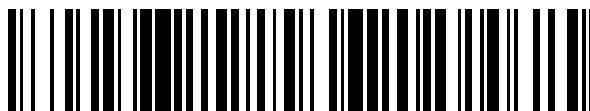


19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 720 662**

51 Int. Cl.:

H04N 19/70 (2014.01)

H04N 19/30 (2014.01)

H04N 19/573 (2014.01)

H04N 19/58 (2014.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

86 Fecha de presentación y número de la solicitud internacional: **23.10.2014** **PCT/US2014/061988**

87 Fecha y número de publicación internacional: **30.04.2015** **WO15061580**

96 Fecha de presentación y número de la solicitud europea: **23.10.2014** **E 14795743 (5)**

97 Fecha y número de publicación de la concesión europea: **16.01.2019** **EP 3061250**

54 Título: **Diseños de formato de archivo de vídeo multicapa**

30 Prioridad:

23.10.2013 US 201361894886 P
22.10.2014 US 201414521153

45 Fecha de publicación y mención en BOPI de la traducción de la patente:
23.07.2019

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;
CHEN, YING;
RAMASUBRAMONIAN, ADARSH KRISHNAN y
HENDRY, FNU

74 Agente/Representante:

FORTEA LAGUNA, Juan José

ES 2 720 662 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

Diseños de formato de archivo de vídeo multicapa

5 CAMPO TÉCNICO

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

ANTECEDENTES

10 [0002] Las capacidades de vídeo digital pueden incorporarse a una amplia gama de dispositivos, que incluye televisores digitales, sistemas de radiodifusión digital directa, sistemas de radiodifusión inalámbrica, asistentes digitales personales (PDA), ordenadores portátiles o de escritorio, ordenadores de tableta, lectores de libros electrónicos, cámaras digitales, dispositivos de grabación digital, reproductores de medios digitales, dispositivos de videojuegos, consolas de videojuegos, teléfonos celulares o de radio por satélite, los denominados «teléfonos inteligentes», dispositivos de videoconferencia, dispositivos de transmisión de vídeo en tiempo real y similares. Los dispositivos de vídeo digital implementan técnicas de compresión de vídeo, tales como las descritas en las normas definidas por MPEG-2, MPEG-4, ITU-T H.263, ITU-T H.264/MPEG-4, parte 10, codificación de vídeo avanzada (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 con más eficiencia implementando dichas técnicas de compresión de vídeo.

25 [0003] Las técnicas de compresión de vídeo realizan predicción espacial (dentro de las imágenes) y/o predicción temporal (entre imágenes) para reducir o eliminar la redundancia intrínseca a las secuencias de vídeo. Para la codificación de vídeo basada en bloques, un sector de vídeo (por ejemplo, una trama de vídeo o una parte de una trama de vídeo) puede dividirse en bloques de vídeo, que también pueden denominarse bloques de árbol, unidades de codificación (CU) y/o nodos de codificación. Los bloques de vídeo de un sector intracodificado (I) de una imagen se codifican usando predicción espacial con respecto a unas muestras de referencia de bloques contiguos de la misma imagen. Los bloques de vídeo de un sector intercodificado (P o B) de una imagen pueden usar predicción espacial con respecto a unas muestras de referencia de bloques contiguos de la misma imagen o predicción temporal con respecto a unas muestras de referencia de otras imágenes de referencia. Las imágenes pueden denominarse tramas, y las imágenes de referencia pueden denominarse tramas de referencia.

35 [0004] La predicción espacial o temporal da como resultado un bloque predictivo para un bloque que se va a codificar. Los datos residuales representan diferencias de píxeles entre el bloque original que se va 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 del píxel a un dominio de transformada, dando como resultado coeficientes de transformada residuales, que a continuación se pueden cuantificar. Los coeficientes de transformada cuantificados, dispuestos inicialmente en una matriz bidimensional, pueden explorarse con el fin de generar un vector unidimensional de coeficientes de transformada, y puede aplicarse codificación de entropía para lograr aún más compresión.

45 [0005] En el documento «MV-HEVC/SHVC HLS: On non-HEVC base layer [MV-HEVC/SHVC HLS: Sobre capa base no HEVC]» de Hannuksela MM, 15.^a reunión JCT-VC; Ginebra; documento n.º JCTVC-O0166, 15 de octubre de 2013, se analiza una necesidad de disponer de un formato de flujo de bits que encapsule una capa base codificada no HEVC y unas capas de mejora MV-HEVC/SHVC y se afirma que no se prevé dicha necesidad.

50 [0006] En «Test Model 1 of Video Coding for Browsers [Modelo de prueba 1 de codificación de vídeo para navegadores]», 105.^a reunión MPEG; Viena; documento n.º N13777, 31 de agosto de 2013, se presenta una visión técnica general del modelo de prueba de codificación de vídeo para navegadores (VCB) (es decir, el codificador).

SUMARIO

55 [0007] En general, esta divulgación se refiere al almacenamiento de contenido de vídeo en un archivo basándose en un formato base de archivos de medios (ISO/BMFF) de la Organización Internacional de Normalización (ISO). Algunos ejemplos de esta divulgación se refieren a procedimientos para almacenar flujos de vídeo que contienen múltiples capas codificadas, donde cada capa puede ser una capa escalable, una vista de textura, una vista de profundidad, etc., y los procedimientos pueden aplicarse al almacenamiento de codificación de vídeo de alta eficiencia multivista (MV-HEVC), HEVC escalable (SHVC), HEVC de 3 dimensiones (3D-HEVC) y otros tipos de datos de vídeo.

65 [0008] En un aspecto, esta divulgación describe un procedimiento de procesamiento de datos de vídeo multicapa, comprendiendo el procedimiento: generar un archivo que comprende una caja de datos de medios que contiene un contenido de medios, comprendiendo el contenido de medios una secuencia de muestras, siendo cada una de las muestras una unidad de acceso de los datos de vídeo multicapa, en el que generar el archivo comprende: como

respuesta a una determinación de que al menos una unidad de acceso de un flujo de bits de los datos de vídeo multicapa incluye una imagen codificada que tiene un indicador [flag] de salida de imagen igual a un primer valor y una imagen codificada que tiene un indicador de salida de imagen igual a un segundo valor, usar al menos una primera pista y una segunda pista para almacenar el flujo de bits en el archivo, en el que: para cada pista respectiva de la primera y la segunda pistas, todas las imágenes codificadas de cada muestra de la pista respectiva tienen el mismo valor del indicador de salida de imagen; y se permite que unas imágenes que tienen indicadores de salida de imagen iguales al primer valor se faciliten y se permite que unas imágenes que tienen indicadores de salida de imagen iguales al segundo valor se usen como imágenes de referencia, pero no se permite que se faciliten.

[0009] En otro aspecto, esta divulgación describe un procedimiento de procesamiento de datos de vídeo multicapa, comprendiendo el procedimiento: obtener, a partir de un archivo, una primera caja de pista y una segunda caja de pista, conteniendo la primera caja de pista metadatos para una primera pista del archivo, conteniendo la segunda caja de pista metadatos para una segunda pista del archivo, en el que: cada una de la primera pista y la segunda pista comprende una secuencia de muestras, siendo cada una de las muestras una unidad de acceso de vídeo de los datos de vídeo multicapa, para cada pista respectiva de la primera pista y la segunda pista, todas las imágenes codificadas de cada muestra de la pista respectiva tienen el mismo valor de un indicador de salida de imagen, y se permite que unas imágenes que tienen indicadores de salida de imagen iguales a un primer valor se faciliten y se permite que unas imágenes que tienen indicadores de salida de imagen iguales a un segundo valor se usen como imágenes de referencia, pero no se permite que se faciliten.

[0010] En otro aspecto, esta divulgación describe un dispositivo de vídeo que comprende: un medio de almacenamiento de datos configurado para almacenar datos de vídeo multicapa; y uno o más procesadores configurados para: generar un archivo que comprende una caja de datos de medios que contiene un contenido de medios, comprendiendo el contenido de medios una secuencia de muestras, siendo cada una de las muestras una unidad de acceso de los datos de vídeo multicapa, en el que para generar el archivo, el uno o más procesadores están configurados para: como respuesta a una determinación de que al menos una unidad de acceso de un flujo de bits de los datos de vídeo multicapa incluye una imagen codificada que tiene un indicador de salida de imagen igual a un primer valor y una imagen codificada que tiene un indicador de salida de imagen igual a un segundo valor, usar al menos una primera pista y una segunda pista para almacenar el flujo de bits en el archivo, en el que: para cada pista respectiva de la primera y la segunda pistas, todas las imágenes codificadas de cada muestra de la pista respectiva tienen el mismo valor del indicador de salida de imagen; y se permite que unas imágenes que tienen indicadores de salida de imagen iguales al primer valor se faciliten y se permite que unas imágenes que tienen indicadores de salida de imagen iguales al segundo valor se usen como imágenes de referencia, pero no se permite que se faciliten.

[0011] En otro aspecto, esta divulgación describe un dispositivo de vídeo que comprende: un medio de almacenamiento de datos configurado para almacenar datos de vídeo multicapa; y uno o más procesadores configurados para: obtener, a partir de un archivo, una primera caja de pista y una segunda caja de pista, conteniendo la primera caja de pista metadatos para una primera pista del archivo, conteniendo la segunda caja de pista metadatos para una segunda pista del archivo, en el que: cada una de la primera pista y la segunda pista comprenden una secuencia de muestras, siendo cada una de las muestras una unidad de acceso de vídeo de los datos de vídeo multicapa, para cada pista respectiva de la primera pista y la segunda pista, todas las imágenes codificadas de cada muestra de la pista respectiva tienen el mismo valor de un indicador de salida de imagen, y se permite que unas imágenes que tienen indicadores de salida de imagen iguales a un primer valor se faciliten y se permite que unas imágenes que tienen indicadores de salida de imagen iguales a un segundo valor se usen como imágenes de referencia, pero no se permite que se faciliten.

[0012] En otro aspecto, esta divulgación describe un dispositivo de vídeo que comprende: medios para generar un archivo que comprende una caja de datos de medios que contiene un contenido de medios, comprendiendo el contenido de medios una secuencia de muestras, siendo cada una de las muestras una unidad de acceso de datos de vídeo multicapa, en el que generar el archivo comprende: como respuesta a una determinación de que al menos una unidad de acceso de un flujo de bits de los datos de vídeo multicapa incluye una imagen codificada que tiene un indicador de salida de imagen igual a un primer valor y una imagen codificada que tiene un indicador de salida de imagen igual a un segundo valor, usar al menos una primera pista y una segunda pista para almacenar el flujo de bits en el archivo, en el que: para cada pista respectiva de la primera y la segunda pistas, todas las imágenes codificadas de cada muestra de la pista respectiva tienen el mismo valor del indicador de salida de imagen; y se permite que unas imágenes que tienen indicadores de salida de imagen iguales al primer valor se faciliten y se permite que unas imágenes que tienen indicadores de salida de imagen iguales al segundo valor se usen como imágenes de referencia, pero no se permite que se faciliten.

[0013] En otro aspecto, esta divulgación describe un dispositivo de vídeo que comprende: medios para obtener, a partir de un archivo, una primera caja de pista y una segunda caja de pista, conteniendo la primera caja de pista metadatos para una primera pista del archivo, conteniendo la segunda caja de pista metadatos para una segunda pista del archivo, en el que: cada una de la primera pista y la segunda pista comprenden una secuencia de muestras, siendo cada una de las muestras una unidad de acceso de vídeo de datos de vídeo multicapa, para cada pista respectiva de la primera pista y la segunda pista, todas las imágenes codificadas de cada muestra de la pista respectiva tienen el mismo valor de un indicador de salida de imagen, y se permite que unas imágenes que tienen indicadores

de salida de imagen iguales a un primer valor se faciliten y se permite que unas imágenes que tienen indicadores de salida de imagen iguales a un segundo valor se usen como imágenes de referencia, pero no se permite que se faciliten.

[0014] En otro aspecto, esta divulgación describe un medio de almacenamiento de datos legible por ordenador que tiene instrucciones almacenadas en el mismo que cuando se ejecutan hacen que uno o más procesadores: generen un archivo que comprende una caja de datos de medios que contiene un contenido de medios, comprendiendo el contenido de medios una secuencia de muestras, siendo cada una de las muestras una unidad de acceso de datos de vídeo multicapa, en el que para generar el archivo, las instrucciones hacen que el uno o más procesadores: como respuesta a una determinación de que al menos una unidad de acceso de un flujo de bits de los datos de vídeo multicapa incluye una imagen codificada que tiene un indicador de salida de imagen igual a un primer valor y una imagen codificada que tiene un indicador de salida de imagen igual a un segundo valor, use al menos una primera pista y una segunda pista para almacenar el flujo de bits en el archivo, en el que: para cada pista respectiva de la primera pista y la segunda pista, todas las imágenes codificadas de cada muestra de la pista respectiva tienen el mismo valor del indicador de salida de imagen; y se permite que unas imágenes que tienen indicadores de salida de imagen iguales al primer valor se faciliten y se permite que unas imágenes que tienen indicadores de salida de imagen igual al segundo valor se usen como imágenes de referencia, pero no se permite que se faciliten.

[0015] En otro aspecto, esta divulgación describe un medio de almacenamiento de datos legible por ordenador que tiene instrucciones almacenadas en el mismo que cuando se ejecutan hacen que uno o más procesadores: obtengan, a partir de un archivo, una primera caja de pista y una segunda caja de pista, conteniendo la primera caja de pista metadatos para una primera pista del archivo, conteniendo la segunda caja de pista metadatos para una segunda pista del archivo, en el que: cada una de la primera pista y la segunda pista comprende una secuencia de muestras, siendo cada una de las muestras una unidad de acceso de vídeo de datos de vídeo multicapa, para cada pista respectiva de la primera pista y la segunda pista, todas las imágenes codificadas de cada muestra de la pista respectiva tienen el mismo valor de un indicador de salida de imagen, y se permite que unas imágenes con indicadores de salida de imagen iguales a un primer valor se faciliten y se permite que unas imágenes con indicadores de salida de imagen iguales a un segundo valor se usen como imágenes de referencia, pero no se permite que se faciliten.

[0016] Los detalles de uno o más ejemplos de la divulgación se exponen en los dibujos adjuntos y en la descripción siguiente. Otras características, objetivos y ventajas resultarán evidentes a partir de la descripción, de los dibujos y de las reivindicaciones. La presente invención se expone en las reivindicaciones adjuntas.

BREVE DESCRIPCIÓN DE LOS DIBUJOS

[0017]

La figura 1 es un diagrama de bloques que ilustra un ejemplo de sistema de codificación y decodificación de vídeo que puede usar las técnicas descritas en esta divulgación.

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

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

La figura 4 es un diagrama de bloques que ilustra un ejemplo de conjunto de dispositivos que forman parte de una red.

La figura 5 es un diagrama conceptual que ilustra un ejemplo de estructura de un archivo, de acuerdo con una o más técnicas de esta divulgación.

La figura 6 es un diagrama conceptual que ilustra un ejemplo de estructura de un archivo, de acuerdo con una o más técnicas de esta divulgación.

La figura 7 es un diagrama de flujo que ilustra un ejemplo de operación de un dispositivo de generación de archivos, de acuerdo con una o más técnicas de esta divulgación.

La figura 8 es un diagrama de flujo que ilustra un ejemplo de operación en el que un dispositivo informático realiza un acceso aleatorio y/o un cambio de nivel, de acuerdo con una o más técnicas de esta divulgación.

La figura 9 es un diagrama de flujo que ilustra un ejemplo de operación de un dispositivo de generación de archivos, de acuerdo con una o más técnicas de esta divulgación.

La figura 10 es un diagrama de flujo que ilustra un ejemplo de operación de un dispositivo informático, de acuerdo con una o más técnicas de esta divulgación.

La figura 11 es un diagrama de flujo que ilustra un ejemplo de operación de un dispositivo de generación de archivos, de acuerdo con una o más técnicas de esta divulgación.

La figura 12 es un diagrama de flujo que ilustra un ejemplo de operación de un dispositivo de destino, de acuerdo con una o más técnicas de esta divulgación.

DESCRIPCIÓN DETALLADA

[0018] El formato base de archivos de medios ISO (ISOBMFF) es un formato de archivo para almacenar datos de medios. El ISOBMFF es ampliable para admitir el almacenamiento de datos de vídeo de conformidad con unas normas de codificación de vídeo particulares. Por ejemplo, el ISOBMFF se ha ampliado previamente para admitir el almacenamiento de datos de vídeo de conformidad con las normas de codificación de vídeo H.264/AVC y de codificación de vídeo de alta eficiencia (HEVC). Además, el ISOBMFF se ha ampliado previamente para admitir el almacenamiento de datos de vídeo de conformidad con las ampliaciones de codificación multivista (MVC) y de codificación de vídeo escalable (SVC) de la norma H.264/AVC. Las normas MV-HEVC, 3D-HEVC y SHVC son ampliaciones de la norma de codificación de vídeo HEVC que admiten datos de vídeo multicapa. Las características agregadas al ISOBMFF para el almacenamiento de datos de vídeo de conformidad con las ampliaciones MVC y SVC de H.264/AVC no son suficientes para el almacenamiento efectivo de datos de vídeo de conformidad con MV-HEVC, 3D-HEVC y SHVC. En otras palabras, pueden surgir diversos problemas si se intenta usar la ampliación del ISOBMFF para el almacenamiento de datos de vídeo de conformidad con las ampliaciones MVC y SVC de H.264/AVC para el almacenamiento de datos de vídeo de conformidad con MV-HEVC, 3D-HEVC y SHVC.

[0019] Por ejemplo, a diferencia de un flujo de bits de conformidad con las ampliaciones MVC o SVC de H.264/AVC, un flujo de bits de conformidad con MV-HEVC, 3D-HEVC o SHVC puede incluir unidades de acceso que contienen imágenes de punto de acceso aleatorio (IRAP) e imágenes no IRAP. Se puede usar una unidad de acceso que contiene imágenes IRAP e imágenes no IRAP para un acceso aleatorio en MV-HEVC, 3D-HEVC y SHVC. Sin embargo, el ISOBMFF y las ampliaciones existentes del mismo no proporcionan una forma de identificar dichas unidades de acceso. Esto puede dificultar la capacidad de un dispositivo informático para realizar el acceso aleatorio y el cambio de capa.

[0020] En consecuencia, de acuerdo con un ejemplo de esta divulgación, un dispositivo informático puede generar un archivo que comprende una caja de pista que contiene metadatos para una pista del archivo. Los datos de medios para la pista comprenden una secuencia de muestras. Cada una de las muestras puede ser una unidad de acceso de vídeo de datos de vídeo multicapa (por ejemplo, datos de vídeo MV-HEVC, 3D-HEVC o SHVC). Como parte de la generación del archivo, el dispositivo informático puede generar, en el archivo, una caja adicional que documenta todas las muestras que contienen al menos una imagen IRAP. La capacidad de determinar muestras que contienen imágenes IRAP basándose en información de la caja adicional puede permitir que los dispositivos informáticos que reciben el archivo realicen un acceso aleatorio y un cambio de capa sin analizar ni interpretar las unidades NAL. Esto puede reducir la complejidad y los tiempos de procesamiento.

[0021] Además, los datos de vídeo multicapa, como los datos de vídeo MV-HEVC, 3D-HEVC y SHVC, pueden incluir múltiples imágenes codificadas para cada unidad de acceso. Sin embargo, el ISOBMFF y las ampliaciones existentes del mismo no proporcionan información relativa a imágenes codificadas individuales dentro de una unidad de acceso cuando hay múltiples imágenes codificadas en la unidad de acceso. Por tanto, en los ejemplos donde un dispositivo informático (por ejemplo, un servidor de transmisión en tiempo real) determina si debe reenviar las unidades NAL en un archivo, el dispositivo informático puede tener que analizar e interpretar la información almacenada en las unidades NAL a fin de determinar si debe reenviar las unidades NAL. El análisis e interpretación de la información almacenada en las unidades NAL puede aumentar la complejidad del dispositivo informático y puede aumentar el retraso de transmisión.

[0022] En consecuencia, de acuerdo con un ejemplo de esta divulgación, un dispositivo informático puede generar un archivo que comprende una caja de pista que contiene metadatos para una pista del archivo. Los datos de medios para la pista comprenden una secuencia de muestras. Cada una de las muestras es una unidad de acceso de vídeo de los datos de vídeo multicapa. Como parte de la generación del archivo, el dispositivo informático genera, en el archivo, una caja de información de submuestra que contiene indicadores que especifican un tipo de información de submuestra proporcionada en la caja de información de submuestra. Cuando los indicadores tienen un valor particular, una submuestra correspondiente a la caja de información de submuestra contiene exactamente una imagen codificada y cero o más unidades NAL no de capa de codificación de vídeo (VCL) asociadas con la imagen codificada. De esta manera, los dispositivos informáticos que reciben el archivo pueden usar la información de submuestra proporcionada en la caja de información de submuestra para tomar determinaciones con respecto a imágenes codificadas individuales situadas dentro de una muestra del archivo. Las unidades no VCL NAL asociadas con la imagen codificada pueden incluir unidades NAL para conjuntos de parámetros (por ejemplo, PPS, SPS, VPS) y SEI aplicables a la imagen codificada.

[0023] En los datos de vídeo multicapa, una unidad de acceso puede incluir imágenes codificadas que están marcadas para salida e imágenes codificadas que están marcadas no para salida. Un descodificador de vídeo puede

usar imágenes codificadas que están marcadas no para salida como imágenes de referencia para descodificar imágenes que están marcadas para salida. Una cabecera de unidad NAL para una unidad NAL de un sector de una imagen puede incluir un indicador de salida de imagen (por ejemplo, `pic_output_flag` en HEVC) que indica si la imagen está marcada para salida. En un archivo ISOBMFF, se requiere que cada muestra esté asociada con un tiempo de salida (por ejemplo, un tiempo de composición) que indique cuándo se va a facilitar una muestra. Sin embargo, las imágenes que están marcadas no para salida no tienen tiempos de salida. Por tanto, la presencia de imágenes que están marcadas no para salida puede infringir este requisito de ISOBMFF o puede requerir técnicas alternativas no estándar.

[0024] En consecuencia, de acuerdo con una o más técnicas de esta divulgación, un dispositivo informático puede generar un archivo que comprende una caja de datos de medios que contiene contenido de medios. El contenido de medios comprende una secuencia de muestras. Cada una de las muestras comprende una unidad de acceso de los datos de vídeo multicapa. Como parte de la generación del archivo, el dispositivo informático, como respuesta a una determinación de que al menos una unidad de acceso de un flujo de bits de los datos de vídeo multicapa incluye una imagen codificada que tiene un indicador de salida de imagen igual a un primer valor (por ejemplo, 1) y una imagen codificada que tiene un indicador de salida de imagen igual a un segundo valor (por ejemplo, 0), puede usar al menos dos pistas para almacenar el flujo de bits en el archivo. Para cada pista respectiva de las al menos dos pistas, todas las imágenes codificadas de cada muestra de la pista respectiva tienen el mismo valor del indicador de salida de imagen. Las imágenes que tienen indicadores de salida de imagen iguales al primer valor (por ejemplo, 1) se pueden facilitar y las imágenes que tienen indicadores de salida de imagen iguales al segundo valor (por ejemplo, 0) se pueden usar como imágenes de referencia, pero no se pueden facilitar. El uso de al menos dos pistas puede resolver el problema descrito anteriormente, ya que a cada muestra de cada pista se le puede asignar un tiempo de salida adecuado, y un descodificador de vídeo no puede facilitar imágenes de la pista que contiene las muestras que no está permitido facilitar.

[0025] Si bien gran parte de la descripción de las técnicas de esta divulgación se refiere a las normas MV-HEVC, 3D-HEVC y SHVC, el lector apreciará que las técnicas de esta divulgación pueden ser aplicables a otras normas de codificación de vídeo y/o ampliaciones de las mismas.

[0026] La figura 1 es un diagrama de bloques que ilustra un ejemplo de sistema de codificación y descodificación de vídeo 10 que puede usar las técnicas descritas en esta divulgación. Como se muestra en la figura 1, 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 comprender cualquiera de entre una amplia gama de dispositivos, incluidos ordenadores de escritorio, ordenadores plegables (es decir, portátiles), ordenadores de tableta, descodificadores, teléfonos tales como los denominados teléfonos «inteligentes», los denominados paneles «inteligentes», televisores, cámaras, dispositivos de visualización, reproductores de medios digitales, consolas de videojuegos, dispositivos de transmisión de vídeo en tiempo real o similares. En algunos casos, el dispositivo de origen 12 y el dispositivo de destino 14 pueden estar equipados para la comunicación inalámbrica. El dispositivo de origen 12 y el dispositivo de destino 14 pueden considerarse dispositivos de vídeo.

[0027] En el ejemplo de la figura 1, el dispositivo de origen 12 incluye una fuente de vídeo 18, un codificador de vídeo 20 y una interfaz de salida 22. En algunos casos, la interfaz de salida 22 puede incluir un modulador/desmodulador (módem) y/o un transmisor. En el dispositivo de origen 12, la fuente de vídeo 18 puede incluir una fuente tal como un dispositivo de captación de vídeo, por ejemplo, una videocámara, un archivo de vídeo que contiene vídeo captado previamente, una interfaz de señal de vídeo en tiempo real para recibir vídeo desde un proveedor de contenido de vídeo y/o un sistema de gráficos por ordenador para generar datos de gráficos por ordenador como el vídeo de origen, o una combinación de dichas fuentes. Sin embargo, las técnicas descritas en esta divulgación pueden ser aplicables a la codificación de vídeo en general, y se pueden aplicar a aplicaciones inalámbricas y/o alámbricas.

[0028] El codificador de vídeo 20 puede codificar el vídeo captado, precaptado o generado por ordenador. El dispositivo de origen 12 puede transmitir los datos de vídeo codificados directamente a un dispositivo de destino 14 a través de una interfaz de salida 22 de un dispositivo de origen 12. Los datos de vídeo codificados se pueden almacenar de forma adicional (o alternativa) en el dispositivo de almacenamiento 33 para un posterior acceso por el dispositivo de destino 14 u otros dispositivos, para su descodificación y/o reproducción.

[0029] El dispositivo de destino 14 incluye una interfaz de entrada 28, un descodificador de vídeo 30 y un dispositivo de visualización 32. En algunos casos, la interfaz de entrada 28 puede incluir un receptor y/o un módem. La interfaz de entrada 28 del dispositivo de destino 14 recibe los datos de vídeo codificados a través del enlace 16. Los datos de vídeo codificados transmitidos a través del enlace 16, o proporcionados en el dispositivo de almacenamiento 33, 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 se pueden incluir con los datos de vídeo codificados transmitidos en un medio de comunicación, almacenar en un medio de almacenamiento o almacenar en un servidor de archivos.

[0030] El dispositivo de visualización 32 puede estar integrado con, o ser externo a, el dispositivo de destino 14. En algunos ejemplos, el dispositivo de destino 14 puede incluir un dispositivo de visualización integrado y también puede estar configurado para interactuar con un dispositivo de visualización externo. En otros ejemplos, el dispositivo de destino 14 puede ser un dispositivo de visualización. En general, el dispositivo de visualización 32 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.

[0031] Tanto el codificador de vídeo 20 como el descodificador de vídeo 30 pueden implementarse como cualquiera de entre una variedad de circuitos codificadores adecuados, tales como uno o más microprocesadores, procesadores de señales digitales (DSP), circuitos integrados específicos de la aplicación (ASIC), matrices de puertas programables *in situ* (FPGA), una 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 legible por ordenador no transitorio adecuado, y ejecutar las instrucciones en hardware usando uno o más procesadores para realizar las técnicas de esta divulgación. Tanto el codificador de vídeo 20 como el descodificador de vídeo 30 pueden estar incluidos en uno o más codificadores o descodificadores, cualesquiera de los cuales puede estar integrado como parte de un codificador/descodificador (CÓDEC) combinado en un dispositivo respectivo.

[0032] El dispositivo de destino 14 puede recibir los datos de vídeo codificados que se van a descodificar a través de 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 hasta el dispositivo de destino 14. En un ejemplo, el enlace 16 puede comprender un medio de comunicación para permitir que el dispositivo de origen 12 transmita datos de vídeo codificados directamente a un dispositivo de destino 14 en tiempo real. Los datos de vídeo codificados se pueden modular de acuerdo con una norma de comunicación, tal como un protocolo de comunicación inalámbrica, y transmitir al dispositivo de destino 14. El medio de comunicación puede comprender cualquier medio de comunicación, inalámbrica o alámbrica, tal como un espectro de radiofrecuencia (RF) o una o más líneas físicas de transmisión. El medio de comunicación puede formar parte de una red basada en paquetes, tal como una red de área local, una red de área amplia o una red global tal como Internet. El medio de comunicación puede incluir encaminadores, conmutadores, estaciones base o cualquier otro equipo que pueda ser útil para facilitar la comunicación desde el dispositivo de origen 12 hasta el dispositivo de destino 14.

[0033] De forma alternativa, la interfaz de salida 22 puede facilitar datos codificados a un dispositivo de almacenamiento 33. De manera similar, la interfaz de entrada 28 puede acceder al dispositivo de almacenamiento de datos codificados 33. El dispositivo de almacenamiento 33 puede incluir cualquiera de una variedad de medios de almacenamiento de datos de acceso distribuido o local, tales como una unidad de disco duro, unos discos Blu-ray, unos DVD, unos CD-ROM, una memoria flash, una memoria volátil o no volátil o cualquier otro medio de almacenamiento digital adecuado para almacenar datos de vídeo codificados. En otro ejemplo, el dispositivo de almacenamiento 33 puede corresponder a un servidor de archivos o a otro dispositivo de almacenamiento intermedio que puede contener el vídeo codificado generado por el dispositivo de origen 12. El dispositivo de destino 14 puede acceder a datos de vídeo almacenados en el dispositivo de almacenamiento 33 a través de transmisión en tiempo real 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. Los ejemplos de servidores de archivos incluyen un servidor web (por ejemplo, para un sitio web), un servidor FTP, dispositivos de almacenamiento conectados en red (NAS) o una unidad de disco local. El dispositivo de destino 14 puede acceder a los datos de vídeo codificados a través de cualquier conexión de datos estándar, incluida una conexión a Internet. Esto puede incluir un canal inalámbrico (por ejemplo, una conexión wifi), 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 33 puede ser una transmisión en tiempo real, una transmisión de descarga o una combinación de ambas.

[0034] Las técnicas de esta divulgación no están limitadas necesariamente a aplicaciones o configuraciones inalámbricas. Las técnicas se pueden aplicar a la codificación de vídeo como apoyo a cualquiera de una variedad de aplicaciones multimedia, tales como radiodifusiones de televisión por aire, transmisiones de televisión por cable, transmisiones de televisión por satélite, transmisiones de vídeo en tiempo real, por ejemplo, a través de 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 estar configurado para admitir una transmisión de vídeo unidireccional o bidireccional, a fin de admitir aplicaciones tales como la transmisión de vídeo en tiempo real, la reproducción de vídeo, la radiodifusión de vídeo y/o la videotelefonía.

[0035] Además, en el ejemplo de la figura 1, el sistema de codificación de vídeo 10 incluye un dispositivo de generación de archivos 34. El dispositivo de generación de archivos 34 puede recibir datos de vídeo codificados generados por el dispositivo de origen 12. El dispositivo de generación de archivos 34 puede generar un archivo que incluye los datos de vídeo codificados. El dispositivo de destino 14 puede recibir el archivo generado por el dispositivo de generación de archivos 34. En diversos ejemplos, el dispositivo de generación de archivos 34 puede incluir diversos tipos de dispositivos informáticos. Por ejemplo, el dispositivo de generación de archivos 34 puede comprender un

elemento de red sensible a los medios (MANE), un dispositivo informático de servidor, un dispositivo informático personal, un dispositivo informático de propósito especial, un dispositivo informático comercial u otro tipo de dispositivo informático. En algunos ejemplos, el dispositivo de generación de archivos 34 forma parte de una red de entrega de contenido. El dispositivo de destino 34 puede recibir datos de vídeo codificados desde el dispositivo de origen 12 a través de un canal, tal como un enlace 16. Además, el dispositivo de destino 14 puede recibir el archivo desde el dispositivo de generación de archivos 34 a través de un canal, tal como el enlace 16. El dispositivo de generación de archivos 34 puede considerarse un dispositivo de vídeo.

[0036] En otros ejemplos, el dispositivo de origen 12 u otro dispositivo informático puede generar un archivo que incluye los datos de vídeo codificados. Sin embargo, para facilitar la explicación, en esta divulgación se indica que el dispositivo de generación de archivos 34 genera el archivo. No obstante, debería entenderse que dichas descripciones son aplicables a los dispositivos informáticos en general.

[0037] El codificador de vídeo 20 y el descodificador de vídeo 30 pueden funcionar de acuerdo con una norma de compresión de vídeo, tal como la norma de codificación de vídeo de alta eficiencia (HEVC) o una ampliación de la misma. La norma HEVC también puede denominarse ISO/IEC 23008-2. Recientemente, el Equipo Colaborativo Conjunto sobre Codificación de Vídeo (JCT-VC) del Grupo de Expertos en Codificación de Vídeo (VCEG) de ITU-T y el Grupo de Expertos en Imágenes en Movimiento (MPEG) de ISO/IEC han terminado el diseño de la HEVC. El borrador de especificación HEVC más reciente, denominado en lo sucesivo HEVC WD en el presente documento, está disponible en http://phenix.int-evry.fr/jct/doc_end_user/documents/14_Vienna/wg11/JCTVC-N1003-v1.zip. El JCT-3V también está desarrollando la ampliación multivista para HEVC, en concreto, la MV-HEVC. Un reciente borrador de trabajo (WD) de MV-HEVC, titulado «MV-HEVC Draft Text 5 [Texto 5 de borrador de trabajo MV-HEVC]» y denominado en lo sucesivo MV-HEVC WD5 en el presente documento, está disponible en http://phenix.int-evry.fr/jct/doc_end_user/documents/14_Vienna/wg11/JCTVC-N1005-v3.zip. El JCT-VC también está elaborando la ampliación escalable de HEVC, denominada SHVC. Un reciente borrador de trabajo (WD) de SHVC, titulado «High efficiency video coding (HEVC) scalable extension draft 3 [Borrador 3 de ampliación escalable de codificación de vídeo de alta eficiencia (HEVC)]» y denominado en lo sucesivo SHVC WD3 en el presente documento, está disponible en http://phenix.int-evry.fr/jct/doc_end_user/documents/14_Vienna/wg11/JCTVC-N1008-v3.zip. Un reciente borrador de trabajo (WD) de la ampliación de rango de HEVC, está disponible en http://phenix.int-evry.fr/jct/doc_end_user/documents/14_Vienna/wg11/JCTVC-N1001-v3.zip. El codificador de vídeo 20 y el descodificador de vídeo 30 pueden funcionar de acuerdo con una o más de estas normas.

[0038] De forma alternativa, el codificador de vídeo 20 y el descodificador de vídeo 30 pueden funcionar de acuerdo con otras normas de propiedad o industriales, tales como la norma ITU-T H.264, denominada de forma alternativa MPEG-4, parte 10, codificación avanzada de vídeo (AVC), o ampliaciones de dichas normas. Sin embargo, las técnicas de esta divulgación no están limitadas a ninguna norma de codificación particular. Otros ejemplos normas de compresión de vídeo incluyen ITU-T H.261, ISO/IEC MPEG-1 Visual, ITU-T H.262 o ISO/IEC MPEG-2 Visual, ITU-T H.263, ISO/IEC MPEG-4 Visual e ITU-T H.264 (también conocida como ISO/IEC MPEG-4 AVC), incluidas sus ampliaciones de codificación de vídeo escalable (SVC) y de codificación de vídeo multivista (MVC).

