

19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 753 960**

51 Int. Cl.:

H04N 19/30 (2014.01)

H04N 19/42 (2014.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

86 Fecha de presentación y número de la solicitud internacional: **14.10.2014 PCT/US2014/060413**

87 Fecha y número de publicación internacional: **23.04.2015 WO15057656**

96 Fecha de presentación y número de la solicitud europea: **14.10.2014 E 14792941 (8)**

97 Fecha y número de publicación de la concesión europea: **07.08.2019 EP 3058731**

54 Título: **Tabla de consulta tridimensional basada en la escalabilidad de la gama de colores en la codificación de vídeo multicapa**

30 Prioridad:

14.10.2013 US 201361890843 P
10.10.2014 US 201414512177

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

15.04.2020

73 Titular/es:

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

72 Inventor/es:

LI, XIANG;
RAPAKA, KRISHNAKANTH;
CHEN, JIANLE y
KARCZEWICZ, MARTA

74 Agente/Representante:

FORTEA LAGUNA, Juan José

ES 2 753 960 T3

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

DESCRIPCIÓN

Tabla de consulta tridimensional basada en la escalabilidad de la gama de colores en la codificación de vídeo multicapa

5

SOLICITUDES RELACIONADAS

CAMPO TÉCNICO

10 **[0001]** Esta divulgación se refiere a la codificación de vídeo.

ANTECEDENTES

15 **[0002]** Las capacidades del vídeo digital pueden incorporarse en una amplia gama de dispositivos, incluyendo televisores digitales, sistemas de difusión directa digital, sistemas de difusión inalámbrica, asistentes digitales personales (PDA), ordenadores portátiles o de escritorio, tabletas electrónicas, lectores de libros electrónicos, cámaras digitales, dispositivos de grabación digital, reproductores de medios digitales, dispositivos de videojuegos, consolas de videojuegos, teléfonos celulares o de radio por satélite, los denominados "teléfonos inteligentes", dispositivos de videoconferencia, dispositivos de emisión de vídeo en tiempo real y similares. Los dispositivos de vídeo digital implementan técnicas de codificación de vídeo, tales como las descritas en las normas definidas por MPEG-2, MPEG-4, ITU-T H.263, ITU-T H.264/MPEG-4, Parte 10, Codificación de Vídeo Avanzada (AVC), la norma de Codificación de Vídeo de Alta Eficiencia (HEVC) y las extensiones de dichas normas. Los dispositivos de vídeo pueden transmitir, recibir, codificar, decodificar y/o almacenar información de vídeo digital más eficazmente, implementando tales técnicas de codificación de vídeo.

25

30 **[0003]** Las técnicas de codificación de vídeo incluyen la predicción espacial (intraimagen) y/o la predicción temporal (entre imágenes) para reducir o eliminar la redundancia intrínseca en las secuencias de vídeo. Para la codificación de vídeo basada en bloques, un fragmento de vídeo (por ejemplo, una trama de vídeo o una parte de una trama de vídeo) puede dividirse en bloques de vídeo, que también pueden denominarse bloques de árbol, unidades de codificación (CU) y/o nodos de codificación. Los bloques de vídeo en un fragmento intracodificado (1) de una imagen se codifican usando la predicción espacial con respecto a muestras de referencia en bloques contiguos de la misma imagen. Los bloques de vídeo en un fragmento intercodificado (P o B) de una imagen pueden usar la predicción espacial con respecto a muestras de referencia en bloques próximos en la misma imagen o la predicción temporal con respecto a muestras de referencia en otras imágenes de referencia. Las imágenes pueden denominarse tramas, y las imágenes de referencia pueden denominarse tramas de referencia.

35

40 **[0004]** La predicción espacial o temporal da como resultado un bloque predictivo para un bloque a codificar. Los datos residuales representan diferencias de píxeles entre el bloque original a codificar y el bloque predictivo. Un bloque intercodificado se codifica de acuerdo con un vector de movimiento que apunta a un bloque de muestras de referencia que forman el bloque predictivo y los datos residuales que indican la diferencia entre el bloque codificado y el bloque predictivo. Un bloque intracodificado se codifica de acuerdo con un modo de intracodificación y los datos residuales. Para una mayor compresión, los datos residuales pueden transformarse desde el dominio de píxel a un dominio de transformada, dando como resultado coeficientes de transformada residuales, que a continuación se pueden cuantificar. Los coeficientes de transformada cuantificados, dispuestos inicialmente en una formación bidimensional, pueden explorarse con el fin de generar un vector unidimensional de coeficientes de transformada, y puede aplicarse codificación por entropía para lograr aún más compresión.

45

50

SUMARIO

55 **[0005]** En general, esta divulgación describe técnicas para una tabla de consulta tridimensional (3D) basada en la escalabilidad de la gama de colores en la codificación de vídeo multicapa. Como ejemplo, la codificación de vídeo multicapa puede estar de acuerdo con la norma de codificación de vídeo de alta eficiencia (HEVC), incluyendo cualquier extensión de codificación de vídeo escalable, una extensión de codificación de vídeo multivista, una extensión de codificación de vídeo 3D (es decir, codificación de vídeo multivista más profundidad) u otras extensiones de codificación de vídeo multicapa para HEVC. Los codificadores de vídeo y/o los decodificadores de vídeo pueden usar técnicas para la escalabilidad de la gama de colores para generar imágenes de referencia entre capas cuando una gama de colores para una capa inferior de datos de vídeo sea diferente de una gama de colores para una capa superior de datos de vídeo.

60

65 **[0006]** Por ejemplo, un codificador de vídeo y/o un decodificador de vídeo pueden realizar primero la predicción de color usando una tabla de consulta 3D para la escalabilidad de la gama de colores para convertir los datos de color de una imagen de referencia para la capa inferior en la gama de colores para la capa superior, y entonces generar imágenes de referencia entre capas basándose en los datos de color convertidos. De acuerdo con las técnicas descritas en esta divulgación, el codificador y/o decodificador de vídeo genera al menos una tabla de consulta 3D que tiene un tamaño que es diferente para el componente de luma que para el primer componente de croma y el segundo componente de croma. El codificador de vídeo y/o el decodificador de vídeo pueden

generar esta tabla de consulta 3D asimétrica usando un número diferente de segmentos para la dimensión de luma de la tabla de consulta 3D. De esta manera, el tamaño total de la tabla puede reducirse mientras se mantiene un buen rendimiento de la codificación.

5 **[0007]** En un ejemplo, esta divulgación está dirigida a un procedimiento de procesamiento de datos de vídeo multicapa, el procedimiento que comprende generar al menos una tabla de consulta 3D para la escalabilidad de la gama de colores, en el que la, al menos una, tabla de consulta 3D tiene un tamaño que es diferente para un componente de luma que para cada uno de un primer componente de croma y un segundo componente de croma; realizar predicciones de color utilizando la, al menos una, tabla de consulta 3D para convertir datos de color de una imagen de referencia en una primera gama de colores para una capa inferior de datos de vídeo a una segunda gama de colores para una capa superior de datos de vídeo; y generar al menos una imagen de referencia entre capas para la capa superior de los datos de vídeo basándose en los datos de color convertidos.

15 **[0008]** En otro ejemplo, esta divulgación está dirigida a un dispositivo de procesamiento de vídeo para procesar datos de vídeo multicapa, el dispositivo que comprende una memoria configurada para almacenar los datos de vídeo multicapa, y uno o más procesadores en comunicación con la memoria. Los uno o más procesadores están configurados para generar al menos una tabla de consulta 3D para la escalabilidad de la gama de colores, en el que la, al menos una, tabla de consulta 3D tiene un tamaño que es diferente para un componente de luma que para un primer componente de croma y un segundo componente de croma, realizar predicciones de color utilizando la, al menos una, tabla de consulta 3D para convertir datos de color de una imagen de referencia en una primera gama de colores para una capa inferior de datos de vídeo a una segunda gama de colores para una capa superior de datos de vídeo, y generar al menos una imagen de referencia entre capas para la capa superior de los datos de vídeo basándose en los datos de color convertidos.

25 **[0009]** En otro ejemplo, esta divulgación está dirigida a un dispositivo de procesamiento de vídeo para procesar datos de vídeo multicapa, el dispositivo que comprende medios para generar al menos una tabla de consulta 3D para la escalabilidad de la gama de colores, en el que la, al menos una, tabla de consulta 3D tiene un tamaño que es diferente para un componente de luma que para un primer componente de croma y un segundo componente de croma; medios para realizar predicciones de color utilizando la, al menos una, tabla de consulta 3D para convertir datos de color de una imagen de referencia en una primera gama de colores para una capa inferior de datos de vídeo a una segunda gama de colores para una capa superior de datos de vídeo; y medios para generar al menos una imagen de referencia entre capas para la capa superior de los datos de vídeo basándose en los datos de color convertidos.

35 **[0010]** En otro ejemplo, esta divulgación está dirigida a un medio de almacenamiento legible por ordenador que almacena instrucciones para procesar datos de vídeo multicapa que, cuando se ejecutan, hacen que uno o más procesadores generen al menos una tabla de consulta 3D para la escalabilidad de la gama de colores, en el que al menos una tabla de consulta 3D tiene un tamaño diferente para un componente de luma que para un primer componente de croma y un segundo componente de croma; realizar predicciones de color utilizando la, al menos una, tabla de consulta 3D para convertir datos de color de una imagen de referencia en una primera gama de colores para una capa inferior de datos de vídeo a una segunda gama de colores para una capa superior de datos de vídeo; y generar al menos una imagen de referencia entre capas para la capa superior de los datos de vídeo basándose en los datos de color convertidos.

45 **[0011]** Los detalles de uno o más ejemplos se exponen en los dibujos adjuntos y en la siguiente descripción. Otras características, objetos y ventajas resultarán evidentes a partir de la descripción y de los dibujos, y a partir de las reivindicaciones.

BREVE DESCRIPCIÓN DE LOS DIBUJOS

50 **[0012]**

La FIG. 1 es un diagrama de bloques que ilustra un sistema de codificación y de decodificación de vídeo de ejemplo que puede utilizar técnicas para una tabla de consulta 3D basada en la escalabilidad de la gama de colores.

La FIG. 2 es una ilustración conceptual que muestra un ejemplo de escalabilidad en tres dimensiones diferentes.

60 La FIG. 3 es una ilustración conceptual que muestra una estructura de ejemplo de un flujo de bits de codificación de vídeo escalable.

La FIG. 4 es una ilustración conceptual que muestra unidades de acceso de codificación de vídeo escalables de ejemplo en orden de flujo de bits.

65

La FIG. 5 es un diagrama de bloques que ilustra una extensión de codificación de vídeo escalable de ejemplo para el codificador HEVC (SHVC).

La FIG. 6 es un gráfico que ilustra una gama de colores de ejemplo de una secuencia de vídeo de muestra.

La FIG. 7 es un diagrama de bloques que ilustra la conversión de la gama de colores BT.709 de alta definición (HD) en la gama de colores BT.2020 de ultraalta definición (UHD).

La FIG. 8 es un diagrama de bloques que ilustra un codificador de gama de colores escalable que incluye una unidad de procesamiento de predicción de color que puede generar una imagen de referencia entre capas cuando una gama de colores de capa base y una gama de colores de capa de mejora son diferentes.

