

CONMUTACIÓN ADAPTATIVA DE ESPACIOS DE COLOR,
FRECUENCIAS DE MUESTREO DE COLOR Y/O PROFUNDIDADES
DE BITS

5 Antecedentes de la invención

Los ingenieros usan la compresión (también denominada cifrado fuente o codificación fuente) para reducir la velocidad binaria del video digital. La compresión disminuye el costo de almacenamiento y transmisión de la información de video mediante la
10 conversión de la información en una forma de menor velocidad binaria. La descompresión (también denominada decodificación) reconstruye una versión de la información original a partir de forma comprimida. Un "códec" es un sistema codificador-decodificador.

En las últimas dos décadas, se han adoptado diversos
15 estándares de códec de video, que incluyen los estándares UIT-T H.261, H.262 (MPEG-2 o ISO/IEC 13818-2), H.263 y H.264 (MPEG-4 A VC o ISO/IEC 14496-10), los estándares MPEG-1 (ISO/IEC 11172-2) y los estándares MPEG-4 Visual (ISO/IEC 14496-2), y el estándar SMPTE 421M (VC-1). Más recientemente, se ha aprobado el
20 estándar H.265/HEVC (ITU-T H.265 o ISO/IEC 23008-2). Las extensiones para el estándar H.265/HEVC (por ejemplo, para la codificación-decodificación de video escalable, para la codificación-decodificación de video con mayor fidelidad en cuanto a la profundidad de bits de muestra o frecuencia de muestreo de croma,
25 para el contenido de captura de pantalla o para la codificación-

decodificación de vistas múltiples) están actualmente en desarrollo.
 Un estándar códec de video comúnmente define las opciones para la
 sintaxis de una corriente de bits de video codificado, que detallan los
 parámetros en la corriente de bits cuando se usan las características
 5 particulares de codificación y decodificación. En muchos casos, un
 estándar de códec de video también proporciona detalles con
 respecto a las operaciones de decodificación que un decodificador
 debería realizar para obtener los resultados concordantes en la
 decodificación. Además de los estándares de códec, varios formatos
 10 de códec propietario definen otras opciones para la sintaxis de una
 corriente de bits de video codificado y las operaciones de
 decodificación correspondientes.

Una fuente de video como una cámara, salida de animación,
 módulo de captura de pantalla, etcétera, normalmente proporciona el
 15 video en un espacio de color particular, con componentes de color
 del video submuestreado de acuerdo con una frecuencia de muestreo
 de color particular y con valores de muestra que tienen una
 profundidad de bits particular. En general, un espacio de color
 (ocasionalmente denominado modelo de color) es un modelo que
 20 representa los colores como valores n por posición física, para $n \geq 1$,
 donde cada uno de los valores n proporciona un valor de componente
 de color para esa posición. Por ejemplo, en un espacio de color YUV,
 un valor de componente de luma (o Y) representa un brillo
 aproximado en una posición y los múltiples valores de componente
 25 de croma (o U y V) representan las diferencias de color en la

posición. O, en un espacio de color RGB, un valor de componente de rojo (R) representa una intensidad de rojo, un valor de componente de verde (G) representa una intensidad de verde y un valor de componente de azul (B) representa una intensidad de azul en una posición. Históricamente, los diferentes espacios de color tienen ventajas para diferentes aplicaciones como visualización, impresión, emisión y codificación-decodificación. Los valores de muestra pueden convertirse entre espacios de color usando las operaciones de transformación de espacio de color.

La frecuencia de muestreo de color (ocasionalmente denominada frecuencia de muestreo de croma) se refiere a la resolución espacial relativa entre los componentes de color. Por ejemplo, para una frecuencia de muestreo de color de 4:4:4, la información para los componentes secundarios (por ejemplo, los componentes U y V para YUV) tiene la misma resolución espacial que la información para un componente primario (por ejemplo, el componente Y para YUV). Para una frecuencia de muestreo de color de 4:2:2 o 4:2:0, la información para los componentes secundarios se someterá a reducción de muestra con relación a la información para el componente primario. El formato YUV a 4:2:0 se usa comúnmente para la codificación-decodificación. Como un principio de diseño, la decisión de usar un formato YUV a 4:2:0 para la codificación-decodificación, se basa en el entendimiento de que, para la mayoría de casos de uso, los espectadores no perciben muchas diferencias visuales entre el video codificado-decodificado en el formato YUV a

4:2:0 y el video codificado-decodificado en el formato YUV a 4:4:4. Las ventajas de compresión para el formato YUV a 4:2:0, que tiene menos muestras por fotograma, por lo tanto, son convincentes.

La profundidad de bits se refiere al número de bits por valor de muestra. Las profundidades de bits son de 8 bits por muestra, 10 bits por muestra y 12 bits por muestra. En general, tener más bits por muestra permite una graduación más precisa de los colores para el video, pero usa más espacio de almacenamiento para el video. Tener menos bits por muestra comúnmente reduce la velocidad binaria a costa de menor calidad.

Muchos codificadores y decodificadores de video comercialmente disponibles solo son compatibles con un formato YUV a 4:2:0. Otros codificadores y decodificadores comercialmente disponibles (por ejemplo, para el estándar H.264/AVC o el estándar H.265/HEVC) permiten que un codificador especifique un espacio de color, frecuencia de muestreo de color y profundidad de bits para una determinada secuencia. El espacio de color especificado, la frecuencia de muestreo de color y la profundidad de bits se usan para toda la secuencia de video. Estos métodos no proporcionan la flexibilidad suficiente para un sistema de códecs de uso general que puede procesar muy diferentes tipos de contenido de video en una sola secuencia de video.

Breve descripción de la invención

En síntesis, la descripción detallada presenta las innovaciones

en el área de la codificación y decodificación adaptativa. Por ejemplo, algunas de las innovaciones se refieren a un codificador que conmuta los espacios de color entre las unidades en una secuencia de video durante la codificación, así como a un
5 decodificador correspondiente. Otras innovaciones se refieren a un codificador que conmuta las frecuencias de muestreo de color entre las unidades en una secuencia de video durante la codificación, así como a un decodificador correspondiente. Aún otras innovaciones se refieren a un codificador que conmuta las profundidades de bits entre
10 las unidades en una secuencia de video durante la codificación, así como a un decodificador correspondiente. Estas innovaciones pueden mejorar la eficacia de codificación en muchos contextos.

De acuerdo con un primer aspecto de las innovaciones descritas en el presente documento, un codificador de imagen o
15 video codifica el video en una secuencia. Como parte de la codificación, el codificador conmuta los espacios de color, las frecuencias de muestreo de color y/o las profundidades de bits de manera espacial o temporal entre la menos algunas unidades (por ejemplo, imágenes de la secuencia, secciones de una determinada
20 imagen de la secuencia, bloques de una determinada imagen de la secuencia) del video en la secuencia. El codificador emite los datos codificados en una corriente de bits. Los datos codificados incluyen una o más señales que indican cómo los espacios de color, las frecuencias de muestreo de color y/o las profundidades de bits se
25 conmutan entre al menos algunas unidades del video en la

secuencia.

Un decodificador correspondiente recibe los datos codificados en una corriente de bits. Los datos codificados incluyen una o más señales que indican cómo espacios de color, frecuencias de muestreo de color y/o profundidades de bits se conmutan entre al menos algunas unidades (por ejemplo, imágenes de la secuencia, secciones de una determinada imagen de la secuencia, bloques de una determinada imagen de la secuencia) del video en una secuencia. El decodificador decodifica los datos codificados. Como parte de la decodificación, el decodificador conmuta los espacios de color, las frecuencias de muestreo de color y la profundidad de bits de manera espacial y/o temporal entre al menos algunas unidades de video en la secuencia.

Por ejemplo, durante la codificación o decodificación, la conmutación incluye el cambio entre dos de los espacios de color usando las operaciones de transformación de espacio de color. Para las operaciones de transformación de espacio de color, los espacios de color pueden incluir al menos un espacio de color tipo YUV y al menos un espacio de color tipo RGB.

O, como otro ejemplo, durante la codificación o decodificación, la conmutación incluye el cambio entre dos de los espacios de color usando las operaciones de reordenamiento del espacio de color. Para las operaciones de reordenamiento de espacio de color, los espacios de color pueden incluir múltiples espacios de color tipo RGB.

O, como otro ejemplo, durante la codificación o decodificación, la conmutación incluye el cambio entre dos de las frecuencias de muestreo de color. Las frecuencias de muestreo de color pueden incluir dos o más de 4:4:4, 4:2:2, 4:2:0 y 4:0:0.

5 O, como otro ejemplo, durante la codificación o decodificación, la conmutación incluye el cambio entre dos de las profundidades de bits. Las profundidades de bits pueden incluir dos o más de 12 bits por muestra, 10 bits por muestra y 8 bits por muestra. Además, los diferentes componentes de color de una determinada unidad del
10 video pueden tener diferentes profundidades de bits.

La sintaxis y semántica de las señales que indican cómo los espacios de color, las frecuencias de muestreo de color y/o las profundidades de bits se conmutan, dependen de la implementación. Por ejemplo, las señales incluyen un valor indicador para una
15 determinada unidad del video. El valor indicador indica una selección entre dos opciones para los espacios de color, las frecuencias de muestreo de color o profundidades de bits. O, las señales incluyen un elemento de sintaxis para una determinada unidad del video. El elemento de sintaxis indica una selección entre tres o más opciones
20 para los espacios de color, las frecuencias de muestreo de color o las profundidades de bits. Una señal puede codificarse de manera predictiva, en cuyo caso un valor delta para un parámetro de una determinada unidad del video representa una diferencia entre un valor existente del parámetro y su predictor. El predictor puede ser
25 un parámetro de capa de secuencia. O, el predictor puede basarse

en uno o más valores existentes del parámetro para una o más unidades anteriores del video.

Además de las señales, los datos codificados pueden incluir la información que identifica o define los espacios de color, las
5 frecuencias de muestreo de color y/o las profundidades de bits. Por ejemplo, la información puede incluir una matriz de valores que pueden usarse para las operaciones de transformación de espacio de color entre dos de los espacios de color.

La codificación o decodificación puede incluir la predicción de
10 los valores de muestra de un bloque actual en una imagen actual con base, al menos en parte, en un conjunto de valores de muestra del contenido previamente reconstruido. Para la predicción intraimagen, el conjunto de valores de muestra del contenido previamente reconstruido es parte de la imagen actual. Para la predicción
15 interimagen, el conjunto de valores de muestra del contenido previamente reconstruido es parte de la imagen de referencia. Como parte de la predicción, el codificador o decodificador puede evaluar si el conjunto de valores de muestra del contenido previamente reconstruido y los valores de muestra del bloque actual están en
20 diferentes formatos.

En algunos casos, el conjunto de valores de muestra del contenido previamente reconstruido se almacena en un primer formato, y los valores de muestra del bloque actual se codifican en un segundo formato diferente al primer formato. En este caso, como
25 parte de la predicción, el codificador o decodificador puede (a)

convertir el conjunto de valores de muestra del contenido previamente reconstruido del primer formato al segundo formato (por ejemplo, conversión entre dos espacios de color, conversión entre dos frecuencias de muestreo de color o conversión entre dos profundidades de bits), (b) predecir los valores de muestra del bloque actual usando el conjunto de valores de muestra convertidos del contenido previamente reconstruido y (c) reconstruir los valores de muestra del bloque actual con base, al menos en parte, en los valores de muestra previstos del bloque actual. O, como parte de la predicción, el codificador o decodificador puede (a) predecir los valores de muestra del bloque actual usando el conjunto de valores de muestra del contenido previamente reconstruido, (b) convertir los valores de muestra previstos del bloque actual del primer formato al segundo formato y (c) reconstruir los valores de muestra del bloque actual con base, al menos en parte, en los valores de muestra previstos convertidos del bloque actual. Después de la predicción, el codificador o decodificador puede (d) convertir los valores de muestra reconstruidos del bloque actual del segundo formato al primer formato, y (e) almacenar los valores de muestra reconstruidos convertidos del bloque actual (en el primer formato) como parte del contenido previamente reconstruido.

La codificación o decodificación puede incluir también el desbloqueo del contenido previamente reconstruido de acuerdo con una o más reglas. Por ejemplo, como parte del desbloqueo, el codificador o decodificador ajusta la intensidad de un filtro de

desbloqueo dependiendo de si los componentes primarios de los dos bloques adyacentes tienen valores residuales distintos de cero. O, como parte del desbloqueo, el codificador o decodificador ajusta la intensidad de un filtro de desbloqueo dependiendo de si los componentes correspondientes de los dos bloques adyacentes tienen valores residuales distintos de cero. Por lo tanto, las reglas pueden determinar los diferentes espacios de color de los dos bloques adyacentes.

Las innovaciones para la conmutación adaptativa de los espacios de color, frecuencias de muestreo de color y/o profundidades de bits, pueden implementarse como parte de un método, como parte de un dispositivo de proceso de datos adaptado para llevar a cabo el método o como parte de un medio legible por computadora tangible que almacena las instrucciones ejecutables en computadora para hacer que un dispositivo de proceso de datos lleve a cabo el método. Varias innovaciones pueden usarse en combinación o por separado. En particular, un codificador y decodificador de video pueden conmutar de manera adaptativa solo los espacios de color en una secuencia de video (con una frecuencia de muestreo de color fija y una profundidad de bits fija). O, un codificador y decodificador de video pueden conmutar de manera adaptativa solo las frecuencias de muestreo de color en una secuencia de video (con un espacio de color fijo y una profundidad de bits fija). O, un codificador y decodificador de video pueden conmutar de manera adaptativa solo las profundidades de bits en

una secuencia de video (con un espacio de color fijo y una frecuencia de muestreo de color fija). O, un codificador y decodificador de video pueden conmutar de manera adaptativa (1) los espacios de color y las frecuencias de muestreo de color, pero no las profundidades de bits, en una secuencia de video, (2) los espacios de color y las profundidades de bits, pero no las frecuencias de muestreo de color, en una secuencia de video, o (3) las frecuencias de muestreo de color y las profundidades de bits, pero no los espacios de color, en una secuencia de video. O, un codificador y decodificador de video pueden conmutar de manera adaptativa los espacios de color, las frecuencias de muestreo de color y las profundidades de bits en una secuencia de video.

Los anteriores y otros objetivos, características y ventajas de la invención serán más evidentes a partir de la siguiente descripción detallada, que procede con referencia a las figuras adjuntas.

Breve descripción de los dibujos

La figura 1 es un diagrama de un sistema de proceso de datos ejemplar en el cual pueden implementarse algunas modalidades descritas.

Las figuras 2a y 2b son diagramas de los entornos de red ejemplares en los cuales pueden implementarse algunas modalidades descritas.

La figura 3 es un diagrama de un sistema codificador ejemplar junto con el cual pueden implementarse algunas modalidades

descritas.

La figura 4 es un diagrama de un sistema decodificador ejemplar junto con el cual pueden implementarse algunas modalidades descritas.

5 Las figuras 5a y 5b son diagramas que ilustran un codificador de video ejemplar junto con el cual pueden implementarse algunas modalidades descritas.

La figura 6 es un diagrama que ilustra un decodificador de video ejemplar junto con el cual pueden implementarse algunas
10 modalidades descritas.

La figura 7 es un diagrama que ilustra un entorno de escritorio de computadora con el contenido que puede proporcionar la salida para la captura de pantalla.

La figura 8 es un diagrama que ilustra el video compuesto con el contenido de video natural y el contenido de video artificial.
15

La figura 9 es un diagrama que ilustra los espacios de color, las frecuencias de muestreo de color y/o las profundidades de bits de adaptación a la imagen para las imágenes en una secuencia.

La figura 10 es un diagrama que ilustra los espacios de color, las frecuencias de muestreo de color y/o las profundidades de bits de adaptación a la sección para las secciones en una secuencia.
20

La figura 11 es un diagrama que ilustra los espacios de color, las frecuencias de muestreo de color y/o las profundidades de bits de adaptación al bloque para los bloques de una sección de una imagen
25 en una secuencia.

La figura 12 es un diagrama de flujo que ilustra una técnica generalizada para conmutar de manera adaptativa los espacios de color, las frecuencias de muestreo de color y/o las profundidades de bits durante la codificación, y la figura 13 es un diagrama de flujo que ilustra una técnica ejemplar más detallada para conmutar de manera adaptativa en una base de unidad por unidad durante la codificación.

La figura 14 es un diagrama de flujo que ilustra una técnica generalizada para conmutar de manera adaptativa los espacios de color, las frecuencias de muestreo de color y/o las profundidades de bits durante la decodificación, y la figura 15 es un diagrama de flujo que ilustra una técnica ejemplar más detallada para conmutar de manera adaptativa en una base de unidad por unidad durante la decodificación.

La figura 16 es un diagrama de flujo que ilustra una técnica para la predicción de los valores de muestra de un bloque codificados en un formato a partir de un conjunto de valores de muestra del contenido previamente reconstruido en un diferente formato.

20

Descripción detallada de la invención

La descripción detallada presenta las innovaciones en el área de la codificación y decodificación adaptativa. Por ejemplo, algunas de las innovaciones se refieren a un codificador que conmuta los espacios de color entre las unidades en una secuencia de video

25

durante la codificación, así como a un decodificador correspondiente. Otras innovaciones se refieren a un codificador que conmuta las frecuencias de muestreo de color entre las unidades en una secuencia de video durante la codificación, así como a un decodificador correspondiente. Aún otras innovaciones se refieren a un codificador que conmuta las profundidades de bits entre las unidades en una secuencia de video durante la codificación, así como a un decodificador correspondiente. Estas innovaciones pueden mejorar la eficacia de codificación en muchos contextos.