[0039] Aunque no se muestra en la figura 1, en algunos aspectos, tanto el codificador de vídeo 20 como el descodificador de vídeo 30 pueden estar integrados con un codificador y descodificador de audio, y pueden incluir unidades MUX-DEMUX adecuadas, u otro tipo de hardware y software, para ocuparse de la codificación tanto de audio como de vídeo en un flujo de datos común o en flujos de datos separados. Si procede, en algunos ejemplos, las unidades MUX-DEMUX pueden ajustarse al protocolo de multiplexador ITU H.223 o a otros protocolos, tales como el protocolo de datagramas de usuario (UDP).

[0040] El equipo JCT-VC está trabajando en la elaboración de la norma HEVC. La iniciativa de normalización HEVC se basa en un modelo en evolución de un dispositivo de codificación de vídeo denominado modelo de prueba HEVC (HM). El HM supone varias capacidades adicionales de los dispositivos de codificación de vídeo respecto a los dispositivos existentes de acuerdo con, por ejemplo, la norma ITU-T H.264/AVC. Por ejemplo, mientras que la norma H.264/AVC proporciona nueve modos de codificación mediante intrapredicción, el HM puede proporcionar hasta treinta y tres modos de codificación mediante intrapredicción.

[0041] En general, el modelo de explotación del HM especifica que una trama o imagen de vídeo puede dividirse en una secuencia de bloques de árbol o unidades de codificación más grandes (LCU), que incluyen tanto muestras de luma como de croma. Los bloques de árbol también pueden denominarse «unidades de árbol de codificación» (CTU). Un bloque de árbol tiene un fin similar a un macrobloque de la norma H.264/AVC. Un sector incluye un número de bloques de árbol consecutivos en orden de codificación. Una trama o imagen de vídeo puede dividirse en uno o más sectores. Cada bloque de árbol puede dividirse en unidades de codificación (CU) de acuerdo con un árbol cuaternario. Por ejemplo, un bloque de árbol, como un nodo raíz del árbol cuaternario, puede dividirse en cuatro nodos hijo, y cada nodo hijo puede, a su vez, ser un nodo padre y dividirse en otros cuatro nodos hijo. Un nodo hijo final no dividido, como un nodo hoja del árbol cuaternario, comprende un nodo de codificación, es decir, un bloque de vídeo codificado.

Los datos sintácticos asociados a un flujo de bits codificado pueden definir un número máximo de veces que puede dividirse un bloque de árbol, y también pueden definir un tamaño mínimo de los nodos de codificación.

[0042] Una CU incluye un nodo de codificación y unidades de predicción (PU) y unidades de transformada (TU) asociadas al nodo de codificación. Un tamaño de la CU corresponde a un tamaño del nodo de codificación y debe ser de forma cuadrada. El tamaño de la CU puede variar desde 8x8 píxeles hasta el tamaño del bloque de árbol, con un máximo de 64x64 píxeles o más. Cada CU puede contener una o más PU y una o más TU. Los datos sintácticos asociados a una CU pueden describir, por ejemplo, la división de la CU en una o más PU. Los modos de división pueden diferir dependiendo de si la CU está codificada en modo de salto o directo, codificada en modo de intrapredicción o codificada en modo de interpredicción. Las PU pueden dividirse para tener forma no cuadrada. Los datos sintácticos asociados a una CU también pueden describir, por ejemplo, la división de la CU en una o más TU de acuerdo con un árbol cuaternario. Una TU puede tener forma cuadrada o no cuadrada.

[0043] La norma HEVC admite transformaciones de acuerdo con las TU, que pueden ser diferentes para diferentes CU. El tamaño de las TU se basa típicamente en el tamaño de las PU dentro de una CU definida dada para una LCU dividida, aunque puede que no sea siempre así. Las TU son típicamente del mismo tamaño o de un tamaño más pequeño que las PU. En algunos ejemplos, las muestras residuales correspondientes a una CU pueden subdividirse en unidades más pequeñas mediante una estructura de árbol cuaternario conocida como «árbol cuaternario residual» (RQT). Los nodos hoja del RQT pueden denominarse TU. Los valores de diferencia de píxel asociados a las TU pueden transformarse para generar coeficientes de transformada, que pueden cuantificarse.

[0044] En general, una PU incluye datos relacionados con el proceso de predicción. Por ejemplo, cuando la PU se codifica en modo intra, la PU puede incluir datos que describen un modo de intrapredicción para la PU. En otro ejemplo, cuando la PU se codifica en modo inter, la PU puede incluir datos que definen un vector de movimiento para la PU. Los datos que definen el vector de movimiento para una PU pueden describir, por ejemplo, una componente horizontal del vector de movimiento, una componente vertical del vector de movimiento, una resolución para el vector de movimiento (por ejemplo, precisión de un cuarto de píxel o precisión de un octavo de píxel), una imagen de referencia a la que apunta el vector de movimiento y/o una lista de imágenes de referencia (por ejemplo, lista 0, lista 1 o lista C) para el vector de movimiento.

[0045] En general, se usa una TU para los procesos de transformada y cuantificación. Una CU determinada que tiene una o más PU también puede incluir una o más unidades de transformada (TU). Tras la predicción, el codificador de vídeo 20 puede calcular valores residuales correspondientes a la PU. Los valores residuales comprenden valores de diferencia de píxel que se pueden transformar en coeficientes de transformada, cuantificar y explorar usando las TU, para generar coeficientes de transformada en serie para codificación de entropía. La presente divulgación usa típicamente el término «bloque de vídeo» para referirse a un nodo de codificación (es decir, un bloque de codificación) de una CU. En algunos casos específicos, esta divulgación también puede usar el término «bloque de vídeo» para referirse a un bloque de árbol, es decir, una LCU o una CU, que incluye un nodo de codificación y unas PU y TU.

[0046] 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 número de imágenes incluidas en el GOP. Cada sector de imagen puede incluir datos sintácticos de sector que describen un modo de codificación para el sector respectivo. El codificador de vídeo 20 actúa típicamente sobre bloques de vídeo de sectores de vídeo individuales con el fin de codificar los datos de vídeo. Un bloque de vídeo puede corresponder a un nodo de codificación de una CU. Los bloques de vídeo pueden tener tamaños fijos o variables y pueden diferir en tamaño de acuerdo con una norma de codificación especificada.

[0047] En un ejemplo, el HM admite predicción en diversos tamaños de PU. Si se supone que el tamaño de una CU particular es 2Nx2N, el HM admite intrapredicción en tamaños de PU de 2Nx2N o NxN e interpredicción en tamaños de PU simétricas de 2Nx2N, 2NxN, Nx2N o NxN. El HM también admite la división asimétrica para interpredicción en tamaños de PU de 2NxN_U, 2NxN_D, nLx2N y nRx2N. En la división asimétrica, una dirección de una CU no está dividida, mientras que la otra dirección está dividida en un 25 % y un 75 %. La parte de la CU correspondiente a la división del 25 % está indicada por una «n» seguida por una indicación de «arriba», «abajo», «izquierda» o «derecha». Así pues, por ejemplo, «2NxN_U» se refiere a una CU 2Nx2N que está dividida horizontalmente y tiene una PU 2Nx0,5N encima y una PU 2Nx1,5N debajo.

[0048] En la presente divulgación, «NxN» y «N por N» pueden usarse indistintamente para hacer referencia a las dimensiones de píxeles de un bloque de vídeo en términos de dimensiones verticales y horizontales, por ejemplo, 16x16 píxeles o 16 por 16 píxeles. En general, un bloque de 16x16 tiene 16 píxeles en una dirección vertical (y = 16) y 16 píxeles en una dirección horizontal (x = 16). Asimismo, un bloque de NxN tiene, en general, N píxeles en una dirección vertical y N píxeles en una dirección horizontal, donde N representa un valor entero no negativo. Los píxeles de un bloque se pueden disponer en filas y columnas. Además, no es necesario que los bloques tengan necesariamente el mismo número de píxeles en la dirección horizontal que en la dirección vertical. Por ejemplo, los bloques pueden comprender NxM píxeles, donde M no es necesariamente igual a N.

[0049] Tras la codificación intrapredictiva o interpredictiva usando las PU de una CU, el codificador de vídeo 20 puede calcular datos residuales para las TU de la CU. Las PU pueden comprender datos de píxeles en el dominio espacial (también denominado dominio del píxel) y las TU pueden comprender coeficientes en el dominio de la transformada tras la aplicación de una transformada, por ejemplo, una transformada discreta de coseno (DCT), una transformada de enteros, una transformada de *wavelet* o una transformada conceptualmente similar a los datos de vídeo residuales. Los datos residuales pueden corresponder a diferencias de píxeles entre píxeles de la imagen no codificada y unos valores de predicción correspondientes a las PU. El codificador de vídeo 20 puede formar las TU incluyendo los datos residuales para la CU y, a continuación, transformar las TU para generar coeficientes de transformada para la CU.

[0050] Tras cualquier transformada para generar coeficientes de transformada, el codificador de vídeo 20 puede realizar una cuantificación de los coeficientes de transformada. La cuantificación se refiere, en general, a un proceso en el que unos coeficientes de transformada se cuantifican para reducir posiblemente la cantidad de datos usados para representar los coeficientes, proporcionando compresión adicional. El proceso de cuantificación puede reducir la profundidad de bits asociada a algunos, o a la totalidad, de los coeficientes. Por ejemplo, un valor de n bits puede redondearse a la baja hasta un valor de m bits durante la cuantificación, donde n es mayor que m .

[0051] En algunos ejemplos, el codificador de vídeo 20 puede usar un orden de exploración predefinido para explorar los coeficientes de transformada cuantificados a fin de generar un vector en serie que se puede someter a codificación de entropía. En otros ejemplos, el codificador de vídeo 20 puede realizar una exploración adaptativa. Después de explorar los coeficientes de transformada cuantificados para formar un vector unidimensional, el codificador de vídeo 20 puede realizar la codificación de entropía del vector unidimensional, por ejemplo, de acuerdo con la codificación de longitud variable adaptativa al contexto (CAVLC), la codificación aritmética binaria adaptativa al contexto (CABAC), la codificación aritmética binaria adaptativa al contexto basada en la sintaxis (SBAC), la codificación de entropía de división de intervalo de probabilidad (PIPE) u otra metodología de codificación de entropía. El codificador de vídeo 20 también puede realizar la codificación de entropía de los elementos sintácticos asociados a los datos de vídeo codificados, para su uso por el descodificador de vídeo 30 en la descodificación de los datos de vídeo.

[0052] Para realizar la codificación CABAC, el codificador de vídeo 20 puede asignar un contexto de un modelo de contexto a un símbolo que se va a transmitir. El contexto se puede referir, por ejemplo, a si los valores contiguos del símbolo son distintos de cero o no. Para realizar la CAVLC, el codificador de vídeo 20 puede seleccionar un código de longitud variable para un símbolo que se va a transmitir. Las palabras de código en la codificación de longitud variable (VLC) se pueden construir de modo que los códigos relativamente más cortos correspondan a los símbolos más probables, mientras que los códigos más largos correspondan a los símbolos menos probables. De esta forma, el uso de la VLC puede lograr un ahorro en bits con respecto, por ejemplo, al uso de palabras de código de igual longitud para cada símbolo que se va a transmitir. La determinación de la probabilidad se puede basar en un contexto asignado al símbolo.

[0053] El codificador de vídeo 20 puede facilitar un flujo de bits que incluye una secuencia de bits que forman una representación de imágenes codificadas y datos asociados. El término «flujo de bits» puede ser un término colectivo usado para referirse a un flujo de unidades de capa de abstracción de red (NAL) (por ejemplo, una secuencia de unidades NAL) o un flujo de bytes (por ejemplo, una encapsulación de un flujo de unidades NAL que contiene unos prefijos de código de inicio y unas unidades NAL como se especifica en el Anexo B de la norma HEVC). Una unidad NAL es una estructura sintáctica que contiene una indicación del tipo de datos de la unidad NAL y los bytes que contienen dichos datos en forma de una carga útil de secuencia de bytes sin procesar (RBSP) entremezclados según sea necesario con bits de prevención de emulación. Cada una de las unidades NAL puede incluir una cabecera de unidad NAL y puede encapsular una RBSP. La cabecera de unidad NAL puede incluir un elemento sintáctico que indica un código de tipo de unidad NAL. El código de tipo de unidad NAL especificado por la cabecera de unidad NAL de una unidad NAL indica el tipo de la unidad NAL. Una RBSP puede ser una estructura sintáctica que contiene un número entero de bytes que está encapsulada dentro de una unidad NAL. En algunos casos, una RBSP incluye cero bits.

[0054] Diferentes tipos de unidades NAL pueden encapsular diferentes tipos de RBSP. Por ejemplo, un primer tipo de unidad NAL puede encapsular una RBSP para un conjunto de parámetros de imagen (PPS), un segundo tipo de unidad NAL puede encapsular una RBSP para un segmento de sector, un tercer tipo de unidad NAL puede encapsular una RBSP para información de mejora complementaria (SEI), y así sucesivamente. Las unidades NAL que encapsulan RBSP para datos de codificación de vídeo (a diferencia de las RBSP para conjuntos de parámetros y mensajes SEI) pueden denominarse unidades NAL de capa de codificación de vídeo (VCL). Las unidades NAL que contienen conjuntos de parámetros (por ejemplo, conjuntos de parámetros de vídeo (VPS), conjuntos de parámetros de secuencia (SPS), PPS, etc.) pueden denominarse unidades NAL de conjuntos de parámetros.

[0055] Esta divulgación puede referirse a una unidad NAL que encapsula una RBSP para un sector de segmento como una unidad NAL de sector codificado. Tal como se define en el HEVC WD, un segmento de sector es un número entero de CTU ordenadas consecutivamente en una exploración en mosaico y contenidas en una sola unidad NAL. En cambio, en el HEVC WD, un sector puede ser un número entero de CTU contenidas en un segmento de sector independiente y todos los segmentos de sector dependientes subsiguientes (si los hubiera) que preceden al siguiente

segmento de sector independiente (si lo hubiera) dentro de la misma unidad de acceso. Un segmento de sector independiente es un segmento de sector para el cual los valores de los elementos sintácticos de la cabecera del segmento de sector no se deducen a partir de los valores para un segmento de sector anterior. Un segmento de sector dependiente es un segmento de sector para el cual los valores de algunos elementos sintácticos de la cabecera del segmento de sector se deducen a partir de los valores para el segmento de sector independiente anterior en orden de descodificación. Una RBSP de una unidad NAL de sector codificado puede incluir una cabecera de segmento de sector y datos de sector. Una cabecera de segmento de sector es una parte de un segmento de sector codificado que contiene los elementos de datos pertenecientes a la primera o todas las CTU representadas en el segmento de sector. Una cabecera de sector es una cabecera de segmento de sector del segmento de sector independiente que es un segmento de sector actual o el segmento de sector independiente más reciente que precede a un segmento de sector dependiente actual en orden de descodificación.

[0056] Un VPS es una estructura sintáctica que comprende elementos sintácticos que se aplican a cero o más secuencias de vídeo codificadas completas (CVS). Un SPS también es una estructura sintáctica que comprende elementos sintácticos que se aplican a cero o más CVS completas. Un SPS puede incluir un elemento sintáctico que identifica un VPS que está activo cuando el SPS está activo. Por tanto, los elementos sintácticos de un VPS pueden ser más aplicables en general que los elementos sintácticos de un SPS.

[0057] Un conjunto de parámetros (por ejemplo, un VPS, SPS, PPS, etc.) puede contener una identificación a la que se hace referencia, de forma directa o indirecta, desde una cabecera de sector de un sector. El proceso de referencia se conoce como «activación». Por tanto, cuando el descodificador de vídeo 30 está descodificando un sector particular, se dice que un conjunto de parámetros, al que un elemento sintáctico de una cabecera de sector del sector particular hace referencia de forma directa o indirecta, se «activa». Dependiendo del tipo de conjunto de parámetros, la activación puede producirse en imágenes o en secuencias. Por ejemplo, una cabecera de sector de un sector puede incluir un elemento sintáctico que identifica un PPS. Por tanto, cuando un codificador de vídeo codifica el sector, el PPS se puede activar. Además, el PPS puede incluir un elemento sintáctico que identifica un SPS. Por tanto, cuando se activa un PPS que identifica el SPS, el SPS se puede activar. El SPS puede incluir un elemento sintáctico que identifica un VPS. Por tanto, cuando se activa un SPS que identifica el VPS, el VPS se activa.

[0058] El descodificador de vídeo 30 puede recibir un flujo de bits generado por el codificador de vídeo 20. Además, el descodificador de vídeo 30 puede analizar el flujo de bits para obtener elementos sintácticos a partir del flujo de bits. El descodificador de vídeo 30 puede reconstruir las imágenes de los datos de vídeo basándose, al menos en parte, en los elementos sintácticos obtenidos a partir del flujo de bits. El proceso para reconstruir los datos de vídeo puede ser, en general, recíproco al proceso realizado por el codificador de vídeo 20. Por ejemplo, el descodificador de vídeo 30 puede usar vectores de movimiento de las PU para determinar bloques predictivos para las PU de una CU actual. Además, el descodificador de vídeo 30 puede realizar una cuantificación inversa de los bloques de coeficientes de unas TU de la CU actual. El descodificador de vídeo 30 puede realizar transformadas inversas en los bloques de coeficientes para reconstruir los bloques de transformada de las TU de la CU actual. El descodificador de vídeo 30 puede reconstruir los bloques de codificación de la CU actual añadiendo las muestras de los bloques predictivos para las PU de la CU actual a las muestras correspondientes de los bloques de transformada de las TU de la CU actual. Mediante la reconstrucción de los bloques de codificación para cada CU de una imagen, el descodificador de vídeo 30 puede reconstruir la imagen.

[0059] En el HEVC WD, una CVS puede comenzar a partir de una imagen de regeneración instantánea de descodificación (IDR) o una imagen de acceso de enlace roto (BLA), o una imagen de acceso aleatorio limpio (CRA) que es la primera imagen del flujo de bits, incluidas todas las imágenes subsiguientes que no son imágenes IDR o BLA. Una imagen IDR solo contiene sectores I (es decir, sectores en los que solo se usa intrapredicción). Una imagen BLA puede ser la primera imagen del flujo de bits en orden de descodificación, o puede aparecer más adelante en el flujo de bits. Cada imagen IDR es la primera imagen de una CVS en orden de descodificación. En el HEVC WD, una imagen IDR puede ser una imagen de punto de acceso aleatorio interno (IRAP) para la cual cada unidad VCL NAL tiene un nal_unit_type igual a IDR_W_RADL o IDR_N_LP.

[0060] Las imágenes IDR pueden usarse para acceso aleatorio. Sin embargo, las imágenes que siguen a una imagen IDR en orden de descodificación no pueden usar imágenes descodificadas antes de la imagen IDR como referencia. En consecuencia, los flujos de bits que dependen de imágenes IDR para acceso aleatorio pueden tener una eficiencia de codificación significativamente menor que los flujos de bits que usan tipos adicionales de imágenes de acceso aleatorio. En al menos algunos ejemplos, una unidad de acceso IDR es una unidad de acceso que contiene una imagen IDR.

[0061] El concepto de imágenes CRA se introdujo en la HEVC para permitir que las imágenes que siguen a una imagen CRA en orden de descodificación, pero preceden a la imagen CRA en orden de salida, usen imágenes descodificadas antes de la imagen CRA como referencia. Las imágenes que siguen a una imagen CRA en orden de descodificación, pero preceden a la imagen CRA en orden de salida, se denominan imágenes iniciales asociadas con la imagen CRA (o imágenes iniciales de la imagen CRA). Es decir, para mejorar la eficiencia de codificación, el concepto de imágenes de CRA se introdujo en la HEVC para permitir que las imágenes que siguen a una imagen CRA en orden de descodificación, pero preceden a la imagen CRA en orden de salida, usen imágenes descodificadas antes

de la imagen CRA como referencia. Una unidad de acceso CRA es una unidad de acceso en la que la imagen codificada es una imagen CRA. En el HEVC WD, una imagen CRA es una imagen de acceso aleatorio interno para la cual cada unidad VCL NAL tiene un nal_unit_type igual a CRA_NUT.

- 5 **[0062]** Las imágenes iniciales de una imagen CRA se pueden descodificar correctamente si la descodificación comienza a partir de una imagen IDR o una imagen CRA que aparece antes de la imagen CRA en orden de descodificación. Sin embargo, las imágenes iniciales de una imagen CRA pueden no ser descodificables cuando se produce un acceso aleatorio desde la imagen CRA. En consecuencia, un descodificador de vídeo típicamente descodifica las imágenes iniciales de una imagen CRA durante la descodificación de acceso aleatorio. Para evitar la propagación de errores por imágenes de referencia que pueden no estar disponibles dependiendo de dónde comience la descodificación, ninguna imagen que siga a una imagen CRA tanto en orden de descodificación como en orden de salida puede usar ninguna imagen que preceda a la imagen CRA en orden de descodificación u orden de salida (lo cual incluye las imágenes iniciales) como referencia.
- 10
- 15 **[0063]** El concepto de una imagen BLA se introdujo en HEVC después de la introducción de las imágenes CRA, y se basa en el concepto de las imágenes CRA. Una imagen BLA se origina típicamente a partir de un empalme de flujo de bits en la posición de una imagen CRA y, en el flujo de bits empalmado, la imagen CRA del punto de empalme se convierte en una imagen BLA. Por tanto, las imágenes BLA pueden ser imágenes CRA en los flujos de bits originales, y el empalmador de flujos de bits cambia una imagen CRA para convertirla en una imagen BLA después del empalme de flujos de bits en la ubicación de la imagen CRA. En algunos casos, una unidad de acceso que contiene una imagen RAP puede denominarse unidad de acceso RAP en el presente documento. Una unidad de acceso BLA es una unidad de acceso que contiene una imagen BLA. En el HEVC WD, una imagen BLA puede ser una imagen de acceso aleatorio interno para la cual cada unidad VCL NAL tiene un nal_unit_type igual a BLA_W_LP, BLA_W_RADL o BLA_N_LP.
- 20
- 25 **[0064]** En general, una imagen IRAP contiene solo sectores I, y puede ser una imagen BLA, una imagen CRA o una imagen IDR. Por ejemplo, el HEVC WD indica que una imagen IRAP puede ser una imagen codificada para la cual cada unidad VCL NAL tiene un nal_unit_type en el intervalo de BLA_W_LP a RSV_IRAP_VCL23, inclusive. Además, el HEVC WD indica que la primera imagen del flujo de bits en orden de descodificación debe ser una imagen IRAP. La tabla 7-1 de HEVC WD muestra los códigos de tipo de unidad NAL y las clases de tipo de unidad NAL. La tabla 7-1 de HEVC WD se reproduce a continuación.
- 30

Tabla 7-1 - Códigos de tipo de unidad NAL y clases de tipo de unidad NAL

nal_unit_type	Nombre de nal_unit_type	Contenido de unidad NAL y estructura sintáctica de RBSP	Clase de tipo de unidad NAL
0	TRAIL_N	Segmento de sector codificado de una slice_segment_layer_rbsp() de imagen posterior no TSA ni STSA	VCL
1	TRAIL_R		
2	TSA_N	Segmento de sector codificado de una slice_segment_layer_rbsp() de imagen TSA	VCL
3	TSA_R		
4	STSA_N	Segmento de sector codificado de una slice_segment_layer_rbsp() de imagen STSA	VCL
5	STSA_R		
6	RADL_N	Segmento de sector codificado de una slice_segment_layer_rbsp() de imagen RADL	VCL
7	RADL_R		
8	RASL_N	Segmento de sector codificado de una slice_segment_layer_rbsp() de imagen RASL	VCL
9	RASL_R		
10	RSV_VCL_N10	Tipos de unidades VCL NAL no de referencia de subcapa no IRAP reservados	VCL
12	RSV_VCL_N12		

nal_unit_type	Nombre de nal_unit_type	Contenido de unidad NAL y estructura sintáctica de RBSP	Clase de tipo de unidad NAL
14	RSV_VCL_N14		
11	RSV_VCL_R11	Tipos de unidades VCL NAL de referencia de subcapa no IRAP reservados	VCL
13	RSV_VCL_R13		
15	RSV_VCL_R15		
16	BLA_W_LP	Segmento de sector codificado de una slice_segment_layer_rbs() de imagen BLA	VCL
17	BLA_W_RADL		
18	BLA_N_LP		
19	IDR_W_RADL	Segmento de sector codificado de una slice_segment_layer_rbsp() de imagen IDR	VCL
20	IDR_N_LP		
21	CRA_NUT	Segmento de sector codificado de una slice_segment_layer_rbsp() de imagen CRA	VCL
22	RSV_IRAP_VCL22	Tipos de unidades VCL NAL de IRAP reservados	VCL
23	RSV_IRAP_VCL23		
24...31	RSV_VCL24..RSV_VCL31	Tipos de unidades VCL NAL no de IRAP reservados	VCL
32	VPS_NUT	video_parameter_set_rbsp() de conjunto de parámetros de vídeo	no VCL
33	SPS_NUT	seq_parameter_set_rbsp() de conjunto de parámetros de secuencia	no VCL
34	PPS_NUT	pic_parameter_set_rbsp() de conjunto de parámetros de imagen	no VCL
35	AUD_NUT	access_unit_delimiter_rbsp() de delimitador de unidad de acceso	no VCL
36	EOS_NUT	end_of_seq_rbsp() de fin de secuencia	no VCL
37	EOB_NUT	end_of_bitstream_rbsp() de fin de flujo de bits	no VCL
38	FD_NUT	filler_data_rbsp() de datos de relleno	no VCL
39	PREFIX_SEI_NUT	sei_rbsp() de información de mejora complementaria	no VCL
40	SUFFIX_SEI_NUT		
41...47	RSV_NVCL41..RSV_NVCL47	Reservado	no VCL
48...63	UNSPEC48...UNSPEC63	Sin especificar	no VCL

[0065] Una diferencia entre las imágenes BLA y las imágenes CRA es la siguiente. Para una imagen CRA, las imágenes iniciales asociadas pueden descodificarse correctamente si la descodificación comienza desde una imagen RAP antes de la imagen CRA en orden de descodificación. Sin embargo, las imágenes iniciales asociadas con una

imagen CRA pueden no descodificarse correctamente cuando se produce un acceso aleatorio desde la imagen CRA (es decir, cuando la descodificación comienza por la imagen CRA, o en otras palabras, cuando la imagen CRA es la primera imagen del flujo de bits). En cambio, puede que no haya un contexto en el que las imágenes iniciales asociadas con una imagen BLA sean descodificables, incluso cuando la descodificación comienza por una imagen RAP antes de la imagen BLA en orden de descodificación.

[0066] Algunas de las imágenes iniciales asociadas con una imagen CRA particular o una imagen BLA particular pueden descodificarse correctamente incluso cuando la imagen CRA particular o la imagen BLA particular es la primera imagen de un flujo de bits. Estas imágenes iniciales pueden denominarse imágenes iniciales descodificables (DLP) o imágenes iniciales descodificables de acceso aleatorio (RADL). En el HEVC WD, una imagen RADL puede ser una imagen codificada para la cual cada unidad VCL NAL tiene un `nal_unit_type` igual a `RADL_R` o `RADL_N`. Además, el HEVC WD indica que todas las imágenes RADL son imágenes iniciales y que las imágenes RADL no se usan como imágenes de referencia para el proceso de descodificación de imágenes posteriores de la misma imagen IRAP asociada. Cuando están presentes, todas las imágenes RADL preceden, en orden de descodificación, a todas las imágenes posteriores de la misma imagen IRAP asociada. El HEVC WD indica que una unidad de acceso RADL puede ser una unidad de acceso en la que la imagen codificada es una imagen RADL. Una imagen posterior puede ser una imagen que sigue a la imagen IRAP asociada (es decir, la imagen IRAP anterior en orden de descodificación) en orden de salida.

[0067] Otras imágenes iniciales pueden denominarse imágenes iniciales no descodificables (NLP) o imágenes iniciales omitidas de acceso aleatorio (RASL). En el HEVC WD, una imagen RASL puede ser una imagen codificada para la cual cada unidad VCL NAL tiene un `nal_unit_type` igual a `RASL_R` o `RASL_N`. Todas las imágenes RASL son imágenes iniciales de una imagen BLA o CRA asociada.

[0068] Siempre que los conjuntos de parámetros necesarios estén disponibles cuando deban activarse, una imagen IRAP y todas las imágenes no RASL subsiguientes en orden de descodificación pueden descodificarse correctamente sin realizar el proceso de descodificación de ninguna imagen que preceda a la imagen IRAP en orden de descodificación. Es posible que haya imágenes en un flujo de bits que contengan solo sectores I que no son imágenes IRAP.

[0069] En la codificación multivista, puede haber múltiples vistas de la misma escena desde diferentes puntos de vista. El término «unidad de acceso» puede usarse para referirse al conjunto de imágenes que corresponden al mismo instante de tiempo. Por tanto, unos datos de vídeo pueden conceptualizarse como una serie de unidades de acceso que aparecen a lo largo del tiempo. Una «componente de vista» puede ser una representación codificada de una vista en una única unidad de acceso. En esta divulgación, una «vista» puede referirse a una secuencia o un conjunto de componentes de vista asociadas al mismo identificador de vista. Una componente de vista puede contener una componente de vista de textura y una componente de vista de profundidad. En esta divulgación, una «vista» puede referirse a un conjunto o una secuencia de una o más componentes de vistas asociadas al mismo identificador de vista.

[0070] Una componente de vista de textura (es decir, una imagen de textura) puede ser una representación codificada de la textura de una vista de una única unidad de acceso. Una vista de textura puede ser una secuencia de componentes de vista de textura asociadas con un valor idéntico de índice de orden de vista. Un índice de orden de vista de una vista puede indicar una posición de cámara de la vista con respecto a otras vistas. Una componente vista de profundidad (es decir, una imagen de profundidad) puede ser una representación codificada de la profundidad de una vista de una única unidad de acceso. Una vista de profundidad puede ser un conjunto o una secuencia de componentes de vista de profundidad asociadas con un valor idéntico de índice de orden de vistas.

[0071] En MV-HEVC, 3D-HEVC y SHVC, un codificador de vídeo puede generar un flujo de bits que comprende una serie de unidades NAL. Diferentes unidades NAL del flujo de bits pueden estar asociadas a diferentes capas del flujo de bits. Una capa se puede definir como un conjunto de unidades VCL NAL y unidades asociadas no VCL NAL que tienen el mismo identificador de capa. Una capa puede ser equivalente a una vista en la codificación de vídeo multivista. En la codificación de vídeo multivista, una capa puede contener todas las componentes de vista de la misma capa con diferentes instantes de tiempo. Cada componente de vista puede ser una imagen codificada de la escena de vídeo que pertenece a una vista específica en un instante de tiempo específico. En algunos ejemplos de codificación de vídeo 3D, una capa puede contener todas las imágenes de profundidad codificadas de una vista específica o unas imágenes de textura codificadas de una vista específica. En otros ejemplos de codificación de vídeo 3D, una capa puede contener tanto componentes de vista de textura como componentes de vista de profundidad de una vista específica. De forma similar, en el contexto de la codificación de vídeo escalable, una capa corresponde típicamente a unas imágenes codificadas que tienen características de vídeo diferentes a unas imágenes codificadas de otras capas. Dichas características de vídeo típicamente incluyen la resolución espacial y el nivel de calidad (por ejemplo, la relación señal-ruido). En HEVC y en sus ampliaciones, la escalabilidad temporal se puede lograr dentro de una capa definiendo un grupo de imágenes con un nivel temporal particular como una subcapa.

[0072] Para cada capa respectiva del flujo de bits, unos datos de una capa inferior pueden descodificarse sin referencia a unos datos de cualquier capa superior. En la codificación de vídeo escalable, por ejemplo, unos datos de

una capa base pueden descodificarse sin referencia a unos datos de una capa de mejora. En general, las unidades NAL solo pueden encapsular datos de una única capa. Por tanto, las unidades NAL que encapsulan datos de la capa restante más alta del flujo de bits pueden eliminarse del flujo de bits sin afectar a la descodificabilidad de los datos de las capas restantes del flujo de bits. En la codificación multivista y 3D-HEVC, las capas superiores pueden incluir componentes de vista adicionales. En SHVC, las capas superiores pueden incluir datos de mejora de relación señal-ruido (SNR), datos de mejora espacial y/o datos de mejora temporal. En MV-HEVC, 3D-HEVC y SHVC, una capa puede denominarse «capa base» si un descodificador de vídeo puede descodificar imágenes de la capa sin hacer referencia a datos de cualquier otra capa. La capa base puede ajustarse a la especificación base de la HEVC (por ejemplo, el HEVC WD).

[0073] En SVC, las capas distintas de la capa base pueden denominarse «capas de mejora» y pueden proporcionar información que mejora la calidad visual de los datos de vídeo descodificados a partir del flujo de bits. La SVC puede mejorar la resolución espacial, la relación señal-ruido (es decir, la calidad) o la velocidad temporal. En la codificación de vídeo escalable (por ejemplo, SHVC), una «representación de capa» puede ser una representación codificada de una capa espacial de una única unidad de acceso. Para facilitar la explicación, esta divulgación puede referirse a componentes de vista y/o representaciones de capa como «componentes de vista/representaciones de capa» o simplemente «imágenes».

[0074] Para implementar las capas, las cabeceras de las unidades NAL pueden incluir unos elementos sintácticos `nuh_reserved_zero_6bits`. En el HEVC WD, el elemento sintáctico `nuh_reserved_zero_6bits` está reservado. Sin embargo, en MV-HEVC, 3D-HEVC y SVC, el elemento sintáctico `nuh_reserved_zero_6bits` se conoce como elemento sintáctico `nuh_layer_id`. El elemento sintáctico `nuh_layer_id` especifica un identificador de una capa. Las unidades NAL de un flujo de bits que tienen elementos sintácticos de ID de capa `nuh` que especifican valores diferentes pertenecen a diferentes capas del flujo de bits.

[0075] En algunos ejemplos, el elemento sintáctico de ID de capa `nuh` de una unidad NAL es igual a 0 si la unidad NAL se refiere a una capa base en codificación multivista (por ejemplo MV-HEVC), codificación 3DV (por ejemplo, 3D-HEVC) o codificación de vídeo escalable (por ejemplo, SHVC). Los datos de una capa de base de un flujo de bits pueden descodificarse sin referencia a los datos de ninguna otra capa del flujo de bits. Si una unidad NAL no se refiere a una capa base en codificación multivista, 3DV o codificación de vídeo escalable, el elemento sintáctico de ID de capa `nuh` de la unidad NAL puede tener un valor distinto de cero.

[0076] Además, algunas componentes de vista/representaciones de capa de una capa pueden descodificarse sin referencia a otras componentes de vista/representaciones de capa de la misma capa. Por tanto, las unidades NAL que encapsulan datos de ciertos componentes de vista/representaciones de capa pueden eliminarse del flujo de bits sin afectar a la descodificabilidad de otras componentes de vista/representaciones de capa de la capa. La eliminación de unidades NAL que encapsulan datos de dichas componentes de vista/representaciones de capa puede reducir la velocidad de tramas del flujo de bits. Un subconjunto de componentes de vista/representaciones de capa de una capa que puede descodificarse sin referencia a otras componentes de vista/representaciones de capa dentro de la capa se puede denominar en el presente documento «subcapa» o «subcapa temporal».