La FIG. 9 es una ilustración conceptual que muestra una tabla de consulta 3D de ejemplo para la escalabilidad de la gama de colores.

La FIG. 10 es una ilustración conceptual que muestra la interpolación trilineal con una tabla de consulta 3D para la escalabilidad de la gama de colores.

La FIG. 11 es un diagrama de bloques que ilustra un ejemplo de un codificador de vídeo que puede implementar técnicas para utilizar una tabla de consulta 3D basada en la escalabilidad de la gama de colores en la codificación de vídeo multicapa.

La FIG. 12 es un diagrama de bloques que ilustra un ejemplo de un codificador de vídeo que puede implementar técnicas para utilizar una tabla de consulta 3D basada en la escalabilidad de la gama de colores en la codificación de vídeo multicapa.

La FIG. 13 es un diagrama de flujo que ilustra una operación de ejemplo de generación de imágenes de referencia entre capas que utilizan una tabla de consulta 3D basada en la escalabilidad de la gama de colores.

DESCRIPCIÓN DETALLADA

[0013] Esta divulgación describe técnicas para una tabla de consulta tridimensional (3D) basada en la escalabilidad de la gama de colores en la codificación de vídeo multicapa. La codificación de vídeo multicapa puede estar de acuerdo con la norma de codificación de vídeo de alta eficiencia (HEVC), incluyendo cualquier extensión de codificación de vídeo escalable, una extensión de codificación de vídeo multivista, una extensión de codificación de vídeo 3D (es decir, codificación de vídeo multivista más profundidad) u otras extensiones de codificación de vídeo multicapa para HEVC. Los codificadores de vídeo y/o los decodificadores de vídeo pueden usar técnicas para generar imágenes de referencia entre capas cuando una gama de colores para una capa inferior de datos de vídeo sea diferente de una gama de colores para una capa superior de datos de vídeo.

[0014] Una gama de colores comprende una gama completa de colores que se pueden reproducir para una imagen, por ejemplo, en una imagen, fragmento, bloque o capa de datos de vídeo. Convencionalmente, en la codificación de vídeo multicapa, una capa inferior de datos de vídeo (por ejemplo, una capa base) y una capa superior de los datos de vídeo (por ejemplo, una capa de mejora) incluyen datos de color en la misma gama de colores, por ejemplo, gama de colores BT.709 de alta definición (HD). En este caso, un codificador de vídeo y/o un decodificador de vídeo pueden generar imágenes de referencia entre capas para la capa superior de los datos de vídeo como versiones muestreadas en sentido ascendente de imágenes de referencia cúbicas para la capa inferior de los datos de vídeo.

[0015] En algunos ejemplos, sin embargo, una capa inferior de los datos de vídeo puede incluir datos de color en una primera gama de colores, por ejemplo, BT.709, y una capa superior de los datos de vídeo pueden incluir datos de color en una segunda gama de colores diferente, por ejemplo, una gama de colores BT.2020 de ultraalta definición (UHD). En este ejemplo, a fin de generar imágenes de referencia entre capas para la capa superior de los datos de vídeo, un codificador de vídeo y/o un decodificador de vídeo primero debe realizar una predicción de color para convertir los datos de color de una imagen de referencia en la primera gama de colores para la capa inferior de los datos de vídeo a la segunda gama de colores para la capa superior de los datos de vídeo.

[0016] El codificador de vídeo y/o el decodificador de vídeo puede realizar la predicción de color usando una tabla de consulta 3D para la escalabilidad de la gama de colores. En algunos ejemplos, se puede generar una tabla de consulta 3D independiente para cada uno de los componentes de color, es decir, un componente de luma (Y), un primer componente de croma (U) y un segundo componente de croma (V). Cada una de las tablas de consulta 3D incluye una dimensión de luma (Y), una primera dimensión de croma (U) y una segunda dimensión de croma (V), y se indexa usando los tres componentes de color independientes (Y, U, V).

[0017] Convencionalmente, las tablas de consulta 3D siempre son simétricas, de manera que las tablas de consulta 3D tienen el mismo tamaño para el componente de luma, el primer componente de croma y el segundo componente de croma. Además, convencionalmente, las tablas de consulta 3D siempre están equilibradas de manera que un tamaño de cada dimensión de las tablas de consulta 3D es siempre el mismo. Esto da como resultado tamaños de tablas grandes con alta complejidad de cálculo y altos costes de señalización. Por ejemplo, los tamaños de tablas pueden ser de hasta $9 \times 9 \times 9$ o $17 \times 17 \times 17$.

[0018] De acuerdo con las técnicas descritas en esta divulgación, el codificador y/o decodificador de vídeo genera al menos una tabla de consulta 3D que tiene un tamaño que es diferente para el componente de luma que para cada uno del primer componente de croma y el segundo componente de croma. El codificador de vídeo y/o el decodificador de vídeo pueden generar esta tabla de consulta 3D asimétrica usando un número diferente de segmentos para la dimensión de luma de la tabla de consulta 3D. Por ejemplo, el codificador de vídeo y/o el decodificador de vídeo pueden generar la tabla de consulta 3D para que tenga un tamaño mayor, es decir, más segmentos para el componente de luma que para cada uno de los primero y segundo componentes de croma. En un ejemplo, la tabla de consulta 3D puede tener un tamaño de hasta $8 \times 2 \times 2$. De esta manera, el tamaño total de la tabla puede reducirse mientras se mantiene un buen rendimiento de la codificación con una resolución más alta para el componente de luma que para los primero y segundo componentes de croma.

[0019] En algunos ejemplos, cada una de las tablas de consulta 3D puede tener el mismo tamaño $M \times N \times K$, en el que el tamaño (M) de la dimensión de luma de la tabla de consulta 3D es diferente de cada tamaño (N) de la primera dimensión de croma de la tabla de consulta 3D y el tamaño (K) de la segunda dimensión de croma de la tabla de consulta 3D. Por ejemplo, el tamaño de la dimensión de luma puede ser mayor que los tamaños de las dimensiones de croma, $M > N$ y $M > K$, y los tamaños de las dimensiones de croma de la tabla de consulta 3D pueden ser los mismos, $N=K$. En un ejemplo, cada una de las tablas de consulta 3D puede tener el mismo tamaño de $8 \times 2 \times 2$.

[0020] En otros ejemplos, una tabla de consulta 3D de componentes de luma (LUT_Y) puede tener un primer tamaño, por ejemplo, $M \times M \times M$ o $M \times N \times K$, que es diferente de un segundo tamaño de una primera tabla de consulta 3D de componentes de croma (LUT_U), por ejemplo, $N \times N \times N$, y un tercer tamaño de una segunda tabla de consulta 3D de componentes de croma (LUT_V), por ejemplo, $K \times K \times K$. Por ejemplo, el tamaño de la tabla de consulta 3D del componente de luma puede ser mayor que los tamaños de las tablas de consulta 3D del componente de croma. Las tablas de consulta 3D del primer y segundo componente de croma pueden ser del mismo tamaño o de tamaños diferentes. En un ejemplo, la LUT_Y puede tener el tamaño de $8 \times 2 \times 2$, y cada una de la LUT_U y la LUT_V puede tener el tamaño de $2 \times 2 \times 2$.

[0021] En otro ejemplo, la precisión de la tabla de consulta 3D del componente de luma también puede ser diferente de la precisión de cada una de la primera tabla de consulta 3D del componente de croma y la segunda tabla de consulta 3D del componente de croma. Por ejemplo, la precisión de la tabla de consulta 3D del componente de luma puede ser mayor que la precisión de las tablas de consulta 3D del componente de croma. En algunos casos, para reducir aún más la complejidad, el codificador de vídeo y/o el decodificador de vídeo solo pueden generar una tabla de consulta 3D del componente de luma, realizar la predicción del componente de luma usando la tabla de consulta 3D del componente de luma y realizar la primera y segunda predicción del componente de croma usando una transformación lineal unidimensional (1D) o una transformación lineal por tramos.

[0022] Para mencionar otro ejemplo, de acuerdo con las técnicas de esta descripción, el codificador de vídeo y/o decodificador de vídeo puede generar una tabla de consulta 3D desequilibrada utilizando más segmentos para la dimensión de la tabla de consulta 3D asociada con el componente de color utilizado como índice de tabla para la tabla de consulta 3D. Por ejemplo, la tabla de consulta 3D del componente de luma puede tener una dimensión de luma más grande que cada una de una primera dimensión de croma y una segunda dimensión de croma basándose en el componente de luma que se utiliza como un índice de tabla para la tabla de consulta 3D del componente de luma. En este ejemplo, la LUT_Y puede tener el tamaño $M \times N \times N$, donde $M > N$. Los tamaños de las tablas de consulta 3D del primer y segundo componente de croma se pueden determinar de manera similar. Por ejemplo, la LUT_U puede tener el tamaño $N \times M \times N$ basándose en el primer componente de croma que se utiliza en el índice de tabla, y la LUT_V puede tener el tamaño $N \times N \times M$ basándose en el segundo componente de croma que se utiliza en el índice de tabla. De esta manera, el tamaño total de cada una de las tablas puede reducirse mientras se mantiene un buen rendimiento de la codificación con una resolución más alta para el componente de color utilizado como el índice de tabla.

[0023] Entre las normas de codificación de vídeo se incluyen ITU-T H.261, ISO/IEC MPEG-1 Visual, ITU-T H.262 o ISO/IEC MPEG-2 Visual, ITU-T H.263, ISO/IEC MPEG-4 Visual e ITU-T H.264 (también conocida como ISO/IEC MPBEG-4 AVC), incluidas sus extensiones de codificación de vídeo escalable (SVC) y de codificación de vídeo multivista (MVC).

[0024] Ha finalizado el diseño de una nueva norma de codificación de vídeo, concretamente la HEVC, por parte el Equipo de Colaboración Conjunta en Codificación de Vídeo (JCT-VC) del Grupo de Expertos en Codificación

de Vídeo (VCEG) de ITU-T y el Grupo de Expertos en Imágenes en Movimiento (MPEG) de ISO/IEC. Un borrador de memoria descriptiva de la norma HEVC denominado Borrador de trabajo 10 (WD10) de la HEVC, de Bross y col., "High efficiency video coding (HEVC) text specification draft 10 (for FDIS & Last Call)", Equipo de Colaboración Conjunta en Codificación de Vídeo (JCT-VC) de ITU-T SG16 WP3 e ISO/IEC JTC1/SC29/WG11, 12ª conferencia: Ginebra, CH. 14 a 23 de enero de 2013, JCTVC-L1003v34, está disponible en http://phenix.int-evry.fr/jct/doc_end_user/documents/12_Geneva/wg11/JCTVC-L1003-v34.zip. La norma HEVC finalizada se denomina versión 1 de HEVC.

[0025] Un informe por defecto, Wang y col., "High efficiency video coding (HEVC) Defect Report", Equipo de Colaboración Conjunta en Codificación de Vídeo (JCT-VC) de ITU-T SG16 WP3 e ISO/IEC JTC1/SC29/WG11. 14ª reunión: Viena, AT, 25 de julio al 2 de agosto de 2013, JCTVC-N1003v1, está disponible en http://phenix.int-evry.fr/jct/doc_end_user/documents/14_Vienna/wg11/JCTVC-n1003-VL.zip. El documento estándar finalizado se publica como ITU-T H.265, Serie H: Sistemas audiovisuales y de multimedia, Infraestructura de servicios audiovisuales - Codificación de vídeo en movimiento, Codificación de vídeo de alta eficiencia, Sector de normalización de las telecomunicaciones de la Unión Internacional de Telecomunicaciones (ITU), abril de 2013.

