parte de VUI de la estructura sintáctica del SPS, el codificador de vídeo 20 genera el flujo de bits codificado para señalar los elementos sintácticos que definen la condición en el VPS y/o la estructura sintáctica de VUI sin referencia a otra estructura sintáctica potencialmente incorporada dentro de las estructuras sintácticas del VPS y/o de la VUI. Los elementos sintácticos pueden incluir el elemento sintáctico indicador_inf_temporización_presente que se especifica en el HEVC WD9 como un elemento sintáctico de la estructura sintáctica de parámetros de HRD. Como resultado, las técnicas pueden reducir y, potencialmente, eliminar la ambigüedad dentro de la memoria descriptiva HEVC WD9, especificando claramente en la sintaxis el origen de los elementos sintácticos que definen la condición.

El codificador de vídeo 20 puede probar el flujo de bits codificado en cuanto a su conformidad con los requisitos especificados, como una o más pruebas de conformidad de flujo de bits, definidas en una memoria descriptiva de codificación de vídeo, tal como el HEVC WD9 o una memoria descriptiva sucesora, tal como el HEVC WD10. El codificador de vídeo 20 puede incluir o utilizar de otro modo un decodificador de referencia hipotético para probar el flujo de bits codificado en cuanto a su conformidad. Según las técnicas descritas en este documento, el codificador de vídeo 20 puede probar el flujo de bits codificado en cuanto a su conformidad, mediante la decodificación del flujo de bits codificado para determinar, a partir de la estructura sintáctica del VPS o en la parte de VUI de la estructura sintáctica del SPS para una secuencia de vídeo codificado, los elementos sintácticos que definen la condición para la señalización del número de pulsos de reloj que corresponden a una diferencia de valores del POC igual a 1. Si se cumple la condición según los valores de los elementos sintácticos, el codificador de vídeo 20 puede determinar el número de pulsos de reloj que corresponden a una diferencia de valores del POC igual a 1, y utilizar el número determinado de pulsos de reloj como entrada para, por ejemplo, la determinación del desborde inferior o superior de la CPB durante la decodificación de imágenes codificadas incluidas en el flujo de bits codificado.

En algunas ocasiones, en el dispositivo de destino 14, un decodificador de vídeo 30 bajo prueba (o VUT) puede, en algunos casos, recibir una representación del flujo de bits codificado, generado por el codificador de vídeo 20, para señalar directamente, en la estructura sintáctica del VPS o en la parte de VUI de la estructura sintáctica del SPS para una secuencia de vídeo codificado, todos los elementos sintácticos que definen la condición para la señalización de una serie de pulsos de reloj que corresponden a una diferencia de valores del recuento de orden de imágenes (POC) igual a 1. El decodificador de vídeo 30 puede decodificar el flujo de bits codificado para determinar, a partir de la estructura sintáctica del VPS o en la parte de VUI de la estructura sintáctica del SPS para una secuencia de vídeo codificado, los elementos sintácticos que definen la condición para la señalización del número de pulsos de reloj que corresponden a una diferencia de valores del POC igual a 1. Si la condición se cumple según los valores de los elementos sintácticos, el decodificador de vídeo 30 puede determinar el número de pulsos de reloj que corresponden a una diferencia de valores del POC igual a 1, y utilizar el número determinado de pulsos de reloj como entrada para, por ejemplo, la determinación del desborde inferior o superior de la CPB durante la decodificación de imágenes codificadas incluidas en el flujo de bits codificado.

En algunos ejemplos, el codificador de vídeo 20 puede generar el flujo de bits codificado para señalar la escala temporal y el número de unidades en un pulso de reloj, como máximo, una vez en cada una de las estructuras sintácticas del VPS y de la VUI para una secuencia dada de vídeo codificado. Es decir, en una estructura sintáctica del VPS dada para el flujo de bits codificado, el codificador de vídeo 20 puede señalar los elementos sintácticos de la escala temporal y del número de unidades en un pulso de reloj, como máximo, una vez. Del mismo modo, en una estructura sintáctica dada de VUI (por ejemplo, la parte de VUI de una estructura sintáctica del SPS) para el flujo de bits codificado, el codificador de vídeo 20 puede señalar los elementos sintácticos de la escala temporal y del número de unidades en un pulso de reloj, como máximo, una vez. Como resultado, el codificador de vídeo 20 que funciona según las técnicas descritas en este documento puede reducir un número de casos del elemento sintáctico de escala temporal (escala_temporal según el WD9) y del elemento sintáctico del número de unidades en un pulso de reloj (núm_unidades_en_pulso según el WD9) en el flujo de bits codificado. Además, el codificador de vídeo 20 puede, en algunos casos, generar el flujo de bits codificado para señalar directamente la escala temporal y el número de unidades en un pulso de reloj en cada una de las estructuras sintácticas del VPS y de la VUI para una secuencia dada de vídeo codificado, en lugar de en una estructura sintáctica de parámetros de HRD, incorporada dentro de una estructura sintáctica del VPS y/o de la VUI.

Según las técnicas descritas en este documento, el codificador de vídeo 20 puede probar un flujo de bits codificado, generado por el codificador de vídeo 20, para señalar la escala temporal y el número de unidades en un pulso de reloj, como máximo, una vez en cada una de las estructuras sintácticas del VPS y de la VUI para una secuencia dada de vídeo codificado, en cuanto a su conformidad, mediante la decodificación del flujo de bits codificado para determinar la escala temporal y el número de unidades en un pulso de reloj, a partir de una estructura sintáctica del VPS del flujo de bits codificado que codifica los elementos sintácticos de la escala temporal y del número de unidades en un pulso de reloj, como máximo, una vez en la estructura sintáctica del VPS. En algunos casos, el

codificador de vídeo 20 puede probar el flujo de bits codificado en cuanto a su conformidad, mediante la decodificación del flujo de bits codificado para determinar la escala temporal y el número de unidades en un pulso de reloj a partir de una estructura sintáctica de VUI del flujo de bits codificado que codifica los elementos sintácticos de la escala temporal y del número de unidades en un pulso de reloj, como máximo, una vez en la estructura sintáctica de la VUI. La escala temporal y el número de unidades en un pulso de reloj se pueden señalar en un lugar que no sea una estructura sintáctica de parámetros de HRD, incorporada dentro de la estructura sintáctica del VPS y/o de la VUI. El codificador de vídeo 20 puede usar la escala temporal determinada y el número determinado de unidades en un pulso de reloj como entrada para, por ejemplo, la determinación del desborde inferior o superior de la CPB durante la decodificación de imágenes codificadas incluidas en el flujo de bits codificado.

En algunas ocasiones, en el dispositivo de destino 14, un decodificador de vídeo 30 bajo prueba puede, en algunos casos, recibir una representación del flujo de bits codificado, generado por el codificador de vídeo 20 para señalar la escala temporal y el número de unidades en un pulso de reloj, como máximo, una vez en cada una de las estructuras sintácticas del VPS y de la VUI para una secuencia dada de vídeo codificado. El decodificador de vídeo 30 puede decodificar el flujo de bits codificado para determinar la escala temporal y el número de unidades en un pulso de reloj a partir de una estructura sintáctica del VPS del flujo de bits codificado que codifica los elementos sintácticos de la escala temporal y del número de unidades en un pulso de reloj, como máximo, una vez en la estructura sintáctica del VPS. En algunos casos, el decodificador de vídeo 30 puede probar el flujo de bits codificado en cuanto a su conformidad, mediante la decodificación del flujo de bits codificado para determinar la escala temporal y el número de unidades en un pulso de reloj a partir de una estructura sintáctica de VUI del flujo de bits codificado, que codifica los elementos sintácticos de la escala temporal y del número de unidades en un pulso de reloj, como máximo, una vez en la estructura sintáctica de VUI. La escala temporal y el número de unidades en un pulso de reloj se pueden señalar en un lugar que no sea una estructura sintáctica de parámetros de HRD, incorporada dentro de la estructura sintáctica del VPS y/o de la VUI. El decodificador de vídeo 30 puede usar la escala temporal determinada y el número determinado de unidades en un pulso de reloj como entrada para, por ejemplo, la determinación del desborde inferior o superior de la CPB durante la decodificación de imágenes codificadas incluidas en el flujo de bits codificado.

En algunos ejemplos, el codificador de vídeo 20 puede generar el flujo de bits codificado para señalar, en una estructura sintáctica del VPS para una o más secuencias de vídeo codificado, un indicador que indica si el valor del POC para cada imagen en una secuencia de vídeo codificado, que no sea la primera imagen en la secuencia de vídeo codificado, en orden de decodificación, es o no proporcional al tiempo de salida de la imagen en relación con el tiempo de salida de la primera imagen en la secuencia de vídeo codificado. Este indicador puede, alternativamente, denominarse el indicador de POC proporcional a la indicación de temporización. Como resultado, el codificador de vídeo 20 puede reducir un número de casos de la indicación en la información de temporización señalizada para múltiples capas de una secuencia de vídeo codificado y/o para un flujo de bits de vídeo ajustable a escala que tenga múltiples capas. En algunos casos, el codificador de vídeo 20 puede incluir este indicador en la estructura sintáctica del VPS solo si también se incluyen los elementos sintácticos de la escala temporal y del número de unidades en un pulso de reloj. El codificador de vídeo 20 puede, de esta manera, evitar la señalización de esta información de temporización particular (es decir, si el valor del POC para cada imagen en una secuencia de vídeo codificado, que no sea la primera imagen en la secuencia de vídeo codificado, en orden de decodificación, es o no proporcional al tiempo de salida de la imagen en relación con el tiempo de salida de la primera imagen en la secuencia de vídeo codificado) si la información de pulso de reloj necesaria para usar la indicación de POC proporcional a la temporización tampoco está presente.

Según las técnicas descritas en este documento, el codificador de vídeo 20 puede comprobar la conformidad de un flujo de bits codificado, generado por el codificador de vídeo 20, para señalar, en una estructura sintáctica del VPS para una o más secuencias de vídeo codificado, el indicador de POC proporcional a indicación de temporización. El codificador de vídeo 20 puede probar el flujo de bits codificado en cuanto a su conformidad, mediante la decodificación del flujo de bits codificado, para determinar un valor para el indicador. El codificador de vídeo 20 puede además, o alternativamente, probar un flujo de bits codificado, generado por el codificador de vídeo 20, para señalar el indicador en la estructura sintáctica del VPS solo si también se incluyen los elementos sintácticos de la escala temporal y del número de unidades en un pulso de reloj. El codificador de vídeo 20 puede utilizar el valor determinado del indicador de POC proporcional a la indicación de temporización, y los elementos sintácticos de la escala temporal y del número de unidades en un pulso de reloj, como entrada para, por ejemplo, la determinación del desborde inferior o superior de la CPB durante la decodificación de imágenes codificadas incluidas en el flujo de bits codificado.

En algunas ocasiones, en el dispositivo de destino 14, un decodificador de vídeo 30 bajo prueba, en algunos casos, puede recibir una representación del flujo de bits codificado, generado por el codificador de vídeo 20, para señalar, en una estructura sintáctica del VPS para una o más secuencias de vídeo codificado, un indicador de POC proporcional a la indicación de temporización. El decodificador de vídeo 30 puede probar el flujo de bits codificado en cuanto a su conformidad, mediante la decodificación del flujo de bits codificado, para determinar un valor para el indicador. El decodificador de vídeo 30 puede, adicionalmente, o alternativamente, probar un flujo de bits codificado, generado por el decodificador de vídeo 30, para señalar el indicador en la estructura sintáctica del VPS, solo si también se incluyen los elementos sintácticos de la escala temporal y del número de unidades en un pulso de reloj.

El decodificador de vídeo 30 puede utilizar el valor determinado del indicador de POC proporcional a la indicación de temporización, y los elementos sintácticos de la escala temporal y del número de unidades en un pulso de reloj, como entrada para, por ejemplo, la determinación del desborde inferior o superior de la CPB durante la decodificación de imágenes codificadas incluidas en el flujo de bits codificado.

La FIG. 2 es un diagrama de bloques que ilustra un codificador ejemplar de vídeo 20 que puede implementar las técnicas descritas en esta divulgación. El codificador de vídeo 20 puede realizar la intra-codificación y la inter-codificación de bloques de vídeo dentro de fragmentos de vídeo. La intra-codificación se apoya en la predicción espacial para reducir o eliminar la redundancia espacial en el vídeo dentro de una trama o imagen de vídeo dada. La inter-codificación se apoya en la predicción temporal para reducir o eliminar la redundancia temporal en el vídeo dentro de tramas o imágenes adyacentes de una secuencia de vídeo. La intra-modalidad (modalidad I) puede referirse a cualquiera de varias modalidades de compresión de base espacial. Las inter-modalidades, tales como la predicción bi-direccional (modalidad P) o la bi-predicción (modalidad B), pueden referirse a cualquiera de varias modalidades de compresión de base temporal.

En el ejemplo de la FIG. 2, el codificador de vídeo 20 incluye una unidad de división 35, un módulo de predicción 41, una memoria de imágenes de referencia 64, un sumador 50, un módulo de transformación 52, una unidad de cuantización 54 y una unidad de codificación por 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 intra-predicción 46. Para la reconstrucción de bloques de vídeo, el codificador de vídeo 20 incluye además una unidad de cuantización inversa 58, un módulo de transformación inversa 60 y un sumador 62. También puede incluirse un filtro de desbloqueo (no mostrado en la FIG. 2) para filtrar límites de bloque, para eliminar distorsiones de efecto pixelado del vídeo reconstruido. Si se desea, el filtro de desbloqueo filtrará habitualmente la salida del sumador 62. También pueden usarse filtros de bucle adicionales (en bucle o tras el bucle), además del filtro de desbloqueo.

Como se muestra en la FIG. 2, el codificador de vídeo 20 recibe datos de vídeo y la unidad de división 35 divide los datos en bloques de vídeo. Esta división también puede incluir la división en fragmentos, mosaicos u otras unidades mayores, así como la división de bloques de vídeo, por ejemplo, según una estructura de árbol cuádruple de las LCU y las CU. El codificador de vídeo 20 ilustra en general los componentes que codifican bloques de vídeo dentro de un fragmento de vídeo a codificar. El fragmento puede dividirse en múltiples bloques de vídeo (y, posiblemente, en conjuntos de bloques de vídeo mencionados como mosaicos). El módulo de predicción 41 puede seleccionar una entre una pluralidad de posibles modalidades de codificación, tal como una entre una pluralidad de modalidades de intra-codificación, o una entre una pluralidad de modalidades de inter-codificación, para el bloque de vídeo actual, en base a resultados de errores (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 para su uso como una imagen de referencia.

El módulo de intra-predicción 46, dentro del módulo de predicción 41, puede realizar la codificación intra-predictiva del bloque de vídeo actual con respecto a uno o más bloques vecinos en la misma trama o fragmento que el bloque que va a codificarse, para proporcionar compresión espacial. La unidad de estimación de movimiento 42 y la unidad de compensación de movimiento 44, dentro del módulo de predicción 41, realizan la codificación inter-predictiva del bloque de vídeo actual con respecto a uno o más bloques predictivos en una o más imágenes de referencia, para proporcionar compresión temporal.