[0077] Las unidades NAL pueden incluir elementos sintácticos de ID temporal que especifican identificadores temporales (es decir, `TemporalId`) de las unidades NAL. El identificador temporal de una unidad NAL identifica una subcapa a la que pertenece la unidad NAL. Por tanto, cada subcapa de un flujo de bits puede tener un identificador temporal diferente. En general, si el identificador temporal de una primera unidad NAL de una capa es menor que el identificador temporal de una segunda unidad NAL de la misma capa, los datos encapsulados por la primera unidad NAL pueden descodificarse sin referencia a los datos encapsulados por la segunda unidad NAL.

[0078] Un flujo de bits puede estar asociado a una pluralidad de puntos de operación. Cada punto de operación de un flujo de bits está asociado con un conjunto de identificadores de capa (es decir, un conjunto de valores `nuh_layer_id`) y un identificador temporal. El conjunto de identificadores de capa se puede denotar como `OpLayerIdSet` y el identificador temporal se puede denotar como `TemporalId`. Si el identificador de capa de una unidad NAL se encuentra en un conjunto de identificadores de capa de un punto de operación y el identificador temporal de la unidad NAL es menor o igual que el identificador temporal del punto de operación, la unidad NAL está asociada al punto de operación. Por tanto, un punto de operación puede corresponder a un subconjunto de unidades NAL del flujo de bits.

[0079] Tal como se ha indicado anteriormente, esta divulgación se refiere al almacenamiento de contenido de vídeo en un archivo basándose en el formato base de archivo de medios ISO (ISOBMFF). En particular, esta divulgación describe diversas técnicas para el almacenamiento de flujos de vídeo que contienen múltiples capas codificadas, en el que cada capa puede ser una capa escalable, una vista de textura, una vista de profundidad u otros tipos de capas o vistas. Las técnicas de esta divulgación pueden aplicarse, por ejemplo, al almacenamiento de datos de vídeo MV-HEVC, datos de vídeo SHVC, datos de vídeo 3D-HEVC y/u otros tipos de datos de vídeo.

[0080] Los formatos de archivo y las normas de formato de archivo se analizarán brevemente a continuación. Las normas de formato de archivo incluyen el formato base de archivo de medios ISO (ISOBMFF, ISO/IEC 14496-12, en lo sucesivo, «ISO/IEC 14996-12») y otras normas de formato de archivo derivados de ISOBMFF, incluido 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, en lo sucesivo «ISO/IEC 14496-15»). Por tanto, la norma ISO/IEC 14496-12 especifica el formato base de archivo de medios ISO. Otros documentos amplían el formato base de archivo de medios ISO para aplicaciones específicas. Por ejemplo, la norma ISO/IEC 14496-15 describe el transporte de vídeo con estructura de unidad NAL en el formato base de archivo de medios ISO. Las normas H.264/AVC y HEVC, así como sus ampliaciones, son ejemplos de vídeo con estructura de unidad NAL. La norma ISO/IEC 14496-15 incluye secciones que describen el transporte de unidades H.264/AVC NAL. Adicionalmente, la sección 8 de ISO/IEC 14496-15 describe el transporte de unidades HEVC NAL.

[0081] El ISOBMFF se usa como la base para muchos formatos de encapsulación de códec, como el formato de archivo AVC, así como para muchos formatos de contenedor multimedia, como el formato de archivo MPEG-4, el formato de archivo 3GPP (3GP) y el formato de archivo DVB. Además de los medios continuos, como el audio y el vídeo, los medios estáticos, como las imágenes, así como los metadatos, se pueden almacenar en un archivo de conformidad con el ISOBMFF. Los archivos estructurados de acuerdo con el ISOBMFF se pueden usar para muchos fines, incluida la reproducción local de archivos de medios, la descarga progresiva de un archivo remoto, segmentos para la transmisión en tiempo real adaptativa y dinámica a través de HTTP (DASH), contenedores para contenido que se va a transmitir y sus instrucciones de empaquetado y grabación de flujos de medios recibidos en tiempo real. Por tanto, aunque originalmente se diseñó para el almacenamiento, el ISOBMFF ha demostrado ser valioso para la transmisión en tiempo real, por ejemplo, para descargas progresivas o DASH. Para fines de transmisión en tiempo real, se pueden usar los segmentos de película definidos en ISOBMFF.

[0082] Un archivo de conformidad con el formato de archivo HEVC puede comprender una serie de objetos, denominados cajas. Una caja puede ser un elemento básico orientado a objetos definido por un identificador de tipo exclusivo y una longitud. Por ejemplo, una caja puede ser la estructura sintáctica elemental en el ISOBMFF, que incluye un tipo de caja codificada de cuatro caracteres, un recuento de bytes de la caja y una carga útil. En otras palabras, una caja puede ser una estructura sintáctica que comprende un tipo de caja codificada, un recuento de bytes de la caja y una carga útil. En algunos casos, todos los datos de un archivo de conformidad con el formato de archivo HEVC pueden estar contenidos dentro de cajas y no puede haber datos en el archivo que no estén en una caja. Por tanto, un archivo ISOBMFF puede consistir en una secuencia de cajas, y las cajas pueden contener otras cajas. Por ejemplo, la carga útil de una caja puede incluir una o más cajas adicionales. La figura 5 y la figura 6, descritas en detalle en otra parte de esta divulgación, muestran ejemplos de cajas de un archivo, de acuerdo con una o más técnicas de esta divulgación.

[0083] Un archivo de conformidad con el ISOBMFF puede incluir diversos tipos de cajas. Por ejemplo, un archivo de conformidad con el ISOBMFF puede incluir una caja de tipo de archivo, una caja de datos de medios, una caja de película, una caja de fragmento de película, etc. En este ejemplo, una caja de tipo de archivo incluye el tipo de archivo e información de compatibilidad. Una caja de datos de medios puede contener muestras (por ejemplo, imágenes codificadas). Una caja de película («moov») contiene metadatos para flujos de medios continuos presentes en el archivo. Cada uno de los flujos de medios continuos se puede representar en el archivo como una pista. Por ejemplo, una caja de película puede contener metadatos relacionados con una película (por ejemplo, relaciones lógicas y de tiempo entre muestras, y también punteros a ubicaciones de muestras). Las cajas de películas pueden incluir varios tipos de subcajas. Las subcajas de una caja de película pueden incluir una o más cajas de pista. Una caja de pista puede incluir información sobre una pista individual de una película. Una caja de pista puede incluir una caja de cabecera de pista que especifica información general de una sola pista. Además, una caja de pista puede incluir una caja de medios que contiene una caja de información de medios. La caja de información de medios puede incluir una caja de tabla de muestras que contiene una indexación de datos de muestras de medios de la pista. La información de la caja de tabla de muestras se puede usar para localizar muestras en el tiempo y, para cada una de las muestras de la pista, un tipo, tamaño, contenedor y desplazamiento en ese contenedor de la muestra. Por tanto, los metadatos para una pista están contenidos en una caja de pista («trak»), mientras que el contenido de medios de una pista está contenido en una caja de datos de medios («mdat») o directamente en un archivo separado. El contenido de medios para unas pistas comprende (es decir, consiste en) una secuencia de muestras, como unas unidades de acceso de audio o vídeo.

[0084] El ISOBMFF especifica los siguientes tipos de pistas: una pista de medios, que contiene un flujo de medios elemental, una pista de notas, que incluye instrucciones de transmisión de medios o representa un flujo de paquetes recibidos, y una pista de metadatos temporizados, que comprende metadatos sincronizados en el tiempo. Los metadatos para cada pista incluyen una lista de entradas de descripción de muestra, cada una de las cuales proporciona el formato de codificación o encapsulación usado 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 muestra de la pista.

[0085] El ISOBMFF permite especificar metadatos específicos de muestra con diversos mecanismos. Las cajas específicas de la caja de tabla de muestras («stbl») se han normalizado para responder a unas necesidades comunes. Por ejemplo, una caja de muestras de sincronización («stss») es una caja situada dentro de una caja de tabla de muestras. La caja de muestras de sincronización se usa para enumerar las muestras de acceso aleatorio de la pista. Esta divulgación puede referirse a una muestra enumerada por la caja de muestras de sincronización como muestra de sincronización. En otro ejemplo, un mecanismo de agrupación de muestras permite la correlación de muestras de

acuerdo con un tipo de agrupación de cuatro caracteres con grupos de muestras que comparten la misma propiedad especificada como una entrada de descripción de grupo de muestras en el archivo. Se han especificado varios tipos de agrupación en el ISOBMFF.

[0086] Una caja de tabla de muestras puede incluir una o más cajas SampleToGroup y una o más cajas de descripción de grupo de muestras (es decir, cajas SampleGroupDescription). Se puede usar una caja SampleToGroup para determinar un grupo de muestras al que pertenece una muestra, junto con una descripción asociada del grupo de muestras. En otras palabras, una caja SampleToGroup puede indicar un grupo al que pertenece una muestra. Una caja SampleToGroup puede tener un tipo de caja de «sbgp». Una caja SampleToGroup puede incluir un elemento de tipo de agrupación (por ejemplo, grouping_type). El elemento de tipo de agrupación puede ser un número entero que identifica un tipo (es decir, un criterio usado para formar los grupos de muestras) de una agrupación de muestras. Además, una caja SampleToGroup puede incluir una o más entradas. Cada entrada de una caja SampleToGroup puede asociarse con una serie diferente, no superpuesta de muestras consecutivas de la pista. Cada entrada puede indicar un elemento de recuento de muestras (por ejemplo, sample_count) y un elemento de índice de descripción de grupo (por ejemplo, group_description_index). El elemento de recuento de muestras de una entrada puede indicar un número de muestras asociadas con la entrada. En otras palabras, el elemento de recuento de muestras de la entrada puede ser un número entero que proporciona el número de muestras consecutivas con el mismo descriptor de grupo de muestras. El elemento de índice de descripción de grupo puede identificar una caja SampleGroupDescription que contiene una descripción de las muestras asociadas con la entrada. Los elementos de índice de descripción de grupo de múltiples entradas pueden identificar la misma caja SampleGroupDescription.

[0087] Los diseños de formato de archivo actuales pueden tener uno o más problemas. Para almacenar contenido de vídeo de un códec de vídeo particular basándose en el ISOBMFF, puede ser necesaria una especificación de formato de archivo para ese códec de vídeo. Para el almacenamiento de flujos de vídeo que contienen múltiples capas, como MV-HEVC y SHVC, es posible reusar algunos de los conceptos del formato de archivo SVC y MVC. Sin embargo, muchas partes no se pueden usar directamente para flujos de vídeo SHVC y MV-HEVC. La aplicación directa del formato de archivo HEVC presenta al menos las siguientes deficiencias: Los flujos de bits SHVC y MV-HEVC pueden comenzar con una unidad de acceso que contiene una imagen IRAP en la capa base, pero también puede contener otras imágenes no IRAP en otras capas o viceversa. La muestra de sincronización actualmente no permite la indicación de dicho punto para acceso aleatorio.

[0088] Esta divulgación describe soluciones potenciales a los problemas anteriores, y también proporciona otras mejoras potenciales, para permitir un almacenamiento eficiente y flexible de flujos de vídeo que contienen múltiples capas. Las técnicas descritas en esta divulgación se aplican potencialmente a cualquier formato de archivo para almacenar dicho contenido de vídeo codificado por cualquier códec de vídeo, aunque la descripción es específica para el almacenamiento de flujos de vídeo SHVC y MV-HEVC basándose en el formato de archivo HEVC, que se especifica en cláusula 8 de ISO/IEC 14496-15.

[0089] La implementación detallada de las técnicas de esta divulgación se analizará en detalle a continuación. Las técnicas de esta divulgación se pueden resumir en los siguientes ejemplos. Los siguientes ejemplos pueden usarse por separado. De forma alternativa, se pueden usar conjuntamente diversas combinaciones de los siguientes ejemplos.

[0090] En un primer ejemplo, Compressorname es un valor especificado en una caja VisualSampleEntry. Como se describe en la sección 8.5.2.1 de ISO/IEC 14496-12, una caja VisualSampleEntry es un tipo de caja de tabla de muestras para pistas de vídeo que almacena información detallada sobre el tipo de codificación usado y cualquier información de inicialización necesaria para esa codificación. Compressorname indica el nombre de un compresor usado para generar datos de medios. Un descodificador de vídeo puede usar el valor de Compressorname para determinar cómo y/o si se deben descodificar los datos de vídeo del archivo. Como se define en la sección 8.5.3 de ISO/IEC 14496-12, Compressorname se formatea como un campo fijo de 32 bytes, con el primer byte establecido en el número de bytes que se va a visualizar, seguido de ese número de bytes de datos visualizables y, a continuación, de un relleno hasta completar un total de 32 bytes (incluido el byte de tamaño).

[0091] El primer ejemplo permite dos nuevos valores de Compressorname. El primer valor nuevo de Compressorname es «013SHVC Coding», para el archivo que contiene flujos de vídeo SHVC. El segundo nuevo valor de Compressorname es «016MV-HEVC Coding» para el archivo que contiene flujos de vídeo MV-HEVC. Este primer ejemplo se puede implementar como se muestra en las secciones 9.5.3.1.3 y 10.5.3.2, proporcionadas más adelante.

[0092] Como se ha descrito de forma breve anteriormente, un archivo puede incluir una caja de película que contiene metadatos para las pistas del archivo. La caja de película puede incluir una caja de pista para cada pista del archivo. Además, una caja de pista puede incluir una caja de información de medios que contiene todos los objetos que declaran información característica de los medios de la pista. La caja de información de medios puede incluir una caja de tabla de muestras. La caja de tabla de muestras puede especificar metadatos específicos de muestra. Por ejemplo, la caja de tabla de muestras puede incluir una pluralidad de cajas de descripción de muestra. Cada una de las cajas de descripción de muestra puede ser un caso de entrada de muestra. En ISO/IEC 14496-12, los casos de clase VisualSampleEntry se pueden usar como entradas de muestra. Las clases de entradas de muestra específicas para

normas de codificación de vídeo particulares pueden ampliar la clase VisualSampleEntry. Por ejemplo, una clase de entradas de muestra específica para HEVC puede ampliar la clase VisualSampleEntry. Por tanto, esta divulgación puede referirse a las diferentes clases que amplían la clase VisualSampleEntry como diferentes tipos de entradas de muestra.

[0093] En un segundo ejemplo, se definen dos nuevos tipos de entrada de muestra (es decir, de «muestra»), «hev2» y «hvc2», para las pistas HEVC. Los dos nuevos tipos de entradas de muestra permiten el uso de agregadores y extractores. En general, un agregador agrega múltiples unidades NAL en forma de una sola unidad de datos agregada. Por ejemplo, un agregador puede contener múltiples unidades NAL y/o puede concatenar virtualmente múltiples unidades NAL. En general, un extractor indica un tipo de datos obtenidos a partir de otras pistas. Por ejemplo, el almacenamiento de datos de medios (por ejemplo, datos HEVC) en múltiples pistas puede dar como resultado archivos compactos, ya que se puede evitar la duplicación de datos haciendo referencia a los datos a través de pistas de medios usando unidades de datos relativamente pequeñas denominadas extractores, que están insertadas como una unidad NAL en los datos de medios. Este segundo ejemplo puede implementarse como se muestra en las secciones 9.5.3.1.1, 9.5.3.1.2, 9.5.4, 9.5.6, 10.4.5, 10.5.3.1.1.1 y 10.5.3.2 proporcionadas más adelante.

[0094] En un tercer ejemplo, la definición de una entrada de muestra que se asocia con un requisito particular en el almacenamiento de conjuntos de parámetros para un flujo de bits multicapa se modifica a fin de permitir el acceso aleatorio conveniente a una capa particular o un punto de operación particular. Por ejemplo, cuando una pista SHVC, MV-HEVC o 3D-HEVC tiene la entrada de muestra y cuando una muestra contiene al menos una imagen IRAP, todos los parámetros necesarios para descodificar esa muestra se incluirán en la entrada de muestra o en esa propia muestra. En este ejemplo, cuando una muestra no contiene ninguna imagen IRAP, todos los conjuntos de parámetros (por ejemplo, VPS, SPS, PPS) necesarios para descodificar esa muestra se incluirán en la entrada de muestra o en cualquiera de las muestras desde la muestra anterior que contiene al menos una imagen IRAP hasta esa propia muestra, inclusive. Este tercer ejemplo puede implementarse como se muestra en la sección 9.5.3.1.1 siguiente.

[0095] En una versión alternativa del tercer ejemplo, cuando una pista SHVC, MV-HEVC o 3D-HEVC tiene la entrada de muestra y cuando una imagen de una muestra es una imagen IRAP, todos los conjuntos de parámetros necesarios para descodificar esa imagen se incluirán en la entrada de muestra o en esa propia muestra. Además, en esta alternativa, cuando la muestra no contiene ninguna imagen IRAP, todos los conjuntos de parámetros necesarios para descodificar la imagen se incluirán en la entrada de muestra o en cualquiera de las muestras que siguen a la muestra anterior que contiene al menos una imagen IRAP en la misma capa hasta esa propia muestra, inclusive.

[0096] En un cuarto ejemplo, se definen los siguientes casos para los tipos de entrada de muestra existentes. En este ejemplo, las muestras que pertenecen a los tipos de entrada de muestra «hev1» y «hvc1» contienen las configuraciones HEVC, SHVC y MV-HEVC para las pistas SHVC y MV-HEVC con unidades HEVC VCL NAL. Además, se definen los tipos de entrada de muestra «hev1» y «hvc1» que contienen configuraciones SHVC y MV-HEVC para las pistas SHVC y MV-HEVC sin unidades HEAL NAL pero con unidades VCL NAL con nuh_layer_id mayor que 0, cuando los extractores no están permitidos. Este cuarto ejemplo se puede implementar como se muestra en la sección 9.5.3.1.1 siguiente.

[0097] En un quinto ejemplo, una muestra de sincronización de una pista SHVC, MV-HEVC o 3D-HEVC se define como una muestra que contiene imágenes que son todas imágenes IRAP. Este quinto ejemplo se puede implementar como se muestra en las secciones 9.5.5 y 10.4.3 siguientes. Como se especifica en la sección 9.5.5 siguiente, se considera que una muestra SHVC es una muestra de sincronización si cada imagen codificada de la unidad de acceso es una imagen IRAP, tal como se define en el HEVC WD. Además, como se especifica en la sección 10.4.3 siguiente, una muestra MV-HEVC se considera una muestra de sincronización si cada imagen codificada de la unidad de acceso es una imagen IRAP sin imágenes RASL, como se define en el HEVC WD.

[0098] Por tanto, en el quinto ejemplo, como parte de la generación de un archivo, el dispositivo de generación de archivos 34 puede generar una caja de muestras de sincronización que incluye una tabla de muestras de sincronización que documenta las muestras de sincronización de una pista de datos de vídeo multicapa. Cada muestra de sincronización de la pista es una muestra de acceso aleatorio de la pista. Una muestra de codificación de vídeo escalable es una muestra de sincronización si cada imagen codificada de una unidad de acceso es una imagen IRAP. Una muestra de codificación de vídeo multivista es una muestra de sincronización si cada imagen codificada de la unidad de acceso es una imagen IRAP sin imágenes RASL.

[0099] En una versión alternativa del quinto ejemplo, una muestra de sincronización de una pista SHVC, MV-HEVC o 3D-HEVC se define como una muestra que contiene imágenes que son todas imágenes IRAP sin imágenes RASL. La tabla de muestras de sincronización documenta las muestras de sincronización. Opcionalmente, un grupo de muestras de muestras de sincronización documenta las muestras de sincronización. En otras palabras, un grupo de muestras de muestras de sincronización incluye información que identifica muestras de sincronización.

[0100] En un sexto ejemplo, se define un grupo de muestras «rap» para contener las muestras que contienen imágenes que son todas imágenes IRAP (con o sin imágenes RASL). Este sexto ejemplo se puede implementar como se muestra en la sección 9.5.5 siguiente. De forma alternativa, en el sexto ejemplo, se define el grupo de muestras

«rap» para contener las muestras que contienen imágenes que son todas imágenes IRAP, pero excluir las muestras que se indican como muestras de sincronización.

[0101] En un séptimo ejemplo, se define un nuevo grupo de muestras o una nueva caja que documentan todas las muestras que contienen al menos una imagen IRAP, el tipo de unidad NAL de las unidades VCL NAL de las imágenes IRAP de la muestra, tanto si todas las imágenes codificadas de la muestra son imágenes IRAP, como si no, el número de imágenes IRAP de la muestra y los valores de ID de capa de estas imágenes IRAP de la muestra.

[0102] Por tanto, en este séptimo ejemplo, el dispositivo de generación de archivos 34 puede generar un archivo que comprende una caja de pista que contiene metadatos para una pista del archivo. Los datos de medios para la pista comprenden una secuencia de muestras. Cada una de las muestras puede ser una unidad de acceso de datos de vídeo multicapa. Como parte de la generación del archivo, el dispositivo de generación de archivos 34 genera, en el archivo, una caja adicional que documenta todas las muestras que contienen al menos una imagen IRAP.

[0103] Este séptimo ejemplo puede implementarse en parte como se muestra en la sección 9.5.5.1 siguiente. Como se muestra en la sección 9.5.5.1 siguiente, una clase de entradas de muestra accesibles aleatoriamente amplía una clase VisualSampleGroupEntry. Los ejemplos de clase de entradas de muestra accesibles aleatoriamente (es decir, cajas de entradas de muestra accesibles aleatoriamente) corresponden a muestras que contienen al menos una imagen IRAP. Además, una caja de entradas de muestra accesibles aleatoriamente incluye un valor `all_pics_are_IRAP` que especifica si todas las imágenes codificadas de la muestra correspondiente son imágenes IRAP.

[0104] Por tanto, en el séptimo ejemplo, el dispositivo de generación de archivos 34 puede generar una entrada de muestra que incluye un valor (por ejemplo, `all_pics_are_IRAP`). Si el valor es igual a 1, especifica que cada imagen codificada de una muestra es una imagen IRAP. Si el valor es igual a 0, especifica que no todas las imágenes codificadas de la muestra son imágenes IRAP.

[0105] Además, de acuerdo con el séptimo ejemplo, cuando no todas las imágenes codificadas de una muestra son imágenes IRAP, el dispositivo de generación de archivos 34 puede incluir, en la entrada de muestra correspondiente a la muestra, un valor que indica un número de imágenes IRAP de cada muestra del grupo de muestras. Adicionalmente, cuando no todas las imágenes codificadas de la muestra son imágenes IRAP, el dispositivo de generación de archivos 34 puede incluir, en la entrada de muestra correspondiente a la muestra, valores que indican los identificadores de capa de las imágenes IRAP de la muestra.

[0106] De forma alternativa, en el séptimo ejemplo, el nuevo grupo de muestras o la nueva caja documenta dichas muestras, pero excluye las que están indicadas como muestras de sincronización o miembros del grupo de muestras «rap».

[0107] Este séptimo ejemplo puede resolver uno o más problemas que pueden surgir cuando los datos de vídeo multicapa se almacenan usando el ISOBMFF o las ampliaciones existentes del mismo. Por ejemplo, en la codificación de vídeo de una sola capa, típicamente solo hay una imagen codificada por unidad de acceso. Sin embargo, en la codificación de vídeo multicapa, típicamente hay más de una imagen codificada por unidad de acceso. El ISOBMFF y las ampliaciones existentes del mismo no proporcionan una forma de indicar qué muestras incluyen una o más imágenes IRAP. Esto puede dificultar la capacidad de un dispositivo informático de localizar puntos de acceso aleatorio en un archivo o de realizar un cambio de capa. Por ejemplo, en ausencia de información que indica cuál de las muestras contiene una o más imágenes IRAP, el dispositivo informático puede necesitar analizar e interpretar unidades NAL a fin de determinar si se puede usar una unidad de acceso como punto de acceso aleatorio y/o para un cambio de capa. El análisis y la interpretación de unidades NAL puede agregar complejidad al dispositivo informático y puede consumir tiempo y recursos de procesamiento. Además, algunos dispositivos informáticos que realizan un acceso aleatorio y/o cambio de capa, como los servidores de transmisión en tiempo real, no están configurados para analizar ni interpretar unidades NAL.

[0108] En un octavo ejemplo, se incluye la introducción de un nuevo tipo de submuestra, donde cada submuestra contiene una imagen codificada y sus unidades asociadas no VCL NAL. Este octavo ejemplo puede implementarse como se muestra en la sección 9.5.8 siguiente. Por tanto, en este octavo ejemplo, el dispositivo de generación de archivos 34 puede generar un archivo que comprende una caja de pista que contiene metadatos para una pista del archivo. Los datos de medios para la pista comprenden una secuencia de muestras. Cada una de las muestras es una unidad de acceso de datos de vídeo multicapa. Como parte de la generación del archivo, el dispositivo de generación de archivos 34 genera, en el archivo, una caja de información de submuestra que contiene indicadores que especifican un tipo de información de submuestra que se proporciona en la caja de información de submuestra. Cuando los indicadores tienen un valor particular, una submuestra correspondiente a la caja de información de submuestra contiene exactamente una imagen codificada y cero o más unidades no VCL NAL asociadas con la imagen codificada.

[0109] Este octavo ejemplo puede resolver uno o más problemas que pueden surgir cuando los datos de vídeo multicapa se almacenan usando el ISOBMFF o las ampliaciones existentes del mismo. Por ejemplo, en la codificación de vídeo multicapa, puede haber múltiples imágenes codificadas por muestra. Por ejemplo, puede haber una o más imágenes en la muestra para cada capa. Sin embargo, en la ampliación de ISOBMFF para H.264/AVC y HEVC, la

caja de información de submuestra no proporciona información sobre imágenes individuales de una muestra cuando la muestra incluye múltiples imágenes. Las técnicas de este octavo ejemplo pueden resolver este problema proporcionando un nuevo tipo de caja de información de submuestra que proporciona información sobre una submuestra que contiene solo una imagen codificada y unidades no VCL NAL asociadas con la imagen codificada.

Proporcionar información sobre una imagen codificada individual en la estructura del archivo, en lugar de solo proporcionar dicha información dentro de las unidades NAL asociadas con la imagen codificada, puede permitir que un dispositivo informático determine la información sobre la imagen codificada sin tener que interpretar las unidades NAL. En algunos casos, para disminuir la complejidad del dispositivo informático y/o aumentar el rendimiento del dispositivo informático, el dispositivo informático no está configurado para interpretar las unidades NAL. En algunos ejemplos en los que un dispositivo informático está transmitiendo unidades NAL almacenadas en un archivo, el dispositivo informático puede usar la información de la caja de información de submuestra para determinar si se deben enviar las unidades NAL de la submuestra a un dispositivo cliente.

[0110] Un noveno ejemplo se refiere al tratamiento de muestras no de salida en un contexto multicapa. Particularmente, en el noveno ejemplo, cuando una unidad de acceso contiene algunas imágenes codificadas que tienen `pic_output_flag` igual a 1 y otras imágenes codificadas que tienen `pic_output_flag` igual a 0, se deben usar al menos dos pistas para almacenar el flujo, de modo que dentro de cada pista todas las imágenes codificadas de cada muestra tienen el mismo valor de `pic_output_flag`. Este noveno ejemplo puede implementarse como se muestra en la sección 9.5.9 siguiente.

[0111] Por tanto, en este noveno ejemplo, el dispositivo de generación de archivos 34 puede generar un archivo que comprende una caja de datos de medios que contiene contenido de medios. El contenido de medios comprende una secuencia de muestras. Cada una de las muestras es una unidad de acceso de los datos de vídeo multicapa. Como respuesta a una determinación de que al menos una unidad de acceso de un flujo de bits de los datos de vídeo multicapa incluye una imagen codificada que tiene un indicador de salida de imagen (por ejemplo, `pic_output_flag`) igual a 1 y una imagen codificada que tiene un indicador de salida de imagen igual a 0, el dispositivo de generación de archivos 34 puede usar al menos dos pistas para almacenar el flujo de bits en el archivo. Para cada pista respectiva de las al menos dos pistas, todas las imágenes codificadas de cada muestra de la pista respectiva tienen el mismo valor del indicador de salida de imagen.

[0112] Este noveno ejemplo puede resolver uno o más problemas que pueden surgir cuando se almacenan datos de vídeo multicapa usando el ISO/BMFF o las ampliaciones existentes del mismo. Por ejemplo, si se usara una sola pista para almacenar imágenes codificadas que tuvieran indicadores de salida de imagen iguales a 0 e indicadores de salida de imagen iguales a 1, se infringirían diversas reglas de formateo de archivos. Por ejemplo, las reglas de formateo de archivos típicamente requieren que solo haya una muestra en una pista por instante de tiempo. Si una sola pista almacenara imágenes codificadas que tuvieran indicadores de salida de imagen iguales a 0 e indicadores de salida de imagen iguales a 1, habría múltiples muestras en una pista por instante de tiempo. Forzar que las imágenes codificadas que tienen diferentes valores del indicador de salida de imagen estén en diferentes pistas de un archivo puede resolver este problema.

[0113] A continuación, se describe un ejemplo de implementación de algunas técnicas de esta divulgación. El ejemplo de implementación que se describe a continuación se basa en la última especificación integrada de 14496-15 en el documento de salida MPEG W13478. Los cambios al Anexo A (mostrados mediante subrayado) y las secciones añadidas (Sección 9 para SHVC y Sección 10 para MV-HEVC) se incluyen más adelante. En otras palabras, los ejemplos particulares de esta divulgación pueden modificar el Anexo A de ISO/IEC 14496-15 y pueden añadir las secciones 9 y/o 10 a ISO/IEC 14496-15. El texto que se muestra con subrayado y doble subrayado puede ser de particular relevancia para los ejemplos de esta divulgación. Aunque el término SHVC se usa en diversos lugares de los ejemplos descritos en el presente documento, el diseño de esta divulgación en realidad no es solo para admitir el códec SHVC, sino que también se pueden admitir todos los códec multicapa, incluidos MV-HEVC y 3D-HEVC, a menos que se indique explícitamente lo contrario.

9 flujos elementales y definiciones de muestras SHVC

9.1 Introducción

[0114] Esta cláusula especifica el formato de almacenamiento de datos SHVC. Amplía las definiciones del formato de almacenamiento de HEVC de la cláusula 8.

[0115] El formato de archivo para el almacenamiento de contenido SHVC, definido en esta cláusula y los Anexos A a D, usa las capacidades existentes del formato base de archivo de medios ISO y el formato de archivo HEVC simple (es decir, el formato de archivo especificado en la cláusula 8). Además, se usan las siguientes estructuras o ampliaciones, entre otras, para admitir características específicas de SHVC.

[0116] Agregador:

estructura para permitir una agrupación escalable eficiente de unidades NAL transformando patrones irregulares de unidades NAL en patrones regulares de unidades de datos agregadas.

[0117] Extractor:

estructura para permitir la extracción eficiente de unidades NAL de otras pistas que no son la que contiene los datos de medios.

[0118] Sentencias de metadatos temporales:

estructuras para almacenar información alineada en el tiempo de muestras de medios.

[0119] Compatibilidad HEVC:

disposición para almacenar un flujo de bits SHVC de una manera compatible con HEVC, de modo que la capa base compatible con HEVC pueda ser usada por cualquier lector compatible con el formato de archivo HEVC simple.

9.2 Estructura de flujo elemental

[0120] Los flujos SHVC se almacenan de acuerdo con 8.2, con la siguiente definición de flujo elemental de vídeo SHVC:

- **Los flujos elementales de vídeo SHVC** contendrán todas las unidades NAL relacionadas con la codificación de vídeo (es decir, las unidades NAL que contengan datos de vídeo o una estructura de vídeo de señalización) y podrán contener unidades NAL no relacionadas con la codificación de vídeo, tales como mensajes SEI y unidades NAL delimitadoras de unidades de acceso. También pueden estar presentes agregadores (véase A.2) o extractores (véase A.3). Los agregadores y extractores se procesarán como se define en esta norma internacional (por ejemplo, no se colocarán directamente en la memoria intermedia de salida al acceder al archivo). Otras unidades NAL que no están expresamente prohibidas pueden estar presentes, y si no se reconocen deberían ignorarse (por ejemplo, no se colocan en la memoria intermedia de salida al acceder al archivo).

[0121] Los flujos SHVC no se almacenarán usando flujos de conjuntos de parámetros asociados.

[0122] Puede haber unidades VCL NAL con nuh_layer_id igual a 0, unidades VCL NAL con nuh_layer_id mayor que 0 y unidades no VCL NAL presentes en un flujo elemental de vídeo SHVC. Adicionalmente, puede haber unidades NAL de agregador y unidades NAL de extractor presentes en un flujo elemental de vídeo SHVC.

9.3 Uso de formato de archivo HEVC simple

[0123] El formato de archivo SHVC es una ampliación del formato de archivo HEVC simple definido en la cláusula 8.

9.4 Definiciones de muestra y configuración

9.4.1 Introducción

[0124] Muestra SHVC: una muestra SHVC también es una unidad de acceso como la definida en el Anexo H de ISO/IEC 23008-2.

9.4.2 Orden canónico y restricciones

9.4.2.1 Restricciones

[0125] Las siguientes restricciones se aplican a los datos SHVC además de los requisitos de 8.3.2.

- **Unidades VCL NAL:** Todas las unidades VCL NAL de una unidad de acceso estarán contenidas en la muestra cuyo tiempo de composición sea el de la imagen representada por la unidad de acceso. Una muestra SHVC contendrá al menos una unidad VCL NAL.
- **Agregadores/extractores:** El orden de todas las unidades NAL incluidas en un agregador o a las que hace referencia un extractor es exactamente el orden de descodificación que tendrían si estas unidades NAL estuvieran presentes en una muestra que no contuviera agregadores/extractores. Después de procesar el agregador o el extractor, todas las unidades NAL deben estar en un orden de descodificación válido como se especifica en ISO/IEC 23008-2.

9.4.2.2 Registro de configuración de descodificador

[0126] Cuando el registro de configuración de descodificador definido en 8.3.3.1 se usa para un flujo que puede interpretarse como un flujo SHVC o HEVC, el registro de configuración de descodificador HEVC reflejará las propiedades de la capa base compatible con HEVC, por ejemplo, contendrá solo conjuntos de parámetros necesarios para descodificar la capa base HEVC.

[0127] El SHVCDecoderConfigurationRecord es estructuralmente idéntico a un HEVCDecoderConfigurationRecord. La sintaxis es la siguiente:

```
aligned(8) class SHVCDecoderConfigurationRecord {
    // same fields as in HEVCDecoderConfigurationRecord syntax
    structure
}
```

[0128] La semántica de los campos de SHVCDecoderConfigurationRecord es la misma que la definida para un HEVCDecoderConfigurationRecord.

9.5 Obtención a partir del formato base de archivo de medios ISO

9.5.1 Estructura de pista SHVC

[0129] Una secuencia de vídeo escalable se representa mediante una o más pistas de vídeo en un archivo. Cada pista representa uno o más puntos operativos del flujo escalable. Un flujo escalable puede, por supuesto, reducirse aún más, si se desea.

[0130] Se va a suponer que el punto de operación más bajo es el uno de la totalidad de los puntos operativos que contiene unidades NAL con nuh_layer_id igual a 0 solo y TemporalId igual a 0 solo. Una pista que contiene el punto operativo más bajo se designará como la «pista escalable base». Todas las demás pistas que forman parte de la misma información codificada escalable se vincularán a esta pista base mediante una referencia de pista del tipo «sbas» (base escalable).