[0026] La extensión multivista a la HEVC (MV-HEVC) y otra extensión de la HEVC para la codificación de vídeo 3D más avanzada (3D-HEVC) se están desarrollando por el JCT-3V. Un borrador de memoria descriptiva de MV-HEVC, denominado Borrador de Trabajo 5 (WD5) de MV-HEVC, Tech y col., "MV-HEVC Draft Text 5", Equipo de Colaboración Conjunta en el Desarrollo de Extensión de Codificación de Vídeo 3D (JCT-3V) del ITU-T SG16 WP3 e ISO/IEC JTC1/SC29/WG11, 5ª conferencia: Viena, AT, 27 de julio-2 de agosto de 2013, JCT3V-E1004v6, está disponible en http://phenix.int-evry.fr/jct/doc_end_user/documents/5_Vienna/wg11/JCT3V-E1004-v6.zip. Un borrador de memoria descriptiva de la norma 3D-HEVC, denominado Borrador de Trabajo 1 (WD1) de la norma 3D-HEVC y descrito en el artículo de Tech y col., "3D-HEVC Draft Text 1", Equipo de Colaboración Conjunta en el Desarrollo de Extensión de Codificación de Vídeo 3D (JCT-3V) del ITU-T SG 16 WP 3 e ISO/IEC JTC 1/SC 29/WG 11, 5ª conferencia: Viena, AT, 27 de julio a 2 de agosto de 2013, JCT3V-E1001v3, está disponible en http://phenix.it-sudparis.eu/jct2/doc_end_user/documents/5_Vienna/wg11/JCT3V-E1001-v3.zip.

[0027] La extensión escalable a HEVC (SHVC) también se está desarrollando por el JCT-VC. Un borrador de memoria descriptiva de la SHVC, denominada Borrador de Trabajo 3 (WD3) de la SHVC, Chen y col., "SHVC Draft 3", Equipo de Colaboración Conjunta en la Codificación de Vídeo (JCT-VC) del ITU-T SG16 WP3 e ISO/IEC JTC1/SC29/WG11, 14ª conferencia: Viena, AT, 25 de julio-2 de agosto de 2013, JCTVC-N1008v3, está disponible en http://phenix.int-evry.fr/jct/doc_end_user/documents/14_Vienna/wg11/JCTVC-N1008-v3.zip.

[0028] La FIG. 1 es un diagrama de bloques que ilustra un sistema de codificación y de decodificación de vídeo de ejemplo 10 que puede utilizar técnicas para una tabla de consulta 3D basada en la escalabilidad de la gama de colores. Como se muestra en la FIG. 1, el sistema 10 incluye un dispositivo de origen 12 que proporciona datos de vídeo codificados, a decodificar en un momento posterior por parte de un dispositivo de destino 14. En particular, el dispositivo de origen 12 proporciona los datos de vídeo al dispositivo de destino 14 mediante un medio legible por ordenador 16. El dispositivo de origen 12 y el dispositivo de destino 14 pueden comprender cualquiera entre una amplia gama de dispositivos, incluyendo ordenadores de mesa, ordenadores plegables (es decir, portátiles), tabletas electrónicas, cajas de conexión, aparato telefónico con microteléfono tales como los denominados teléfonos "inteligentes", los denominados paneles "inteligentes", televisores, cámaras, dispositivos de visualización, reproductores de medios digitales, consolas de videojuegos, un dispositivo de emisión de vídeo en tiempo real o similares. En algunos casos, el dispositivo de origen 12 y el dispositivo de destino 14 pueden estar equipados para la comunicación inalámbrica.

[0029] El dispositivo de destino 14 puede recibir los datos de vídeo codificados a decodificar mediante el medio legible por ordenador 16. El medio legible por ordenador 16 puede comprender cualquier tipo de medio o dispositivo capaz de transportar los datos de vídeo codificados desde el dispositivo de origen 12 al dispositivo de destino 14. En un ejemplo, el medio legible por ordenador 16 puede comprender un medio de comunicación para habilitar el dispositivo de origen 12 para transmitir datos de vídeo codificados directamente al dispositivo de destino 14 en tiempo real. Los datos de vídeo codificados se pueden modular de acuerdo con una norma de comunicación, tal como un protocolo de comunicación inalámbrica, y transmitirse al dispositivo de destino 14. El medio de comunicación puede comprender cualquier medio de comunicación, inalámbrico o cableado, tal como un espectro de radiofrecuencia (RF) o una o más líneas físicas de transmisión. El medio de comunicación puede formar parte de una red basada en paquetes, tal como una red de área local, una red de área amplia o una red global tal como Internet. El medio de comunicación puede incluir encaminadores, conmutadores, estaciones base o cualquier otro equipo que pueda ser útil para facilitar la comunicación desde el dispositivo de origen 12 hasta el dispositivo de destino 14.

[0030] En algunos ejemplos, pueden emitirse datos codificados desde la interfaz de salida 22 hasta un dispositivo de almacenamiento. De forma similar, se puede acceder a los datos codificados desde el dispositivo de almacenamiento mediante una interfaz de entrada. El dispositivo de almacenamiento puede incluir cualquiera entre varios medios de almacenamiento de datos, distribuidos o de acceso local, tales como un disco duro, discos Blu-ray, DVD, CD-ROM, memoria flash, memoria volátil o no volátil, o cualquier otro medio de

almacenamiento digital adecuado para almacenar datos de vídeo codificados. En un ejemplo adicional, el dispositivo de almacenamiento puede corresponder a un servidor de ficheros o a otro dispositivo de almacenamiento intermedio que pueda almacenar el vídeo codificado generado por el dispositivo de origen 12. El dispositivo de destino 14 puede acceder a datos de vídeo almacenados desde el dispositivo de almacenamiento, mediante transmisión continua o descarga. El servidor de archivos puede ser cualquier tipo de servidor capaz de almacenar datos de vídeo codificados y transmitir esos datos de vídeo codificados al dispositivo de destino 14. Los servidores de archivos de ejemplo incluyen un servidor web (por ejemplo, para un sitio web), un servidor de FTP, dispositivos de almacenamiento conectados en red (NAS) o una unidad de disco local. El dispositivo de destino 14 puede acceder a los datos de vídeo codificados a través de cualquier conexión de datos estándar, incluida una conexión a Internet. Esto puede incluir un canal inalámbrico (por ejemplo, una conexión de Wi-Fi), una conexión cableada (por ejemplo, DSL, módem de cable, etc.) o una combinación de ambos que sea adecuada para acceder a datos de vídeo codificados almacenados en un servidor de archivos. La transmisión de datos de vídeo codificados desde el dispositivo de almacenamiento puede ser una transmisión continua, una transmisión de descarga o una combinación de ambas.

[0031] Las técnicas de esta divulgación no están limitadas necesariamente a aplicaciones o configuraciones inalámbricas. Las técnicas se pueden aplicar a la codificación de vídeo como soporte de cualquiera entre varias aplicaciones de multimedia, tales como difusiones de televisión por el aire, transmisiones de televisión por cable, transmisiones de televisión por satélite, transmisiones continuas de vídeo por Internet, tales como la transmisión continua adaptativa dinámica por HTTP (DASH), vídeo digital que se codifica en un medio de almacenamiento de datos, decodificación de vídeo digital almacenado en un medio de almacenamiento de datos u otras aplicaciones. En algunos ejemplos, el sistema 10 puede estar configurado para prestar soporte a una transmisión de vídeo unidireccional o bidireccional, a fin de prestar soporte a aplicaciones tales como la transmisión continua de vídeo, la reproducción de vídeo, la radiodifusión de vídeo y/o la videotelefonía.

[0032] En el ejemplo de la FIG. 1, el dispositivo de origen 12 incluye una fuente de vídeo 18, un codificador de vídeo 20 y una interfaz de salida 22. El dispositivo de destino 14 incluye una interfaz de entrada 28, un decodificador de vídeo 30 y un dispositivo de visualización 32. De acuerdo con esta divulgación, el codificador de vídeo 20 del dispositivo de origen 12 puede estar configurado para aplicar las técnicas para el procesamiento de datos de vídeo en paralelo. En otros ejemplos, un dispositivo de origen y un dispositivo de destino pueden incluir otros componentes o disposiciones. Por ejemplo, el dispositivo de origen 12 puede recibir datos de vídeo desde una fuente de vídeo externa 18 como una cámara externa. Del mismo modo, el dispositivo de destino 14 puede interactuar con un dispositivo de visualización externo, en lugar de incluir un dispositivo de visualización integrado.

[0033] El sistema ilustrado 10 de la FIG. 1 es simplemente un ejemplo. Las técnicas para el procesamiento de datos de vídeo en paralelo pueden realizarse mediante cualquier dispositivo de codificación y/o de decodificación de vídeo digital. Aunque, por lo general, las técnicas de esta divulgación se llevan a cabo mediante un dispositivo de codificación de vídeo, las técnicas también pueden llevarse a cabo mediante un codificador/decodificador de vídeo, denominado típicamente "CÓDEC". Además, las técnicas de esta divulgación también se pueden llevar a cabo mediante un preprocesador de vídeo. El dispositivo de origen 12 y el dispositivo de destino 14 son simplemente ejemplos de dichos dispositivos de codificación en los que el dispositivo de origen 12 genera datos de vídeo codificados para su transmisión al dispositivo de destino 14. En algunos ejemplos, los dispositivos 12, 14 pueden funcionar de manera sustancialmente simétrica, de modo que cada uno de los dispositivos 12, 14 incluya componentes de codificación y decodificación de vídeo. De ahí que el sistema 10 pueda prestar soporte a la transmisión de vídeo unidireccional o bidireccional entre los dispositivos de vídeo 12, 14, por ejemplo, para la transmisión continua de vídeo, la reproducción de vídeo, la radiodifusión de vídeo o la videotelefonía.

[0034] La fuente de vídeo 18 del dispositivo de origen 12 puede incluir un dispositivo de captura de vídeo, tal como una cámara de vídeo, un archivo de vídeo que contiene vídeo capturado previamente y/o una interfaz de suministro de vídeo para recibir vídeo desde un proveedor de contenidos de vídeo. Como otra alternativa, la fuente de vídeo 18 puede generar datos basados en gráficos de ordenador como la fuente de vídeo, o una combinación de vídeo en directo, vídeo archivado y vídeo generado por ordenador. En algunos casos, si la fuente de vídeo 18 es una cámara de vídeo, el dispositivo de origen 12 y el dispositivo de destino 14 pueden formar los llamados teléfonos con cámara o videoteléfonos. Sin embargo, como se ha mencionado anteriormente, las técnicas descritas en esta divulgación pueden ser aplicables a la codificación de vídeo en general, y se pueden aplicar a aplicaciones inalámbricas y/o cableadas. En cada caso, el codificador de vídeo 20 puede codificar el vídeo capturado, precapturado o generado por ordenador. Después, la información de vídeo codificada puede ser proporcionada por la interfaz de salida 22 a un medio legible por ordenador 16.