La unidad de estimación de movimiento 42 puede ser configurada para determinar la modalidad de inter-predicción para un fragmento de vídeo, según un patrón predeterminado para una secuencia de vídeo. El patrón predeterminado puede designar fragmentos de vídeo en la secuencia como fragmentos P, fragmentos B o fragmentos GPB. La unidad de estimación de movimiento 42 y la unidad de compensación de movimiento 44 pueden estar sumamente integradas, pero se ilustran por separado con fines conceptuales. La estimación de movimiento, realizada por la unidad de estimación de movimiento 42, es el proceso de generación de vectores de movimiento, que estiman el movimiento para los bloques de vídeo. Un vector de movimiento, por ejemplo, puede indicar el desplazamiento de una PU de un bloque de vídeo dentro de una trama o imagen de vídeo actual, con respecto a un bloque predictivo dentro de una imagen de referencia.

Un bloque predictivo es un bloque del que se descubre que se corresponde estrechamente con la PU del bloque de vídeo a codificar, en lo que respecta a la diferencia de píxeles, que puede determinarse mediante la suma de una diferencia absoluta (SAD), una suma de diferencia de cuadrados (SSD) u otras métricas de diferencia. En algunos ejemplos, el codificador de vídeo 20 puede calcular valores para posiciones de fracciones de píxel de imágenes de referencia almacenadas en la memoria de imágenes de referencia 64. Por ejemplo, el codificador de vídeo 20 puede interpolar valores de posiciones de un cuarto de píxel, posiciones de un octavo de píxel u otras posiciones fraccionarias de píxel de la imagen de referencia. Por lo tanto, la unidad de estimación de movimiento 42 puede realizar una búsqueda de movimiento con respecto a las posiciones de píxeles completos y a las posiciones de fracciones de píxel, y emitir un vector de movimiento con una precisión de fracciones de píxel.

La unidad de estimación de movimiento 42 calcula un vector de movimiento para una PU de un bloque de vídeo en

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

La compensación de movimiento, realizada por la unidad de compensación de movimiento 44, puede implicar extraer o generar el bloque predictivo en base al vector de movimiento determinado por la estimación de movimiento, realizando posiblemente interpolaciones hasta la precisión de sub-píxeles. Tras recibir el vector de movimiento para la PU del bloque de vídeo actual, la unidad de compensación de movimiento 44 puede localizar el bloque predictivo al que apunta el vector de movimiento en una de las listas de imágenes de referencia. El codificador de vídeo 20 forma un bloque de vídeo residual restando los valores de píxeles del bloque predictivo a los valores de píxeles del bloque de vídeo actual que está codificándose, generando valores de diferencia de píxel. Los valores de diferencia de píxel forman datos residuales para el bloque, y pueden incluir componentes de diferencias, tanto de luma como de croma. El sumador 50 representa el componente o los componentes que realizan esta operación de sustracción. La unidad de compensación de movimiento 44 también puede generar elementos sintácticos 55 asociados a los bloques de vídeo y al fragmento de vídeo para su uso por parte del decodificador de vídeo 30 en la decodificación de los bloques de vídeo del fragmento de vídeo.

La unidad de compensación de movimiento 44 puede generar elementos sintácticos 55 que se ajustan a una estructura sintáctica según técnicas descritas en esta divulgación. En algunos ejemplos, el codificador de vídeo 20 puede generar elementos sintácticos 55 para señalar directamente, en la estructura sintáctica del conjunto de parámetros de vídeo (VPS) o en la parte de la información de facilidad de uso de vídeo (VUI) de la estructura sintáctica del conjunto de parámetros de secuencia (SPS), asociada a los bloques de vídeo, todos los elementos sintácticos que definen la condición para señalar una serie de pulsos de reloj que corresponden a una diferencia de valores del recuento de orden de imágenes (POC) igual a 1. En otras palabras, en lugar de señalar los elementos sintácticos para la condición para señalar una serie de pulsos de reloj que corresponden a una diferencia de valores del recuento de orden de imágenes (POC) igual a 1 en otra estructura sintáctica (tal como la estructura sintáctica de parámetros de HRD) que está incorporada en una estructura sintáctica del VPS o en la parte de VUI de la estructura sintáctica del SPS, la unidad de compensación de movimiento 44 genera el flujo de bits codificado para señalar los elementos sintácticos para los elementos sintácticos que definen la condición en la estructura sintáctica del VPS y/o de la VUI, sin referencia a otra estructura sintáctica potencialmente incorporada dentro de cualquiera de las estructuras sintácticas del VPS y de la VUI, o de ambas.

En algunos ejemplos, la unidad de compensación de movimiento 44 puede generar los elementos sintácticos 55 para señalar la escala temporal y el número de unidades en un pulso de reloj, como máximo, una vez en cada una de las estructuras sintácticas del VPS y la VUI para una secuencia dada de vídeo codificado. Es decir, en una estructura sintáctica dada del VPS para el flujo de bits codificado, la unidad de compensación de movimiento 44 puede generar los elementos sintácticos 55 para señalar los elementos sintácticos de la escala temporal y del número de unidades en un pulso de reloj, como máximo, una vez. Del mismo modo, en una estructura sintáctica dada de VUI (por ejemplo, la parte de VUI de una estructura sintáctica del SPS) para el flujo de bits codificado, la unidad de compensación de movimiento 44 puede generar los elementos sintácticos 55 para señalar los elementos sintácticos de la escala temporal y del número de unidades en un pulso de reloj, como máximo, una vez. Además, la unidad de compensación de movimiento 44 puede, en algunos casos, generar los elementos sintácticos 55 para señalar directamente la escala temporal y el número de unidades en un pulso de reloj en cada una de las estructuras sintácticas del VPS y de la VUI para una secuencia dada de vídeo codificado, en lugar de en una estructura sintáctica de parámetros de HRD, incorporada dentro de una estructura sintáctica del VPS y / o la VUI.

En algunos ejemplos, la unidad de compensación de movimiento 44 puede generar los elementos sintácticos 55 para señalar, en una estructura sintáctica del VPS para una o más secuencias de vídeo codificado, un indicador que indica si el valor del POC para cada imagen en una secuencia de vídeo codificado, que no sea la primera imagen en la secuencia de vídeo codificada, en orden de decodificación, es o no proporcional al tiempo de salida de la imagen en relación con el tiempo de salida de la primera imagen en la secuencia de vídeo codificado. Este indicador, alternativamente, puede denominarse el indicador de POC proporcional a la indicación de temporización. Como resultado, la unidad de compensación de movimiento 44 puede reducir un número de casos de la indicación en la información de temporización señalizada para múltiples capas de una secuencia de vídeo codificado y/o para un flujo de bits de vídeo ajustable a escala que tiene múltiples capas. En algunos casos, la unidad de compensación de movimiento 44 puede incluir este indicador en la estructura sintáctica del VPS, solo si también se incluyen los elementos sintácticos de la escala temporal y del número de unidades en un pulso de reloj. La unidad de compensación de movimiento 44 puede, de esta manera, evitar la señalización de esta información de temporización particular (es decir, si el valor del POC para cada imagen en una secuencia de vídeo codificado, que no sea la primera imagen en la secuencia de vídeo codificado, en orden de decodificación, es o no proporcional al tiempo de salida de la imagen en relación con el tiempo de salida de la primera imagen en la secuencia de vídeo codificado) si la información de pulsos de reloj necesaria para utilizar la indicación de POC proporcional a la temporización tampoco está presente.

Algunos cambios ejemplares en el texto del HEVC WD9 para efectuar las técnicas anteriores para generar elementos sintácticos 55 son los siguientes (otras partes no mencionadas pueden quedar sin modificarse según el HEVC WD9):

El siguiente es un ejemplo de una estructura sintáctica de RBSP del conjunto de parámetros de vídeo, modificada para resolver uno o más de los problemas anteriores (la sintaxis subrayada es una adición a la estructura sintáctica de la RBSP del conjunto de parámetros de vídeo del HEVC WD9; otra sintaxis puede estar sin cambios en relación con el HEVC WD9):

rbbsp_conjunto_de_parámetros_de_vídeo () {	Descriptor
...	
<u>vps indicador inf temporización presente</u>	u(1)
si(vps indicador inf temporización presente) {	
<u>vps núm unidades en pulso</u>	u(32)
<u>vps escala temporal</u>	u(32)
<u>vps indicador poc proporcional a temporización</u>	ue(1)
si (vps indicador poc proporcional a temporización)	
<u>vps núm pulsos dif poc uno menos1</u>	ue(v)
}	
<u>vps núm parámetros hrd</u>	ue(v)
para(i = 0; i < vps_núm_parámetros_hrd; i++) {	
<u>indicador_cprms_presente[i]</u>	u(1)
parámetros_hrd(indicador_cprms_presente[i], vps_máx_sub_capas_menos1)	
}	
...	
}	

Tabla 4: Ejemplo de estructura sintáctica de RBSP del conjunto de parámetros de vídeo

La tabla 4 define los elementos sintácticos agregados recientemente según la semántica siguiente de la RBSP del conjunto de parámetros de vídeo (VPS):

El vps indicador inf temporización presente igual a 1 indica que vps_núm_unidades_en_pulso, vps_escala_temporal y vps_indicador_poc_proporcional_a_temporización están presentes en el conjunto de parámetros de vídeo. El vps indicador inf temporización presente igual a 0 especifica que vps_núm_unidades_en_pulso, vps_escala_temporal y vps_indicador_poc_proporcional_a_temporización no están presentes en el conjunto de parámetros de vídeo.

El vps_núm_unidades_en_pulso es el número de unidades de tiempo de un reloj que funciona en la frecuencia de vps_escala_temporal Hz, que corresponde a un incremento (llamado un pulso de reloj) de un contador de pulsos de reloj. El valor de vps_núm_unidades_en_pulso será mayor que 0. Un pulso de reloj, en unidades de segundos, es igual al cociente de vps_núm_unidades_en_pulso dividido entre vps_escala_temporal. Por ejemplo, cuando la frecuencia de imágenes de una señal de vídeo es de 25 Hz, vps_escala_temporal puede ser igual a 27.000.000 y vps_núm_unidades_en_pulso puede ser igual a 1.080.000 y, en consecuencia, un pulso de reloj puede ser de 0,04 segundos.

La vps_escala_temporal es el número de unidades de tiempo que pasan en un segundo. Por ejemplo, un sistema de coordenadas de tiempo que mide el tiempo usando un reloj de 27 MHz tiene una vps_escala_temporal de 27.000.000. El valor de vps_escala_temporal será mayor que 0.

El vps indicador_poc_proporcional_a_temporización igual a 1 indica que el valor del recuento de orden de imágenes para cada imagen en la secuencia de vídeo codificada, que no sea la primera imagen en la secuencia de vídeo codificado, en orden de decodificación, es proporcional al tiempo de salida de la imagen en relación con el tiempo de salida de la primera imagen en la secuencia de vídeo codificado. El vps indicador_poc_proporcional_a_temporización igual a 0 indica que el valor del recuento de orden de imágenes para cada imagen en la secuencia de vídeo codificado, que no sea la primera imagen en la secuencia de vídeo

codificado, en orden de decodificación, puede o no ser proporcional al tiempo de salida de la imagen en relación con el tiempo de salida de la primera imagen en la secuencia de vídeo codificado.

5 El **vps_núm_pulsos_dif_poc_uno_menos1** más 1 especifica el número de pulsos de reloj que corresponden a una diferencia de valores del recuento de orden de imágenes igual a 1. El valor de **vps_núm_pulsos_dif_poc_uno_menos1** estará en la gama entre 0 y $2^{32} - 1$, inclusive.

10 Lo siguiente es un ejemplo de una estructura sintáctica de parámetros de VUI, modificada para resolver uno o más de los problemas anteriores (la sintaxis subrayada es una adición a la estructura sintáctica de parámetros de VUI del HEVC WD9; la sintaxis en cursiva se elimina de la estructura sintáctica de parámetros de VUI del HEVC WD9; otras partes de la tabla de sintaxis se mantienen sin cambios en relación con el HEVC WD9):

parámetros_vui() {	Descriptor
...	
<u>sps indicador inf temporización presente</u>	<u>u(1)</u>
<i>si(sps indicador inf temporización presente) {</i>	
<u>sps núm unidades en pulso</u>	<u>u(32)</u>
<u>sps escala temporal</u>	<u>u(32)</u>
<u>sps indicador poc proporcional a temporización</u>	<u>u(1)</u>
<i>si(sps indicador poc proporcional a temporización)</i>	
<u>sps núm pulsos dif poc uno menos1</u>	<u>ue(v)</u>
<i>}</i>	
indicador parámetros hrd presentes	u(1)
<i>si(indicador parámetros hrd presentes)</i>	
<i>parámetros hrd(1, sps_máx_sub_capas_menos1)</i>	
<i>indicador poc proporcional a temporización</i>	u(1)
<i>si(indicador poc proporcional a temporización && indicador inf temporización presente)</i>	
<i>núm pulsos dif poc uno menos1</i>	<u>ue(v)</u>
...	
}	

Tabla 5: Estructura sintáctica ejemplar modificada de parámetros de VUI

15 La tabla 5 define elementos sintácticos agregados recientemente según la siguiente semántica de parámetros de VUI (la semántica para los elementos sintácticos eliminados se ha eliminado del mismo modo):

20 el **sps_indicador_inf_temporización_presente** igual a 1 especifica que el **sps_núm_unidades_en_pulso**, la **sps_escal_a_temporal** y el **sps_indicador_poc_proporcional_a_temporización** están presentes en la estructura sintáctica **parámetros_vui()**. El **sps_indicador_inf_temporización_presente** igual a 0 especifica que **sps_núm_unidades_en_pulso**, **sps_escal_a_temporal** y **sps_indicador_poc_proporcional_a_temporización** no están presentes en la estructura sintáctica **parámetros_vui()**.

25 El **sps_núm_unidades_en_pulso** es el número de unidades de tiempo de un reloj que funciona en la frecuencia de **sps_escal_a_temporal** Hz, que corresponde a un incremento (llamado un pulso de reloj) de un contador de pulsos de reloj. El **sps_núm_unidades_en_pulso** deberá ser mayor que 0. Un pulso de reloj, en unidades de segundos, es igual al cociente de **sps_núm_unidades_en_pulso** dividido entre **sps_escal_a_temporal**. Por ejemplo, cuando la frecuencia de imágenes de una señal de vídeo es de 25 Hz, **sps_escal_a_temporal** puede ser igual a 27.000.000 y **sps_núm_unidades_en_pulso** puede ser igual a 1.080.000 y, en consecuencia, un pulso de reloj puede ser igual a 0,04 segundos (véase la ecuación (1)). Cuando **vps_núm_unidades_en_pulso** esté presente en el conjunto de parámetros de vídeo al que hace referencia el conjunto de parámetros de secuencia, **sps_núm_unidades_en_pulso**, cuando esté presente, será igual a **vps_núm_unidades_en_pulso**.

