

19



OFICINA ESPAÑOLA DE  
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 692 190**

51 Int. Cl.:

**H04N 19/70** (2014.01)

**H04N 19/174** (2014.01)

**H04N 19/46** (2014.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

86 Fecha de presentación y número de la solicitud internacional: **30.01.2013 PCT/US2013/023867**

87 Fecha y número de publicación internacional: **08.08.2013 WO13116348**

96 Fecha de presentación y número de la solicitud europea: **30.01.2013 E 13705054 (8)**

97 Fecha y número de publicación de la concesión europea: **25.07.2018 EP 2810441**

54 Título: **Procedimiento de codificación de vídeo y almacenamiento de contenido de vídeo**

30 Prioridad:

**30.01.2012 US 201261592462 P**  
**29.01.2013 US 201313753278**

45 Fecha de publicación y mención en BOPI de la traducción de la patente:  
**30.11.2018**

73 Titular/es:

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

72 Inventor/es:

**WANG, YE-KUI y**  
**CHEN, YING**

74 Agente/Representante:

**FORTEA LAGUNA, Juan José**

ES 2 692 190 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

Procedimiento de codificación de vídeo y almacenamiento de contenido de vídeo

### 5 SOLICITUDES RELACIONADAS

#### CAMPO TÉCNICO

10 [1] Esta divulgación se refiere a la codificación de vídeo y al almacenamiento de contenido de vídeo y, más particularmente, a técnicas para encapsular conjuntos de parámetros y otros parámetros dentro de un archivo de vídeo.

#### ANTECEDENTES

15 [2] Las capacidades del vídeo digital pueden incorporarse en una amplia gama de dispositivos, incluyendo televisores digitales, sistemas de radiodifusión directa digital, sistemas de radiodifusión inalámbrica, asistentes digitales personales (PDA), ordenadores portátiles o de escritorio, ordenadores tipo tableta, lectores de libros electrónicos, cámaras digitales, dispositivos de grabación digital, reproductores de medios digitales, dispositivos de videojuegos, consolas de videojuegos, teléfonos celulares o de radio por satélite, los denominados "teléfonos inteligentes", dispositivos de videoconferencia, dispositivos de transmisión continua de vídeo y similares. Los dispositivos de vídeo digitales 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), la norma de Codificación de Vídeo de Alta Eficiencia (HEVC) actualmente en desarrollo y las ampliaciones de dichas normas. Los dispositivos de vídeo pueden transmitir, recibir, codificar, descodificar y/o almacenar información de vídeo digital más eficazmente, implementando dichas técnicas de compresión de vídeo.

30 [3] Las técnicas de compresión de vídeo realizan la predicción espacial (intraimagen) 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 (es decir, una trama de vídeo o una porción de una trama de vídeo) puede dividirse en bloques de vídeo, que también pueden denominarse bloques de árbol, unidades de árbol de codificación (CTU), bloques de árbol de codificación (CTB), unidades de codificación (CU) y/o nodos de codificación. Los bloques de vídeo en un fragmento intracodificado (I) de una imagen se codifican mediante predicción espacial con respecto a muestras de referencia en bloques contiguos de la misma imagen. Los bloques de vídeo en un fragmento intercodificado (P o B) de una imagen pueden usar la predicción espacial con respecto a muestras de referencia en bloques contiguos de 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.

40 [4] La predicción espacial o temporal da como resultado un bloque predictivo para un bloque a codificar. Los datos residuales representan diferencias de píxeles entre el bloque original a codificar y el bloque predictivo. Un bloque intercodificado se codifica de acuerdo con un vector de movimiento que apunta a un bloque de muestras de referencia que forman el bloque predictivo, y los datos residuales que indican la diferencia entre el bloque codificado y el bloque predictivo. Un bloque intracodificado se codifica de acuerdo con un modo de intracodificación y los datos residuales. Para una mayor compresión, los datos residuales pueden transformarse desde el dominio de los píxeles al dominio de las transformadas, dando como resultado unos coeficientes de transformada residuales, que posteriormente se pueden cuantificar. Los coeficientes de transformada cuantificados, dispuestos inicialmente en una matriz bidimensional, pueden explorarse con el fin de generar un vector unidimensional de coeficientes de transformada, y puede aplicarse la codificación por entropía para lograr aún más compresión.

50 [5] ISO/IEC 14496-15: 2010 "Information Technology - Coding of Audio-visual Objects - Part 15: Advanced Video Coding (AVC) File Format [Tecnología de la información - Codificación de objetos audiovisuales] - Parte 15: Formato de archivo de Codificación Avanzada de Vídeo (AVC)", ISO/IEC JTC1/SC29/WG11, 22 de enero de 2010 describe el formato de archivo AVC.

55 [6] El artículo "High-Level Syntax for Bitstream Extraction [Sintaxis de alto nivel para la extracción de flujos de bits]" de Sjöberg y otros, JCTVC-G607, 8 de noviembre de 2011, describe los cambios en la cabecera de la unidad NAL y SPS.

#### SUMARIO

60 [7] La invención se define en las reivindicaciones a las que está dirigida ahora la referencia. Esta divulgación se refiere a diversas técnicas para almacenar datos de vídeo. En particular, esta divulgación describe una técnica para encapsular conjuntos de parámetros y otros parámetros dentro de un archivo de vídeo.

65 [8] 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 de los dibujos, y a partir de las

reivindicaciones.

**BREVE DESCRIPCIÓN DE LOS DIBUJOS**

5 **[9]**

La FIG. 1 es un diagrama conceptual que ilustra la estructura de un archivo multimedia de ejemplo que puede generarse de acuerdo con las técnicas descritas en esta divulgación.

10 La FIG. 2 ilustra un ejemplo de una unidad de acceso HEVC que puede incluirse en un archivo multimedia generado de acuerdo con las técnicas descritas en esta divulgación.

La FIG. 3 es un diagrama de bloques que ilustra un sistema de codificación y decodificación de vídeo de ejemplo que puede utilizar las técnicas descritas en esta divulgación.

15 La FIG. 4 es un diagrama de bloques que ilustra un ejemplo de un codificador de vídeo.

La FIG. 5 es un diagrama de bloques que ilustra un ejemplo de un módulo de encapsulación que puede implementar las técnicas descritas en esta divulgación.

20 La FIG. 6 es un diagrama conceptual que ilustra la estructura de un archivo multimedia de ejemplo que puede generarse de acuerdo con las técnicas descritas en esta divulgación.

25 La FIG. 7 es un diagrama de flujo que ilustra un ejemplo de generación de un archivo de vídeo que incluye contenido de vídeo codificado de acuerdo con las técnicas de esta divulgación.

**DESCRIPCIÓN DETALLADA**

30 **[10]** En general, esta divulgación se refiere a la codificación de vídeo y al almacenamiento de contenido de vídeo codificado en un archivo. Esta divulgación describe diversas técnicas para almacenar contenido de vídeo que incluyen, por ejemplo, almacenar contenido de vídeo codificado de acuerdo con la norma de vídeo HEVC en un archivo basado en un formato de archivo multimedia basado en ISO (ISO/BMFF). En particular, esta divulgación describe técnicas para encapsular conjuntos de parámetros de secuencia (SPS), conjuntos de parámetros de imagen (PPS) y conjuntos de parámetros adaptativos (APS) incluidos en el contenido de vídeo HEVC en un archivo multimedia. En el caso en que todos los conjuntos de parámetros puedan conocerse al comienzo del archivo y ninguno necesite actualización (es decir, en algunos casos los conjuntos de parámetros con el mismo ID de conjunto de parámetros se actualizan para incluir contenido diferente), los conjuntos de parámetros pueden incluirse en la descripción de la muestra, lo que permite la transmisión de parámetros fuera de banda. En el caso en que no se puedan conocer todos los conjuntos de parámetros al comienzo del archivo, o al menos uno necesite ser actualizado, los conjuntos de parámetros pueden incluirse en las muestras del flujo, o de forma alternativa, el flujo se puede dividir en múltiples subflujos, para cada uno de los cuales los conjuntos de parámetros se pueden incluir en una descripción de la muestra correspondiente. Al permitir que los conjuntos de parámetros se incluyan en la descripción de la muestra o en las muestras del flujo a las que se aplica la descripción de la muestra, no se necesitan pistas adicionales para almacenar solo los conjuntos de parámetros, y un módulo de creación de archivos de vídeo puede negociar entre el número de descripciones de la muestra con la posibilidad de transmisión del conjunto de parámetros fuera de banda.

50 **[11]** Entre las normas de codificación de vídeo se 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 multivista (MVC). Además, existe una nueva norma de codificación de vídeo, concretamente la Codificación de Vídeo de Alta Eficiencia (HEVC) que está siendo desarrollada por el Equipo de Colaboración Conjunta en Codificación de Vídeo (JCT-VC) del Grupo de Expertos en Codificación de Vídeo (VCEG) de ITU-T y el Grupo de Expertos en Imágenes en Movimiento (MPEG) de ISO/IEC. Un borrador de trabajo reciente (WD) de la norma HEVC, denominado "HEVC Working Draft 5 [Borrador de trabajo de HEVC 5]" o "WD5", se describe en el documento JCTVC-G1103 de Bross y otros, titulado "WD5: Working Draft 5 of High-Efficiency Video Coding (HEVC) [WD5: Borrador de trabajo 5 de Codificación de Vídeo de Alta Eficiencia (HEVC)]" Equipo de Colaboración Conjunta en Codificación de Vídeo (JCT-VC) de ITU-T SG16 WP3 e ISO/IEC JTC1/SC29/WG11, séptima reunión: Ginebra, CH, noviembre de 2012. Además, otro borrador de trabajo reciente de la norma HEVC, Borrador de trabajo 7, se describe en el documento HCTVC-HCTVC-11003, de Bross y otros, titulado "High Efficiency Video Coding (HEVC) Text Specification Draft 7 [Codificación de Vídeo de Alta Eficiencia (HEVC)] Borrador de trabajo 7" Equipo de Colaboración Conjunta en Codificación de Vídeo (JCT-VC) de ITU-T SG16 WP3 e ISO/IEC JTC1/SC29/WG11, novena reunión: Ginebra, Suiza, 27 de abril de 2012 al 7 de mayo de 2012. La norma HEVC también puede denominarse ISO/IEC 23008-HEVC, que se pretende que sea el número de la norma para la versión publicada de HEVC.

**[12]** Para almacenar contenido de vídeo codificado de acuerdo con una norma de vídeo particular, se necesita una especificación de formato de archivo correspondiente a la norma de vídeo particular. El último borrador del formato de archivo HEVC utiliza ISOBMFF como base para su formato de encapsulación. El último borrador del formato de archivo HEVC, denominado en el presente documento HEVCFF, se define según ISO/IEC 14496-15: 2010/Amd.1:2011(E) e incluye las siguientes referencias normativas:

(1) ISO/IEC 14496-1:2001, Tecnología de la información - Codificación de objetos audiovisuales - Parte 1: Sistemas;

(2) ISO/IEC 23008-HEVC; y

(3) ISO/IEC 14496-12, Tecnología de la información - Codificación de objetos audiovisuales - Parte 12: formato de archivo multimedia basado en ISO (es decir, ISOBMFF)

cada uno de los cuales son incorporados por referencia en su totalidad.

**[13]** Cabe señalar que ISO/IEC 14496-12 es técnicamente idéntica a ISO / IEC 15444-12. El ISOBMFF se utiliza como base para muchos formatos de encapsulación de la norma de codificación de vídeo, así como para muchos formatos de contenedor multimedia. Ejemplos de otros formatos de archivo basados en ISOBMFF, incluyen, por ejemplo, el formato de archivo MPEG-4 (ISO/IEC 14496-14), el formato de archivo 3GPP (3GPP TS 26.244) y el formato de archivo AVC (ISO/IEC 14496-15) (es decir, AVCFF). Además, debe observarse que aunque las técnicas divulgadas en el presente documento se describen en algunos ejemplos con respecto a HEVC y pueden hacer referencia a formatos de archivo particulares, pueden ser igualmente aplicables a otras normas de codificación de vídeo y formatos de archivo.

**[14]** Los archivos estructurados de acuerdo con ISOBMFF se pueden usar para muchos fines, incluida la reproducción local de archivos multimedia, la descarga progresiva de un archivo remoto, segmentos para la transmisión en tiempo real adaptativa dinámica sobre HTTP (DASH), contenedores para transmitir contenido y sus instrucciones de empaquetado y grabación de flujos de medios recibidos en tiempo real. En general, ISOBMFF permite que los datos multimedia continuos, como los datos de audio y vídeo, se almacenen independientemente de los metadatos. En un ejemplo, los datos multimedia pueden incluir un grupo de imágenes codificadas de acuerdo con HEVC. Los metadatos incluyen típicamente información de presentación y temporización de modo que porciones de datos de medios se pueden descodificar independientemente.

**[15]** ISOBMFF usa un esquema orientado a objetos en el cual elementos sintácticos elementales se usan como bloques de construcción para formar un archivo multimedia. En ISOBMFF estos elementos sintácticos elementales se conocen como "cajas". Sin embargo, los formatos de archivo específicos basados en ISOBMFF pueden referirse a cajas que usan terminología diferente, como "átomos". Además, debe observarse que, cuando se describen formatos de archivos específicos basados en ISOBMFF, la terminología específica a menudo se puede usar indistintamente con la terminología ISOBMFF de una manera no limitativa. Por ejemplo, cuando se describe HEVCFF, el término "unidad de acceso" puede corresponderse con el término "muestra" en el ISOBMFF y una unidad de acceso puede describirse usando cualquiera de los términos. En ISOBMFF, una caja incluye un tipo de cuatro caracteres, el recuento de bytes de la caja y la carga útil. El tipo de caja define la relación lógica entre una caja en particular y otros tipos de cajas. El tipo de caja también puede describir qué datos son obligatorios para el tipo de caja particular y qué datos son opcionales para el tipo de caja particular. Una caja puede ser parte de una secuencia o grupo de cajas y puede contener otras cajas, que pueden denominarse subcajas.

**[16]** En ISOBMFF, una caja de película ("moov") contiene los metadatos de los flujos continuos de medios presentes en el archivo. Los metadatos para los flujos de medios continuos presentes en el archivo se pueden encerrar en cajas adicionales en la caja de película. Los flujos de medios continuos se pueden representar en un archivo ISOMBFF como una pista, donde una pista es una estructura lógica que hace referencia a metadatos que pueden estar encerrados dentro de la caja de película y en la que el flujo de medios consiste en una secuencia de muestras. En ISOBMFF, un flujo de medios puede incluirse dentro de cajas, como una caja de Datos de medios ("mdat") y subcajas de las mismas. Además, en algunos ejemplos, un flujo de medios puede almacenarse directamente en un archivo ISOBMFF. El contenido multimedia de una pista consiste en una secuencia de muestras, como unidades de acceso de vídeo. Una secuencia de muestras se puede referir como un flujo. El ISOBMFF especifica una pista de medios que contiene un flujo de medios elemental. Los metadatos para cada pista incluyen una lista de entradas de descripción de la muestra, cada una de las cuales proporciona el formato de codificación o encapsulación utilizado en la pista y los datos de inicialización necesarios para procesar ese formato. Cada muestra está asociada con una de las entradas de descripción de la muestra de la pista. Algunos tipos de archivos basados en ISOBMFF también incluyen pistas de indicación. Las pistas de indicación incluyen información que facilita la reconstrucción del contenido multimedia de modo que la calidad y la fiabilidad mejoran cuando el contenido multimedia se transmite a través de una red.

**[17]** Similar a un archivo almacenado de acuerdo con ISOBMFF, un archivo HEVC puede incluir una serie de flujos elementales que hacen referencia a los metadatos. En el HEVCFF, una muestra puede denominarse una

unidad de acceso como se define en ISO/IEC 23008-HEVC. En el HEVCFF, los metadatos también se pueden incluir en las entradas de descripción de la muestra. El HEVCFF puede permitir indicaciones, pero debe tenerse en cuenta que lo que las pistas de indicación denominan "tramas B" en algunos formatos de archivo, en realidad, pueden ser imágenes "desechables" o imágenes no de referencia como se define en ISO/IEC 23008-HEVC. Además, un formato de archivo utilizado para el almacenamiento de contenido de vídeo HEVC puede usar las capacidades existentes de ISOBMFF, pero también puede definir ampliaciones para admitir características de contenido de vídeo codificado HEVC específicas tales como: conjuntos de parámetros, escalabilidad temporal y punto de recuperación de acceso aleatorio.

**[18]** El término conjuntos de parámetros puede referirse a un conjunto de parámetros de secuencia, un conjunto de parámetros de imagen o un conjunto de parámetros de adaptación, como se define en ISO/IEC 23008-HEVC. Los mecanismos de conjunto de parámetros de secuencia e imagen pueden desacoplar la transmisión de información que cambia poco frecuentemente de la transmisión de datos de bloque codificados. De acuerdo con HEVC, cada fragmento que contiene datos de bloque codificados puede hacer referencia a un conjunto de parámetros de imagen que contiene sus parámetros de descodificación. Además, un conjunto de parámetros de imagen puede hacer referencia a un conjunto de parámetros de secuencia que contiene información de parámetros de descodificación de nivel de secuencia. Además, HEVC también admite conjuntos de parámetros de adaptación. Los conjuntos de parámetros de adaptación pueden contener parámetros de descodificación que se espera que cambien con más frecuencia que los parámetros de codificación en los conjuntos de parámetros de imagen. Un conjunto de parámetros de adaptación también puede ser referenciado por un fragmento, cuando sea necesario. Cabe señalar que AVC no admite actualmente conjuntos de parámetros de adaptación.

**[19]** HEVCFF puede admitir la agrupación de muestras de escalabilidad temporal al incluir un mecanismo de estructuración y agrupamiento para indicar la asociación de unidades de acceso con diferentes niveles de jerarquización de escalabilidad temporal. La escalabilidad temporal, en general, permite que el contenido de vídeo codificado se reproduzca a diversas velocidades de trama (por ejemplo, 30 tramas por segundo (fps) o 60 fps) a varias velocidades de bits. En un ejemplo, una pista de vídeo HEVC puede contener cero o una instancia de una SampleToGroupBox con un grouping\_type 'tscl'. Esta instancia de SampleToGroupBox puede representar la asignación de muestras en la pista a capas temporales. Una instancia que acompaña a la SampleGroupDescriptionBox con el mismo tipo de agrupación puede, si está presente, contener entradas de grupo de muestra TemporalLayerEntry que describen las capas temporales. Por lo tanto, una entrada de grupo de muestra de capa temporal se puede definir en HEVCFF como tipo de Grupo: 'tscl'; Contenedor: Caja de descripción de grupo de muestra ('sgpd'); Obligatorio: No; y Cantidad: cero o más muestras.