[0131] Todas las pistas que comparten la misma pista base escalable deben compartir la misma escala de tiempo que la pista base escalable.

9.5.2 Uso compartido y extracción de datos

[0132] Diferentes pistas pueden compartir datos lógicamente. Este uso compartido puede adoptar una de las dos formas siguientes:

- a) Los datos de muestra se copian de una pista a otra pista (y posiblemente se compactan o se vuelven a intercalar con otros datos, como audio). Esto crea archivos generales más grandes, pero los datos de baja velocidad de bits se pueden compactar y/o intercalar con otro material, para facilitar la extracción.
- b) Puede haber instrucciones sobre cómo realizar esta copia en el momento en que se lee el archivo.

[0133] Para el segundo caso, se usan extractores (definidos en A.3).

9.5.3 Definición de flujo de vídeo SHVC

9.5.3.1 Nombre y formato de entrada de muestra

9.5.3.1.1 Definición

[0134]

Tipos: «hvc2», «hev2», «shc1», «shv1», «shcC»

Contenedor: Caja de descripción de muestra («std») «stsd»

Obligatorio: Es obligatoria una entrada de muestra «hvc1», «hev1», «hvc2», «hev2», «shc1» o «shv1»

Cantidad: Pueden estar presentes una o más entradas de muestra

[0135] Cuando el nombre de entrada de muestra es «shc1», el valor predeterminado y obligatorio de array_completeness es 1 para matrices de todos los tipos de conjuntos de parámetros, y 0 para todas las demás matrices. Cuando el nombre de entrada de muestra es «shv1», el valor predeterminado del grado de compleción de la matriz es 0 para todas las matrices.

[0136] Cuando el nombre de entrada de muestra es «shv1», se aplica lo siguiente:

- Si una muestra contiene al menos una imagen IRAP como la definida en ISO/IEC 23008-2, todos los conjuntos de parámetros necesarios para descodificar esa muestra se incluirán en la entrada de muestra o en esa propia muestra.
- De lo contrario (la muestra no contiene ninguna imagen IRAP), todos los conjuntos de parámetros necesarios para descodificar esa muestra se incluirán en la entrada de la muestra o en cualquiera de las muestras, desde la muestra anterior que contiene al menos una imagen IRAP hasta esa propia muestra, inclusive.

[0137] De forma alternativa, cuando el nombre de entrada de muestra es «shv1», se aplica lo siguiente:

- Si una imagen codificada de una muestra es una imagen IRAP como la definida en ISO/IEC 23008-2, todos los conjuntos de parámetros necesarios para descodificar esa imagen codificada se incluirán en la entrada de muestra o en esa propia muestra.
- De lo contrario (la imagen codificada de la muestra no es una imagen IRAP), todos los conjuntos de parámetros necesarios para descodificar esa imagen codificada se incluirán en la entrada de muestra o en cualquiera de las muestras, desde la muestra anterior que contiene una imagen IRAP en la misma capa que esa imagen codificada hasta esa propia muestra, inclusive.

[0138] Si un flujo SHVC elemental contiene una capa base compatible con HEVC usable, se usará una entrada de muestra visual HEVC («hvc1» o «hev1»). En este caso, la entrada contendrá inicialmente una caja de configuración HEVC, posiblemente seguida de una caja de configuración SHVC como la definida a continuación. La caja de configuración HEVC documenta el perfil, la capa, el nivel y posiblemente también los conjuntos de parámetros pertenecientes a la capa base compatible con HEVC, según lo definido por el HEVCDecoderConfigurationRecord. La caja de configuración SHVC documenta el perfil, la capa, el nivel y, posiblemente, también los conjuntos de parámetros pertenecientes al flujo completo que contiene las capas de mejora compatibles con SHVC, según lo definido por el HEVCDecoderConfigurationRecord, almacenado en la SHVCConfigurationBox.

[0139] Si el flujo elemental SHVC no contiene ninguna capa base HEVC usable, se usará una entrada de muestra visual SHVC («shc1» o «shv1»). La entrada de muestra visual SHVC contendrá una caja de configuración SHVC, como la definida a continuación. Esto incluye un SHVCDecoderConfigurationRecord, como el definido en esta norma internacional.

[0140] El campo lengthSizeMinusOne de las configuraciones SHVC y HEVC de cualquier entrada de muestra dada tendrá el mismo valor.

[0141] Se pueden usar extractores o agregadores para unidades NAL con nuh_layer_id mayor que 0 en las pistas «hvc1», «hev1», «hvc2», «hev2», «shc1» o «shv1». Las «extra_boxes» de una entrada de muestra «hvc2» o «hev2» pueden ser una SHVCConfigurationBox u otras cajas de ampliación.

NOTA: Cuando se indica la compatibilidad con HEVC, puede ser necesario indicar un nivel poco realista para la capa base HEVC, a fin de adaptarse a la velocidad de bits del flujo completo, ya que todas las unidades NAL se consideran incluidas en la capa base HEVC y, en consecuencia, pueden facilitarse al descodificador, que se espera que descarte las unidades NAL que no reconoce. Este caso tiene lugar cuando se usa la entrada de muestra «hvc1» o «hev1» y las configuraciones HEVC y SHVC están presentes.

[0142] Una SHVCConfigurationBox puede estar presente en una entrada de muestra «hvc1» o «hev1». En este caso, se aplica la siguiente definición de HEVC SHVC SampleEntry.

[0143] La siguiente tabla muestra, para una pista de vídeo, todos los usos posibles de las entradas de muestra, las configuraciones y las herramientas SHVC (excluidos los metadatos temporizados, que siempre se usan en otra pista):

Tabla 10 - Uso de entradas de muestra para pistas HEVC y SHVC

nombre de entrada de muestra	de de	con registros de configuración	Significado
«hvc1» o «hev1»		Solo configuración HEVC	Una pista HEVC simple sin unidades NAL con nuh_layer_id mayor que 0; no estarán presentes ni extractores ni agregadores.
«hvc1» o «hev1»		Configuraciones HEVC y SHVC	Una pista SHVC con unidades NAL con un ID de capa nuh igual a 0 y unidades NAL con nuh_layer_id mayor que 0; pueden estar presentes extractores y agregadores; los extractores no harán referencia a unidades NAL con un ID de capa nuh igual a 0; los agregadores no contendrán, pero pueden hacer referencia a, unidades NAL con nuh_layer_id igual a 0.
«hvc2» o «hev2»		Solo configuración HEVC	Una pista HEVC simple sin unidades NAL con nuh_layer_id mayor que 0; pueden estar presentes y usarse extractores para hacer referencia a unidades NAL; pueden estar presentes agregadores para contener y hacer referencia a unidades NAL.
«hvc2» o «hev2»		Configuraciones HEVC y SHVC	Una pista SHVC con unidades NAL con un ID de capa nuh igual a 0 y unidades NAL con nuh_layer_id mayor que 0; pueden estar presentes extractores y agregadores; los extractores pueden hacer referencia a cualquier unidad NAL; los agregadores pueden contener, y hacer referencia a, cualquier unidad NAL.
«shc1» o «shv1»		Configuración SHVC	Una pista SHVC sin unidades NAL con ID de capa nuh igual a 0; pueden estar presentes y usarse extractores para hacer referencia a unidades NAL; pueden estar presentes agregadores para contener y hacer referencia a unidades NAL.

9.5.3.1.2 Sintaxis

[0144]

5

```

class SHVCConfigurationBox extends Box('shcC') {
    SHVCDecoderConfigurationRecord() SHVCConfig;
}
class HEVC2SampleEntry() extends VisualSampleEntry ('hev2' or 'hvc2') {
    HEVCConfigurationBox hevconfig;
    MPEG4BitRateBox (); // optional
    MPEG4ExtensionDescriptorsBox (); // optional
} extra_boxes boxes; // optional
class HEVC2SHVC2SampleEntry() extends HEVC2SampleEntry {
    SHVCConfigurationBox shvcconfig;
}
class HEVC2SHVC2SampleEntry() extends HEVC2SampleEntry() {
} SHVCConfigurationBox shvcconfig; // optional
// Use this if track is not HEVC compatible
class SHVC2SampleEntry() extends VisualSampleEntry ('shv1' or 'shc1') {
    SHVCConfigurationBox shvcconfig;
    MPEG4BitRateBox (); // optional
    MPEG4ExtensionDescriptorsBox (); // optional
} extra_boxes boxes; // optional

```

9.5.3.1.3 Semántica

- 10 [0145] Cuando el flujo al que se aplica la entrada de muestra contiene unidades NAL con nuh_layer_id mayor que 0, Compressorname en la clase base VisualSampleEntry indica el nombre del compresor usado con el valor «\013SHVC Coding» que se recomienda (\013 es 11, la longitud de la cadena «SHVC Coding» en bytes).

9.5.4 Anchura y altura visual SHVC

15

- [0146] La anchura y la altura visual documentadas en una VisualSampleEntry de un flujo que contiene unidades NAL con un ID de capa nuh mayor que 0 es la anchura y la altura visual de la capa base HEVC, si el flujo se describe

mediante una entrada de muestra de tipo «hvc1», «hev1», «hvc2» o «hev2»; de lo contrario, es la anchura y la altura visual de las imágenes descodificadas de la capa más alta al descodificar toda la secuencia.

9.5.5 Muestra de sincronización

[0147] Una muestra SHVC se considera una muestra de sincronización si cada imagen codificada de la unidad de acceso es una imagen IRAP, como la definida en ISO/IEC 23008-2. Las muestras de sincronización se documentan mediante la tabla de muestras de sincronización, y pueden documentarse adicionalmente mediante el grupo de muestras de muestras de sincronización y el grupo de muestras «rap».

9.5.5.1 Grupo de muestras de muestras accesibles aleatoriamente

9.5.5.1.1 Definición

[0148]

Tipos de grupos: «ras»

Contenedor: Caja de descripción de grupo de muestras («ras»)

Obligatorio: No

Cantidad: Cero o mas

[0149] Un grupo de muestras de muestras accesibles aleatoriamente identifica muestras que contienen al menos una imagen IRAP.

9.5.5.1.2 Sintaxis

[0150]

```
class RandomAccessibleSampleEntry() extends VisualSampleGroupEntry ('ras') {
    unsigned int(1) reserved = 0;
    unsigned int(1) all_pics_are_IRAP
    unsigned int(6) IRAP_nal_unit_type
    if(all_pics_are_IRAP) {
        unsigned int(2) reserved = 0;
        unsigned int(6) num_IRAP_pics;
        for(i=0; i< num_IRAP_pics; i++) {
            unsigned int(2) reserved = 0;
            unsigned int(6) IRAP_pic_layer_id
        }
    }
}
```

9.5.5.1.3 Semántica

[0151] all_pics_are_IRAP igual a 1 especifica que todas las imágenes codificadas de cada muestra del grupo son imágenes IRAP. Cuando el valor es igual a 0, la restricción anterior puede o no aplicarse.

[0152] IRAP_nal_unit_type especifica el tipo de unidad NAL de las imágenes IRAP de cada muestra del grupo. El valor de IRAP_nal_unit_type deberá estar dentro del intervalo de 16 a 23, inclusive.

[0153] num_IRAP_pics especifica el número de imágenes IRAP de cada muestra del grupo. IRAP_pic_layer_id especifica el valor del ID de capa nuh de la iésima imagen IRAP de cada muestra del grupo.

9.5.6 Grupos de muestras en puntos de recuperación de acceso aleatorio y puntos de acceso aleatorio

[0154] Para datos de vídeo descritos mediante una entrada de muestra de tipo «hvc1», «hev1», «hvc2» o «hev2», el grupo de muestras de recuperación de acceso aleatorio y el grupo de muestras de punto de acceso aleatorio identifican puntos de recuperación de acceso aleatorio y puntos de acceso aleatorio, respectivamente, tanto para un descodificador HEVC como para un descodificador SHVC (si lo hubiera) que operan con todo el flujo de bits. Para datos de vídeo descritos mediante una entrada de muestra de tipo «shc1» o «shv1», el grupo de muestras de recuperación de acceso aleatorio identifica la recuperación de acceso aleatorio en todo el flujo de bits SHVC y el grupo de muestras de punto de acceso aleatorio identifica puntos de acceso aleatorio en todo el flujo de bits SHVC.

[0155] Una muestra SHVC se considera un punto de acceso aleatorio si cada imagen codificada de la unidad de acceso es una imagen IRAP (con o sin imágenes RASL) como la definida en ISO/IEC 23008-2, y las muestras iniciales en ISO/IEC 14496-2 son muestras en las que todas las imágenes son imágenes RASL como las definidas en ISO/IEC 23008-2.

9.5.7 Caja de muestras desechables independientes

[0156] Si se usa en una pista que es compatible tanto con HEVC como con SHVC, entonces debe procurarse que las sentencias sean verdaderas, independientemente de qué subconjunto válido de los datos SHVC (posiblemente solo los datos HEVC) se use. Los valores «desconocidos» (valor 0 de los campos `sample_depends_on`, `sample_is_dependent_on` y `sample_has_redundancy`) pueden ser necesarios si la información varía.

9.5.8 Definición de una submuestra para SHVC

[0157] Esta subcláusula amplía la definición de submuestra de HEVC en 8.4.8.

[0158] Para el uso de la caja de información de submuestra (8.7.7 de ISO/CEI 14496-12) de un flujo SHVC, una submuestra se define sobre la base del valor de los indicadores de la caja de información de submuestra como se especifica a continuación. La presencia de esta caja es opcional; sin embargo, si está presente en una pista que contiene datos SHVC, tendrá la semántica definida aquí.

[0159] “flags” especifica el tipo de información de submuestra que se proporciona en esta caja de la siguiente manera:

0: Submuestras basadas en unidades NAL. Una submuestra contiene una o más unidades NAL contiguas.

1: Submuestras basadas en unidades de decodificación. Una submuestra contiene exactamente una unidad de decodificación.

2: Submuestras basadas en mosaico. Una submuestra contiene un mosaico y las unidades no VCL NAL asociadas, si las hubiera, de la(s) unidad(es) VCL NAL que contiene(n) el mosaico, o contiene una o más unidades no VCL NAL.

3: Submuestras basadas en filas de CTU. Una submuestra contiene una fila de CTU dentro de un sector y las unidades no VCL NAL asociadas, si las hubiera, de la(s) unidad(es) VCL NAL que contiene(n) la fila CTU o contiene(n) una o más unidades no VCL NAL. Este tipo de información de submuestra no se usará cuando `entropy_coding_sync_enabled_flag` sea igual a 0.

4: Submuestras basadas en sectores. Una submuestra contiene un sector (donde cada sector puede contener uno o más segmentos de sector, cada uno de los cuales es una unidad NAL) y las unidades no VCL NAL asociadas, si las hubiera, o contiene una o más unidades no VCL NAL.

5: Submuestras basadas en imágenes. Una submuestra contiene una imagen codificada y las unidades no VCL NAL asociadas.

[0160] Otros valores de indicadores [flags] están reservados.

[0161] El campo `subsample_priority` se establecerá en un valor de acuerdo con la especificación de este campo en ISO/IEC 14496-12.

[0162] El campo `discardable` se establecerá en 1 solo si esta muestra todavía puede descodificarse si esta submuestra se descarta (por ejemplo, la submuestra consiste en una unidad SEI NAL).

[0163] Cuando el primer byte de una unidad NAL está incluido en una submuestra, el campo de longitud anterior también debe estar incluido en la misma submuestra.

```
if(flags 0) {
    unsigned int(1) SubLayerRefNalUnitFlag;
    unsigned int(1) RapNalUnitFlag;
    unsigned int(1) VclNalUnitFlag;
    unsigned int(1) DiscardableFlag;
```

```

    unsigned int(1) NoInterLayerPredFlag;
    unsigned int(6) LayerId;
5    unsigned int(3) TempId;
    unsigned int(18) reserved = 0;
    }else if(flags==1)
10    unsigned int(32) reserved = 0;
    else if(flags==2) {
15    unsigned int(2) vcl_idc;
    unsigned int(2) reserved = 0;
    unsigned int(4) log2_min_luma_ctb;
20    unsigned int(12) ctb_x;
    unsigned int(12) ctb_y;
25    } else if(flags == 3 || flags == 4) {
    unsigned int(2) vcl_idc;
    unsigned int(30) reserved = 0;
30    } else if(flags == 5) {
    unsigned int(1) DiscardableFlag;
35    unsigned int(6) VclNalUnitType;
    unsigned int(6) LayerId;
    unsigned int(3) TempId;
40    unsigned int(1) NoInterLayerPredFlag;
    unsigned int(1) SubLayerRefNalUnitFlag;
45    unsigned int(14) reserved=0;
    }

```

[0164] SubLayerRefNalUnitFlag igual a 0 indica que todas las unidades NAL de la submuestra son unidades VCL NAL de una imagen no de referencia de subcapa como la especificada en ISO/IEC 23008-2. El valor 1 indica que todas las unidades NAL de la submuestra son unidades VCL NAL de una imagen de referencia de subcapa como la especificada en ISO/IEC 23008-2. RapNalUnitFlag igual a 0 indica que ninguna de las unidades NAL de la submuestra tiene nal_unit_type igual a IDR_W_RADL, IDR_N_LP, CRA_NUT, BLA_W_LP, BLA_WRADL, BLA_N_LP, RSV_IRAP_VCL22 o RSV_IRAP_VCL23 como se indica en ISO/IEC 23008-2. El valor 1 indica que todas las unidades NAL de la submuestra tienen un nal_unit_type igual a IDR_W_RADL, IDR_N_LP, CRA_NUT, BLA_W_LP, BLA_WRADL, BLA_N_LP, RSV_IRAP_VCL22 o RSV_IRAP_VCL23 como se especifica en ISO/IEC 23008-2.

[0165] VclNalUnitFlag igual a 0 indica que todas las unidades NAL de la submuestra son unidades no VCL NAL. El valor 1 indica que todas las unidades NAL de la submuestra son unidades VLC NAL.

[0166] DiscardableFlag indica el valor de indicador descartable de las unidades VCL NAL de la submuestra. Todas las unidades VCL NAL de la submuestra tendrán el mismo valor de indicador descartable.

OBSÉRVESE que esta no es la misma definición que la del campo descartable de la caja de información de submuestra.

[0167] NoInterLayerPredFlag indica el valor del indicador de pred. entre capas habilitada de las unidades VCL NAL de la submuestra. Todas las unidades VCL NAL de la submuestra tendrán el mismo valor del indicador de pred. entre capas habilitada.

[0168] LayerId indica el valor de ID de capa nuh de las unidades NAL de la submuestra. Todas las unidades NAL de la submuestra tendrán el mismo valor de ID de capa nuh.

[0169] TemporalId indica el valor de TemporalId de las unidades NAL de la submuestra. Todas las unidades NAL de la submuestra tendrán el mismo valor de TemporalId.

[0170] vcl_idc indica si la submuestra contiene datos de capa de codificación de vídeo (VCL), datos no VCL, o ambos, de la siguiente manera:

0: la submuestra contiene datos VCL y no contiene datos no VCL

1: la submuestra no contiene datos VCL y contiene datos no VCL

2: la submuestra puede contener datos VCL y no VCL, que se asociarán entre sí. Por ejemplo, una submuestra puede contener un mensaje SEI de información de unidad de descodificación seguido del conjunto de unidades NAL asociadas con el mensaje SEI.

3: reservados

[0171] log2_min_luma_ctb indica la unidad de ctb_x y ctb_y, especificada de la siguiente manera:

0: 8 muestras de luma

1: 16 muestras de luma

2: 32 muestras de luma

3: 64 muestras de luma

[0172] ctb_x especifica la coordenada basada en 0 de las muestras de luma más a la derecha del mosaico asociado con la submuestra cuando flags es igual a 2 y vcl_idc es igual a 1 o 2, en unidades derivadas de log2_min_luma_ctb como se ha especificado anteriormente.

[0173] ctb_y especifica la coordenada basada en 0 de las muestras de luma más abajo del mosaico asociado con la submuestra cuando flags es igual a 2 y vcl_idc es igual a 1 o 2, en unidades derivadas de log2_min_luma_ctb como se ha especificado anteriormente.

[0174] VclNalUnitType indica el valor de tipo de unidad NAL de las unidades VCL NAL de la submuestra. Todas las unidades VCL NAL de la submuestra tendrán el mismo valor de tipo de unidad NAL.

9.5.9 Tratamiento de muestras no de salida

[0175] La especificación de 8.4.9 se aplica reemplazando «HEVC» por «SHVC», y una muestra no de salida se define como una muestra en la que la imagen o las imágenes de las capas de salida de destino tienen pic_output_flag igual a 0. Cuando una unidad de acceso contiene algunas imágenes codificadas que tienen el indicador pic_output igual a 1 y algunas otras imágenes codificadas que tienen el indicador pic_output igual a 0, se deben usar al menos dos pistas para almacenar el flujo, de tal forma que dentro de cada pista todas las imágenes codificadas de cada muestra tiene el mismo valor de pic_output_flag.

10 10 Definiciones de flujo elemental y muestras MV-HEVC

10.1 Introducción

[0176] Esta cláusula especifica el formato de almacenamiento de datos MV-HEVC. Amplía las definiciones del formato de almacenamiento de HEVC de la cláusula 8.

[0177] El formato de archivo para el almacenamiento de contenido de MV-HEVC, definido en esta cláusula y en los Anexos A a D, usa las capacidades existentes del formato base de archivo de medios ISO y el formato de archivo HEVC simple (es decir, el formato de archivo especificado en la cláusula 8). Además, se usan las siguientes estructuras o ampliaciones, entre otras, para admitir características específicas de MV-HEVC.

[0178] Agregador:

estructura para permitir una agrupación escalable eficiente de unidades NAL transformando patrones irregulares de unidades NAL en patrones regulares de unidades de datos agregadas.

5 **[0179]** Extractor:

estructura para permitir la extracción eficiente de unidades NAL de otras pistas que no son la que contiene los datos de medios.

10 **[0180]** Compatibilidad HEVC:

una disposición para almacenar un flujo de bits MV-HEVC de una manera compatible con HEVC, de tal forma que cualquier lector compatible con el formato de archivo HEVC simple pueda usar la capa base compatible con HEVC.

15 **[0181]** El soporte para MV-HEVC incluye un número de herramientas, y hay diversos «modelos» de cómo podrían usarse. En particular, un flujo MV-HEVC se puede colocar en unas pistas en un número de maneras, entre las que se encuentran las siguientes:

- 20 1. todas las vistas en una pista, etiquetadas con grupos de muestras;
2. cada vista en su propia pista, etiquetada en las entradas de muestra;
3. un híbrido de una pista que contiene todas las vistas y una o más pistas de vista única, cada una de las cuales contiene una vista que puede codificarse de forma independiente;
- 25 4. los puntos operativos esperados cada uno en una pista (por ejemplo, la base HEVC, un par estéreo, una escena multivista).

30 **[0182]** El formato de archivo MV-HEVC permite el almacenamiento de una o más vistas en una pista, de manera similar al soporte para SHVC en la cláusula 9. Se puede usar el almacenamiento de múltiples vistas por pista, por ejemplo, cuando un proveedor de contenido desea proporcionar un flujo de bits multivista que no está destinado a subconjuntos o cuando el flujo de bits se ha creado para unos pocos conjuntos predefinidos de vistas de salida (como 1, 2, 5 o 9 vistas), pudiéndose crear las pistas como corresponda. Si más de una vista se almacena en una pista y hay varias pistas (más de una) que representan el flujo de bits MV-HEVC, se recomienda el uso del mecanismo de agrupación de muestras.

35 **[0183]** Cuando un flujo de bits MV-HEVC está representado por múltiples pistas y un reproductor usa un punto operativo que contiene datos en múltiples pistas, el reproductor debe reconstruir las unidades de acceso MV-HEVC antes de pasarlas al descodificador MV-HEVC. Un punto operativo MV-HEVC puede estar representado explícitamente por una pista, es decir, una unidad de acceso se reconstruye simplemente determinando todas las unidades NAL de extractor y agregador de una muestra. Si el número de puntos operativos es grande, crear una pista para cada punto operativo puede consumir espacio y ser poco práctico. En dicho caso, las unidades de acceso MV-HEVC se reconstruyen como se especifica en 10.5.2. El registro de configuración de descodificador MV-HEVC contiene un campo que indica si las muestras asociadas usan una reconstrucción de unidades de acceso explícita o implícita (véase el campo `explicit_au_track`).

10.2 Estructura de pista MV-HEVC

50 **[0184]** Los flujos MV-HEVC se almacenan de acuerdo con 8.2, con la siguiente definición de un flujo elemental de vídeo MV-HEVC:

- **Los flujos elementales de vídeo MV-HEVC** contendrán todas las unidades NAL relacionadas con la codificación de vídeo (es decir, las unidades NAL que contengan una estructura de datos de vídeo o de señalización de vídeo) y pueden contener unidades NAL no relacionadas con la codificación de vídeo, como mensajes SEI y unidades NAL delimitadoras de unidades de acceso. También pueden estar presentes agregadores (véase A.2) o extractores (véase A.3). Los agregadores y extractores se procesarán como se define en esta norma internacional (por ejemplo, no se colocarán directamente en la memoria intermedia de salida al acceder al archivo). Otras unidades NAL que no están expresamente prohibidas pueden estar presentes, y si no se reconocen deberían ignorarse (por ejemplo, no se colocan en la memoria intermedia de salida al acceder al archivo).

60 **[0185]** Los flujos MV-HEVC no se almacenarán usando flujos de conjuntos de parámetros asociados, cuando sea necesario.

[0186] Puede haber unidades VCL NAL con nuh_layer_id igual a 0, unidades VCL NAL con nuh_layer_id mayor que 0, y otras unidades no VCL NAL, presentes en un flujo elemental de vídeo MV-HEVC. Adicionalmente, puede haber unidades NAL de agregador o extractor presentes en un flujo elemental de vídeo MV-HEVC.

5 10.3 Uso del formato de archivo HEVC simple

[0187] El formato de archivo MV-HEVC es una ampliación del formato de archivo HEVC simple definido en la cláusula 8.

10 10.4 Definición de muestra y configuración

10.4.1 Introducción

[0188] Muestra MV-HEVC: Una muestra MV-HEVC también es una unidad de acceso como la definida en el Anexo F de ISO/IEC 23008-2.

10.4.2 Orden canónico y restricción

10.4.2.1 Restricciones

[0189] Las siguientes restricciones se aplican a los datos MV-HEVC además de los requisitos de la cláusula 8.3.2.

- **Unidades VCL NAL:** Todas las unidades VCL NAL de una unidad de acceso estarán contenidas en la muestra cuyo tiempo de composición sea el de la imagen representada por la unidad de acceso. Una muestra MV-HEVC contendrá al menos una unidad VCL NAL.
- **Agregadores/extractores:** El orden de todas las unidades NAL incluidas en un agregador o a las que hace referencia un extractor es exactamente el orden de descodificación que tendrían si estas unidades NAL estuvieran presentes en una muestra que no contuviera agregadores/extractores. Después de procesar el agregador o el extractor, todas las unidades NAL deben estar en un orden de descodificación válido como se especifica en ISO/IEC 23008-2.

10.4.2.2 Registro de configuración de descodificador

[0190] Cuando el registro de configuración de descodificador definido en la cláusula 8.3.3.1 se usa para un flujo que puede interpretarse como un flujo MV-HEVC o HEVC, el registro de configuración de descodificador HEVC reflejará las propiedades de la vista base compatible con HEVC, por ejemplo, contendrá solo los conjuntos de parámetros necesarios para descodificar la vista base HEVC.

[0191] El MVHEVCDecoderConfigurationRecord es estructuralmente idéntico a un HEVCDecoderConfigurationRecord. La sintaxis es la siguiente:

```
aligned(8) class MVHEVCDecoderConfigurationRecord {
    // same fields as in HEVCDecoderConfigurationRecord syntax
    structure
}
```

[0192] La semántica de los campos de MVHEVCDecoderConfigurationRecord es la misma que la definida para un HEVCDecoderConfigurationRecord.

10.4.3 Muestra de sincronización

[0193] Una muestra MV-HEVC se considera una muestra de sincronización si cada imagen codificada de la unidad de acceso es una imagen IRAP sin imágenes RASL, como se define en ISO/IEC 23008-2. Las muestras de sincronización se documentan mediante la tabla de muestras de sincronización, y pueden documentarse adicionalmente mediante el grupo de muestras de muestras sincronizadas y el grupo de muestras «rap» definido de manera similar a como se hace en SHVC.

10.4.4 Caja de muestras independientes y desechables

[0194] Si se usa en una pista que es compatible tanto con HEVC como con MV-HEVC, deberá procurarse que las sentencias sean verdaderas, independientemente de qué subconjunto válido de datos de MV-HEVC (posiblemente solo los datos de HEVC) se use. Los valores «desconocidos» (valor 0 de los campos sample_depends_on, sample_is_dependent_on y sample_has_redundancy) pueden ser necesarios si la información varía.

10.4.5 Grupos de muestras en puntos de recuperación de acceso aleatorio y puntos de acceso aleatorio

[0195] Para los datos de vídeo descritos mediante una entrada de muestra de tipo «hvc1», «hev1», «hvc2» o «hev2», el grupo de muestras de recuperación de acceso aleatorio y el grupo de muestras de punto de acceso aleatorio identifican puntos de recuperación de acceso aleatorio y puntos de acceso aleatorio, respectivamente, tanto para un decodificador HEVC como para un decodificador MV-HEVC (si lo hubiera) que operan con todo el flujo de bits.

[0196] Para los datos de vídeo descritos mediante un tipo de entrada de muestra MV-HEVC, el grupo de muestras de recuperación de acceso aleatorio identifica la recuperación de acceso aleatorio en todo el flujo de bits MV-HEVC y el grupo de muestras de punto de acceso aleatorio identifica puntos de acceso aleatorio en todo el flujo de bits MV-HEVC.

10.5 Obtención a partir del formato base de archivo de medios ISO

10.5.1 Estructura de pista MV-HEVC

[0197] Un flujo de vídeo multivista está representado por una o más pistas de vídeo en un archivo. Cada pista representa una o más vistas del flujo.

[0198] Hay un conjunto mínimo de una o más pistas que, cuando se toman juntas, contienen el conjunto completo de información codificada. Todas estas pistas tendrán el indicador «complete_representation» establecido en todas sus entradas de muestra. Este grupo de pistas que forman la información codificada completa se denomina «subconjunto completo».

[0199] Se va a suponer que el punto de operación más bajo es el uno de la totalidad de los puntos operativos que contiene unidades NAL con nuh_layer_id igual a 0 solo y TemporalId igual a 0 solo. Una pista que contenga el punto operativo más bajo se designará como la «pista de vista base». Todas las demás pistas que forman parte del mismo flujo se vincularán a esta pista base mediante una referencia de pista del tipo «sbas» (base de vista).

[0200] Todas las pistas que comparten la misma pista de vista base deben compartir la misma escala de tiempo que la pista de vista base.

[0201] Si una vista representada por una pista usa otra vista representada por otra pista como referencia de predicción entre vistas, se incluirá una referencia de pista del tipo «scal» en la pista que se refiere a la pista de origen para la predicción entre vistas.

[0202] Si se aplican ediciones a las pistas que contienen componentes de vista de un flujo de bits MV-HEVC, las listas de edición serán coherentes en todas las pistas afectadas por las ediciones.

10.5.2 Reconstrucción de una unidad de acceso

[0203] A fin de reconstruir una unidad de acceso a partir de muestras de una o más pistas MV-HEVC, puede ser necesario determinar primero las vistas de salida de destino.

[0204] Puede llegarse a una conclusión sobre las vistas que se requieren para decodificar las vistas de salida de destino determinadas, a partir de los identificadores de vista de referencia incluidos en la caja de identificadores de vista o las referencias de pista «scal».

[0205] Si varias pistas contienen datos para la unidad de acceso, la alineación de las muestras respectivas en las pistas se realiza en el tiempo de decodificación, es decir, usando la tabla de tiempo-muestras solo sin considerar las listas de edición.

[0206] Una unidad de acceso se reconstruye a partir de las muestras respectivas de las pistas requeridas disponiendo sus unidades NAL en un orden de conformidad con ISO/IEC 23008-02. El siguiente orden proporciona un resumen del procedimiento para formar una unidad de acceso adecuada:

- Todas las unidades NAL del conjunto de parámetros (de las pistas de conjuntos de parámetros asociadas y de las pistas de flujo elemental asociadas).
- Todas las unidades SEI NAL (de las pistas de conjunto de parámetros asociadas y de las pistas de flujo elemental asociadas).
- Componentes de vista en orden ascendente de valor de índice de orden de vista. Las unidades NAL de un componente de vista están en la muestra por su orden de aparición.

10.5.3 Entrada de muestra

10.5.3.1 Cajas para entrada de muestra

10.5.3.1.1 Caja de identificador de vista

10.5.3.1.1.1 Definición

[0207]

Tipo de ventana: «vwid»

Contenedor: Entrada de muestra («hev1», «hvc1», «hev2», «hvc2», «mhc1», «mhv1») o MultiviewGroupEntry

Obligatorio: Sí (para entradas de muestra)

Cantidad: Exactamente una (para entradas de muestra)

[0208] Cuando está incluida en una entrada de muestra, esta caja indica las vistas incluidas en la pista. Este caja también indica el índice de orden de vista para cada vista enumerada. Adicionalmente, la caja incluye los valores mínimo y máximo de temporal_id incluidos en la pista cuando la caja de identificador de vista está incluida en una entrada de muestra. Por otra parte, la caja indica las vistas a las que se hace referencia requeridas para descodificar las vistas incluidas en la pista.

10.5.3.1.1.2 Sintaxis

[0209]

```
class ViewIdentifierBox extends FullBox ('vwid', version=0, flags)
{
    unsigned int(2) reserved6 = 0;
    unsigned int(3) min_temporal_id;
    unsigned int(3) max_temporal_id;
    unsigned int(16) num_views;
    for (i=0; i<num_views; i++) {
        unsigned int(6) reserved1 = 0;
        unsigned int(6) layer_id[i];
        unsigned int(10) view_id[i];
        unsigned int(2) base_view_type;
        for (j = 0; j < layer_id[i]; j++) {
            unsigned int(1) depend_layer[i][j];
        }
    }
}
```

10.5.3.1.1.3 Semántica

[0210] min_temporal_id y max_temporal_id toman el valor mínimo y máximo, respectivamente, del elemento sintáctico temporal_id que está presente en la ampliación de cabecera de unidad NAL de las unidades NAL correlacionadas con la pista o capa cuando la caja de identificador de vista está incluida en una entrada de muestra, respectivamente. Para los flujos AVC este adopta el valor que hay, o habría, en la unidad NAL de prefijo.

[0211] num_views, cuando la caja de identificador de vista está presente en una entrada de muestra, indica el número de vistas incluidas en la pista.

[0212] layer_id[i] indica el valor del elemento sintáctico nuh_layer_id en la cabecera de unidad NAL de una capa incluida en la pista cuando la caja de identificador de vista está incluida en una entrada de muestra.

[0213] view_id indica el identificador de vista de la i-ésima capa con el ID de capa nuh igual a layer_id[i], como se especifica en el Anexo F de ISO/IEC 23008-2.

[0214] base_view_type indica si la vista es una vista base (virtual o no). Adopta los siguientes valores:

0 indica que la vista no es ni una vista base ni una vista base virtual.

1 se usará para etiquetar la vista base no virtual del flujo de bits MV-HEVC.

2 es un valor reservado y no se usará.