35 La fórmula para obtener la variable PulsoReloj (también denominada en este documento "pulso de reloj") se modifica para que sea de la forma siguiente:

$$\text{PulsoReloj} = \frac{\text{sps_núm_unidades_en_pulso}}{\text{sps_escala_temporal}}$$

Ecuación (1)

5 **sps_escala_temporal** es el número de unidades de tiempo que pasan en un segundo. Por ejemplo, un sistema de coordenadas de tiempo que mide el tiempo usando un reloj de 27 MHz tiene una *sps_escala_temporal* de 27.000.000. El valor de *sps_escala_temporal* será mayor que 0. Cuando *vps_escala_temporal* esté presente en el conjunto de parámetros de vídeo al que hace referencia el conjunto de parámetros de secuencia, *sps_escala_temporal*, cuando esté presente, será igual a *vps_escala_temporal*.

10 El **sps_indicador_poc_proporcional_a_temporización** igual a 1 indica que el valor del recuento de orden de imágenes para cada imagen en la secuencia de vídeo codificado, que no sea la primera imagen en la secuencia de vídeo codificado, en orden de decodificación, es proporcional al tiempo de salida de la imagen en relación con el tiempo de salida de la primera imagen en la secuencia de vídeo codificado. El *sps_indicador_poc_proporcional_a_temporización* igual a 0 indica que el valor del recuento de orden de imágenes para cada imagen en la secuencia de vídeo codificado, que no sea la primera imagen en la secuencia de vídeo codificado, en orden de decodificación, puede o no ser proporcional al tiempo de salida de la imagen con respecto al tiempo de salida de la primera imagen en la secuencia de vídeo codificado. Cuando el *vps_indicador_poc_proporcional_a_temporización* esté presente en el conjunto de parámetros de vídeo al que hace referencia el conjunto de parámetros de secuencia, el *sps_indicador_poc_proporcional_a_temporización*, cuando esté presente, será igual al *vps_indicador_poc_proporcional_a_temporización*.

20 El **sps_núm_pulsos_dif_poc_uno_menos1** más 1 especifica el número de pulsos de reloj que corresponden a una diferencia de valores del recuento de orden de imágenes igual a 1. El valor de *sps_núm_pulsos_dif_poc_uno_menos1* estará en la gama entre 0 y $2^{32} - 1$, inclusive. Cuando el *vps_núm_pulsos_dif_poc_uno_menos1* esté presente en el conjunto de parámetros de vídeo al que hace referencia el conjunto de parámetros de secuencia, el *sps_núm_pulsos_dif_poc_uno_menos1*, cuando esté presente, será igual al *sps_núm_pulsos_dif_poc_uno_menos1*.

30 El siguiente es un ejemplo de una estructura sintáctica de parámetros de HRD, modificada para resolver uno o más de los problemas anteriores (la sintaxis en cursiva se elimina de la estructura sintáctica de parámetros de HRD del HEVC WD9):

parámetros_hrd(indicadorInfComúnPresente, máxNúmSubCapasMenos1) {	Descriptor
si(indicadorInfComúnPresente) {	
<i>indicador_inf_temporización_presente</i>	<i>u(1)</i>
<i>si(indicador_inf_temporización_presente) {</i>	
<i>núm_unidades_en_pulso</i>	<i>u(32)</i>
<i>escala_temporal</i>	<i>u(32)</i>
}	
...	
}	

Tabla 6: Estructura sintáctica ejemplar modificada de los parámetros de HRD

35 La semántica de los elementos sintácticos eliminados según la estructura sintáctica ejemplar modificada de parámetros de HRD de la tabla 6 se elimina del mismo modo.

40 El módulo de intra-predicción 46 puede intra-predicir un bloque actual, como alternativa a la inter-predicción llevada a cabo por la unidad de estimación de movimiento 42 y la unidad de compensación de movimiento 44, como se ha descrito anteriormente. En particular, el módulo de intra-predicción 46 puede determinar una modalidad de intra-predicción a usar para codificar un bloque actual. En algunos ejemplos, el módulo de intra-predicción 46 puede codificar un bloque actual usando varias modalidades de intra-predicción, por ejemplo, durante diferentes pasadas de codificación, y el módulo de intra-predicción 46 (o la unidad de selección de modalidad 40, en algunos ejemplos) puede seleccionar una modalidad de intra-predicción adecuada para usar, entre las modalidades probadas. Por ejemplo, el módulo de intra-predicción 46 puede calcular valores de velocidad-distorsión usando un análisis de velocidad-distorsión para las diversas modalidades de intra-predicción probadas, y seleccionar la modalidad de intra-predicción que tenga las mejores características de velocidad-distorsión entre las modalidades probadas. El análisis de velocidad-distorsión determina generalmente una magnitud de distorsión (o error) entre un bloque codificado y un bloque original no codificado, que fue codificado para generar el bloque codificado, así como una velocidad de bits (es decir, el número de bits) usado para generar el bloque codificado. El módulo de intra-predicción 46 puede calcular razones a partir de las distorsiones y velocidades para los diversos bloques codificados, para determinar

qué modalidad de intra-predicción presenta el mejor valor de velocidad-distorsión para el bloque.

En cualquier caso, tras seleccionar una modalidad de intra-predicción para un bloque, el módulo de intra-predicción 46 puede proporcionar información que indica la modalidad de intra-predicción seleccionada para el bloque, a la unidad de codificación por entropía 56. La unidad de codificación por entropía 56 puede codificar la información indicando la modalidad de intra-predicción seleccionada según las técnicas de esta divulgación. El codificador de vídeo 20 puede incluir datos de configuración en el flujo de bits transmitido, que pueden incluir una pluralidad de tablas de índices de modalidad de intra-predicción y una pluralidad de tablas de índices de modalidades de intra-predicción modificadas (también denominadas tablas de correlación de palabras de código), definiciones de contextos de codificación para varios bloques e indicaciones de la modalidad de intra-predicción más probable, una tabla de índices de modalidades de intra-predicción y una tabla modificada de índices de modalidades de intra-predicción, a usar en cada uno de los contextos.

Después de que el módulo de predicción 41 genere el bloque predictivo para el bloque de vídeo actual, ya sea mediante la inter-predicción o la intra-predicción, el codificador de vídeo 20 forma un bloque de vídeo residual restando el bloque predictivo al bloque de vídeo actual. Los datos de vídeo residual en el bloque residual pueden incluirse en una o más TU y aplicarse al módulo de transformación 52. El módulo de transformación 52 transforma los datos de vídeo residual en coeficientes de transformación residual usando 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 vídeo residual, desde un dominio 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 cuantización 54. La unidad de cuantización 54 cuantiza los coeficientes de transformación para reducir adicionalmente la velocidad de bits. El proceso de cuantización puede reducir la profundidad de bits asociada a algunos de, o a todos, los coeficientes. El grado de cuantización puede modificarse ajustando un parámetro de cuantización. En algunos ejemplos, la unidad de cuantización 54 puede realizar a continuación un escaneo de la matriz, incluyendo los coeficientes de transformación cuantizados. Como alternativa, la unidad de codificación por entropía 56 puede realizar el escaneo.

Tras la cuantización, la unidad de codificación por entropía 56 codifica por entropía los coeficientes de transformación cuantizados. Por ejemplo, la unidad de codificación por entropía 56 puede realizar una codificación de longitud variable adaptativa al contexto (CAVLC), una codificación aritmética binaria adaptativa al contexto (CABAC), una codificación aritmética binaria adaptativa al contexto y basada en la sintaxis (SBAC), una codificación por entropía mediante división en intervalos de probabilidad (PIPE) u otra metodología o técnica de codificación por entropía. Tras la codificación por entropía realizada por la unidad de codificación por entropía 56, el flujo de bits codificado puede transmitirse al decodificador de vídeo 30, o archivarse para su posterior transmisión o recuperación por parte del decodificador de vídeo 30. La unidad de codificación por entropía 56 también puede codificar por entropía los vectores de movimiento y los otros elementos sintácticos para el fragmento de vídeo actual que está siendo codificado.

La unidad de cuantización inversa 58 y el módulo de transformación inversa 60 aplican una cuantización inversa y una transformación inversa, respectivamente, para reconstruir el bloque residual en el dominio de píxeles, para su uso posterior como un bloque de referencia de una imagen de referencia. La unidad de compensación de movimiento 44 puede calcular un bloque de referencia añadiendo el bloque residual a un bloque predictivo de una de las imágenes de referencia dentro de una de las listas de imágenes de referencia. La unidad de compensación de movimiento 44 también puede aplicar uno o más filtros de interpolación al bloque residual reconstruido para calcular valores de fracciones de píxel para su uso en la estimación de movimiento. El sumador 62 añade el bloque residual reconstruido al bloque de predicción compensado por movimiento, generado por la unidad de compensación de movimiento 44 para generar un bloque de referencia, para su almacenamiento en la memoria de imágenes de referencia 64 (a veces llamada memoria intermedia de imágenes decodificadas (DPB)). El bloque de referencia puede ser usado por la unidad de estimación de movimiento 42 y la unidad de compensación de movimiento 44 como un bloque de referencia para inter-predecir un bloque en una trama o imagen de vídeo posterior.

El codificador de vídeo 20 puede incluir optativamente un decodificador de referencia hipotético (HRD) 57 (ilustrado como optativo mediante el uso de líneas de trazos) para comprobar los flujos de bits codificados, generados por elementos del codificador de vídeo 20 en cuanto a su conformidad con el modelo de memoria intermedia definido para el HRD 57. El HRD 57 puede comprobar flujos de bits o subconjuntos de flujos de bits de tipo I y/o tipo II en cuanto a conformidad de HRD. Los conjuntos de parámetros necesarios para el funcionamiento del HRD 57 son señalizados por uno de los dos tipos de conjuntos de parámetros de HRD, parámetros de HRD de NAL y parámetros de HRD de VCL. Como se ha descrito anteriormente, los conjuntos de parámetros de HRD se pueden incorporar dentro de la estructura sintáctica del SPS y/o la estructura sintáctica del VPS.

El HRD 57 puede probar los bloques de vídeo y elementos sintácticos asociados 55 en cuanto a su conformidad con los requisitos especificados, como una o más pruebas de conformidad de flujo de bits, definidas en una memoria descriptiva de codificación de vídeo, tal como el HEVC WD9 o una memoria descriptiva sucesora, tal como el HEVC

WD 10. Por ejemplo, el HRD 57 puede probar el flujo de bits codificado en cuanto a su conformidad, mediante el procesamiento de los elementos sintácticos 55 para determinar, a partir de la estructura sintáctica del VPS o en la parte de VUI de la estructura sintáctica del SPS para una secuencia de vídeo codificado, los elementos sintácticos que definen la condición para la señalización del número de pulsos de reloj que corresponden a una diferencia de valores del POC igual a 1. Si la condición se cumple según los valores de los elementos sintácticos, el HRD 57 puede determinar el número de pulsos de reloj que corresponden a una diferencia de valores del POC igual a 1 y utilizar el número determinado de pulsos de reloj como entrada para, por ejemplo, la determinación del desborde inferior o superior de la CPB durante la decodificación de imágenes codificadas incluidas en el flujo de bits codificado. El uso en este documento del término "procesamiento" en relación con una serie de elementos sintácticos puede referirse a la extracción, decodificación y extracción, la lectura, el análisis sintáctico y cualquier otra operación de servicio, o combinación de operaciones, para obtener los elementos sintácticos en una forma utilizable por un decodificador / HRD 57.

Como otro ejemplo, el HRD 57 puede probar el flujo de bits codificado en cuanto a su conformidad, mediante la decodificación del flujo de bits codificado para determinar la escala temporal y el número de unidades en un pulso de reloj, a partir de una estructura sintáctica del VPS de elementos sintácticos 55 que codifica los elementos sintácticos de la escala temporal y del número de unidades en un pulso de reloj, como máximo, una vez en la estructura sintáctica del VPS. En algunos casos, el HRD 57 puede probar el flujo de bits codificado en cuanto a su conformidad mediante la decodificación de los elementos sintácticos 55 para determinar la escala temporal y el número de unidades en un pulso de reloj a partir de una estructura sintáctica de VUI del flujo de bits codificado que codifica los elementos sintácticos de la escala temporal y del número de unidades en un pulso de reloj, como máximo, una vez en la estructura sintáctica de la VUI. La escala temporal y el número de unidades en un pulso de reloj se pueden señalar en un lugar que no sea una estructura sintáctica de parámetros de HRD incorporada dentro de la estructura sintáctica del VPS y/o de la VUI. El HRD 57 puede usar la escala temporal determinada y el número determinado de unidades en un pulso de reloj como entrada para, por ejemplo, la determinación del desborde inferior o superior de la CPB durante la decodificación de imágenes codificadas incluidas en el flujo de bits codificado.

Según las técnicas descritas en este documento, el HRD 57 puede probar el flujo de bits codificado en cuanto a su conformidad, mediante la decodificación, de una estructura sintáctica del VPS de los elementos sintácticos 55 para una o más secuencias de vídeo codificado, del valor de un indicador de POC proporcional a la indicación de temporización. El HRD 57 puede, adicionalmente o alternativamente, probar un flujo de bits codificado en cuanto a su conformidad decodificando el valor de un indicador de POC proporcional a la indicación de temporización en la estructura sintáctica del VPS solo si también se incluyen los elementos sintácticos de la escala temporal y del número de unidades en un pulso de reloj. El HRD 57 puede utilizar el valor determinado del indicador de POC proporcional al indicador de temporización y los elementos sintácticos de la escala temporal y del número de unidades en un pulso de reloj como entrada para, por ejemplo, la determinación del desborde inferior o superior de la CPB durante la decodificación de imágenes codificadas incluidas en el flujo de bits codificado.

La FIG. 3 es un diagrama de bloques que ilustra un decodificador ejemplar de vídeo 76 que puede implementar las técnicas descritas en esta divulgación. En el ejemplo de la FIG. 3, el decodificador de vídeo 76 incluye una memoria intermedia de imágenes codificadas (CPB) 78, una unidad de decodificación por entropía 80, un módulo de predicción 81, una unidad de cuantización inversa 86, una unidad de transformación inversa 88, un sumador 90 y una memoria intermedia 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. En algunos ejemplos, el decodificador de vídeo 76 puede realizar una pasada de decodificación generalmente recíproca a la pasada de codificación descrita con respecto al codificador de vídeo 20 de la FIG. 2. El decodificador de vídeo 76 puede representar un caso ejemplar de decodificador de vídeo 30 del dispositivo de destino 14 o del decodificador de referencia hipotético 57 de la FIG. 2.

La CPB 78 almacena imágenes codificadas del flujo de bits de imágenes codificadas. En un ejemplo, la CPB 78 es una memoria intermedia del tipo 'primero en entrar, primero en salir', que contiene unidades de acceso (AU) en orden de decodificación. Una AU es un conjunto de unidades de capa de abstracción de red (NAL) que están asociadas entre sí según una regla de clasificación especificada, son consecutivas en el orden de decodificación y cada una contiene exactamente una imagen codificada. El orden de decodificación es el orden en el que se decodifican las imágenes, y puede ser diferente al orden en que se muestran las imágenes (es decir, el orden de visualización). El funcionamiento de la CPB puede ser especificado por un decodificador de referencia hipotético (HRD).