**[20]** En algunos casos, una entrada de grupo de muestra de capa temporal puede definir la información de capa temporal para todas las muestras en una capa temporal. Las capas temporales se pueden numerar con enteros no negativos, cada capa temporal se puede asociar con un valor particular de temporal\_id como se define en ISO/IEC 23008-HEVC. En un ejemplo, una capa temporal asociada con un valor temporal\_id mayor que 0 puede depender de todas las capas temporales asociadas con valores temporal\_id menores. Una representación de capa temporal (que también puede denominarse la representación de una capa temporal) asociada con un valor temporal\_id particular puede consistir en todas las capas temporales asociadas con valores temporal\_id iguales o mayores que el valor temporal\_id particular.

**[21]** El siguiente es un ejemplo de sintaxis que puede usarse para una entrada de capa temporal:

```

class TemporalLayerEntry() extends VisualSampleGroupEntry ('tscl')
{
    unsigned int(8) temporalLayerId;
    unsigned int(8) tlProfileIndication;
    unsigned int(8) tlProfileCompatibility;
    unsigned int(8) tlLevelIndication;
    unsigned int(16) tlMaxBitRate;
    unsigned int(16) tlAvgBitRate;
    unsigned int(8) tlConstantFrameRate;
    unsigned int(16) tlAvgFrameRate;
}

```

**[22]** En la sintaxis de entrada de la capa temporal de ejemplo anterior, la semántica se puede definir de la

siguiente manera:

**temporalLayerId** puede dar el ID de la capa temporal. En algunos ejemplos, para todas las muestras que son miembros de este grupo de muestra, las unidades de la capa de abstracción de red (NCL) de la capa de codificación de vídeo (VCL) (las unidades NAL de VCL se describen con más detalle a continuación) tendrán `temporal_id`, como se define en ISO/IEC 23008-HEVC, igual a `temporalLayerId`.

**tiProfileIndication** puede contener el código de perfil (`profile_idc`) como se define en ISO/IEC 23008-HEVC, para la representación de la capa temporal identificada por `temporalLayerId`.

**tiProfileCompatibility** puede ser un byte definido exactamente igual que el byte que se produce entre el `profile_idc` y el `level_idc` en un SPS, como se define en ISO/IEC 23008-HEVC, para la representación de la capa temporal identificada por `temporalLayerId`.

**LevelIndication** puede contener el código de nivel (`level_idc`) como se define en ISO/IEC 23008-10, para la representación de la capa temporal identificada por `temporalLayerId`.

**tiMaxBitrate** puede dar la velocidad máxima en 1000 bits por segundo sobre cualquier ventana de un segundo, para la representación de la capa temporal identificada por `temporalLayerId`.

**tiAvgBitRate** puede dar la velocidad de bits promedio en unidades de 1000 bits por segundo, para la representación de la capa temporal identificada por `temporalLayerId`.

**tiConstantFrameRate** puede ser igual a 1 para indicar que la representación de la capa temporal identificada por `temporalLayerId` es de velocidad de trama constante. El valor cero puede indicar que la representación de la capa temporal identificada por `temporalLayerId` puede ser o no de velocidad de trama constante.

**tiAvgFrameRate** puede dar la velocidad de trama promedio en unidades de tramas/(256 segundos), para la representación de la capa temporal identificada por `temporalLayerId`.

[23] Además de la agrupación de muestras de escalabilidad temporal, el borrador de trabajo de HEVC también incluye el concepto de "actualización de descodificación gradual" o punto de recuperación de acceso aleatorio. El punto de recuperación de acceso aleatorio se puede señalar en el flujo de bits utilizando el mensaje de mejora complementaria (SEI) del punto de recuperación. Este mensaje se puede encontrar al comienzo del acceso aleatorio y puede indicar cuántos datos se deben descodificar después de la unidad de acceso en la posición del mensaje SEI antes de que se complete la recuperación. El concepto de recuperación gradual se admite en el HEVCFF utilizando la agrupación de muestra del tipo "rollo" como se define en ISO/IEC 14496-12. La membresía del grupo puede marcar la muestra que contiene el mensaje SEI de tal manera que la "distancia de rollo" se limita a ser solo positiva (es decir, un postrollo). Cabe señalar que la forma en que el grupo de rollos cuenta las muestras en el formato de archivo puede no coincidir con la forma en que se representan las distancias en el mensaje SEI. Por lo tanto, dentro de un flujo, puede ser necesario marcar el comienzo del prerollo, de modo que un descodificador de flujo puede comenzar a descodificar desde allí. Sin embargo, en un archivo, cuando se realiza un acceso aleatorio, puede desearse una búsqueda determinista para la trama anterior más cercana que se puede descodificar perfectamente (por ejemplo, una muestra de sincronización o el final de un prerollo).

[24] La FIG. 1 es un diagrama conceptual que ilustra la estructura de un archivo multimedia HEVC de ejemplo basado en el ISOBMFF y el HEVCFF actual. Debe observarse que el archivo 142 multimedia HEVC de ejemplo de la FIG. 1 pretende ilustrar la relación lógica entre flujos de datos de vídeo y metadatos. En aras de la brevedad, no se ha ilustrado la estructura de encapsulación completa del archivo 142 multimedia HEVC. Sin embargo, el archivo 142 multimedia HEVC puede utilizar cajas y estructuras de encapsulación definidas de acuerdo con HEVCFF. En el ejemplo ilustrado en la FIG. 1, el archivo 142 multimedia HEVC incluye una caja 144 de metadatos y flujos 146A-146N de datos de vídeo. La caja 144 de metadatos puede ser similar a una caja "moov" de ISOBMFF y puede contener metadatos para cada uno de los flujos 146A-146N de datos de vídeo. En un ejemplo, los metadatos pueden incluirse dentro de las tablas de metadatos 148A-148N. Las tablas de metadatos 148A-148N pueden incluir entradas de muestra que corresponden a los flujos 146A-146N de datos de vídeo. Un ejemplo de una tabla de metadatos es una caja de tabla de muestra. Cada uno de los flujos 146A-146N de datos de vídeo puede hacer referencia a una o más de las tablas de metadatos 148A-148N.

[25] Los flujos 146A-146N de datos de vídeo pueden corresponder a las pistas de vídeo o visuales descritas en ISOBMFF. Por lo tanto, los flujos 146A-146N de datos de vídeo pueden usar, por lo tanto: (1) un `handler_type` de 'vide' en la `HandlerBox`; (2) una cabecera de medios de vídeo 'vmhd'; y (3) una derivada de `VisualSampleEntry`. Un flujo de vídeo puede estar representado por una pista de vídeo en el archivo 142 multimedia HEVC. Cada uno de los flujos 146A-146N de datos de vídeo puede incluir una o más unidades de acceso 150A-150N. Las unidades de acceso 150A-150N del archivo 142 multimedia HEVC se pueden definir de acuerdo con HEVCFF. Las unidades de acceso 150A-150N pueden corresponder a una muestra en ISOBMFF. Las unidades de acceso 150A-150N pueden disponerse en tramas externamente y tienen un tamaño suministrado por esa disposición en tramas externa. En

algunos casos, la caja de metadatos 144 puede incluir información que define el tamaño de las unidades de acceso 150A-150N. Las unidades de acceso 150A-150N pueden incluir un conjunto de unidades NAL (es decir, capa de abstracción de red) 152A-152N. Las unidades NAL 152A-152N se pueden definir de acuerdo con HEVCFF.

5 [26] Cada unidad NAL 152A-152N puede incluir un campo de longitud que indica la longitud de la unidad NAL en bytes. En algunos ejemplos, el campo de longitud se puede configurar para ser uno, dos o cuatro bytes. Los campos de longitud permiten un escaneo fácil de las unidades NAL dentro de una unidad de acceso. Las unidades NAL 152A-152N pueden contener datos de unidad NAL como se especifica en ISO/IEC 23008-HEVC. Además de otros tipos de unidades NAL, HEVCFF define los siguientes tipos de unidades NAL: Las unidades NAL de datos de vídeo pueden denominarse unidades NAL de VCL, unidades NAL de información de mejora complementaria (SEI) y unidades NAL delimitadoras de la unidad de acceso (AU), así como otros tipos de unidades NAL. Cabe señalar que los tipos de unidades NAL que están reservados en ISO/IEC 23008-HEVC pueden adquirir una definición en el futuro. Algunos lectores de archivos pueden configurarse para ignorar unidades NAL con valores reservados de tipo de unidad NAL cuando están presentes en estas matrices. Cabe señalar que este comportamiento "tolerante" está diseñado para que no se generen errores, lo que permite la posibilidad de ampliaciones compatibles con versiones anteriores de estas matrices en futuras especificaciones.

20 [27] Las unidades NAL de VCL pueden formatearse de forma tal que todas las unidades NAL de fragmento codificadas para una sola imagen estén contenidas dentro de la unidad de acceso cuyo tiempo de descodificación y tiempo de composición son los de la imagen. Además, se puede requerir que las unidades de acceso 150A-150N contengan al menos una unidad NAL de datos de vídeo.

25 [28] Los flujos 146A-146N de datos de vídeo también pueden formatearse de manera que todas las unidades NAL de SEI estén contenidas en la unidad de acceso cuyo tiempo de descodificación se produce antes del momento en que los mensajes SEI entran en vigor. El orden de los mensajes SEI dentro de las unidades de acceso 150A-150N puede ser como se define según ISO/IEC 23008-HEVC. Además, el archivo HEVC 142 puede requerir que los mensajes SEI para una imagen estén incluidos en la unidad de acceso que contiene esa imagen y que los mensajes SEI pertenecientes a una secuencia de imágenes estén incluidos en la unidad de acceso que contiene la primera imagen de la secuencia a la que mensaje SEI pertenece. Las unidades NAL del delimitador de unidades de acceso pueden estar restringidas de acuerdo con ISO/IEC 23008-HEVC. El archivo HEVC 142 puede requerir que la secuencia de unidades NAL en los flujos 146A-146N de datos de vídeo y dentro de una única unidad de acceso estén dispuestas en un orden de descodificación válido como se especifica en ISO/IEC 23008-HEVC.

35 [29] La FIG. 2 ilustra un ejemplo de una unidad de acceso HEVC que incluye unidades NAL. Como se ilustra en la FIG. 2, la unidad de acceso 150 incluye una unidad NAL de delimitador AU, una unidad NAL de SEI y dos unidades NAL de fragmento. Además, debe observarse que el HEVCFF puede soportar el concepto de submuestras descrito en ISOMBFF. Una submuestra se define como una o más unidades NAL contiguas dentro de una muestra y que tienen el mismo valor de los siguientes campos; RefPicFlag y VclNalUnitFlag. En el ejemplo, la unidad de acceso 150 ilustrada en la FIG. 2, las dos unidades NAL de fragmento pueden formar una submuestra. En la Sección 8.7.7 de ISO/IEC 14496-12, se describe un ejemplo de una caja de información de submuestra que puede usarse para un flujo de HEVC. Una submuestra puede incluir múltiples unidades NAL y su campo o campos de longitud de unidad NAL precedente. Cuando se usan submuestras, la presencia de una caja de información de submuestra puede ser opcional. Sin embargo, de acuerdo con HEVCFF, si la caja de información de submuestra está presente en una pista que contiene datos HEVC, puede requerirse que el campo de subsample\_priority se establezca en un valor de acuerdo con la especificación de este campo en ISO/IEC 14496-12. Además, puede requerirse que el campo descartable se establezca a 1 solo si esta muestra todavía puede descodificarse si esta submuestra se descarta (por ejemplo, la submuestra consiste en una unidad NAL de SEI). El campo reservado se define en HEVCFF de acuerdo con la siguiente sintaxis de ejemplo:

50 *unsigned int(1) RefPicFlag;*  
*unsigned int(1) VclNalUnitFlag;*  
*unsigned int(30) reserved = 0;*

55 [30] En la sintaxis de ejemplo anterior, la semántica se puede definir de la siguiente manera:

**RefPicFlag** igual a 0 puede indicar que todas las unidades NAL en la submuestra tienen nal\_ref\_flag igual a 0. RefPicFlag igual a 1 puede indicar que todas las unidades NAL en la submuestra tienen nal\_ref\_flag igual a 1.

60 **VclNalUnitFlag** igual a 0 puede indicar que todas las unidades NAL en la submuestra son unidades NAL no VCL. El valor 1 puede indicar que todas las unidades NAL en la submuestra son unidades NAL de VLC.

[31] Con referencia de nuevo a la FIG. 1, el archivo 142 multimedia HEVC puede formatearse de manera que toda la información de temporización sea externa a los flujos 146A-146N de datos de vídeo. Sin embargo, en algunos ejemplos, el archivo HEVC 142 puede incluir mensajes SEI de temporización de imagen que definen marcas

de tiempo de presentación o composición en los flujos 146A-146N de datos de vídeo, ya que el mensaje SEI de temporización de imagen puede contener información distinta a la temporización y puede requerirse para verificación de conformidad. El archivo 142 multimedia HEVC puede almacenar toda la información de temporización en las tablas de metadatos 148A-148N de la muestra. La información de temporización almacenada en las tablas de metadatos 148A-148N puede anular cualquier temporización proporcionada en los flujos 146A-146N de datos de vídeo. Además, el archivo 142 multimedia HEVC puede configurarse de modo que la información de temporización proporcionada dentro de un flujo sea ignorada por un descodificador ya que puede contradecir la información de temporización proporcionada en otro archivo 142 multimedia HEVC y, además, en algunos casos la información de temporización proporcionada dentro de un flujo puede no ser correcta o coherente dentro de sí misma. Debe observarse que las restricciones de información de temporización pueden ser impuestas debido al hecho de que la edición de post-compresión, combinación o re-temporización de un flujo en el nivel de formato de archivo puede invalidar o hacer incoherente cualquier información de temporización incorporada presente dentro del flujo.

**[32]** En algunos ejemplos, el archivo 142 multimedia HEVC puede configurarse de modo que las unidades de acceso 150A-150N se almacenen en orden de descodificación. Por lo tanto, si no se utiliza el reordenamiento de la imagen y el tiempo de descodificación (DTS) y el tiempo de composición (CTS) son los mismos, entonces la presentación es igual que el orden de descodificación y solo se necesita usar una tabla de tiempos-de-muestra 'stts'. Cabe señalar que cualquier tipo de imagen puede reordenarse en vídeo HEVC, no solo en imágenes B. Si el tiempo de descodificación y el tiempo de composición difieren, la tabla "ctts" del tiempo-de-muestra de composición también se puede usar junto con la tabla 'stts'. En algunos ejemplos, una unidad de acceso puede considerarse como una muestra de sincronización si las unidades NAL de VCL en la unidad de acceso indican que la imagen codificada contenida en la unidad de acceso es una imagen de actualización de descodificación instantánea (IDR).

**[33]** En algunos casos, el archivo 142 multimedia HEVC puede configurarse de manera que se requiera que los flujos 146A-146N de datos de vídeo no incluyan los códigos de inicio. Sin embargo, los sistemas que desean entregar un flujo para el archivo 142 multimedia HEVC usando códigos de inicio pueden hacerlo formateando de nuevo los flujos de vídeo 146A-146N para introducir códigos de inicio. Además, los datos de vídeo pueden representarse naturalmente como una tasa de bits variable en el archivo 142 multimedia HEVC y pueden llenarse para la transmisión si es necesario. Por lo tanto, el archivo 142 multimedia HEVC puede configurarse de modo que los flujos 146A-146N de datos de vídeo no requieren incluir unidades NAL de datos de relleno y mensajes SEI de datos de relleno. Debe observarse que si el archivo 142 multimedia HEVC no contiene unidades NAL de datos de relleno, códigos de inicio, elementos sintácticos de cero bytes y/o mensajes SEI de datos de relleno, las características de flujo de bits del archivo 142 multimedia HEVC pueden cambiar con respecto a la conformidad con el descodificador de referencia hipotético (HRD) cuando se opera el HRD en modo de velocidad constante de bits (CBR) como se especifica en ISO/IEC 23008-HEVC, Anexo C.

**[34]** Similar a HEVCFF, AVCFF también incluye ISO/IEC 14496-1:2001 como referencia normativa y AVCFF incluye un diseño similar de la interfaz del sistema (también conocido como sintaxis de alto nivel) como el borrador de trabajo de HEVC. Por lo tanto, algunas partes del archivo 142 multimedia HEVC pueden formatearse de una manera similar a un archivo AVC. Sin embargo, algunos aspectos del AVCFF pueden no proporcionar un formato óptimo del contenido de vídeo codificado HEVC.

**[35]** Un aspecto del AVCFF es que, dado que los conjuntos de parámetros pueden incluirse en la descripción de la muestra o en una pista de conjunto de parámetros separada, puede haber un aumento innecesario del número de pistas o del número de descripciones de la muestra, donde cada una corresponde a un subconjunto del flujo de vídeo completo en una pista de vídeo. Cabe señalar que el HEVCFF actual no permite pistas de conjuntos de parámetros por separado.

**[36]** Otro aspecto del AVCFF es que se agrupan diferentes tipos de conjuntos de parámetros, ya estén incluidos en la descripción de la muestra o en una pista de conjunto de parámetros por separado. Por lo tanto, en los archivos AVC que no usan una pista de parámetros separada, no es posible incluir algunos tipos de conjuntos de parámetros en una descripción de la muestra mientras se excluyen otros. Por lo tanto, en este caso, SPS y PPS se almacenan en la descripción de la muestra, aunque PPS puede cambiar más frecuentemente. De esta manera, PPS no se puede transmitir independientemente de SPS.

**[37]** Otro aspecto del AVCFF es que una descripción de la muestra incluye una matriz para un SPS y otra matriz para PPS. En AVCFF estas matrices también pueden contener unidades NAL de SEI que contienen mensajes SEI declarativos. Por lo tanto, un analizador de archivos debería verificar la cabecera de la unidad NAL de cada unidad NAL incluida en una matriz SPS o PPS para determinar si la unidad NAL es una unidad NAL de conjunto de parámetros o una unidad NAL de SEI.

**[38]** Otro aspecto del formato de archivo AVC es que las propiedades de transmisión como la resolución espacial y la velocidad de tramas que son importantes para la selección de pista o flujo (por ejemplo, en aplicaciones de transmisión adaptativa basadas en archivos) no se pueden almacenar en la descripción de la muestra y por lo tanto no son convenientemente accesibles con el diseño AVCFF.