[0035] El medio legible por ordenador 16 puede incluir medios transitorios, tales como una radiodifusión inalámbrica o una transmisión de red cableada, o medios de almacenamiento (es decir, medios de almacenamiento no transitorios), tales como un disco duro, una unidad de memoria flash, un disco compacto, un disco de vídeo digital, un disco Blu-ray u otros medios legibles por ordenador. En algunos ejemplos, un servidor de red (no se muestra) puede recibir datos de vídeo codificados desde el dispositivo de origen 12 y proporcionar los datos de vídeo codificados al dispositivo de destino 14, por ejemplo, mediante transmisión por red. De forma

similar, un dispositivo informático de una utilidad de producción de un medio, tal como una utilidad de grabación de discos, puede recibir datos de vídeo codificados desde el dispositivo de origen 12 y producir un disco que contiene los datos de vídeo codificados. Por lo tanto, se puede entender que el medio legible por ordenador 16 incluye uno o más medios legibles por ordenador de diversas formas, en varios ejemplos.

[0036] La interfaz de entrada 28 del dispositivo de destino 14 recibe información desde el medio legible por ordenador 16. La información del medio legible por ordenador 16 puede incluir información de sintaxis definida por el codificador de vídeo 20, que también se usa por el decodificador de vídeo 30, que incluye elementos de sintaxis que describen características y/o procesamiento de bloques y otras unidades codificadas, por ejemplo, grupos de imágenes (GOP). El dispositivo de visualización 32 muestra los datos de vídeo decodificados a un usuario, y puede comprender cualquiera entre varios dispositivos de visualización, tales como un tubo de rayos catódicos (CRT), una pantalla de cristal líquido (LCD), una pantalla de plasma, una pantalla de diodos orgánicos emisores de luz (OLED) u otro tipo de dispositivo de visualización.

[0037] El codificador de vídeo 20 y el decodificador de vídeo 30 pueden implementarse como cualquiera entre una variedad de circuitos de codificadores adecuados, tales como uno o más microprocesadores, procesadores de señales digitales (DSP), circuitos integrados de aplicación específica (ASIC), matrices de puertas programables in situ (FPGA), lógica discreta, software, hardware, firmware o cualquier combinación de estos. Cuando las técnicas se implementan parcialmente en software, un dispositivo puede almacenar instrucciones para el software en unos medios legibles por ordenador no transitorios adecuados, y ejecutar las instrucciones en hardware mediante uno o más procesadores que realizan las técnicas de esta divulgación. Tanto el codificador de vídeo 20 como el decodificador de vídeo 30 pueden estar incluidos en uno o más codificadores o decodificadores, donde cualquiera de ambos puede estar integrado como parte de un codificador/decodificador (CÓDEC) combinado en un dispositivo respectivo.

[0038] En algunos ejemplos, el codificador de vídeo 20 y el decodificador de vídeo 30 funcionan de acuerdo con una norma de compresión de vídeo, tal como la ISO/IEC MPEG-4 Visual y la ITU-T H.264 (también conocida como ISO/IEC MPEG-4 AVC), incluida su extensión de codificación de vídeo escalable (SVC), su extensión de codificación de vídeo multivista (MVC) y su extensión de vídeo tridimensional (3DV) basada en MVC. En algunos casos, cualquier flujo de bits que se ajuste a 3DV basada en MVC siempre contiene un subflujo de bits que es compatible con un perfil MVC, por ejemplo, un alto perfil estéreo. Además, se están emprendiendo unas iniciativas para generar una extensión de codificación 3DV para H.264/AVC, concretamente, 3DV basada en AVC. En otros ejemplos, el codificador de vídeo 20 y el decodificador de vídeo 30 pueden funcionar de acuerdo con la ITU-T H.261, ISO/IEC MPEG-1 Visual, ITU-T H.262 o la ISO/IEC MPEG-2 Visual y la ITU-T H.264, ISO/IEC Visual.

[0039] En el ejemplo de la FIG. 1, el codificador de vídeo 20 y el decodificador de vídeo 30 pueden funcionar de acuerdo con la norma de Codificación de Vídeo de Alta Eficiencia (HEVC) elaborada por el Equipo de Colaboración Conjunta en Codificación de Vídeo (JCT-VC) del Grupo de Expertos en Codificación de Vídeo (VCEG) de ITU-T y del Grupo de Expertos en Imágenes en Movimiento (MPEG) de ISO/IEC. El borrador de memoria descriptiva de la HEVC, mencionada anteriormente, se denomina Borrador de Trabajo 10 (WD10) de HEVC, y la versión finalizada de la norma de HEVC se denomina versión 1 de HEVC. El MV-HEVC y 3D-HEVC se están desarrollando por el JCT-3V. Un borrador de memoria descriptiva reciente de la MV-HEVC se denomina MV-HEVC WD5, y un borrador de memoria descriptiva reciente de la 3D-HEVC se denomina 3D-HEVC WD1. La SHVC se está desarrollando por el JCT-VC. Un borrador de memoria descriptiva reciente de la SHVC se denomina SHVC WD3.

[0040] En la HEVC y en otras normas de codificación de vídeo, una secuencia de vídeo incluye típicamente una serie de imágenes. Las imágenes también pueden denominarse "tramas". Una imagen puede incluir tres matrices de muestras, denominadas S_L , S_{Cb} y S_{Cr} . S_L es una matriz bidimensional (es decir, un bloque) de muestras de luma. S_{Cb} es una matriz bidimensional de muestras de crominancia Cb. S_{Cr} es una matriz bidimensional de muestras de crominancia Cr. Las muestras de crominancia también se pueden denominar en el presente documento muestras de "croma". En otros casos, una imagen puede ser monocromática y puede solo incluir una matriz de muestras de luma.

[0041] El codificador de vídeo 20 puede generar un conjunto de unidades de árbol de codificación (CTU). Cada una de las CTU puede comprender un bloque de árbol de codificación de muestras de luma, dos bloques correspondientes de árbol de codificación de muestras de croma y estructuras de sintaxis usadas para codificar las muestras de los bloques de árbol de codificación. En una imagen monocromática o en una imagen que tenga tres planos de color separados, una CTU puede comprender un solo bloque de árbol de codificación y estructuras de sintaxis usadas para codificar las muestras del bloque de árbol de codificación. Un bloque de árbol de codificación puede ser un bloque de muestras de tamaño $N \times N$. Una CTU también puede denominarse "bloque de árbol" o "unidad de codificación de máximo tamaño" (LCU). Las CTU de la HEVC pueden ser ampliamente análogas a los macrobloques de otras normas de codificación de vídeo, tales como H.264/AVC. Sin embargo, una CTU no está limitada necesariamente a un tamaño particular y puede incluir una o más unidades de

codificación (CU). Un fragmento puede incluir un número entero de CTU ordenadas consecutivamente en el barrido de trama.

- 5 **[0042]** Esta divulgación puede usar el término "unidad de vídeo", "bloque de vídeo" o "bloque" para referirse a uno o más bloques de muestras y estructuras de sintaxis usadas para codificar muestras de los uno o más bloques de muestras. Los tipos de ejemplo de unidades de vídeo pueden incluir CTU, CU, PU, unidades de transformada (TU) en HEVC, O macrobloques, particiones de macrobloque, y así sucesivamente en otras normas de codificación de vídeo.
- 10 **[0043]** Para generar una CTU codificada, el codificador de vídeo 20 puede realizar de forma recursiva una partición de árbol cuaternario en los bloques de árbol de codificación de una CTU para dividir los bloques de árbol de codificación en bloques de codificación, de ahí el nombre "unidades de árbol de codificación". Un bloque de codificación es un bloque de muestras de tamaño NxN. Una CU puede comprender un bloque de codificación de muestras de luma y dos bloques de codificación correspondientes de muestras de croma de una imagen que
15 tenga una matriz de muestras de luma, una matriz de muestras de Cb y una matriz de muestras de Cr, y estructuras de sintaxis usadas para codificar las muestras de los bloques de codificación. En una imagen monocromática o en una imagen que tenga tres planos de color separados, una CU puede comprender un solo bloque de codificación y estructuras de sintaxis usadas para codificar las muestras del bloque de codificación.
- 20 **[0044]** El codificador de vídeo 20 puede dividir un bloque de codificación de una CU en uno o más bloques de predicción. Un bloque de predicción puede ser un bloque rectangular (es decir, cuadrado o no cuadrado) de muestras en las que se aplica la misma predicción. Una unidad de predicción (PU) de una CU puede ser un bloque de predicción de muestras de luma, dos bloques correspondientes de predicción de muestras de croma de una imagen y estructuras de sintaxis usadas para predecir las muestras de bloques de predicción. En una
25 imagen monocromática o una imagen que tenga tres planos de color separados, una PU puede comprender un solo bloque de predicción y estructuras de sintaxis usadas para predecir las muestras de bloque de predicción. El codificador de vídeo 20 puede generar bloques de predicción de luma, Cb y Cr para bloques de predicción de luma, Cb y Cr de cada PU de la CU.
- 30 **[0045]** El codificador de vídeo 20 puede usar intrapredicción o interpredicción para generar los bloques predictivos para una PU. Si el codificador de vídeo 20 usa la intrapredicción para generar los bloques predictivos de una PU, el codificador de vídeo 20 puede generar los bloques predictivos de la PU basándose en muestras decodificadas de la imagen asociada a la PU.
- 35 **[0046]** Si el codificador de vídeo 20 usa la interpredicción para generar los bloques predictivos de una PU, el codificador de vídeo 20 puede generar los bloques predictivos de la PU basándose en muestras decodificadas de una o más imágenes distintas a la imagen asociada a la PU. La interpredicción puede ser una interpredicción unidireccional (es decir, unipredicción) o una interpredicción bidireccional (es decir, bipredicción). Para realizar la unipredicción o la bipredicción, el codificador de vídeo 20 puede generar una primera lista de imágenes de
40 referencia (RefPicList0) y una segunda lista de imágenes de referencia (RefPicList1) para un fragmento actual.
- [0047]** Cada una de las listas de imágenes de referencia puede incluir una o más imágenes de referencia. Cuando se utiliza unipredicción, el codificador de vídeo 20 puede buscar las imágenes de referencia en RefPicList0, RefPicList1 o en ambas para determinar una ubicación de referencia dentro de una imagen de
45 referencia. Además, cuando se usa la unipredicción, el codificador de vídeo 20 puede generar, basándose al menos en parte en las muestras correspondientes a la ubicación de referencia, los bloques predictivos de muestras para la PU. Además, cuando se usa la unipredicción, el codificador de vídeo 20 puede generar un único vector de movimiento que indique un desplazamiento espacial entre un bloque de predicción de la PU y la ubicación de referencia. Para indicar el desplazamiento espacial entre un bloque de predicción de la PU y la
50 ubicación de referencia, un vector de movimiento puede incluir un componente horizontal que especifique un desplazamiento horizontal entre el bloque de predicción de la PU y la ubicación de referencia y puede incluir un componente vertical que especifique un desplazamiento vertical entre el bloque de predicción de la PU y la ubicación de referencia.
- 55 **[0048]** Cuando se utiliza la predicción doble para codificar un bloque, el codificador de vídeo 20 puede determinar una primera ubicación de referencia en una imagen de referencia en RefPicList0 y una segunda ubicación de referencia en una imagen de referencia en RefPicList1. El codificador de vídeo 20 puede generar entonces, basándose al menos en parte en muestras correspondientes a las primera y segunda ubicaciones de
60 referencia, los bloques predictivos para la PU. Además, cuando se usa la bipredicción para codificar la PU, el codificador de vídeo 20 puede generar un primer movimiento que indique un desplazamiento espacial entre un bloque de muestra de la PU y la primera ubicación de referencia y un segundo movimiento que indique un desplazamiento espacial entre el bloque de predicción de la PU y la segunda ubicación de referencia.
- 65 **[0049]** Después de que el codificador de vídeo 20 genera bloques de predicción de luma, Cb y Cr para una o más PU de una CU, el codificador de vídeo 20 puede generar un bloque residual de luma para la CU. Cada muestra en el bloque residual de luma de la CU indica una diferencia entre una muestra de luma en uno de los