En general, un espacio de color (ocasionalmente denominado modelo de color) es un modelo que representa los colores como valores n por posición física, para $n \geq 1$, donde cada uno de los valores n proporciona un valor de componente de color para esa posición.

Por ejemplo, en un espacio de color YUV, un valor de componente de luma (o Y) representa un brillo aproximado en una posición y los múltiples valores de componente de croma (o U y V) representan las diferencias de color en la posición. Las definiciones precisas de los valores de diferencia de color (y las operaciones de conversión a/de un espacio de color YUV a otro espacio de color como RGB) dependen de la implementación. Por lo general, a efectos de codificación y decodificación, el componente Y es el componente primario, y los componentes U e V son los componentes secundarios. En general, tal como se usa en el presente documento, el término espacio de color tipo YUV indica cualquier espacio de

color con un componente de luma (o luminosidad) y uno o más componentes de croma (o crominancia), que incluyen $Y'UV$, YIQ , $Y'IQ$ e $YDbDr$, así como variaciones como $YCbCr$ e $YCoCg$.

Las medidas de señal de componente que se usan pueden
5 ajustarse mediante la aplicación de una función de las características de transferencia no lineal (generalmente conocida como "precompensación de gamma" y se representa a menudo por el uso de un símbolo de virgulilla, aunque el símbolo de virgulilla se omite a menudo por conveniencia tipográfica). O, las medidas de
10 señal de componente pueden estar en un dominio que tiene una relación lineal con la amplitud de luz. Las señales de componente de luma y croma pueden estar bien alineadas con la percepción de brillo y color para el sistema visual humano, o las señales de componente de luma y croma pueden desviarse un poco de tales medidas (por
15 ejemplo, como en la variación de $YCoCg$, donde se aplican las fórmulas que simplifican el proceso de datos de los valores de componente de color).

O, en un espacio de color RGB, un valor de componente de rojo (R) representa una intensidad de rojo, un valor de componente
20 de verde (G) representa una intensidad de verde y un valor de componente de azul (B) representa una intensidad de azul en una posición. Tal como se usa en el presente documento, el término espacio de color tipo RGB indica un espacio de color con los componentes de color R, G y B en cualquier orden. Los ejemplos
25 incluyen los espacios de color RGB, BGR y GBR, que se diferencian

en cuanto al componente principal para propósitos de codificación y decodificación. El componente primario se indica con la primera letra del espacio de color (por ejemplo, R para RGB).

La frecuencia de muestreo de color (ocasionalmente denominada frecuencia de muestreo de croma) se refiere a la resolución espacial relativa entre los componentes de color. Por ejemplo, para una frecuencia de muestreo de color de 4:4:4, la información para los componentes secundarios (por ejemplo, los componentes U y V para YUV) tiene la misma resolución espacial que la información para un componente primario (por ejemplo, el componente Y para YUV). Para una frecuencia de muestreo de color de 4:2:2 o 4:2:0, la información para los componentes secundarios se someterá a reducción de muestra con relación a la información para el componente primario. Un formato YUV a 4:2:0 es un formato que submuestrea la información de croma en comparación con un formato YUV a 4:4:4, de modo que la resolución de croma sea la mitad de la resolución de luma en sentido tanto horizontal como vertical. Un formato YUV a 4:2:2 es un formato que submuestrea la información de croma en comparación horizontal con un formato YUV a 4:4:4, de modo que la resolución de croma sea la mitad de la resolución de luma en sentido horizontal. Otros ejemplos de las frecuencias de muestreo de color son 4:1:1 (los componentes secundarios tienen un cuarto de resolución en sentido horizontal) y 4:0:0 (los componentes secundarios se descartan). El submuestreo de color por lo general se aplica a los espacios de color tipo YUV. Los espacios de color tipo

RGB suelen tener una frecuencia de muestreo de color de 4:4:4, pero pueden tener una diferente frecuencia de muestreo de color dependiendo de que componentes de color secundarios sean submuestreados.

5 Aunque el formato YUV a 4:2:0 se usa convencionalmente para la codificación y decodificación de video, existen algunos casos de uso para cuyo video tiene una información de color más abundante, y puede justificarse la mayor fidelidad de color. En tales casos de uso, las diferencias entre los formatos de muestreo de croma YUV a 4:4:4
10 y YUV a 4:2:0 las perciben más fácilmente los espectadores. Por ejemplo, para la codificación-decodificación del contenido de texto de la pantalla de computadora, contenido de video animado con límites artificiales muy nítidos, o ciertas características del contenido de video de manera más general (como desplazamiento de rótulos y
15 gráficos muy nítidos, o un video con información concentrada en los canales de croma), puede preferirse el formato 4:4:4 en lugar de un formato 4:2:0.

La profundidad de bits se refiere al número de bits por valor de muestra. Las profundidades de bits son de 8 bits por muestra, 10 bits
20 por muestra y 12 bits por muestra. Otras posibles profundidades de bits incluyen 4 bits por muestra y 16 bits por muestra.

Aunque las operaciones descritas en el presente documento, en algunos sitios se describen como realizadas por un codificador o decodificador de video, en muchos casos las operaciones pueden
25 realizarse por otro tipo de herramienta de procesamiento de medios

(por ejemplo, codificador o decodificador de imagen). Por ejemplo, las operaciones pueden realizarse para aplicaciones como codificación-decodificación de imagen estática, codificación-decodificación de contenido de exploración médica, codificación-decodificación de contenido de imágenes multiespectrales, etcétera.

Algunas de las innovaciones descritas en el presente documento se ilustran con referencia a elementos de sintaxis y operaciones específicas para el estándar H.265/HEVC. Por ejemplo, se hace referencia a la versión preliminar JCTVC-P1005 del estándar H.265/HEVC - High Efficiency Video Coding (HEVC) Range Extensions Text Specification: Draft 6, JCTVC-P1005_v1, febrero de 2014. También pueden implementarse las innovaciones descritas en el presente documento para otros estándares o formatos.

Más generalmente, son posibles diversas alternativas a los ejemplos descritos en el presente documento. Por ejemplo, algunos de los métodos descritos en el presente documento pueden modificarse al cambiar el orden de las acciones descritas del método, al dividir, repetir u omitir ciertas acciones del método, etcétera. Los diversos aspectos de la tecnología divulgada pueden usarse en combinación o por separado. Diferentes modalidades usan una o más de las innovaciones descritas. Algunas de las innovaciones descritas en el presente documento abordan uno o más de los problemas observados anteriormente. Comúnmente, una determinada técnica/herramienta no resuelve todos esos problemas.

I. Sistemas de proceso de datos ejemplares.

La figura 1 ilustra un ejemplo generalizado de un sistema de proceso de datos conveniente (100) en el cual pueden implementarse varias de las innovaciones descritas. El sistema de proceso de datos (100) no pretende sugerir ninguna limitación en cuanto al ámbito de uso o funcionalidad, puesto que las innovaciones pueden implementarse en diversos sistemas de proceso de datos de uso general o especial.

Con referencia a la figura 1, el sistema de proceso de datos (100) incluye una o más unidades de procesamiento (110, 115) y una memoria (120, 125). Las unidades de procesamiento (110, 115) ejecutan las instrucciones ejecutables en computadora. Una unidad de procesamiento puede ser una unidad central de procesamiento ("CPU", por sus siglas en inglés) de uso general, un procesador en un circuito integrado de aplicación específica ("ASIC", por sus siglas en inglés) o cualquier otro tipo de procesador. En un sistema de múltiple procesamiento, las múltiples unidades de procesamiento ejecutan las instrucciones ejecutables en computadora para aumentar la potencia de procesamiento. Por ejemplo, la figura 1 muestra una unidad central de procesamiento (110), así como una unidad de procesamiento gráfico o una unidad de coprocesamiento (115). La memoria tangible (120, 125) puede ser una memoria transitoria (por ejemplo, registros, memoria caché, RAM), memoria permanente (por ejemplo, ROM, EEPROM, memoria Flash, etcétera), o alguna combinación de dos, accesibles por las unidades de procesamiento. La memoria (120, 125) almacena el *software* (180)

que implementa una o más innovaciones para la conmutación adaptativa de los espacios de color, frecuencias de muestreo de color y/o profundidades de bits, en forma de instrucciones ejecutables en computadora adecuadas para la ejecución por medio
5 de las unidades de procesamiento.

Un sistema de proceso de datos puede tener características adicionales. Por ejemplo, el sistema de proceso de datos (100) incluye un almacenamiento (140), uno o más dispositivos de entrada (150), uno o más dispositivos de salida (160) y una o más
10 conexiones de comunicación (170). Un mecanismo de interconexión (no ilustrado) como un bus, controlador o red, interconecta los componentes del sistema de proceso de datos (100). Por lo general, *software* de sistema operativo (no ilustrado) proporciona un entorno operativo para otro *software* que se ejecuta en el sistema de proceso
15 de datos (100) y coordina las actividades de los componentes del sistema de proceso de datos (100).

El almacenamiento tangible (140) puede ser extraíble o no extraíble e incluye discos magnéticos, cintas magnéticas o casetes, CD-ROM, DVD o cualquier otro medio que pueda usarse para
20 almacenar información y que pueda consultarse en el sistema de proceso de datos (100). El almacenamiento (140) almacena las instrucciones para el *software* (180) que implementa una o más innovaciones para la conmutación adaptativa de espacios de color, frecuencias de muestreo de color o profundidades de bits.

25 Los dispositivos de entrada (150) pueden ser un dispositivo de

entrada táctil como un teclado, ratón, lápiz, bola rastreadora, un dispositivo de entrada de voz, un dispositivo de exploración u otro dispositivo que proporcione la entrada al sistema de proceso de datos (100). Para el video, los dispositivos de entrada (150) pueden ser una cámara, tarjeta de video, tarjeta sintonizadora de TV, módulo de captura de pantalla o dispositivo similar que acepte la entrada de video en forma analógica o digital, o un CD-ROM o CD-RW que lee la entrada de video en el sistema de proceso de datos (100). Los dispositivos de salida (160) pueden ser una pantalla, impresora, altavoz, grabadora de CD u otro dispositivo que proporcione la salida desde el sistema de proceso de datos (100).

Las conexiones de comunicación (170) permiten la comunicación en un medio de comunicación a otra entidad de proceso de datos. El medio de comunicación transporta la información como las instrucciones ejecutables en computadora, entrada o salida de audio o video, u otros datos en una señal de datos modulada. Una señal de datos modulada es una señal que tiene una o más de sus características establecidas o modificadas de tal manera que codifique la información en la señal. A modo de ejemplo y sin limitación, los medios de comunicación pueden usar un soporte eléctrico, óptico, RF, u otro soporte.

Las innovaciones pueden describirse en el contexto general de los medios de legibles por computadora. Los medios legibles por computadora son cualquier medio tangible que pueda consultarse dentro de un entorno de proceso de datos. A modo de ejemplo y sin

limitación, con el sistema de proceso de datos (100), los medios legibles por computadora incluyen una memoria (120, 125), almacenamiento (140) y combinaciones de cualquiera de los anteriores.

5 Las innovaciones pueden describirse en el contexto general de las instrucciones ejecutables en computadora, como aquellas incluidas en los módulos de programa, que se ejecutan en un sistema de proceso de datos en un procesador de destino existente o virtual. Generalmente, los módulos de programa incluyen rutinas,
10 programas, bibliotecas, objetos, clases, componentes, estructuras de datos, etcétera, que realizan las tareas particulares o implementan los tipos de datos abstractos particulares. La funcionalidad de los módulos de programa puede combinarse o dividirse entre los módulos de programa según se desee en diferentes modalidades.
15 Las instrucciones ejecutables en computadora para los módulos de programa pueden ejecutarse dentro de un sistema de proceso de datos local o distribuido.

Los términos "sistema" y "dispositivo" se usan de manera indistinta en el presente documento. A menos que el contexto
20 claramente indique otra cosa, el término no implica limitación alguna en un tipo de sistema de proceso de datos o dispositivo de proceso de datos. En general, un sistema de proceso de datos o dispositivo de proceso de datos puede ser local o distribuido, y puede incluir cualquier combinación de *hardware* de uso especial y/o *hardware* de
25 uso general con el *software* que implementa la funcionalidad descrita

en el presente documento.

También pueden implementarse los métodos divulgados usando el *hardware* de proceso de datos especializado configurado para realizar cualquiera de los métodos divulgados. Por ejemplo, pueden
5 implementarse los métodos divulgados por medio de un circuito integrado (por ejemplo, un ASIC (como un procesador de señales digitales ("DSP", por sus siglas en inglés) con ASIC, una unidad de procesamiento gráfico ("GPU", por sus siglas en inglés) o un dispositivo lógico programable ("PLD", por sus siglas en inglés),
10 como un dispositivo multicircuito integrado programable por el usuario ("FPGA", por sus siglas en inglés)) especialmente diseñado o configurado para implementar cualquiera de los métodos divulgados.

En aras de la presentación, la descripción detallada usa términos como "determinar" y "usar" para describir las operaciones
15 de la computadora en un sistema de proceso de datos. Estos términos son abstracciones de nivel avanzado para las operaciones realizadas por una computadora y no deben confundirse con las acciones realizadas por un ser humano. Las operaciones existentes de la computadora correspondientes a estos términos varían
20 dependiendo de la implementación.

II. Entornos de red ejemplares.

Las figuras 2a y 2b muestran los entornos de red ejemplares (201, 202) que incluyen los codificadores de video (220) y los decodificadores de video (270). Los codificadores (220) y los
25 decodificadores (270) están conectados en una red (250) usando un

protocolo de comunicación adecuado. La red (250) puede incluir el Internet u otra red de computadoras.

En el entorno de red (201) mostrado en la figura 2a, cada herramienta de comunicación en tiempo real ("RTC", por sus siglas en inglés) (210) incluye tanto un codificador (220) como un decodificador (270) para la comunicación bidireccional. Un determinado codificador (220) puede producir la salida compatible con una variación o extensión del estándar H.265/HEVC, estándar SMPTE 421M, estándar ISO-IEC 14496-10 (también denominado como H.264 o AVC), otro estándar o un formato propietario, con un decodificador correspondiente (270) que acepta los datos codificados desde el codificador (220). La comunicación bidireccional puede ser parte de una videoconferencia, videollamada u otro contexto de comunicación bipartita o multipartita. Aunque el entorno de red (201) en la figura 2a incluye dos herramientas de comunicación en tiempo real (210), el entorno de red (201) en cambio puede incluir tres o más herramientas de comunicación en tiempo real (210) que participan en la comunicación multipartita.

Una herramienta de comunicación en tiempo real (210) controla la codificación mediante un codificador (220). La figura 3 muestra un sistema codificador ejemplar (300) que puede incluirse en la herramienta de comunicación en tiempo real (210). Por otra parte, la herramienta de comunicación en tiempo real (210) usa otro sistema codificador. Una herramienta de comunicación en tiempo real (210) también gestiona la decodificación por medio de un decodificador

(270). La figura 4 muestra un sistema decodificador ejemplar (400) que puede incluirse en la herramienta de comunicación en tiempo real (210). Por otra parte, la herramienta de comunicación en tiempo real (210) usa otro sistema decodificador.

5 En el entorno de red (202) mostrado en la figura 2b, una herramienta de codificación (212) incluye un codificador (220) que codifica el video para el suministro a las múltiples herramientas de reproducción (214), las cuales pueden incluir los decodificadores (270). La comunicación unidireccional puede proporcionarse por un

10 sistema de videovigilancia, sistema de monitoreo por cámara web, módulo de captura de pantalla, presentación de conferencias de escritorio remoto u otro contexto en el cual el video se codifica y envía de una ubicación a una o más otras ubicaciones. Aunque el entorno de red (202) en la figura 2b incluye dos herramientas de

15 reproducción (214), el entorno de red (202) puede incluir más o menos herramientas de reproducción (214). En general, una herramienta de reproducción (214) se comunica con la herramienta de codificación (212) para determinar una emisión en continuo de video para que la reciba la herramienta de reproducción (214). La

20 herramienta de reproducción (214) recibe la emisión en continuo, introduce en memoria intermedia los datos codificados recibidos durante un período adecuado y comienza la decodificación y la reproducción.

 La figura 3 muestra un sistema codificador ejemplar (300) que

25 puede incluirse en la herramienta de codificación (212). Por otra

parte, la herramienta de codificación (212) usa otro sistema codificador. La herramienta de codificación (212) también puede incluir la lógica de controlador del servidor para la gestión de las conexiones con una o más herramientas de reproducción (214). La figura 4 muestra un sistema decodificador ejemplar (400), que puede incluirse en la herramienta de reproducción (214). Por otra parte, la herramienta de reproducción (214) usa otro sistema decodificador. Una herramienta de reproducción (214) también puede incluir la lógica de control del cliente para gestionar las conexiones con la herramienta de codificación (212).