**[39]** Otro aspecto del formato de archivo AVC es que las propiedades de escalabilidad temporal, como el perfil, el nivel y la velocidad de trama, que son importantes para la selección de un subconjunto temporal particular de un flujo, faltan para cada representación de capa temporal.

**[40]** La FIG. 3 es un diagrama de bloques que ilustra un sistema de codificación y transmisión de vídeo de ejemplo que puede utilizar las técnicas descritas en esta divulgación. En un ejemplo, el sistema 10 puede configurarse para encapsular un PPS dentro de las unidades de acceso de un flujo de vídeo, en el que la descripción de la muestra incluye un indicador que identifica el número de PPS almacenados dentro de una o más unidades de acceso del flujo de vídeo. Como se muestra en la FIG. 3, el sistema 10 incluye un dispositivo de origen 12 que genera datos de vídeo codificados que un dispositivo de destino 14 va a descodificar en un momento posterior. El dispositivo de origen 12 y el dispositivo de destino 14 pueden configurarse de manera que el archivo 142 multimedia HEVC, descrito con respecto a la FIG. 1 puede generarse en el dispositivo de origen 12 y transmitirse al 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 tipo tableta, descodificadores, equipos telefónicos tales como los denominados teléfonos "inteligentes", los denominados paneles "inteligentes", televisores, cámaras, dispositivos de visualización, reproductores multimedia digitales, consolas de videojuegos, un dispositivo de transmisión continua 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.

**[41]** En el ejemplo de la FIG. 3, el dispositivo de origen 12 incluye una fuente de vídeo 18, codificador de vídeo 20, unidad de encapsulación 21 y una interfaz de salida 22. La fuente de vídeo 18 puede incluir una fuente tal como un dispositivo de captura de vídeo, por ejemplo, una videocámara, un archivo de vídeo que contiene 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 vídeo de origen, o una combinación de dichas fuentes. El vídeo capturado, precapturado o generado por ordenador puede ser codificado por el codificador de vídeo 20. El codificador de vídeo 20 puede funcionar de acuerdo con una norma de compresión de vídeo, tal como HEVC y puede ajustarse al modelo de prueba HEVC (HM). Además, el codificador de vídeo 20 puede funcionar de acuerdo con otras normas de vídeo descritas en el presente documento.

**[42]** Para la codificación de vídeo de acuerdo con el HEVC, una trama de vídeo puede dividirse en unidades de codificación. Una unidad de codificación (CU) se refiere en general a una región de imagen que sirve de unidad básica a la que se aplican diversas herramientas de codificación para la compresión de vídeo. Una CU por lo general tiene una componente de luminancia, denominada Y, y dos componentes de crominancia, denominadas U y V. Según el formato de muestreo de vídeo, el tamaño de las componentes U y V, en términos del número de muestras, puede ser el mismo o diferente al tamaño de la componente Y. Una CU es típicamente cuadrada, y puede considerarse similar a un llamado macrobloque, por ejemplo, en virtud de otras normas de codificación de vídeo, tales como AVC. La codificación, de acuerdo con algunos de los aspectos propuestos actualmente de la norma HEVC en proceso de elaboración, se describirá en esta solicitud con fines ilustrativos. Sin embargo, las técnicas descritas en esta divulgación pueden ser útiles para otros procesos de codificación de vídeo, tales como los que se definen de acuerdo con AVC u otras normas o procesos de codificación de vídeo patentados. Los esfuerzos de normalización de la HEVC se basan en un modelo de un dispositivo de codificación de vídeo denominado Modelo de Prueba de HEVC (HM). El HM supone varias capacidades de los dispositivos de codificación de vídeo respecto a dispositivos de acuerdo con, por ejemplo, la norma ITU-T H.264/AVC. Por ejemplo, mientras que AVC proporciona nueve modos de codificación por intrapredicción, HM proporciona hasta treinta y cuatro modos de codificación por intrapredicción.

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

**[44]** De acuerdo con el HM, una CU puede incluir una o más unidades de predicción (PU) y/o una o más unidades de transformada (TU). Los datos sintácticos dentro de un flujo de bits pueden definir una unidad de codificación máxima (LCU), que es la CU más grande en términos del número de píxeles. En general, una CU tiene un propósito similar a un macrobloque de la norma H.264, excepto en que una CU no tiene una distinción de tamaño. Por tanto, una CU puede dividirse en sub-CU. En general, las referencias de esta divulgación a una CU pueden referirse a la unidad de codificación máxima de una imagen o a una sub-CU de una LCU. Una LCU puede dividirse en sub-CU y cada sub-CU puede dividirse además en sub-CU. Los datos sintácticos para un flujo de bits pueden definir un número máximo de veces en que puede dividirse una LCU, denominada profundidad de CU. Por consiguiente, un flujo de bits también puede definir una unidad de codificación de mínimo tamaño (SCU). Esta divulgación también usa el término "bloque" o "porción" para referirse a cualquiera entre una CU, una PU o una TU.

En general, "porción" puede referirse a cualquier subconjunto de una trama de vídeo.

**[45]** Una LCU puede asociarse a una estructura de datos de árbol cuaternario. En general, una estructura de datos en árbol cuaternario incluye un nodo por CU, donde un nodo raíz corresponde a la LCU. Si una CU se divide en cuatro sub-CU, el nodo correspondiente a la CU incluye cuatro nodos hoja, cada uno de los cuales corresponde a una de las sub-CU. Cada nodo de la estructura de datos en árbol cuaternario puede proporcionar datos sintácticos para la CU correspondiente. Por ejemplo, un nodo en el árbol cuaternario puede incluir un indicador de división, que indica si la CU correspondiente al nodo está dividida o no en varias sub-CU. Los elementos sintácticos para una CU pueden definirse de manera recursiva y pueden depender de si la CU está dividida en varias sub-CU. Si una CU no está dividida adicionalmente, se denomina CU hoja. En esta divulgación, 4 sub-CU de una CU hoja también se denominarán CU hoja aunque no haya una división explícita de la CU hoja original. Por ejemplo, si una CU con un tamaño de 16x16 no se divide adicionalmente, las cuatro sub-CU de tamaño 8x8 también se denominarán CU hoja aunque la CU de tamaño 16x16 no se haya dividido nunca.

**[46]** Además, las TU de las CU hoja también pueden asociarse a respectivas estructuras de datos de árbol cuaternario. Es decir, una CU hoja puede incluir un árbol cuaternario que indica cómo la CU hoja está dividida en varias TU. Esta divulgación hace referencia al árbol cuaternario que indica cómo una LCU está dividida como un árbol cuaternario de CU, e indicando el árbol cuaternario cómo una CU hoja está dividida en varias TU como un árbol cuaternario de TU. El nodo raíz de un árbol cuaternario de TU corresponde en general a una CU hoja, mientras que el nodo raíz de un árbol cuaternario de CU corresponde en general a una LCU. Las TU del árbol cuaternario de TU que no están divididas se denominan TU hoja.

**[47]** Una CU hoja puede incluir una o más unidades de predicción (PU). En general, una PU representa la totalidad o una parte de la CU correspondiente, y puede incluir datos para recuperar una muestra de referencia para la PU. Por ejemplo, cuando la PU se codifica de forma intermodal, la PU puede incluir datos que definan un vector de movimiento para la PU. Los datos que definen el vector de movimiento pueden describir, por ejemplo, una componente horizontal del vector de movimiento, una componente vertical del vector de movimiento, una resolución para el vector de movimiento (por ejemplo, con una precisión de un cuarto de píxel o con una precisión de un octavo de píxel), una trama de referencia a la que apunte el vector de movimiento y/o una lista de referencia (por ejemplo, lista 0 o lista 1) para el vector de movimiento. Los datos de la CU hoja que definan la(s) PU también pueden describir, por ejemplo, una división de la CU en una o más PU. Las modalidades de división pueden diferir en función de si la CU está sin codificar, codificada en la modalidad de intrapredicción o codificada en la modalidad de interpredicción. Para la intracodificación, una PU puede tratarse de la misma forma que una unidad de transformada de hoja descrita a continuación.

**[48]** En un ejemplo, el HM soporta la predicción en diversos tamaños de PU. Suponiendo que el tamaño de una CU particular es  $2N \times 2N$ , el HM admite intrapredicción en tamaños de PU de  $2N \times 2N$  o  $N \times N$  e interpredicción en tamaños de PU simétricos de  $2N \times 2N$ ,  $2N \times N$ ,  $N \times 2N$  o  $N \times N$ . El HM también admite la división asimétrica para la interpredicción en tamaños de PU de  $2N \times nU$ ,  $2N \times nD$ ,  $nL \times 2N$  y  $nR \times 2N$ . En la división asimétrica, una dirección de una CU no está dividida, mientras que la otra dirección está dividida en el 25 % y el 75 %. La parte de la CU correspondiente a la división del 25 % está indicada por una "n" seguida de una indicación de "arriba", "abajo", "izquierda" o "derecha". Así pues, por ejemplo, " $2N \times nU$ " se refiere a una CU de tamaño  $2N \times 2N$  que está dividida horizontalmente con una PU de tamaño  $2N \times 0,5N$  encima y una PU de tamaño  $2N \times 1,5N$  debajo.

**[49]** En esta divulgación, " $N \times N$ " y " $N$  por  $N$ " pueden usarse indistintamente para hacer referencia a las dimensiones de píxeles de un bloque de vídeo en términos de dimensiones verticales y horizontales, por ejemplo,  $16 \times 16$  píxeles o 16 por 16 píxeles. En general, un bloque de tamaño  $16 \times 16$  tendrá 16 píxeles en la dirección vertical ( $y = 16$ ) y 16 píxeles en la dirección horizontal ( $x = 16$ ). Asimismo, un bloque de tamaño  $N \times N$  presenta, en general,  $N$  píxeles en una dirección vertical y  $N$  píxeles en una dirección horizontal, donde  $N$  representa un valor entero no negativo. Los píxeles en un bloque pueden disponerse en filas y columnas. Además, no es necesario que los bloques presenten 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  $N$ .

**[50]** Para codificar un bloque (por ejemplo, una unidad de predicción de datos de vídeo), se obtiene primero un predictor para el bloque. El predictor, también denominado bloque predictivo, puede obtenerse, ya sea mediante la intrapredicción (I) (es decir, predicción espacial) o la interpredicción (P o B) (es decir, predicción temporal). Por lo tanto, algunas unidades de predicción pueden estar intracodificadas (I) usando predicción espacial con respecto a las muestras de referencia en bloques de referencia contiguos en la misma trama (o fragmento), y otras unidades de predicción pueden estar intercodificadas unidireccionalmente (P) o intercodificadas bidireccionalmente (B) con respecto a bloques de muestras de referencia en otras tramas (o fragmentos) previamente codificadas. En cada caso, las muestras de referencia se pueden utilizar para formar un bloque predictivo para un bloque que ha de codificarse.

**[51]** Tras la identificación de un bloque predictivo, se determina la diferencia entre el bloque de datos de vídeo original y su bloque predictivo. Esta diferencia puede denominarse datos residuales de predicción, e indica las diferencias de píxeles entre los valores de píxel en el bloque que tiene que codificarse y los valores de píxel en el

bloque predictivo seleccionado para representar al bloque codificado. Para lograr una mejor compresión, los datos residuales de predicción pueden transformarse, por ejemplo, usando una transformada de coseno discreta (DCT), una transformada de enteros, una transformada de Karhunen-Loeve (K-L) u otra transformación.

5 **[52]** Los datos residuales en un bloque de transformación, tales como una TU, pueden estar dispuestos en una matriz bidimensional (2D) de valores de diferencia de píxeles en el dominio espacial de píxeles. Una transformación convierte los valores de los píxeles residuales en una matriz bidimensional de coeficientes de transformación en un dominio de la transformación, tal como un dominio de frecuencia. Para una mayor compresión, los coeficientes de transformación pueden cuantificarse antes de la codificación por entropía. Un codificador por entropía, a continuación, aplica codificación por entropía, tal como la codificación de longitud variable adaptativa de contexto (CAVLC), la codificación aritmética binaria adaptativa al contexto (CABAC), la codificación por entropía de división de intervalos de probabilidad (PIPE), o similares, a los coeficientes de transformación cuantificados.

15 **[53]** Para codificar por entropía un bloque de coeficientes de transformación cuantificados, generalmente se realiza un proceso de escaneo de modo que se reordene la matriz bidimensional (2D) de coeficientes de transformación cuantificados en un bloque, de acuerdo con un orden de escaneo particular, en una matriz ordenada, unidimensional (1D), es decir, un vector, de coeficientes de transformación. A continuación se aplica la codificación por entropía al vector de coeficientes de transformación. El escaneo de los coeficientes de transformación cuantificados en una unidad de transformación serializa la matriz bidimensional de coeficientes de transformación para el codificador por entropía. Puede generarse una correlación de significancia para indicar las posiciones de los coeficientes significativos (es decir, distintos a cero). El escaneo se puede aplicar para escanear niveles de coeficientes significativos (es decir, distintos de cero), y/o para codificar signos de los coeficientes significativos.

25 **[54]** En HEVC, la información de posición de la transformación significativa (por ejemplo, la correlación de significancia) se codifica en primer lugar para una TU para indicar la ubicación del último coeficiente distinto de cero en el orden de escaneo. La correlación de significancia y la información de nivel (los valores absolutos y los signos de los coeficientes) se codifican para cada coeficiente en un orden de escaneo inverso.

30 **[55]** Tras cualquier transformada para generar coeficientes de transformada, el codificador de vídeo 20 puede realizar la cuantificación de los coeficientes de transformada. La cuantificación se refiere, en general, a un proceso en el que los coeficientes de transformada se cuantifican para reducir posiblemente la cantidad de datos usados para representar los coeficientes, proporcionando compresión adicional. El proceso de cuantificación puede reducir la profundidad de bits asociada con algunos o la totalidad de los coeficientes. Por ejemplo, un valor de bit  $n$  se puede redondear a un valor de bit  $m$  durante la cuantificación, 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 cuantificados para producir un vector serializado que puede ser codificado por entropía. En otros ejemplos, el codificador de vídeo 20 puede realizar un escaneo adaptativo.

40 **[56]** La FIG. 4 es un diagrama de bloques que ilustra un ejemplo de codificador de vídeo 20 que puede implementar las técnicas descritas en esta divulgación. El codificador de vídeo 20 puede realizar la intracodificación y la intercodificación de bloques de vídeo dentro de fragmentos de vídeo. La intracodificación se apoya en la predicción espacial para reducir o eliminar la redundancia espacial en el vídeo dentro de una trama o imagen de vídeo determinada. La intercodificación se basa en la predicción temporal para reducir o eliminar la redundancia temporal en el vídeo dentro de tramas o imágenes adyacentes de una secuencia de vídeo. El modo intra (modo I) puede referirse a cualquiera de varios modos de compresión espacial. Los modos inter, tales como la predicción unidireccional (modo P) o la bipredicción (modo B), pueden referirse a cualquiera de varios modos de compresión de base temporal.

50 **[57]** En el ejemplo de la FIG. 4, el codificador de vídeo 20 incluye un módulo de división 35, un módulo de predicción 41, una memoria de imágenes de referencia 64, un sumador 50, un módulo de transformación 52, un módulo de cuantificación 54 y un módulo de codificación por entropía 56. El módulo de predicción 41 incluye un módulo de estimación de movimiento 42, un módulo de compensación de movimiento 44 y un módulo de intrapredicción 46. El módulo de predicción 41 también puede incluir el módulo de división 35. Para la reconstrucción de bloques de vídeo, el codificador de vídeo 20 incluye además un módulo de cuantificación inversa 58, un módulo de transformación inversa 60 y un sumador 62. También puede incluirse un filtro de eliminación de bloques (no mostrado en la FIG. 4) para filtrar límites de bloque a fin de eliminar distorsiones de efecto pixelado del vídeo reconstruido. Si se desea, el filtro de eliminación de bloques filtrará típicamente la salida del sumador 62. También pueden usarse filtros de bucle adicionales (en bucle o tras el bucle), además del filtro de eliminación de bloques.

60 **[58]** Como se representa en la FIG. 4, el codificador de vídeo 20 recibe datos de vídeo, y el módulo de división 35 divide los datos en bloques de vídeo. Esta división también puede incluir la división de datos de vídeo en fragmentos, mosaicos u otras unidades más grandes, así como divisiones de bloques de vídeo, por ejemplo, de acuerdo con una estructura de árbol cuaternario residual de LCU y CU para producir PU y TU. El módulo de división 35 puede dividir bloques de datos de vídeo tales como LCU en subbloques tales como CU, PU y TU. Por ejemplo, el módulo de división 35 puede dividir bloques de datos de vídeo basándose en la evaluación de esquemas de división previos en pasadas de codificación previas. Por ejemplo, el módulo de división 35 puede dividir inicialmente una

trama o un fragmento en LCU, y dividir cada una de las LCU en sub-CU basándose en un análisis de velocidad-distorsión (por ejemplo, optimización de velocidad-distorsión) junto con el módulo de predicción 41. Por consiguiente, aunque el módulo de división 35 se muestra en la FIG. 4 como un bloque separado para fines ilustrativos, las funciones de división realizadas por el módulo de división 35 pueden integrarse con las funciones de selección y predicción de modo del módulo de predicción 41, particularmente cuando el análisis de velocidad-distorsión realizado por el módulo de predicción 41 puede usarse para determinar, al menos en parte, aspectos de la división aplicada a una LCU para producir CU, sub-CU, PU y TU. En general, el módulo de predicción 41, junto con el módulo de división 35, puede producir una estructura de datos de árbol cuaternario indicativa de la división de una LCU en sub-CU. Las CU de nodos de hojas del árbol cuaternario pueden incluir una o más PU y una o más TU.

**[59]** El módulo de predicción 41, representado por el módulo de división 35, puede proporcionar diversas técnicas para definir un esquema de división para una de la pluralidad de componentes de vídeo independientemente de otros componentes de vídeo, como se describe en esta divulgación. En otro aspecto, el módulo de predicción 41 puede determinar si se define un esquema de división para al menos uno de la pluralidad de componentes de vídeo de un bloque de vídeo independientemente de los otros componentes de vídeo basándose en una o más propiedades del bloque de vídeo. En algunos ejemplos, las propiedades pueden incluir al menos uno de tamaño de bloque, tipo de imagen, división de bloque o información de movimiento.