bloques de predicción de luma de la CU y una muestra correspondiente en el bloque de codificación de luma original de la CU. Además, el codificador de vídeo 20 puede generar un bloque residual de Cb para la CU. Cada muestra en el bloque residual de Cb de la CU puede indicar una diferencia entre una muestra Cb en uno de los bloques Cb predictivos de la CU y una muestra correspondiente en el bloque de codificación Cb original de la CU. El codificador de vídeo 20 también puede generar un bloque residual de Cr para la CU. Cada muestra en el bloque residual de Cr de la CU puede indicar una diferencia entre una muestra de Cr en uno de los bloques predictivos de Cr de la CU y una muestra correspondiente en el bloque de codificación de Cr original de la CU.

[0050] Además, el codificador de vídeo 20 puede usar la partición en árbol cuádruple para descomponer los bloques residuales de luma, Cb y Cr de una CU en uno o más bloques de transformada de luma, Cb y Cr. Un bloque de transformada puede ser un bloque rectangular de muestras en el que se aplique la misma transformada. Una unidad de transformada (TU) de una CU puede comprender un bloque de transformada de muestras de luma, dos bloques de transformada correspondientes de muestras de croma y estructuras de sintaxis usadas para transformar las muestras de bloques de transformada. En imágenes monocromáticas o imágenes que tengan tres planos de color independientes, una PU puede comprender un único bloque de transformada y estructuras de sintaxis usadas para transformar las muestras del bloque de transformada. De este modo, cada TU de una CU puede estar asociada a un bloque de transformada de luma, un bloque de transformada de Cb y un bloque de transformada de Cr. El bloque de transformada de luma asociado con la TU puede ser un subbloque del bloque residual de luma de la CU. El bloque de transformada de Cb puede ser un subbloque del bloque residual de Cb de la CU. El bloque de transformada de Cr puede ser un subbloque del bloque residual de Cr de la CU.

[0051] El codificador de vídeo 20 puede aplicar una o más transformadas a un bloque de transformada de luma de una TU con el fin de generar un bloque de coeficientes de luma para la TU. Un bloque de coeficientes puede ser una matriz bidimensional de coeficientes de transformada. Un coeficiente de transformada puede ser una cantidad escalar. El codificador de vídeo 20 puede aplicar una o más transformadas a un bloque de transformada de Cb de una TU para generar un bloque de coeficientes de Cb para la TU. El codificador de vídeo 20 puede aplicar una o más transformadas a un bloque de transformada de Cr de una TU para generar un bloque de coeficientes de Cr para la TU.

[0052] Después de generar un bloque de coeficientes (por ejemplo, un bloque de coeficientes de luma, un bloque de coeficientes de Cb o un bloque de coeficientes de Cr), el codificador de vídeo 20 puede cuantificar el bloque de coeficientes. La cuantificación se refiere, en general, a un proceso en el que los coeficientes de transformada se cuantifican para reducir, posiblemente, la cantidad de datos usados para representar los coeficientes de transformada, proporcionando una compresión adicional. Además, el codificador de vídeo 20 puede cuantificar inversamente los coeficientes de transformada y aplicar una transformada inversa a los coeficientes de transformada para reconstruir bloques de transformada de TU de las CU de una imagen. La unidad de reconstrucción 20 puede usar los valores residuales de los bloques de transformada de las TU de una CU y los bloques predictivos de las PU de la CU para reconstruir los bloques de codificación de la CU. Mediante la reconstrucción de los bloques de codificación para cada CU de una imagen, el decodificador de vídeo 20 puede reconstruir la imagen. El codificador de vídeo 20 puede almacenar imágenes reconstruidas en una memoria intermedia de imágenes decodificadas (DPB). El codificador de vídeo 20 puede usar imágenes reconstruidas en la DPB para la interpredicción y la intrapredicción.

[0053] Después de que el codificador de vídeo 20 cuantifique un bloque de coeficientes, el codificador de vídeo 20 puede codificar por entropía elementos de sintaxis que indican los coeficientes de transformada cuantificados. Por ejemplo, el codificador de vídeo 20 puede realizar la Codificación Aritmética Binaria Adaptativa al Contexto (CABAC) en los elementos de sintaxis que indican los coeficientes de transformada cuantificados. El codificador de vídeo 20 puede emitir los elementos de sintaxis codificados por entropía en un flujo de bits.

[0054] El codificador de vídeo 20 puede emitir un flujo de bits que incluye una secuencia de bits que forma una representación de imágenes codificadas y datos asociados. El flujo de bits puede comprender una secuencia de unidades de capa de abstracción de red (NAL). Cada una de las unidades de NAL incluye un encabezamiento de unidad de NAL y encapsula una carga útil de secuencia de octetos sin procesar (RBSP). El encabezamiento de la unidad de NAL puede incluir un elemento de sintaxis que indique un código de tipo de unidad de NAL. El código de tipo de unidad de NAL especificado por el encabezamiento de unidad de NAL de una unidad de NAL indica el tipo de la unidad de NAL. Una RBSP puede ser una estructura de sintaxis que contenga un número entero de octetos que se encapsule dentro de una unidad de NAL. En algunos casos, una RBSP incluye cero bits.

[0055] Diferentes tipos de unidades de NAL pueden encapsular diferentes tipos de RBSP. Por ejemplo, un primer tipo de unidad de NAL puede encapsular una RBSP para un conjunto de parámetros de imagen (PPS), un segundo tipo de unidad de NAL puede encapsular una RBSP para un fragmento codificado, un tercer tipo de unidad de NAL puede encapsular una RBSP para información de mejora complementaria (SEI), y así sucesivamente. Un PPS es una estructura de sintaxis que puede contener elementos de sintaxis que se aplican a cero o más imágenes codificadas completas. Las unidades de NAL que encapsulan las RBSP para datos de

codificación de vídeo (a diferencia de las RBSP para conjuntos de parámetros y mensajes de SEI) pueden denominarse unidades de NAL de la capa de codificación de vídeo (VCL). Una unidad NAL que encapsula un fragmento codificado puede denominarse en el presente documento unidad NAL de fragmento codificado. Una RBSP para un fragmento codificado puede incluir un encabezamiento de fragmento y datos de fragmento.

5

[0056] El decodificador de vídeo 30 puede recibir un flujo de bits. Además, el decodificador de vídeo 30 puede analizar el flujo de bits para decodificar elementos de sintaxis del flujo de bits. El decodificador de vídeo 30 puede reconstruir las imágenes de los datos de vídeo basándose, al menos en parte, en los elementos de sintaxis decodificados a partir del flujo de bits. El proceso para reconstruir los datos de vídeo puede ser, en general, recíproco al proceso realizado por el codificador de vídeo 20. Por ejemplo, el decodificador de vídeo 30 puede usar vectores de movimiento de las PU para determinar bloques predictivos para las PU de una CU actual. El decodificador de vídeo 30 puede usar un vector de movimiento o vectores de movimiento de las PU para generar bloques predictivos para las PU.

10

[0057] Además, el decodificador de vídeo 30 puede cuantificar inversamente bloques de coeficientes asociados a las TU de la CU actual. El decodificador de vídeo 30 puede realizar transformadas inversas en los bloques de coeficientes para reconstruir los bloques de transformada asociados a las TU de la CU actual. El decodificador de vídeo 30 puede reconstruir los bloques de codificación de la CU actual añadiendo las muestras de los bloques predictivos de muestras para las PU de la CU actual a las muestras correspondientes de los bloques de transformada de las TU de la CU actual. Mediante la reconstrucción de los bloques de codificación para cada CU de una imagen, el decodificador de vídeo 30 puede reconstruir la imagen. El decodificador de vídeo 30 puede almacenar imágenes decodificadas en una memoria intermedia de imágenes decodificadas, para su salida y/o para su uso en la decodificación de otras imágenes.

20

[0058] En MV-HEVC, 3D-HEVC y SHVC, un codificador de vídeo puede generar un flujo de bits multicapa que comprende una serie de unidades de capa de abstracción de red (NAL). Diferentes unidades de NAL del flujo de bits pueden estar asociadas a diferentes capas del flujo de bits. Una capa se puede definir como un conjunto de unidades de NAL de capa de codificación de vídeo (VCL) y unidades asociadas de NAL que no son de VCL que tengan el mismo identificador de capa. Una capa puede ser equivalente a una vista en la codificación de vídeo multivista. En la codificación de vídeo multivista, una capa puede contener todos los componentes de vista de la misma capa con diferentes instancias de tiempo. Cada componente de vista puede ser una imagen codificada de la escena de vídeo que pertenece a una vista específica en una instancia de tiempo específico. En la codificación de vídeo 3D, una capa puede contener todas las imágenes de profundidad codificadas de una vista específica o imágenes de texturas codificadas de una vista específica. De forma similar, en el contexto de la codificación de vídeo escalable, una capa corresponde típicamente a imágenes codificadas que tienen características de vídeo diferentes de las imágenes codificadas en otras capas. Dichas características de vídeo típicamente incluyen la resolución espacial y el nivel de calidad (relación señal-ruido). En la norma HEVC y en sus extensiones, la escalabilidad temporal se puede lograr dentro de una capa definiendo un grupo de imágenes con un nivel temporal particular como una subcapa.

25

30

35

40

[0059] Para cada capa respectiva del flujo de bits, los datos en una capa inferior pueden decodificarse sin referencia a unos datos en cualquier capa superior. En la codificación de vídeo escalable, por ejemplo, los datos en una capa base pueden decodificarse sin referencia a datos en una capa de mejora. Las unidades de NAL solo encapsulan datos de una única capa. En la SHVC, una vista puede denominarse "capa base" si un decodificador de vídeo puede decodificar imágenes en la vista sin referencia a imágenes en cualquier otra capa. La capa base puede ajustarse a la memoria descriptiva base de la HEVC. Por tanto, las unidades de NAL que encapsulan datos de la capa restante más alta del flujo de bits pueden eliminarse del flujo de bits sin afectar a la decodificabilidad de los datos en las capas restantes del flujo de bits. En las normas MV-HEVC y 3D-HEVC, las capas superiores pueden incluir componentes de vista adicionales. En SHVC, las capas superiores pueden incluir datos de mejora de relación señal-ruido (SNR), datos de mejora espacial y/o datos de mejora temporal.