3 indica que la vista con view_id[i] es una vista base virtual. La respectiva vista no base codificada independientemente con view_id[i] reside en otra pista. Cuando base_view_type es igual a 3, las subsiguientes num_ref_views serán iguales a 0.

[0215] depdent_layer[i][j] indica si la j-ésima capa con nuh_layer_id igual a j puede ser una capa a la que se hace referencia de forma directa o indirecta de la capa con nuh_layer_id igual a layer_id[i]. Cuando la caja de identificador de vista está incluida en una entrada de muestra, se recomienda indicar las vistas a las que se hace referencia en la misma entrada de muestra.

10.5.3.2 Definición de entrada de muestra

[0216]

Tipos de entradas de muestra: «hvc2», «hev2», «mhc1», «mhv1», «mhcC»,
 Contenedor: Caja de descripción de muestra («stsd»)
 Obligatorio: Una de las cajas «hvc1», «hev1», «hvc2», «hev2», «mhc1» o «mhv1» es obligatoria
 Cantidad: Pueden estar presentes una o más entradas de muestra

[0217] Si un flujo elemental MV-HEVC contiene una vista base compatible con HEVC usable, entonces se usará una entrada de muestra visual HEVC («hvc1», «hev1», «hvc2», «hev2»). Aquí, la entrada contendrá inicialmente una caja de configuración HEVC, posiblemente seguida de una caja de configuración MV-HEVC como se define a continuación. La caja de configuración HEVC documenta el perfil, el nivel y, posiblemente también, los conjuntos de parámetros pertenecientes a la vista base compatible con HEVC, definidos por el HEVCDecoderConfigurationRecord. La caja de configuración MV-HEVC documenta el perfil, el nivel y la información del conjunto de parámetros pertenecientes al flujo completo que contiene las vistas no base tal como se define en el MVHEVCDecoderConfigurationRecord, almacenado en la MVHEVCConfigurationBox. Para todas las entradas de muestra «hvc1», «hev1», «hvc2» y «hev2», los campos de anchura y altura de la entrada de muestra documentan la capa base HEVC. Para una entrada de muestra MV-HEVC («mhc1», «mhv1»), la anchura y la altura documentan la resolución lograda al descodificar cualquier vista individual del flujo completo.

[0218] Si el flujo elemental MV-HEVC no contiene ninguna vista base HEVC usable, se usará una entrada de muestra visual MV-HEVC («mhc1», «mhv1»). La entrada de muestra visual MV-HEVC contendrá una caja de configuración MV-HEVC, como se define a continuación. Esta incluye un MVHEVCDecoderConfigurationRecord, como el definido en esta norma internacional.

[0219] El campo lengthSizeMinusOne de las configuraciones MV-HEVC y HEVC de cualquier entrada de muestra dada tendrá el mismo valor.

[0220] Los requisitos para los tipos de entrada de muestra «hvc1» y «hev1» documentados en 6.5.3.1.1 también se aplican aquí.

[0221] La MVHEVCConfigurationBox puede estar presente en una entrada de muestra «hvc1», «hev1», «hvc2», «hev2». En estos casos, se aplica la definición HEVCMVHEVCsSampleEntry o HEVC2MVHEVCsSampleEntry siguientes, respectivamente. Compressorname en la clase base VisualSampleEntry indica el nombre del compresor usado, recomendándose el valor «\014MV-HEVC Coding» (\016 es 14, la longitud de la cadena «MV-HEVC Coding» en bytes).

[0222] Los conjuntos de parámetros requeridos para descodificar una unidad NAL que está presente en los datos de muestra de un flujo de vídeo, ya sea directamente o por referencia de un extractor, estarán presentes en la configuración del descodificador de ese flujo de vídeo o en el flujo del conjunto de parámetros asociado (si se usara).

[0223] La siguiente tabla muestra para una pista de vídeo todos los usos posibles de las entradas de muestra cuando un flujo elemental MV-HEVC se almacena en una o más pistas, configuraciones y herramientas MV-HEVC.

Tabla 14 Uso de entradas de muestra para pistas HEVC y MV-HEVC

nombre de entrada de muestra	con registros de configuración	Significado
«hvc1» o «hev1»	Solo configuración HEVC	Una pista HEVC simple sin unidades NAL con nuh_layer_id mayor que 0; no estarán presentes ni extractores ni agregadores.

nombre de entrada de muestra	con registros de configuración	Significado
«hvc1» o «hev1»	Configuraciones HEVC y MV-HEVC	Una pista MV-HEVC con unidades NAL con nuh_layer_id igual a 0 y unidades NAL con un ID de capa nuh mayor que 0; pueden estar presentes extractores y agregadores; los extractores no harán referencia a unidades NAL con un ID de capa nuh igual a 0; los agregadores no contendrán, pero podrán hacer referencia a, unidades NAL con nuh_layer_id igual a 0.
«hvc2» o «hev2»	Solo configuración HEVC	Una pista HEVC simple sin unidades NAL con nuh_layer_id mayor que 0; pueden estar presentes y usarse extractores para hacer referencia a unidades NAL; pueden estar presentes agregadores para contener y hacer referencia a unidades NAL.
«hvc2» o «hev2»	Configuraciones HEVC y MV-HEVC	Una pista MV-HEVC con unidades NAL con nuh_layer_id igual a 0 y unidades NAL con un ID de capa nuh mayor que 0; pueden estar presentes extractores y agregadores; los extractores pueden hacer referencia a cualquier unidad NAL; los agregadores pueden contener y hacer referencia a cualquier unidad NAL.
«mhc1» o «mhv1»	Configuración MV-HEVC	Una pista MV-HEVC sin unidades NAL con un ID de capa nuh igual a 0; pueden estar presentes extractores y usarse para hacer referencia a unidades NAL; pueden estar presentes agregadores para contener y hacer referencia a unidades NAL.

[0224] La entrada de muestra tipo mvhevc siguiente es una de {mhv1, mhc1}.

10.5.3.3 Sintaxis

[0225]

```
class MVHEVCConfigurationBox extends Box('mhcC') {
    MVHEVCDecoderConfigurationRecord() MVHEVCConfig;
}

class HEVCMVHEVCSampleEntry() extends HEVCSampleEntry() {
    ViewIdentifierBox view_identifiers; // optional
    MVHEVCConfigurationBox mvhevccconfig; // optional
}

class HEVC2MVHEVCSampleEntry() extends HEVC2SampleEntry() {
    ViewIdentifierBox view_identifiers; // optional
    MVHEVCConfigurationBox mvhevccconfig; // optional
}

// Use this if the track is NOT HEVC compatible
class MVHEVCSampleEntry() extends VisualSampleEntry(mvhevc-type) {
    MVHEVCConfigurationBox mvhevccconfig; // mandatory
    ViewIdentifierBox view_identifiers; // mandatory
    MPEG4BitRateBox bitrate; // optional
    MPEG4ExtensionDescriptorsBox descr; // optional
}
```

10.5.4 Definición de una submuestra para MV-HEVC

[0226] La definición de submuestra para MV-HEVC se define de manera similar a la definida para SHVC.

10.5.5 Tratamiento de muestras no de salida

[0227] El tratamiento de muestras no de salida para MV-HEVC se define de manera similar al definido para SHVC.

[0228] Los cambios al Anexo A se muestran a continuación.

Anexo A (normativa)

Estructuras en el flujo

A.1 Introducción

[0229] Los agregadores y los extractores son estructuras internas de formato de archivo que permiten la agrupación eficiente de unidades NAL o la extracción de unidades NAL a partir de otras pistas.

[0230] Los agregadores y extractores usan la sintaxis de unidad NAL. Estas estructuras son consideradas unidades NAL en el contexto de la estructura de muestra. Al acceder a una muestra, los agregadores deben eliminarse (dejando las unidades NAL que contienen o las unidades NAL a las que hacen referencia) y los extractores deben reemplazarse por los datos a los que hacen referencia. Los agregadores y extractores no deben estar presentes en un flujo fuera del formato de archivo.

[0231] Estas estructuras usan tipos de unidades NAL reservados para la capa de aplicación/transporte según ISO/IEC 14496-10 o ISO/IEC 23008-2.

NOTA: La siguiente nota es de ISO/IEC 14496-10:

«NOTA: los tipos de unidad NAL 0 y 24...31 se pueden usar según lo determine la aplicación. No se especifica ningún proceso de descodificación para estos valores de nal_unit_type en esta recomendación/norma internacional.»

NOTA: La siguiente nota es de ISO/IEC 23008-2:

«NOTA 1: Los tipos de unidades NAL del intervalo UNSPEC48...UNSPEC63 se pueden usar según lo determine la aplicación. En esta especificación no se especifica ningún proceso de descodificación para estos valores de nal_unit_type. Dado que diferentes aplicaciones podrían usar estos tipos de unidades NAL para diferentes propósitos, debe procederse con especial cuidado durante el diseño de codificadores que generan unidades NAL con estos valores nal_unit_type, y durante el diseño de descodificadores que interpretan el contenido de las unidades NAL con estos valores nal_unit_type.»

A.2 Agregadores

A.2.1 Definición

[0232] Esta subcláusula describe los agregadores, que permiten que las entradas NALU-map-group sean coherentes y repetitivas. (Véase Anexo B).

[0233] Los agregadores se usan para agrupar unidades NAL que pertenecen a la misma muestra.

[0234] Para el almacenamiento de vídeo ISO/IEC 14496-10, se aplican las siguientes reglas:

- Los agregadores usan la misma cabecera de unidad NAL que las unidades SVC VCL NAL o las unidades MVC VCL NAL, pero con un valor diferente de tipo de unidad NAL.
- Cuando el svc_extension_flag de la sintaxis de unidad NAL (especificada en 7.3.1 de ISO/IEC 14496-10) de un agregador es igual a 1, la cabecera de unidad NAL de las unidades SVC VCL NAL se usa para el agregador. De lo contrario, la cabecera de unidad NAL de las unidades MVC VCL NAL se usa para el agregador.

[0235] Para el almacenamiento de vídeo ISO/IEC 23008-2, los agregadores usan la cabecera de unidad NAL, definida en ISO/IEC 23008-2, que tiene la misma sintaxis para HEVC simple, SHVC y MV-HEVC.

[0236] Los agregadores pueden agregar, por *inclusión*, unidades NAL dentro de ellos (dentro del tamaño indicado por su longitud) y también agregar, por *referencia*, unidades NAL que los siguen (dentro del área indicada por el campo additional_bytes que se encuentra dentro de ellos). Cuando un lector de archivos AVC o HEVC explora el flujo, solo las unidades NAL incluidas se consideran «internas» al agregador. Esto permite que un lector de archivos AVC o HEVC omita un conjunto completo de unidades NAL no necesarias cuando se agregan por inclusión. Esto también permite que un lector AVC o HEVC no omita unidades NAL necesarias, sino que permita que estas permanezcan en el flujo cuando se agregan por referencia.

[0237] Se pueden usar agregadores para agrupar unidades NAL de capa base o vista base. Si estos agregadores se usan en una pista «avc1», «hvc1» o «hev1», entonces un agregador no usará la inclusión sino la referencia de unidades NAL de capa base o vista base (la longitud del agregador incluye solo su cabecera, y las unidades NAL a las que hace referencia el agregador se especifican mediante bytes adicionales).

[0238] Cuando un extractor con data_length igual a cero o un grupo de muestras de correlación hacen referencia al agregador, el agregador se trata como si agregara tanto los bytes incluidos como los bytes a los que se hace referencia.

[0239] Un agregador puede incluir o hacer referencia a extractores. Un extractor puede extraer de los agregadores. Un agregador no debe incluir ni hacer referencia a otro agregador directamente; sin embargo, un agregador puede incluir o hacer referencia a un extractor que hace referencia a un agregador.

5 **[0240]** Al explorar la secuencia:

a) si el agregador no se reconoce (por ejemplo, mediante un lector o decodificador AVC o HEVC) se descarta fácilmente con su contenido incluido;

10 b) si el agregador no es necesario (es decir, pertenece a una capa no deseada), este y su contenido por inclusión y referencia se descartan fácilmente (usando sus campos de longitud y de bytes adicionales);

c) si se necesita el agregador, su cabecera se descarta fácilmente y su contenido se conserva.

15 **[0241]** Un agregador se almacena dentro de una muestra como cualquier otra unidad NAL.

[0242] Todas las unidades NAL permanecen en orden de decodificación dentro de un agregador.

A.2.2 Sintaxis

20

[0243]

```
class aligned(8) Aggregator (AggregatorSize) {
    NALUnitHeader();
    unsigned int i = sizeof(NALUnitHeader());
    unsigned int((lengthSizeMinusOne+1)*8)
        additional_bytes;
    i += lengthSizeMinusOne+1;
    while (i<AggregatorSize) {
        unsigned int((lengthSizeMinusOne+1)*8)
            NALUnitLength;
        unsigned int(NALUnitLength*8) NALUnit;
        i += NALUnitLength+lengthSizeMinusOne+1;
    };
}
```

25 A.2.3 Semántica

[0244] El valor de la variable AggregatorSize es igual al tamaño de la unidad NAL del agregador, y la función sizeof(X) genera el tamaño del campo X en bytes.

30 NALUnitHeader(): los primeros cuatro bytes de las unidades SVC y MVC VCL NAL, o los primeros dos bytes de las unidades NAL ISO/IEC 23008-2.

nal_unit_type se establecerá en el tipo de unidad NAL de agregador (tipo 30 para vídeo ISO/IEC 14496-10 y tipo 48 para vídeo ISO/IEC 23008-2).

35

Para un agregador que incluye o hace referencia a unidades SVC NAL, se aplicará lo siguiente.

forbidden_zero_bit y reserved_three_2bits se establecerán como se especifica en ISO/IEC 14496-10.

40

Otros campos (nal_ref_idc, idr_flag, priority_id, no_inter_layer_pred_flag, dependency_id, quality_id, temporal id, use_ref_base_pic_flag, discardable_flag e indicador de salida) se establecerán como se especifica en A.4.

Para un agregador que incluye o hace referencia a unidades MVC NAL, se aplicará lo siguiente.

45

forbidden_zero_bit y reserved one bit se establecerán como se especifica en ISO/IEC 14496-10.

Otros campos (nal_ref_idc, non_idr_flag, priority_id, view_id, temporal_id, anchor_pic_flag e inter_view_flag) se establecerán como se especifica en A.5.

50

Para un agregador que incluye o hace referencia a unidades ISO/IEC 23008-2 NAL, se aplicará lo siguiente.

forbidden_zero_bit se establecerá como se especifica en ISO/IEC 23008-2.

Otros campos (ID de capa nuh y nuh_temporal_id_plus1) se establecerán como se especifica en A.6.

additional_bytes: El número de bytes que siguen a esta unidad NAL de agregador que se deberían considerar agregados cuando un extractor con data_length igual a cero o el grupo de muestras de correlación hacen referencia a este agregador.

NALUnitLength: Especifica el tamaño, en bytes, de la siguiente unidad NAL. El tamaño de este campo se especifica con el campo lengthSizeMinusOne.

Unidad NAL: una unidad NAL como la especificada en ISO/IEC 14496-10 o ISO/IEC 23008-2, incluida la cabecera de unidad NAL. El tamaño de la unidad NAL está especificado por NALUnitLength.

A.3 Extractores

A.3.1 Definición

[0245] Esta subcláusula describe los extractores, que permiten la formación compacta de pistas que extraen, por referencia, datos de unidad NAL de otras pistas.

[0246] Un agregador puede incluir o hacer referencia a extractores. Un extractor puede hacer referencia a agregadores. Cuando un lector de archivos que lo requiere procesa un extractor, el extractor se reemplaza lógicamente por los bytes a los que hace referencia. Esos bytes no deben contener extractores; un extractor no debe hacer referencia, de forma directa o indirecta, a otro extractor.

NOTA: La pista a la que se hace referencia puede contener extractores, aunque los datos a los que el extractor hace referencia no deban contenerlos.

[0247] Un extractor contiene una instrucción para extraer datos de otra pista, que está vinculada a la pista en la que reside el extractor, por medio de una referencia de pista de tipo «scal». Los bytes copiados serán uno de los siguientes:

- a) Una unidad NAL completa; obsérvese que cuando se hace referencia a un agregador, se copian tanto los bytes incluidos como los bytes a los que se hace referencia
- b) Más de una unidad NAL completa

[0248] En ambos casos, los bytes extraídos comienzan con un campo de longitud válida y una cabecera de unidad NAL.

[0249] Los bytes se copian solo de la muestra única identificada en la pista a la que se hace referencia a través de la referencia de pista «scal» indicada. La alineación se hace en tiempo de decodificación, es decir, usando solo la tabla tiempo-muestra, seguida de un desplazamiento contado en el número de muestra. Los extractores son un concepto de nivel de medios y, en consecuencia, se aplican a la pista de destino antes de considerar cualquier lista de edición. (Sin embargo, normalmente cabría esperar que las listas de edición de las dos pistas fueran idénticas).

A.3.2 Sintaxis

[0250]

```
class aligned(8) Extractor () {
    NALUnitHeader();
    unsigned int(8) track_ref_index;
    signed int(8) sample_offset;
    unsigned int((lengthSizeMinusOne+1)*8)
        data_offset;
    unsigned int((lengthSizeMinusOne+1)*8)
        data_length;
}
```

A.3.3 Semántica

[0251]

NALUnitHeader(): los primeros cuatro bytes de las unidades SVC y MVC VCL NAL, o los primeros dos bytes de las unidades NAL ISO/IEC 23008-2.

nal_unit_type se establecerá en el tipo de unidad NAL de extractor (tipo 31 para vídeo ISO/IEC 14496-10 y tipo 49 para vídeo ISO/IEC 23008-2).

Para un extractor que hace referencia a unidades SVC NAL, se aplicará lo siguiente.

forbidden_zero_bit y reserved_three_2bits se establecerán como se especifica en ISO/IEC 14496-10.

Otros campos (nal_ref_idc, idr_flag, priority_id, no_inter_layer_pred_flag, dependency_id, quality_id, temporal_id, use_ref_base_pic_flag, indicador de descartable y de salida) se establecerán como se especifica en A.4.

Para un extractor que hace referencia a unidades MVC NAL, se aplicará lo siguiente.

forbidden_zero_bit y reserved_one_bit se establecerán como se especifica en ISO/IEC 14496-10.

Otros campos (nal_ref_idc, non_idr_flag, priority_id, view_id, temporal_id, anchor_pic_flag e inter_view_flag) se establecerán como se especifica en A.5.

Para un extractor que hace referencia a unidades ISO/IEC 23008-2 NAL, se aplicará lo siguiente.

forbidden_zero_bit se establecerá como se especifica en ISO/IEC 23008-2.

Otros campos (nuh_layer_id y nuh_temporal_id_plus1) se establecerán como se especifica en A.6.

track_ref_index especifica el índice de la referencia de pista de tipo «scal» que se va a usar para encontrar la pista de donde se van a extraer los datos. La muestra de la pista de la que se extraen los datos está temporalmente alineada o en la posición anterior más cercana en la línea de tiempo de decodificación de medios, es decir, usando la tabla de tiempo-muestra solo, ajustada por un desplazamiento especificado por sample_offset, conteniendo la muestra el extractor. La primera referencia de pista tiene el valor de índice 1; el valor 0 está reservado.

sample_offset proporciona el índice relativo de la muestra en la pista vinculada que se usará como fuente de información. La muestra 0 (cero) es la muestra con el mismo tiempo de decodificación, o el anterior más cercano, en comparación con el tiempo de decodificación de la muestra que contiene el extractor; la muestra 1 (uno) es la siguiente muestra, la muestra -1 (menos 1) es la muestra anterior, y así sucesivamente.

data_offset: El desplazamiento del primer byte de la muestra de referencia que se va a copiar. Si la extracción comienza con el primer byte de datos de esa muestra, el desplazamiento toma el valor 0. El desplazamiento hará referencia al comienzo de un campo de longitud de unidad NAL.

data_length: El número de bytes que se van a copiar. Si este campo toma el valor 0, entonces se copia la única unidad NAL completa a la que se hace referencia (es decir, la longitud que se va a copiar se toma del campo de longitud al que hace referencia el desplazamiento de datos, aumentado por el campo de bytes adicionales en el caso de los agregadores).

NOTA: Si las dos pistas usan diferentes valores de lengthSizeMinusOne, entonces los datos extraídos deberán formatearse de nuevo para ajustarse al tamaño del campo de longitud de la pista de destino.

A.4 Valores de cabecera de unidad NAL para SVC

[0252] Tanto los extractores como los agregadores usan la ampliación SVC de cabecera de unidad NAL. Las unidades NAL extraídas por un extractor o agregadas por un agregador son todas las unidades NAL a las que se hace referencia o que se incluyen mediante inspección recursiva del contenido de las unidades NAL de agregador o extractor.

[0253] Los campos nal_ref_idc, idr_flag, priority_id, temporal_id, dependency_id, quality_id, el indicador de salida descartable, use_ref_base_pic_flag, y no_inter_layer_pred_flag tomarán los siguientes valores:

nal_ref_idc se establecerá en el valor más alto del campo en todas las unidades NAL extraídas o agregadas.

El indicador idr se establecerá en el valor más alto del campo en todas las unidades NAL extraídas o agregadas.

priority_id, temporal_id, dependency_id y quality_id se establecerán en los valores más bajos de los campos, respectivamente, en todas las unidades NAL extraídas o agregadas.

discardable_flag se establecerá en 1 si y solo si todas las unidades NAL extraídas o agregadas tienen el discardable_flag establecido en 1, y se establecerá en 0 en caso contrario.

el indicador de salida debería establecerse en 1 si al menos una de las unidades NAL agregadas o extraídas tiene este indicador establecido en 1, y debería establecerse en 0 en caso contrario.

use_ref_base_pic_flag se establecerá en 1 si y solo si al menos una de las unidades VCL NAL extraídas o agregadas tiene el use_ref_base_pic_flag establecido en 1, y se establecerá en 0 en caso contrario.

no_inter_layer_pred_flag se establecerá en 1 si y solo si todas las unidades VCL NAL extraídas o agregadas tienen el no_inter_layer_pred_flag establecido en 1, y se establecerá en 0 en caso contrario.

[0254] Si el conjunto de unidades NAL extraídas o agregadas está vacío, entonces cada uno de estos campos toma un valor de conformidad con la descripción de capa correlacionada.

NOTA: Los agregadores podrían agrupar unidades NAL con información de escalabilidad diferente.

NOTA: Los agregadores podrían usarse para agrupar unidades NAL que pertenecen a un nivel de escalabilidad que no puede ser señalado por la cabecera de la unidad NAL (por ejemplo, unidades NAL que pertenecen a una zona de interés). La descripción de dichos agregadores se puede hacer con la descripción de capa y los grupos de correlación de unidades NAL. En este caso, más de un agregador con la misma información de escalabilidad pueden estar presentes en una muestra.

NOTA: Si múltiples pistas escalables hacen referencia a los mismos datos de medios, entonces un agregador debería agrupar unidades NAL con información de escalabilidad idéntica solo. Esto asegura que cada una de las pistas pueda acceder al patrón resultante.

NOTA: Si no existe ninguna unidad NAL de una capa particular en una unidad de acceso, entonces puede existir un agregador vacío (en el que la longitud del agregador solo incluye la cabecera, y additional_bytes es cero).

A.5 Valores de cabecera de unidad NAL para MVC

[0255] Tanto los agregadores como los extractores usan la ampliación MVC de cabecera de unidad NAL. Las unidades NAL extraídas por un extractor o agregadas por un agregador son todas las unidades NAL a las que se hace referencia o que se incluyen mediante inspección recursiva del contenido de las unidades NAL de agregador o extractor.

[0256] Los campos nal_ref_idc, non_idr_flag, priority_id, view_id, temporal_id, anchor_pic_flag e inter_view_flag tomarán los siguientes valores:

nal_ref_idc se establecerá en el valor más alto del campo en todas las unidades NAL agregadas o extraídas.

non_idr_flag se establecerá en el valor más bajo del campo en todas las unidades NAL agregadas o extraídas.

priority_id y el ID temporal se establecerán en los valores más bajos de los campos, respectivamente, en todas las unidades NAL agregadas o extraídas.

El ID de vista se establecerá en el valor de ID de vista de la unidad VCL NAL con el índice de orden de vista más bajo entre todas las unidades VCL NAL agregadas o extraídas.

anchor_pic_flag y el indicador entre vistas se establecerán en el valor más alto de los campos, respectivamente, en todas las unidades VCL NAL agregadas o extraídas.

[0257] Si el conjunto de unidades NAL extraídas o agregadas está vacío, entonces cada uno de estos campos toma un valor de conformidad con la descripción de capa correlacionada.

A.6 Valores de cabecera de unidad NAL para ISO/IEC 23008-2

[0258] Tanto los agregadores como los extractores usan la cabecera de unidad NAL como se especifica en ISO/IEC 23008-2. Las unidades NAL extraídas por un extractor o agregadas por un agregador son todas las unidades NAL a las que se hace referencia o que se incluyen mediante inspección recursiva del contenido de las unidades NAL de agregador o extractor.

[0259] Los campos nuh_layer_id y nuh_temporal_id_plus1 se establecerán de la siguiente manera:

nuh_layer_id se establecerá en el valor más bajo del campo en todas las unidades NAL agregadas o extraídas.

nuh temporal id plus1 se establecerá en el valor más bajo del campo en todas las unidades NAL agregadas o extraídas.

[0260] En un ejemplo alternativo, se define una nueva estructura, tabla o grupo de muestras para documentar todas las unidades de acceso IRAP como se define en el Anexo F de MV-HEVC WD5 o SHVC WD3. De forma alternativa, la nueva estructura, tabla o grupo de muestras se define para documentar todas las unidades de acceso IRAP como se define en el Anexo F de MV-HEVC WD5 o SHVC WD3, pero excluyendo las unidades de acceso donde todas las imágenes codificadas son imágenes IRAP. En otro ejemplo alternativo, la entrada de grupo de muestras de muestras de sincronización SyncSampleEntry se redefine para incluir un sync_flag alineado en uno de los bits reservados que especifica que todas las imágenes de la muestra que pertenecen a este grupo son imágenes IDR, imágenes CRA o imágenes BLA. En otro ejemplo alternativo, se define un formato de archivo común para SHVC y MV-HEVC que incluye todos los aspectos comunes de los formatos de archivo SHVC y MV-HEVC, y solo se redefinen los formatos de archivo SHVC y MV-HEVC para incluir solo los aspectos relacionados solo con esa ampliación. En otro ejemplo alternativo, se definen unas entradas de muestra de metadatos SHVC SHVCMetadataSampleEntry y SHVCMetadataSampleConfigBox, y también se define un tipo de sentencia de muestra de metadatos scalabilityInfoSHVCStatement.

[0261] La figura 2 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 estar configurado para facilitar unos tipos de datos de vídeo de vista única, multivista, escalable, 3D y otros. El codificador de vídeo 20 puede estar configurado para facilitar vídeo a una entidad de procesamiento de posprocesamiento 27. La entidad de procesamiento de posprocesamiento 27 pretende representar un ejemplo de una entidad de vídeo, como un dispositivo MANE o de empalme/edición, que puede procesar datos de vídeo codificados del codificador de vídeo 20. En algunos casos, la entidad de procesamiento de posprocesamiento puede ser un ejemplo de entidad de red. En algunos sistemas de codificación de vídeo, la entidad de posprocesamiento 27 y el codificador de vídeo 20 pueden formar parte de dispositivos separados, mientras que, en otros casos, la funcionalidad descrita con respecto a la entidad de posprocesamiento 27 puede ser desempeñada por el mismo dispositivo que comprende el decodificador de vídeo 20. La entidad de posprocesamiento 27 puede ser un dispositivo de vídeo. En algunos ejemplos, la entidad de posprocesamiento 27 puede ser la misma que el dispositivo de generación de archivos 34 de la figura 1.

[0262] El codificador de vídeo 20 puede realizar intracodificación e intercodificación de bloques de vídeo dentro de sectores de vídeo. La intracodificación se basa en la predicción espacial para reducir o eliminar la redundancia espacial en el vídeo dentro de una trama o imagen de vídeo dada. La intercodificación se basa en la predicción temporal para reducir o eliminar la redundancia temporal en el vídeo dentro de tramas o imágenes adyacentes de una secuencia de vídeo. El modo intra (modo I) puede referirse a cualquiera de varios modos de compresión espacial. Los modos inter, tales como la predicción unidireccional (modo P) o la predicción bidireccional (modo B), pueden referirse a cualquiera de varios modos de compresión temporal.

[0263] En el ejemplo de la figura 2, el codificador de vídeo 20 incluye una unidad de división 35, una unidad de procesamiento de predicción 41, una unidad de filtro 63, una memoria de imágenes de referencia 64, un sumador 50, una unidad de procesamiento de transformada 52, una unidad de cuantificación 54 y una unidad de codificación de entropía 56. La unidad de procesamiento de predicción 41 incluye una unidad de estimación de movimiento 42, una unidad de compensación de movimiento 44 y una unidad de procesamiento de intrapredicción 46. Para la reconstrucción de bloques de vídeo, el codificador de vídeo 20 incluye también una unidad de cuantificación inversa 58, una unidad de procesamiento de transformada inversa 60 y un sumador 62. La unidad de filtro 63 está destinada a representar uno o más filtros de bucle tales como un filtro de eliminación de bloques, un filtro de bucle adaptativo (ALF) y un filtro de desplazamiento adaptativo de muestras (SAO). Aunque la unidad de filtro 63 que se muestra en la figura 2 es un filtro de bucle, en otras configuraciones, la unidad de filtro 63 puede implementarse como un filtro posbucle.

[0264] Una memoria de datos de vídeo del decodificador de vídeo 20 puede almacenar datos de vídeo que los componentes del codificador de vídeo 20 van a codificar. Los datos de vídeo almacenados en la memoria de datos de vídeo se pueden obtener, por ejemplo, a partir de la fuente de vídeo 18. La memoria de imágenes de referencia 64 puede ser una memoria de imágenes de referencia que almacena datos de vídeo de referencia para su uso en la codificación de datos de vídeo por el codificador de vídeo 20, por ejemplo, en modos de intracodificación o de intercodificación. La memoria de datos de vídeo y la memoria imágenes de referencia 64 pueden estar formadas por cualquiera de una variedad de dispositivos de memoria, tales como una memoria dinámica de acceso aleatorio (DRAM), que incluye DRAM síncrona (SDRAM), RAM magnetorresistiva (MRAM), RAM resistiva (RRAM) u otros tipos de dispositivos de memoria. El mismo dispositivo de memoria o unos dispositivos de memoria separados pueden proporcionar la memoria de datos de vídeo y la memoria de imágenes de referencia 64. En diversos ejemplos, la memoria de datos de vídeo puede estar en el chip con otros componentes del codificador de vídeo 20, o fuera del chip con respecto a esos componentes.

[0265] Como se representa en la figura 2, el codificador de vídeo 20 recibe datos de vídeo, y la unidad de división 35 divide los datos en bloques de vídeo. Esta división también puede incluir la división en sectores, mosaicos u otras unidades mayores, así como la división de bloques de vídeo, por ejemplo, de acuerdo con una estructura de árbol

cuaternario de LCU y CU. El codificador de vídeo 20 ilustra, en general, los componentes que codifican bloques de vídeo dentro de un sector de vídeo que se va a codificar. El sector se puede dividir en múltiples bloques de vídeo (y, posiblemente, en conjuntos de bloques de vídeo denominados mosaicos). La unidad de procesamiento 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). La unidad de procesamiento de predicción 41 puede proporcionar el bloque intracodificado o intercodificado resultante al sumador 50 para generar datos de bloque residuales y al sumador 62 para reconstruir el bloque codificado para su uso como imagen de referencia.

[0266] La unidad de procesamiento de intrapredicción 46, situada dentro de la unidad de procesamiento 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 sector que el bloque actual que se va a codificar, para proporcionar compresión espacial. La unidad de estimación de movimiento 42 y la unidad de compensación de movimiento 44 de la unidad de procesamiento de predicción 41 realizan la codificación interpredictiva del bloque de vídeo actual con respecto a uno o más bloques predictivos de una o más imágenes de referencia para proporcionar una compresión temporal.

[0267] La unidad de estimación de movimiento 42 puede estar configurada para determinar el modo de interpredicción para un sector de vídeo de acuerdo con un patrón predeterminado para una secuencia de vídeo. El patrón predeterminado puede designar sectores de vídeo de la secuencia como sectores P, sectores B o sectores GPB. La unidad de estimación de movimiento 42 y la unidad de compensación de movimiento 44 pueden estar sumamente integradas, pero se ilustran por separado con fines conceptuales. La estimación del movimiento, realizada por la unidad de estimación de movimiento 42, es el proceso de generación de vectores de movimiento, que estiman el movimiento para los bloques de vídeo. Un vector de movimiento, por ejemplo, puede indicar el desplazamiento de una PU de un bloque de vídeo de una trama o imagen de vídeo actual en relación con un bloque predictivo de una imagen de referencia.

[0268] Un bloque predictivo es un bloque que resulta corresponder estrechamente con la PU del bloque de vídeo que se va a codificar en términos de diferencia de píxel, que puede determinarse mediante una suma de una diferencia absoluta (SAD), una suma de diferencia al cuadrado (SSD) u otras métricas de diferencia. En algunos ejemplos, el codificador de vídeo 20 puede calcular los valores para posiciones de píxeles de subentero de imágenes de referencia almacenadas en la memoria de imágenes de referencia 64. Por ejemplo, el codificador de vídeo 20 puede interpolar valores de posiciones de un cuarto de píxel, posiciones de un octavo de píxel u otras posiciones fraccionarias de píxel de la imagen de referencia. Por lo tanto, la unidad de estimación de movimiento 42 puede realizar una búsqueda de movimiento relativa a las posiciones de píxel completo y las posiciones de píxel fraccionario, y facilitar un vector de movimiento con una precisión de píxel fraccionario.

[0269] La unidad de estimación de movimiento 42 calcula un vector de movimiento para una PU de un bloque de vídeo de un sector intracodificado comparando la posición de la PU con la posición de un bloque predictivo de una imagen de referencia. La imagen de referencia puede seleccionarse a partir de una primera lista (lista 0) de imágenes de referencia o una segunda lista (lista 1) de imágenes de referencia, cada una de las cuales identifica una o más imágenes de referencia almacenadas en una memoria de imágenes de referencia 64. La unidad de estimación de movimiento 42 envía el vector de movimiento calculado a la unidad de codificación de entropía 56 y a la unidad de compensación de movimiento 44.

[0270] La compensación de movimiento, realizada por la unidad de compensación de movimiento 44, puede implicar extraer o generar el bloque predictivo basándose en el vector de movimiento determinado mediante estimación de movimiento, realizando posiblemente interpolaciones hasta una precisión de subpíxel. Tras recibir el vector de movimiento para la PU del bloque de vídeo actual, la unidad de compensación de movimiento 44 puede localizar el bloque predictivo al que apunta el vector de movimiento en una de las listas de imágenes de referencia. El codificador de vídeo 20 puede formar un bloque de vídeo residual restando valores de píxel del bloque predictivo a los valores de píxel del bloque de vídeo actual que se está codificando, formando valores de diferencia de píxel. Los valores de diferencia de píxel forman datos residuales para el bloque, y pueden incluir componentes de diferencia tanto de luma como de croma. El sumador 50 representa el componente o los componentes que realizan esta operación de resta. La unidad de compensación de movimiento 44 también puede generar elementos sintácticos asociados a los bloques de vídeo y al sector de vídeo para su uso por el descodificador de vídeo 30 en la descodificación de los bloques de vídeo del sector de vídeo.