**[60]** El codificador de vídeo 20 ilustra, en general, los componentes que codifican bloques de vídeo de un fragmento de vídeo que se va a codificar. El fragmento puede dividirse en varios bloques de vídeo (y, posiblemente, en conjuntos de bloques de vídeo denominados mosaicos). El módulo de predicción 41 puede seleccionar uno entre una pluralidad de posibles modos de codificación, tal como uno entre una pluralidad de modos de intracodificación, o uno entre una pluralidad de modos de intercodificación, para el bloque de vídeo actual, basándose en resultados de errores (por ejemplo, la velocidad de codificación y el nivel de distorsión). El módulo de predicción 41 puede proporcionar el bloque intracodificado o intercodificado 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.

**[61]** El módulo de intrapredicción 46, dentro del módulo de predicción 41, puede realizar la codificación intrapredictiva del bloque de vídeo actual con respecto a uno o más bloques contiguos en la misma trama o fragmento que el bloque que va a codificarse, para proporcionar compresión espacial. El módulo de estimación de movimiento 42 y el módulo de compensación de movimiento 44, dentro del módulo de predicción 41, realizan la codificación interpredictiva del bloque de vídeo actual con respecto a uno o más bloques predictivos en una o más imágenes de referencia, para proporcionar compresión temporal.

**[62]** Si se selecciona la interpredicción en lugar de la intrapredicción para un bloque de vídeo actual, el módulo 42 de estimación de movimiento puede configurarse para determinar el modo de interpredicción para un fragmento de vídeo de acuerdo con 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 P/B generalizados (GPB). El módulo de estimación de movimiento 42 y el módulo de compensación de movimiento 44 pueden estar altamente integrados, pero se ilustran por separado con fines conceptuales. La estimación de movimiento, realizada mediante el módulo de estimación de movimiento 42, es el proceso de generar 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 de una trama o imagen de vídeo actual con respecto a un bloque predictivo de una imagen de referencia.

**[63]** En intercodificación, un bloque predictivo es un bloque del que se descubre que se corresponde estrechamente con la PU del bloque de vídeo que se va a codificar en términos de diferencia de píxeles, que puede determinarse mediante la suma de la diferencia absoluta (SAD), suma de la diferencia al cuadrado (SSD) u otras métricas de diferencia. En algunos ejemplos, el codificador de vídeo 20 puede calcular valores para posiciones fraccionarias de píxeles de imágenes de referencia almacenadas en la memoria de imágenes de referencia 64. Por ejemplo, el codificador de vídeo 20 puede interpolar valores de posiciones de un cuarto de píxel, posiciones de un octavo de píxel u otras posiciones de fracciones de píxel de la imagen de referencia. Por lo tanto, el módulo 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.

**[64]** El módulo de estimación de movimiento 42 calcula un vector de movimiento para una PU de un bloque de vídeo en un fragmento intercodificado, comparando la posición de la PU con la posición de un bloque predictivo de una imagen de referencia. La imagen de referencia puede seleccionarse de una primera lista de imágenes de referencia (Lista 0) o de una segunda lista de imágenes de referencia (Lista 1), cada una de las cuales identifica una o más imágenes de referencia almacenadas en la memoria de imágenes de referencia 64. El módulo de estimación de movimiento 42 envía el vector de movimiento calculado al módulo de codificación por entropía 56 y al módulo de compensación de movimiento 44, por ejemplo, junto con otros elementos sintácticos.

**[65]** La compensación de movimiento, realizada por el módulo de compensación de movimiento 44, puede implicar extraer o generar el bloque predictivo basándose en el vector de movimiento determinado por la estimación

de movimiento, realizando posiblemente interpolaciones hasta la precisión de subpíxel. Tras recibir el vector de movimiento para la PU del bloque de vídeo actual, el módulo 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 se está codificando, generando valores de diferencia de píxel. Los valores de diferencia de píxel forman datos residuales para el bloque, y pueden incluir componentes de diferencia de luminancia y crominancia. El sumador 50 representa el componente o los componentes que realizan esta operación de resta. El módulo de compensación de movimiento 44 también puede generar elementos sintácticos asociados a los bloques de vídeo y al fragmento de vídeo para su uso mediante el descodificador de vídeo 30 en la descodificación de los bloques de vídeo del fragmento de vídeo.

**[66]** El módulo de intrapredicción 46 puede intrapredecir un bloque actual, como alternativa a la interpredicción llevada a cabo por el módulo de estimación de movimiento 42 y el módulo de compensación de movimiento 44, como se ha descrito anteriormente. En particular, el módulo de intrapredicción 46 puede determinar un modo de intrapredicción a usar para codificar un bloque actual. En algunos ejemplos, el módulo de intrapredicción 46 puede codificar un bloque actual para generar un bloque predictivo usando varios modos de intrapredicción, por ejemplo, durante pasadas de codificación independientes, y el módulo de intrapredicción 46 (o el módulo de selección de modo 40, en algunos ejemplos) puede seleccionar un modo adecuado de intrapredicción a usar, entre los modos probados. Por ejemplo, el módulo de intrapredicción 46 puede calcular valores de velocidad-distorsión usando un análisis de velocidad-distorsión para los diversos modos de intrapredicción probados, y seleccionar el modo de intrapredicción que tenga las mejores características de velocidad-distorsión entre los modos probados. El análisis de velocidad-distorsión determina, en general, una magnitud de distorsión (o error) entre un bloque codificado y un bloque original no codificado que se codificó para generar el bloque codificado, así como una velocidad de transferencia de bits (es decir, un número de bits) utilizada para generar el bloque codificado. El módulo de intrapredicción 46 puede calcular proporciones a partir de las distorsiones y velocidades para los diversos bloques codificados, para determinar qué modo de intrapredicción presenta el mejor valor de velocidad-distorsión para el bloque.

**[67]** En cualquier caso, tras seleccionar un modo de intrapredicción para un bloque, el módulo de intrapredicción 46 puede proporcionar información que indica el modo de intrapredicción seleccionado para el bloque al módulo de codificación por entropía 56. Por ejemplo, el módulo de intrapredicción 46 puede proporcionar señalización, tal como uno o más elementos sintácticos, para indicar un modo intra seleccionado. El módulo de codificación por entropía 56 puede codificar la información que indica el modo de intrapredicción seleccionado. 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 modos de intrapredicción y una pluralidad de tablas de índices de modos de intrapredicción modificadas (también denominadas tablas de correlación de palabras de código), definiciones de contextos de codificación para varios bloques e indicaciones del modo de intrapredicción más probable, una tabla de índices de modos de intrapredicción y una tabla de índices de modos de intrapredicción modificadas a utilizar para cada uno de los contextos.

**[68]** Después de que el módulo de predicción 41 genere el bloque predictivo para el bloque de vídeo actual, el codificador de vídeo 20 forma un bloque de vídeo residual al restar el bloque predictivo del bloque de vídeo actual. Los datos de vídeo residual del bloque residual pueden ser incluidos en una o más TU y ser aplicados al módulo de transformada 52. El módulo de transformada 52 transforma los datos de vídeo residual en coeficientes de transformada residual, usando una transformada, tal como una transformada de coseno discreta (DCT) o una transformada conceptualmente similar. El módulo de transformada 52 puede convertir los datos de vídeo residual, desde un dominio de píxeles a un dominio de transformada, tal como un dominio de frecuencia.

**[69]** El módulo de transformada 52 puede enviar los coeficientes de transformación resultantes al módulo de cuantificación 54. El módulo de cuantificación 54 cuantifica los coeficientes de transformación para reducir adicionalmente la velocidad de transmisión de bits. El proceso de cuantificación puede reducir la profundidad de bits asociada con algunos o la totalidad de los coeficientes. El grado de cuantificación puede modificarse ajustando un parámetro de cuantificación. En algunos ejemplos, el módulo de cuantificación 54 puede realizar, a continuación, un escaneo de la matriz que incluye los coeficientes de transformación cuantificados. De forma alternativa, el módulo de codificación por entropía 56 puede realizar el escaneo.

**[70]** Tras la cuantificación, el módulo de codificación por entropía 56 codifica por entropía los coeficientes de transformación cuantificados. Por ejemplo, el módulo 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 basándose en la sintaxis (SBAC), una codificación por entropía por división de intervalos de probabilidad (PIPE) u otra metodología o técnica de codificación por entropía. Tras la codificación por entropía realizada por el módulo de codificación por entropía 56, el flujo de bits codificado puede transmitirse al descodificador de vídeo 30, o archivarse para su posterior transmisión o recuperación mediante el descodificador de vídeo 30. El módulo de codificación por entropía 56 también puede realizar la codificación por entropía de los vectores de movimiento y de los otros elementos sintácticos para el fragmento de vídeo actual que se está codificando.

**[71]** El módulo de cuantificación inversa 58 y el módulo de transformación inversa 60 aplican una cuantificación inversa y una transformación inversa, respectivamente, para reconstruir el bloque residual en el dominio de los píxeles, para su uso posterior como un bloque de referencia de una imagen de referencia. El módulo de compensación de movimiento 44 puede calcular un bloque de referencia sumando el bloque residual a un bloque predictivo de una de las imágenes de referencia de una de las listas de imágenes de referencia. El módulo de compensación de movimiento 44 también puede aplicar uno o más filtros de interpolación al bloque residual reconstruido para calcular valores de píxeles fraccionarios para su uso en la estimación de movimiento. El sumador 62 añade el bloque residual reconstruido al bloque predictivo con compensación de movimiento generado por el módulo de compensación de movimiento 44 para generar un bloque de referencia para su almacenamiento en la memoria de imágenes de referencia 64. El módulo de estimación de movimiento 42 y el módulo de compensación de movimiento 44 pueden usar el bloque de referencia como un bloque de referencia para realizar la interpredicción de un bloque en una imagen o trama de vídeo posterior.

**[72]** El codificador de vídeo 20 de la FIG. 4 representa un ejemplo de un codificador de vídeo configurado para emitir contenido de vídeo codificado HEVC, que puede incluir muestras codificadas de contenido de vídeo, conjuntos de parámetros y mensajes SEI. Como se describió anteriormente, con respecto al archivo 142 multimedia HEVC ilustrado en la FIG. 1, los tipos de conjuntos de parámetros utilizados con la norma de codificación de vídeo HEVC incluyen SPS, PPS y APS. Cada conjunto de parámetros tiene un identificador, y cada fragmento codificado de acuerdo con la norma HEVC, como se describió anteriormente, puede hacer referencia al conjunto de parámetros con el que se codificó utilizando el identificador del conjunto de parámetros.

**[73]** Con referencia de nuevo a la FIG. 3, el módulo de encapsulación 21 puede recibir el contenido de vídeo codificado del codificador de vídeo 20 y generar un archivo de vídeo de acuerdo con las técnicas descritas en el presente documento. En un ejemplo, el módulo de encapsulación 21 puede recibir contenido de vídeo codificado HEVC y generar un archivo de vídeo usando un formato de archivo basado en el ISOBMFF y la versión actual del HEVCFF.

**[74]** La FIG. 5 es un diagrama de bloques que ilustra componentes de un ejemplo de módulo de encapsulación 21. En el ejemplo de la FIG. 5, el módulo de encapsulación 21 incluye la interfaz 70 de entrada de vídeo, el módulo 72 de creación de archivos de vídeo y la interfaz 76 de salida de archivo de vídeo. El módulo 70 de creación de archivos de vídeo, en este ejemplo, incluye el constructor 74 de la unidad de la capa de abstracción de red (NAL).

**[75]** La interfaz 70 de entrada de vídeo puede recibir contenido de vídeo codificado. La interfaz 70 de entrada de vídeo puede recibir contenido de vídeo codificado desde el codificador 20 de vídeo o puede recuperar contenido de vídeo codificado desde un dispositivo de almacenamiento. El contenido de vídeo codificado se puede codificar de acuerdo con la norma HEVC y puede incluir muestras de contenido de vídeo, conjuntos de parámetros y mensajes SEI. Al recibir el contenido de vídeo codificado, la interfaz 70 de entrada de vídeo puede enviar el contenido de vídeo codificado al módulo 72 de creación de archivos de vídeo para ensamblarlo en un archivo de vídeo. En algunos casos, la interfaz 70 de entrada de vídeo puede facilitar el ensamblaje de un archivo de vídeo organizando o almacenando en memoria intermedia contenido de vídeo codificado antes de que se emita al módulo 72 de creación de archivos de vídeo.

**[76]** En general, el módulo 72 de creación de archivos de vídeo puede generar un archivo de vídeo que incluye el contenido de vídeo codificado recibido. En un ejemplo, el módulo 72 de creación de archivos de vídeo puede crear un archivo de vídeo, tal como, el archivo 142 multimedia HEVC descrito con respecto a la FIG. 1. El módulo 72 de creación de archivos de vídeo puede corresponder a una unidad de control que incluye hardware, software y/o firmware configurada para realizar las funciones y procedimientos atribuidos a los mismos. La unidad de control puede realizar adicionalmente las funciones atribuidas al módulo de encapsulación 21 en general. Para ejemplos en los que el módulo 72 de creación de archivos de vídeo está incorporado en software y/o firmware, el módulo de encapsulación 21 puede incluir un medio legible por ordenador que comprende instrucciones para el módulo 72 de creación de archivos de vídeo y una unidad de procesamiento para ejecutar las instrucciones. Los submódulos del módulo 72 de creación de archivos de vídeo (constructor 74 de unidades NAL en este ejemplo) pueden implementarse como módulos de hardware individuales y/o módulos de software, y pueden integrarse funcionalmente o separarse adicionalmente en submódulos adicionales. El módulo 72 de creación de archivos de vídeo puede corresponder a cualquier unidad de procesamiento o circuitos de procesamiento adecuados, como, por ejemplo, uno o más microprocesadores, circuitos integrados específicos de la aplicación (ASIC), matriz de puertas programables in situ (FPGA), procesadores de señales digitales (DSP), o cualquier combinación de los mismos. El módulo 72 de creación de archivos de vídeo puede incluir, además, instrucciones de almacenamiento de medios legibles por ordenador no transitorias para el constructor 74 de unidades NAL, así como también un procesador para ejecutar las instrucciones.

**[77]** En un ejemplo, el módulo 72 de creación de archivos de vídeo puede configurarse para generar datos de vídeo codificados de acuerdo con un formato de archivo donde los conjuntos de parámetros pueden incluirse en la descripción de la muestra o en las muestras o unidades de acceso del flujo al que se aplica la descripción de la muestra. En un ejemplo, si se incluye un tipo de conjunto de parámetros en la descripción de la muestra, es posible que no se incluya también dentro de las propias muestras. Es decir, en este ejemplo particular, para un flujo de vídeo

al que se aplica una descripción de la muestra particular, un tipo particular de conjunto de parámetros, si está presente, puede almacenarse en la descripción de la muestra o en las muestras, pero nunca en ambas. En un ejemplo alternativo, se puede incluir un tipo de conjunto de parámetros tanto en la descripción de la muestra como en las muestras del flujo al que se aplica la descripción de la muestra. Cuando se permite la inclusión de conjuntos de parámetros en la descripción de la muestra o en las propias muestras, se puede incluir una indicación en la descripción de la muestra para indicar si existen conjuntos de parámetros incluidos en las muestras además de los incluidos en la descripción de la muestra. En otro ejemplo, el módulo 72 de creación de archivos de vídeo puede configurarse para generar datos de vídeo codificados de acuerdo con un formato de archivo en el que solo uno o más de un tipo particular de conjunto de parámetros puede incluirse tanto en la descripción de la muestra como en las muestras del flujo al que se aplica la descripción de la muestra. Por ejemplo, en algunos casos, un formato de archivo puede permitir que se incluya PPS tanto en la descripción de la muestra como en las muestras, mientras que SPS solo puede incluirse en la descripción de la muestra. Además, en este caso, solo se puede permitir que APS se almacene en muestras. Además, se puede incluir una indicación en la descripción de la muestra para indicar si puede haber conjuntos de parámetros del tipo particular incluido en las muestras además de los incluidos en la descripción de la muestra.

**[78]** Además, el módulo 72 de creación de archivos de vídeo puede configurarse de manera que cada descripción de la muestra HEVC, que contiene la información específica del descodificador de flujo de vídeo HEVC, pueda incluir un grupo del tipo particular de conjuntos de parámetros. Un grupo de conjuntos de parámetros puede funcionar de manera similar a un libro de códigos. Cada conjunto de parámetros puede tener un identificador, y cada fragmento codificado puede hacer referencia a un identificador de conjunto de parámetros. Según lo definido por un formato de archivo, cada configuración de conjuntos de parámetros se puede representar por separado. De acuerdo con algunos formatos de archivo, un conjunto de parámetros no se puede actualizar sin causar que se use una descripción de la muestra diferente. En este caso, los sistemas que deseen enviar actualizaciones del conjunto de parámetros pueden necesitar comparar las dos configuraciones para encontrar las diferencias a fin de enviar las actualizaciones apropiadas del conjunto de parámetros. Cabe señalar, que HEVCFF típicamente recomienda que cuando se utilizan varios conjuntos de parámetros y se desea actualizar el conjunto de parámetros, los conjuntos de parámetros se incluyan en las muestras del flujo. También se debe tener en cuenta que los descodificadores que cumplen con HEVCFF pueden ser necesarios para admitir conjuntos de parámetros almacenados en las muestras así como conjuntos de parámetros almacenados en las entradas de descripción de la muestra, a menos que esté restringido por otro formato de archivo basado en HEVCFF.

**[79]** En un ejemplo, el módulo 72 de creación de archivos de vídeo puede configurarse para generar datos de vídeo codificados de acuerdo con un formato de archivo donde los diferentes tipos de conjuntos de parámetros no se agrupan cuando se incluyen en la descripción de la muestra o en las muestras del flujo al que se aplica la descripción de la muestra. En este ejemplo, la inclusión de un tipo de conjunto de parámetros en la descripción de la muestra o en las muestras puede ser independiente de si se incluyen otros tipos de conjuntos de parámetros en la descripción de la muestra o en las muestras. En otro ejemplo, el módulo de creación de archivos de vídeo puede configurarse para generar datos de vídeo codificados de acuerdo con un formato de archivo donde solo un subconjunto de todos los tipos de conjuntos de parámetros se agrupan cuando se trata de si están incluidos en la descripción de muestra o en las muestras del flujo al que se aplica la descripción de la muestra. Por ejemplo, SPS y PPS pueden agruparse, mientras que la inclusión de APS en la descripción de la muestra o en las muestras puede ser independiente de la inclusión de SPS y PPS en la descripción de la muestra o en las muestras.