Durante el proceso de decodificación, el decodificador de vídeo 76 recibe un flujo de bits de vídeo codificado que representa bloques de vídeo de un fragmento de vídeo codificado y elementos sintácticos asociados, desde el codificador de vídeo 20. La unidad de decodificación por entropía 80 del decodificador de vídeo 76 decodifica el flujo de bits para generar coeficientes cuantizados, vectores de movimiento y otros elementos sintácticos 55. La unidad de decodificación por entropía 80 remite los vectores de movimiento y otros elementos sintácticos 55 al módulo de predicción 81. El decodificador de vídeo 76 puede recibir los elementos sintácticos 55 a nivel del fragmento de vídeo y/o a nivel del bloque de vídeo. El flujo de bits de vídeo codificado puede incluir información de temporización señalizada según las técnicas descritas a continuación. Por ejemplo, el flujo de bits de vídeo codificado puede incluir

un conjunto de parámetros de vídeo (VPS), un conjunto de parámetros de secuencia (SPS), o cualquier combinación de los mismos, que tengan estructuras sintácticas según las técnicas descritas en este documento para señalar parámetros para operaciones de HRD.

5 Cuando el fragmento de vídeo se codifica como un fragmento intra-codificado (I), el módulo de intra-predicción 84 del módulo de predicción 81 puede generar datos de predicción para un bloque de vídeo del fragmento de vídeo actual, en base a una modalidad de intra-predicción señalizada, y datos de bloques previamente decodificados de la trama o imagen actual. Cuando la trama de vídeo se codifica como un fragmento inter-codificado (es decir, B, P o GPB), la unidad de compensación de movimiento 82 del módulo de predicción 81 genera bloques predictivos para un bloque de vídeo del fragmento de vídeo actual, en base a los vectores de movimiento y otros elementos sintácticos 55 recibidos desde la unidad de decodificación por entropía 80. Los bloques predictivos pueden ser generados a partir de una de las imágenes de referencia dentro de una de las listas de imágenes de referencia. El decodificador de vídeo 76 puede construir las listas de tramas de referencia, la lista 0 y la lista 1, usando técnicas de construcción por omisión, en base a las imágenes de referencia almacenadas en la DPB 92.

15 La unidad de compensación de movimiento 82 determina la información de predicción para un bloque de vídeo del fragmento de vídeo actual, analizando los vectores de movimiento y otros elementos sintácticos 55, y usa la información de predicción para generar los bloques predictivos para el bloque de vídeo actual que está siendo decodificado. Por ejemplo, la unidad de compensación de movimiento 82 utiliza algunos de los elementos sintácticos recibidos 55 para determinar una modalidad de predicción (por ejemplo, intra-predicción o inter-predicción), usada para codificar los bloques de vídeo del fragmento de vídeo, un tipo de fragmento de inter-predicción (por ejemplo, fragmento B, fragmento P o fragmento GPB), información de construcción para una o más de las listas de imágenes de referencia para el fragmento, vectores de movimiento para cada bloque de vídeo inter-codificado del fragmento, el estado de inter-predicción para cada bloque de vídeo inter-codificado del fragmento y otra información, para decodificar los bloques de vídeo en el fragmento de vídeo actual.

20 La unidad de compensación de movimiento 82 también puede realizar la interpolación en base a filtros de interpolación. La unidad de compensación de movimiento 82 puede usar filtros de interpolación como los usados por el codificador de vídeo 20 durante la codificación de los bloques de vídeo para calcular valores interpolados para fracciones de píxeles de bloques de referencia. En este caso, la unidad de compensación de movimiento 82 puede determinar los filtros de interpolación usados por el codificador de vídeo 20 a partir de los elementos sintácticos recibidos 55 y usar los filtros de interpolación para generar bloques predictivos.

30 La unidad de cuantización inversa 86 cuantiza de manera inversa, es decir, descuantiza, los coeficientes de transformación cuantizados, proporcionados en el flujo de bits y decodificados por la unidad de decodificación por entropía 80. El proceso de cuantización inversa puede incluir el uso de un parámetro de cuantización calculado por el codificador de vídeo 20 para cada bloque de vídeo en el fragmento de vídeo, para determinar un grado de cuantización y, asimismo, un grado de cuantización inversa que debería aplicarse. La unidad de transformación inversa 88 aplica una transformación inversa, por ejemplo, una DCT inversa, una transformación inversa entera o un proceso de transformación inversa conceptualmente similar, a los coeficientes de transformación con el fin de generar bloques residuales en el dominio de píxeles.

35 Después de que la unidad de compensación de movimiento 82 genere el bloque predictivo para el bloque de vídeo actual, en base a los vectores de movimiento y a otros elementos sintácticos 55, el decodificador de vídeo 76 forma un bloque de vídeo decodificado sumando los bloques residuales procedentes de la unidad de transformación inversa 88 a los correspondientes bloques predictivos generados por la unidad de compensación de movimiento 82. El sumador 90 representa el componente o los componentes que llevan a cabo esta operación de suma. Si se desea, también puede aplicarse un filtro de desbloqueo para filtrar los bloques decodificados, con el fin de eliminar distorsiones de efecto pixelado. Otros filtros de bucle (ya sea en el bucle de codificación o después del bucle de codificación) también pueden ser usados para allanar las transiciones de píxeles, o mejorar de otro modo la calidad del vídeo. A continuación, los bloques de vídeo decodificados en una trama o imagen dada se almacenan en la DPB 92, que almacena imágenes de referencia usadas para la posterior compensación de movimiento. La DPB 92 almacena también vídeo decodificado para su presentación posterior en un dispositivo de visualización, tal como el dispositivo de visualización 32 de la FIG. 1. Como la CPB 78, en un ejemplo, el funcionamiento de la DPB 92 puede ser especificado por el decodificador de referencia hipotético (HRD).

45 El codificador 20 y el decodificador 76, como se describe en esta divulgación, representan ejemplos de dispositivos configurados para llevar a cabo técnicas de temporización de señalización en un proceso de codificación de vídeo tal como se describe en esta divulgación. Según ello, las operaciones descritas en esta divulgación para la señalización del tiempo pueden ser realizadas por el codificador 20, el decodificador 76 o ambos. En algunos casos, el codificador 20 puede señalar la información de temporización y el decodificador 76 puede recibir dicha información de temporización, por ejemplo, para su uso en la definición de una o más funciones, características, parámetros o condiciones del HRD.

60 El decodificador de vídeo 76 puede, en algunos casos, ser un decodificador de vídeo 76 bajo prueba (o VUT). El decodificador de vídeo 76 puede recibir una representación del flujo de bits codificado, generado por un codificador

de vídeo 20 para señalar directamente, en la estructura sintáctica del VPS o en la parte de VUI de la estructura sintáctica del SPS de elementos sintácticos 55 para una secuencia de vídeo codificado, todos los elementos sintácticos que definen la condición para señalar una serie de pulsos de reloj que corresponden a una diferencia de valores del recuento de orden de imágenes (POC) igual a 1. El decodificador de vídeo 76 puede decodificar el flujo de bits codificado para determinar, a partir de la estructura sintáctica del VPS o en la parte de VUI de la estructura sintáctica del SPS para una secuencia de vídeo codificado, los elementos sintácticos que definen la condición para la señalización del número de pulsos de reloj que corresponden a una diferencia de valores del POC igual a 1. Si la condición se cumple según los valores de los elementos sintácticos, el decodificador de vídeo 76 puede determinar el número de pulsos de reloj que corresponden a una diferencia de valores del POC igual a 1, y utilizar el número determinado de pulsos de reloj como entrada para, por ejemplo, la determinación del desborde inferior o superior de la CPB 78 durante la decodificación de imágenes codificadas incluidas en el flujo de bits codificado.

En otro ejemplo, el codificador de vídeo 20 puede recibir una representación del flujo de bits codificado, generado por el codificador de vídeo 20 para indicar la escala temporal y el número de unidades en un pulso de reloj, como máximo, una vez en cada una de las estructuras sintácticas del VPS y de la VUI de los elementos sintácticos 55 para una secuencia dada de vídeo codificado. El decodificador de vídeo 76 puede decodificar el flujo de bits codificado para determinar la escala temporal y el número de unidades en un pulso de reloj a partir de una estructura sintáctica del VPS del flujo de bits codificado que codifica los elementos sintácticos de la escala temporal y del número de unidades en un pulso de reloj, como máximo, una vez en la estructura sintáctica del VPS. En algunos casos, el decodificador de vídeo 76 puede probar el flujo de bits codificado en cuanto a su conformidad, mediante la decodificación del flujo de bits codificado para determinar la escala temporal y el número de unidades en un pulso de reloj a partir de una estructura sintáctica de la VUI del flujo de bits codificado que codifica los elementos sintácticos de la escala temporal y del número de unidades en un pulso de reloj, como máximo, una vez en la estructura sintáctica de la VUI. La escala temporal y el número de unidades en un pulso de reloj se pueden señalar en un lugar que no sea una estructura sintáctica de parámetros de HRD incorporada dentro de la estructura sintáctica del VPS y/o de la VUI. El decodificador de vídeo 76 puede usar la escala temporal determinada y el número determinado de unidades en un pulso de reloj como entrada para, por ejemplo, la determinación del desborde inferior o superior de la CPB 78 durante la decodificación de imágenes codificadas incluidas en el flujo de bits codificado.

En otro ejemplo, el decodificador de vídeo 76 puede recibir una representación del flujo de bits codificado, generado por el codificador de vídeo 20 para señalar, en una estructura sintáctica del VPS de los elementos sintácticos 55 para una o más secuencias de vídeo codificado, un indicador de POC proporcional a indicación de temporización. El decodificador de vídeo 76 puede probar el flujo de bits codificado en cuanto a su conformidad, mediante la decodificación del flujo de bits codificado para determinar un valor para el indicador. El decodificador de vídeo 76 puede, adicionalmente o alternativamente, probar un flujo de bits codificado, generado por el decodificador de vídeo 76 para señalar el indicador en la estructura sintáctica del VPS solo si también se incluyen los elementos sintácticos de la escala temporal y del número de unidades en un pulso de reloj. El decodificador de vídeo 76 puede utilizar el valor determinado del indicador de POC proporcional a indicación de temporización y los elementos sintácticos de la escala temporal y del número de unidades en un pulso de reloj como entrada para, por ejemplo, la determinación del desborde inferior o superior de la CPB 78 durante la decodificación de las imágenes codificadas incluidas en el flujo de bits codificado.

La FIG. 4 es un diagrama de bloques que ilustra una estructura ejemplar de codificación 100 para un conjunto de imágenes de referencia. La estructura de codificación 100 incluye los fragmentos 102A a 102E (colectivamente, "fragmentos 102"). El recuento de orden de imágenes 108 asociado a la estructura de codificación 100 indica el orden de salida del fragmento correspondiente en el conjunto de imágenes de referencia. Por ejemplo, el fragmento l 102A tiene que emitirse en primer lugar (valor de POC 0), mientras que el fragmento b 102B tiene que emitirse en segundo lugar (valor de POC 1). El orden de decodificación 110 asociado a la estructura de codificación 100 denota el orden de decodificación para el fragmento correspondiente en el conjunto de imágenes de referencia. Por ejemplo, el fragmento l 102A tiene que emitirse en primer lugar (orden de decodificación 1), mientras que el fragmento b 102B tiene que emitirse en segundo lugar (orden de decodificación 2).

La flecha 104 indica un tiempo de salida de las imágenes a lo largo de un continuo temporal t . El intervalo de tiempo 106 representa un intervalo de tiempo que corresponde a una diferencia de valores del recuento de orden de imágenes (POC) igual a 1. El intervalo de tiempo 106 puede incluir un número de pulsos de reloj, que puede depender de la escala temporal (correspondiente, por ejemplo, a una frecuencia de oscilador - tal como 27 MHz - que define un sistema de coordenadas de tiempo para la información señalizada) y del número de unidades de tiempo de un reloj que funciona en la escala temporal que corresponde a un incremento de un contador de pulsos de reloj, que se denomina "pulso de reloj". Según las técnicas descritas en el presente documento, un codificador de vídeo 20 puede generar un flujo de bits para señalar directamente, en la estructura sintáctica del conjunto de parámetros de vídeo (VPS) o en la parte de información de facilidad de uso de vídeo (VUI) de la estructura sintáctica del conjunto de parámetros de secuencia (SPS) para una secuencia de vídeo codificado, elementos sintácticos que definen una condición para la señalización de una serie de pulsos de reloj que corresponden a una diferencia de valores del recuento de orden de imágenes (POC) igual a 1.

La FIG. 5 es un diagrama de flujo que muestra un procedimiento ejemplar de funcionamiento según las técnicas descritas en esta divulgación. Un codificador de vídeo 20 codifica las imágenes de una secuencia de vídeo para generar una secuencia de vídeo codificado (200). El codificador de vídeo 20 genera, además, los conjuntos de parámetros para la secuencia de vídeo codificado. Los conjuntos de parámetros pueden incluir parámetros codificados según una estructura sintáctica del conjunto de parámetros de secuencia (SPS) y/o según una estructura sintáctica del conjunto de parámetros de vídeo (VPS). Según las técnicas descritas en este documento, el codificador de vídeo 20 codifica los elementos sintácticos para el número de unidades en un pulso de reloj y la escala temporal, directamente en la estructura sintáctica del VPS y/o directamente en la estructura sintáctica del SPS para la secuencia de vídeo codificado (202). El término "directamente" indica que tal codificación puede generarse sin incorporar, en la estructura sintáctica del VPS o la estructura sintáctica del SPS (según el caso), los elementos sintácticos para el número de unidades en un pulso de reloj y la escala temporal definidos para una estructura sintáctica independiente de un conjunto de parámetros, tal como la correspondiente al conjunto de parámetros del decodificador de referencia hipotético (HRD) definido en el HEVC WD9.

Además, el codificador de vídeo 20 codifica, directamente en la estructura sintáctica del VPS y/o la estructura sintáctica del SPS de la secuencia de vídeo codificado, una condición para la señalización de una serie de pulsos de reloj que corresponden a una diferencia de valores del recuento de orden de imágenes (POC) igual a uno (204). La condición puede incluir uno o más elementos sintácticos que representan las variables de una fórmula booleana, en cuyo caso el codificador de vídeo 20 puede codificar cada uno de dichos elementos sintácticos directamente en la estructura sintáctica del VPS y/o la estructura sintáctica del SPS de la secuencia de vídeo codificado. El codificador de vídeo 20 emite la secuencia de vídeo codificado y la estructura sintáctica del VPS y/o la estructura sintáctica del SPS para la secuencia de vídeo codificado (206). En algunos casos, el codificador de vídeo 20 emite estas estructuras a un HRD del codificador de vídeo 20.