45

50

[0060] En algunos ejemplos, los datos en una capa superior pueden decodificarse en referencia a los datos en una o más capas inferiores. Las capas inferiores se pueden usar como imágenes de referencia para comprimir la capa superior usando la predicción entre capas. Los datos de las capas inferiores se pueden muestrear en sentido ascendente para tener la misma resolución que las capas superiores. En general, el codificador de vídeo 20 y el decodificador de vídeo 30 pueden realizar una predicción entre capas de manera similar a la interpredicción descrita anteriormente, excepto que una o más capas inferiores muestreadas en sentido ascendente pueden usarse como imágenes de referencia en contraposición a una o más imágenes adyacentes.

55

[0061] La FIG. 2 es una ilustración conceptual que muestra un ejemplo de escalabilidad en tres dimensiones diferentes. En una estructura de codificación de vídeo escalable, las escalabilidades se habilitan en tres dimensiones. En el ejemplo de la FIG. 2, las escalabilidades están habilitadas en una dimensión espacial (S) 100, en una dimensión temporal (T) 102 y en una relación señal-ruido (SNR) o dimensión de calidad (Q) 104. En la dimensión temporal 102, las frecuencias de trama con 7,5 Hz (T0), 15 Hz (T1) o 30 Hz (T2), por ejemplo, pueden ser compatibles con la escalabilidad temporal. Cuando se admite la escalabilidad espacial, diferentes resoluciones como QCIF (S0), CIF (S1) y 4CIF (S2), por ejemplo, se habilitan en la dimensión espacial 100. Para

60

65

cada resolución espacial y frecuencia de tramas específicas, se pueden añadir capas SNR (Q1) en la dimensión SNR 104 para mejorar la calidad de imagen.

[0062] Una vez que el contenido de vídeo se ha codificado de dicha manera escalable, se puede usar una herramienta de extracción para adaptar el contenido real suministrado de acuerdo con los requisitos de la aplicación, que pueden depender, por ejemplo, de los clientes o del canal de transmisión. En el ejemplo que se muestra en la FIG. 2, cada cúbico contiene imágenes con la misma frecuencia de trama (nivel temporal), resolución espacial y capas SNR. Se puede lograr una mejor representación añadiendo cubos (es decir, imágenes) en cualquiera de las dimensiones 100, 102 o 104. Se admite la escalabilidad combinada cuando hay dos, tres o incluso más escalabilidades habilitadas.

[0063] En las normas de codificación de vídeo escalables, tales como la extensión SVC a H.264 o SHVC, las imágenes con la capa espacial y SNR más baja son compatibles con el códec de vídeo de una única capa, y las imágenes en el nivel temporal más bajo forman la capa base temporal, que se puede mejorar con imágenes en niveles temporales superiores. Además de la capa base, se pueden añadir varias capas de mejora espacial y/o de SNR para proporcionar escalabilidades espaciales y/o de calidad. Cada capa de mejora espacial o de SNR puede ser temporalmente escalable, con la misma estructura de escalabilidad temporal que la capa base. Para una capa de mejora espacial o de SNR, la capa inferior de la que depende también puede denominarse capa base de esa capa de mejora espacial o de SNR específica.

[0064] La FIG. 3 es una ilustración conceptual que muestra una estructura de ejemplo 110 de un flujo de bits de codificación de vídeo escalable. La estructura de flujo de bits 110 incluye una capa 0 112 que incluye imágenes o fragmentos 10, P4 y P8, y una capa 1 114 que incluye imágenes o fragmentos B2, B6 y B10. Además, la estructura de flujo de bits 110 incluye una capa 2 116 y una capa 3 117 que incluyen cada una imágenes 0, 2, 4, 6, 8 y 10, y una capa 4 118 que incluye imágenes 0 a 11.

[0065] Una capa base tiene la capa espacial y de calidad más baja (es decir, imágenes en la capa 0 112 y de la capa 1 114 con resolución QCIF). Entre ellas, esas imágenes del nivel temporal más bajo forman la capa base temporal, como se muestra en la capa 0 112 de la FIG. 3. La capa base temporal (capa 0) 112 se puede mejorar con imágenes de un nivel temporal superior, por ejemplo, la capa 1 114 con una frecuencia de trama de 15 Hz o la capa 4 118 con una frecuencia de trama de 30 Hz.

[0066] Además de la capa base 112, 114, se pueden añadir varias capas de mejora espacial y/o de SNR para proporcionar escalabilidades espaciales y/o de calidad. Por ejemplo, la capa 2 116 con resolución CIF puede ser una capa de mejora espacial en la capa base 112, 114. En otro ejemplo, la capa 3 117 puede ser una capa de mejora SNR en la capa base 112, 114 y en la capa 2 116. Como se muestra en la FIG. 3, cada capa de mejora espacial o de SNR puede ser temporalmente escalable, con la misma estructura de escalabilidad temporal que la capa base 112, 114. Además, una capa de mejora puede mejorar tanto la resolución espacial como la frecuencia de trama. Por ejemplo, la capa 4 118 proporciona una capa de mejora de resolución 4CIF, que aumenta aún más la frecuencia de trama de 15 Hz a 30 Hz.

[0067] La FIG. 4 es una ilustración conceptual que muestra un ejemplo de unidades de acceso de codificación de vídeo escalables 120A-120E ("unidades de acceso 120") en orden de flujo bits. Como se muestra en la FIG. 4, las imágenes o fragmentos codificados en la misma instancia de tiempo son sucesivos en el orden de flujo de bits y forman una unidad de acceso en el contexto de una norma de codificación de vídeo escalable, tal como la extensión SVC a H.264 o SHVC. Las unidades de acceso 120 siguen entonces el orden de decodificación, que podría ser diferente del orden de visualización y determinada, por ejemplo, por la relación de predicción temporal entre las unidades de acceso 120.

[0068] Por ejemplo, la unidad de acceso 120A incluye la imagen 10 de la capa 0 112, la imagen 0 de la capa 2 116, la imagen 0 de la capa 3 117 y la imagen 0 de la capa 4 118. La unidad de acceso 120B incluye la imagen P4 de la capa 0 112, la imagen 4 de la capa 2 116, la imagen 4 de la capa 3 117 y la imagen 4 de la capa 4 118. La unidad de acceso 120C incluye la imagen B2 de la capa 1 114, la imagen 2 de la capa 2 116, la imagen 2 de la capa 3 117 y la imagen 2 de la capa 4 118. La unidad de acceso 120D incluye la imagen 1 de la capa 4 118, y la unidad de acceso 120E incluye la imagen 3 de la capa 4 118.

[0069] La FIG. 5 es un diagrama de bloques que ilustra un codificador 122 de SHVC de 3 capas de ejemplo. Como se ilustra en la FIG. 5, el codificador de SHVC 122 incluye un codificador de capa base 124, un primer codificador de capa de mejora 125 y un segundo codificador de capa de mejora 126. En la sintaxis de alto nivel de solo SHVC, no hay nuevas herramientas de codificación a nivel de bloque en comparación con la codificación de capa única de HEVC. En SHVC, solo se permiten cambios de sintaxis del nivel de fragmento y por encima, tal como el filtrado de imágenes o el muestreo ascendente.

[0070] Para reducir la redundancia entre las capas, las imágenes de capa de referencia cúbicas muestreadas en sentido ascendente para una capa base/inferior pueden generarse y almacenarse en una memoria intermedia de referencia para una capa superior/de mejora de modo que la predicción entre capas se

puede conseguir de la misma manera que la predicción entre tramas dentro de una única capa. Como se ilustra en la FIG. 5, se genera una imagen de referencia entre capas (ILR) 128 remuestreada a partir de una imagen de referencia en el codificador 124 de la capa base y se almacena en el primer codificador de capa de mejora 125. De forma similar, una imagen ILR 129 remuestreada se genera a partir de una imagen de referencia en el primer codificador de capa de mejora 125 y se almacena en el segundo codificador de capa de mejora 126. En SHVC WD3, la imagen ILR se marca como una imagen de referencia a largo plazo para la capa de mejora. La diferencia de vector de movimiento asociada con una imagen de referencia entre capas se limita a cero.

[0071] El próximo despliegue de los dispositivos de televisión de ultraalta definición (UHDTV) y de contenido usará una gama de colores diferente de los dispositivos heredados. Específicamente, HD usa la recomendación BT.709, la Recomendación ITU-R BT.709 "Parameter values for the HDTV standards for production and international programme exchange" diciembre de 2010, mientras que UHDTV usará la recomendación BT.2020, la Recomendación ITU-R BT.2020 "Parameter values for UHDTV systems for production and international programme exchange" abril de 2012. Una gama de colores comprende una gama completa de colores que se pueden reproducir para una imagen, por ejemplo, en una imagen, fragmento, bloque o capa de datos de vídeo. Una diferencia clave entre estos sistemas es que la gama de colores de UHDTV es significativamente mayor que la HD. Se afirma que UHDTV proporcionará una experiencia de visualización más realista, que sea congruente con otras características de UHDTV, como la alta resolución.

[0072] La FIG. 6 es un gráfico que ilustra una gama de colores de ejemplo de una secuencia de vídeo de muestra 130. Como se ilustra en la FIG. 6, la secuencia de vídeo de muestra SWG1 130 se indica como un grupo de puntos dentro de un contorno de línea de la gama de colores UHD BT.2020 132. Para fines comparativos, un contorno de la gama de colores BT.709 HD 134 y un contorno del espacio de color lineal CIE-XYZ 136 de la Comisión Internacional de Iluminación superponen la secuencia de vídeo de muestra SWG1 130. Se observa fácilmente desde la FIG. 6 que la gama de colores BT.2020 UHD 132 es mucho más grande que la gama de colores BT.709 HD 134. Tenga en cuenta el número de píxeles en la secuencia de vídeo de muestra SWG1 130 que se encuentran fuera de la gama de colores 134 de BT.709.

[0073] La FIG. 7 es un diagrama de bloques que ilustra la conversión de la gama de colores BT.709 HD 134 a la gama de colores BT.2020 UHD 132. Tanto la gama de colores BT.709 HD 134 como la gama de colores BT.2020 UHD 132 definen representaciones de píxeles de color en componentes de luma y de croma (por ejemplo, YCbCr o YUV). Cada gama de colores define la conversión hacia y desde el espacio de color lineal CIE-XYZ 136. Este espacio de color intermedio común se puede usar para definir la conversión de los valores de luma y de croma en la gama de colores BT.709 HD 134 a los valores de luma y de croma correspondientes en la gama de colores BT.2020 UHD 132.

[0074] Más detalles con respecto a la gama de colores de la secuencia de muestra ilustrada en la FIG. 6 y la conversión de la gama de colores ilustrada en la FIG. 7 se pueden encontrar en L. Kerofsky, A. Segall, S.-H. Kim, K. Misra, "Color Gamut Scalable Video Coding: New Results", JCTVC-L0334, Ginebra, CH, 14-23 de enero de 2013 (en lo sucesivo denominado "JCTVC-L0334").