III. Sistemas codificadores ejemplares.

La figura 3 es un diagrama de bloques de un sistema codificador ejemplar (300) junto con el cual pueden implementarse algunas modalidades descritas. El sistema codificador (300) puede ser una herramienta de codificación de uso general capaz de operar en cualquiera de los múltiples modos de codificación como un modo de codificación de baja latencia para la comunicación en tiempo real, un modo de transcodificación y un modo de codificación de alta latencia para producir los medios para la reproducción a partir de un archivo o una emisión en continuo, o puede ser una herramienta de codificación de uso especial adaptada para tal modo de codificación. El sistema codificador (300) puede adaptarse para la codificación de un tipo particular de contenido (por ejemplo, contenido de captura de pantalla), o puede adaptarse para la codificación de cualquiera de varios tipos diferentes de contenido (por ejemplo, contenido de

captura de pantalla y de video natural). El sistema codificador (300) puede implementarse como un módulo del sistema operativo, como parte de una biblioteca de aplicación o como una aplicación independiente. En general, el sistema codificador (300) recibe una
5 secuencia de fotogramas de video de origen (311) desde una fuente de video (310) y produce los datos codificados como salida a un canal (390). La salida de datos codificados al canal puede incluir el contenido codificado con la conmutación adaptativa de los espacios de color, frecuencias de muestreo de color y/o profundidades de bits.

10 La fuente de video (310) puede ser una cámara, tarjeta sintonizadora, medios de almacenamiento, módulo de captura de pantalla u otra fuente de video digital. La fuente de video (310) produce una secuencia de fotogramas de video a una frecuencia de fotogramas de, por ejemplo, 30 fotogramas por segundo. Tal como se
15 usa en el presente documento, el término "fotograma" generalmente se refiere a los datos de imagen de origen codificados o reconstruidos. Para el video de exploración progresiva, un fotograma es un fotograma de video de exploración progresiva. Para el video entrelazado, en las modalidades ejemplares, un fotograma de video
20 entrelazado puede desentrelazarse antes de la codificación. Por otra parte, dos campos de video entrelazado complementarios se codifican juntos como un solo fotograma de video o se codifican como dos campos codificados por separado. Además de indicar un fotograma de video de exploración progresiva o fotograma de video
25 de exploración entrelazada, el término "fotograma" o "imagen" puede

indicar un único campo de video no pareado, un par complementario de campos de video, un plano de objeto de video que representa un objeto de video en un determinado momento, o una región de interés en una imagen más grande. El plano o región de objeto de video
5 puede ser parte de una imagen más grande que incluye múltiples objetos o regiones de una escena.

Un siguiente fotograma de origen (311) se almacena en un área de almacenamiento de memoria temporal de fotograma de origen (320) que incluye múltiples áreas de almacenamiento de memoria
10 intermedia de fotograma (321, 322, ..., 32n). Una memoria intermedia de fotograma (321, 322, etcétera) contiene un fotograma de origen en el área de almacenamiento de fotograma de origen (320). Después de que uno o más de los fotogramas de origen (311) se han almacenado en las memorias intermedias de fotograma (321, 322,
15 etcétera), un selector de fotograma (330) selecciona un fotograma de origen individual en el área de almacenamiento de fotograma original (320). El orden en el cual los fotogramas se seleccionan por el selector de fotograma (330) para la entrada al codificador (340) puede diferir del orden en el cual los fotogramas se producen por la
20 fuente de video (310), por ejemplo, la codificación de algunos fotogramas puede retrasarse en orden, para permitir que algunos fotogramas posteriores se codifiquen primero y así facilitar la predicción temporal regresiva. Antes del codificador (340), el sistema codificador (300) puede incluir un preprocessor (no mostrado) que
25 realiza un preprocesamiento (por ejemplo, filtración) del fotograma

seleccionado (331) antes de la codificación.

El codificador (340) codifica el fotograma seleccionado (331) para producir un fotograma codificado (341) y también produce las señales de la operación de control de gestión de memoria ("MMCO", por sus siglas en inglés) (342) o la información del conjunto de imágenes de referencia ("RPS", por sus siglas en inglés). El RPS es el conjunto de fotogramas que puede usarse para la referencia en la compensación de movimiento para un fotograma actual o cualquier fotograma subsiguiente. Si el fotograma actual no es el primer fotograma que se ha codificado, cuando se realiza su proceso de codificación, el codificador (340) puede usar uno o más fotogramas previamente codificados-descodificados (369) que se han almacenado en un área de almacenamiento de memoria temporal de fotograma decodificado (360). Tales fotogramas decodificados almacenados (369) se usan como fotogramas de referencia para la predicción interfotograma del contenido del fotograma de origen actual (331). La información de MMCO/RPS (342) indica a un decodificador cuales fotogramas reconstruidos puede usar como fotogramas de referencia, y, por lo tanto, debe almacenar en un área de almacenamiento de fotograma.

El codificador (340) acepta el video en un espacio de color particular (por ejemplo, un espacio de color tipo YUV, un espacio de color tipo RGB), con una frecuencia de muestreo de color particular (por ejemplo, 4:4:4) y un número particular de bits por muestra (por ejemplo, 12 bits por muestra). Durante la codificación, para

diferentes imágenes, secciones, bloques u otras unidades de video, el codificador (340) puede realizar las conversiones de espacio de color para transformar un espacio de color tipo YUV y un espacio de color tipo RGB, o a/de algún otro espacio de color. El codificador (340) también puede realizar las conversiones de espacio de color para reordenar los componentes de color, cambiar el componente de color que es el componente primario (por ejemplo, convertir entre los formatos RGB, BGR y GBR). Durante la codificación, el codificador (340) también puede realizar el procesamiento de remuestreo para cambiar las frecuencias de muestreo de color (por ejemplo, los formatos 4:4:4, 4:2:2 y 4:2:0) para diversas imágenes, secciones, bloques u otras unidades del video. El codificador (340) también puede cambiar las profundidades de bits (por ejemplo, entre 12 bits por muestra, 10 bits por muestra y 8 bits por muestra) durante la codificación de diferentes imágenes, secciones, bloques u otras unidades del video. En algunas implementaciones ejemplares, el codificador (340) puede conmutar espacios de color, frecuencias de muestreo de color y/o profundidades de bits en una base de imagen por imagen durante la codificación.

Generalmente, el codificador (340) incluye múltiples módulos de codificación que realizan las tareas de codificación como segmentar en piezas, adaptación de espacio de color, frecuencia de muestreo de color y/o profundidad de bits, estimación intrapredicción y predicción, estimación y compensación de movimiento, transformación de frecuencia, cuantificación y codificación por

entropía. Las operaciones exactas realizadas por el codificador (340) pueden variar dependiendo del formato de compresión. El formato de los datos codificados de salida puede ser una variación o extensión del formato H.265/HEVC, formato de Windows Media Video, formato VC-1, formato MPEG-x (por ejemplo, MPEG-1, MPEG-2 o MPEG-4),
5 formato H.26x (por ejemplo, H.261, H.262, H.263, H.264), u otro formato.

El codificador (340) puede segmentar un fotograma en múltiples piezas del mismo tamaño o diferentes tamaños. Por
10 ejemplo, el codificador (340) divide el fotograma a lo largo de filas de piezas y columnas de piezas que, con límites de fotograma, definen los límites horizontales y verticales de las piezas en el fotograma, donde cada pieza es una región rectangular. Las piezas se usan a menudo para proporcionar las opciones para el
15 procesamiento en paralelo. Un fotograma también puede organizarse como uno o más secciones, donde una sección puede ser un fotograma completo o una región del fotograma. Una sección puede decodificarse independientemente de otras secciones en un fotograma, que mejora la capacidad de recuperación de error. El
20 contenido de una sección o una pieza adicional se segmenta en bloques o en otros conjuntos de valores de muestra para fines de codificación y decodificación. En algunas implementaciones ejemplares, el codificador (340) puede conmutar espacios de color, frecuencias de muestreo de color y/o profundidades de bits en una
25 base de sección por sección durante la codificación.

Para la sintaxis de acuerdo con el estándar H.265/HEVC, el codificador divide el contenido de un fotograma (o sección o pieza) en unidades de codificación en árbol. Una unidad de codificación en árbol ("CTU", por sus siglas en inglés) incluye los valores de muestras de luma organizados como un bloque de codificación en árbol ("CTB", por sus siglas en inglés) de luma y los valores de muestra de croma correspondientes organizados como dos CTB de croma. El tamaño de una CTU (y sus CTB) se selecciona por el codificador y puede ser, por ejemplo, valores de prueba de 64 x 64, 32 x 32 o 16 x 16. Una CTU incluye una o más unidades de codificación. Una unidad de codificación ("CU", por sus siglas en inglés) tiene un bloque de codificación ("CB", por sus siglas en inglés) de luma y dos CB de croma correspondiente. Por ejemplo, una CTU con un CTB de luma de 64 x 64 y dos CTB de croma de 64 x 64 (formato YUV a 4:4:4) puede dividirse en cuatro CU, cada CU incluye un CB de luma de 32 x 32 y dos CB de croma de 32 x 32 y cada CU posiblemente se divide aún más en CU más pequeñas. O, como otro ejemplo, una CTU con un CTB de luma de 64 x 64 y dos CTB de croma de 32 x 32 (formato YUV a 4:2:0) puede dividirse en cuatro CU, cada CU incluye un CB de luma de 32 x 32 y dos CB de croma de 16 x 16 y cada CU posiblemente se divide aún más en CU más pequeñas. El tamaño más pequeño permitido de la CU (por ejemplo, 8 x 8, 16 x 16) puede señalizarse en la corriente de bits.

Generalmente, una CU tiene un modo de predicción como entre varios elementos o dentro de varios elementos. Una CU incluye una

o más unidades de predicción para los propósitos de señalización de la información de predicción (como detalles del modo de predicción, valores de desplazamiento, etcétera) y/o procesamiento de predicción. Una unidad de predicción ("PU", por sus siglas en inglés) tiene un bloque de predicción ("PB", por sus siglas en inglés) de luma y dos PB de croma. Para una CU intraprevista, la PU tiene el mismo tamaño que la CU, a menos que la CU tenga el tamaño más pequeño (por ejemplo, 8 x 8). En ese caso, la CU puede dividirse en cuatro PU más pequeñas (por ejemplo, cada 4 x 4 Si el tamaño de CU más pequeño es 8 x 8) o la PU puede tener el tamaño de CU más pequeño, tal como se indica por un elemento de sintaxis para la CU. Una CU también tiene una o más unidades de transformación para los propósitos de codificación-decodificación residual, donde una unidad de transformación ("TU", por sus siglas en inglés) tiene un bloque de transformación ("TB", por sus siglas en inglés) y dos TB de croma. Una PU en una CU intraprevista puede contener una sola TU (igual en tamaño a la PU) o múltiples TU. El codificador decide cómo segmentar el video en UTC, CU, TU, PU, etcétera.

En las implementaciones de H.265/HEVC, una sección puede incluir un solo segmento de sección (segmento de sección independiente) o dividirse en múltiples segmentos de sección (segmento de sección independiente y uno o más segmentos de sección dependientes). Un segmento de sección es un número entero de las CTU ordenadas consecutivamente en una exploración de pieza, contenido en una unidad de capa de abstracción de red

("NAL", por sus siglas en inglés). Para un segmento de sección independiente, un encabezado de segmento de sección incluye los valores de los elementos de sintaxis que se aplican para el segmento de sección independiente. Para un segmento de sección dependiente, un encabezado de segmento de sección truncado incluye algunos valores de los elementos de sintaxis que se aplican para ese segmento de sección dependiente y los valores de los otros elementos de sintaxis para el segmento de sección dependiente se deducen a partir de los valores para el segmento de sección independiente anterior en el orden de decodificación.

Tal como se usa en el presente documento, el término "bloque" puede indicar un macrobloque, una unidad de predicción, una unidad de datos residuales, o un CB, PB o TB, o algún otro conjunto de valores de muestra, dependiendo del contexto. En algunas implementaciones ejemplares, el codificador (340) puede conmutar espacios de color, frecuencias de muestreo de color y/o profundidades de bits en una base de bloque por bloque durante la codificación.

Volviendo a la figura 3, el codificador representa un bloque intracodificado de un fotograma de origen (331) en términos de predicción a partir de otros valores de muestra previamente reconstruidos en el fotograma (331). Para la predicción dentro de las copias de bloque ("BC", por sus siglas en inglés), un estimador intraimagen estima el desplazamiento de un bloque con respecto a otros valores de muestra previamente reconstruidos. Una región de

referencia de predicción intrafotograma es una región de los valores de muestra en el fotograma que se usan para generar los valores de predicción de BC para el bloque. La región de predicción intrafotograma puede indicarse con un valor de vector de bloque ("BV", por sus siglas en inglés) (determinado en la estimación de BV). Para la predicción intraespacial para un bloque, el estimador intrainimagen estima la extrapolación de los valores de muestra reconstruidos próximos en el bloque. El estimador intrainimagen puede emitir la información de predicción (como valores de BV para la predicción dentro de BC, o el modo (dirección) de predicción para la predicción intraespacial), que se codifica por entropía. Un predictor de predicción intrafotograma aplica la información de predicción para determinar los valores intrapredicción.

El codificador (340) representa un bloque previsto codificado interfotograma de un fotograma de origen (331) en términos de la predicción a partir de los fotogramas de referencia. Un estimador de movimiento calcula el movimiento del bloque con respecto a uno o más fotogramas de referencia (369). Cuando se usan múltiples fotogramas de referencia, los múltiples fotogramas de referencia pueden ser de diferentes direcciones temporales o de la misma dirección temporal. Una región de referencia de predicción con compensación de movimiento es una región de los valores de muestra en los fotogramas de referencia que se usan para generar los valores de predicción con compensación de movimiento para un bloque de los valores de muestra de un fotograma actual. El

estimador de movimiento emite la información de movimiento como la información de vector de movimiento ("MV", por sus siglas en inglés), que se codifica por entropía. Un compensador de movimiento aplica los MV a los fotogramas de referencia (369) para determinar los valores de predicción compensada para la predicción interfotograma.

El codificador puede determinar las diferencias (si existen) entre los valores de predicción de un bloque (dentro de ellos o entre ellos) y los valores originales correspondientes. Estos valores de predicción residuales además se codifican usando una transformación de frecuencia, cuantificación y la codificación entrópica. Por ejemplo, el codificador (340) establece los valores para el parámetro de cuantificación ("QP", por sus siglas en inglés) para una imagen, pieza, sección y/u otra porción del video y cuantifica por consiguiente los coeficientes de transformación. El codificador entrópico del codificador (340) comprime los valores cuantificados de coeficiente de transformación, así como cierta información secundaria (por ejemplo, información de MV, valores de índice para predictores de BV, diferenciales de BV, valores de QP, decisiones de modo, opciones de parámetro). Las técnicas de codificación entrópica comunes incluyen codificación Golomb exponencial, codificación Golomb-Rice, codificación aritmética, codificación diferencial, codificación de Huffman, codificación de longitud de ejecución, codificación de longitud variable a longitud variable ("V2V", por sus siglas en inglés), codificación de longitud variable a longitud fija ("V2F", por sus siglas en inglés), codificación

Lempel-Ziv "(LZ", por sus siglas en inglés), codificación de diccionario, codificación entrópica de fraccionamiento de intervalo de probabilidad ("PIPE", por sus siglas en inglés) y combinaciones de las anteriores. El codificador entrópico puede usar diferentes técnicas de codificación para diferentes tipos de información, puede
5 aplicar múltiples técnicas en combinación (por ejemplo, al aplicar la codificación Golomb-Rice seguida de la codificación aritmética) y puede elegir entre múltiples tablas de código dentro de una técnica de codificación particular.

10 Un filtro de desbloqueo adaptativo se incluye dentro del bucle de compensación de movimiento en el codificador (340) para ajustar las discontinuidades a través de las filas y/o columnas del límite de bloque en un fotograma decodificado. Otra filtración (como, filtración de eliminación de distorsión, filtración de bucle adaptativo ("ALF",
15 por sus siglas en inglés), o filtración de desplazamiento de adaptación a muestra ("SAO", por sus siglas en inglés); no mostradas) puede alternativamente o además aplicarse como operaciones de filtración en el bucle.

Los datos codificados producidos por el codificador (340)
20 incluyen los elementos de sintaxis para varias capas de sintaxis de corriente de bits. Para la sintaxis de acuerdo con el estándar H.265/HEVC, por ejemplo, un conjunto de parámetros de imagen ("PPS", por sus siglas en inglés) es una estructura de sintaxis que contiene los elementos de sintaxis que pueden asociarse con una
25 imagen. En algunas implementaciones ejemplares, un PPS puede

incluir una o más señales que indican un espacio de color, frecuencia de muestreo de color y/o profundidad de bits que se aplica para una imagen (o múltiples imágenes que usan los PPS), así como otra información que identifica o define los espacios de color disponibles, frecuencias de muestreo de color disponibles y/o profundidades de bits disponibles. Un PPS puede usarse para una sola imagen, o un PPS puede usarse nuevamente para múltiples imágenes en una secuencia. Un PPS se señala comúnmente separado de los datos codificados para una imagen (por ejemplo, una unidad de NAL de un PPS y una o más unidades de NAL para los datos codificados para una imagen). Dentro de los datos codificados para una imagen, un elemento de sintaxis indica cual PPS usar para la foto. Del mismo modo, para la sintaxis de acuerdo con el estándar H.265/HEVC, un conjunto de parámetros de secuencia ("SPS", por sus siglas en inglés) es una estructura de sintaxis que contiene los elementos de sintaxis que pueden asociarse con una secuencia de imágenes. Una corriente de bits puede incluir un solo SPS o múltiples SPS. Un SPS se señala comúnmente separado de otros datos para la secuencia, y un elemento de sintaxis en los otros datos indica cual SPS usar. En algunas implementaciones ejemplares, un SPS para una secuencia puede incluir la información que identifica o define los espacios de color disponibles, las frecuencias de muestreo de color disponibles y/o las profundidades de bit disponibles, a los cuales se hace referencia al momento de la conmutación de los espacios de color, frecuencias de muestreo de color y/o

profundidades de bits en la secuencia.