[0271] La unidad de procesamiento de intrapredicción 46 puede realizar la intrapredicción de un bloque actual, como alternativa a la interpredicción llevada a cabo por la unidad de estimación de movimiento 42 y la unidad de compensación de movimiento 44, como se ha descrito anteriormente. En particular, la unidad de procesamiento de intrapredicción 46 puede determinar un modo de intrapredicción que se va a usar para codificar un bloque actual. En algunos ejemplos, la unidad de procesamiento de intrapredicción 46 puede codificar un bloque actual usando diversos modos de intrapredicción, por ejemplo, durante pasadas de codificación independientes, y la unidad de procesamiento de intrapredicción 46 (o la unidad de selección de modo 40, en algunos ejemplos) puede seleccionar un modo de intrapredicción adecuado para su uso a partir de los modos probados. Por ejemplo, la unidad de procesamiento 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 tiene las mejores características de velocidad-distorsión entre los modos probados. El análisis de velocidad-distorsión determina, en general, una cantidad de distorsión (o error) entre un bloque codificado y un bloque original no codificado que se ha codificado para generar el bloque codificado, así como una velocidad de bits (es decir, un número de bits) usada para generar el bloque codificado. La unidad de procesamiento 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.

[0272] En cualquier caso, tras seleccionar un modo de intrapredicción para un bloque, la unidad de procesamiento de intrapredicción 46 puede proporcionar información que indica el modo de intrapredicción seleccionado para el bloque, a la unidad de codificación de entropía 56. La unidad de codificación de entropía 56 puede codificar la información que indica el modo de intrapredicción seleccionado de acuerdo con las técnicas de esta divulgación. El codificador de vídeo 20 puede incluir, en el flujo de bits transmitido, datos de configuración que pueden incluir una pluralidad de tablas de índices de modo de intrapredicción y una pluralidad de tablas de índices de modo de intrapredicción modificadas (también denominadas tablas de correlación de palabras de código), definiciones de contextos de codificación para diversos bloques e indicaciones de un modo de intrapredicción más probable, una tabla de índices de modo de intrapredicción y una tabla de índices de modo de intrapredicción modificada para su uso para cada uno de los contextos.

[0273] Después de que la unidad de procesamiento de predicción 41 haya generado el bloque predictivo para el bloque de vídeo actual, ya sea mediante interpredicción o intrapredicción, el codificador de vídeo 20 puede formar un bloque de vídeo residual restando el bloque predictivo al bloque de vídeo actual. Los datos de vídeo residuales del bloque residual pueden estar incluidos en una o más TU y aplicarse a la unidad de procesamiento de transformada 52. La unidad de procesamiento de transformada 52 transforma los datos de vídeo residuales en coeficientes de transformada residuales usando una transformada, tal como una transformada de coseno discreta (DCT) o una transformada conceptualmente similar. La unidad de procesamiento de transformada 52 puede convertir los datos de vídeo residuales de un dominio de píxel a un dominio de transformada, tal como un dominio de frecuencia.

[0274] La unidad de procesamiento de transformada 52 puede enviar los coeficientes de transformada resultantes a la unidad de cuantificación 54. La unidad de cuantificación 54 cuantifica los coeficientes de transformada para reducir todavía más la velocidad de bits. El proceso de cuantificación puede reducir la profundidad de bits asociada a algunos, o a la totalidad, de los coeficientes. El grado de cuantificación se puede modificar ajustando un parámetro de cuantificación. En algunos ejemplos, la unidad 54 de cuantificación puede realizar, a continuación, una exploración de la matriz que incluye los coeficientes de transformada cuantificados. De forma alternativa, la unidad de codificación de entropía 56 puede realizar la exploración.

[0275] Después de la cuantificación, la unidad de codificación de entropía 56 puede realizar la codificación de entropía de los elementos sintácticos que representan los coeficientes de transformada cuantificados. Por ejemplo, la unidad de codificación de 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 basada en la sintaxis (SBAC), una codificación de entropía por división de intervalos de probabilidad (PIPE) u otros procedimientos o técnicas de codificación de entropía. Tras la codificación de entropía por la unidad de codificación de entropía 56, el flujo de bits codificado se puede transmitir al decodificador de vídeo 30, o archivar para su posterior transmisión o recuperación por el decodificador de vídeo 30. La unidad de codificación de entropía 56 también puede realizar la codificación de entropía de los vectores de movimiento y los otros elementos sintácticos para el sector de vídeo actual que se está codificando.

[0276] La unidad de cuantificación inversa 58 y la unidad de procesamiento de transformada inversa 60 aplican una cuantificación inversa y una transformada inversa, respectivamente, para reconstruir el bloque residual en el dominio del píxel, para su posterior uso como bloque de referencia de una imagen de referencia. La unidad de compensación de movimiento 44 puede calcular un bloque de referencia añadiendo el bloque residual a un bloque predictivo de una de las imágenes de referencia de una de las listas de imágenes de referencia. La unidad de compensación de movimiento 44 también puede aplicar uno o más filtros de interpolación al bloque residual reconstruido para calcular valores de píxel de subentero para su uso en la estimación de movimiento. El sumador 62 añade el bloque residual reconstruido al bloque de predicción con compensación de movimiento generado por la unidad de compensación de movimiento 44 para generar un bloque de referencia para su almacenamiento en la memoria de imágenes de referencia 64. La unidad de estimación de movimiento 42 y la unidad de compensación de movimiento 44 pueden usar el bloque de referencia como bloque de referencia para realizar la interpredicción de un bloque en una trama o imagen de vídeo subsiguiente.

[0277] El codificador de vídeo 20 representa un ejemplo de codificador de vídeo configurado para generar datos de vídeo que pueden almacenarse usando las técnicas de formato de archivo descritas en esta divulgación.

[0278] La figura 3 es un diagrama de bloques que ilustra un ejemplo de decodificador de vídeo 30 que puede implementar las técnicas descritas en esta divulgación. El decodificador de vídeo 30 puede estar configurado para

descodificar datos de vídeo de vista única, multivista, escalable, 3D y de otros tipos. En el ejemplo de la figura 3, el descodificador de vídeo 30 incluye una unidad de descodificación de entropía 80, una unidad de procesamiento de predicción 81, una unidad de procesamiento de cuantificación inversa 86, una unidad de procesamiento de transformada inversa 88, un sumador 90, una unidad de filtro 91 y una memoria de imágenes de referencia 92. La

[0279] Una memoria intermedia de imágenes codificadas (CPB) 79 puede recibir y almacenar datos de vídeo codificados (por ejemplo, unidades NAL) de un flujo de bits. Los datos de vídeo almacenados en la CPB 79 pueden obtenerse, por ejemplo, a partir del enlace 16, por ejemplo, desde una fuente de vídeo local, tal como una cámara, a través de una transmisión por red alámbrica o inalámbrica de datos de vídeo, o accediendo a medios físicos de almacenamiento de datos. La CPB 79 puede formar una memoria de datos de vídeo que almacena datos de vídeo codificados de un flujo de bits de vídeo codificado. La CPB 79 puede ser una memoria de imágenes de referencia que almacena datos de vídeo de referencia para su uso en la descodificación de datos de vídeo por el descodificador de vídeo 30, por ejemplo, en modos de intracodificación o intercodificación. La CPB 79 y la memoria de imágenes de referencia 92 pueden estar formadas por cualesquiera de una variedad de dispositivos de memoria, tales como memoria dinámica de acceso aleatorio (DRAM), que incluye DRAM síncrona (SDRAM), RAM magnetorresistiva (MRAM), RAM resistiva (RRAM) u otros tipos de dispositivos de dispositivos de memoria. El mismo dispositivo de memoria o unos dispositivos de memoria independientes pueden proporcionar la CPB 79 y la memoria de imágenes de referencia 92. En diversos ejemplos, la CPB 79 puede estar en un chip con otros componentes del descodificador de vídeo 30, o fuera de chip con respecto a esos componentes.

[0280] Durante el proceso de descodificación, el descodificador de vídeo 30 recibe un flujo de bits de vídeo codificado que representa bloques de vídeo de un sector de vídeo codificado y elementos sintácticos asociados, desde el codificador de vídeo 20. El descodificador de vídeo 30 puede recibir el flujo de bits de vídeo codificado desde una entidad de red 29. La entidad de red 29 puede ser, por ejemplo, un servidor, un MANE, un editor/empalmador de vídeo u otro dispositivo similar configurado para implementar una o más de las técnicas descritas anteriormente. La entidad de red 29 puede incluir o no un codificador de vídeo, tal como el codificador de vídeo 20. La entidad de red 29 puede implementar algunas de las técnicas descritas en esta divulgación, antes de que la entidad de red 29 transmita el flujo de bits de vídeo codificado al descodificador de vídeo 30. En algunos sistemas de descodificación de vídeo, la entidad de red 29 y el descodificador de vídeo 30 pueden formar parte de dispositivos separados, mientras que en otros casos el mismo dispositivo que comprende el descodificador de vídeo 30 puede desempeñar la funcionalidad descrita con respecto a la entidad de red 29. Se puede considerar que la entidad de red 29 es un dispositivo de vídeo. Además, en algunos ejemplos, la entidad de red 29 es el dispositivo de generación de archivos 34 de la figura 1.

[0281] La unidad de descodificación de entropía 80 del descodificador de vídeo 30 realiza la descodificación de entropía de elementos sintácticos particulares del flujo de vídeo para generar coeficientes cuantificados, vectores de movimiento y otros elementos sintácticos. La unidad de descodificación de entropía 80 envía los vectores de movimiento y otros elementos sintácticos a la unidad de procesamiento de predicción 81. El descodificador de vídeo 30 puede recibir los elementos sintácticos en el nivel de sector de vídeo y/o el nivel de bloque de vídeo.

[0282] Cuando el sector de vídeo se codifica como un sector intracodificado (I), la unidad de procesamiento de intrapredicción 84 de la unidad de procesamiento de predicción 81 puede generar datos de predicción para un bloque de vídeo del sector de vídeo actual basándose en un modo de intrapredicción señalado, y datos de bloques previamente descodificados de la trama o imagen actual. Cuando la trama de vídeo se codifica como un sector intercodificado (es decir, B, P o GPB), la unidad de compensación de movimiento 82 de la unidad de procesamiento de predicción 81 genera bloques predictivos para un bloque de vídeo del sector de vídeo actual basándose en los vectores de movimiento y en otros elementos sintácticos recibidos desde la unidad de descodificación de entropía 80. Los bloques predictivos se pueden generar a partir de una de las imágenes de referencia de una de las listas de imágenes de referencia. El descodificador de vídeo 30 puede construir las listas de tramas de referencia, lista 0 y lista 1, mediante técnicas de construcción predeterminadas, basándose en imágenes de referencia almacenadas en la memoria de imágenes de referencia 92.

[0283] La unidad de compensación de movimiento 82 determina la información de predicción para un bloque de vídeo del sector de vídeo actual, analizando los vectores de movimiento y otros elementos sintácticos, y usa la información de predicción para generar los bloques predictivos para el bloque de vídeo actual que se está descodificando. Por ejemplo, la unidad de compensación de movimiento 82 usa algunos de los elementos sintácticos recibidos para determinar un modo de predicción (por ejemplo, intrapredicción o interpredicción) usado para codificar los bloques de vídeo del sector de vídeo, un tipo de sector de interpredicción (por ejemplo, sector B, sector P o sector GPB), información de construcción para una o más de las listas de imágenes de referencia para el sector, vectores de movimiento para cada bloque de vídeo intercodificado del sector, un estado de interpredicción para cada bloque de vídeo intercodificado del sector y otro tipo de información para descodificar los bloques de vídeo en el sector de vídeo actual.

[0284] La unidad de compensación de movimiento 82 también puede realizar la interpolación basándose en unos filtros de interpolación. La unidad de compensación de movimiento 82 puede usar filtros de interpolación como los usados por el codificador de vídeo 20 durante la codificación de los bloques de vídeo, para calcular valores interpolados para píxeles de subentero de los bloques de referencia. En este caso, la unidad de compensación de movimiento 82 puede determinar los filtros de interpolación usados por el codificador de vídeo 20 a partir de los elementos sintácticos recibidos y puede usar los filtros de interpolación para generar bloques predictivos.

[0285] La unidad de cuantificación inversa 86 realiza la cuantificación inversa, es decir, la descuantificación, de los coeficientes de transformada cuantificados proporcionados en el flujo de bits y descodificados por la unidad de descodificación de entropía 80. El proceso de cuantificación inversa puede incluir el uso de un parámetro de cuantificación calculado por el codificador de vídeo 20 para cada bloque de vídeo del sector de vídeo con el fin de determinar un grado de cuantificación y, asimismo, un grado de cuantificación inversa que se debería aplicar. La unidad de procesamiento de transformada inversa 88 aplica una transformada inversa, por ejemplo una DCT inversa, una transformada inversa de enteros o un proceso de transformada inversa conceptualmente similar, a los coeficientes de transformada, con el fin de generar bloques residuales en el dominio del píxel.

[0286] Una vez que la unidad de compensación de movimiento 82 ha generado el bloque predictivo para el bloque de vídeo actual basándose en los vectores de movimiento y en otros elementos sintácticos, el descodificador de vídeo 30 forma un bloque de vídeo descodificado sumando los bloques residuales procedentes de la unidad de procesamiento de transformada inversa 88 a los correspondientes bloques predictivos generados por la unidad de compensación de movimiento 82. El sumador 90 representa el componente o los componentes que realizan esta operación de suma. Si se desea, también pueden usarse filtros de bucle (ya sea en el bucle de codificación o después del bucle de codificación) para suavizar las transiciones de píxeles o mejorar de otro modo la calidad del vídeo. La unidad de filtro 91 está destinada a representar uno o más filtros de bucle tales como un filtro de eliminación de bloques, un filtro de bucle adaptativo (ALF) y un filtro de desplazamiento adaptativo de muestras (SAO). Aunque la unidad de filtro 91 que se muestra en la figura 3 es un filtro de bucle, en otras configuraciones, la unidad de filtro 91 puede implementarse como un filtro posbucle. Los bloques de vídeo descodificados de una trama o imagen determinada se almacenan a continuación en la memoria de imágenes de referencia 92, que almacena imágenes de referencia usadas para una subsiguiente compensación de movimiento. La memoria de imágenes de referencia 92 almacena también vídeo descodificado para su presentación posterior en un dispositivo de visualización, tal como el dispositivo de visualización 32 de la figura 1.

[0287] El descodificador de vídeo 30 de la figura 3 representa un ejemplo de descodificador de vídeo configurado para descodificar datos de vídeo que pueden almacenarse mediante las técnicas de formato de archivo descritas en esta divulgación.

[0288] La figura 4 es un diagrama de bloques que ilustra un ejemplo de conjunto de dispositivos que forman parte de la red 100. En este ejemplo, la red 100 incluye unos dispositivos de encaminamiento 104A, 104B (dispositivos de encaminamiento 104) y un dispositivo de transcodificación 106. Los dispositivos de encaminamiento 104 y el dispositivo de transcodificación 106 están concebidos para representar un pequeño número de dispositivos que pueden formar parte de la red 100. Otros dispositivos de red, tales como conmutadores, concentradores, pasarelas, cortafuegos, puentes y otros dispositivos de ese tipo también pueden estar incluidos en la red 100. Además, pueden proporcionarse dispositivos de red adicionales a lo largo de un trayecto de red entre un dispositivo servidor 102 y un dispositivo cliente 108. El dispositivo servidor 102 puede corresponder a un dispositivo de origen 12 (figura 1), mientras que el dispositivo cliente 108 puede corresponder a un dispositivo de destino 14 (figura 1), en algunos ejemplos.

[0289] En general, los dispositivos de encaminamiento 104 implementan uno o más protocolos de encaminamiento para intercambiar datos de red a través de la red 100. En algunos ejemplos, los dispositivos de encaminamiento 104 pueden estar configurados para realizar operaciones de *proxy* o de memoria caché. Por lo tanto, en algunos ejemplos, los dispositivos de encaminamiento 104 pueden denominarse dispositivos *proxy*. En general, los dispositivos de encaminamiento 104 ejecutan protocolos de encaminamiento para descubrir rutas a través de la red 100. Al ejecutar dichos protocolos de encaminamiento, un dispositivo de encaminamiento 104B puede descubrir una ruta de red desde sí mismo hasta el dispositivo servidor 102, a través de un dispositivo de encaminamiento 104A.

[0290] Unos dispositivos de red tales como dichos dispositivos de encaminamiento 104 y dispositivo de transcodificación 106 pueden implementar las técnicas de esta divulgación, aunque un dispositivo cliente 108 también las puede implementar. De esta manera, los dispositivos de encaminamiento 104, el dispositivo de transcodificación 106 y el dispositivo cliente 108 representan ejemplos de dispositivos configurados para realizar las técnicas de esta divulgación. Por otro lado, los dispositivos de la figura 1, y el codificador 20 ilustrado en la figura 2 y el descodificador 30 ilustrado en la figura 3 son también ejemplos de dispositivos que pueden estar configurados para realizar una o más de las técnicas de esta divulgación.

[0291] La figura 5 es un diagrama conceptual que ilustra un ejemplo de estructura de un archivo 300 de acuerdo con una o más técnicas de esta divulgación. En el ejemplo de la figura 5, el archivo 300 incluye una caja de película 302 y una pluralidad de cajas de datos de medios 304. Aunque en el ejemplo de la figura 5 se ilustran en el mismo archivo, en otros ejemplos, la caja de película 302 y la caja de datos de medios 304 pueden estar en archivos separados. Como

se ha indicado anteriormente, una caja puede ser un elemento básico orientado a objetos definido por un identificador de tipo exclusivo y una longitud. Por ejemplo, una caja puede ser la estructura sintáctica elemental en el ISO BMFF, que incluye un tipo de caja codificada de cuatro caracteres, un recuento de bytes de la caja y una carga útil.

[0292] La caja de película 302 puede contener metadatos para las pistas del archivo 300. Cada pista del archivo 300 puede comprender un flujo continuo de datos de medios. Cada una de las cajas de datos de medios 304 puede incluir una o más muestras 305. Cada una de las muestras 305 puede comprender una unidad de acceso de audio o vídeo. Como se describe en otra parte de esta divulgación, cada unidad de acceso puede comprender múltiples imágenes codificadas en codificación multivista (por ejemplo, MV-HEVC y 3D-HEVC) y codificación de vídeo escalable (por ejemplo, SHVC). Por ejemplo, una unidad de acceso puede incluir una o más imágenes codificadas para cada capa.

[0293] Además, en el ejemplo de la figura 5, la caja de película 302 incluye una caja de pista 306. La caja de pista 306 puede contener metadatos para una pista del archivo 300. En otros ejemplos, la caja de película 302 puede incluir múltiples cajas de pista para diferentes pistas del archivo 300. La caja de pista 306 incluye una caja de medios 307. La caja de medios 307 puede contener todos los objetos que declaran información sobre los datos de medios de la pista. La caja de medios 307 incluye una caja de información de medios 308. La caja de información de medios 308 puede contener todos los objetos que declaran información característica de los medios de la pista. La caja de información de medios 308 incluye una caja de tabla de muestras 309. La caja de tabla de muestras 309 puede especificar metadatos específicos de muestra.

[0294] En el ejemplo de la figura 5, la caja de tabla de muestras 309 incluye una caja SampleToGroup 310 y una caja de SampleGroupDescription 312. En otros ejemplos, la caja de tabla de muestras 309 puede incluir otras cajas, además de la caja SampleToGroup 310 y la caja SampleGroupDescription 312, y/o puede incluir múltiples cajas SampleToGroup y cajas SampleGroupDescription. La caja SampleToGroup 310 puede correlacionar muestras (por ejemplo, muestras particulares de las muestras 305) con un grupo de muestras. La caja SampleGroupDescription 312 puede especificar una propiedad compartida por las muestras del grupo de muestras (es decir, el grupo de muestras). Además, la caja de tabla de muestras 309 puede incluir una pluralidad de cajas de entradas de muestra 311. Cada una de las cajas de entradas de muestra 311 puede corresponder a una muestra del grupo de muestras. En algunos ejemplos, las cajas de entradas de muestra 311 son ejemplos de una clase de entrada de muestra accesible aleatoriamente que amplía una clase de descripción de grupo de muestras base como se define en la sección 9.5.5.1.2 anterior.

[0295] De acuerdo con una o más técnicas de esta divulgación, la caja SampleGroupDescription 312 puede especificar que cada muestra del grupo de muestra contiene al menos una imagen IRAP. De esta manera, el dispositivo de generación de archivos 34 puede generar un archivo que comprende una caja de pista 306 que contiene metadatos para una pista del archivo 300. Los datos de medios para la pista comprenden una secuencia de muestras 305. Cada una de las muestras puede ser una unidad de acceso de vídeo de datos de vídeo multicapa (por ejemplo, datos de vídeo SHVC, MV-HEVC o 3D-HEVC). Además, como parte de la generación del archivo 300, el dispositivo de generación de archivos 34 puede generar, en el archivo 300, una caja adicional (es decir, una caja de tabla de muestras 309) que documenta todas las muestras 305 que contienen al menos una imagen IRAP. En otras palabras, la caja adicional identifica todas las muestras 305 que contienen al menos una imagen IRAP. En el ejemplo de la figura 5, la caja adicional define un grupo de muestras que documenta (por ejemplo, identifica) todas las muestras 305 que contienen al menos una imagen IRAP. En otras palabras, la caja adicional especifica que las muestras 305 que contienen al menos una imagen IRAP pertenecen a un grupo de muestras.

[0296] Además, de acuerdo con una o más técnicas de esta divulgación, cada una de las cajas de entradas de muestra 311 puede incluir un valor (por ejemplo, `all_pics_are_IRAP`) que indica si todas las imágenes codificadas de la muestra correspondiente son imágenes IRAP. En algunos ejemplos, el valor que es igual a 1 especifica que no todas las imágenes codificadas de la muestra son imágenes IRAP. Cuando el valor es igual a 0, especifica que no se requiere que cada imagen codificada de cada muestra del grupo de muestras sea una imagen IRAP.

[0297] En algunos ejemplos, cuando no todas las imágenes codificadas de una muestra particular son imágenes IRAP, el dispositivo de generación de archivos 34 puede incluir, en una de las cajas de entradas de muestra 311 para la muestra particular, un valor (por ejemplo, `num_IRAP_pics`) que indica un número de imágenes IRAP de la muestra particular. Adicionalmente, el dispositivo de generación de archivos 34 puede incluir, en la entrada de muestra para la muestra en particular, unos valores que indican los identificadores de capa de las imágenes IRAP de la muestra particular. El dispositivo de generación de archivos 34 también puede incluir, en la entrada de muestra para la muestra particular, un valor que indica un tipo de unidad NAL de las unidades VCL NAL de las imágenes IRAP de la muestra particular.

[0298] Además, en el ejemplo de la figura 5, la caja de tabla de muestras 309 incluye una caja de información de submuestra 314. Aunque el ejemplo de la figura 5 solo muestra una caja de información de submuestra, la caja de tabla de muestras 309 puede incluir múltiples cajas de información de submuestra. En general, una caja de información de submuestra está diseñada para contener información de submuestra. Una submuestra es un intervalo contiguo de bytes de una muestra. La norma ISO/CEI 14496-12 indica que la definición específica de una submuestra debe facilitarse para un sistema de codificación determinado, como H.264/AVC o HEVC.

[0299] La sección 8.4.8 de ISO/IEC 14496-15 especifica una definición de una submuestra para HEVC. En particular, la sección 8.4.8 de ISO/IEC 14496-15 especifica que para el uso de la caja de información de submuestra (sección 8.7.7 de ISO/IEC 14496-12) en un flujo HEVC, se define una submuestra sobre la base de un valor de un campo de indicadores de la caja de información de submuestra. De acuerdo con una o más técnicas de esta divulgación, si el campo de indicadores de la caja de información de submuestra 314 es igual a 5, una submuestra correspondiente a la caja de información de submuestra 314 contiene una imagen codificada y las unidades no VCL NAL asociadas. Las unidades no VCL NAL asociadas pueden incluir unidades NAL que contienen mensajes SEI aplicables a la imagen codificada y unidades NAL que contienen conjuntos de parámetros (por ejemplo, VPS, SPS, PPS, etc.) aplicables a la imagen codificada.

[0300] Por tanto, en un ejemplo, el dispositivo de generación de archivos 34 puede generar un archivo (por ejemplo, el archivo 300) que comprende una caja de pista (por ejemplo, la caja de pista 306) que contiene metadatos para una pista del archivo. En este ejemplo, los datos de medios para la pista comprenden una secuencia de muestras, cada una de las cuales es una unidad de acceso de vídeo de datos de vídeo multicapa (por ejemplo, datos de vídeo de SHVC, MV-HEVC o 3D-HEVC). Además, en este ejemplo, como parte del dispositivo de generación de archivos 34 que genera el archivo, el dispositivo de generación de archivos 34 puede generar, en el archivo, una caja de información de submuestra (por ejemplo, la caja de información de submuestra 314) que contiene indicadores que especifican un tipo de información de submuestra dada en la caja de información de submuestra. Cuando los indicadores tienen un valor particular, una submuestra correspondiente a la caja de información de submuestra contiene exactamente una imagen codificada y cero o más unidades no VCL NAL asociadas con la imagen codificada.

[0301] Además, de acuerdo con una o más técnicas de esta divulgación, si el campo de indicadores de la caja de información de submuestra 314 es igual a 0, la caja de información de submuestra 314 incluye además un valor DiscardableFlag, un valor NoInterLayerPredFlag, un valor LayerId y un valor TempId. Si el campo de indicadores de la caja de información de submuestra 314 es igual a 5, la caja de información de submuestra 314 puede incluir un valor DiscardableFlag, un valor VclNalUnitType, un valor LayerId, un valor TempId, un valor NoInterLayerPredFlag, un valor SubLayerRefNalUnitFlag y un valor reservado.

[0302] SubLayerRefNalUnitFlag igual a 0 indica que todas las unidades NAL de la submuestra son unidades VCL NAL de una imagen no de referencia de subcapa como se especifica en ISO/IEC 23008-2 (es decir, HEVC). SubLayerRefNalUnitFlag igual a 1 indica que todas las unidades NAL de la submuestra son unidades VCL NAL de una imagen de referencia de subcapa como se especifica en ISO/IEC 23008-2 (es decir, HEVC). Por tanto, cuando el dispositivo de generación de archivos 34 genera una caja de información de submuestra 314 y los indicadores tienen un valor particular (por ejemplo, 5), el dispositivo de generación de archivos 34 incluye, en la caja de información de submuestra 314, un indicador adicional que indica si todas las unidades NAL de la submuestra son unidades VCL NAL de una imagen de no de referencia de subcapa.

[0303] El valor DiscardableFlag indica un valor de un valor discardable_flag de las unidades VCL NAL de la submuestra. Como se especifica en la sección A.4 de ISO/IEC 14496-15, el valor del indicador descartable se establecerá en 1 si y solo si todas las unidades NAL extraídas o agregadas tienen el discardable_flag establecido en 1, y se establecerá en 0 en caso contrario. Una unidad NAL puede tener un discardable_flag establecido en 1 si un flujo de bits que contiene la unidad NAL puede descodificarse correctamente sin la unidad NAL. Por tanto, una unidad NAL puede ser «descartable» si un flujo de bits que contiene la unidad NAL puede descodificarse correctamente sin la unidad NAL. Todas las unidades VCL NAL de la submuestra tendrán el mismo valor discardable_flag. Por tanto, cuando el dispositivo de generación de archivos 34 genera una caja de información de submuestra 314 y los indicadores tienen un valor particular (por ejemplo, 5), el dispositivo de generación de archivos 34 incluye, en la caja de información de submuestra 314, un indicador adicional (por ejemplo, discardable_flag) que indica si todas las unidades VCL NAL de la submuestra son descartables.

[0304] El valor NoInterLayerPredFlag indica el valor del inter_layer_pred_enabled_flag de las unidades VCL NAL de la submuestra. El inter_layer_pred_enabled_flag se establecerá en 1 si y solo si todas las unidades VCL NAL extraídas o agregadas tienen el inter_layer_pred_enabled_flag establecido en 1, y se establecerá en 0 de lo contrario. Todas las unidades VCL NAL de la submuestra tendrán el mismo valor de inter_layer_pred_enabled_flag. Por tanto, cuando el dispositivo de generación de archivos 34 genera una caja de información de submuestra 314 y los indicadores tienen un valor particular (por ejemplo, 5), el dispositivo de generación de archivos 34 incluye, en la caja de información de submuestra 314, un valor adicional (por ejemplo, inter_layer_pred_enabled_flag) que indica si la predicción entre capas está habilitada para todas las unidades VCL NAL de la submuestra.

[0305] LayerId indica el valor nuh_layer_id de las unidades NAL de la submuestra. Todas las unidades NAL de la submuestra tendrán el mismo valor nuh_layer_id. Por tanto, cuando el dispositivo de generación de archivos 34 genera una caja de información de submuestra 314 y los indicadores tienen un valor particular (por ejemplo, 5), el dispositivo de generación de archivos 34 incluye, en la caja de información de submuestra 314, un valor adicional (por ejemplo, LayerId) que indica un identificador de capa de cada unidad NAL de la submuestra.

[0306] TempId indica el valor de TemporalId de las unidades NAL de la submuestra. Todas las unidades NAL de la submuestra tendrán el mismo valor de TemporalId. Por tanto, cuando el dispositivo de generación de archivos 34 genera una caja de información de submuestra 314 y los indicadores tienen un valor particular (por ejemplo, 5), el dispositivo de generación de archivos 34 incluye, en la caja de información de submuestra 314, un valor adicional (por ejemplo, TempId) que indica un identificador temporal de cada unidad NAL de la submuestra.

[0307] VclNalUnitType indica el elemento sintáctico nal_unit_type de las unidades VCL NAL de la submuestra. El elemento sintáctico nal_unit_type es un elemento sintáctico de una cabecera de unidad NAL de una unidad NAL. El elemento sintáctico nal_unit_type especifica el tipo de RBSP contenida en la unidad NAL. Todas las unidades VCL NAL de nal_unit_type de la submuestra tendrán el mismo valor nal_unit_type. Por tanto, cuando el dispositivo de generación de archivos 34 genera una caja de información de submuestra 314 y los indicadores tienen un valor particular (por ejemplo, 5), el dispositivo de generación de archivos 34 incluye, en la caja de información de submuestra 314, un valor adicional (por ejemplo, VclNalUnitType) que indica un tipo de unidad NAL de las unidades VCL NAL de la submuestra. Todas las unidades VCL NAL de la submuestra tienen el mismo tipo de unidad NAL.

[0308] La figura 6 es un diagrama conceptual que ilustra un ejemplo de estructura de un archivo 300 de acuerdo con una o más técnicas de esta divulgación. Como se especifica en la sección 8.4.9 de ISO/IEC 14496-15, la HEVC permite muestras de formato de archivo que se usan solo como referencia y no como salida. Por ejemplo, la HEVC permite una imagen de referencia no visualizada en vídeo.

[0309] Además, la sección 8.4.9 de ISO/IEC 14496-15 especifica que cuando cualquiera de dichas muestras no de salida está presente en una pista, el archivo se restringirá de la siguiente manera.

1. A una muestra no de salida se le dará un tiempo de composición que está fuera del intervalo de tiempo de las muestras que se facilitan.
2. Se usará una lista de edición para excluir los tiempos de composición de las muestras no de salida.
3. Cuando la pista incluye una CompositionOffsetBox («ctts»),
 - a. se usará la versión 1 de la CompositionOffsetBox,
 - b. el valor de sample_offset se establecerá para ser igual a -2^{31} para cada muestra no de salida,
 - c. La CompositionToDecodeBox («cslg») debe estar contenida en la SampleTableBox («stbl») de la pista, y
 - d. cuando la CompositionToDecodeBox está presente para la pista, el valor del campo leastDecodeToDisplayDelta de la caja será igual al desplazamiento de composición más pequeño de la CompositionOffsetBox excluidos los valores de sample_offset para muestras no de salida.

NOTA: Por tanto, leastDecodeToDisplayDelta es mayor que -2^{31} .

[0310] Como se especifica en ISO/IEC 14496-12, la CompositionOffsetBox proporciona el desplazamiento entre el tiempo de decodificación y el tiempo de composición. La CompositionOffsetBox incluye un conjunto de valores de sample_offset. Cada uno de los valores de sample_offset es un entero no negativo que proporciona el desplazamiento entre el tiempo de composición y el tiempo de decodificación. El tiempo de composición se refiere a un momento en el que se va a facilitar una muestra. El tiempo de decodificación se refiere a un momento en el que se va a decodificar una muestra.

[0311] Como se ha indicado anteriormente, una unidad NAL de sector codificado puede incluir una cabecera de segmento de sector. La cabecera de segmento de sector puede formar parte de un segmento de sector codificado y puede contener elementos de datos pertenecientes a la primera o todas las CTU del segmento de sector. En HEVC, la cabecera de segmento de sector incluye un elemento sintáctico pic_output_flag. En general, el elemento sintáctico pic_output_flag está incluido en una primera cabecera de segmento de sector de un sector de una imagen. Por tanto, esta divulgación puede referirse al indicador pic_output de la primera cabecera de segmento de sector del sector de la imagen como el pic_output_flag de la imagen.

[0312] Como se especifica en la sección 7.4.7.1 del HEVC WD, el elemento sintáctico pic_output_flag afecta la salida de imagen decodificada y los procesos de eliminación, tal como se especifica en el Anexo C del HEVC WD. En general, si el elemento sintáctico pic_output_flag de una cabecera de segmento de sector para un segmento de sector es 1, se facilita una imagen que incluye un sector correspondiente a la cabecera de segmento de sector. De lo contrario, si el elemento sintáctico del indicador pic_output de la cabecera de segmento de sector para un segmento de sector es 0, la imagen que incluye el sector correspondiente a la cabecera de segmento de sector puede decodificarse para su uso como imagen de referencia, pero no se facilita.

[0313] De acuerdo con una o más técnicas de esta divulgación, las referencias de la sección 8.4.9 de ISO/IEC 14496-15 a HEVC pueden reemplazarse por las referencias correspondientes a SHVC, MV-HEVC o 3D-HEVC. Además, de acuerdo con una o más técnicas de esta divulgación, cuando una unidad de acceso contiene algunas imágenes codificadas que tienen el indicador `pic_output` igual a 1 y otras imágenes codificadas que tienen `pic_output_flag` igual a 0, al menos se deben usar dos pistas para almacenar el flujo. Para cada una respectiva de las pistas, todas las imágenes codificadas de cada muestra de la pista respectiva tienen el mismo valor del indicador `pic_output`. Por tanto, todas las imágenes codificadas de una primera de las pistas tienen `pic_output_flag` igual a 0 y todas las imágenes codificadas de una segunda de las pistas tienen un indicador `pic_output` igual a 1.

[0314] En consecuencia, en el ejemplo de la figura 6, el dispositivo de generación de archivos 34 puede generar un archivo 400. De manera similar al archivo 300 del ejemplo de la figura 5, el archivo 400 incluye una caja de película 402 y una o más cajas de datos de medios 404. Cada una de las cajas de datos de medios 404 puede corresponder a una pista diferente del archivo 400. La caja de película 402 puede contener metadatos para las pistas del archivo 400. Cada pista del archivo 400 puede comprender un flujo continuo de datos de medios. Cada una de las cajas de datos de medios 404 puede incluir una o más muestras 405. Cada una de las muestras 405 puede comprender una unidad de acceso de audio o vídeo.