[0075] La FIG. 8 es un diagrama de bloques que ilustra un codificador de gama de colores escalable 140 que incluye una unidad de procesamiento de predicción de color 144 que puede generar una imagen de referencia entre capas cuando una gama de colores de capa base y una gama de colores de capa de mejora son diferentes. La unidad de procesamiento de predicción de color 144 puede usarse por un codificador de vídeo, tal como el codificador de vídeo 20 o el decodificador de vídeo 30 de la FIG. 1, para realizar la codificación de vídeo escalable en la gama de colores, en la que la gama de colores de la base y de la capa de mejora sea diferente

[0076] En el ejemplo que se ilustra en la FIG. 8, un bucle de codificación de capa base 142 realiza la codificación de vídeo de imágenes que incluyen datos de color en una primera gama de colores, por ejemplo, BT.709, y un bucle de codificación de capa de mejora 146 realiza la codificación de vídeo de imágenes que incluyen datos de color en una segunda gama de colores, por ejemplo, BT.2020. La unidad de procesamiento de predicción de color 144 realiza la predicción de color para asignar o convertir datos de color de una imagen de referencia de capa base en la primera gama de colores a la segunda gama de colores, y genera una imagen de referencia entre capas para la capa de mejora basándose en los datos de color asignados de la imagen de referencia de la capa base.

[0077] Para lograr una codificación de alta eficiencia, la unidad de proceso de predicción de color 144 está configurada para realizar la predicción de color específica al generar las imágenes de referencia entre capas. Como se describe con más detalle a continuación, la unidad de procesamiento de predicción de color 144 puede configurarse para realizar la predicción de color de acuerdo con cualquiera de los modelos de predicción lineal, un modelo de predicción lineal por tramos o un modelo de predicción de color basada en tablas de consulta 3D.

[0078] Un modelo de predicción lineal se propone en JCTVC-L0334, mencionado anteriormente. En general, el proceso de predicción de color del modelo de predicción lineal puede describirse como un modelo de ganancia y compensación. El modelo de predicción lineal funciona en planos de color individuales. Para facilitar el cálculo de

números enteros, un parámetro describe el número de bits fraccionados usados en el cálculo usando el parámetro numFractionBits. Para cada canal, se especifican una ganancia [c] y un desplazamiento [c]. El modelo de predicción lineal se define de la siguiente manera:

$$\text{Pred}[c][x][y] = (\text{gan}[c] * \text{In}[x][y] + (1 \ll (\text{numFractionBits} - 1))) \gg \text{numFractionBits} + \text{desplaz}[c]$$

5

[0079] Un modelo de predicción lineal por tramos se propone en C. Auyeung, K. Sato, "AHG14: "Color gamut scalable video coding with piecewise linear predictions and shift-offset model", JCTVC-N0271, Viena, Austria, julio de 2013, basado en JCTVC-L0334, mencionado anteriormente. El proceso de predicción de color del modelo de predicción lineal por tramos también se puede describir como un modelo de ganancia y de compensación. El modelo de predicción lineal por tramos se define de la siguiente manera:

10

Asignar $d[c][x][y] = \text{In}[c][x][y] - \text{nodo}[c]$. Si $d[c][x][y] \leq 0$

$$\text{Pred}[c][x][y] = (\text{gan1}[c] * d[c][x][y] + \text{desplaz}[c] + (1 \ll (\text{numFractionBits} - 1))) \gg \text{numFractionBits}$$

si no

15

$$\text{Pred}[c][x][y] = (\text{gan2}[c] * d[c][x][y] + \text{desplaz}[c] + (1 \ll (\text{numFractionBits} - 1))) \gg \text{numFractionBits}$$

Los parámetros de predicción nudo [c], compensación[c], ganancia1[c] y ganancia2 [c] pueden codificarse en el flujo de bits.

20

[0080] La FIG. 9 es una ilustración conceptual que muestra una tabla de consulta 3D 150 de ejemplo para la escalabilidad de la gama de colores. Un modelo de predicción de color basado en tablas de consulta 3D se propone en P. Bordes, P. Andrivon, F. Hiron, "AHG14: Color Gamut Scalable Video Coding using 3D LUT: New Results", JCTVC-N0168, Viena, Austria, julio de 2013 (en lo sucesivo denominado "JCTVC-N0168"). El principio de la tabla de consulta 3D para la escalabilidad de la gama de colores se muestra en la FIG. 9. La tabla de consulta 3D 150 puede considerarse como una submuestra de un primer espacio de color 3D, por ejemplo, la gama de colores BT.709 HD, donde cada vértice está asociado con un triplete de color (y, u, v) correspondiente a valores de un segundo espacio de color 3D (es decir, predictivos), por ejemplo, gama de colores BT.2020 UHD).

25

[0081] En general, la primera gama de colores puede dividirse en octantes o cuboides en cada dimensión de color (es decir, Y, U, y V), y los vértices de los octantes están asociados con el triplete de color correspondiente a la segunda gama de colores y usado para introducir datos en la tabla de consulta 3D 150. El número de vértices o segmentos en cada dimensión de color indica el tamaño de la tabla de consulta 3D. La FIG. 9(a) ilustra los vértices o los puntos de red de intersección de los octantes en cada dimensión de color. La FIG. 9(b) ilustra los diferentes valores de color asociados con cada uno de los vértices. Como se ilustra, en la FIG. 9(a) cada dimensión de color tiene cuatro vértices y en la FIG. 9(b) cada dimensión de color incluye cuatro valores de color.

30

35

[0082] La FIG. 10 es una ilustración conceptual que muestra la interpolación trilineal con una tabla de consulta 3D 150 para la escalabilidad de la gama de colores. Para una muestra de color de la capa base en la primera gama de colores, el cálculo de su predicción en la segunda gama de colores para una capa de mejora se realiza mediante la interpolación trilineal de acuerdo con la siguiente ecuación:

40

$$\overline{\text{valor}}_y = K \times \sum_{i=0,1} \sum_{j=0,1} \sum_{k=0,1} s_i(y) \times s_j(u) \times s_k(v) \times \text{LUT}[y_i][u_j][v_k].y$$

45

$$K = \frac{1}{(y_1 - y_0) \times (u_1 - u_0) \times (v_1 - v_0)}$$

Donde:

$$S_0(y) = y_1 - y \quad \text{y} \quad S_1(y) = y - y_0$$

50

y_0 es el índice del vértice inferior submuestreado más cercano a y ,

y_1 es el índice del vértice superior submuestreado más cercano a y .

5 **[0083]** En algunos ejemplos, se puede generar una tabla de consulta 3D independiente para cada uno de los componentes de color, es decir, un componente de luma (Y), un primer componente de croma (U) y un segundo componente de croma (V). Cada una de las tablas de consulta 3D incluye una dimensión de luma (Y), una primera dimensión de croma (U) y una segunda dimensión de croma (V), y se indexa usando los tres componentes de color independientes (Y, U, V).

10 **[0084]** En un ejemplo, se puede definir una función de asignación para cada componente de color basada en la tabla de consulta 3D. Una función de asignación de ejemplo para un valor de píxel de luma (Y) se presenta en la siguiente ecuación:

$$Y_E = LUT_Y(Y_B, U_B, V_B) * Y_B + LUT_U(Y_B, U_B, V_B) * U_B + LUT_V(Y_B, U_B, V_B) * V_B + LUT_C(Y_B, U_B, V_B)$$

15 En la ecuación anterior, Y_E representa el valor de píxel de luma en la capa de mejora, (Y_B, U_B, V_B) representa un valor de píxel de capa base, y LUT_Y , LUT_U , LUT_V y LUT_C representan la tabla de consulta 3D para cada componente de color Y, U, V y una constante. Se pueden definir funciones de asignación similares para un primer valor de píxel de croma (U) y un segundo valor de píxel de croma (V) en la capa de mejora.

20 **[0085]** Más detalles de la tabla de consulta 3D que se ilustra en la FIG. 9 y la interpolación trilineal con la tabla de consulta 3D que se ilustra en la FIG. 10 se pueden encontrar en JCTVC-N0168, mencionado anteriormente.

25 **[0086]** En general, una tabla de consulta 3D basada en la escalabilidad de la gama de colores da como resultado un buen rendimiento de la codificación. Sin embargo, el tamaño de la tabla de consulta 3D puede ser preocupante, ya que la tabla de consulta 3D se almacena típicamente en la memoria caché en una implementación de hardware. Una tabla de consulta 3D con un tamaño de tabla grande puede dar como resultado una alta complejidad de cálculo y un alto coste de señalización. Por ejemplo, convencionalmente, las tablas de consulta 3D siempre son simétricas, de manera que las tablas de consulta 3D tienen el mismo tamaño para el componente de luma, el primer componente de croma y el segundo componente de croma. Además, convencionalmente, las tablas de consulta 3D siempre están equilibradas de manera que un tamaño de cada dimensión de las tablas de consulta 3D es siempre el mismo. Esto da como resultado tamaños de tablas grandes con alta complejidad de cálculo y altos costes de señalización. Por ejemplo, los tamaños de tablas pueden ser de hasta 9x9x9 o 17x17x17.

35 **[0087]** Se proponen los siguientes procedimientos para reducir tanto el coste de la señalización como el coste de cálculo para la tabla de consulta 3D basada en la escalabilidad de la gama de colores.

40 **[0088]** El primer procedimiento incluye la generación de una tabla de consulta 3D asimétrica de manera que los componentes de luma (Y) y de croma (U y V) tienen diferentes tamaños. En algunos casos, la tabla de consulta 3D puede tener un tamaño más grande, es decir, más segmentos, para el componente de luma que para cada uno de los primero y segundo componentes de croma. En este caso, los componentes de croma pueden usar una tabla de consulta más gruesa y el componente de luma puede usar una tabla de consulta más perfeccionada. En otros casos, la tabla de consulta 3D puede tener un tamaño mayor para uno o ambos componentes de croma que para el componente de luma

45 **[0089]** En general, la predicción de color basada en tablas de consulta 3D es un tipo de predicción lineal por tramos en 3D. Teóricamente, cuantos más segmentos se usen en cada dimensión de color (es decir, Y, U y V) de la tabla de consulta 3D, mayor será la exactitud de la predicción de color. Sin embargo, una gran cantidad de segmentos puede conducir a un alto coste de señalización y una alta complejidad de cálculo (es decir, un tamaño de tabla más grande). En la práctica, la señal de croma puede ser menos importante y más fácil de predecir. En este caso, puede ser mejor tener una tabla de consulta de alta resolución para el componente de luma (Y) y tener tablas de consulta de baja resolución para cada uno del primer componente de croma (U) y el segundo componente de croma (V).

55 **[0090]** De acuerdo con las técnicas de esta divulgación, un codificador de vídeo, como el codificador de vídeo 20 y/o el decodificador de vídeo 30, puede generar la tabla de consulta 3D asimétrica utilizando un número diferente de segmentos para la dimensión de luma de la tabla de consulta 3D que para la primera dimensión de croma y la segunda dimensión de croma de la tabla de consulta 3D. Por ejemplo, un codificador de vídeo 20 y/o el decodificador de vídeo 30 pueden generar la tabla de consulta 3D para que tenga un tamaño mayor, es decir, más segmentos para el componente de luma que para cada uno de los primero y segundo componentes de croma. En un ejemplo, la tabla de consulta 3D puede tener un tamaño de hasta 8x2x2. De esta manera, el tamaño total de la tabla puede reducirse y el coste de señalización también puede reducirse mientras se mantiene un buen rendimiento de la codificación con una resolución más alta para el componente de luma que para los primero y segundo componentes de croma. Los tamaños de tabla para la tabla de consulta 3D se pueden señalar en un flujo de bits, tal como en un conjunto de parámetros de vídeo (VPS), conjunto de

parámetros de secuencia (SPS), conjunto de parámetros de imagen (PPS), encabezamiento de fragmento o extensiones relacionadas.