Para la capa de sección, un encabezado de sección (por ejemplo, encabezado de segmento de sección) incluye los valores de los elementos de sintaxis que se aplican para una sección (por ejemplo, una sección independiente y cualquier sección dependiente
5 siguiente). En algunas implementaciones ejemplares, un encabezado de sección puede incluir una o más señales que indican un espacio de color, frecuencia de muestreo de color o profundidad de bits que se aplican para una sección. En algunas implementaciones
10 ejemplares, un encabezado de sección también puede incluir la información que identifica o define los espacios de color disponibles, las frecuencias de muestreo de color disponibles y/o las profundidades de bit disponibles, a los cuales se hace referencia al momento de la conmutación de los espacios de color, frecuencias de
15 muestreo de color y/o profundidades de bits dentro de la sección. Para la capa de bloque (por ejemplo, para una CTU), una estructura de sintaxis incluye los valores de los elementos de sintaxis que se aplican para un bloque. En algunas implementaciones ejemplares, la estructura de sintaxis para un bloque puede incluir una o más
20 señales que indican un espacio de color, frecuencia de muestreo de color y/o profundidad de bits que se aplican para el bloque.

Los fotogramas codificados (341) y la información de MMCO/RPS (342) (o información equivalente a la información de MMCO/RPS (342)), puesto que ya se conocen las dependencias y las
25 estructuras de ordenamiento de los fotogramas en el codificador

(340)) se procesan por un emulador de proceso de decodificación (350). El emulador de proceso de decodificación (350) implementa alguna funcionalidad de un decodificador, por ejemplo, las tareas de decodificación para reconstruir los fotogramas de referencia. De acuerdo con la información de MMCO/RPS (342), el emulador de procesos de decodificación (350) determina si un determinado fotograma codificado (341) deben reconstruirse y almacenarse para su uso como un fotograma de referencia en la predicción interfotograma de los fotogramas posteriores que se codificarán. Si un fotograma codificado (341) debe almacenarse, el emulador de proceso de decodificación (350) modela el proceso de decodificación que se llevaría a cabo por un decodificador que recibe el fotograma codificado (341) y produce un fotograma decodificado correspondiente (351). De esta manera, cuando el codificador (340) ha usado los fotogramas decodificados (369) que se han almacenado en el área de almacenamiento de fotogramas decodificados (360), el emulador de proceso de decodificación (350) también usa los fotogramas descodificados (369) del área de almacenamiento (360) como parte del proceso de decodificación.

El área de almacenamiento de memoria temporal de fotograma decodificado (360) incluye múltiples áreas de almacenamiento de memoria intermedia de fotograma (361, 362, ..., 36n). De acuerdo con la información de MMCO/RPS (342), el emulador de proceso de decodificación (350) gestiona el contenido del área de almacenamiento (360) con el fin de identificar cualquier memoria

intermedia de fotograma (361, 362, etcétera) donde los fotogramas ya no son necesarios por el codificador (340) para su uso como fotogramas de referencia. Después de modelar el proceso de decodificación, el emulador de proceso de decodificación (350) 5 almacena un fotograma recientemente decodificado (351) en una memoria intermedia de fotograma (361, 362, etcétera) que se ha identificado de esta manera.

Los fotogramas codificados (341) y la información de MMCO/RPS (342) se introducen en memoria intermedia en un área 10 temporal de datos codificados (370). Los datos codificados que se agregan en el área de datos codificados (370) contienen, como parte de la sintaxis de una corriente de bits elemental del video codificado, los datos codificados para una o más imágenes. Los datos codificados que se agregan en el área de datos codificados (370) 15 también pueden incluir los metadatos de medios que se relacionan con los datos de video codificados (por ejemplo, como uno o más parámetros en uno o más mensajes de información complementaria de mejora ("SEI", por sus siglas en inglés) o mensajes de información de facilidad de uso de video ("VUI", por sus siglas en 20 inglés)).

Los datos agregados (371) desde el área temporal de datos codificados (370) se procesan por un codificador de canal (380). El codificador de canal (380) puede agrupar y/o multiplexar los datos agregados para la transmisión o almacenamiento como una 25 secuencia de medios (por ejemplo, de acuerdo con un formato de

emisión en continuo de programa de medios o de emisión en continuo de transporte como ITU-T H.222.0 | ISO/IEC 13818-1 o un formato de protocolo de transporte en tiempo real en Internet como IETF RFC 3550), en cuyo caso el codificador de canal (380) puede

5 agregar elementos de sintaxis como parte de la sintaxis de la emisión en continuo de transmisión de medios. O, el codificador de canal (380) puede organizar los datos agregados para almacenarlos como un archivo (por ejemplo, de acuerdo con un formato de contenedor de medios como ISO/IEC 14496-12), en cuyo caso el

10 codificador de canal (380) puede agregar los elementos de sintaxis como parte de la sintaxis del archivo de almacenamiento de medios. O, más generalmente, el codificador de canal (380) puede implementar uno o más protocolos de multiplexación de sistema o protocolos de transporte, en cuyo caso el codificador de canal (380)

15 puede agregar los elementos de sintaxis como parte de la sintaxis de los protocolos. El codificador de canal (380) proporciona la salida a un canal (390), que representa el almacenamiento, una conexión de comunicaciones u otro canal para la salida. El codificador de canal (380) o canal (390) también puede incluir otros elementos (no

20 mostrados), por ejemplo, para la codificación de corrección anticipada de errores ("FEC", por sus siglas en inglés) y la modulación de señal analógica.

IV. Sistemas decodificadores ejemplares.

La figura 4 es un diagrama de bloques de un sistema

25 decodificador ejemplar (400) junto con el cual pueden implementarse

algunas modalidades descritas. El sistema decodificador (400) puede ser una herramienta de decodificación de uso general capaz de operar en cualquiera de los múltiples modos de decodificación como un modo de decodificación de baja latencia para la comunicación en tiempo real y un modo de decodificación de alta latencia para la reproducción de los medios a partir de un archivo o una emisión en continuo, o puede ser una herramienta de decodificación de uso especial adaptada para tal modo de decodificación. El sistema decodificador (400) puede adaptarse para la decodificación de un tipo particular de contenido (por ejemplo, contenido de captura de pantalla), o puede adaptarse para la decodificación de cualquiera de varios tipos diferentes de contenido (por ejemplo, contenido de captura de pantalla y video natural). El sistema codificador (400) puede implementarse como un módulo del sistema operativo, como parte de una biblioteca de aplicación o como una aplicación independiente. En general, el sistema decodificador (400) recibe los datos codificados desde un canal (410) y produce los fotogramas reconstruidos como salida para un destino de salida (490). Los datos codificados pueden incluir el contenido codificado con la conmutación adaptativa de los espacios de color, frecuencias de muestreo de color y/o profundidades de bits.

El sistema decodificador (400) incluye un canal (410), que puede representar un almacenamiento, una conexión de comunicaciones u otro canal para los datos codificados como entrada. El canal (410) produce los datos codificados que se han

codificado en el canal. Un decodificador de canal (420) puede procesar los datos codificados. Por ejemplo, el codificador de canal (420) desagrupa y/o desmultiplexa los datos que se han agregado para la transmisión o almacenamiento como una emisión en continuo de medios (por ejemplo, de acuerdo con un formato de emisión en continuo de programa de medios o de emisión en continuo de transporte como ITU-T H.222.0 | ISO/ISO/IEC 13818-1 o un formato de protocolo de transporte en tiempo real en Internet como IETF RFC 3550), en cuyo caso el codificador de canal (420) puede analizar los elementos de sintaxis como parte de la sintaxis de la emisión en continuo de transmisión de medios. O, el decodificador de canal (420) separa los datos de video codificados que se han agregado para el almacenamiento como un archivo (por ejemplo, de acuerdo con un formato de contenedor de medios como ISO/IEC 14496-12), en cuyo caso el decodificador de canal (420) puede analizar los elementos de sintaxis como parte de la sintaxis del archivo de almacenamiento de medios. O, más generalmente, el decodificador de canal (420) puede implementar uno o más protocolos de desmultiplexación de sistema de medios o protocolos de transporte, en cuyo caso el decodificador de canal (420) puede analizar los elementos de sintaxis como parte de la sintaxis de los protocolos. El canal (410) o el decodificador de canal (420) también puede incluir otros elementos (no mostrados), por ejemplo, para la decodificación de FEC y la modulación de señal analógica.

Los datos codificados (421) que se emiten desde el

decodificador de canal (420) se almacenan en un área temporal de datos codificados (430) hasta que se haya recibido una cantidad suficiente de tales datos. Los datos codificados (421) incluyen los fotogramas codificados (431) y la información de MMCO/RPS (432).

5 Los datos codificados (421) en el área de datos codificados (430) contienen, como parte de la sintaxis de una corriente de bits elemental del video codificado, los datos codificados para una o más imágenes. Los datos codificados (421) en el área de datos codificados (430) también pueden incluir los metadatos de medios

10 que se relacionan con los datos de video codificados (por ejemplo, como uno o más parámetros en uno o más mensajes de SEI o mensajes de VUI).

En general, el área de datos codificados (430) almacena temporalmente los datos codificados (421) hasta que tales datos

15 codificados (421) se usen mediante el decodificador (450). En ese momento, los datos codificados para un fotograma codificado (431) y la información de MMCO/RPS (432) se transfieren desde el área de datos codificados (430) al decodificador (450). A medida que continua la decodificación, nuevos datos codificados se agregan al

20 área de datos codificados (430) y los datos codificados más antiguos restantes en el área de datos codificados (430) se transfieren al decodificador (450).

El decodificador (450) decodifica un fotograma codificado (431) para producir un fotograma decodificado correspondiente (451).

25 Según corresponda, cuando se realiza su proceso de decodificación,

el decodificador (450) puede usar uno o más fotogramas previamente decodificados (469) como los fotogramas de referencia para la predicción interfotograma. El decodificador (450) lee tales fotogramas previamente decodificados (469) desde un área de almacenamiento de memoria temporal de fotograma decodificado (460).

Durante la decodificación, para diferentes imágenes, secciones, bloques u otras unidades de video, el decodificador (450) puede realizar las conversiones de espacio de color para la transformación entre un espacio de color tipo YUV y un espacio de color tipo RGB, o a/de algún otro espacio de color. El decodificador (450) también puede realizar las conversiones de espacio de color para reordenar los componentes de color para diferentes imágenes, secciones, bloques u otras unidades de video, cambiando cual componente de color es el componente primario (por ejemplo, convertir entre los formatos RGB, BGR y GBR). Durante la decodificación, el decodificador (450) también puede realizar el procesamiento de remuestreo para cambiar las frecuencias de muestreo de color y/o cambiar las profundidades de bits para diferentes imágenes, secciones, bloques u otras unidades de video. En algunas implementaciones ejemplares, el decodificador (450) puede conmutar espacios de color, frecuencias de muestreo de color y/o profundidades de bits en una base de imagen por imagen durante la decodificación. Por otra parte, el decodificador (450) puede conmutar espacios de color, frecuencias de muestreo de color y/o

profundidades de bits en una base de sección por sección durante la decodificación.

Generalmente, el codificador (450) incluye múltiples módulos de decodificación que realizan las tareas de decodificación como
5 decodificación entrópica, adaptación de espacio de color, frecuencia de muestreo de color y/o profundidad de bits, predicción intrafotograma, predicción interfotograma con compensación de movimiento, cuantificación invertida, transformación de frecuencia invertida y combinación de las piezas. Las operaciones exactas
10 realizadas por el decodificador (450) pueden variar dependiendo del formato de compresión.

Por ejemplo, el decodificador (450) recibe los datos codificados para un fotograma comprimido o secuencia de fotogramas y produce la salida que incluye el fotograma decodificado (451). En el
15 decodificador (450), una memoria intermedia recibe los datos codificados para un fotograma comprimido y, en el momento adecuado, facilita los datos codificados recibidos a un decodificador entrópico. La entropía del decodificador entrópico decodifica los datos cuantificados codificados de manera por entropía, así como la
20 información secundaria de codificación entrópica, normalmente aplicando lo contrario a la codificación entrópica en el codificador. Un compensador de movimiento aplica la información de movimiento a uno o más fotogramas de referencia para formar los valores de predicción con compensación de movimiento para cualquier bloque
25 intercodificado del fotograma que se reconstruye. Un módulo de

predicción intrafotograma puede predecir de manera espacial los valores de muestra de un bloque actual a partir de los valores de muestra próximos previamente reconstruidos o, para la predicción intra-BC, predecir los valores de muestra de un bloque actual usando

5 los valores de muestra previamente reconstruidos de una región de referencia de predicción intrafotograma en el fotograma. La región de referencia puede indicarse con un valor de BV. El decodificador (450) también reconstruye los valores residuales de predicción. Un

10 cuantificador invertido cuantifica de manera invertida los datos decodificados por entropía. Por ejemplo, el decodificador (450) establece los valores para el QP para una imagen, sección, pieza, sección y/u otra porción del video con base en los elementos de sintaxis en la corriente de bits, y por consiguiente cuantifica de manera invertida los coeficientes de transformación. Un

15 transformador de frecuencia invertida convierte los datos cuantificados de dominio de frecuencia, en datos de dominio espacial. Para un bloque previsto interfotograma, el decodificador (450) combina los valores residuales de predicción reconstruidos con los valores de predicción con compensación de movimiento. El

20 decodificador (450) del mismo modo puede combinar los valores residuales de predicción con los valores de predicción a partir de la intrapredicción. Un filtro de desbloqueo adaptativo se incluye dentro del bucle de compensación de movimiento en el codificador (450) para ajustar las discontinuidades a través de las filas y/o columnas

25 del límite de bloque en el fotograma decodificado (451). Otra

filtración (como, filtración de eliminación de distorsión, ALF, o filtración de SAO; no mostradas) puede alternativamente o además aplicarse como operaciones de filtración en el bucle.

El área de almacenamiento de memoria temporal de fotograma decodificado (460) incluye múltiples áreas de almacenamiento de memoria intermedia de fotograma (461, 462, ..., 46n). El área de almacenamiento de fotograma decodificado (460) es un ejemplo de una memoria intermedia de imagen descodificada. El decodificador (450) usa la información de MMCO/RPS (432) para identificar una memoria intermedia de fotograma (461, 462, etcétera) en la cual puede almacenar un fotograma decodificado (451). El decodificador (450) almacena el fotograma decodificado (451) en esa memoria intermedia de fotograma.

Un secuenciador de salida (480) identifica cuando el siguiente fotograma, que se producirá en orden de salida, está disponible en el área de almacenamiento de fotograma decodificado (460). Cuando el siguiente fotograma (481), que se producirá en el orden de salida, está disponible en el área de almacenamiento de fotograma decodificado (460), se lee por el secuenciador de salida (480) y se emite al destino de salida (490) (por ejemplo, se despliega). En general, el orden en el cual se emiten los fotogramas desde el área de almacenamiento de fotograma decodificado (460) por el secuenciador de salida (480), puede diferir del orden en el cual los fotogramas se decodifican por el decodificador (450).

V. Codificadores de video ejemplares.

Las figuras 5a y 5b son diagramas de bloques que ilustran un codificador de video generalizado (500) junto con el cual pueden implementarse algunas modalidades descritas. El codificador (500) recibe una secuencia de imágenes de video que incluye una imagen actual como una señal de entrada de video (505) y produce los datos codificados en una corriente de bits de video codificado (595) como salida.

El codificador (500) se basa en bloques y usa un formato de bloque que depende de la implementación. Los bloques pueden subdividirse en diferentes etapas, por ejemplo, en etapas de predicción, transformación de frecuencia y/o codificación por entropía. Por ejemplo, una imagen puede dividirse en bloques de 64 x 64, bloques de 32 x 32 o bloques de 16 x 16, que a su vez pueden dividirse en bloques más pequeños de los valores de muestra para la codificación y decodificación. En las implementaciones de codificación para el estándar H.265/HEVC, el codificador segmenta una imagen en CTU (CTB), CU (CB), PU (PB) y TU (TB).

El codificador (500) comprime las imágenes usando la codificación intraimagen y/o la codificación interimagen. Muchos de los componentes del codificador (500) se usan para la codificación intraimagen y la codificación interimagen. Las operaciones exactas realizadas por esos componentes pueden variar dependiendo del tipo de información que se comprime.

Un módulo de división en piezas (510) segmenta opcionalmente una imagen en múltiples piezas del mismo tamaño o de diferentes

tamaños. Por ejemplo, el módulo de división en piezas (510) divide la imagen a lo largo de filas de piezas y columnas de piezas que, con límites de imagen, definen los límites horizontales y verticales de las piezas en la imagen, donde cada pieza es una región rectangular. En las implementaciones de H.265/HEVC, el codificador (500) fragmenta una imagen en una o más secciones, donde cada sección incluye uno o más segmentos de sección.