Las FIGs. 6A a 6B son diagramas de flujo que ilustran procedimientos ejemplares de funcionamiento según las técnicas descritas en esta divulgación. En la FIG. 6A, un codificador de vídeo 20 codifica las imágenes de una secuencia de vídeo para generar una secuencia de vídeo codificado (300). El codificador de vídeo 20 genera, además, los conjuntos de parámetros para la secuencia de vídeo codificado. Los conjuntos de parámetros pueden incluir parámetros codificados según una estructura sintáctica del conjunto de parámetros de vídeo (VPS). Según las técnicas descritas en el presente documento, el codificador de vídeo 20 codifica los elementos sintácticos para el número de unidades en un pulso de reloj y la escala temporal directamente y, como máximo, una vez, en una estructura sintáctica del VPS para la secuencia de vídeo codificado (302). En algunos casos, incluso en los casos en los que la estructura sintáctica del VPS incluye múltiples casos de parámetros de HRD, mediante la codificación de los elementos sintácticos directamente en la estructura sintáctica del VPS (como máximo, una vez) y no en los conjuntos de parámetros del HRD (o cualquier otra estructura sintáctica de conjunto de parámetros incorporados), la estructura sintáctica del VPS puede incluir un elemento sintáctico único para cada uno entre el número de unidades en un pulso de reloj y la escala temporal. El codificador de vídeo 20 emite la secuencia de vídeo codificado y la estructura sintáctica del VPS para la secuencia de vídeo codificado (304). En algunos casos, el codificador de vídeo 20 emite estas estructuras a un HRD del codificador de vídeo 20.

En la FIG. 6B, un codificador de vídeo 20 codifica las imágenes de una secuencia de vídeo para generar una secuencia de vídeo codificado (310). El codificador de vídeo 20 genera, además, los conjuntos de parámetros para la secuencia de vídeo codificado. Los conjuntos de parámetros pueden incluir parámetros codificados según una estructura sintáctica del conjunto de parámetros de vídeo (SPS). Según las técnicas descritas en el presente documento, el codificador de vídeo 20 codifica los elementos sintácticos para el número de unidades en un pulso de reloj y la escala temporal directamente y, como máximo, una vez, en una estructura sintáctica del SPS para la secuencia de vídeo codificado (312). En algunos casos, incluso en los casos en los que la estructura sintáctica del SPS incluye múltiples casos de parámetros de HRD, mediante la codificación de los elementos sintácticos directamente en la estructura sintáctica del SPS (como máximo, una vez) y no en los conjuntos de parámetros del HRD (o cualquier otra estructura sintáctica de conjunto de parámetros incorporados), la estructura sintáctica del SPS puede incluir un elemento sintáctico único para cada uno entre el número de unidades en un pulso de reloj y la escala temporal. El codificador de vídeo 20 emite la secuencia de vídeo codificado y la estructura sintáctica del SPS para la secuencia de vídeo codificado (314). En algunos casos, el codificador de vídeo 20 emite estas estructuras a un HRD del codificador de vídeo 20. En algunos casos, el codificador de vídeo 20 puede codificar los elementos sintácticos para el número de unidades en un pulso de reloj y la escala temporal, tanto para una estructura sintáctica del VPS como para una estructura sintáctica del SPS para la secuencia de vídeo codificado.

La FIG. 7 es un diagrama de flujo que ilustra un procedimiento ejemplar de funcionamiento según las técnicas descritas en esta divulgación. Un codificador de vídeo 20 codifica las imágenes de una secuencia de vídeo para generar una secuencia de vídeo codificado (400). El codificador de vídeo 20 genera, además, los conjuntos de parámetros para la secuencia de vídeo codificado. Los conjuntos de parámetros pueden incluir parámetros codificados según una estructura sintáctica del conjunto de parámetros de vídeo (VPS). Si la información de temporización tiene que incluirse, por ejemplo, para la definición de un modelo de memoria intermedia del HRD (rama SÍ en 402), el codificador de vídeo 20 codifica, directamente en una estructura sintáctica del VPS para la secuencia de vídeo codificado, un elemento sintáctico que tiene un valor que especifica si el valor del recuento de orden de imágenes (POC) para cada imagen en la secuencia de vídeo codificada, que no sea la primera imagen en

la secuencia de vídeo codificada, en orden de decodificación, es o no proporcional al tiempo de salida de la imagen en relación con el tiempo de salida de la primera imagen en la secuencia de vídeo codificado (404). El elemento sintáctico puede ser semánticamente similar al indicador_poc_proporcional_a_temporización definido por el HEVC WD9. La información de temporización puede representar el número de unidades en un pulso de reloj y la escala temporal.

Si el valor del elemento sintáctico es verdadero (rama SÍ en 406), el codificador de vídeo 20 codifica también un elemento sintáctico para el número de pulsos de reloj correspondientes a una diferencia de valores del recuento de orden de imágenes igual a uno (408). Debido a que el codificador de vídeo 20 codifica los elementos sintácticos en el VPS, los valores de los elementos sintácticos pueden aplicarse a todas las capas o a todos los posibles subconjuntos de flujos de bits de un flujo de bits de vídeo ajustable a escala, pues el VPS representa el conjunto de parámetros de la capa más alta y describe las características generales de secuencias de imágenes codificadas.

Si la información de temporización no se va a incluir en la estructura sintáctica del VPS (rama NO de 402), el codificador de vídeo 20 no codifica ni el elemento sintáctico para indicar que el POC es proporcional a la información de temporización ni el elemento sintáctico para el número de pulsos de reloj que corresponden a una diferencia de valores del recuento de orden de imágenes igual a uno. Si el POC no es proporcional a la información de temporización (es decir, el valor es falso) (rama NO de 406), el codificador de vídeo 20 no codifica el elemento sintáctico para el número de pulsos de reloj que corresponden a una diferencia de valores del recuento de orden de imágenes igual a uno.

El codificador de vídeo 20 emite la secuencia de vídeo codificado y la estructura sintáctica del VPS para la secuencia de vídeo codificado (410). En algunos casos, el codificador de vídeo 20 emite estas estructuras a un HRD del codificador de vídeo 20.

La FIG. 8 es un diagrama de flujo que ilustra un procedimiento ejemplar de funcionamiento según las técnicas descritas en esta divulgación. Un dispositivo decodificador de vídeo 30, o decodificador de referencia hipotético 57, de un dispositivo codificador de vídeo 20 (en adelante, "el decodificador") recibe una secuencia de vídeo codificado y una estructura sintáctica del conjunto de parámetros de vídeo (VPS) y/o una estructura sintáctica del conjunto de parámetros de secuencia (SPS) para la secuencia de vídeo codificado (500). La secuencia de vídeo codificado y/o la(s) estructura(s) sintáctica(s) pueden codificarse en un flujo de bits que incluye una o más imágenes codificadas.

El decodificador procesa la estructura sintáctica del VPS y/o la estructura sintáctica del SPS para extraer un elemento sintáctico que especifica, directamente en la estructura sintáctica del VPS y/o la estructura sintáctica del SPS, una condición para la señalización de una serie de pulsos de reloj que corresponden a una diferencia de valores del recuento de orden de imágenes (POC) igual a uno (502). La condición puede incluir uno o más elementos sintácticos que representan variables de una fórmula booleana, en cuyo caso el decodificador puede procesar cada uno de dichos elementos sintácticos directamente desde la estructura sintáctica del VPS y/o la estructura sintáctica del SPS de la secuencia de vídeo codificado.

El decodificador, además, procesa la estructura sintáctica del VPS y/o la estructura sintáctica del SPS para extraer elementos sintácticos para el número de unidades en un pulso de reloj y la escala temporal, directamente desde la estructura sintáctica del VPS y/o directamente desde la estructura sintáctica del SPS para la secuencia de vídeo codificado (504). El decodificador puede entonces verificar la conformidad de la secuencia de vídeo codificado con un modelo de memoria intermedia de vídeo que está definido, al menos en parte, mediante los valores para la condición, el número de unidades en un pulso de reloj y la escala temporal, tal como se extraen de la estructura sintáctica del VPS y/o la estructura sintáctica del SPS, y como se leen desde los elementos sintácticos correspondientes (506).

Las FIGs. 9A a 9B son diagramas de flujo que ilustran procedimientos ejemplares de funcionamiento según las técnicas descritas en esta divulgación. En la FIG. 9A, un dispositivo decodificador de vídeo 30 o un decodificador de referencia hipotético 57 de un dispositivo codificador de vídeo 20 (en adelante, "el decodificador") recibe una secuencia de vídeo codificado y una estructura sintáctica del conjunto de parámetros de vídeo (VPS) para la secuencia de vídeo codificado (600). La secuencia de vídeo codificado y/o la estructura sintáctica del VPS pueden codificarse en un flujo de bits que incluye una o más imágenes codificadas.

Según las técnicas descritas en este documento, el decodificador procesa la estructura sintáctica del VPS para extraer elementos sintácticos para el número de unidades en un pulso de reloj y la escala temporal, lo que se produce directamente y, como máximo, una vez en la estructura sintáctica del VPS para la secuencia de vídeo codificado (602). El decodificador puede entonces verificar la conformidad de la secuencia de vídeo codificado con un modelo de memoria intermedia de vídeo que se define, al menos en parte, mediante los valores para el número de unidades en un pulso de reloj y la escala temporal, tal como se extraen de la estructura sintáctica del VPS y como se leen desde los elementos sintácticos correspondientes (604).

En la FIG. 9B, el decodificador recibe una secuencia de vídeo codificado y una estructura sintáctica del conjunto de parámetros de vídeo (SPS) para la secuencia de vídeo codificado (610). La secuencia de vídeo codificado y/o las

estructuras sintácticas del SPS pueden codificarse en un flujo de bits que incluye una o más imágenes codificadas.

Según las técnicas descritas en este documento, el decodificador procesa la estructura sintáctica del SPS para extraer elementos sintácticos para el número de unidades en un pulso de reloj y la escala temporal, lo que se produce directamente y, como máximo, una vez en la estructura sintáctica del SPS para la secuencia de vídeo codificado (612). El decodificador puede entonces verificar la conformidad de la secuencia de vídeo codificado con un modelo de memoria intermedia de vídeo que se define, al menos en parte, mediante los valores para el número de unidades en un pulso de reloj y la escala temporal, tal como se extraen de la estructura sintáctica del SPS y como se leen desde los elementos sintácticos correspondientes (614).

La FIG. 10 es un diagrama de flujo que ilustra un procedimiento ejemplar de funcionamiento según las técnicas descritas en esta divulgación. En la FIG. 10, un dispositivo decodificador de vídeo 30 o un decodificador de referencia hipotético 57 de un dispositivo codificador de vídeo 20 (en adelante, "el decodificador") recibe una secuencia de vídeo codificado y una estructura sintáctica del conjunto de parámetros de vídeo (VPS) para la secuencia de vídeo codificado (700). La secuencia de vídeo codificado y/o la estructura sintáctica del VPS pueden codificarse en un flujo de bits que incluye una o más imágenes codificadas.

El decodificador procesa la estructura sintáctica del VPS para extraer un elemento sintáctico que especifica si el valor del recuento de orden de imágenes para cada imagen en la secuencia de vídeo codificado, que no sea la primera imagen en la secuencia de vídeo codificado, en orden de decodificación, es o no proporcional al tiempo de salida de la imagen en relación con el tiempo de salida de la primera imagen en la secuencia de vídeo codificado (702). Si el valor del elemento sintáctico es verdadero, entonces el decodificador procesa adicionalmente la estructura sintáctica del VPS para extraer un elemento sintáctico para el número de pulsos de reloj que corresponden a una diferencia de valores del recuento de orden de imágenes igual a uno (706). El decodificador puede entonces verificar la conformidad de la secuencia de vídeo codificado con un modelo de memoria intermedia de vídeo que se define, al menos en parte, mediante el valor para el número de pulsos de reloj que corresponden a una diferencia de valores del recuento de orden de imágenes igual a uno, extraído de la estructura sintáctica del VPS y tal como se lee desde el elemento sintáctico correspondiente (708).

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

A modo de ejemplo, y no de manera limitativa, tales medios de almacenamiento legibles por ordenador pueden comprender RAM, ROM, EEPROM, CD-ROM u otro almacenamiento de disco óptico, almacenamiento de disco magnético u otros dispositivos de almacenamiento magnético, memoria flash o cualquier otro medio que pueda usarse para almacenar código de programa deseado en forma de instrucciones o estructuras de datos y al que pueda accederse mediante un ordenador. Además, cualquier conexión puede denominarse adecuadamente un medio legible por ordenador. Por ejemplo, si las instrucciones se transmiten desde una sede de la Red, un servidor u otra fuente remota usando un cable coaxial, un cable de fibra óptica, un par trenzado, una línea de abonado digital (DSL) o tecnologías inalámbricas tales como infrarrojos, radio y microondas, entonces el cable coaxial, el cable de fibra óptica, el par trenzado, la DSL o las tecnologías inalámbricas tales como infrarrojos, radio y microondas se incluyen en la definición de medio. Sin embargo, debería entenderse que los medios de almacenamiento legibles por ordenador y los medios de almacenamiento de datos no incluyen conexiones, ondas portadoras, señales u otros medios transitorios, sino que, en cambio, se orientan a medios de almacenamiento tangibles no transitorios. Los discos, como se usan en el presente documento, incluyen el disco compacto (CD), el disco de láser, el disco óptico, el disco versátil digital (DVD), el disco flexible y el disco Blu-ray, donde algunos discos normalmente reproducen datos de manera magnética, mientras que otros discos reproducen los datos de manera óptica con láser. Las combinaciones de lo anterior también deberían incluirse dentro del 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), microprocesadores de propósito general, circuitos integrados específicos de la aplicación (ASIC), formaciones lógicas programables en el terreno (FPGA) u otros circuitos lógicos integrados o discretos equivalentes. Por consiguiente, el término "procesador", como se usa en el presente documento, puede referirse a cualquier estructura anterior 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

proporcionarse dentro de hardware especializado y/o módulos de software configurados para la codificación y la decodificación, o incorporarse en un códec combinado. Además, las técnicas podrían implementarse completamente en uno o más circuitos o elementos lógicos.

5 Las técnicas de esta divulgación se pueden implementar en una gran variedad de dispositivos o aparatos, incluyendo un equipo de mano inalámbrico, un circuito integrado (IC) o un conjunto de IC (por ejemplo, un conjunto de chips). Varios componentes, módulos o unidades se describen en esta divulgación para enfatizar aspectos funcionales de dispositivos configurados para realizar las técnicas divulgadas, pero no requieren necesariamente la realización mediante diferentes unidades de hardware. Más bien, como se ha descrito anteriormente, pueden
10 combinarse diversas unidades en una unidad de hardware de códec o ser proporcionadas por una colección de unidades de hardware inter-operativas, incluyendo uno o más procesadores, como se ha descrito anteriormente, junto con el software y/o firmware adecuado.

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

REIVINDICACIONES

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

5 recibir (600, 610) una secuencia de vídeo codificado que comprende imágenes codificadas de una secuencia de vídeo; y

10 recibir parámetros de temporización para la secuencia de vídeo codificado que incluyen un primer elemento sintáctico en una estructura sintáctica de un conjunto de parámetros de vídeo, VPS, a la que hace referencia la secuencia de vídeo codificado que proporciona una escala temporal, y un segundo elemento sintáctico en la estructura sintáctica del VPS, que proporciona un número de unidades en un pulso de reloj, estando los elementos sintácticos primero y segundo en la estructura sintáctica del VPS incluidos directamente y, como máximo, una vez, en la estructura sintáctica del VPS, y siendo aplicables a todas las capas especificadas por la estructura sintáctica del VPS,

15 en el que los parámetros de temporización para la secuencia de vídeo codificado incluyen un primer elemento sintáctico en una parte de información de facilidad de uso de vídeo, VUI, de una estructura sintáctica de un conjunto de parámetros de secuencia, SPS (612), a la que hace referencia la secuencia de vídeo codificado que proporciona la escala temporal, y un segundo elemento sintáctico en la parte de VUI de la estructura sintáctica de SPS que proporciona el número de unidades en un pulso de reloj, estando los elementos sintácticos primero y segundo en la parte de VUI de la estructura sintáctica del SPS incluidos directamente y, como máximo, una vez, en la parte de VUI de la estructura sintáctica del SPS, y siendo aplicables a todas las capas que están dentro de la secuencia de vídeo codificado y que hacen referencia a la estructura sintáctica del SPS.