[0315] Como se ha indicado anteriormente, en algunos ejemplos, cuando una unidad de acceso contiene algunas imágenes codificadas que tienen `pic_output_flag` igual a 1 y otras imágenes codificadas que tienen `pic_output_flag` igual a 0, se deben usar al menos dos pistas para almacenar el flujo. En consecuencia, en el ejemplo de la figura 6, la caja de película 402 incluye una caja de pista 406 y una caja de pista 408. Cada una de las cajas de pista 406 y 408 contiene metadatos para una pista diferente del archivo 400. Por ejemplo, la caja de pista 406 puede contener metadatos para una pista que tiene imágenes codificadas con el indicador `pic_output` igual a 0, y ninguna imagen con el indicador `pic_output` igual a 1. La caja de pista 408 puede contener metadatos para una pista que tiene imágenes codificadas con `pic_output_flag` igual a 1, y ninguna imagen con `pic_output_flag` igual a 0.

[0316] Por tanto, en un ejemplo, el dispositivo de generación de archivos 34 puede generar un archivo (por ejemplo, el archivo 400) que comprende una caja de datos de medios (por ejemplo, la caja de datos de medios 404) que contiene (por ejemplo, comprende) contenido de medios. El contenido de medios comprende una secuencia de muestras (por ejemplo, las muestras 405). Cada una de las muestras puede ser una unidad de acceso de datos de vídeo multicapa. En este ejemplo, cuando el dispositivo de generación de archivos 34 genera el archivo, como respuesta a una determinación de que al menos una unidad de acceso del flujo de bits incluye una imagen codificada que tiene un indicador de salida de imagen igual a 1 y una imagen codificada que tiene un indicador de salida de imagen igual a 0, el dispositivo de generación de archivos 34 puede usar al menos dos pistas para almacenar el flujo de bits en el archivo. Para cada pista respectiva de las al menos dos pistas, todas las imágenes codificadas de cada muestra de la pista respectiva tienen el mismo valor del indicador de salida de imagen. Las imágenes que tienen indicadores de salida de imagen iguales a 1 pueden facilitarse y las imágenes que tienen indicadores de salida de imagen igual a 0 pueden usarse como imágenes de referencia, pero no pueden facilitarse.

[0317] La figura 7 es un diagrama de flujo que ilustra un ejemplo de operación del dispositivo de generación de archivos 34, de acuerdo con una o más técnicas de esta divulgación. La operación de la figura 7, junto con las operaciones ilustradas en otros diagramas de flujo de esta divulgación, son ejemplos. Otros ejemplos de operaciones de acuerdo con las técnicas de esta divulgación pueden incluir más, menos o diferentes acciones.

[0318] En el ejemplo de la figura 7, el dispositivo de generación de archivos 34 genera un archivo (500). Como parte de la generación del archivo, el dispositivo de generación de archivos 34 genera una caja de pista que contiene metadatos para una pista del archivo (502). De esta manera, el dispositivo de generación de archivos 34 genera un archivo que comprende una caja de pista que contiene metadatos para una pista del archivo. Los datos de medios para la pista comprenden una secuencia de muestras. Cada una de las muestras es una unidad de acceso de vídeo de los datos de vídeo multicapa. En algunos ejemplos, el dispositivo de generación de archivos 34 codifica los datos de vídeo multicapa.

[0319] Además, como parte de la generación del archivo, el dispositivo de generación de archivos 34 identifica todas las muestras que contienen al menos una imagen IRAP (504). Además, el dispositivo de generación de archivos 34 puede generar, en el archivo, una caja adicional que documenta todas las muestras que contienen al menos una imagen IRAP (506). En algunos ejemplos, la caja adicional es una caja nueva que no está definida en el ISO/BMFF o en las ampliaciones existentes del mismo. En algunos ejemplos, la caja adicional define un grupo de muestras que documenta todas las muestras que contienen al menos una imagen IRAP. Por ejemplo, la caja adicional puede ser o comprender una caja de tabla de muestras que incluye una caja `SampleToGroup` y una caja `SampleGroupDescription`. La caja `SampleToGroup` identifica las muestras que contienen al menos una imagen IRAP. La caja `SampleGroupDescription` indica que el grupo de muestras es un grupo de muestras que contiene al menos una imagen IRAP.

[0320] Además, en el ejemplo de la figura 7, el dispositivo de generación de archivos 34 puede generar una entrada de muestra para una de las muestras que incluye al menos una imagen IRAP (508). En algunos ejemplos, el dispositivo de generación de archivos 34 puede generar una entrada de muestra para cada una respectiva de las muestras que

incluye al menos una imagen IRAP. La entrada de muestra puede ser una RandomAccessibleSampleEntry como la definida en la sección 9.5.5.1.2 anterior.

[0321] Como se ilustra en el ejemplo de la figura 7, como parte de la generación de la entrada de muestra para la muestra particular, el dispositivo de generación de archivos 34 puede incluir, en la entrada de muestra para la muestra particular, un valor que indica si todas las imágenes codificadas de la muestra particular son imágenes IRAP (510). De esta manera, el dispositivo de generación de archivos 34 puede generar, en el archivo, una entrada de muestra que incluye un valor que indica si todas las imágenes codificadas de una muestra particular de la secuencia de muestras son imágenes IRAP. Además, el dispositivo de generación de archivos 34 puede incluir, en la entrada de muestra para la muestra particular, un valor que indica un tipo de unidad NAL de las unidades VCL NAL de las imágenes IRAP de la muestra particular (512).

[0322] Además, el dispositivo de generación de archivos 34 puede determinar si todas las imágenes codificadas de la muestra particular son imágenes IRAP (514). Cuando no todas las imágenes codificadas de la muestra particular son imágenes IRAP («NO» en 514), el dispositivo de generación de archivos 34 puede incluir, en la entrada de muestra para la muestra particular, un valor que indica un número de imágenes IRAP de la muestra particular (516). Adicionalmente, el dispositivo de generación de archivos 34 puede incluir, en la entrada de muestra para la muestra particular, valores que indican identificadores de capa (por ejemplo, unos ID de capa nuh) de las imágenes IRAP de la muestra particular.

[0323] Como se ha indicado anteriormente, la figura 7 se proporciona como ejemplo. Otros ejemplos no incluyen cada acción de la figura 7. Por ejemplo, algunos ejemplos excluyen las etapas 502, 504 y 508. Además, algunos ejemplos excluyen diversas de las etapas 510-518. Además, algunos ejemplos incluyen una o más acciones adicionales. Por ejemplo, algunos ejemplos incluyen una acción adicional de generar, como parte de la generación del archivo, una caja de muestras de sincronización que incluye una tabla de muestras de sincronización que documenta muestras de sincronización de una pista de los datos de vídeo multicapa. Cada muestra de sincronización de la pista es una muestra de acceso aleatorio de la pista. En este ejemplo, una muestra de codificación de vídeo escalable es una muestra de sincronización si cada imagen codificada de una unidad de acceso es una imagen IRAP. Además, en este ejemplo, una muestra de codificación de vídeo multivista es una muestra de sincronización si cada imagen codificada de la unidad de acceso es una imagen IRAP sin imágenes RASL.

[0324] La figura 8 es un diagrama de flujo que ilustra un ejemplo de operación en el que un dispositivo informático realiza un acceso aleatorio y/o un cambio de nivel, de acuerdo con una o más técnicas de esta divulgación. En el ejemplo de la figura 8, un dispositivo informático recibe un archivo (550). En el ejemplo de la figura 8, el dispositivo informático puede ser un dispositivo de red intermedio (por ejemplo, un MANE, un servidor de transmisión en tiempo real), un dispositivo de descodificación (por ejemplo, un dispositivo de destino 14) u otro tipo de dispositivo de vídeo. En algunos ejemplos, el dispositivo informático puede formar parte de una red de entrega de contenido.

[0325] En el ejemplo de la figura 8, el dispositivo informático puede obtener, a partir del archivo, una caja de pista que contiene metadatos para una pista del archivo (552). Los datos de medios para la pista comprenden una secuencia de muestras. En el ejemplo de la figura 8, cada una de las muestras es una unidad de acceso de vídeo de datos de vídeo multicapa.

[0326] Además, en el ejemplo de la figura 8, el dispositivo informático puede obtener una caja adicional a partir del archivo (554). La caja adicional documenta todas las muestras que contienen al menos una imagen IRAP. Por tanto, el dispositivo informático puede determinar, basándose en la información de la caja adicional, todas las muestras que contienen al menos una imagen IRAP (556).

[0327] Además, en algunos ejemplos, el dispositivo informático puede obtener, a partir del archivo, una entrada de muestra que incluye un valor que indica si todas las imágenes codificadas de una muestra particular de la secuencia de muestras son imágenes IRAP. Cuando no todas las imágenes codificadas de la muestra particular son imágenes IRAP, el dispositivo informático puede obtener, a partir de la entrada de muestra, un valor que indica un número de imágenes IRAP de la muestra particular. Adicionalmente, el dispositivo informático puede obtener, a partir de la entrada de muestra, unos valores que indican identificadores de capa de imágenes IRAP de la muestra particular. Además, en algunos ejemplos, el dispositivo informático puede obtener, a partir de la entrada de muestra, un valor que indica un tipo de unidad NAL de las unidades VCL NAL de las imágenes IRAP de la muestra particular. Adicionalmente, en algunos ejemplos, el dispositivo informático puede obtener, a partir del archivo, una caja de muestras de sincronización que incluye una tabla de muestras de sincronización que documenta muestras de sincronización de una pista de los datos de vídeo. En dichos ejemplos, cada muestra de sincronización de la pista es una muestra de acceso aleatorio de la pista, una muestra de codificación de vídeo escalable es una muestra de sincronización si cada imagen codificada de una unidad de acceso es una imagen IRAP, y una muestra de codificación de vídeo multivista es una muestra de sincronización si cada imagen codificada de la unidad de acceso es una imagen IRAP sin imágenes RASL.

[0328] Adicionalmente, en el ejemplo de la figura 8, el dispositivo informático puede comenzar a reenviar o descodificar unidades NAL de una muestra que contiene al menos una imagen IRAP sin reenviar o descodificar unidades NAL del archivo anteriores a la muestra (558) en orden de descodificación. De esta manera, el dispositivo

informático puede realizar un acceso aleatorio o un cambio de capa. Por ejemplo, el dispositivo informático puede comenzar a descodificar datos de vídeo multicapa en una de la una o más muestras que contienen al menos una imagen IRAP.

[0329] La figura 9 es un diagrama de flujo que ilustra un ejemplo de operación del dispositivo de generación de archivos 34, de acuerdo con una o más técnicas de esta divulgación. En el ejemplo de la figura 9, el dispositivo de generación de archivos 34 puede generar un archivo que comprende una caja de pista que contiene metadatos para una pista del archivo (600). Los datos de medios para la pista comprenden una secuencia de muestras. En el ejemplo de la figura 9, cada una de las muestras es una unidad de acceso de vídeo de los datos de vídeo multicapa. En algunos ejemplos, el dispositivo de generación de archivos 34 codifica los datos de vídeo multicapa.

[0330] Como parte de la generación del archivo, el dispositivo de generación de archivos 34 puede determinar si una submuestra contiene exactamente una imagen codificada y cero o más unidades no VCL NAL asociadas con la imagen codificada (602). Como respuesta a la determinación de que la submuestra contiene exactamente una imagen codificada y cero o más unidades no VCL NAL asociadas con la imagen codificada («SÍ» en 602), el dispositivo de generación de archivos 34 puede generar, en el archivo, una caja de información de submuestra que contiene indicadores que tienen un valor (por ejemplo, 5) que indica que la submuestra contiene exactamente una imagen codificada y cero o más unidades no VCL NAL asociadas con la imagen codificada (604). De lo contrario («NO» en 602), el dispositivo de generación de archivos 34 puede generar, en el archivo, la caja de información de submuestra que contiene indicadores que tienen otro valor (por ejemplo, 0, 1, 2, 3, 4) (606).

[0331] De esta manera, el dispositivo de generación de archivos 34 puede generar un archivo que comprende una caja de pista que contiene metadatos para una pista del archivo. Los datos de medios para la pista comprenden una secuencia de muestras, siendo cada una de las muestras una unidad de acceso de vídeo de datos de vídeo multicapa. Como parte de la generación del archivo, el dispositivo de generación de archivos 34 genera, en el archivo, una caja de información de submuestra que contiene indicadores que especifican un tipo de información de submuestra que se proporciona en la caja de información de submuestra. Cuando los indicadores tienen un valor particular, una submuestra correspondiente a la caja de información de submuestra contiene exactamente una imagen codificada y cero o más unidades no VCL NAL asociadas con la imagen codificada.

[0332] La figura 10 es un diagrama de flujo que ilustra un ejemplo de operación de un dispositivo informático, de acuerdo con una o más técnicas de esta divulgación. En el ejemplo de la figura 10, un dispositivo informático recibe un archivo (650). En el ejemplo de la figura 10, el dispositivo informático puede ser un dispositivo de red intermedio, tal como un MANE o un servidor de transmisión en tiempo real. En algunos ejemplos, el dispositivo informático puede formar parte de una red de entrega de contenido. Además, en el ejemplo de la figura 10, el dispositivo informático puede obtener una caja de pista a partir del archivo (651). La caja de pista contiene metadatos para una pista del archivo. Los datos de medios para la pista comprenden una secuencia de muestras. En el ejemplo de la figura 10, cada una de las muestras es una unidad de acceso de vídeo de datos de vídeo multicapa.

[0333] Además, en el ejemplo de la figura 10, el dispositivo informático puede obtener una caja de información de submuestra a partir del archivo (652). El dispositivo informático usa información de la caja de información de submuestra para extraer un subflujo de bits (654). El subflujo de bits puede comprender cada unidad NAL de un punto de operación de un flujo de bits almacenado en el archivo. En otras palabras, las unidades NAL del subflujo de bits pueden ser un subconjunto de las unidades NAL almacenadas en el archivo. El dispositivo informático puede obtener la caja de información de submuestra del archivo y puede extraer el subflujo de bits sin analizar ni interpretar las unidades NAL incluidas en la secuencia de muestras. No analizar ni interpretar las unidades NAL al extraer el subflujo de bits puede reducir la complejidad del dispositivo informático y/o puede acelerar el proceso de extracción del subflujo de bits.

[0334] Además, en algunos ejemplos, cuando los indicadores tienen un valor particular, el dispositivo informático puede obtener, a partir de la caja de información de submuestra, uno o más de:

- un indicador adicional que indica si todas las unidades VCL NAL de la submuestra son descartables,
- un valor adicional que indica un tipo de unidad NAL de las unidades VCL NAL de la submuestra, en el que todas las unidades VCL NAL de la submuestra tienen el mismo tipo de unidad NAL,
- un valor adicional que indica un identificador de capa de cada unidad NAL de la submuestra,
- un valor adicional que indica un identificador temporal de cada unidad NAL de la submuestra,
- un indicador adicional que indica si la predicción entre capas está habilitada para todas las unidades VCL NAL de la submuestra, o
- un indicador adicional que indica si todas las unidades NAL de la submuestra son unidades VCL NAL de una imagen no de referencia de subcapa.

[0335] En el ejemplo de la figura 10, como parte de la extracción del subflujo de bits, el dispositivo informático puede determinar si un valor «flags» de la caja de información de submuestra tiene un valor particular (por ejemplo, 5) que indica que la caja de información de submuestra corresponde exactamente a una imagen codificada y cero o más

unidades no VCL NAL asociadas con la imagen codificada (656). Cuando el valor «flags» de la caja de información de submuestra tiene un valor particular («SÍ» en 656), el dispositivo informático puede determinar, basándose en la información especificada en la caja de información de submuestra, si la imagen codificada se requiere a fin de descodificar el punto de operación (658). Por ejemplo, el dispositivo informático puede determinar, basándose en un indicador de descartable, un indicador de tipo de unidad VCL NAL, un identificador de capa, un identificador temporal, un indicador de predicción no entre capas y/o un indicador de unidad NAL de referencia de subcapa, si la imagen codificada se requiere para descodificar el punto de operación. Cuando la imagen codificada se requiere para descodificar el punto de operación («SÍ» en 658), el dispositivo informático puede incluir unidades NAL de la submuestra en el subflujo de bits (660). De lo contrario, en el ejemplo de la figura 10, cuando la imagen codificada no se requiere para descodificar el punto de operación («NO» en 658), el dispositivo informático no incluye las unidades NAL de la submuestra en el subflujo de bits (662).

[0336] Además, en el ejemplo de la figura 10, el dispositivo informático puede facilitar el subflujo de bits (664). Por ejemplo, el dispositivo informático puede almacenar el subflujo de bits en un medio de almacenamiento legible por ordenador o transmitir el subflujo de bits a otro dispositivo informático.

[0337] Como se ha indicado anteriormente, la figura 10 es un ejemplo. Otros ejemplos pueden incluir u omitir acciones particulares de la figura 10. Por ejemplo, algunos ejemplos omiten las acciones 650, 651, 654 y/o 664. Además, algunos ejemplos omiten una o más de las acciones 656-662.

[0338] La figura 11 es un diagrama de flujo que ilustra un ejemplo de operación del dispositivo de generación de archivos 34, de acuerdo con una o más técnicas de esta divulgación. En el ejemplo de la figura 11, el dispositivo de generación de archivos 34 puede generar un archivo que comprende una caja de datos de medios que contiene contenido de medios (700). El contenido de medios puede comprender una secuencia de muestras, siendo cada una de las muestras una unidad de acceso de datos de vídeo multicapa. En diversos ejemplos, los datos de vídeo multicapa pueden ser datos SHVC, datos MV-HEVC o datos 3D-HEVC. En algunos ejemplos, el dispositivo de generación de archivos 34 codifica los datos de vídeo multicapa.

[0339] En el ejemplo de la figura 11, como parte de la generación del archivo, el dispositivo de generación de archivos 34 puede determinar si al menos una unidad de acceso de un flujo de bits de los datos de vídeo multicapa incluye una imagen codificada que tiene un indicador de salida de imagen igual a un primer valor (por ejemplo, 1) y una imagen codificada que tiene un indicador de salida de imagen igual a un segundo valor (por ejemplo, 0) (702). Las imágenes que tienen indicadores de salida de imagen iguales al primer valor (por ejemplo, 1) pueden facilitarse y las imágenes que tienen indicadores de salida de imagen iguales al segundo valor (por ejemplo, 0) se pueden usar como imágenes de referencia, pero no se pueden facilitar. En otros ejemplos, otros dispositivos pueden determinar si al menos una unidad de acceso de un flujo de bits de datos de vídeo multicapa incluye una imagen codificada que tiene un indicador de salida de imagen igual al primer valor y una imagen codificada que tiene un indicador de salida de imagen igual al segundo valor.

[0340] Como respuesta a una determinación de que al menos una unidad de acceso de un flujo de bits de los datos de vídeo multicapa incluye una imagen codificada que tiene un indicador de salida de imagen igual al primer valor y una imagen codificada que tiene un indicador de salida de imagen igual al segundo valor («SÍ» en 702), el dispositivo de generación de archivos 34 usa al menos una primera pista y una segunda pista para almacenar el flujo de bits en el archivo (704). Para cada pista respectiva de la primera y la segunda pistas, todas las imágenes codificadas de cada muestra de la pista respectiva tienen el mismo valor del indicador de salida de imagen.

[0341] Además, en el ejemplo de la figura 11, como respuesta a la determinación de que ninguna unidad de acceso del flujo de bits incluye una imagen codificada que tiene un indicador de salida de imagen igual al primer valor (por ejemplo, 1) y una imagen codificada que tiene un indicador de salida de imagen igual al segundo valor (por ejemplo, 0) («NO» en 702), el dispositivo de generación de archivos 34 puede usar una sola pista para almacenar el flujo de bits en el archivo (706). En otros ejemplos, el dispositivo de generación de archivos 34 puede generar el archivo con múltiples pistas, incluso cuando ninguna unidad de acceso del flujo de bits incluye una imagen codificada que tiene un indicador de salida de imagen igual al primer valor (por ejemplo, 1) y una imagen codificada que tiene un indicador de salida de imagen igual al segundo valor (por ejemplo, 0).

[0342] Como se ha indicado anteriormente, la figura 11 es un ejemplo. Otros ejemplos pueden incluir menos acciones. Por ejemplo, algunos ejemplos omiten las acciones 702 y 706.

[0343] La figura 12 es un diagrama de flujo que ilustra un ejemplo de operación de un dispositivo de destino 14, de acuerdo con una o más técnicas de esta divulgación. En el ejemplo de la figura 12, el dispositivo de destino 14 recibe un archivo (750). El archivo puede comprender una caja de datos de medios que contiene contenido de medios, comprendiendo el contenido de medios una secuencia de muestras. Cada una de las muestras puede ser una unidad de acceso de datos de vídeo multicapa. En diversos ejemplos, los datos de vídeo multicapa pueden ser datos SHVC, datos MV-HEVC o datos 3D-HEVC. Además, en el ejemplo de la figura 12, el dispositivo de destino 14 puede obtener, a partir del archivo, una primera caja de pista y una segunda caja de pista (751). La primera caja de pista contiene metadatos para una primera pista del archivo. La segunda caja de pista contiene metadatos para una segunda pista

del archivo. Para cada pista respectiva de la primera pista y la segunda pista, todas las imágenes codificadas de cada muestra de la pista respectiva tienen el mismo valor del indicador de salida de imagen. Las imágenes que tienen indicadores de salida de imagen iguales a un primer valor (por ejemplo, 1) pueden facilitarse, y las imágenes que tienen indicadores de salida de imagen iguales a un segundo valor (por ejemplo, 0) pueden usarse como imágenes de referencia, pero no pueden facilitarse.

[0344] El descodificador de vídeo 30 del dispositivo de destino 14 puede descodificar imágenes de la pista para imágenes con indicadores de salida de imagen iguales a un primer valor (por ejemplo, 1) y puede descodificar imágenes de la pista para imágenes con indicadores de salida de imagen iguales a un segundo valor (por ejemplo, 0) (752). En algunos casos, el descodificador de vídeo 30 puede usar imágenes con indicadores de salida de imagen iguales a 1 para descodificar imágenes con indicadores de salida de imagen iguales a 0, y viceversa. El dispositivo de destino 14 puede facilitar las imágenes con indicadores de salida de imagen iguales al primer valor (754). El dispositivo de destino 14 no facilita las imágenes con indicadores de salida de imagen iguales al segundo valor (756). De esta manera, para cada pista respectiva de la primera y la segunda pista, el dispositivo de destino 14 puede descodificar las imágenes codificadas de cada muestra de la pista respectiva y facilitar las imágenes descodificadas que tienen indicadores de salida de imagen iguales al primer valor.

[0345] Como se ha indicado anteriormente, la figura 12 se proporciona como ejemplo. Otros ejemplos pueden omitir acciones particulares de la figura 12, como las acciones 752-756.

[0346] En uno o más ejemplos, las funciones descritas pueden implementarse en hardware, software, firmware o cualquier combinación de los mismos. Si se implementan en software, las funciones se pueden almacenar en, o transmitir a través de, un medio legible por ordenador como una o más instrucciones o código y ejecutarse mediante una unidad de procesamiento basada en hardware. Los medios legibles por ordenador pueden incluir medios de almacenamiento legibles por ordenador, que corresponden a un medio tangible tal como unos medios de almacenamiento de datos, o unos medios de comunicación que incluyen cualquier medio que facilita la transferencia de un programa informático de un lugar a otro, por ejemplo, de acuerdo con un protocolo de comunicación. De esta manera, los medios legibles por ordenador pueden corresponder en general a (1) medios de almacenamiento tangibles legibles por ordenador que 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 pueda acceder desde uno o más ordenadores o uno o más procesadores para recuperar instrucciones, código y/o estructuras de datos para la implementación de las técnicas descritas en esta divulgación. Un producto de programa informático puede incluir un medio legible por ordenador.

[0347] A modo de ejemplo, y no de limitación, dichos medios de almacenamiento legibles por ordenador pueden comprender RAM, ROM, EEPROM, CD-ROM u otro almacenamiento de disco óptico, almacenamiento de disco magnético u otros dispositivos de almacenamiento magnético, memoria flash o cualquier otro medio que pueda usarse para almacenar un código de programa deseado en forma de instrucciones o estructuras de datos y al que un ordenador pueda acceder. 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 otra fuente remota mediante un cable coaxial, un cable de fibra óptica, un par trenzado, una línea de abonado digital (DSL) o unas 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 están incluidos 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 su lugar, están dirigidos a medios de almacenamiento tangibles no transitorios. El término disco, como se usa 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, de los cuales los discos flexibles normalmente reproducen datos magnéticamente, mientras que el resto de discos reproducen datos ópticamente con láseres. Las combinaciones de los anteriores también deberían incluirse dentro del alcance de los medios legibles por ordenador.

[0348] 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), matrices de puertas programables *in situ* (FPGA) u otros circuitos lógicos, integrados o discretos equivalentes. En consecuencia, el término «procesador», como se usa en el presente documento, se puede referir a cualquiera de las estructuras anteriores o a cualquier otra estructura adecuada para la implementación de las técnicas descritas en el presente documento. Además, en algunos aspectos, la funcionalidad descrita en el presente documento se puede proporcionar dentro de módulos de hardware y/o software dedicados, configurados para codificar y descodificar, o incorporar en un códec combinado. Además, las técnicas se podrían implementar totalmente en uno o más circuitos o elementos lógicos.

[0349] Las técnicas de la presente divulgación se pueden implementar en una amplia variedad de dispositivos o aparatos, incluidos un teléfono inalámbrico, un circuito integrado (CI) o un conjunto de CI (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 por diferentes unidades de hardware. En su lugar, como se ha descrito anteriormente, diversas unidades

se pueden combinar en una unidad de hardware de códec, o proporcionar mediante un grupo de unidades de hardware interoperativas, incluido uno o más procesadores, como se ha descrito anteriormente, conjuntamente con software y/o firmware adecuados.

- 5 **[0350]** Se han descrito diversos ejemplos. Estos y otros ejemplos están dentro del alcance de las siguientes reivindicaciones.

REIVINDICACIONES

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

generar (700) un archivo (300, 400) que comprende una caja de película (302, 402) y al menos una caja de datos de medios (304, 404), conteniendo la caja de película metadatos para flujos de medios presentes en el archivo, conteniendo la caja de datos de medios contenido de medios que comprende una secuencia de unidades de acceso de los datos de vídeo multicapa, incluyendo cada unidad de acceso una pluralidad de imágenes codificadas, en el que generar el archivo comprende:

como respuesta a una determinación (702) de que al menos una unidad de acceso de un flujo de bits de los datos de vídeo multicapa incluye una imagen codificada que tiene un indicador de salida de imagen igual a un primer valor y una imagen codificada que tiene un indicador de salida de imagen igual a un segundo valor, usar (704) al menos una primera pista y una segunda pista para almacenar el flujo de bits en el archivo, en el que, para cada pista respectiva de la primera y segunda pistas, todas las imágenes codificadas en cada unidad de acceso de la pista respectiva tienen el mismo valor del indicador de salida de imagen, en el que unas imágenes que tienen indicadores de salida de imagen iguales al primer valor se pueden facilitar y unas imágenes que tienen indicadores de salida de imagen iguales al segundo valor se pueden usar como imágenes de referencia pero no se pueden facilitar; y

para cada pista respectiva del archivo, incluir una caja de pista respectiva (306, 406) en la caja de película e incluir una caja de datos de medios respectiva (304, 404), de la al menos una caja de datos de medios, conteniendo la caja de pista respectiva metadatos para la pista respectiva, incluyendo la caja de datos de medios respectiva para la pista respectiva una secuencia respectiva de unidades de acceso de los datos de vídeo multicapa.

2. El procedimiento de acuerdo con la reivindicación 1, en el que generar el archivo comprende:

como respuesta a una determinación (702) de que ninguna unidad de acceso del flujo de bits incluye una imagen codificada que tiene un indicador de salida de imagen igual al primer valor y una imagen codificada que tiene un indicador de salida de imagen igual al segundo valor, usar (706) una sola pista para almacenar el flujo de bits en el archivo.

3. Un procedimiento para procesamiento de datos de vídeo multicapa, comprendiendo el procedimiento:

obtener (751), a partir de un archivo (300, 400), una primera caja de pista (306, 406) y una segunda caja de pista (408), estando la primera caja de pista y la segunda caja de pista en una caja de película (302, 402) que contiene metadatos para unas pistas presentes en el archivo, en el que para cada pista respectiva del archivo, una secuencia respectiva de unidades de acceso de los datos de vídeo multicapa está contenida en una respectiva caja de datos de medios (304, 404) del archivo, incluyendo cada unidad de acceso una pluralidad de imágenes codificadas, en el que la primera caja de pista contiene metadatos para una primera pista del archivo y la segunda caja de pista contiene metadatos para una segunda pista del archivo, en el que para cada pista respectiva de la primera pista y la segunda pista, todas las imágenes codificadas en cada unidad de acceso de la pista respectiva tienen el mismo valor de un indicador de salida de imagen, y en el que unas imágenes que tienen indicadores de salida de imagen iguales a un primer valor se pueden facilitar y unas imágenes que tienen indicadores de salida de imagen iguales a un segundo valor se pueden usar como imágenes de referencia pero no se pueden facilitar.

4. El procedimiento de acuerdo con la reivindicación 1 o la reivindicación 3, en el que los datos de vídeo multicapa son datos de codificación de vídeo de alta eficiencia escalables (SHVC), datos de codificación de vídeo de alta eficiencia multivista (MV-HEVC) y datos de codificación de vídeo de alta eficiencia de 3 dimensiones (3D-HEVC), y en el que los datos de vídeo multicapa incluyen una capa base de codificación de vídeo de alta eficiencia (HEVC).

5. El procedimiento de acuerdo con la reivindicación 1 o la reivindicación 3, en el que el primer valor es igual a 1 y el segundo valor es igual a 0.

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

para cada pista respectiva de la primera y la segunda pista:

descodificar (752) las imágenes codificadas en cada unidad de acceso de la pista respectiva; y

facilitar (754) las imágenes descodificadas que tienen identificadores de salida de imagen iguales al primer valor.

7. Un dispositivo de vídeo para procesar datos de vídeo multicapa, comprendiendo el dispositivo de vídeo:

un medio de almacenamiento de datos configurado para almacenar datos de vídeo multicapa; y

uno o más procesadores configurados para:

generar (700) un archivo (300, 400) que comprende una caja de película (302, 402) y al menos una caja de datos de medios (304, 404), conteniendo la caja de película metadatos para flujos de medios presentes en el archivo, conteniendo la caja de datos de medios contenido de medios que comprende una secuencia de unidades de acceso de los datos de vídeo multicapa, incluyendo cada unidad de acceso una pluralidad de imágenes codificadas, en el que para generar el archivo, los uno o más procesadores están configurados para:

como respuesta a una determinación (702) de que al menos una unidad de acceso de un flujo de bits de los datos de vídeo multicapa incluye una imagen codificada que tiene un indicador de salida de imagen igual a un primer valor y una imagen codificada que tiene un indicador de salida de imagen igual a un segundo valor, usar (704) al menos una primera pista y una segunda pista para almacenar el flujo de bits en el archivo, en el que para cada pista respectiva de la primera y segunda pistas, todas las imágenes codificadas en cada unidad de acceso de la pista respectiva tienen el mismo valor del indicador de salida de imagen, en el que unas imágenes que tienen indicadores de salida de imagen iguales al primer valor se pueden facilitar y unas imágenes que tienen indicadores de salida de imagen iguales al segundo valor se pueden usar como imágenes de referencia pero no se pueden facilitar, para cada pista respectiva del archivo, incluir una caja de pista respectiva (306, 406) en la caja de película e incluir una caja de datos de medios respectiva (304, 404), de dicha al menos una caja de datos de medios, conteniendo la caja de pista respectiva metadatos para la pista respectiva, incluyendo la caja de datos de medios respectiva para la pista respectiva una secuencia respectiva de unidades de acceso de los datos de vídeo multicapa.

8. El dispositivo de vídeo de acuerdo con la reivindicación 7, en el que para generar el archivo, los uno o más procesadores están configurados para:

como respuesta a una determinación (702) de que ninguna unidad de acceso del flujo de bits incluye una imagen codificada que tiene un indicador de salida de imagen igual al primer valor y una imagen codificada que tiene un indicador de salida de imagen igual al segundo valor, usar (706) una sola pista para almacenar el flujo de bits en el archivo.

9. El dispositivo de vídeo de acuerdo con la reivindicación 7, en el que los uno o más procesadores están configurados para codificar los datos de vídeo multicapa.

10. Un dispositivo de vídeo para procesar datos de vídeo multicapa, comprendiendo el dispositivo de vídeo:

un medio de almacenamiento de datos configurado para almacenar datos de vídeo multicapa; y

uno o más procesadores configurados para:

obtener (751), a partir de un archivo (300, 400), una primera caja de pista (306, 406) y una segunda caja de pista (408), estando la primera caja de pista y la segunda caja de pista en una caja de película (302, 402) que contiene metadatos para flujos de medios presentes en el archivo, en el que para cada pista respectiva del archivo, una secuencia respectiva de unidades de acceso de los datos de vídeo multicapa está contenida en una caja de datos de medios (304, 404) del archivo, incluyendo cada unidad de acceso una pluralidad de imágenes codificadas, en el que la primera caja de pista contiene metadatos para una primera pista del archivo y la segunda caja de pista contiene metadatos para una segunda pista del archivo, en el que para cada pista respectiva de la primera pista y la segunda pista, todas las imágenes codificadas en cada unidad de acceso de la pista respectiva tienen el mismo valor de un indicador de salida de imagen, y en el que unas imágenes que tienen indicadores de salida de imagen iguales a un primer valor se pueden facilitar y unas imágenes que tienen indicadores de salida de imagen iguales a un segundo valor se pueden usar como imágenes de referencia pero no se pueden facilitar.

11. El dispositivo de vídeo de acuerdo con la reivindicación 10, en el que el uno o más procesadores están configurados para, para cada pista respectiva de la primera y la segunda pista, descodificar (752) las imágenes

codificadas en cada unidad de acceso de la pista respectiva, y facilitar (754) la imágenes descodificadas que tienen identificadores de salida de imagen iguales al primer valor.

- 5 **12.** El dispositivo de vídeo de acuerdo con la reivindicación 7 o la reivindicación 10, en el que los datos de vídeo multicapa son uno de: datos de codificación de vídeo de alta eficiencia escalables, SHVC, datos de codificación de vídeo de alta eficiencia multivista, MV-HEVC, y datos de codificación de vídeo de alta eficiencia de 3 dimensiones, 3D-HEVC, y en el que los datos de vídeo multicapa incluyen una capa base de codificación de vídeo de alta eficiencia, HEVC.
- 10 **13.** El dispositivo de vídeo de acuerdo con la reivindicación 7 o la reivindicación 10, en el que el primer valor es igual a 1 y el segundo valor es igual a 0.
- 15 **14.** El dispositivo de vídeo de acuerdo con la reivindicación 7 o la reivindicación 10, en el que el dispositivo de vídeo comprende al menos uno de:
 - un circuito integrado;
 - un microprocesador; o
 - 20 un dispositivo de comunicación inalámbrica.
- 25 **15.** Un medio de almacenamiento de datos legible por ordenador que tiene instrucciones almacenadas en el mismo que cuando se ejecutan hacen que uno o más procesadores lleven a cabo un procedimiento de acuerdo con cualquiera de las reivindicaciones 1 a 6.