**[80]** En un ejemplo, el módulo 72 de creación de archivos de vídeo puede configurarse para generar datos de vídeo codificados de acuerdo con un formato de archivo donde una descripción de la muestra incluye una matriz para cada tipo de conjunto de parámetros, y la matriz para un tipo particular de conjunto de parámetros solo puede incluir unidades NAL de conjunto de parámetros de ese tipo en particular. Además, la descripción de la muestra también puede incluir una matriz para unidades NAL de SEI que contienen mensajes SEI declarativos. Por lo tanto, en este caso, un analizador de archivos no necesita verificar la cabecera de la unidad NAL de cada unidad NAL incluida en una matriz para determinar el tipo de la unidad NAL. En un ejemplo alternativo, el módulo 72 de creación de archivos de vídeo puede configurarse para generar datos de vídeo codificados de acuerdo con un formato de archivo donde se incluye una matriz en la descripción de la muestra, y la matriz puede incluir cualquier tipo de conjuntos de parámetros así como unidades NAL de SEI que contienen mensajes SEI declarativos.

**[81]** En un ejemplo, el módulo 72 de creación de archivos de vídeo puede configurarse para emitir datos de vídeo codificados según un formato de archivo donde las propiedades del flujo como resolución espacial y velocidad de trama que son importantes para la selección de pista o flujo (por ejemplo, en aplicaciones de transmisión adaptativa basadas en archivos) están incluidas en la descripción de la muestra. Además, en otro ejemplo, el módulo 72 de creación de archivos de vídeo puede configurarse para emitir datos de vídeo codificados de acuerdo con un formato de archivo donde la descripción de la muestra incluye la profundidad de bits de salida (es decir, la profundidad de bits, posiblemente la misma o diferente para diferentes componentes de color, de las imágenes enviadas desde el codificador de vídeo 20). La inclusión de las propiedades de flujo que son importantes para la selección de la pista o flujo en la descripción de la muestra puede permitir el acceso conveniente a estas propiedades de flujo. En otro ejemplo, el módulo 72 de creación de archivos de vídeo puede configurarse para generar datos de vídeo codificados de acuerdo con un formato de archivo donde la descripción de la muestra incluye

información sobre un conjunto de mosaicos (como se define en HEVC) que se puede descodificar independientemente de otros mosaicos en todo el flujo, por ejemplo, la región de geometría cubierta por el conjunto de mosaicos.

5 **[82]** En un ejemplo, el módulo 72 de creación de archivos de vídeo puede configurarse para generar datos de vídeo codificados de acuerdo con un formato de archivo donde las propiedades de escalabilidad temporal tales como perfil, nivel y velocidad de trama, que son importantes para la selección de un subconjunto temporal particular de un flujo, se incluyen para cada representación de capa temporal, a través del mecanismo de agrupación de muestra de escalabilidad temporal.

10 **[83]** En un ejemplo, el módulo 72 de creación de archivos de vídeo puede configurarse para almacenar un flujo de vídeo completo en una pista, independientemente de cuántas capas temporales estén incluidas en la transmisión de vídeo. En un ejemplo alternativo, un flujo de vídeo completo puede almacenarse en más de una pista, donde cada pista contiene un subconjunto de capas temporales consecutivas. En el caso en que se almacena un flujo de vídeo completo en más de una pista, puede haber una indicación en el archivo para indicar cada pista para la cual la capa temporal contenida más baja tiene un `temporal_id`, denotado como `tId`, mayor que 0, e indicar la pista que contiene la capa con `temporal_id` igual a `tId - 1`. Estas pistas se pueden designar respectivamente como pista A y pista B. Dicha indicación puede ser una referencia de pista de tipo 'tscl' incluida en la pista A, donde la referencia de pista se refiere a la pista B de la que depende la pista A. Las propiedades de flujo de la pista A pueden describir que el flujo consiste en la pista A y todas las pistas que colectivamente contienen las capas temporales con valores de `temporal_id` menores que `tId`, como si todas estas capas temporales también estuvieran en la pista A. En un ejemplo, el módulo 72 de creación de archivos de vídeo puede estar configurado para realizar cualquiera y todas las combinaciones de los ejemplos descritos anteriormente. Por ejemplo, el módulo 72 de creación de archivos de vídeo puede configurarse como salida de datos de vídeo codificados de acuerdo con un formato de archivo que incluye cualquiera y todas las combinaciones de los formatos de archivo descritos en el presente documento.

15 **[84]** El módulo 72 de creación de archivos de vídeo puede incluir información de configuración del descodificador en un registro de configuración de información del descodificador almacenado dentro de un archivo. La información del descodificador puede definirse mediante la norma HEVC y el registro de configuración del descodificador puede basarse en el registro de configuración del descodificador actualmente definido en el HEVCFF, pero también puede incluir información adicional coherente con los ejemplos descritos en el presente documento. De esta manera, una o más de las configuraciones de ejemplo del módulo 72 de creación de archivos de vídeo descrito anteriormente pueden implementarse de acuerdo con un formato de archivo que define la información de configuración del descodificador. La información de configuración del descodificador puede incluirse en un registro de configuración del descodificador como se describe adicionalmente a continuación. Por lo tanto, en un ejemplo, una descripción de la muestra como se describió anteriormente puede incorporarse en un registro de configuración del descodificador. Debe observarse que aunque metadatos, tales como conjuntos de parámetros e indicaciones respectivas de los mismos, se describen como incluidos en un registro de configuración de descodificador o en una descripción de la muestra, esto no debe interpretarse como limitación y los metadatos descritos anteriormente con respectivas configuraciones de ejemplo del módulo 72 de creación de archivos de vídeo pueden almacenarse en otras partes de un archivo generado por el módulo 72 de creación de archivos de vídeo, donde las otras partes de un archivo generado por el módulo 72 de creación de archivos de vídeo son distintas de los flujos de vídeo.

20 **[85]** En un ejemplo, el módulo 72 de creación de archivos de vídeo puede generar un archivo que incluye un registro de configuración del descodificador que puede contener el tamaño del campo de longitud utilizado en cada unidad de acceso para indicar la longitud de las unidades NAL contenidas en una unidad de acceso. Además, en un ejemplo, el registro de configuración del descodificador generado por el módulo 72 de creación de archivos de vídeo también puede incluir información con respecto a conjuntos de parámetros y unidades NAL de SEI declarativas, tales como, por ejemplo: (1) una indicación de si los conjuntos de parámetros están almacenados en la descripción de la muestra o dentro de las unidades de acceso (2) el número de conjuntos de parámetros almacenados en una descripción de la muestra o dentro de las unidades de acceso de un flujo; (3) información con respecto a las unidades NAL de conjunto de parámetros; (4) información con respecto a las unidades NAL de SEI declarativas; y/o (5) información de resolución espacial.

25 **[86]** Un registro de configuración de descodificador de ejemplo puede ser generado por el módulo 72 de creación de archivos de vídeo de acuerdo con la siguiente sintaxis de ejemplo. Debe observarse que la sintaxis siguiente puede modificarse de acuerdo con las otras configuraciones del módulo 72 de creación de archivos de vídeo de ejemplo descritas anteriormente.

30 *aligned(8) class HEVCDecoderConfigurationRecord {*  
*unsigned int(8) configurationVersion = 1;*  
*unsigned int(8) ProfileIndication;*  
*unsigned int(8) profileCompatibility;*



```

        unsigned int(8) LevelIndication;
        bit(6) reserved = '111111' b;
        unsigned int(2) chromaFormat;

5      bit(5) reserved = '11111' b;
        unsigned int(3) bitDepthLumaMinus8;
        bit(5) reserved = '11111' b;
        unsigned int(3) bitDepthChromaMinus8;
        bit(16) picture WidthInLumaSamples;
10     bit(16) pictureHeightInLumaSamples;
        bit(16) avgFrameRate;
        bit(1) constantFrameRate;
        bit(3) numTemporalLayers;
        bit(2) reserved = '11' b;
15     unsigned int(2) lengthSizeMinusOne;
        unsigned int(8) numOfSequenceParameterSets;
        for (i=0; i< numOfSequenceParameterSets; i++) {
            unsigned int(16) sequenceParameterSetLength ;
            bit(8*sequenceParameterSetLength) sequenceParameterSetNalUnit;
20     }
        unsigned int(8) numOfPictureParameterSets;
        for (i=0; i< numOfPictureParameterSets; i++) {
            unsigned int(16) pictureParameterSetLength;
            bit(8*pictureParameterSetLength) pictureParameterSetNalUnit;
25     }
        unsigned int(8) numOfAdaptationParameterSets;
        for (i=0; i< numOfAdaptationParameterSets; i++) {
            unsigned int(16) adaptationParameterSetLength;
            bit (8*adaptationParameterSetLength) adaptationParameterSetNALUnit;
30     }
        unsigned int(8) numOfDeclarativeSeiNalUnits;
        for (i=0; i< numOfDeclarativeSeiNalUnits; i++) {
            unsigned int(16) declarativeSeiNalUnitLength;
            bit(8*declarativeSeiNalUnitLength) declarativeSeiNalUnit;
35     }
    }

```

**[87]** En la sintaxis anterior de ejemplo, la semántica se puede definir de la siguiente manera:

40 **ProfileIndication** puede contener el código de perfil (profile\_idc) como se define en ISO/IEC 23008-HEVC, para el flujo al que se aplica este registro de configuración.

- profileCompatibility** puede ser un byte definido exactamente igual que el byte que se produce entre `profile_idc` y `level_idc` en un conjunto de parámetros de secuencia (SPS), como se define en ISO/IEC 23008-HEVC, para el flujo al que se aplica este registro de configuración.
- 5 **LevelIndication** puede contener el código de nivel (`level_idc`) como se define en ISO/IEC 23008-10.
- chromaFormat** puede contener el indicador `chroma_format` tal como se define en el parámetro `chroma_format_idc` en ISO/IEC 23008-HEVC, para el flujo al que se aplica este registro de configuración.
- 10 **bitDepthLumaMinus8** puede contener el indicador de profundidad de bits de luma definido por el parámetro `bit_depth_luma_minus8` en ISO/IEC 23008-HEVC, para el flujo al que se aplica este registro de configuración.
- bitDepthChromaMinus8** puede contener el indicador de profundidad de bits de croma definido por `bit_depth_chroma_minus8` en ISO/IEC 23008-HEVC, para el flujo al que se aplica este registro de configuración.
- 15 **pictureWidthInLumaPixels** puede indicar el ancho de la imagen descodificada en unidades de píxeles luma, para el flujo al que se aplica este registro de configuración.
- pictureHeightInLumaPixels** puede indicar la altura de la imagen descodificada en unidades de píxeles luma, para el flujo al que se aplica este registro de configuración.
- 20 **avgFrameRate** puede dar la velocidad de trama promedio en unidades de tramas/(256 segundos), para el flujo al que se aplica este registro de configuración. El valor 0 puede indicar una velocidad de trama promedio no especificada.
- 25 **constantFrameRate** igual a 1 puede indicar que el flujo al que se aplica este registro de configuración es de velocidad de trama constante. El valor 2 puede indicar que la representación de cada capa temporal en el flujo es de velocidad de trama constante. El valor 0 puede indicar que el flujo puede ser o no de velocidad de trama constante.
- 30 **numTemporalLayers** mayor que 1 puede indicar que el flujo al que se aplica este registro de configuración es temporalmente escalable y el número de capas contenidas es igual a `numTemporalLayers`. El valor 1 puede indicar que el flujo no es escalable temporalmente. El valor 0 puede indicar que se desconoce si el flujo es temporalmente escalable.
- 35 **lengthSizeMinusOne** más 1 puede indicar la longitud en bytes del campo `NALUnitLength` en una muestra de vídeo HEVC en el flujo al que se aplica este registro de configuración. Por ejemplo, un tamaño de un byte se puede indicar con un valor de 0. El valor de este campo puede ser uno de 0, 1 o 3, correspondiente a una longitud codificada con 1, 2 o 4 bytes, respectivamente.
- 40 **numOfSequenceParameterSets** puede indicar el número de SPS que se utilizan para descodificar el flujo al que se aplica este registro de configuración. En un ejemplo, si el valor es mayor que 0, no habrá SPS incluido en las muestras del flujo. En un ejemplo, si el valor es igual a 0, debe haber al menos un SPS incluido en las muestras del flujo.
- 45 **sequenceParameterSetLength** puede indicar la longitud en bytes de las unidades NAL de SPS, como se define en ISO/IEC 23008-HEVC.
- 50 **sequenceParameterSetNalUnit** puede contener una unidad NAL de SPS, como se especifica en ISO/IEC 23008-HEVC.
- numOfPictureParameterSets** puede indicar el número de conjuntos de parámetros de imagen que se utilizan para descodificar el flujo al que se aplica este registro de configuración. En un ejemplo, si el valor es mayor que 0, no habrá ningún PPS incluido en las muestras del flujo. En un ejemplo, si el valor es igual a 0, debe haber al menos un PPS incluido en las muestras del flujo.
- 55 **pictureParameterSetLength** puede indicar la longitud en bytes de la unidad NAL de PPS como se define en ISO/IEC 23008-HEVC.
- 60 **pictureParameterSetNalUnit** puede contener una unidad NAL de PPS, como se especifica en ISO/IEC 23008-HEVC.
- numOfAdaptationParameterSets** puede indicar el número de conjuntos de parámetros de adaptación (APS) que se utilizan para descodificar el flujo al que se aplica este registro de configuración. En un ejemplo, si el valor es mayor que 0, no se incluirá ningún APS en las muestras del flujo. En un ejemplo, si el valor es igual a 0, puede haber o no APS incluidos en las muestras del flujo.
- 65

**adaptationParameterSetLength** puede indicar la longitud en bytes de la unidad NAL de APS como se define en ISO/IEC 23008-HEVC.

5 **adaptationParameterSetNALUnit** puede contener una unidad NAL de APS, como se especifica en ISO/IEC 23008-HEVC.

10 **numOfDeclarativeSeiNALUnits** puede indicar el número de unidades NAL de SEI declarativas para el flujo al que se aplica este registro de configuración. Una unidad NAL de SEI declarativa puede contener mensajes SEI de naturaleza 'declarativa', es decir, aquellos que proporcionan información sobre el flujo como un todo. Un ejemplo de dicho SEI es un SEI de datos de usuario.

**declarativeSeiNALUnitLength** puede indicar la longitud en bytes de la unidad NAL de SEI declarativa.

15 **declarativeSeiNALUnit** puede contener una unidad NAL de SEI declarativa.

**numOfSequenceParameterSetExt** puede indicar el número de Ampliaciones de conjunto de parámetros de secuencia que se utilizan para descodificar el flujo elemental de AVC.

20 **sequenceParameterSetExtLength** puede indicar la longitud en bytes de la unidad NAL de ampliación de SPS como se define en ISO/IEC 14496-10.

25 **sequenceParameterSetExtNALUnit** puede contener una unidad NAL de ampliación de SPS, como se especifica en ISO/IEC 14496-10.

30 **[88]** Además del ejemplo de sintaxis y semántica descrito anteriormente, un formato de archivo utilizado por el módulo 72 de creación de vídeo puede incluir restricciones adicionales con respecto a un registro de configuración del descodificador. Por ejemplo, en algunos casos, los valores para ProfileIndication, LevelIndication y los indicadores que indican la compatibilidad del perfil deben ser válidos para todos los conjuntos de parámetros del flujo descrito por un registro de configuración del descodificador. Además, en algunos casos, la indicación de nivel debe indicar un nivel de capacidad igual o mayor que el nivel más alto indicado en los conjuntos de parámetros incluidos y cada indicador de compatibilidad de perfil solo se puede establecer si todos los conjuntos de parámetros incluidos establecen ese indicador. Además, en algunos casos, la indicación de perfil debe indicar un perfil al que se ajusta el flujo asociado con este registro de configuración. Además, en algunos casos, si los conjuntos de parámetros de secuencia están marcados con diferentes perfiles y los indicadores de compatibilidad de perfil pertinente son todos cero, entonces el flujo puede necesitar un examen para determinar a qué perfil, si corresponde, corresponde todo el flujo. Además, si no se examina todo el flujo, o si el examen revela que no existe un perfil al que se ajuste todo el flujo, entonces el flujo completo debe dividirse en dos o más subflujos con registros de configuración separados en los se pueden cumplir estas reglas.

40 **[89]** Además, en algunos ejemplos, se puede proporcionar una indicación explícita en el registro de configuración del descodificador sobre el formato cromático y la profundidad de bits, así como otra información de formato importante utilizada por el flujo elemental de vídeo HEVC. Se puede requerir que cada tipo de dicha información sea idéntico en todos los conjuntos de parámetros, si está presente, en un único registro de configuración de HEVC. Además, en algunos casos, si dos secuencias difieren en cualquier tipo de dicha información, pueden ser necesarios dos registros de configuración de HEVC diferentes. Además, si las dos secuencias difieren en las indicaciones de espacio de color en su información de VUI, entonces también pueden requerirse dos registros de configuración diferentes.

50 **[90]** Además, el registro de configuración del descodificador generado por el módulo 72 de creación de archivos de vídeo puede disponerse en tramas externamente. Cuando el registro de configuración del descodificador se dispone en tramas externamente, su tamaño puede ser suministrado por la estructura que lo contiene. El registro de configuración del descodificador también puede contener un campo de versión. En algunos casos, las ampliaciones compatibles de este registro de configuración del descodificador pueden extenderlo y no cambiarán el código de la versión de configuración. En algunos casos, los lectores de archivos deben estar preparados para ignorar datos no reconocidos más allá de la definición de los datos que entienden.

60 **[91]** Además de generar un registrador de configuración de descodificador, el módulo 72 de creación de archivos de vídeo puede configurarse adicionalmente para ensamblar unidades de acceso que incluyen todas las unidades NAL para una instancia de tiempo particular. Con referencia de nuevo a la FIG. 5, el constructor 74 de la unidad NAL puede formar unidades NAL que incluyen datos de vídeo codificados. Como se describe anteriormente con respecto al archivo 142 multimedia HEVC descrito en la FIG. 1, los tipos de unidades NAL pueden incluir: unidad NAL de delimitador AU, unidad NAL de SEI y unidades NAL de fragmento. Como se describe adicionalmente con respecto a la sintaxis y la semántica del registro del descodificador de ejemplo, los tipos adicionales de unidades NAL pueden incluir: unidades NAL de SPS, unidades NAL de PPS, unidades NAL de APS y unidades NAL de SEI declarativas.