El control de codificación general (520) recibe las imágenes para la señal video de entrada (505), así como la retroalimentación (no mostrada) desde varios módulos del codificador (500). En general, el control de codificación general (520) proporciona las señales de control (no mostradas) a otros módulos (como módulo de división en piezas (510), transformador/escalador/cuantificador (530), escalador/transformador invertido (535), estimador intraimagen (540), estimador de movimiento (550) e inter/intraconmutador) para establecer y cambiar los parámetros de codificación durante la codificación. En particular, el control de codificación general (520) puede decidir cómo conmutar de manera adaptativa los espacios de color, las frecuencias de muestreo de color y/o las profundidades de bits durante la codificación en imagen por imagen, sección por sección, bloque por bloque o algún otro modo. El control de codificación general (520) también puede evaluar los resultados intermedios durante la codificación. El control de codificación general (520) produce los datos de control general (522) que indican las decisiones tomadas durante la codificación, para que

un decodificador correspondiente pueda tomar las decisiones coherentes. Los datos de control general (522) se proporcionan al formador de encabezado/codificador entrópico (590).

Si la imagen actual se predice usando la predicción
5 interimagen, un estimador de movimiento (550) estima el movimiento de los bloques de valores de muestra de la imagen actual de la señal video de entrada (505) con respecto a una o más imágenes de referencia. La memoria intermedia de imagen decodificada (570) introduce en memoria intermedia una o más imágenes previamente
10 codificadas reconstruidas para uso como imágenes de referencia. Cuando se usan múltiples imágenes de referencia, las múltiples imágenes de referencia pueden ser de diferentes direcciones temporales o de la misma dirección temporal. El estimador de movimiento (550) produce como información secundaria los datos de
15 movimiento (552) como datos de MV, valores de índice de modo de combinación y datos de selección de imagen de referencia. Los datos de movimiento (522) se proporcionan al formador de encabezado/codificador entrópico (590), así como al compensador de movimiento (555).

20 El compensador de movimiento (555) aplica los MV a las imágenes de referencia reconstruidas desde la memoria intermedia de imagen decodificada (570). El compensador de movimiento (555) produce las predicciones con compensación de movimiento para la imagen actual. Cuando los componentes secundarios para una
25 imagen tienen la misma resolución que el componente primario (por

ejemplo, cuando el formato es formato YUV a 4:4:4 o formato RGB a 4:4:4), el valor de MV que se aplica para un bloque de componente secundario puede ser el mismo que el valor de MV aplicado para el bloque de componente primario correspondiente. Por otra parte, cuando los componentes secundarios para una imagen han reducido la resolución en relación con el componente primario (por ejemplo, cuando el formato de YUV a 4:2:0), el valor de MV que se aplica para un bloque de componente secundario puede reducirse y posiblemente redondearse para ajustar la diferencia de resolución (por ejemplo, al dividir los componentes verticales y horizontales del valor de MV en dos y al truncarlos o redondearlos a valores enteros).

En un curso separado en el codificador (500), un estimador intraimagen (540) determina cómo realizar la predicción intraimagen para los bloques de valores de muestra de una imagen actual de la señal de video de entrada (505). La imagen actual puede ser completa o parcialmente codificada usando la codificación intraimagen. Usando los valores de una reconstrucción (538) de la imagen actual, para la predicción intraespacial, el estimador intraimagen (540) determina cómo predecir de manera espacial los valores de muestra de un bloque actual de la imagen actual a partir de los valores de muestra próximos previamente reconstruidos de la imagen actual. O, para predicción intra-BC usando los valores de BV, el estimador intraimagen (540) estima el desplazamiento de los valores de muestra del bloque actual a las diferentes regiones de referencia posibles en la imagen actual. En la figura 5b, las regiones

de referencia posibles incluyen los valores de muestra reconstruidos. Por otra parte, para fines de estimación de BV, las regiones de referencia posibles pueden incluir los valores de muestra de entrada.

El estimador intraimagen (540) produce, como información
5 secundaria, los datos intrapredicción (542), como la información que indica si la intrapredicción usa la predicción espacial o la predicción intra-BC (por ejemplo, un valor indicador por intrabloque), dirección de modo de predicción (para la predicción intraespacial) y los valores de BV (para la predicción intra-BC). Los datos intrapredicción (522)
10 se proporcionan al formador de encabezado/codificador entrópico (590), así como al predictor intraimagen (545).

De acuerdo con los datos intrapredicción (542), el predictor intraimagen (545) predice de manera espacial los valores de muestra de un bloque actual de la imagen actual a partir de los valores de
15 muestra próximos previamente reconstruidos de la imagen actual. O, para la predicción intra-BC, el predictor intraimagen (545) predice los valores de muestra del bloque actual usando los valores de muestra previamente reconstruidos de una región de referencia predicción intraimagen, que se indica por un valor de BV para el bloque actual.
20 En algunos casos, el valor de BV puede ser un predictor de BV (valor previsto de BV). En otros casos, el valor de BV puede ser diferente de su valor de BV previsto, en cuyo caso una BV diferencial indica la diferencia entre el valor de BV previsto y el valor de BV. Cuando los componentes secundarios para una imagen tienen la misma
25 resolución que el componente primario (por ejemplo, cuando el

formato es formato YUV a 4:4:4 o formato RGB a 4:4:4), el valor de BV que se aplica para un bloque de componente secundario puede ser el mismo que el valor de BV aplicado para el bloque de componente primario correspondiente. Por otra parte, cuando los

5 componentes secundarios para una imagen han reducido la resolución en relación con el componente primario (por ejemplo, cuanto el formato YUV a 4:2:0), el valor de BV que se aplica para un bloque de componente secundario puede reducirse y posiblemente redondearse para ajustar la diferencia de resolución (por ejemplo, al

10 dividir los componentes verticales y horizontales del valor de BV en dos y al truncarlos o redondearlos a valores enteros).

El intra/interconmutador selecciona si la predicción (558) para un determinado bloque será una predicción con compensación de movimiento o una predicción intraimagen. La diferencia (si existe)

15 entre un bloque de la predicción (558) y una parte correspondiente de la imagen actual original de la señal video de entrada (505) proporciona los valores de la magnitud residual (518), para un bloque de modo de no omisión. Durante la reconstrucción de la imagen actual, para un bloque de modo de no omisión, los valores

20 residuales reconstruidos se combinan con la predicción (558) para producir una reconstrucción aproximada o exacta (538) del contenido original a partir de la señal de video (505). (En la compresión con pérdida, alguna información se pierde en la señal de video (505).)

En el transformador/escalador/cuantificador (530), un

25 transformador de frecuencia convierte la información de video de

dominio espacial en los datos de dominio de frecuencia (es decir, espectrales, de transformación). Para la codificación de video con base en bloques, el transformador de frecuencia aplica una transformada de coseno discreta "(DCT", por sus siglas en inglés), una aproximación en número entero de la misma u otro tipo de transformación directa de bloque (por ejemplo, una transformada de seno discreta o una aproximación en número entero de la misma) a los bloques de los datos residuales de predicción (o datos de valor de muestra si la predicción (558) es nula), produciendo bloques de coeficientes de transformación de frecuencia. El codificador (500) también puede indicar que tal etapa de transformación se omitirá. El escalador/cuantificador escala y cuantifica los coeficientes de transformación. Por ejemplo, el cuantificador aplica la cuantificación de escala de zonas inactivas a los datos de dominio de frecuencia con un tamaño de etapa de cuantificación que varía en una base imagen por imagen, base de pieza por pieza, base de sección por sección, base de bloque por bloque, base de frecuencia específica u otra base. Los datos cuantificados de coeficiente de transformación (532) se proporcionan al formador de encabezado/codificador entrópico (590).

En el escalador/transformador invertido (535), un escalador/cuantificador invertido realiza el escalamiento invertido y la cuantificación invertida en los coeficientes de transformación cuantificados. Un transformador de frecuencia invertida realiza una transformación de frecuencia invertida, produciendo bloques de

valores residuales de predicción reconstruidos o valores de muestra. Para el bloque de modo de no omisión, el codificador (500) combina los valores residuales reconstruidos con valores de la predicción (558) (por ejemplo, valores de predicción con compensación de movimiento, valores de predicción intraimagen) para formar la
5 reconstrucción (538). Para un bloque de modo de omisión, el codificador (500) usa los valores de la predicción (558) como la reconstrucción (538).

Para la predicción intraimagen, los valores de la reconstrucción
10 (538) pueden retroalimentarse al estimador intraimagen (540) y al predictor intraimagen (545). Además, los valores de la reconstrucción (538) pueden usarse para la predicción con compensación de movimiento de las imágenes posteriores. Los valores de la reconstrucción (538) pueden filtrarse aún más. Un
15 control de filtración (560) determina cómo realizar la filtración de desbloqueo y la filtración de SAO en los valores de la reconstrucción (538), para una determinada imagen de la señal de video (505). El control de filtración (560) produce los datos de control de filtro (562), que se proporcionan al formador de encabezado/codificador
20 entrópico (590) y combinación/ filtro (565).

En la combinación/filtro (565), el codificador (500) combina el contenido de diferentes piezas en una versión reconstruida de la imagen. El codificador (500) realiza selectivamente la filtración de desbloqueo y la filtración de SAO de acuerdo con los datos de
25 control de filtro (562), para ajustar de manera adaptativa las

discontinuidades a través de los límites en las imágenes. Otra filtración (como la filtración de eliminación de distorsión o ALF; no mostrada) puede aplicarse de manera alternativa o adicional. Los límites de la pieza pueden filtrarse selectivamente o no filtrarse, dependiendo de la configuración del codificador (500), y el codificador (500) puede proporcionar la sintaxis en la corriente de bits codificada para indicar si o no se aplica tal filtración. La memoria intermedia de imagen decodificada (570) introduce en memoria intermedia la imagen actual reconstruida para uso en la predicción con compensación de movimiento subsecuente.

El formateador de encabezado/codificador entrópico (590) formatean y/o codifican por entropía los datos de control generales (522), cuantifican los datos de coeficiente de transformación (532), los datos de intrapredicción (542), los datos de movimiento (552) y los datos de control de filtro (562). Los datos de control generales (522) incluyen las señales que indican los espacios de color, las frecuencias de muestreo de color y/o las profundidades de bits para imágenes, secciones, bloques u otras unidades de video. Estas señales pueden incluirse, por ejemplo, en un PPS, encabezado de sección, estructura de sintaxis de bloque u otra estructura de sintaxis, y puede codificarse por entropía o señalizarse como valores de longitud fija. El formateador de encabezado/codificador entrópico (590) puede también formatear y/o codificar por entropía la información que identifica o define los espacios de color disponibles (por ejemplo, lista de espacios de colores predefinidos, matrices

modificadas para operaciones de transformación de espacio de color), la información que identifica o define las frecuencias de muestreo de color disponibles (por ejemplo, lista frecuencias de muestreo de color predefinidas, otras frecuencias de muestreo de color o identificación de operaciones de reducción de muestra e incremento de muestra) y/o la información que identifica o define las profundidades de bits disponibles (por ejemplo, lista de profundidades de bits predefinidas, otras profundidades de bits o identificación de las operaciones de conversión de profundidad de bits), que se usan por un decodificador durante la conmutación adaptativa.

El formador de encabezado/codificador entrópico (590) proporciona los datos codificados en la corriente de bits de video codificado (595). El formato de la corriente de bits de video codificado (595) puede ser una variación o extensión del formato H.265/HEVC, formato de Windows Media Video, formato VC-1, formato MPEG-x (por ejemplo, MPEG-1, MPEG-2 o MPEG-4), formato H.26x (por ejemplo, H.261, H.262, H.263, H.264), u otro formato.

Dependiendo de la implementación y del tipo de compresión deseada, los módulos de un codificador (500) pueden agregarse, omitirse, dividirse en múltiples módulos, combinarse con otros módulos y/o sustituirse con módulos similares. En las modalidades alternativas, los codificadores con diferentes módulos y/u otras configuraciones de módulos realizan una o más de las técnicas descritas. Las modalidades específicas de los codificadores por lo

general usan una variación o versión complementada del codificador (500). Las relaciones mostradas entre los módulos en el codificador (500) indican el flujo general de información en el codificador; otras relaciones no se muestran en aras de simplicidad.

5 VI. Decodificadores de video ejemplares.

Las figuras 5a y 5b son diagramas de bloques que ilustran un decodificador de video generalizado (600) junto con el cual pueden implementarse algunas modalidades descritas. El decodificador (600) recibe los datos codificados en una corriente de bits de video
10 codificado (605) y produce la salida que incluye las imágenes para el video reconstruido (695). El formato de la corriente de bits de video codificado (605) puede ser una variación o extensión del formato H.265/HEVC, formato de Windows Media Video, formato VC-1, formato MPEG-x (por ejemplo, MPEG-1, MPEG-2 o MPEG-4), formato
15 H.26x (por ejemplo, H.261, H.262, H.263, H.264), u otro formato.

El decodificador (600) se basa en bloques y usa un formato de bloque que depende de la implementación. Los bloques pueden subdividirse en diferentes etapas. Por ejemplo, una imagen puede dividirse en bloques de 64 x 64, bloques de 32 x 32 o bloques de 16
20 x 16, que a su vez pueden dividirse en bloques más pequeños de los valores de muestra. En las implementaciones de decodificación para el estándar H.265/HEVC, una imagen se segmenta en CTU (CTB), CU (CB), PU (PB) y TU (TB).

El decodificador (600) descomprime las imágenes usando la
25 decodificación intraimagen y/o la decodificación interimagen. Muchos

de los componentes del decodificador (600) se usan para la decodificación intraimagen y la decodificación intraimagen. Las operaciones exactas realizadas por esos componentes pueden variar dependiendo del tipo de información que se descomprime.

5 Una memoria intermedia recibe los datos codificados en la corriente de bits de video codificado (605) y facilita los datos codificados recibidos al analizador/decodificador entrópico (610). El analizador/decodificador entrópico (610) decodifica por entropía los datos codificados por entropía, normalmente aplicando lo contrario a
10 la codificación entrópica realizada en el codificador (500) (por ejemplo, decodificación aritmética binario de adaptación al contexto). Como resultado del análisis y decodificación entrópica, el analizador/decodificador entrópico (610) produce los datos de control general (622), los datos de coeficiente de transformación
15 cuantificados (632), los datos intrapredicción (642), los datos de movimiento (652) y los datos de control de filtro (662). Los datos de control generales (622) incluyen las señales que indican los espacios de color, las frecuencias de muestreo de color y/o las profundidades de bits para imágenes, secciones, bloques u otras unidades de
20 video. Estas señales pueden incluirse, por ejemplo, en un PPS, encabezado de sección, estructura de sintaxis de bloque u otra estructura de sintaxis, y puede codificarse por entropía o señalizarse como valores de longitud fija. El analizador/decodificador entrópico (610) puede también analizar y/o decodificar por entropía la
25 información que identifica o define los espacios de color disponibles

(por ejemplo, lista de espacios de colores predefinidos, matrices modificadas para operaciones de transformación de espacio de color), la información que identifica o define las frecuencias de muestreo de color disponibles (por ejemplo, lista frecuencias de muestreo de color predefinidas, otras frecuencias de muestreo de color o identificación de operaciones para reducción de muestra e incremento de muestra) y/o la información que identifica o define las profundidades de bits disponibles (por ejemplo, lista de profundidades de bits predefinidas, otras profundidades de bits o identificación de las operaciones para conversión de profundidad de bits), que pueden usarse por el decodificador (600) durante la conmutación adaptativa.

El control de decodificación general (620) recibe los datos de control generales (622) y proporciona las señales de control (no mostradas) a otros módulos (como el escalador/transformador invertido (635), predictor intraimagen (645), compensador de movimiento (655) e intra/interconmutador) para establecer y cambiar los parámetros de decodificación durante la decodificación. En particular, el control de decodificación general (620) puede hacer que el decodificador (600) conmute los espacios de color, las frecuencias de muestreo de color y/o las profundidades de bits durante la decodificación en una base de imagen por imagen, base de sección por sección, base de bloque por bloque o alguna otra base.

Si la imagen actual se predice usando la predicción

interimagen, un compensador de movimiento (655) recibe los datos de movimiento (652), como datos de MV, datos de selección de imagen de referencia y valores de índice de modo de combinación. El compensador de movimiento (655) aplica los MV a las imágenes de referencia reconstruidas desde la memoria intermedia de imagen decodificada (670). El compensador de movimiento (655) produce las predicciones con compensación de movimiento para los bloques intercodificados de la imagen actual. La memoria intermedia de imagen decodificada (670) almacena una o más imágenes previamente reconstruidas para uso como imágenes de referencia.

En un curso separado en el decodificador (600), el predictor de predicción intraimagen (645) recibe los datos intrapredicción (642), como la información que indica si la intrapredicción usa la predicción espacial o la predicción intra-BC (por ejemplo, un valor indicador por intrabloque), dirección de modo de predicción (para la predicción intraespacial) y los valores de BV (para la predicción intra-BC). Para la predicción intraespacial, usando los valores de una reconstrucción (638) de la imagen actual, de acuerdo con los datos de modo de predicción, el predictor intraimagen (645) predice de manera espacial los valores de muestra de un bloque actual de la imagen actual a partir de los valores de muestra próximos previamente reconstruidos de la imagen actual. O, para la predicción intra-BC usando los valores de BV, el predictor intraimagen (645) predice los valores de muestra del bloque actual usando los valores de muestra previamente reconstruidos de una región de referencia de predicción

intraimagen, que se indica por un valor de BV para el bloque actual.