20 2. El procedimiento de la reivindicación 1, en el que la recepción de los parámetros de temporización para la secuencia de vídeo codificado comprende la recepción de la escala temporal y del número de unidades en un pulso de reloj en la estructura sintáctica del VPS.

30 3. El procedimiento de la reivindicación 1, en el que la recepción de los parámetros de temporización para la secuencia de vídeo codificado comprende la recepción de la escala temporal y del número de unidades en un pulso de reloj en la parte de VUI de la estructura sintáctica del SPS.

35 4. El procedimiento de la reivindicación 1, en el que la recepción de la secuencia de vídeo codificado comprende la recepción de un flujo de bits codificado que comprende una secuencia de bits que forma una representación de las imágenes codificadas, comprendiendo además el procedimiento:

40 verificar la conformidad del flujo de bits a un modelo de almacenamiento temporal de vídeo de un almacén temporal de imágenes codificadas y de un almacén temporal de imágenes decodificadas, definido, al menos en parte, por la escala temporal y el número de unidades en un pulso de reloj.

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

45 codificar (300, 310) imágenes de una secuencia de vídeo para generar una secuencia de vídeo codificado que comprende las imágenes codificadas; y

50 señalar parámetros de temporización para la secuencia de vídeo codificado, señalizando (302), al menos en parte, un primer elemento sintáctico en una estructura sintáctica del conjunto de parámetros de vídeo, VPS, a la que hace referencia la secuencia de vídeo codificado que proporciona una escala temporal, y un segundo elemento sintáctico en la estructura sintáctica del VPS que proporciona un número de unidades en un pulso de reloj, estando los elementos sintácticos primero y segundo en la estructura sintáctica del VPS incluidos directamente y, como máximo, una vez, en la estructura sintáctica del VPS, y siendo aplicables a todas las capas especificadas por la estructura sintáctica del VPS, y señalizando (312), al menos en parte, un primer elemento sintáctico en una parte de información de facilidad de uso de vídeo, VUI, de una estructura sintáctica del conjunto de parámetros de secuencia, SPS, a la que hace referencia la secuencia de vídeo codificado que proporciona la escala temporal, y un segundo elemento sintáctico en la parte de VUI de la estructura sintáctica del SPS que proporciona el número de unidades en un pulso de reloj, estando los elementos sintácticos primero y segundo en la parte de VUI de la estructura sintáctica del SPS incluidos directamente y, como máximo, una vez, en la parte de VUI de la estructura sintáctica del SPS, y siendo aplicables a todas las capas que están dentro de la secuencia de vídeo codificado y que hacen referencia a la estructura sintáctica del VPS.

60 6. El procedimiento de la reivindicación 1 o 5, en el que el primer elemento sintáctico en la estructura sintáctica del VPS comprende un elemento sintáctico vps_escala_temporal y el segundo elemento sintáctico en la estructura sintáctica del VPS comprende un elemento sintáctico vps_núm_unidades_en_pulso.

65

- 5
7. El procedimiento de la reivindicación 1 o 5, en el que el primer elemento sintáctico en la parte de VUI de la estructura sintáctica del SPS comprende un elemento sintáctico `sps_escala_temporal` y el segundo elemento sintáctico en la parte de VUI de la estructura sintáctica del SPS comprende un elemento sintáctico `sps_núm_unidades_en_pulso`.
8. El procedimiento de la reivindicación 1 o 5, en el que los elementos sintácticos primero y segundo en la estructura sintáctica del VPS no están señalizados en una estructura sintáctica de parámetros del decodificador de referencia hipotético (HRD), incorporada en la estructura sintáctica del VPS.
- 10
9. El procedimiento de la reivindicación 1 o 5, en el que los elementos sintácticos primero y segundo en la parte de VUI de la estructura sintáctica del SPS no están señalizados en una estructura sintáctica de parámetros del decodificador de referencia hipotético (HRD), incorporada en la estructura sintáctica del SPS.
- 15
10. El procedimiento de la reivindicación 5, que comprende además:
- generar un flujo de bits codificado que comprende una secuencia de bits que forma una representación de las imágenes codificadas,
- en el que los parámetros de temporización definen, al menos en parte, un modelo de almacén temporal de vídeo de un almacén temporal de imágenes codificadas y un almacén temporal de imágenes decodificadas, para verificar la conformidad del flujo de bits.
- 20
11. El procedimiento de la reivindicación 1 o 5, en el que los parámetros de temporización comprenden parámetros de temporización para operaciones de decodificación de referencia hipotética.
- 25
12. El procedimiento de la reivindicación 1 o 5, en el que la estructura sintáctica del VPS especifica una pluralidad de capas.
- 30
13. Un medio de almacenamiento no transitorio, legible por ordenador, que almacena instrucciones para el procesamiento de datos de vídeo que, al ser ejecutadas por uno o más procesadores, hacen que los uno o más procesadores lleven a cabo el procedimiento de una cualquiera de las reivindicaciones 1 a 12.
- 35
14. Un dispositivo para el procesamiento de datos de vídeo que comprende: medios para recibir una secuencia de vídeo codificado que comprende imágenes codificadas de una secuencia de vídeo; y
- medios para recibir parámetros de temporización para la secuencia de vídeo codificado, que incluyen un primer elemento sintáctico en una estructura sintáctica del conjunto de parámetros de vídeo, VPS, a la que hace referencia la secuencia de vídeo codificado que proporciona una escala temporal y un segundo elemento sintáctico en la estructura sintáctica del VPS que proporciona un número de unidades en un pulso de reloj, estando los elementos sintácticos primero y segundo en la estructura sintáctica del VPS incluidos directamente y, como máximo, una vez, en la estructura sintáctica del VPS, y siendo aplicables a todas las capas especificadas por la estructura sintáctica del VPS,
- 40
- en el que los parámetros de temporización para la secuencia de vídeo codificado incluyen un primer elemento sintáctico en una parte de información de facilidad de uso de vídeo, VUI, de una estructura sintáctica del conjunto de parámetros de secuencia, SPS, a la que hace referencia la secuencia de vídeo codificado que proporciona la escala temporal, y un segundo elemento sintáctico en la parte de VUI de la estructura sintáctica del VPS que proporciona el número de unidades en un pulso de reloj, estando los elementos sintácticos primero y segundo en la parte de VUI de la estructura sintáctica del SPS incluidos directamente y, como máximo, una vez, en la parte de VUI de la estructura sintáctica del SPS, y siendo aplicables a todas las capas que están dentro de la secuencia de vídeo codificado y que hacen referencia a la estructura sintáctica del SPS.
- 45
- 50
- 55
15. Un dispositivo para codificar datos de vídeo, que comprende:
- medios para codificar imágenes de una secuencia de vídeo, para generar una secuencia de vídeo codificado que comprende las imágenes codificadas; y
- 60
- medios para señalar parámetros de temporización para la secuencia de vídeo codificado, señalizando, al menos en parte, un primer elemento sintáctico en una estructura sintáctica del conjunto de parámetros de vídeo, VPS, a la que hace referencia la secuencia de vídeo codificado que proporciona una escala temporal, y un segundo elemento sintáctico en la estructura sintáctica del VPS que proporciona un número de unidades en un pulso de reloj, estando los elementos sintácticos primero y segundo en la estructura sintáctica del VPS incluidos directamente y, como máximo, una vez, en la estructura sintáctica del VPS, y siendo aplicables a todas las capas especificadas por la estructura sintáctica del VPS, y señalizando, al menos en parte, un primer elemento sintáctico en una parte de información de facilidad de uso de vídeo, VUI, de una estructura
- 65

- 5 sintáctica del conjunto de parámetros de vídeo, SPS, a la que hace referencia la secuencia de vídeo codificado que proporciona la escala temporal, y un segundo elemento sintáctico en la parte de VUI de la estructura sintáctica del VPS, que proporciona el número de unidades en un pulso de reloj, estando los elementos sintácticos primero y segundo en la parte de VUI de la estructura sintáctica del SPS incluidos directamente y, como máximo, una vez, en la parte de VUI de la estructura sintáctica del SPS, y siendo aplicables a todas las capas que están dentro de la secuencia de vídeo codificado y que hacen referencia a la estructura sintáctica del SPS.