[92] En un ejemplo, el constructor 74 de unidades NAL puede configurarse para definir unidades NAL basándose en un registro de configuración del descodificador de acuerdo con la siguiente sintaxis:

```

5      aligned(8) class HEVCSample
        {
            unsigned int PictureLength = sample_size; //Tamaño de HEVCSample de SampleSizeBox
            for (i=0; i<PictureLength;) // al final de la imagen
                {
10             unsigned int((HEVCDecoderConfigurationRecord.LengthSizeMinusOne+1)*8)
                NALUnitLength;
                bit(NALUnitLength * 8) NALUnit;
                i += (HEVCDecoderConfigurationRecord.LengthSizeMinusOne+1) + NALUnitLength;
                }
15     }

```

[93] En la sintaxis de ejemplo, **NALUnitLength** puede indicar el tamaño de una unidad NAL en bytes. El campo de longitud puede incluir el tamaño de la cabecera NAL y la carga útil de RBSP, pero no necesariamente incluye el campo de longitud en sí. Además, **NALUnit** puede contener una sola unidad NAL. La sintaxis de una unidad NAL puede definirse en ISO/IEC 23008-HEVC y puede incluir tanto la cabecera NAL como la carga útil de flujo de bytes encapsulados de longitud variable.

[94] Además, en algunos ejemplos, el módulo 72 de creación de archivos de vídeo puede configurarse adicionalmente para ensamblar unidades de acceso que incluyen unidades NAL de acuerdo con las restricciones definidas por un formato de archivo. Por ejemplo, en un caso, donde se permite almacenar un conjunto de parámetros dentro de una muestra para ser utilizado para una imagen codificada, debe enviarse antes de la muestra que contiene esa imagen o en la muestra para esa imagen. Con referencia a la FIG. 1, si una imagen está contenida dentro de la unidad de acceso 150B, de acuerdo con esta restricción de ejemplo, para un conjunto de parámetros a usar con la unidad de acceso 150B, debe almacenarse en la caja de metadatos 144, la unidad de acceso 150A o la unidad de acceso 150B. Cabe señalar que el almacenamiento de conjuntos de parámetros en las descripciones de la muestra de un flujo de vídeo proporciona una manera simple y estática de suministrar conjuntos de parámetros. Sin embargo, almacenar conjuntos de parámetros en muestras puede ser más complejo que almacenar conjuntos de parámetros en la descripción de la muestra, pero permite más dinamismo en el caso de las actualizaciones del conjunto de parámetros y en el caso de agregar conjuntos de parámetros adicionales. Una actualización del conjunto de parámetros se refiere al contenido de un conjunto de parámetros en particular, pero el ID del conjunto de parámetros permanece igual.

[95] Además, un formato de archivo utilizado por el módulo 72 de creación de vídeo puede definir el formato de las unidades de acceso en un flujo elemental de vídeo HEVC, que también puede denominarse entradas de la muestra visual de HEVC. En un ejemplo, la sintaxis de una unidad de acceso puede configurarse a través de la configuración específica del descodificador para el flujo elemental de HEVC. Además, en un ejemplo, el nombre de la descripción de la muestra y el formato asociado con una o más unidades de acceso se pueden definir de la siguiente manera: (1) los tipos de cajas se pueden definir como 'hvc1' o 'hvcC'; (2) un contenedor puede definirse como Caja de tabla de muestra ('stbl'); (3) una entrada obligatoria puede incluir una entrada de muestra 'hvcI'; y (4) se puede definir una cantidad para permitir que una o más entradas de la muestra estén presentes. Además, en un ejemplo, puede requerirse que una entrada de la muestra visual de HEVC contenga una caja de configuración de HEVC, como se describe con más detalle a continuación. En algunos ejemplos, una caja de configuración de HEVC puede incluir un registro de configuración de descodificador HEVC, como se describe anteriormente.

[96] Un elemento de sintaxis opcional MPEG4BitRateBox puede estar presente en una entrada de la muestra visual HEVC para señalar la información de tasa de bits del flujo de vídeo HEVC. Los descriptores de ampliación que deberían insertarse en el Descriptor de flujo elemental, cuando se usan en MPEG-4, también pueden estar presentes. En algunos ejemplos, se pueden usar múltiples descripciones de la muestra, según lo permitido por la especificación del formato de archivo multimedia de base ISO, para indicar secciones de vídeo que usan configuraciones o conjuntos de parámetros diferentes.

[97] En algunos ejemplos, el nombre de la entrada de la muestra 'hvcI' solo se puede usar cuando el flujo al que se aplica esta entrada de la muestra es un flujo HEVC dócil y compatible visto por un descodificador HEVC que opera bajo la configuración (incluyendo perfil y nivel) dada en el HEVCConfigurationBox.

[98] En un ejemplo, el módulo 72 de creación de archivos de vídeo puede configurarse para definir entradas de las muestras visuales HEVC de acuerdo con la siguiente sintaxis:

```
// Secuencias visuales
```

```

class HEVCConfigurationBox extends Box('hvcC') {
    HEVCDecoderConfigurationRecord() HEVCConfig;
}
class HEVCSampleEntry() extends VisualSampleEntry ('hvc1'){
5     HEVCConfigurationBox      config;
    MPEG4BitRateBox ();          //opcional
    MPEG4ExtensionDescriptorsBox (); //opcional
    extra_boxes                  boxes;          //opcional
}

```

10 **[99]** Además, en el ejemplo anterior de sintaxis, Compressorname como se define en la clase base VisualSampleEntry puede indicar el nombre del compresor utilizado con el valor "\013HEVC Coding" que se recomienda (012 es 10, la longitud de la cadena en bytes). Los elementos de sintaxis Compressorname se pueden definir de acuerdo con HEVCFF. El elemento de sintaxis config se puede definir de acuerdo con un registro de configuración del descodificador.

20 **[100]** El módulo 72 de creación de archivos de vídeo puede configurarse, además, para ensamblar unidades de acceso, incluidas las unidades NAL, donde los campos de anchura y altura se incluyen en VisualSampleEntry. En un ejemplo, los campos de anchura y altura en VisualSampleEntry deben documentar correctamente las dimensiones de la imagen recortada (tamaño de presentación visual) del flujo HEVC que se describe en esa entrada. En un ejemplo, los campos de anchura y altura no reflejan ningún cambio en el tamaño causado por los mensajes SEI como pan-scan (barrido horizontal). En algunos ejemplos, el manejo visual de los mensajes SEI como pan-scan es opcional y dependiente del terminal. Además, en algunos casos, si la anchura y la altura de la secuencia cambian, entonces se necesita una nueva descripción de la muestra. En algunos ejemplos, los campos de anchura y altura de la cabecera de la pista pueden no ser los mismos que los campos de anchura y altura en una o más de una VisualSampleEntry en la pista de vídeo. Tal como se especifica en el Formato de archivo multimedia de base ISO, si se necesita una presentación visual normalizada, todas las secuencias se pueden normalizar según la anchura y la altura de la pista para la presentación.

30 **[101]** Después de que la unidad 72 de creación de archivos de vídeo genera un archivo de vídeo, la unidad 72 de creación de archivos de vídeo puede pasar el archivo de vídeo a la interfaz 76 de salida de archivos de vídeo. La interfaz 76 de salida de archivos de vídeo puede enviar el archivo de vídeo, por ejemplo, a la interfaz 22 de salida del dispositivo de origen 20. En algunos ejemplos, la interfaz 76 de salida de archivo de vídeo puede enviar el archivo de vídeo a un medio de almacenamiento del dispositivo de origen 20 (no mostrado) o al dispositivo de almacenamiento 32.

40 **[102]** La FIG. 6 es un diagrama conceptual que ilustra la estructura de un archivo multimedia HEVC de ejemplo basado en el ISOBMFF y el HEVCFF actual. El archivo 242 multimedia HEVC representa un ejemplo de un texto de formato de archivo que utiliza las técnicas de esta divulgación donde (1) conjuntos de parámetros pueden incluirse en la descripción de la muestra o en las muestras del flujo al que se aplica la descripción de la muestra; (2) los diferentes tipos de conjuntos de parámetros no están agrupados cuando se trata de si están incluidos en la descripción de la muestra o en las muestras del flujo al que se aplica la descripción de la muestra; (3) una descripción de la muestra incluye una matriz para cada tipo de conjunto de parámetros, y la matriz para un tipo particular de conjunto de parámetros solo puede incluir unidades NAL de conjunto de parámetros de ese tipo particular y la descripción de la muestra también incluye una matriz para unidades NAL de SEI que contienen mensajes SEI declarativos; (4) las propiedades del flujo se incluyen en la descripción de la muestra. Debe observarse que el archivo 242 multimedia HEVC de ejemplo de la FIG. 6 está destinado a ilustrar la relación lógica entre flujos de datos de vídeo y metadatos. En aras de la brevedad, no se ha ilustrado la estructura de encapsulación completa del archivo 242 multimedia HEVC. Sin embargo, el archivo 242 multimedia HEVC puede utilizar cajas y estructuras de encapsulación definidas de acuerdo con HEVCFF. El archivo multimedia HEVC puede ser generado por la unidad 72 de creación de archivos de vídeo de acuerdo con la sintaxis y las restricciones de ejemplo descritas anteriormente.

55 **[103]** En el ejemplo ilustrado en la FIG. 6, el archivo 242 multimedia HEVC incluye una caja de metadatos 244 y un flujo de datos de vídeo 246. La caja de metadatos 244 puede ser similar a una caja ISOBMFF 'moov' y puede contener datos para flujos 266 de datos de vídeo, tal como el registro 250 de configuración del descodificador. En un ejemplo, la tabla de metadatos puede ser una caja de tabla de muestras. Tanto la caja de metadatos 244 como el registro 250 de configuración del descodificador pueden formar porciones de una descripción de la muestra. En un ejemplo, el registro 250 de configuración del descodificador se puede generar usando la sintaxis de ejemplo descrita anteriormente y puede incluir información con respecto a los conjuntos de parámetros y los mensajes de la unidad

NAL de SEI. Además, el registro de configuración del descodificador incluye ELEMENTOS SINTÁCTICOS. Los ELEMENTOS SINTÁCTICOS pueden incluir información sobre las propiedades del flujo, como la resolución espacial. En un ejemplo, el registro 252 de configuración del descodificador puede incluir la matriz 254 de SPS y la matriz 258 de SEI. La matriz 254 de SPS puede incluir la unidad NAL de SPS 255. La matriz 258 de SEI puede incluir la unidad NAL de SEI 259.

**[104]** Los flujos 246 de datos de vídeo pueden corresponder a un vídeo o pista visual descrita en HEVCFF. Por lo tanto, el flujo 246 de datos de vídeo puede, por lo tanto, usar: (1) un handler\_type de 'vide' en la HandlerBox; (2) una cabecera multimedia de vídeo 'vmhd'; y (3) una derivada de VisualSampleEntry. El flujo 246 de datos de vídeo puede incluir una pluralidad de unidades de acceso 260A-260N. Las unidades de acceso 260A-260N pueden ser generadas por la unidad 72 de creación de archivos de vídeo de acuerdo con la sintaxis y las restricciones de ejemplo descritas anteriormente. Como se ilustra en la FIG. 6 la unidad de acceso 260A incluye la unidad NAL de PPS 256A y una unidad NAL de VCL 256B y la unidad de acceso 260N incluye una unidad NAL de APS 256C y NAL de VCL 256D. En un ejemplo, el flujo 246 de datos de vídeo puede incluir una o más capas temporales y la caja de metadatos 244 puede incluir, además, una indicación de si el flujo de datos de vídeo 246 está dividido en una o más pistas.

**[105]** De esta manera, el archivo HEVC 242 representa un archivo de ejemplo donde (1) los conjuntos de parámetros se incluyen en la descripción de la muestra y en las muestras del flujo al que se aplica la descripción de la muestra; (2) los diferentes tipos de conjuntos de parámetros no están agrupados cuando se trata de si están incluidos en la descripción de la muestra o en las muestras del flujo al que se aplica la descripción de la muestra; (3) una descripción de la muestra incluye una matriz para cada tipo de conjunto de parámetros y la descripción de la muestra también incluye una matriz para unidades NAL de SEI que contiene mensajes SEI declarativos; y (4) las propiedades del flujo están incluidas en la descripción de la muestra.

**[106]** Con referencia de nuevo a la FIG. 3, el dispositivo de origen 12 puede configurarse para emitir datos de vídeo codificados de acuerdo con uno o más formatos de archivo y el dispositivo de destino 14 puede configurarse para recibir datos de vídeo codificados en uno o más formatos de archivo, tales como los descritos en el presente documento. Los archivos de vídeo descritos en el presente documento pueden transmitirse directamente al dispositivo de destino 14 a través de la interfaz de salida 22 del dispositivo de origen 12. El archivo de vídeo puede asimismo (o de forma alternativa) almacenarse en el dispositivo de almacenamiento 32 para un acceso posterior por el dispositivo de destino 14 u otros dispositivos, para su descodificación y/o reproducción.

**[107]** El dispositivo de destino 14 incluye una interfaz de entrada 28, un módulo de desencapsulación 29, un descodificador de vídeo 30 y un dispositivo de visualización 34. En algunos casos, la interfaz de entrada 28 puede incluir un receptor y/o un módem. La interfaz de entrada 28 del dispositivo de destino 14 recibe los datos de vídeo codificados por el enlace 16. Los datos de vídeo codificados, 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 descodificador de vídeo, tal como el descodificador de vídeo 30, en la descodificación de los datos de vídeo. Dichos elementos sintácticos pueden incluirse con los datos de vídeo codificados, transmitidos en un medio de comunicación, almacenados en un medio de almacenamiento o almacenados en un servidor de archivos. En algunos ejemplos, dichos elementos sintácticos pueden incluir elementos de sintaxis que señalizan modos de intracodificación como se describe en esta divulgación.

**[108]** El dispositivo de visualización 34 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 estar configurado para interconectarse 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 34 visualiza los datos de vídeo descodificados ante un usuario y puede comprender cualquiera de una variedad de dispositivos de visualización, tales como una pantalla de cristal líquido (LCD), una pantalla de plasma, una pantalla de diodos orgánicos emisores de luz (OLED) u otro tipo de dispositivo de visualización.

**[109]** El dispositivo de destino 14 puede recibir los datos de vídeo codificados que se van a descodificar mediante un enlace 16. El enlace 16 puede comprender cualquier tipo de medio o dispositivo capaz de desplazar los datos de vídeo codificados desde el dispositivo de origen 12 al dispositivo de destino 14. En un ejemplo, el enlace 16 puede comprender un medio de comunicación para permitir al dispositivo de origen 12 transmitir los datos de vídeo codificados directamente al dispositivo de destino 14 en tiempo real. Los datos de vídeo codificados pueden modularse de acuerdo con una norma de comunicación, tal como un protocolo de comunicación inalámbrica, y transmitirse al dispositivo de destino 14. El medio de comunicación puede comprender cualquier medio de comunicación inalámbrica o alámbrica, tal como un espectro de radiofrecuencia (RF) o una o más líneas de transmisión física. El medio de comunicación puede formar parte de una red basada en paquetes, tal como una red de área local, una red de área amplia o una red global tal como Internet. El medio de comunicación puede incluir 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. En algunos casos, la interfaz de salida 22 puede incluir un modulador/desmodulador (módem) y/o un transmisor.

**[110]** De forma 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 28. El dispositivo de almacenamiento 32 puede incluir cualquiera de una diversidad de medios de almacenamiento de datos, de acceso distribuido o local, tales como una unidad de disco duro, unos discos Blu-ray, discos DVD, discos CD-ROM, una memoria flash, memoria volátil o no volátil o cualquier otro medio de almacenamiento digital adecuado, para almacenar datos de vídeo codificados. En un ejemplo adicional, el dispositivo de almacenamiento 32 puede corresponder a un servidor de archivos o a otro dispositivo de almacenamiento intermedio que pueda retener el vídeo codificado generado por el dispositivo de origen 12. El dispositivo de destino 14 puede acceder a los datos de vídeo almacenados desde el dispositivo de almacenamiento 32 a través de transmisión en continuo o descarga. El servidor de archivos puede ser cualquier tipo de servidor capaz de almacenar datos de vídeo codificados y transmitir esos datos de vídeo codificados al dispositivo de destino 14. Ejemplos de servidores de archivos incluyen un servidor web (por ejemplo, para un sitio web), un servidor FTP, dispositivos de almacenamiento conectado en red (NAS) o una unidad de disco local. El dispositivo de destino 14 puede acceder a los datos de vídeo codificados mediante cualquier conexión de datos estándar, incluyendo una conexión a Internet. Esto puede incluir un canal inalámbrico (por ejemplo, una conexión Wi-Fi), una conexión alámbrica (por ejemplo, DSL, módem de cable, etc.), o una combinación de ambos que sea adecuada para acceder a datos de vídeo codificados almacenados en un servidor de archivos. La transmisión de datos de vídeo codificados desde el dispositivo de almacenamiento 32 puede ser una transmisión en continuo, una transmisión de descarga o una combinación de ambas.

**[111]** 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 de una diversidad de aplicaciones de multimedia, tales como radiodifusiones de televisión por el aire, transmisiones de televisión por cable, transmisiones de televisión por satélite, transmisiones de vídeo en continuo, por ejemplo, mediante Internet, codificación de vídeo digital para su almacenamiento en un medio de almacenamiento de datos, descodificación de vídeo digital almacenado en un medio de almacenamiento de datos, u otras aplicaciones. En algunos ejemplos, el sistema 10 puede configurarse para soportar la transmisión de vídeo unidireccional o bidireccional para soportar aplicaciones tales como la transmisión continua de vídeo, la reproducción de vídeo, la radiodifusión de vídeo y/o la videotelefonía.

**[112]** Aunque no se muestra en la FIG. 3, en algunos aspectos, el codificador de vídeo 20 y el descodificador de vídeo 30 pueden estar integrados, cada uno de ellos, en un codificador y descodificador de audio, y pueden incluir unidades MUX-DEMUX adecuadas, u otro tipo de hardware y software, para gestionar la codificación tanto de audio como de vídeo en un flujo de datos común o en flujos de datos diferentes. Si procede, en algunos ejemplos, las unidades MUX-DEMUX pueden conformarse al protocolo de multiplexado ITU H.223 o a otros protocolos, tales como el protocolo de datagramas de usuario (UDP).