El intra/interconmutador selecciona si la predicción (658) para un determinado bloque es una predicción con compensación de movimiento o una predicción intraimagen. Por ejemplo, cuando se aplica la sintaxis de H.265/HEVC, el intra/interconmutador puede controlarse con base en un elemento de sintaxis codificado para una CU de una imagen que puede contener las CU intraprevistas y las CU interprevistas. Para un bloque de modo de no omisión, el decodificador (600) combina la predicción (658) con los valores residuales reconstruidos para producir la reconstrucción (638) del contenido de la señal de video. Para un bloque de modo de omisión, el decodificador (600) usa los valores de la predicción (658) como la reconstrucción (638).

Para reconstruir la magnitud residual para un bloque de modo de no omisión, el escalador/transformador invertido (635) recibe y procesa los datos cuantificados de coeficiente de transformación (632). En el escalador/transformador invertido (635), un escalador/cuantificador invertido realiza el escalamiento invertido y la cuantificación invertida en los coeficientes de transformación cuantificados. Un transformador de frecuencia invertida realiza una transformación de frecuencia invertida, produciendo bloques de valores residuales de predicción reconstruidos o valores de muestra. Por ejemplo, el transformador de frecuencia invertida aplica una transformación de bloque invertida a los coeficientes de transformación de frecuencia, produciendo los datos de valor de

muestra o los datos residuales de predicción. La transformación de frecuencia invertida puede ser una DCT invertida, una aproximación en número entero de la misma u otro tipo de transformación de frecuencia invertida (por ejemplo, una transformada de seno discreta invertida o una aproximación en número entero de la misma).

Para la predicción intraimagen, los valores de la reconstrucción (638) pueden retroalimentarse al predictor intraimagen (645). Para la predicción intraimagen, los valores de la reconstrucción (638) pueden filtrarse aún más. En la combinación/filtro (665), el decodificador (600) combina el contenido de diferentes piezas en una versión reconstruida de la imagen. El decodificador (600) realiza selectivamente la filtración de desbloqueo y la filtración de SAO de acuerdo con los datos de control de filtro (662) y las reglas para la adaptación de filtro, para ajustar de manera adaptativa las discontinuidades a través de los límites en las imágenes. Otra filtración (como la filtración de eliminación de distorsión o ALF; no mostrada) puede aplicarse de manera alternativa o adicional. Los límites de pieza pueden filtrarse selectivamente o no filtrarse en absoluto, dependiendo de la configuración del decodificador (600) o de una indicación de sintaxis en los datos de corriente de bits codificados. La memoria intermedia de imagen decodificada (670) introduce en memoria intermedia la imagen actual reconstruida para uso en la predicción con compensación de movimiento subsecuente.

El decodificador (600) también puede incluir un filtro de posprocesamiento. El filtro de posprocesamiento (608) puede incluir

la filtración de eliminación de distorsión, la filtración Wiener adaptativa, la filtración de reproducción del crono de la película, la filtración de SAO u otro tipo de filtración.

Dependiendo de la implementación y del tipo de descompresión deseada, los módulos de un decodificador (600) pueden agregarse, omitirse, dividirse en múltiples módulos, combinarse con otros módulos y/o sustituirse con módulos similares. En las modalidades alternativas, los decodificadores con diferentes módulos y/u otras configuraciones de módulos realizan una o más de las técnicas descritas. Las modalidades específicas de los decodificadores por lo general usan una variación o versión complementada del decodificador (600). Las relaciones mostradas entre los módulos en el decodificador (600) indican el flujo general de información en el decodificador; otras relaciones no se muestran en aras de simplicidad.

VII. Conmutación adaptativa de espacios de color, frecuencias de muestreo de color y/o profundidades de bits.

Durante la codificación, un codificador puede tratar un componente de color primario más cuidadosamente que los componentes de color secundario en varios aspectos (por ejemplo, más opciones para los modos de codificación, potencialmente menor tamaño de etapa de cuantificación). También, mediante el submuestreo de los componentes de color secundario, un codificador puede asignar relativamente más bits (y, por lo tanto, calidad) al componente de color primario.

Los valores de muestra de diferentes imágenes, secciones o bloques de una secuencia de video pueden tener diferentes características estadísticas. Estas características estadísticas diferentes pueden afectar la eficiencia de la codificación, que puede medirse en términos de rendimiento de la tasa de distorsión (menor o mayor calidad para una determinada velocidad binaria; o menor o mayor velocidad binaria para una determinada calidad).

Esta sección presenta varias características de codificación y decodificación adaptativa para las unidades de una secuencia de video. Una unidad puede ser una imagen, sección, bloque u otra porción de la secuencia de video. Por ejemplo, algunas de las características se refieren a la codificación-decodificación que incluye la conmutación adaptativa de los espacios de color entre las unidades en una secuencia de video. Otras características se refieren a la codificación-decodificación que incluye la conmutación adaptativa de las frecuencias de muestreo de color entre las unidades en una secuencia de video. Aún otras características se refieren a la codificación-decodificación que incluye la conmutación adaptativa de las profundidades de bits entre las unidades en una secuencia de video. Estas características pueden mejorar la eficiencia de codificación en muchos contextos y pueden usarse en combinación o por separado.

En particular, la conmutación adaptativa de espacios de color, frecuencias de muestreo de color y/o profundidades de bits puede mejorar el rendimiento de la tasa de distorsión al codificar ciertos

contenidos de video creados "artificialmente" como el contenido de captura de pantalla, o cuando una se codifica una mezcla de contenido de video natural y de video creado artificialmente. Los contextos comunes para la codificación-decodificación del contenido de pantalla de captura incluyen las conferencias en escritorio remoto y la codificación-decodificación de las superposiciones gráficas en video natural u otro video de "contenido mixto".

A. Tipos de video.

En general, el video de captura de pantalla representa la salida de una pantalla de computadora u otro monitor. La figura 7 muestra un entorno de escritorio de computadora (710) con el contenido que puede proporcionar la salida para la captura de pantalla. Por ejemplo, el video de captura de pantalla puede representar una serie de imágenes de todo el escritorio de computadora (711). O, el video de captura de pantalla puede representar una serie de imágenes para una de las ventanas del entorno de escritorio de computadora, como la ventana de aplicación (713) que incluye el contenido de juego, la ventana del navegador (712) con el contenido de la página Web o la ventana (714) con el contenido del procesador de texto.

Al igual que el contenido de video creado artificialmente generado por computadora, el contenido de captura de pantalla tiende a tener relativamente pocos valores de muestra discretos, en comparación con el contenido de video natural que se captura usando una cámara de video. Por ejemplo, una región de contenido de captura de pantalla a menudo incluye un solo color uniforme,

mientras una región de contenido de video natural incluye más probablemente colores que varían gradualmente. El contenido de captura de pantalla incluye también, por lo general estructuras distintas (por ejemplo, gráficos, caracteres de texto) que se repiten exactamente de fotograma a fotograma, aunque el contenido puede moverse espacialmente (por ejemplo, debido al desplazamiento). El contenido de captura de pantalla a menudo se codifica en un formato (por ejemplo, YUV a 4:4:4 o RGB a 4:4:4) con mayor resolución de muestreo de croma, aunque también puede codificarse en un formato con menor resolución de muestreo de croma (por ejemplo, YUV a 4:2:0, YUV a 4:2:2).

La figura 8 muestra el video compuesto (820) que incluye el video natural (821) y el contenido de video creado artificialmente. El contenido de video creado artificialmente incluye un gráfico (822) al lado del video natural (821) y una cinta de información (823) que se mueve debajo del video natural (821). Al igual que el contenido de captura de pantalla mostrado en la figura 7, el contenido de video creado artificialmente, mostrado en la figura 8 tiende a tener relativamente pocos valores de muestra discretos. También suele tener distintas estructuras (por ejemplo, gráficos, caracteres de texto) que se repiten exactamente de fotograma a fotograma (por ejemplo, debido al desplazamiento).

B. Ejemplos de conmutación adaptativa.

Para la conmutación adaptativa de espacios de color, diferentes unidades de video en una secuencia se codifican en

diferentes espacios de color. Por ejemplo, algunas de las unidades se codifican en un espacio de color tipo YUV (por ejemplo, YCbCr, YCoCg), y otras unidades se codifican en un espacio de color tipo RGB (por ejemplo, RGB, BGR, GBR). En este caso, un codificador o decodificador, según corresponda, realiza las operaciones de transformación de espacio de color en los valores de muestra para conmutar entre un espacio de color tipo YUV y un espacio de color tipo RGB. Por lo general, las operaciones de transformación de espacios de color pueden representarse como operaciones de multiplicación de matriz realizadas en una base de posición por posición, donde los valores de muestra n en un primer espacio de color para una posición, se multiplicarán por una matriz $n \times n$ para producir los valores de muestra n en un segundo espacio de color para la posición. En la práctica, las operaciones de transformación de espacios de color pueden implementarse mediante otra aritmética.

Como otro ejemplo de conmutación adaptativa de espacios de color, diferentes unidades pueden codificarse en diferentes espacios de color tipo RGB, que difieren en cuanto a su componente primario y el orden en el cual están señalizados los componentes (por ejemplo, para los datos residuales). En este caso, un codificador o decodificador, según sea el caso, realiza las operaciones de reordenamiento de espacio de color en los bloques o planos de los valores de muestra para cambiar cual componente de color es el componente de color primario.

En algunas implementaciones ejemplares, para codificación

con pérdida, un codificador puede conmutar entre cualquiera de los diferentes espacios de color. Para la codificación sin pérdida, sin embargo, un codificador solo realiza las conversiones de espacio de color invertibles (por ejemplo, reordenamiento de los componentes de color entre espacios de color RGB, BGR y GBR, o, en algunas implementaciones, conversión a/de YCoCg usando una mayor profundidad de bits intermedia).

Para la conmutación adaptativa de frecuencias de muestreo de color, diferentes unidades de video en una secuencia se codifican con diferentes frecuencias de muestreo de color. Por ejemplo, algunas de las unidades se codifican en un formato 4:2:2 o formato 4:2:0 (como YUV a 4:2:2 e YUV a 4:2:0), mientras otras unidades se codifican en un formato 4:4:4 (como YUV a 4:4:4). Un espacio de color tipo RGB generalmente tiene una frecuencia de muestreo de color de 4:4:4, pero sus componentes de color en cambio pueden submuestrearse de acuerdo con una frecuencia de muestreo de color de 4:2:2 o 4:2:0.

Cuando se realiza la reducción de muestra de los valores de muestra para un componente secundario de manera horizontal o vertical, un codificador o decodificador puede llevar a cabo el submuestreo simple, filtración de paso bajo más submuestreo u otra filtración más submuestreo. Para la reducción de muestra correspondiente de los valores de muestra del componente secundario, un codificador o decodificador invierte el submuestreo usando, por ejemplo, la repetición y/o filtración del valor de muestra.

Para la conmutación adaptativa de las profundidades de bits, diferentes unidades de video en una secuencia se codifican con diferentes profundidades de bits. Por ejemplo, algunas de las unidades se codifican con valores de muestra de 12 bits, mientras
5 otras unidades se codifican con valores de muestra de 10 bits o valores de muestra de 8 bits. Al convertir entre profundidades de bits, un codificador o decodificador puede truncar los valores de muestra de mayor profundidad de bits (con o sin adición de un factor de redondeo) para producir valores de muestra de menor profundidad
10 de bits o graduar los valores de muestra de menor profundidad de bits para producir los valores de muestra de mayor profundidad de bits.

Un codificador normalmente recibe el video de entrada en un determinado formato como RGB a 4:4:4 con valores de muestra de
15 12 bits. El codificador puede convertir entre el formato de entrada y otro formato (con un diferente espacio de color, diferente frecuencia de muestreo de color y/o diferente profundidad de bits) para una determinada unidad del video. La adaptación de espacio de color puede realizarse a nivel de imagen, nivel de sección, nivel de bloque
20 o algún otro nivel, en sí o en combinación con la adaptación de frecuencia de muestreo de color y/o la adaptación de profundidad de bits. Del mismo modo, la adaptación de frecuencia de muestreo de color puede realizarse a nivel de imagen, nivel de sección, nivel de bloque o algún otro nivel, en sí o en combinación con la adaptación
25 de espacio de color y/o la adaptación de profundidad de bit. La

adaptación de profundidad de bits puede realizarse a nivel de imagen, nivel de sección, nivel de bloque o algún otro nivel, en sí o en combinación con la adaptación de espacio de color y/o la adaptación de frecuencia de muestreo de color.

5 La figura 9 muestra los espacios de color, las frecuencias de muestreo de color y/o las profundidades de bits de adaptación a la imagen para las imágenes en una secuencia (900). La secuencia (900) incluye una serie de imágenes. Según sea necesario, el codificador convierte el video de entrada del formato de video de
10 entrada al formato seleccionado para una determinada imagen. El formato de imagen 1 es YCbCr a 4:2:0 con valores de muestra de 8 bits y el formato de imagen 2 es RGB a 4:4:4 con valores de muestra de 8 bits. Las imágenes 3 y 4 son video en BGR a 4:4:4, pero sus valores de muestra tienen diferentes profundidades de bits.

15 La figura 10 muestra los espacios de color, las frecuencias de muestreo de color y/o las profundidades de bits de adaptación a la sección para las secciones de una imagen (1000) en una secuencia. La imagen (1000) incluye cinco secciones, cuyos límites se muestran en líneas discontinuas. Por ejemplo, las secciones 0, 1, 3 y 4 podrían
20 asociarse con el contenido de captura de pantalla u otro contenido de video creado artificialmente, mientras la sección 2 se asocia con el contenido de video natural. El formato de las secciones 0 y 3 es BGR a 4:4:4 con valores de muestra de 10 bits. La sección 1 incluye los valores de muestra de 12 bits de video en GBR a 4:4:4. El
25 formato de la sección 2 es YCbCr a 4:2:0 con valores de muestra de

8 bits y el formato de sección 4 es RGB a 4:4:4 con valores de muestra de 8 bits.

La figura 11 muestra los espacios de color, las frecuencias de muestreo de color y/o las profundidades de bits de adaptación al bloque para los bloques de una sección de una sección (1100) de una imagen en una secuencia. La sección (1100) incluye 16 bloques que tienen tres diferentes tamaños de bloque. El formato de los dos primeros bloques es YCbCr a 4:2:0 con valores de muestra de 8 bits y el formato del último bloque es YCbCr a 4:2:0 con valores de muestra de 10 bits. Los bloques 2-15 incluyen los valores de muestra de 12 bits en un espacio de color tipo RGB con una frecuencia de muestreo de 4:4:4. El espacio de color para los bloques 2-15 varía entre RGB, BGR y GBR.

C. Ejemplos de la conmutación adaptativa durante la codificación.

La figura 12 muestra una técnica generalizada (1200) para la conmutación adaptativa de espacios de color, frecuencias de muestreo de color y/o profundidades de bits durante la codificación. Un codificador de imagen o codificador de video tal como se describe con referencia a la figura 3 o figuras 5a-5b puede realizar la técnica (1200).

El codificador codifica (1210) el video en una secuencia. Como parte de la codificación, el codificador conmuta espacios de color, frecuencias de muestreo de color y/o profundidades de bits entre al menos algunas unidades del video en la secuencia durante la

codificación. El codificador puede conmutar espacios de color, frecuencias de muestreo de color y/o profundidades de bits de manera espacial en una imagen. El codificador puede conmutar espacios de color, frecuencias de muestreo de color y/o profundidades de bits de imagen a imagen. Para propósitos de conmutación adaptativa, las unidades del video pueden ser imágenes de la secuencia, sección de una determinada imagen de la secuencia, bloques de una determinada imagen de la secuencia, u otras porciones de la secuencia de video.

10 Por ejemplo, la conmutación puede incluir cambiar entre dos de los espacios de color usando las operaciones de transformación de espacio de color. Para las operaciones de transformación de espacio de color, los espacios de color pueden incluir al menos un espacio de color tipo YUV y al menos un espacio de color tipo RGB. O, como
15 otro ejemplo, la conmutación puede incluir cambiar entre dos espacios de color usando las operaciones de reordenamiento de espacio de color. Para las operaciones de reordenamiento de espacio de color, los espacios de color pueden incluir múltiples espacios de color tipo RGB. Por otra parte, los espacios de color
20 incluyen otros y/o adicionales espacios de color.

 O, como otro ejemplo, la conmutación puede incluir cambiar entre dos de las frecuencias de muestreo de color. Las frecuencias de muestreo de color pueden incluir dos o más de 4:4:4, 4:2:2, 4:2:0 y 4:0:0. Por otra parte, las frecuencias de muestreo de color incluyen
25 otras y/o adicionales frecuencias de muestreo de color.

O, como otro ejemplo, la conmutación puede incluir cambiar entre dos de las profundidades de bits. Las profundidades de bits pueden incluir dos o más de 12 bits por muestra, 10 bits por muestra y 8 bits por muestra. Los diferentes componentes de color de una determinada unidad del video pueden tener la misma profundidad de bits o diferentes profundidades de bits. Por otra parte, las profundidades de bits incluyen otras y/o adicionales profundidades de bits.

Durante la codificación, el codificador puede predecir los valores de muestra de un bloque actual en una imagen actual con base, al menos en parte, en un conjunto de valores de muestra del contenido previamente reconstruido. Para la predicción intraimagen, el conjunto de valores de muestra es parte de la imagen actual. Para la predicción interimagen, el conjunto de valores de muestra es parte de la imagen de referencia. Cuando los valores de muestra del bloque actual y el conjunto de valores de muestra del contenido previamente reconstruido están en diferentes formatos, el codificador puede realizar las operaciones de conversión para facilitar la predicción. Los ejemplos de tales operaciones de conversión se describen en la sección VII.F.