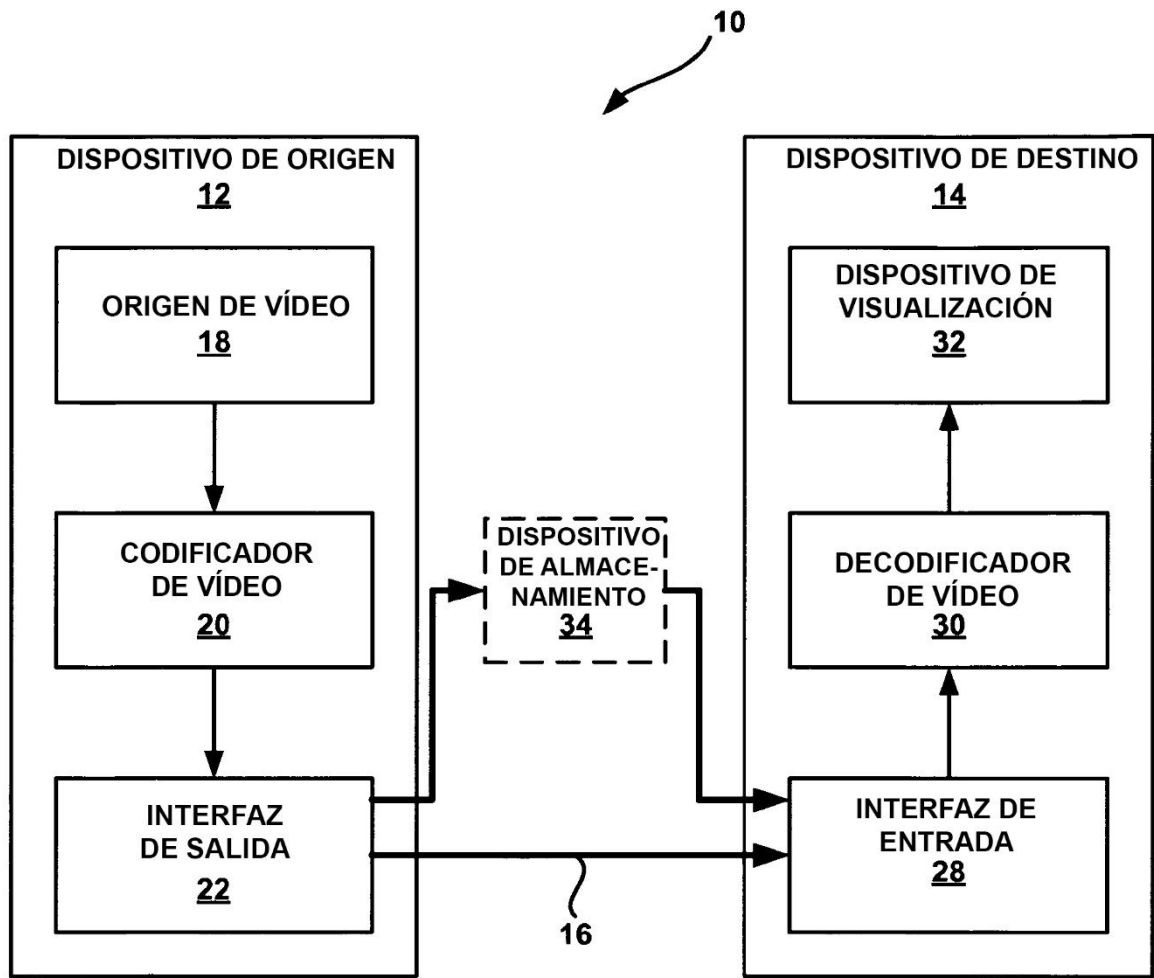


FIG. 1

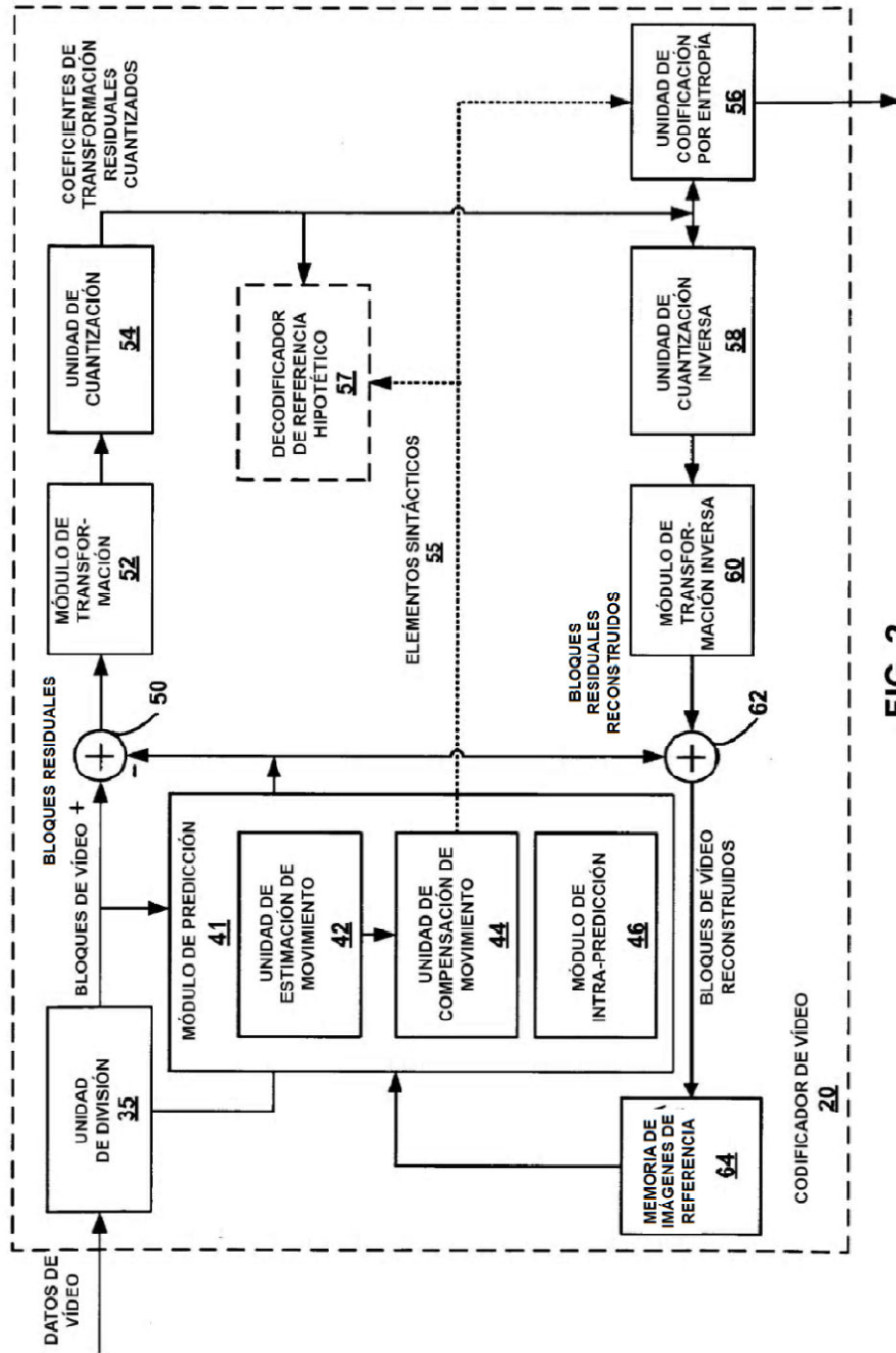


FIG. 2

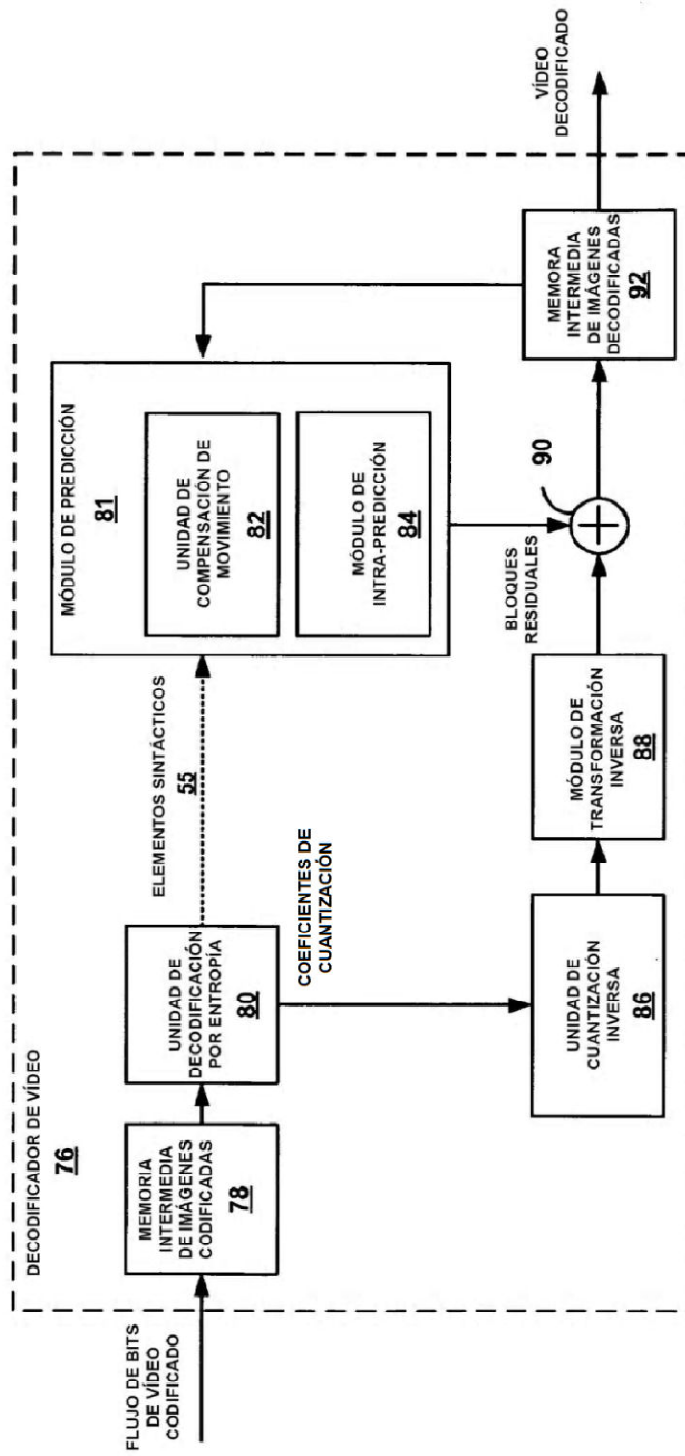


FIG. 3

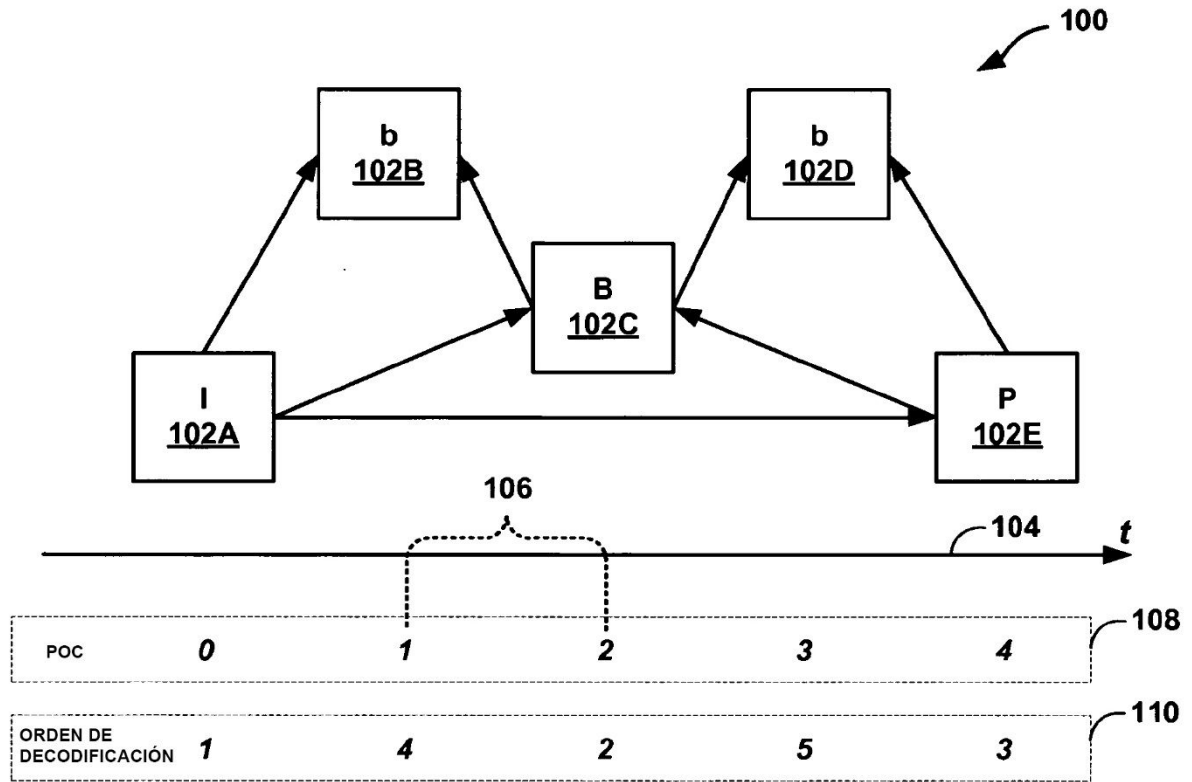


FIG. 4

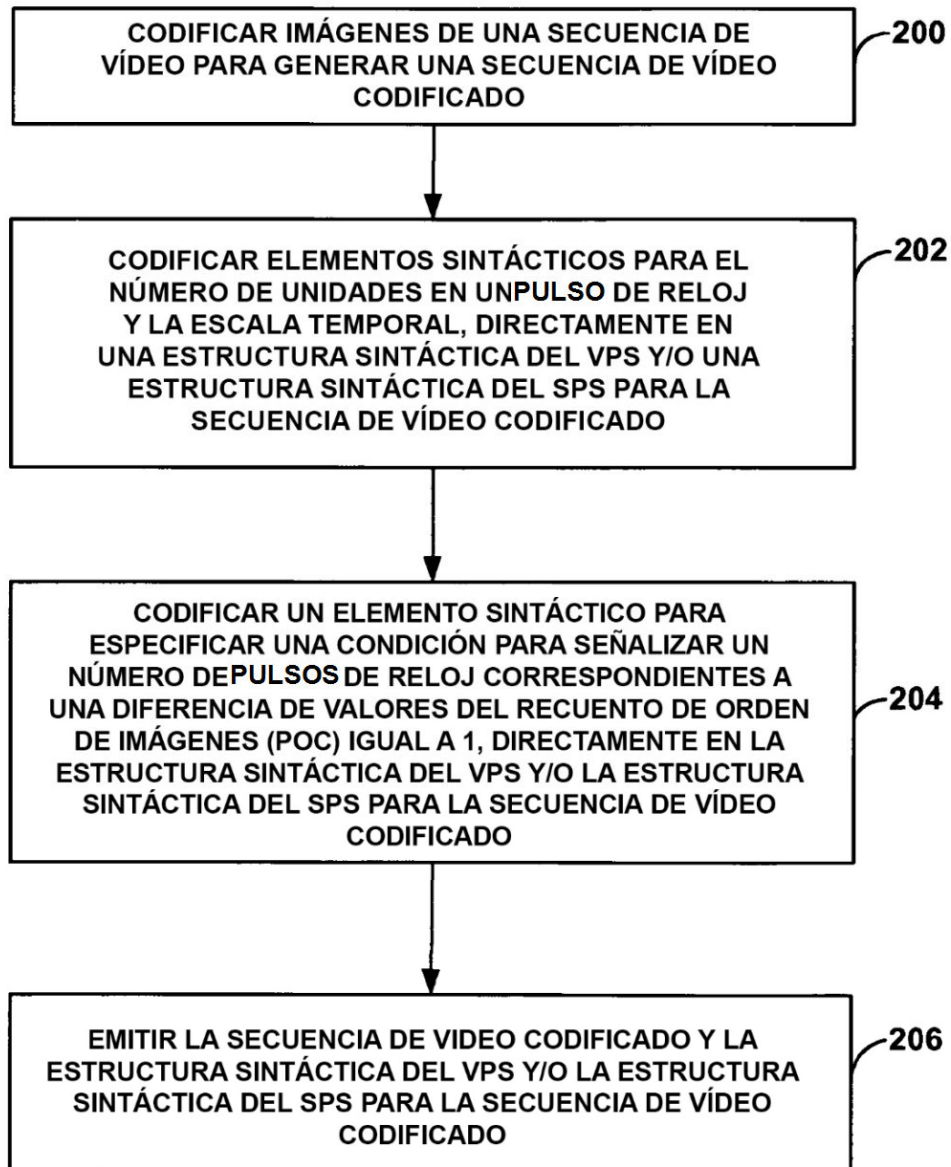


FIG. 5

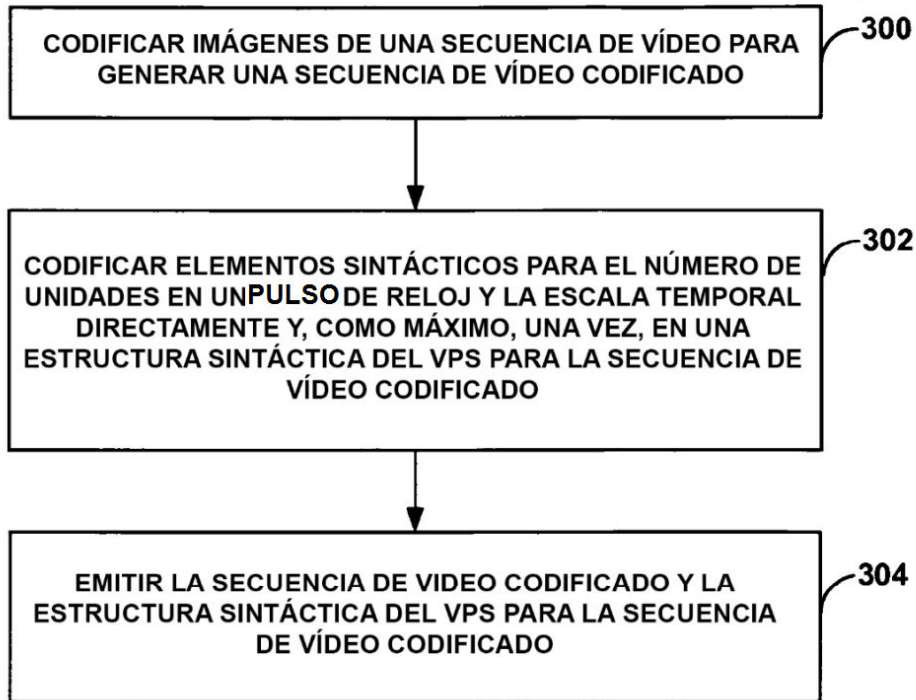


FIG. 6A