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

**[114]** El módulo 29 de desencapsulación puede estar configurado para recibir un archivo de vídeo que analiza el archivo de vídeo de manera que el descodificador de vídeo 30 pueda emitir datos de vídeo descodificados. En algunos ejemplos, el módulo de desencapsulación 29 puede realizar procesos recíprocos del módulo de encapsulación 72. En algunos ejemplos, el descodificador de vídeo 30 puede realizar una pasada de descodificación que en general es recíproca a la pasada de codificación descrita con respecto al codificador de vídeo 20 de la FIG. 4.

**[115]** La FIG. 7 es un diagrama de flujo que ilustra un ejemplo de generación de un archivo de vídeo que incluye contenido de vídeo codificado de acuerdo con las técnicas de esta divulgación. El proceso descrito en la FIG. 7 puede realizarse mediante un dispositivo de codificación de vídeo que puede incluir cualquier combinación de dispositivo de origen 20, módulo de encapsulación 21 y módulo 72 de creación de vídeo descritos en el presente documento. Un dispositivo de codificación de vídeo puede obtener fragmentos de contenido de vídeo codificado (702). El contenido de vídeo codificado puede incluir una pluralidad de fragmentos codificados de acuerdo con HEVC. El dispositivo de codificación de vídeo puede obtener conjuntos de parámetros asociados con los fragmentos de contenido de vídeo (704). Los conjuntos de parámetros pueden incluir SPS, PPS y/o APS. El dispositivo de codificación de vídeo puede encapsular fragmentos de contenido de vídeo codificado dentro de las unidades de acceso (706). Las unidades de acceso se pueden definir como se describen en el presente documento. El dispositivo de codificación de vídeo puede encapsular conjuntos de parámetros dentro de las unidades de acceso (708). Los

conjuntos de parámetros se pueden encapsular dentro de las unidades de acceso de acuerdo con las técnicas descritas en el presente documento. El dispositivo de codificación de vídeo puede encapsular conjuntos de parámetros dentro de una descripción de la muestra (710). En un ejemplo, PPS puede estar encapsulado dentro de una o más unidades de acceso mientras que SPS está encapsulado dentro de una descripción de la muestra. El dispositivo de codificación de vídeo genera el archivo de vídeo (712). El archivo de vídeo puede generarse adicionalmente basándose en un formato de archivo en el que el formato de archivo se basa en las técnicas descritas en el presente documento.

**[116]** En uno o más ejemplos, las funciones descritas pueden implementarse en hardware, software, firmware o cualquier combinación de estos. Si se implementan en software, las funciones pueden almacenarse en o transmitirse a través de, como una o más instrucciones o código, un medio legible por ordenador o 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 unos medios de almacenamiento de datos o unos medios de comunicación que incluyen cualquier medio que facilite la transferencia de un programa informático desde un lugar a otro, por ejemplo, de acuerdo con un protocolo de comunicación. De esta manera, los medios legibles por ordenador pueden corresponder en general a (1) unos medios de almacenamiento tangibles legibles por ordenador que son 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 puede 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.

**[117]** A modo de ejemplo, y no de manera limitativa, tales medios de almacenamiento legibles por ordenador pueden comprender RAM, ROM, EEPROM, CD-ROM u otro almacenamiento de disco óptico, almacenamiento de disco magnético u otros dispositivos de almacenamiento magnético, memoria flash o cualquier otro medio que pueda usarse para almacenar código de programa deseado en forma de instrucciones o estructuras de datos y al que pueda accederse mediante un ordenador. Además, cualquier conexión recibe debidamente la denominación de medio legible por ordenador. Por ejemplo, si las instrucciones se transmiten desde un sitio web, un servidor u otro origen remoto usando un cable coaxial, un cable de fibra óptica, un par trenzado, una línea de abonado digital (DSL) o tecnologías inalámbricas tales como infrarrojos, radio y microondas, entonces el cable coaxial, el cable de fibra óptica, el par trenzado, la DSL o las tecnologías inalámbricas tales como infrarrojos, radio y microondas se incluyen en la definición de medio. Sin embargo, debería entenderse que los medios de almacenamiento legibles por ordenador y los medios de almacenamiento de datos no incluyen conexiones, ondas portadoras, señales u otros medios transitorios, sino que, en cambio, se orientan a medios de almacenamiento tangibles no transitorios. El término disco, tal como se utiliza en el presente documento, incluye un disco compacto (CD), un disco láser, un disco óptico, un disco versátil digital (DVD), un disco flexible y un disco Blu-ray, donde algunos discos habitualmente emiten datos magnéticamente, mientras que otros discos emiten datos ópticamente con láseres. Las combinaciones de los anteriores también deben incluirse dentro del alcance de los medios legibles por ordenador.

**[118]** 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), matriz de puertas programables in situ (FPGA) u otros circuitos lógicos integrados o discretos equivalentes. Por consiguiente, el término «procesador», como se utiliza en el presente documento, puede referirse a cualquiera de las estructuras anteriores o a cualquier otra estructura adecuada para la implementación de las técnicas descritas en el presente documento. Además, en algunos aspectos, la funcionalidad descrita en el presente documento puede proporcionarse dentro de módulos de hardware y/o software dedicados configurados para la codificación y la decodificación, o incorporarse en un códec combinado. Asimismo, las técnicas podrían implementarse por completo en uno o más circuitos o elementos lógicos.

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

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



**REIVINDICACIONES**

1. Un procedimiento para generar un archivo de vídeo HEVC que incluye contenido codificado de vídeo, el procedimiento que comprende:
- 5           obtener (702) una pluralidad de fragmentos de contenido de vídeo codificado;
- obtener (704) una primera pluralidad de conjuntos de parámetros de un primer tipo asociado con la pluralidad de fragmentos de contenido de vídeo y una segunda pluralidad de conjuntos de parámetros de un segundo tipo asociado con la pluralidad de fragmentos de contenido de vídeo;
- 10           encapsular (706) la pluralidad de fragmentos de contenido de vídeo codificado dentro de una pluralidad de unidades de acceso de un flujo de vídeo;
- 15           encapsular (708) la pluralidad de unidades de acceso dentro de una pluralidad de muestras en una pista de archivo;
- encapsular la primera pluralidad de conjuntos de parámetros; y
- 20           encapsular la segunda pluralidad de conjuntos de parámetros; **caracterizado por**
- encapsular la primera pluralidad de conjuntos de parámetros dentro de la pluralidad de muestras y encapsular (710) la segunda pluralidad de conjuntos de parámetros dentro de una descripción de la muestra de la pista de archivo para soportar la transmisión en banda de la primera pluralidad de conjuntos de parámetros mientras se admite la transmisión fuera de banda de la segunda pluralidad de conjuntos de parámetros.
- 25
2. Un procedimiento para procesar un archivo de vídeo HEVC que incluye contenido codificado de vídeo, el procedimiento que comprende:
- 30           obtener una pista de archivo que incluye una pluralidad de muestras;
- desencapsular una pluralidad de unidades de acceso de la pluralidad de muestras de un flujo de vídeo en la pista de archivo;
- 35           desencapsular una pluralidad de fragmentos de contenido de vídeo codificado a partir de la pluralidad de unidades de acceso;
- 40           desencapsular una primera pluralidad de conjuntos de parámetros de un primer tipo asociado con la pluralidad de fragmentos de contenido de vídeo; y
- 45           desencapsular una segunda pluralidad de conjuntos de parámetros de un segundo tipo asociado con la pluralidad de fragmentos de contenido de vídeo; **caracterizado por**
- desencapsular la primera pluralidad de conjuntos de parámetros de la pluralidad de muestras y desencapsular la segunda pluralidad de conjuntos de parámetros de una descripción de la muestra de la pista de archivo para admitir la transmisión en banda de la primera pluralidad de conjuntos de parámetros mientras se admite la transmisión fuera de banda de la segunda pluralidad de conjuntos de parámetros.
- 50
3. El procedimiento según la reivindicación 1 o 2, en el que el primer tipo es un conjunto de parámetros de imagen (PPS) y el segundo tipo es un conjunto de parámetros de secuencia (SPS).
4. El procedimiento según la reivindicación 1 o 2, en el que la descripción de la muestra incluye un indicador que identifica un número de conjuntos de parámetros del segundo tipo almacenado dentro de la descripción de la muestra.
- 55
5. El procedimiento según la reivindicación 1 o 2, en el que la segunda pluralidad de conjuntos de parámetros consiste en conjuntos de parámetros de diferentes tipos.
- 60
6. El procedimiento según la reivindicación 5, en el que las unidades de capa de abstracción de red (NAL) de conjuntos de parámetros de cada tipo se almacenan en una matriz dedicada en la descripción de la muestra.
7. El procedimiento según la reivindicación 6, en el que la descripción de la muestra incluye, además, una matriz que incluye unidades de capa de abstracción de red (NAL) de información de mejora suplementaria (SEI).
- 65
8. El procedimiento según la reivindicación 1 o 2, en el que el primer tipo y el segundo tipo son ambos de un

mismo tipo, de manera que la primera pluralidad de conjuntos de parámetros y la segunda pluralidad de conjuntos de parámetros incluyen conjuntos de parámetros del mismo tipo, y en el que la pista de archivo incluye una indicación de si al menos algunos conjuntos de parámetros del mismo tipo pueden estar encapsulados dentro de una muestra y dentro de la descripción de la muestra.

- 5
9. El procedimiento según la reivindicación 1 o 2, en el que la primera pluralidad de conjuntos de parámetros consiste en conjuntos de parámetros de diferentes tipos.
- 10
10. El procedimiento según la reivindicación 1 o 2, en el que la descripción de la muestra incluye información asociada con al menos una de una velocidad de trama y una resolución espacial del flujo de vídeo.
11. El procedimiento según la reivindicación 1 o 2, en el que el flujo de vídeo incluye múltiples capas temporales y en el que la descripción de la muestra incluye además una indicación de un número de las múltiples capas temporales del flujo de vídeo.
- 15
12. El procedimiento según la reivindicación 1 o 2, en el que la pista de archivo contiene una representación de una asignación de las muestras en la pista de archivo a capas temporales así como una descripción de características para cada una de las capas temporales, en el que la descripción de características incluye al menos uno de una identificación de capa temporal, un perfil, un nivel, una velocidad de bits y una velocidad de trama.
- 20
13. Un medio de almacenamiento no transitorio legible por ordenador que tiene instrucciones almacenadas en el mismo que, durante la ejecución, hace que uno o más procesadores de un dispositivo de codificación de vídeo lleven a cabo el procedimiento de cualquiera de las reivindicaciones 1 a 12.
- 25
14. Un aparato configurado para generar un archivo de vídeo HEVC que incluye contenido de vídeo codificado, el aparato que comprende:
- 30
- medios (70) para obtener una pluralidad de fragmentos de contenido de vídeo codificado;
- medios para obtener una primera pluralidad de conjuntos de parámetros de un primer tipo asociado con la pluralidad de fragmentos de contenido de vídeo y una segunda pluralidad de conjuntos de parámetros de un segundo tipo asociado con la pluralidad de fragmentos de contenido de vídeo;
- 35
- medios (21) para encapsular la pluralidad de fragmentos de contenido de vídeo codificado dentro de una pluralidad de unidades de acceso (260A, 260B, ..., 260N) de un flujo de vídeo (246);
- medios (21) para encapsular la pluralidad de unidades de acceso dentro de una pluralidad de muestras en una pista de archivo;
- 40
- medios (21) para encapsular la primera pluralidad de conjuntos de parámetros;
- y medios (21) para encapsular la segunda pluralidad de conjuntos de parámetros;
- 45
- caracterizado por**
- los medios para encapsular la primera pluralidad de conjuntos de parámetros que se configuran para encapsular la primera pluralidad de conjuntos de parámetros (256A; 256C) dentro de la pluralidad de muestras; y
- 50
- los medios para encapsular la segunda pluralidad de conjuntos de parámetros que se configuran para encapsular la segunda pluralidad de conjuntos de parámetros (255) dentro de una descripción de muestra de la pista de archivo (252).
- 55
15. Un aparato (29) configurado para procesar un archivo de vídeo HEVC que incluye contenido de vídeo codificado, el aparato que comprende:
- medios (28) para obtener una pista de archivo que incluye una pluralidad de muestras;
- 60
- medios (29) para desencapsular una pluralidad de unidades de acceso de la pluralidad de muestras de un flujo de vídeo en la pista de archivo;
- medios (29) para desencapsular una pluralidad de fragmentos de contenido de vídeo codificado de la pluralidad de unidades de acceso;
- 65
- medios (29) para desencapsular una primera pluralidad de conjuntos de parámetros de un primer tipo

asociado con la pluralidad de fragmentos de contenido de vídeo; y

medios (29) para desencapsular una segunda pluralidad de conjuntos de parámetros de un segundo tipo asociado con la pluralidad de fragmentos de contenido de vídeo; **caracterizado por**

5 los medios para desencapsular la primera pluralidad de conjuntos de parámetros que están configurados para desencapsular la primera pluralidad de conjuntos de parámetros a partir de la pluralidad de muestras; y

10 los medios para desencapsular la segunda pluralidad de conjuntos de parámetros que están configurados para desencapsular la segunda pluralidad de conjuntos de parámetros a partir de una descripción de la muestra de la pista de archivo.

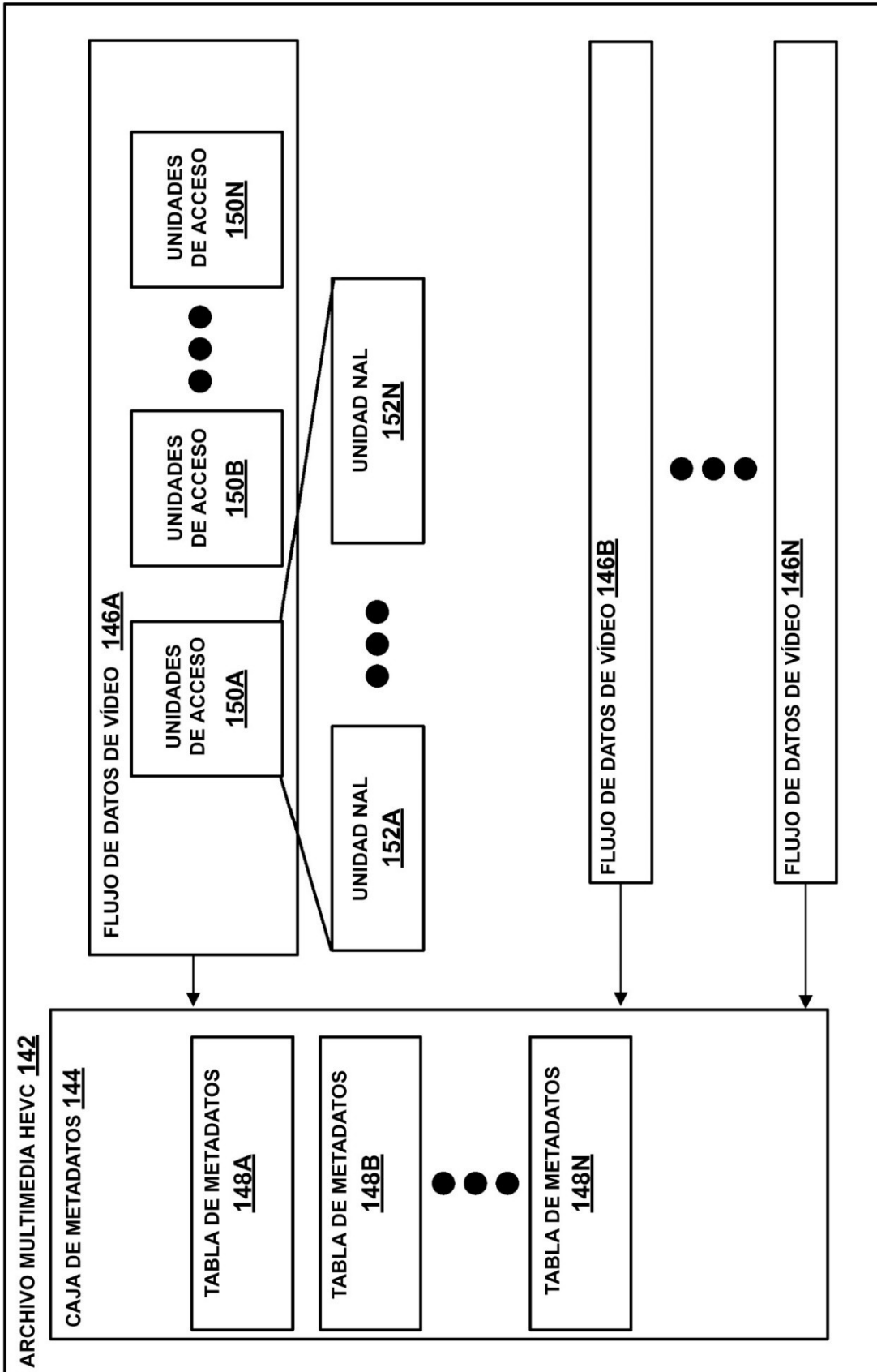
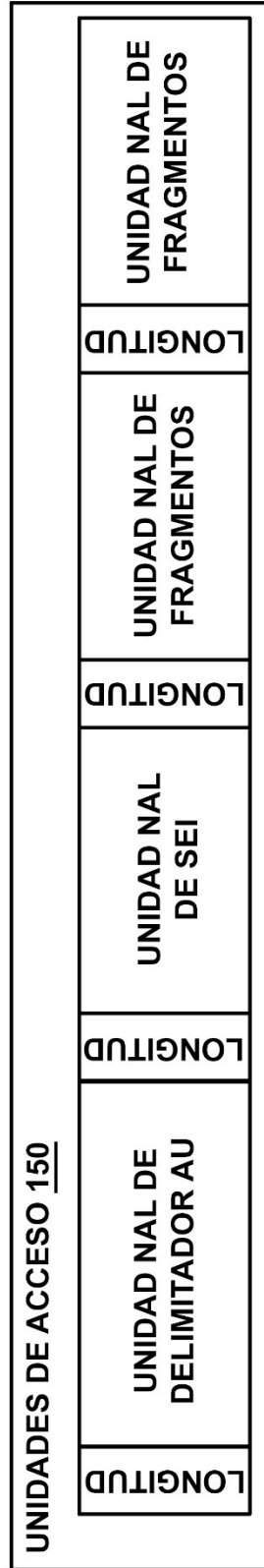