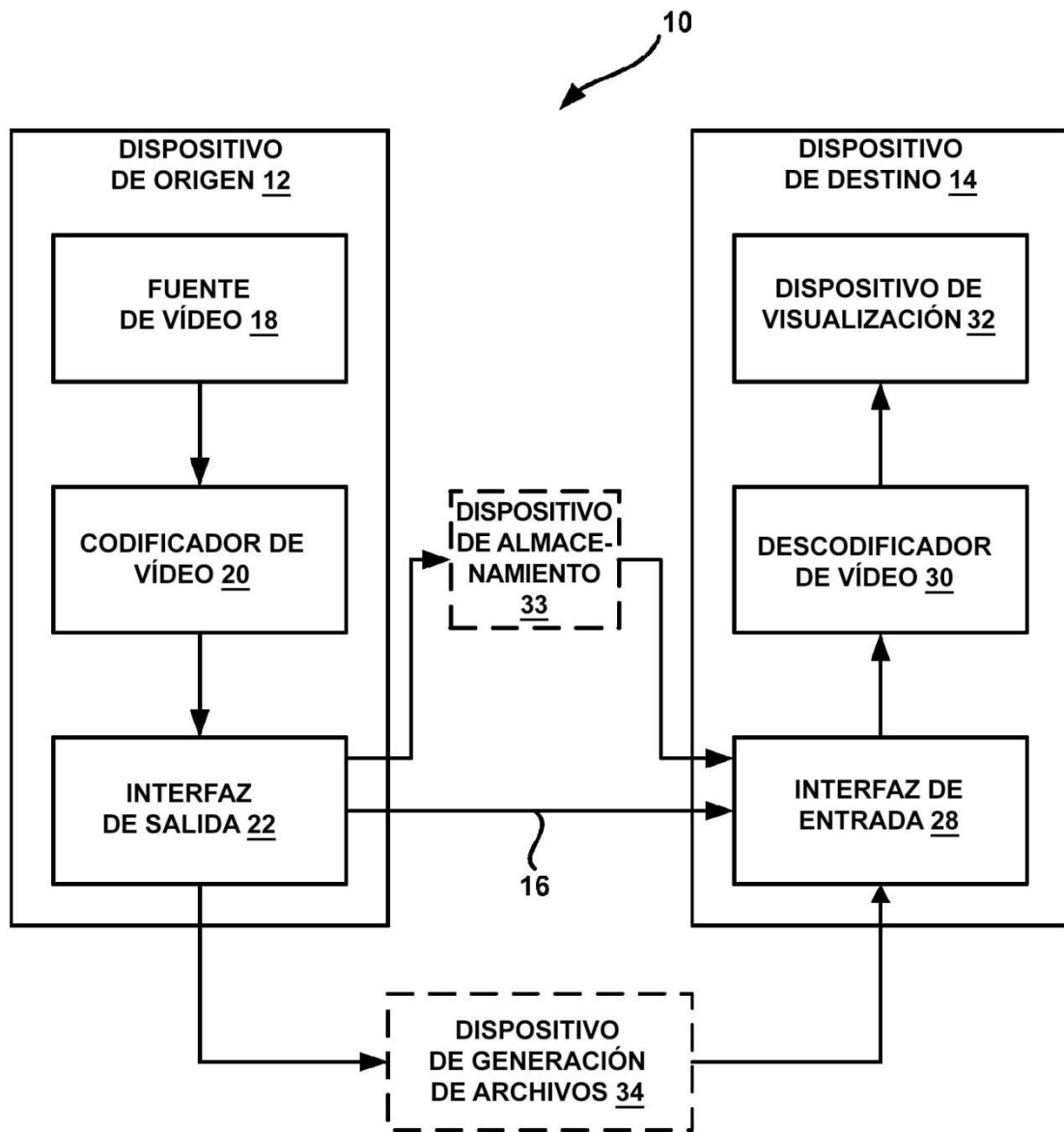


FIG. 1

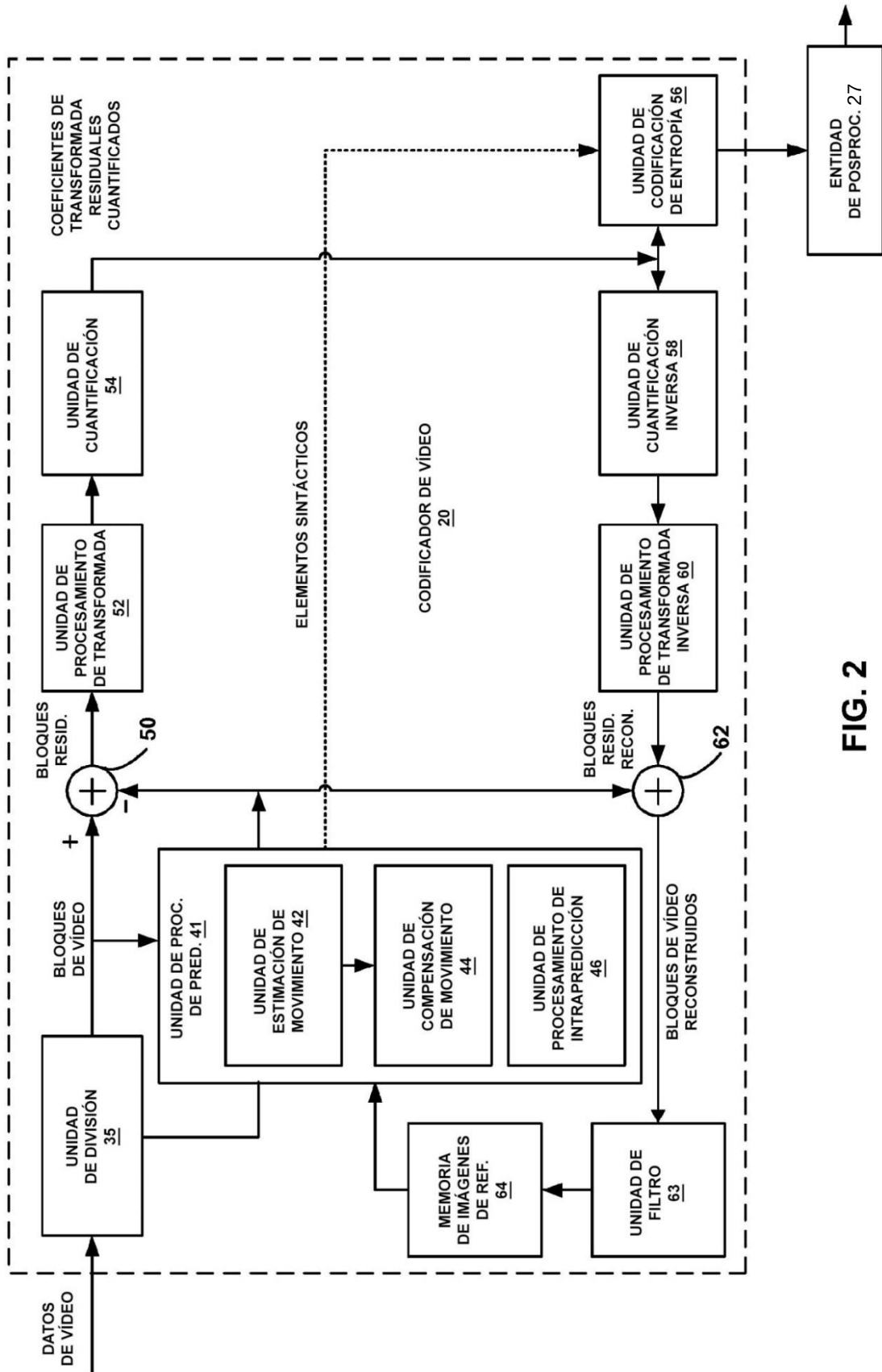


FIG. 2

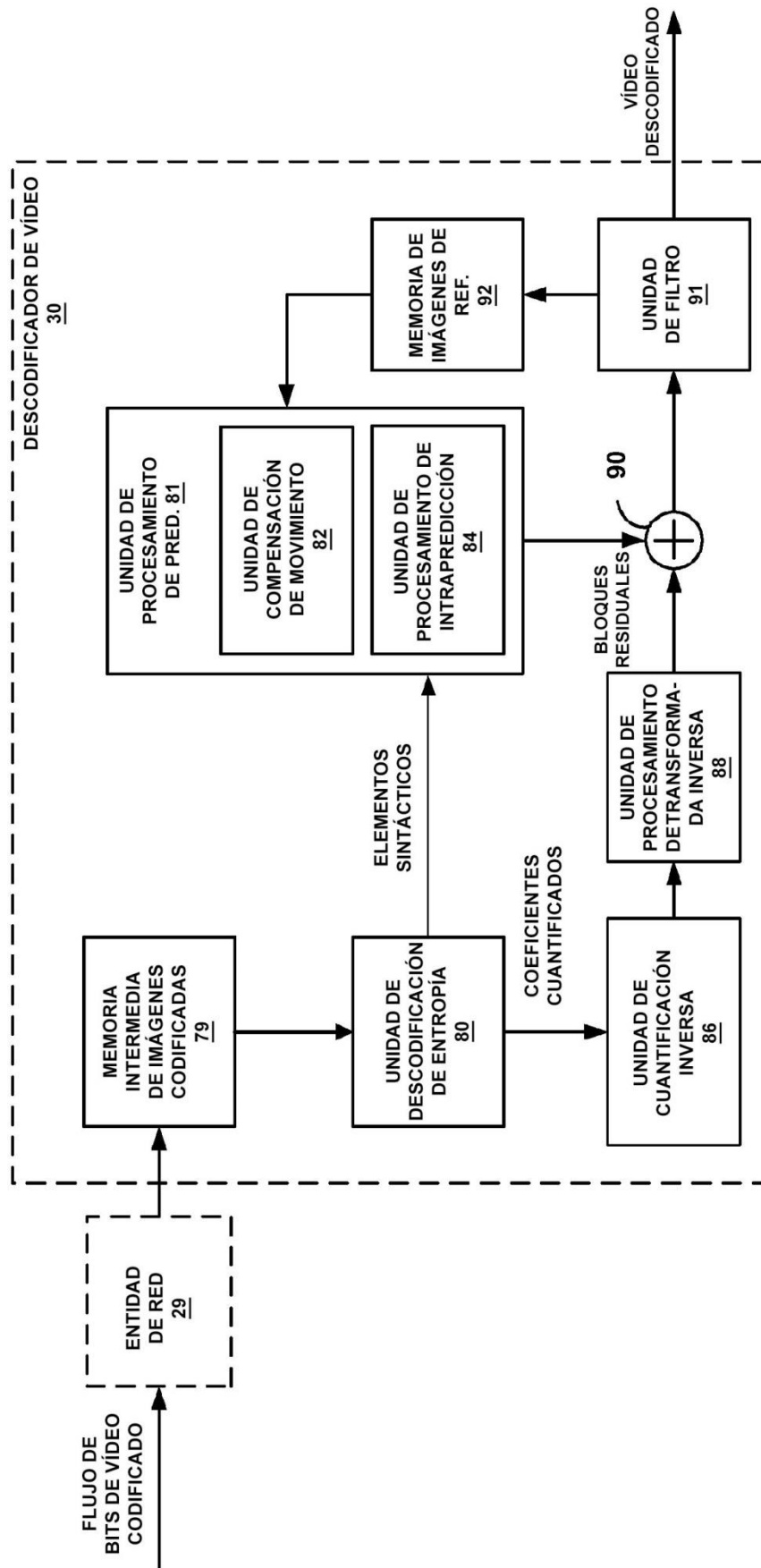


FIG. 3

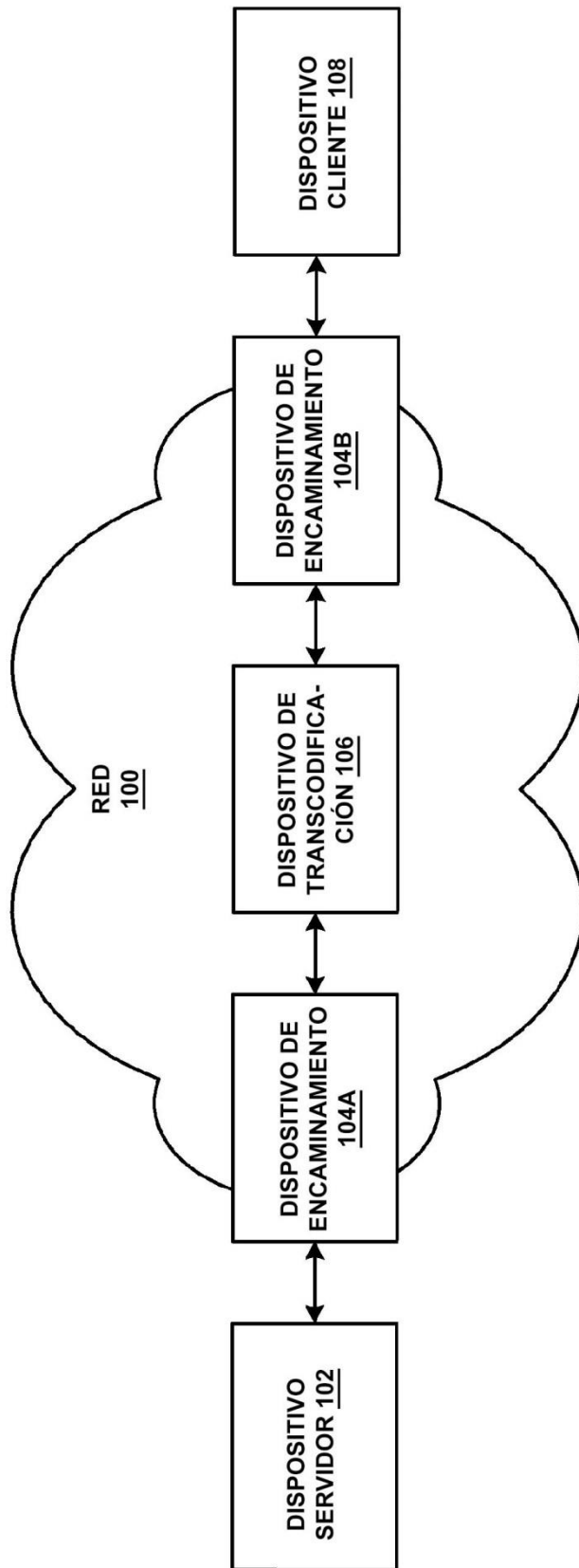


FIG. 4

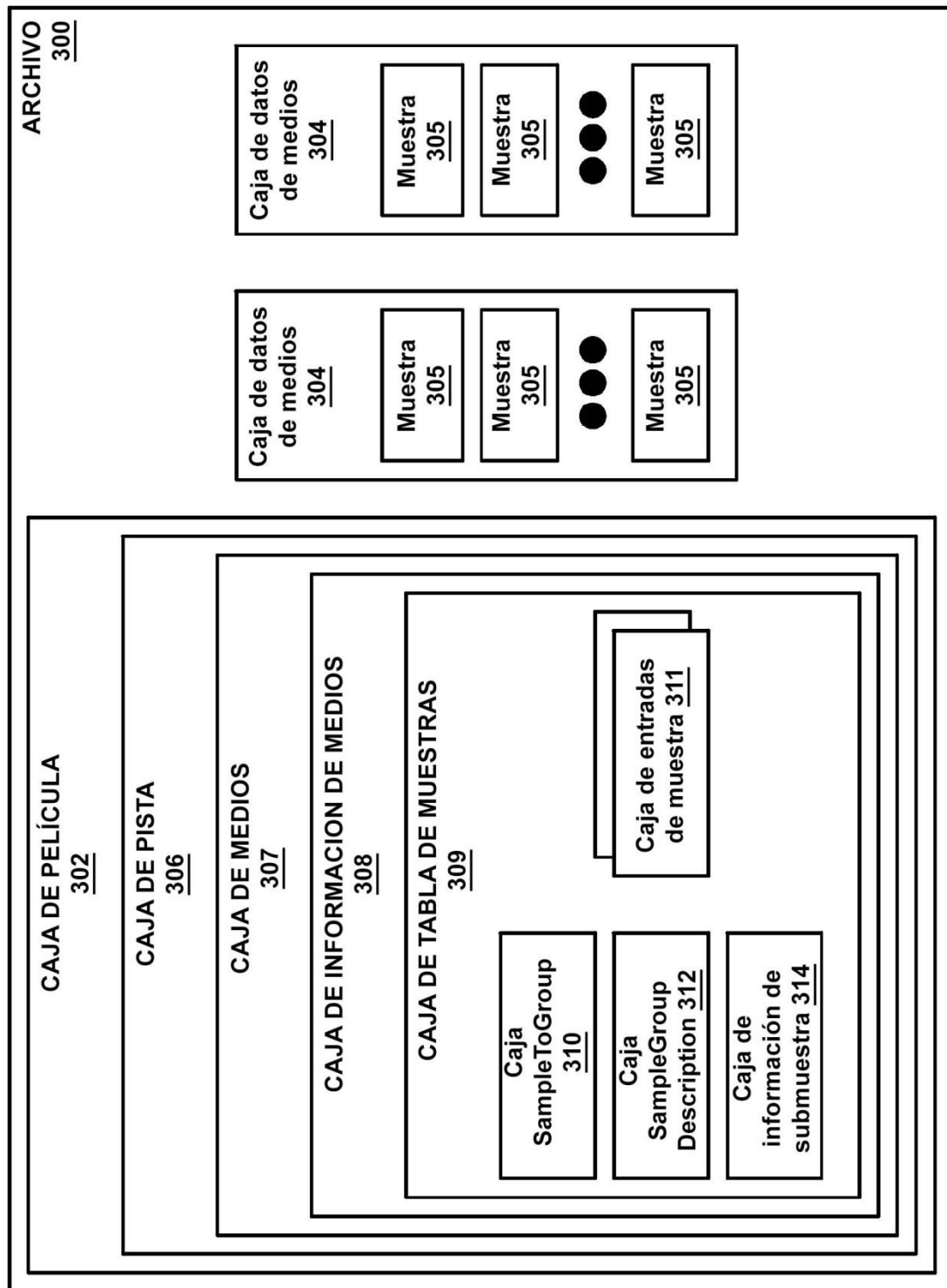


FIG. 5

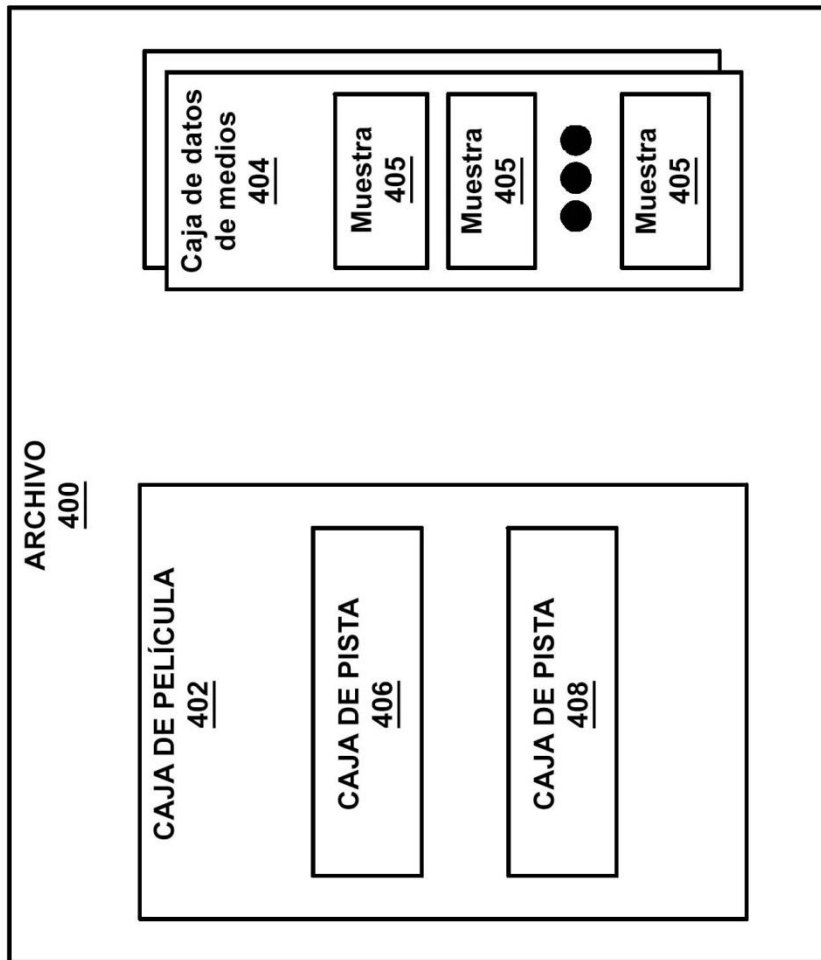


FIG. 6

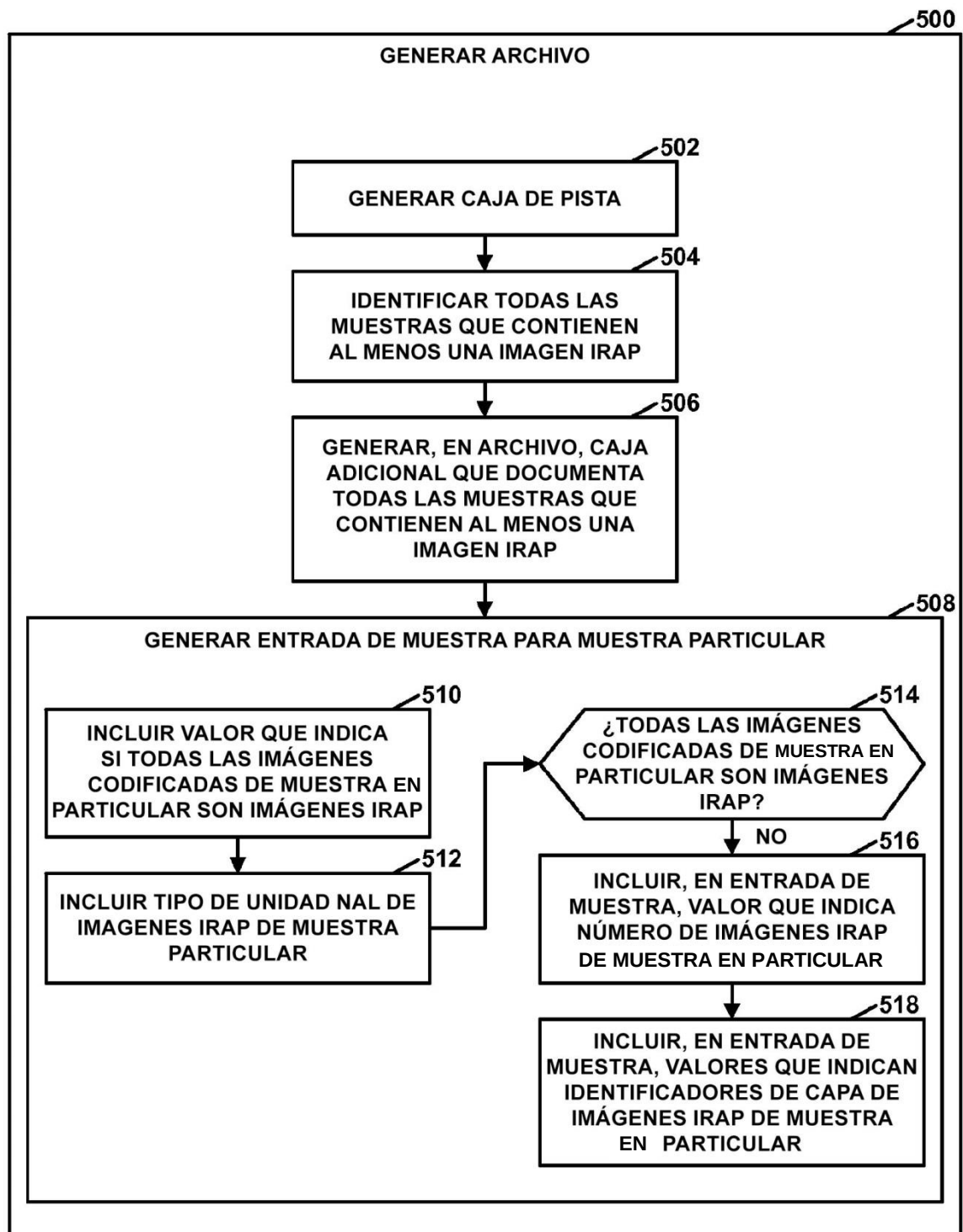


FIG. 7

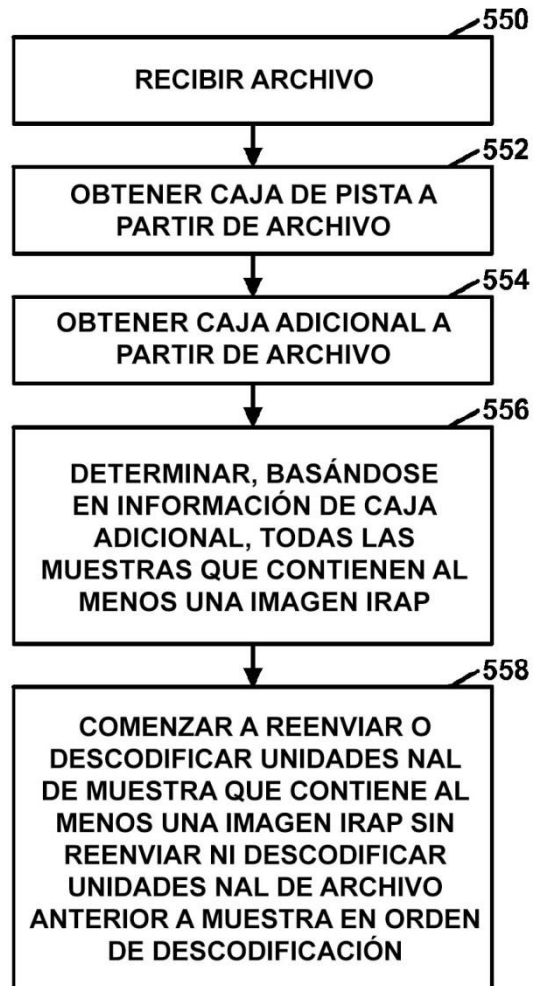


FIG. 8

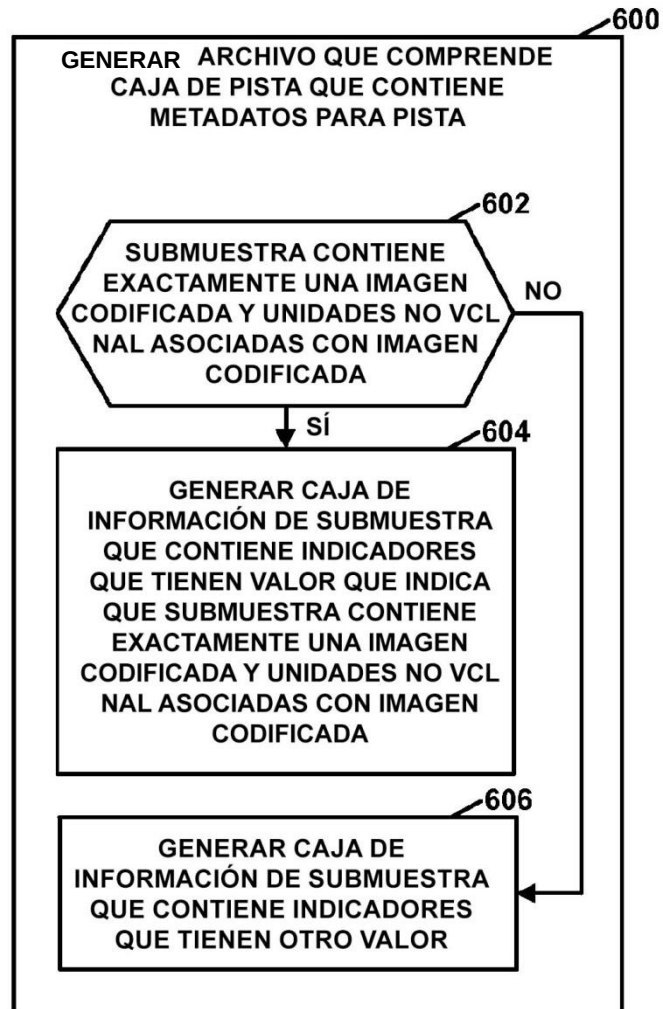


FIG. 9

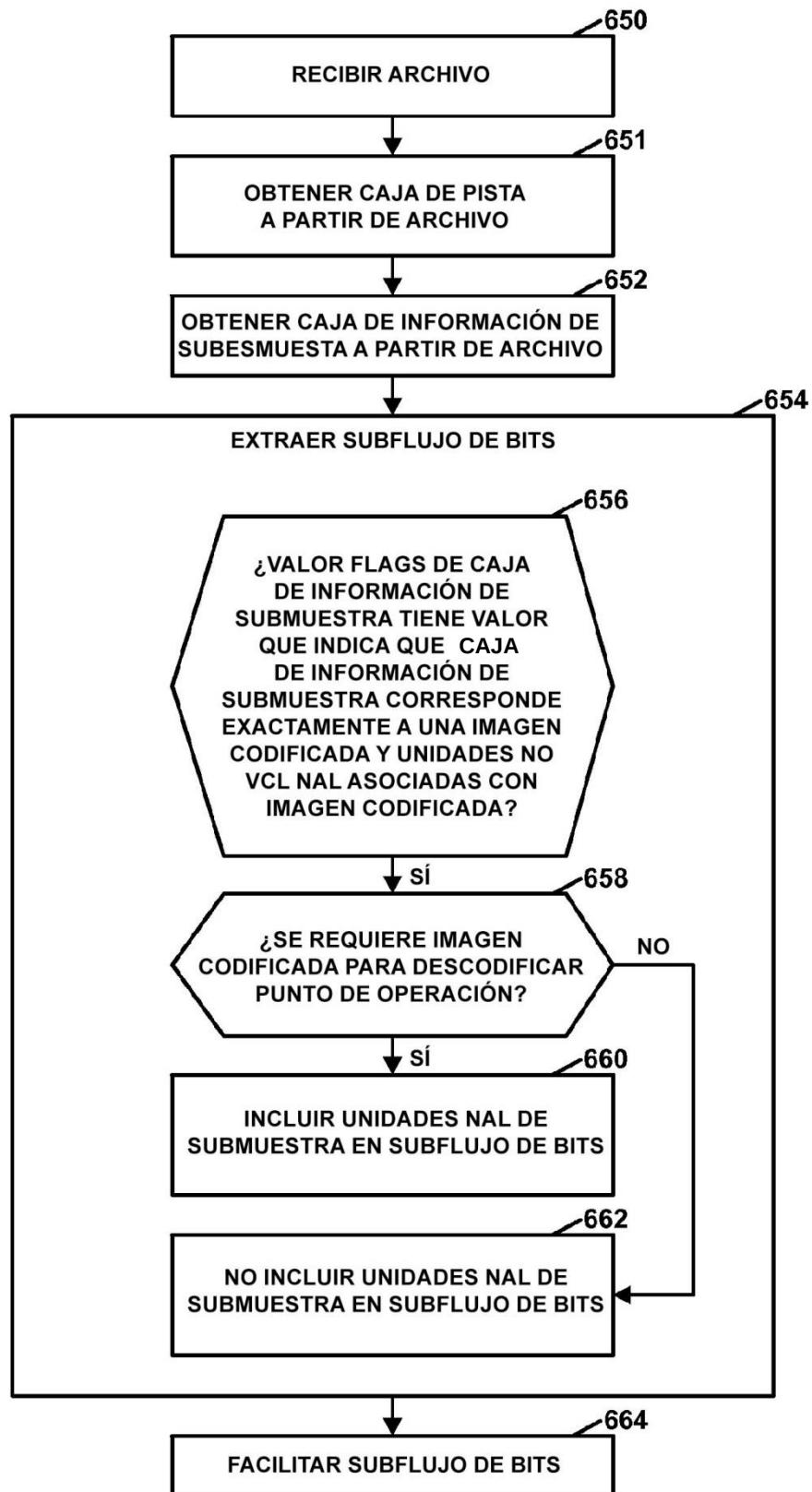


FIG. 10

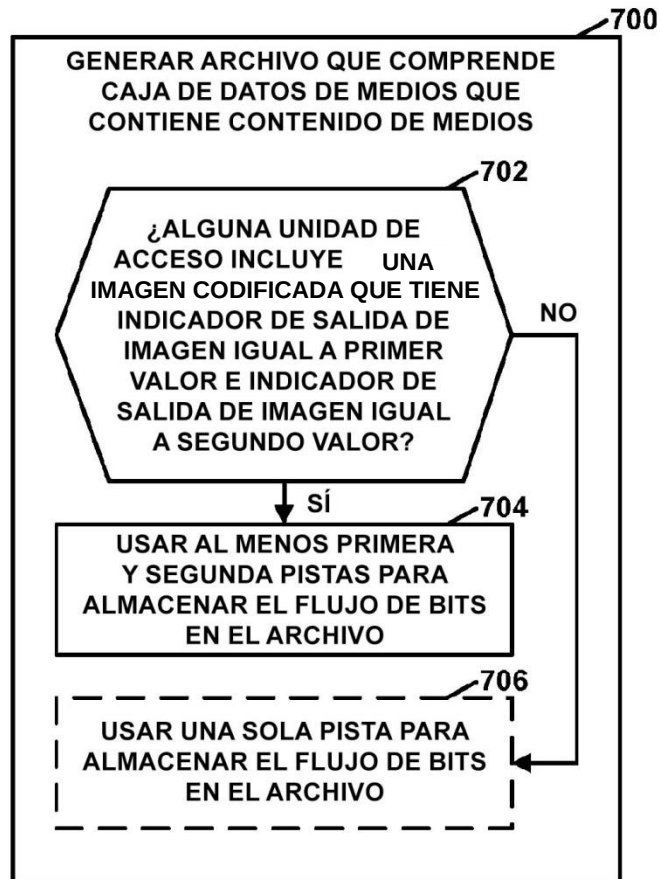


FIG. 11

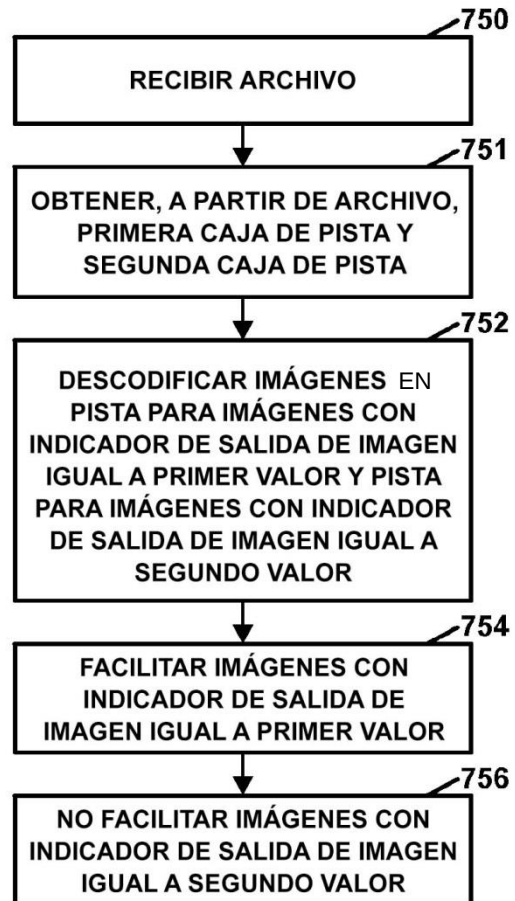


FIG. 12