Durante la codificación, el codificador puede realizar también el desbloqueo adaptativo en el bucle para el contenido previamente reconstruido de acuerdo con una o más reglas. Las reglas pueden determinar los diferentes espacios de color de dos bloques adyacentes en el contenido previamente reconstruido. Los ejemplos

de las reglas para el desbloqueo adaptativo se describen en la sección VII.F.

El codificador emite (1220) los datos codificados en una corriente de bits. Los datos codificados incluyen una o más señales
5 que indican cómo los espacios de color, las frecuencias de muestreo de color y/o las profundidades de bits se conmutan entre al menos algunas unidades del video en la secuencia. Los ejemplos de elementos de sintaxis para una o más señales se describen en la sección VII.D.

10 La figura 13 muestra una técnica más detallada (1300) para la conmutación adaptativa de espacios de color, frecuencias de muestreo de color y/o profundidades de bits durante la codificación. Un codificador de imagen o codificador de video tal como se describe con referencia a la figura 3 o figuras 5a-5b
15 puede realizar la técnica (1300). En particular, la figura 13 muestra las decisiones unidad por unidad tomadas por el codificador, pero las opciones para la codificación son tal como se describe con referencia a la figura 12.

El codificador obtiene (1310) la siguiente unidad (por ejemplo,
20 imagen, sección, bloque) y establece (1320) el espacio de color, frecuencia de muestreo de color y/o profundidad de bits para la unidad. Según sea necesario, el codificador convierte los valores de muestra a otro espacio de color, frecuencia de muestreo de color, y/o profundidad de bits para la unidad, entonces codifica (1330) la
25 unidad. El codificador emite (1340) los datos codificados para la

unidad, que incluyen una o más señales que indican el espacio de color, frecuencia de muestreo de color y/o profundidad de bits para la unidad. El codificador verifica (1350) si continuar con la siguiente unidad y, si es así, obtiene (1310) la siguiente unidad.

5 **C. Ejemplos de conmutación adaptativa durante la decodificación.**

La figura 14 muestra una técnica generalizada (1400) para la conmutación adaptativa de espacios de color, frecuencias de muestreo de color y/o profundidades de bits durante la
10 decodificación. Un decodificador de imagen o decodificador de video tal como se describe con referencia a la figura 4 o figura 6 puede realizar la técnica (1400).

El decodificador recibe (1400) los datos codificados en una corriente de bits. Los datos codificados incluyen una o más señales
15 que indican cómo los espacios de color, las frecuencias de muestreo de color y/o las profundidades de bits se conmutan entre al menos algunas unidades del video en la secuencia. Los ejemplos de elementos de sintaxis para una o más señales se describen en la sección VII.D.

20 El decodificador decodifica (1420) el video en la secuencia. Como parte de la decodificación, el decodificador conmuta espacios de color, frecuencias de muestreo de color y/o profundidades de bits entre al menos algunas unidades del video en la secuencia durante la decodificación. El decodificador puede conmutar espacios de
25 color, frecuencias de muestreo de color y/o profundidades de bits de

manera espacial en una imagen. El decodificador puede conmutar espacios de color, frecuencias de muestreo de color y/o profundidades de bits de imagen a imagen. Para propósitos de conmutación adaptativa, las unidades del video pueden ser imágenes de la secuencia, secciones de una determinada imagen de la secuencia, bloques de una determinada imagen de la secuencia, u otras porciones de la secuencia de video.

Por ejemplo, la conmutación puede incluir cambiar entre dos de los espacios de color usando las operaciones de transformación de espacio de color. Para las operaciones de transformación de espacio de color, los espacios de color pueden incluir al menos un espacio de color tipo YUV y al menos un espacio de color tipo RGB. O, como otro ejemplo, la conmutación puede incluir cambiar entre dos espacios de color usando las operaciones de reordenamiento de espacio de color. Para las operaciones de reordenamiento de espacio de color, los espacios de color pueden incluir múltiples espacios de color tipo RGB. Por otra parte, los espacios de color incluyen otros y/o adicionales espacios de color.

O, como otro ejemplo, la conmutación puede incluir cambiar entre dos de las frecuencias de muestreo de color. Las frecuencias de muestreo de color pueden incluir dos o más de 4:4:4, 4:2:2, 4:2:0 y 4:0:0. Por otra parte, las frecuencias de muestreo de color incluyen otras y/o adicionales frecuencias de muestreo de color.

O, como otro ejemplo, la conmutación puede incluir cambiar entre dos de las profundidades de bits. Las profundidades de bits

pueden incluir dos o más de 12 bits por muestra, 10 bits por muestra y 8 bits por muestra. Los diferentes componentes de color de una determinada unidad del video pueden tener la misma profundidad de bits o diferentes profundidades de bits. Por otra parte, las

5 profundidades de bits incluyen otras y/o adicionales profundidades de bits.

Durante la decodificación, el decodificador puede predecir los valores de muestra de un bloque actual en una imagen actual con base, al menos en parte, en un conjunto de valores de muestra del

10 contenido previamente reconstruido. Para la predicción intraimagen, el conjunto de valores de muestra es parte de la imagen actual. Para la predicción interimagen, el conjunto de valores de muestra es parte de la imagen de referencia. Cuando los valores de muestra del bloque actual y el conjunto de valores de muestra del contenido

15 previamente reconstruido están en diferentes formatos, el decodificador puede realizar las operaciones de conversión para facilitar la predicción. Los ejemplos de tales operaciones de conversión se describen en la sección VII.E.

Durante la decodificación, el decodificador puede realizar

20 también el desbloqueo adaptativo en el bucle para el contenido previamente reconstruido de acuerdo con una o más reglas. Las reglas pueden determinar los diferentes espacios de color de dos bloques adyacentes en el contenido previamente reconstruido. Los ejemplos de las reglas para el desbloqueo adaptativo se describen

25 en la sección VII.F.

La figura 15 muestra una técnica más detallada (1500) para la conmutación adaptativa de espacios de color, frecuencias de muestreo de color y/o profundidades de bits durante la decodificación. Un decodificador de imagen o decodificador de video tal como se describe con referencia a la figura 4 o figura 6 puede realizar la técnica (1500). En particular, la figura 15 muestra las decisiones unidad por unidad tomadas por el decodificador, pero las opciones para la decodificación son tal como se describe con referencia a la figura 14.

El decodificador obtiene (1510) los datos codificados para la siguiente unidad (por ejemplo, imagen, sección, bloque). Los datos codificados incluyen una o más señales que indican el espacio de color, la frecuencia de muestreo de color y/o la profundidad de bits para la unidad. El decodificador determina (1520) el espacio de color, la frecuencia de muestreo de color y/o la profundidad de bits para la unidad. El decodificador descodifica (1530) la unidad, convirtiendo los valores de muestra, según sea necesario, a partir del espacio de color, frecuencia de muestreo de color y/o profundidad de bits establecidos para la unidad de muestreo. El decodificador comprueba (1540) si continuar con la siguiente unidad y, si es así, obtiene (1510) los datos codificados para la siguiente unidad.

D. Ejemplos de señalización de la información de conmutación adaptativa.

La sección describe los ejemplos de las señales que indican

cómo los espacios de color, las frecuencias de muestreo de color y/o las profundidades de bits se conmutan entre al menos algunas unidades del video en la secuencia. En general, una señal en la corriente de bits indica una selección entre dos o más opciones para

5 codificar-decodificar una unidad. Para la adaptación de espacio de color, la señal indica una selección entre dos o más espacios de color usados cuando se codifica-decodifica una unidad. Para la adaptación de la frecuencia de muestreo de color, la señal indica una selección entre dos o más frecuencias de muestreo de color usadas

10 cuando se codifica-decodifica una unidad. Para la adaptación de la profundidad de bits, la señal indica una selección entre dos o más profundidades de bits usadas cuando se codifica-decodifica una unidad. Las diferentes profundidades de bits pueden señalizarse para diferentes componentes de color de una unidad.

15 La señal puede ser un valor indicador para una determinada unidad, lo cual indica una selección entre dos opciones para los espacios de color (o frecuencias de muestreo de color o profundidades de bit) para la codificación-decodificación de la determinada unidad. O, la señal puede ser algún otro elemento de

20 sintaxis para una determinada unidad (o múltiples indicadores para la determinada unidad), que indica una selección entre tres o más opciones para los espacios de color (frecuencias de muestreo de color o profundidades de bit) para la codificación-decodificación de la determinada unidad.

25 La señal para una determinada unidad puede estar presente en

cualquiera de varios niveles de la corriente de bits de sintaxis. Para la adaptación imagen por imagen, por ejemplo, un elemento de sintaxis que indica una selección de un espacio de color (frecuencia de muestreo de color o profundidad de bits) puede ser parte de un PPS. O, para la adaptación sección por sección, un elemento de sintaxis que indica una selección de un espacio de color (frecuencia de muestreo de color o profundidad de bits) puede ser parte de un encabezado de sección. O, para la adaptación bloque por bloque, un elemento de sintaxis que indica una selección de espacio de color (frecuencia de muestreo de color o profundidad de bits) puede ser parte de una estructura de sintaxis para un bloque (por ejemplo, estructura de sintaxis de CU). En general, la adaptación imagen por imagen es menos flexible que la adaptación sección por sección o la adaptación bloque por bloque, pero usa menos bits de señalización. La adaptación bloque por bloque usa más bits de señalización que la adaptación sección por sección, pero es más flexible.

La señal de una determinada unidad puede codificarse por entropía, por ejemplo, usando la CABAC, o formatearse como un valor de longitud fija. La señal para una determinada unidad puede codificarse de manera predictiva. En este caso, un valor delta para un parámetro para la determinada unidad representa una diferencia entre un valor existente del parámetro y su predictor. El predictor puede ser un valor del parámetro señalado a un nivel superior de sintaxis (por ejemplo, como parte de un SPS o como otro parámetro de capa de secuencia). Por ejemplo, un valor de profundidad de bits

para un bloque o sección actual puede señalizarse como un valor delta con relación a una profundidad de bits específica para el SSP. O, el predictor puede basarse en uno o más valores existentes del parámetro para una o más unidades anteriores del video (por ejemplo, el valor existente del parámetro para la última unidad codificada, o una media de los valores existentes del parámetro en una inmediación alrededor de la determinada unidad. Por ejemplo, la profundidad de bits de un bloque actual puede señalizarse como un valor delta con relación a la profundidad de bits del bloque codificado anterior.

Los datos codificados pueden incluir también la información que identifica o define los espacios de color disponibles, la información que identifica o define las frecuencias de muestreo de color disponibles o la información que identifica o define las profundidades de bits disponibles. Tal información puede señalizarse como parte de un SPS, PPS, encabezado de sección u otra estructura de la corriente de bits de sintaxis. Los espacios de color, frecuencias de muestreo de color o profundidades de bits disponibles pueden seleccionarse a partir de entre las opciones predefinidas que el codificador y el decodificador conoce. O, un codificador puede especificar un espacio de color modificado (o frecuencia de muestreo de color o profundidad de bits) en los datos codificados, para el uso por el codificador y el decodificador. Por ejemplo, la información puede incluir una matriz de valores que pueden usarse para las operaciones de transformación de espacio de color entre dos de los

espacios de color. Especificar una opción modificada usa más bits, pero es más flexible que identificar un conjunto de opciones predefinidas. Otra posibilidad es usar las opciones predefinidas conocidas para el codificador y el decodificador, sin ninguna
5 información de identificación señalizada como parte de un SPS, PPS, encabezado de sección, etcétera. Esto no usa ningún bit para la información adicional, pero puede llevar a la codificación menos eficiente de las señales, cuando muchas de las opciones no se usan en absoluto durante la codificación de una secuencia.

10 La información que identifica o define los espacios de color disponibles (o frecuencias de muestreo de color o profundidades de bits) se usa en combinación con las señales que indican la selección de las opciones para las unidades del video. Por ejemplo, un SPS puede incluir la información que identifica los espacios de color
15 disponibles y cuyas conversiones de espacio de color se permiten para una secuencia, y un elemento de sintaxis por PPS (o encabezado de sección o estructura de sintaxis de bloque) indica una selección de la lista especificada de los espacios de color disponibles. O, un encabezado de sección puede incluir la
20 información que identifica las frecuencias de muestreo de color disponibles, y un elemento de sintaxis por estructura de sintaxis de bloque indica una selección a partir de la lista especificada de frecuencias de muestreo de color disponibles. O, puede usarse alguna combinación de las variaciones anteriores. Por ejemplo, un
25 SPS puede indicar una lista de espacios de color predefinidos, con

un PPS o encabezado de sección, que incluye una selección de uno de los espacios de color predefinido o que incluye la información para un espacio de color modificado.

En algunas implementaciones ejemplares, las reglas
5 adicionales aplican para la señalización a nivel de CU de los indicadores de espacio de color. Si la CU actual es una CU prevista interespacial, un indicador de espacio de color puede señalizarse como parte de una estructura de sintaxis de CU. Si la CU actual es una CU prevista intra-BC o una CU prevista interimagen, un
10 indicador de espacio de color puede señalizarse como parte de una estructura de sintaxis de CU cuando hay una magnitud residual, pero no se permiten las conversiones de espacio de color que solo implican el reordenamiento (por ejemplo, entre RGB, BGR y GBR). En este contexto, tales conversiones de espacio de color no mejoran
15 el rendimiento y se desperdiciarían los bits consumidos señalizando cómo reordenar los componentes de color.

E. Ejemplos de las operaciones de conversión para la predicción.

Durante la codificación o decodificación, un codificador o
20 decodificador puede predecir los valores de muestra de un bloque actual en una imagen actual con base, al menos en parte, en un conjunto de valores de muestra del contenido previamente reconstruido. Para la predicción intraimagen, el conjunto de valores de muestra es parte de la imagen actual (por ejemplo, valores de
25 muestra en las posiciones espacialmente adyacentes para la

predicción intraespacial o valores de muestra de los bloques anteriores para la predicción intra-BC). Para la predicción interimagen, el conjunto de valores de muestra es parte de la imagen de referencia. Si los valores de muestra del bloque actual y el conjunto de valores de muestra del contenido previamente reconstruido están en diferentes formatos, el codificador o el decodificador puede realizar las operaciones de conversión para facilitar la predicción de los valores de muestra del bloque actual.

Por ejemplo, si el conjunto de valores de muestra del contenido previamente reconstruido está en el espacio de color GBR y el bloque actual se codifica en el espacio de color YCbCr, el codificador o decodificador puede convertir el conjunto de valores de muestra del contenido previamente reconstruido al espacio de color YCbCr antes de la predicción intraimagen o interimagen usando el conjunto de valores de muestra convertidos.

Para simplificar la implementación, el codificador y decodificador pueden almacenar los valores de muestra del contenido previamente reconstruido en un formato "principal". El formato principal proporciona una representación común del contenido reconstruido en un espacio de color particular, con una frecuencia de muestreo de color particular y los valores de muestra que tienen una profundidad de bits particular. Dependiendo de la implementación, el formato principal puede ser GBR a 4:4:4 con valores de muestra de 12 bits, YUV a 4:4:4 con valores de muestra de 12 bits o alguna otra combinación de espacio de color, frecuencia

de muestreo de color y profundidad de bits. El formato principal está predefinido para el codificador y decodificador. Por otra parte, el formato principal puede especificarse mediante una señal en la corriente de bits, que indica un espacio de color, frecuencia de muestreo de color y/o profundidad de bits para el almacenamiento del contenido reconstruido. El formato principal puede seleccionarse, por ejemplo, que será el formato de codificación más común para las unidades de una secuencia de video, con el fin de evitar las operaciones de conversión en muchos casos. Usar GBR a 4:4:4 con valores de muestra de 12 bits como el formato principal, puede ser eficiente para el contenido de captura de pantalla cuando la conmutación de espacio de color es común, pero la conmutación de frecuencia de muestreo de color es poco común. Usar YUV a 4:4:4 con valores de muestra de 12 bits como el formato principal, puede ser eficiente cuando la conmutación de espacio de color y la conmutación de frecuencia de muestreo de color son comunes.

La figura 16 muestra una técnica (1600) para la predicción de los valores de muestra de un bloque codificados en un formato a partir de un conjunto de valores de muestra del contenido previamente reconstruido en un diferente formato. Un codificador o descodificador puede realizar la técnica (1600) durante la codificación o decodificación.

En primer lugar, el codificador o decodificador obtiene (1610) un conjunto de valores de muestra del contenido previamente reconstruido que se usará para la predicción de los valores de

muestra de un bloque actual de una imagen actual. El conjunto de valores de muestra puede ser parte de la imagen actual o una imagen de referencia, dependiendo del tipo de predicción.

Los valores de muestra del contenido previamente reconstruido se almacenan en un primer formato (por ejemplo, formato principal). El codificador o decodificador comprueba (1620) si el bloque actual está codificado en el primer formato. Más generalmente, el codificador o decodificador puede evaluar si el conjunto de valores de muestra del contenido previamente reconstruido y los valores de muestra del bloque actual están en diferentes formatos.