FIG. 1



**FIG. 2**

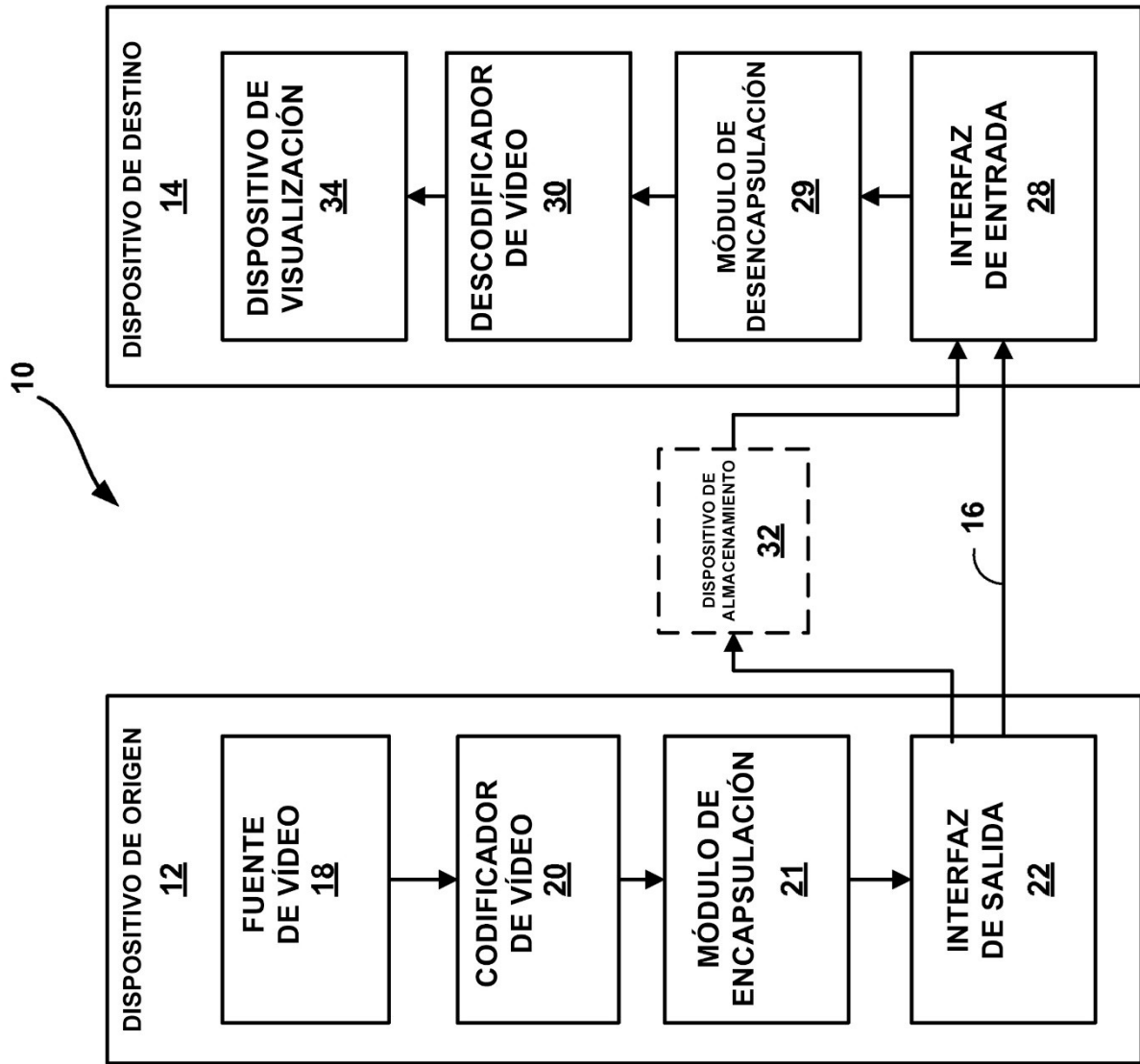


FIG. 3

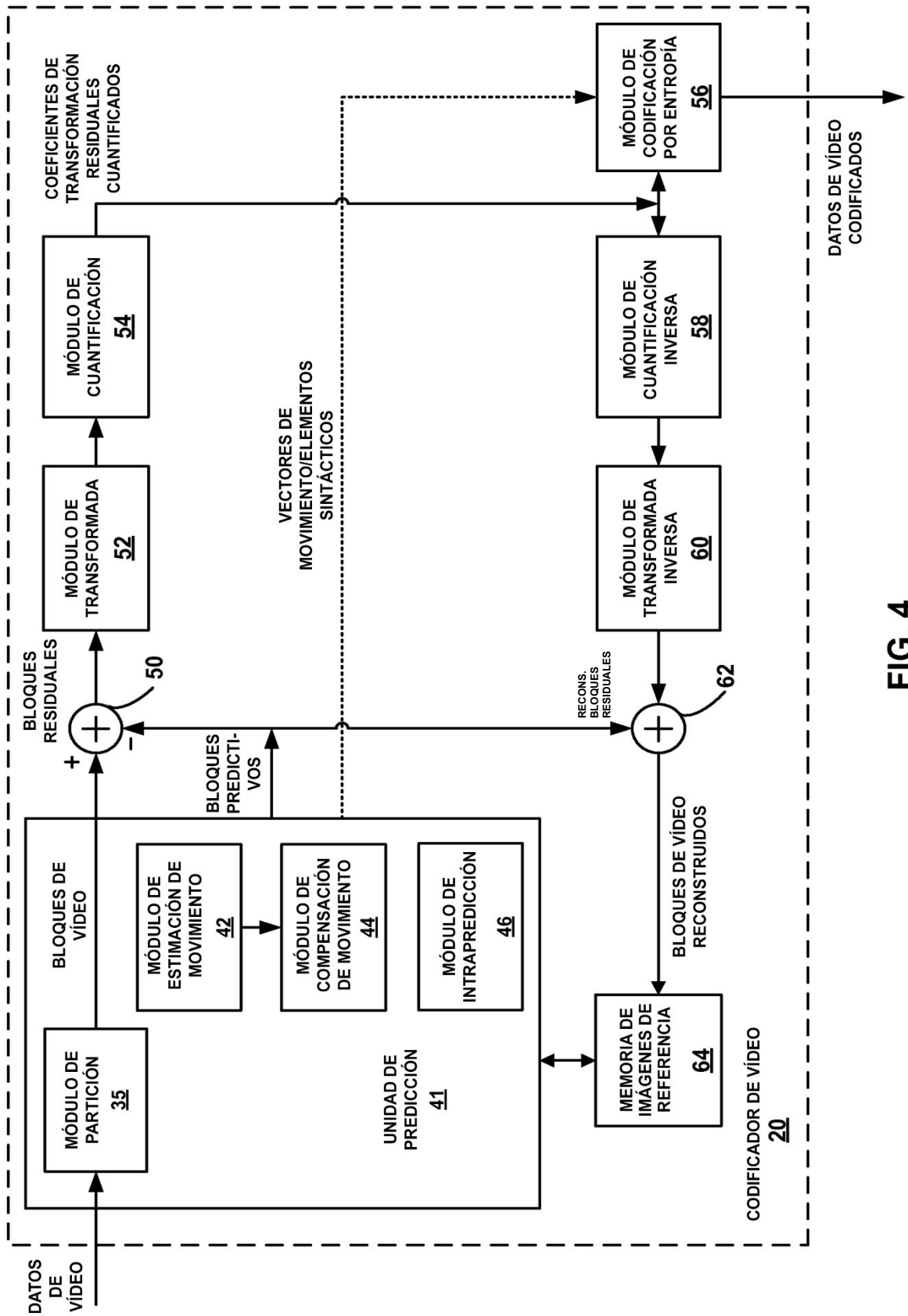


FIG. 4

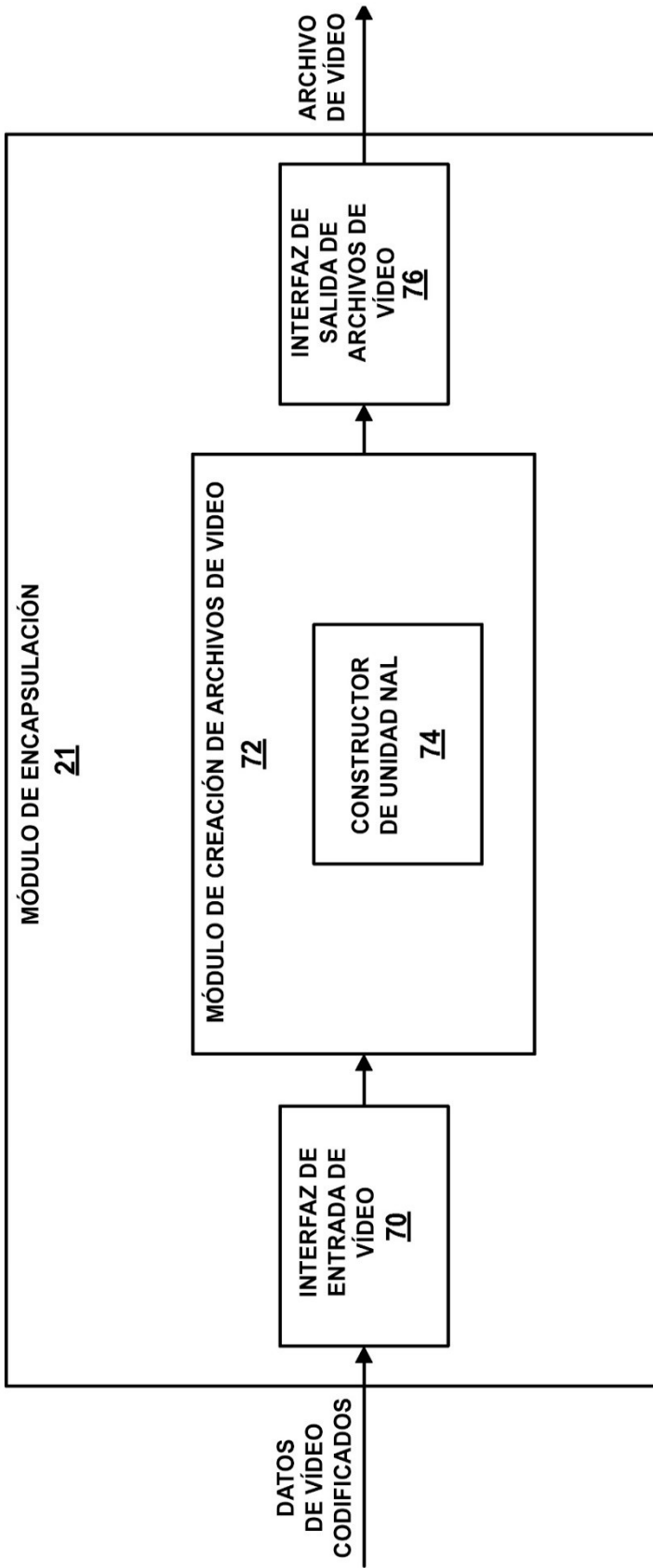


FIG. 5



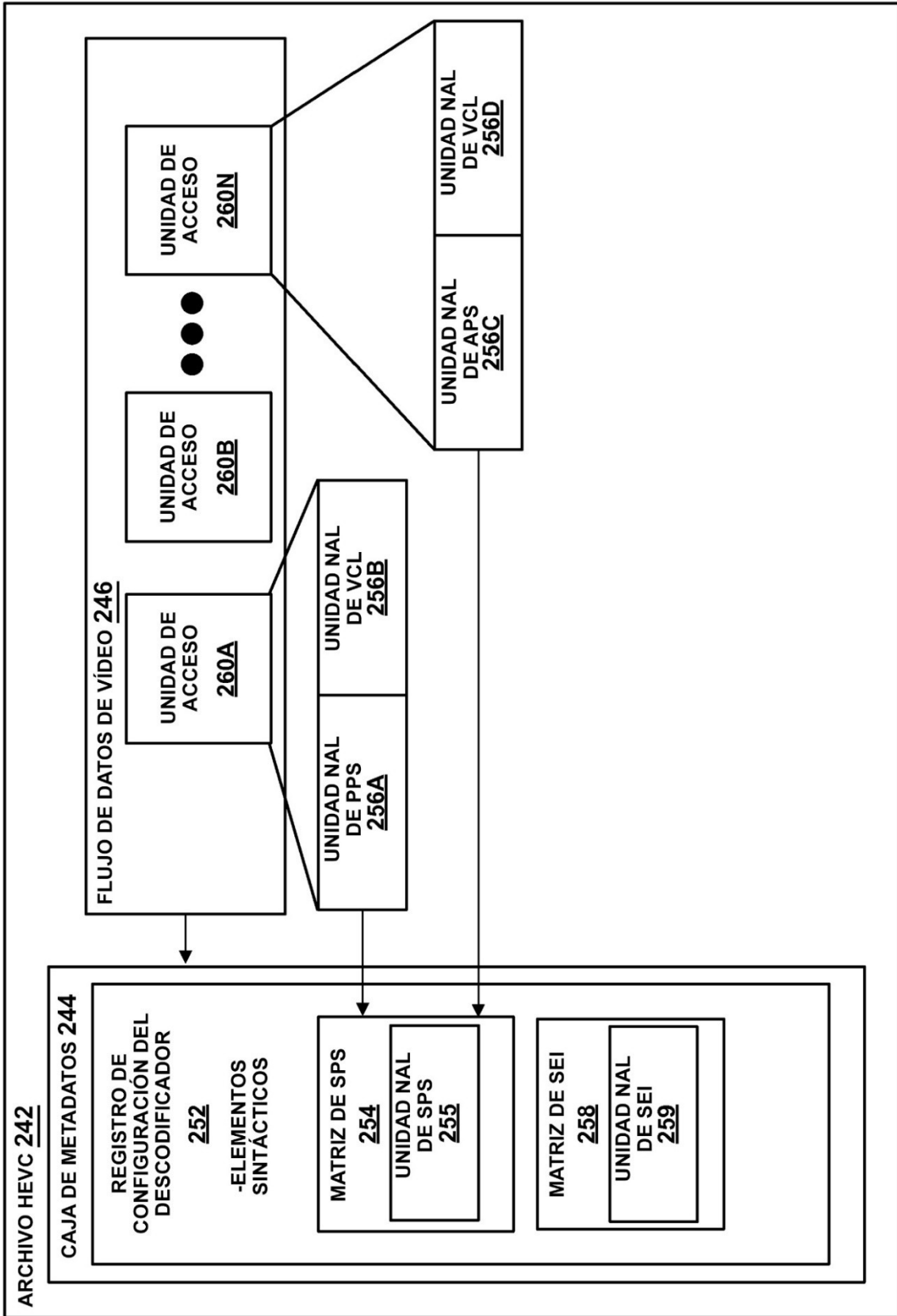
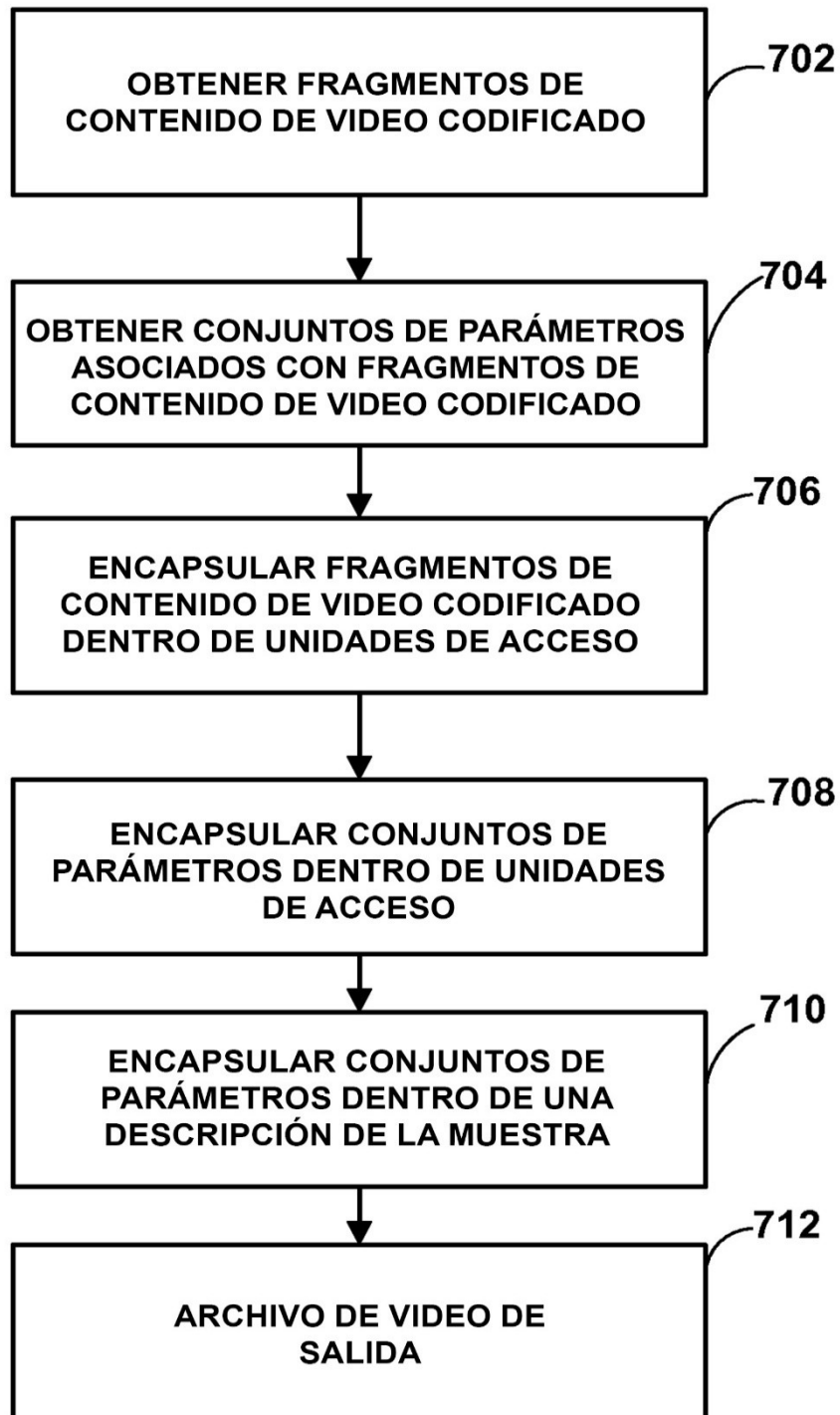


FIG. 6



**FIG. 7**