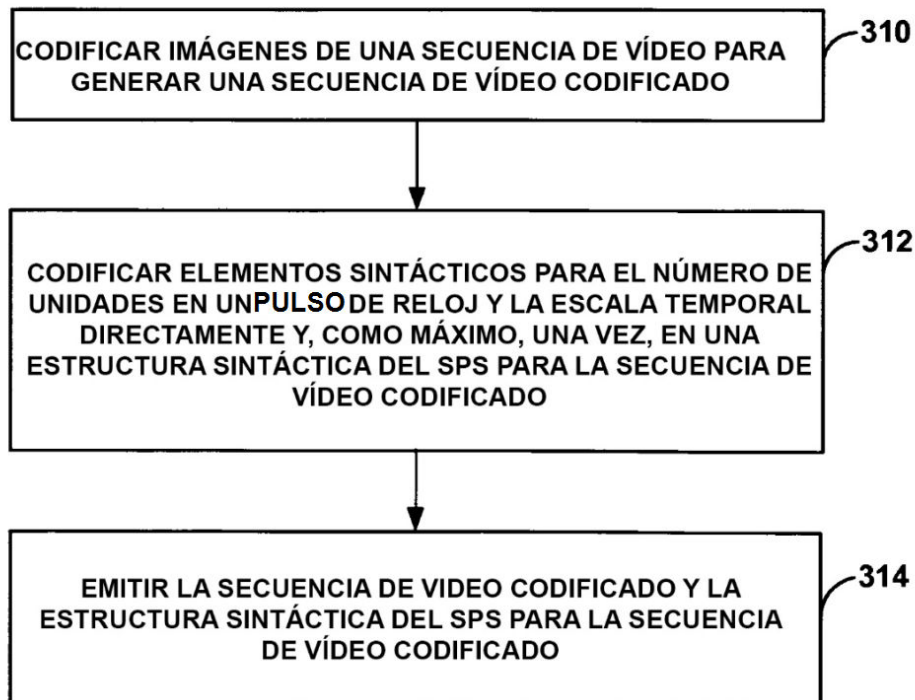


FIG. 6B

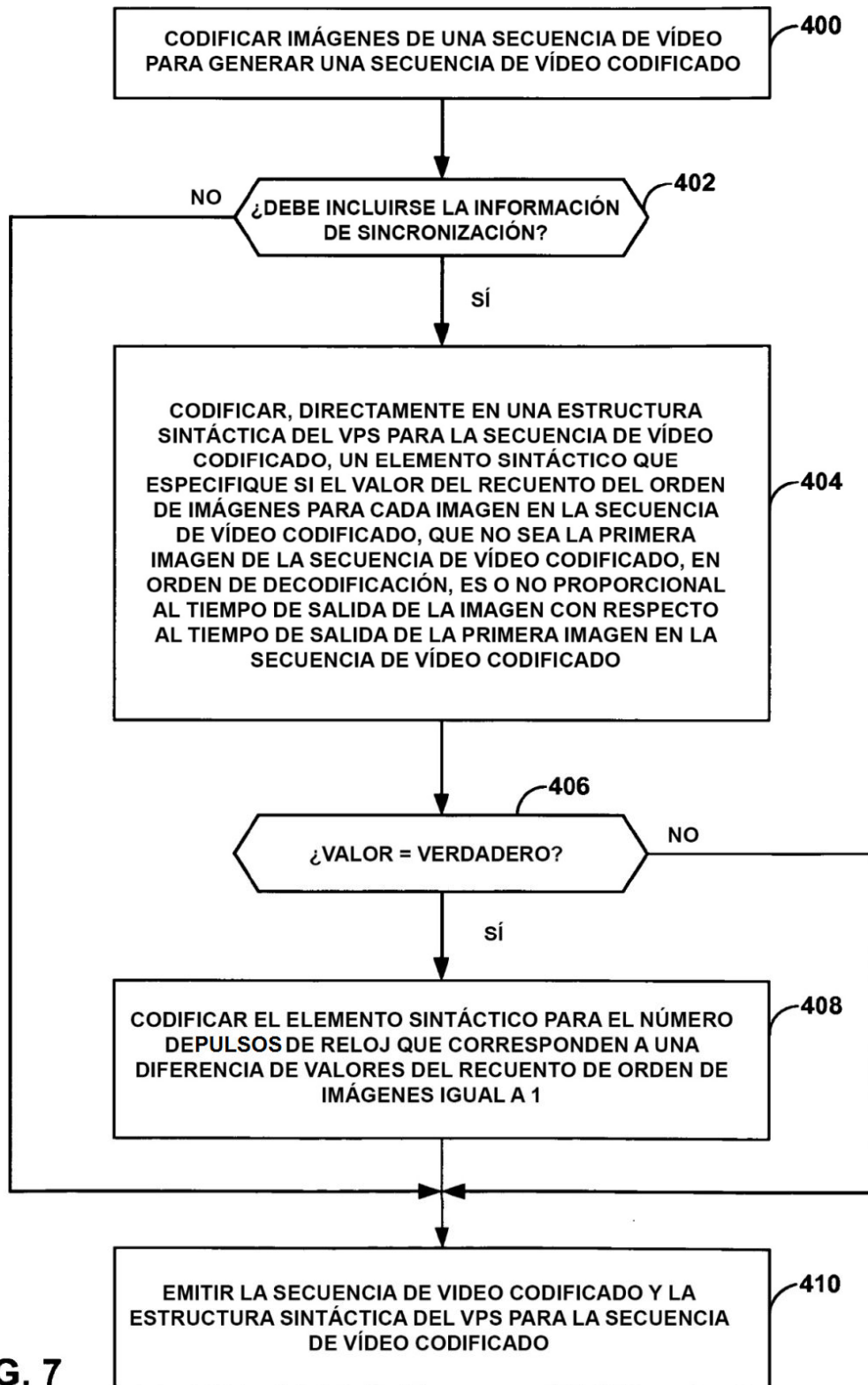


FIG. 7

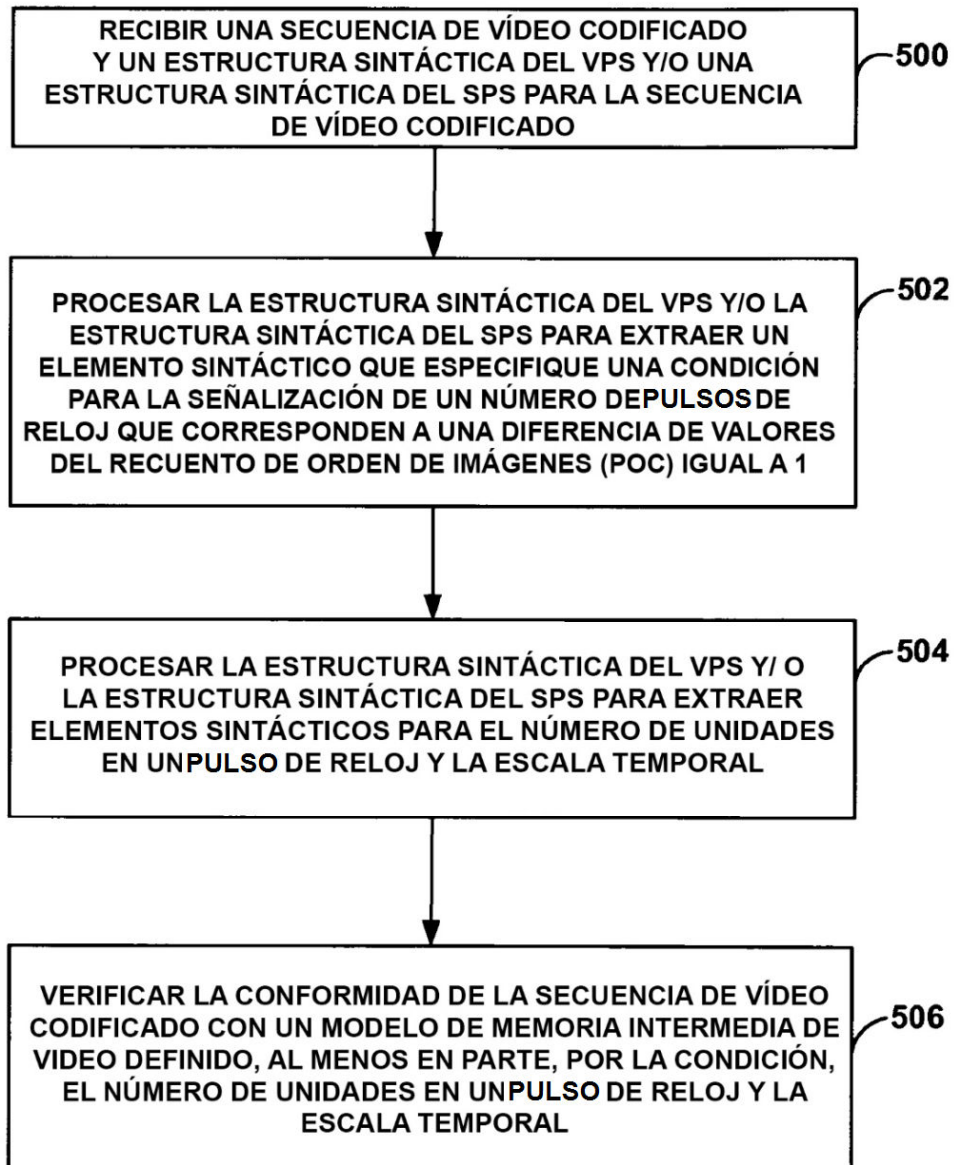


FIG. 8

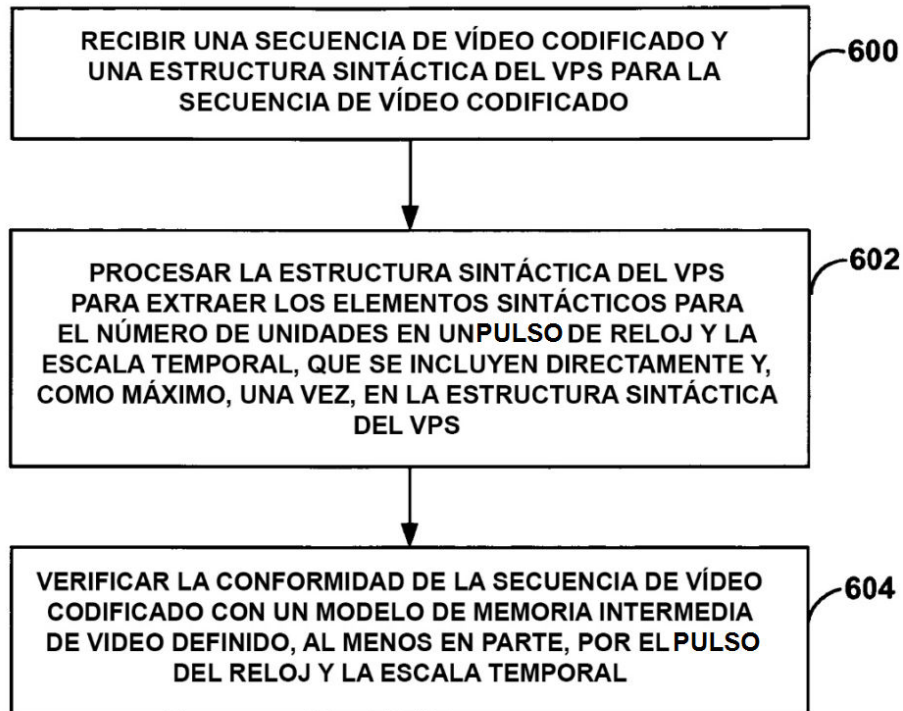


FIG. 9A

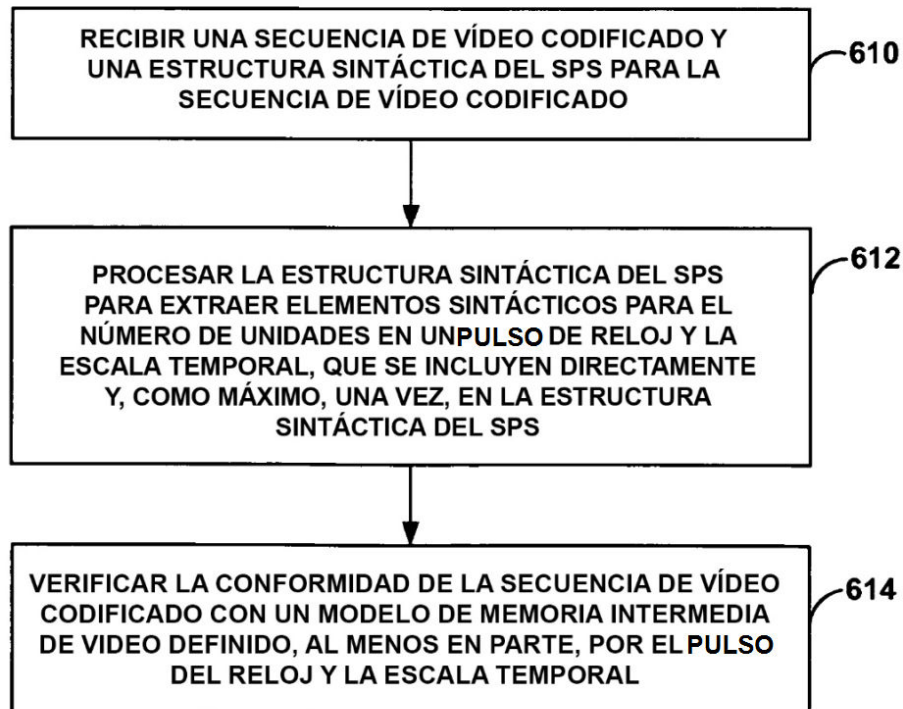


FIG. 9B

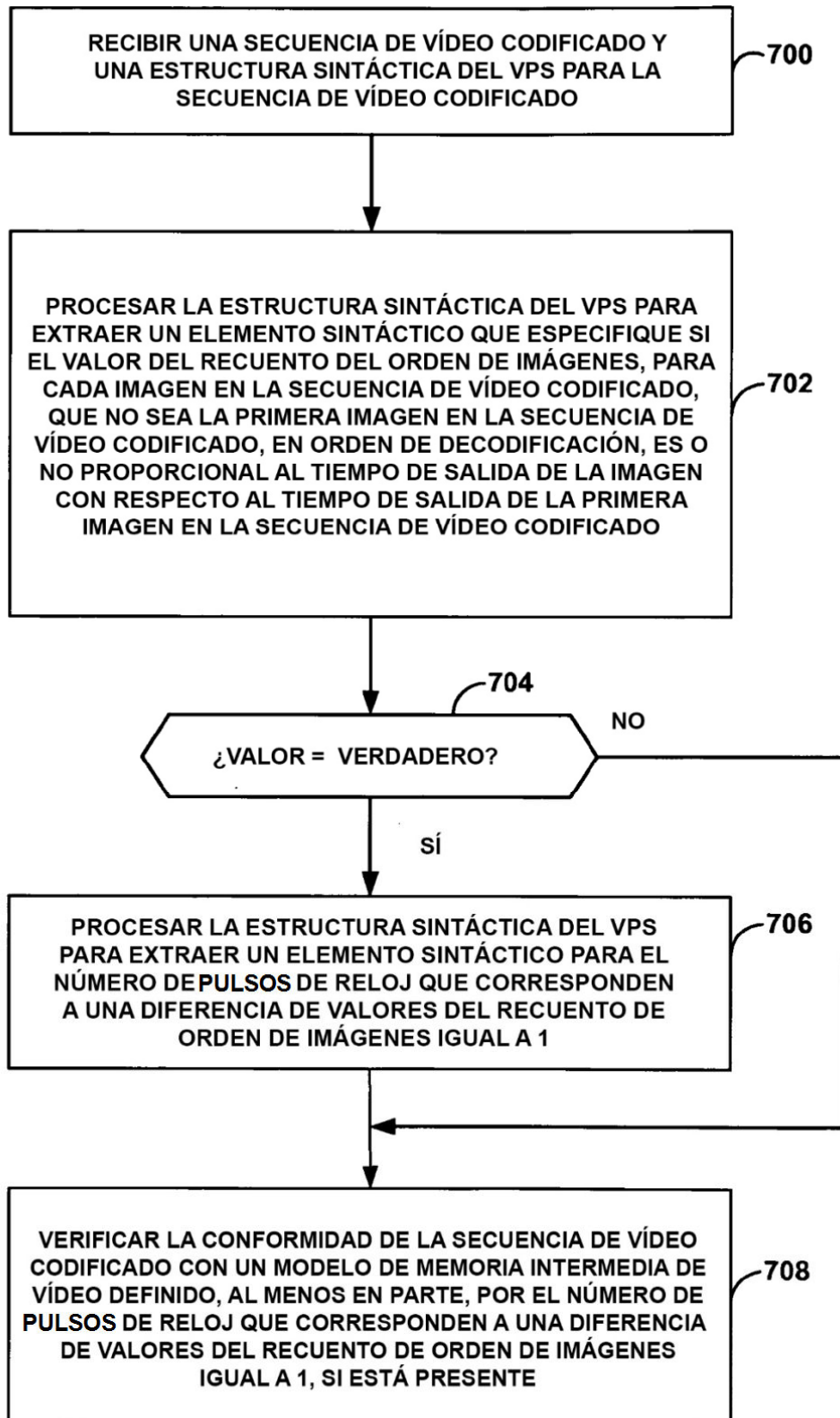


FIG. 10