Si el bloque actual está codificado en el primer formato (por ejemplo, el formato principal), el codificador o decodificador no necesita realizar las operaciones de conversión antes de la predicción. El codificador o decodificador predicen (1640) los valores de muestra del bloque actual usando un conjunto de valores de muestra del contenido previamente reconstruido. Durante la codificación, el codificador puede entonces generar, codificar y reconstruir (1650) los valores residuales. Durante la decodificación, el decodificador puede reconstruir (1650) los valores residuales. El codificador o decodificador entonces reconstruye (1660) los valores de muestra del bloque actual con base, al menos en parte, en los valores de muestra previstos. Para un bloque no omitido, el codificador o decodificador puede combinar los valores de la muestra previstos con los valores residuales reconstruidos para producir los valores de muestra reconstruidos. El codificador o decodificador

entonces almacenan (1670) los valores de muestra reconstruidos del bloque actual como parte del contenido previamente reconstruido. Si el bloque actual está codificado en el primer formato (por ejemplo, el formato principal), el codificador o decodificador no necesita realizar las operaciones de conversión después de la predicción.

Por otro lado, si el bloque actual no se codifica en el primer formato, el codificador o decodificador convierte (1632) el conjunto de valores de muestra del contenido previamente reconstruido del primer formato al segundo formato en el cual se codifica el bloque actual. Por ejemplo, si el primero y segundo formato tienen diferentes espacios de color, el codificador o decodificador realiza las operaciones de transformación de espacio de color y/o las operaciones de reordenamiento de espacio de color para cambiar el espacio de color del conjunto de valores de muestra del contenido previamente reconstruido al segundo formato. Por ejemplo, el codificador o decodificador puede convertir un conjunto P de valores de muestra del contenido previamente reconstruido del espacio de color RGB al espacio de color YCoCg, tal como se detalla a continuación.

Suponiendo que los valores de muestra para una posición en el espacio de color RGB sean PS_R , PS_G y PS_B , respectivamente, y la profundidad de bits actual sea $bit_depth_{current}$. El valor de muestra mínimo es 0, y el valor de muestra máximo es $(1 \ll bit_depth_{current}) - 1$ (por ejemplo, 255 para los valores de muestra de 8 bits). Para ajustar el intervalo dinámico de Co y Cg, el término add_value se

define como $1 \ll (\text{bit_depth}_{\text{current}} - 1)$ (por ejemplo, 128 para los valores de muestra de 8 bits). Los valores de muestra Ps_Y , Ps_{Co} , y Ps_{Cg} para la posición en el espacio de color YCoCg, pueden calcularse por:

$$\begin{aligned} 5 \quad Ps_Y &= (Ps_R + (2 * Ps_G) + Ps_B) \gg 2 \\ Ps_{Co} &= ((Ps_R - Ps_B) \gg 1) + \text{add_value} \\ Ps_{Cg} &= (((2 * Ps_G) - Ps_R - Ps_B) \gg 2) + \text{add_value} \end{aligned}$$

Los valores de muestra Ps_Y , Ps_{Co} , y Ps_{Cg} se limitan al intervalo definido por el valor de muestra mínimo y el valor de muestra
10 máximo.

Más generalmente, la matriz de conversión de color $CC_matrix_{1_to_2}$ para la transformación de un primer espacio de color a un segundo color espacio, puede definirse como:

$$\begin{aligned} &[C_{00}, C_{01}, C_{02} \\ 15 \quad &C_{10}, C_{11}, C_{12} \\ &C_{20}, C_{21}, C_{22}] \end{aligned}$$

Para ajustar el intervalo dinámico de la salida, puede usarse una matriz de desplazamientos $CC_offsets_{1_to_2}$. Los desplazamientos $CC_offsets_{1_to_2}$ pueden definirse como:

$$\begin{aligned} 20 \quad &[o_0, \\ &o_1, \\ &o_2] \end{aligned}$$

Suponiendo que los valores de muestra para una posición en el primer espacio de color sean $Ps_{CC1} = [Ps_{00}, Ps_{01}, Ps_{02}]^T$. Los valores
25 de muestra Ps_{CC2} para la posición en el segundo espacio de color,

pueden calcularse como:

$$Ps_{CC2} = CC_matrix_{1_to_2} * Ps_{CC1} + CC_offsets_{1_to_2}$$

Los valores de muestra Ps_{CC1} se limitan al intervalo adecuado.

Las operaciones de reordenamiento de espacio de color para el
 5 conjunto de P de los valores de muestra del contenido previamente
 reconstruido, pueden implementarse al reordenar los componentes
 de color de un bloque o plano de los valores de muestra, con el fin
 de cambiar el orden de los componentes de color secundarios y
 primarios.

10 Si el primer y segundo formato tienen diferentes frecuencias de
 muestreo de color, el codificador o decodificador realiza la reducción
 de muestra o el incremento de muestra (con o sin filtración), para
 cambiar la frecuencia de muestreo de color del conjunto de valores
 de muestra del contenido previamente reconstruido al segundo
 15 formato. Por ejemplo, el codificador o decodificador puede convertir
 un conjunto P de valores de muestra del contenido previamente
 reconstruido de la frecuencia de muestreo de color a 4:4:4 a la
 frecuencia de muestreo de color a 4:2:0, tal como se detalla a
 continuación.

20 Suponiendo que los valores de muestra del contenido
 previamente reconstruido sean los valores de muestra de YUV a
 4:4:4. La resolución espacial del video de YUV a 4:4:4 es *anchura* x
altura para todos los componentes de color. Después de la
 conversión de frecuencia de muestreo de color, la resolución
 25 espacial del video de YUV a 4:2:0 es *anchura* x *altura* para el

componente Y y $(anchura/2) \times (altura/2)$ para los componentes U y V. Para cada posición (i,j) , para $0 \leq i < anchura$, y $0 \leq j < altura$, un codificador o decodificador puede calcular los valores de muestra para el video de YUV a 4:2:0 como:

$$\begin{aligned}
 5 \quad Y_{420}[i][j] &= Y_{444}[i][j] \\
 U_{420}[i][j] &= (U_{444}[2*i][2*j] + U_{444}[2*i+1][2*j] + U_{444}[2*i][2*j+1] + \\
 &U_{444}[2*i+1][2*j+1]) \gg 2 \\
 V_{420}[i][j] &= (V_{444}[2*i][2*j] + V_{444}[2*i+1][2*j] + V_{444}[2*i][2*j+1] + \\
 &V_{444}[2*i+1][2*j+1]) \gg 2
 \end{aligned}$$

10 En este método, los valores de muestra para los componentes U y V se determinan sin filtración. Por otra parte, el codificador o decodificador puede usar la filtración cuando obtiene los componentes U y V con reducción de muestra.

Si el primer y segundo formato tienen diferentes profundidades de bits, el codificador o decodificador cambia el conjunto de P de los valores de muestra del contenido previamente reconstruido de la profundidad de bits bit_depth_{ref} a la profundidad de bits $bit_depth_{current}$ de los valores de muestra del bloque actual, tal como se detalla a continuación.

```

20   if  $bit\_depth_{ref} > bit\_depth_{current}$ 
      shift =  $bit\_depth_{ref} - bit\_depth_{current}$ 
      add =  $1 \ll (shift - 1)$ 
       $P = (P + add) \gg shift$ 
      else if  $bit\_depth_{ref} < bit\_depth_{current}$ 
25   shift =  $bit\_depth_{current} - bit\_depth_{ref}$ 

```

$P \ll = \text{shift}$

Por ejemplo, suponiendo que la profundidad de bits del primer formato es de 10 bits por muestra, la profundidad de bits del segundo formato es de 8 bits por muestra y el bloque actual usa la predicción intraespacial. El codificador o decodificador obtiene los valores de muestra de 10 bits de las posiciones próximas en la predicción intraespacial. Los valores de muestra de 10 bits se convierten a valores de muestra de 8 bits, que entonces se usan para generar los valores de muestra previstos de 8 bits para el bloque actual.

Después de la conversión (1632), el codificador o decodificador predicen (1642) los valores de muestra del bloque actual usando el conjunto de valores de muestra convertidos del contenido previamente reconstruido. Durante la codificación, el codificador puede entonces generar, codificar y reconstruir (1652) los valores residuales. Durante la decodificación, el decodificador puede reconstruir (1652) los valores residuales. El codificador o decodificador entonces reconstruye (1662) los valores de muestra del bloque actual con base, al menos en parte, en los valores de muestra previstos. Para un bloque no omitido, el codificador o decodificador puede combinar los valores de la muestra previstos con los valores residuales reconstruidos para producir los valores de muestra reconstruidos.

En este momento, los valores de muestra reconstruidos del bloque actual están en el segundo formato. Los valores de muestra

reconstruidos del bloque actual se convierten al primer formato antes del almacenamiento como parte del contenido reconstruido. El codificador o decodificador convierte (1672) los valores de muestra reconstruidos del bloque actual del segundo formato al primer
 5 formato, y almacena (1670) los valores de muestra reconstruidos convertidos del bloque actual ahora en el primer formato (por ejemplo, formato principal), como parte del contenido previamente reconstruido.

Por ejemplo, si el primer y segundo formato tienen diferentes
 10 espacios de color, el codificador o decodificador realiza las operaciones de transformación de espacio de color y/o las operaciones de reordenamiento de espacio de color para cambiar el espacio de color de los valores de muestra reconstruidos del bloque actual al primer formato. (Las operaciones de conversión de espacio
 15 de color del segundo formato al primer formato generalmente replican las operaciones anteriores de conversión de espacio de color del primer formato al segundo formato). Por ejemplo, el codificador o decodificador puede convertir los valores de muestra reconstruidos R de un bloque actual del espacio de color YCoCg al
 20 espacio de color RGB, tal como se detalla a continuación.

Suponiendo que los valores de muestra para una posición en el espacio de color YCoCg son Rs_Y , Rs_{Co} y Rs_{Cg} , respectivamente, y la profundidad de bits actual es $bit_depth_{current}$. Para ajustar el intervalo dinámico, el término add_value se define como $1 \ll (bit_depth_{current} -$
 25 $1)$ (por ejemplo, 128 para los valores de muestra de 8 bits). Los

valores de muestra Rs_R , Rs_G , y Rs_B para la posición en el espacio de color RGB, pueden calcularse por:

$$Rs_{Co} -= \text{add_value}$$

$$Rs_{Cg} -= \text{add_value}$$

$$5 \quad Rs_R = Rs_Y + Rs_{Co} - Rs_{Cg}$$

$$Rs_G = Rs_Y + Rs_{Cg}$$

$$Rs_B = Rs_Y - Rs_{Co} - Rs_{Cg}$$

Los valores de muestra Rs_R , Rs_G , y Rs_B se limitan al intervalo definido por el valor de muestra mínimo y el valor de muestra máximo.

10 Más generalmente, la matriz de conversión de color $CC_matrix_{2_to_1}$ para la transformación de un segundo espacio de color de regreso al primer espacio de color, puede definirse como:

$$[c'_{00}, c'_{01}, c'_{02}$$

$$c'_{10}, c'_{11}, c'_{12}$$

$$15 \quad c'_{20}, c'_{21}, c'_{22}]$$

Para ajustar el intervalo dinámico de la salida, puede usarse una matriz de desplazamientos $CC_offsets_{2_to_1}$. Los desplazamientos $CC_offsets_{2_to_1}$ pueden definirse como:

$$[o'_0,$$

$$20 \quad o'_1,$$

$$o'_2]$$

Suponiendo que los valores de muestra para una posición en el segundo espacio de color sean $Rs_{CC1} = [Rs_{00}, Rs_{01}, Rs_{02}]^T$. Los valores de muestra Rs_{CC1} para la posición en el primer espacio de color, pueden calcularse como:

25

$$Rs_{CC1} = CC_matrix_{2_to_1} * (Rs_{CC2} + CC_offsetS_{2_to_1})$$

Los valores de muestra Rs_{CC1} se limitan al intervalo adecuado.

Las operaciones de reordenamiento de espacio de color para los valores de muestra reconstruidos R pueden implementarse al reordenar los componentes de color de un bloque o plano de los valores de muestra, con el fin de cambiar el orden de los componentes de color secundarios y primarios.

Si el primer y segundo formato tienen diferentes frecuencias de muestreo de color, el codificador o decodificador realiza la reducción de muestra o el incremento de muestra (con o sin filtración), para cambiar la frecuencia de muestreo de color de los valores de muestra reconstruidos del bloque actual al primer formato. (Las operaciones de conversión de frecuencia de muestreo de color del segundo formato al primer formato generalmente replican las operaciones anteriores de conversión de frecuencia de muestreo de color del primer formato al segundo formato). Por ejemplo, el codificador o decodificador puede convertir los valores de muestra reconstruidos R de un bloque actual de la frecuencia de muestreo de color a 4:2:0 a la frecuencia de muestreo de color a 4:4:4 al reducir la muestra de los componentes de color con incremento de muestra.

Si el primer y segundo formato tienen diferentes profundidades de bits, el codificador o decodificador cambia los valores de muestra reconstruidos R del bloque actual de la profundidad de bits $bit_depth_{current}$ a la profundidad de bits bit_depth_{ref} de los valores de muestra del contenido reconstruido, tal como se detalla a

continuación.

$\text{if } \text{bit_depth}_{ref} > \text{bit_depth}_{current}$

$\text{shift} = \text{bit_depth}_{ref} - \text{bit_depth}_{current}$

$R \ll \text{shift}$

5 $\text{else if } \text{bit_depth}_{ref} < \text{bit_depth}_{current}$

$\text{shift} = \text{bit_depth}_{current} - \text{bit_depth}_{ref}$

$\text{add} = 1 \ll (\text{shift} - 1)$

$R = (R + \text{add}) \gg \text{shift}$

El codificador o decodificador puede repetir la técnica (1600)
10 de otros bloques de una sección o imagen.

En la figura 61, la conversión (1632) se lleva a cabo antes de la predicción (1642). Por otra parte, para algunos tipos de predicción (por ejemplo, predicción intraespacial, predicción intra-BC), la conversión puede ir después de la predicción. En este caso, el
15 codificador o decodificador predice los valores de muestra del bloque actual usando el conjunto de valores de muestra del contenido previamente reconstruido, después convierte los valores de muestra previstos del bloque actual del primer formato al segundo formato. El
20 codificador o decodificador reconstruye los valores de muestra del bloque actual con base, al menos en parte, en los valores de muestra previstos convertidos del bloque actual.

Por ejemplo, suponiendo que la profundidad de bits del primer formato es de 10 bits por muestra, la profundidad de bits del segundo formato es de 8 bits por muestra y el bloque actual usa la
25 predicción intraespacial. El codificador o decodificador obtiene los

valores de muestra de 10 bits de las posiciones próximas en la predicción intraespacial. Los valores de muestra de 10 bits se usan para generar los valores de muestra previstos de 10 bits, que entonces se convierten en valores de muestra de 8 bits.

5 Por otra parte, los valores de muestra de las unidades del contenido reconstruido pueden almacenarse en el formato en el cual se codifican las unidades. También se almacena alguna indicación de los formatos de codificación para las unidades respectivas. Los valores de muestra del contenido reconstruido se convierten a otro
10 formato, según sea necesario, para la predicción.

F. Ejemplos de las operaciones de desbloqueo.

Durante la codificación o decodificación, un codificador o decodificador puede realizar el desbloqueo adaptativo en el bucle para el contenido previamente reconstruido de acuerdo con una o
15 más reglas. En algunos casos, el desbloqueo puede aplicarse a través de un límite entre los dos bloques codificados con diferentes espacios de color.

Por ejemplo, en algunas implementaciones del desbloqueo adaptativo, la intensidad de un filtro de desbloqueo para un límite
20 entre los dos bloques se basa, al menos en parte, en si cualquiera de los bloques adyacentes tiene valores residuales distintos de cero. Cuando dos bloques adyacentes tienen diferentes espacios de color, la intensidad del filtro de desbloqueo para el límite entre los bloques, puede depender de si los componentes primarios de los dos bloques
25 adyacentes tienen valores residuales distintos de cero. Por ejemplo,

se evalúa el componente primario del primer bloque adyacente, y se evalúa el componente primario del segundo bloque adyacente, aunque los espacios de color son diferentes para los dos bloques. La intensidad de filtro depende de si cualquier componente primario para los bloques tiene los valores residuales distintos de cero.

Por otra parte, las reglas pueden determinar los diferentes espacios de color de los dos bloques adyacentes. En este caso, cuando dos bloques adyacentes tienen diferentes espacios de color, la intensidad del filtro de desbloqueo para el límite entre los bloques depende de si los componentes correspondientes de los dos bloques adyacentes tienen valores residuales distintos de cero. Por ejemplo, suponiendo que el espacio de color de un bloque adyacente es GBR, y el espacio de color del otro bloque adyacente es RGB. Ambos espacios de color son espacios de color tipo RGB. Cuando se determina la intensidad de filtro para el componente G, el codificador o decodificador considera los valores residuales del primer componente (G) del primer bloque y los valores residuales del segundo componente (G) del segundo bloque. Esta variación de desbloqueo adaptativo puede usarse para los bloques con diferentes espacios de color tipo RGB.

Teniendo en cuenta muchas modalidades posibles a las cuales pueden aplicar los principios de la invención divulgada, debe reconocerse que las modalidades ilustradas son solo ejemplos preferidos de la invención y no deben interpretarse como limitativos del ámbito de la invención. Por el contrario, el ámbito de la invención

se define por las siguientes reivindicaciones. Por lo tanto, reclamamos como nuestra invención todo aquello que se encuentre dentro del ámbito y esencia de estas reivindicaciones.

5

10

15

20

25