[0091] En algunos ejemplos, cada una de las tablas de consulta 3D puede tener el mismo tamaño $M \times N \times K$, en el que el tamaño (M) de la dimensión de luma de la tabla de consulta 3D es diferente del tamaño (N) de la primera dimensión de croma de la tabla de consulta 3D y el tamaño (K) de la segunda dimensión de croma de la tabla de consulta 3D. Por ejemplo, el tamaño de la dimensión de luma puede ser mayor que los tamaños de las dimensiones de croma, $M > N$ y $M > K$, y los tamaños de las dimensiones de croma de la tabla de consulta 3D pueden ser los mismos, $N=K$. En un ejemplo, cada una de las tablas de consulta 3D puede tener el mismo tamaño de $8 \times 2 \times 2$. En otro ejemplo, cada una de las tablas de consulta 3D puede tener el mismo tamaño de $9 \times 6 \times 6$.

[0092] En otros ejemplos, una tabla de consulta 3D de componentes de luma (LUT_Y) puede tener un primer tamaño, por ejemplo, $M \times M \times M$ o $M \times N \times K$, que es diferente de un segundo tamaño de una primera tabla de consulta 3D de componentes de croma (LUT_U), por ejemplo, $N \times N \times N$, y un tercer tamaño de una segunda tabla de consulta 3D de componentes de croma (LUT_V), por ejemplo, $K \times K \times K$. Por ejemplo, el tamaño de la LUT_Y puede ser más grande que los tamaños de LUT_U y LUT_V . Los tamaños de LUT_U y LUT_V pueden ser iguales o diferentes entre sí. En un ejemplo, la LUT_Y puede tener el tamaño de $8 \times 2 \times 2$, y cada una de la LUT_U y la LUT_V puede tener el tamaño de $2 \times 2 \times 2$. En otro ejemplo, la LUT_Y puede tener el tamaño de $9 \times 9 \times 9$ o $9 \times 6 \times 6$, y cada una de la LUT_U y la LUT_V puede tener el tamaño de $9 \times 3 \times 3$ o $3 \times 3 \times 3$.

[0093] En otro ejemplo, la precisión de la tabla de consulta 3D puede depender del componente de color asociado. Por ejemplo, un primer valor de precisión de la tabla de consulta 3D del componente de luma puede ser diferente de un segundo valor de precisión de las tablas de consulta 3D del primer y segundo componente de croma. Por ejemplo, el primer valor de precisión de la tabla de consulta 3D del componente de luma puede ser mayor que el segundo valor de precisión de las tablas de consulta 3D del componente de croma. En un ejemplo, el primer valor de precisión puede ser de 8 bits para el componente de luma y el segundo valor de precisión puede ser de 6 bits para los componentes de croma. En algunos casos, al generar las tablas de consulta 3D, se puede aplicar un desplazamiento adicional para cumplir con una profundidad de bits objetivo de la capa de mejora. El uso de diferentes valores de precisión de un valor de precisión por defecto para las tablas de consulta 3D basándose en el componente de color asociado puede indicarse en el VPS, SPS, PPS, encabezamiento de fragmento o extensiones relacionadas.

[0094] La tabla de consulta 3D propuesta en JCTVC-N0168 siempre está equilibrada de manera que el tamaño de cada dimensión de la tabla de consulta 3D sea siempre el mismo. Como se ha analizado anteriormente, cuantos más segmentos se usen en cada dimensión de color (es decir, Y, U y V) de la tabla de consulta 3D, mejor será la eficacia de predicción de color. Sin embargo, una gran cantidad de segmentos puede conducir a un alto coste de señalización y una alta complejidad de cálculo (es decir, un tamaño de tabla más grande). Teniendo en cuenta que cada componente de color en una primera gama de colores normalmente tiene una mayor correlación con el mismo componente de color en una segunda gama de colores, puede ser más útil mejorar la eficacia de predicción usando más segmentos, es decir, usar un tamaño más grande para la dimensión de la tabla cuando el componente de color asociado se usa como índice de tabla para la tabla de consulta 3D. El tamaño de la dimensión de la tabla puede ser menor cuando se utiliza un componente de color diferente como un índice de tabla para la tabla de consulta 3D.

[0095] De acuerdo con las técnicas de esta descripción, un codificador de vídeo, como el codificador de vídeo 20 y/o el decodificador de vídeo 30, puede generar la tabla de consulta 3D desequilibrada utilizando más segmentos para la dimensión de la tabla de consulta 3D asociada con el componente de color utilizado como índice de tabla para la tabla de consulta 3D. Por ejemplo, la tabla de consulta 3D del componente de luma puede tener una dimensión de luma más grande que una primera dimensión de croma y una segunda dimensión de croma basándose en el componente de luma que se utiliza como un índice de tabla para la tabla de consulta 3D del componente de luma. En un ejemplo, la LUT_Y puede tener el tamaño $M \times N \times N$, donde $M > N$. Los tamaños de las tablas de consulta 3D del primer y segundo componente de croma se pueden determinar de manera similar. Por ejemplo, la LUT_U puede tener el tamaño $N \times M \times N$ basándose en el primer componente de croma que se utiliza en el índice de tabla, y la LUT_V puede tener el tamaño $N \times N \times M$ basada en el segundo componente de croma que se utiliza en el índice de tabla. De esta manera, el tamaño total de cada una de las tablas puede reducirse mientras se mantiene un buen rendimiento de la codificación con una resolución más alta para el componente de color utilizado como el índice de tabla.

[0096] En un ejemplo, cuando el componente Y, U, V se usa como entrada en la tabla de consulta 3D, se puede usar una tabla $8 \times 2 \times 2$ para el componente Y, se puede usar una tabla $2 \times 8 \times 2$ para el componente U y se puede usar una tabla $2 \times 2 \times 8$ para el componente V. En otro ejemplo, cuando el componente Y, U, V se usa como entrada en la tabla de consulta 3D, se puede usar una tabla $9 \times 3 \times 3$ para el componente Y, se puede usar una tabla $3 \times 9 \times 3$ para el componente U y se puede usar una tabla $3 \times 3 \times 9$ para el componente V.

[0097] Los tamaños de tabla para la tabla de consulta 3D se pueden señalar en un flujo de bits tal como en el VPS, SPS, PPS, encabezamiento de fragmento o extensiones relacionadas. En los ejemplos anteriores, solo los valores de M y N pueden señalizarse para indicar los tamaños de la tabla. En algunos casos, en lugar de señalar los valores de M y N, los valores por defecto se pueden establecer para M y N, de modo que no se necesite señalar el tamaño de la tabla. Por ejemplo, N puede establecerse en un valor de 3 o un valor de 2 por defecto.

[0098] La FIG. 11 es un diagrama de bloques que ilustra un ejemplo de un codificador de vídeo 20 que puede implementar técnicas para utilizar una tabla de consulta 3D basada en la escalabilidad de la gama de colores en la codificación de vídeo multicapa. El codificador de vídeo 20 puede realizar la intracodificación e intercodificación de bloques de vídeo dentro de fragmentos de vídeo. La intracodificación se basa en la predicción espacial para reducir o eliminar la redundancia espacial en el vídeo dentro de una trama o imagen de vídeo dada. La intercodificación se basa en la predicción temporal para reducir o eliminar la redundancia temporal en el vídeo dentro de tramas o imágenes adyacentes de una secuencia de vídeo. El intramodo (modo I) puede referirse a cualquiera de varios modos de codificación de base espacial. Los intermodos, tales como la predicción unidireccional (modo P) o la bipredicción (modo B), pueden referirse a cualquiera de varios modos de codificación de base temporal.

[0099] Como se muestra en la FIG. 11, el codificador de vídeo 20 recibe un bloque de vídeo actual dentro de una trama de vídeo que se va a codificar. En el ejemplo de la FIG. 11, el codificador de vídeo 20 incluye una unidad de selección de modo 40, una memoria de datos de vídeo 41, una memoria de imágenes de referencia 64, un sumador 50, una unidad de procesamiento de transformada 52, una unidad de cuantificación 54 y una unidad de codificación por entropía 56. A su vez, la unidad de selección de modo 40 incluye una unidad de compensación de movimiento 44, una unidad de estimación de movimiento 42, una unidad de intrapredicción 46, una unidad de partición 48 y una unidad de procesamiento de predicción de color 66. Para la reconstrucción de bloques de vídeo, el codificador de vídeo 20 incluye también una unidad de cuantificación inversa 58, una unidad de procesamiento de transformada inversa 60 y un sumador 62. También se puede incluir un filtro de eliminación de bloques (no se muestra en la FIG. 11) para filtrar fronteras de bloques, para eliminar distorsiones de efecto pixelado del vídeo reconstruido. Si se desea, el filtro de eliminación de bloques filtrará típicamente la salida del sumador 62. También pueden usarse filtros adicionales (en el bucle o después del bucle), además del filtro de eliminación de bloques. Dichos filtros no se muestran por razones de brevedad pero, si se desea, pueden filtrar la salida del sumador 50 (como un filtro en el bucle).

[0100] La memoria de datos de vídeo 41 puede almacenar datos de vídeo a codificar por parte de los componentes del codificador de vídeo 20. Los datos de vídeo almacenados en la memoria de datos de vídeo 41 se pueden obtener, por ejemplo, a partir de la fuente de vídeo 18. La memoria intermedia de imágenes decodificadas 64 puede ser una memoria de imágenes de referencia que almacena datos de vídeo de referencia para su uso en la codificación de datos de vídeo mediante el codificador de vídeo 20, por ejemplo, en los modos de intracodificación o intercodificación. La memoria de datos de vídeo 41 y la memoria intermedia de imágenes decodificadas 64 pueden estar formadas por cualquiera de una variedad de dispositivos de memoria, tales como memoria dinámica de acceso aleatorio (DRAM), incluyendo DRAM síncrona (SDRAM), RAM magnetorresistiva (MRAM), RAM resistiva (RRAM) u otros tipos de dispositivos de memoria. La memoria de datos de vídeo 41 y la memoria intermedia de imágenes decodificadas 64 pueden proporcionarse mediante el mismo dispositivo de memoria o dispositivos de memoria independientes. En diversos ejemplos, la memoria de datos de vídeo 41 puede estar en un chip con otros componentes del codificador de vídeo 20, o fuera del chip con respecto a esos componentes.

[0101] Durante el proceso de codificación, el codificador de vídeo 20 recibe una trama o un fragmento de vídeo a codificar. La trama o el fragmento pueden dividirse en múltiples bloques de vídeo. La unidad de estimación de movimiento 42 y la unidad de compensación de movimiento 44 llevan a cabo la codificación interpredictiva del bloque de vídeo recibido con respecto a uno o más bloques en una o más tramas de referencia para proporcionar predicción temporal. La unidad de intrapredicción 46, de forma alternativa, puede llevar a cabo la codificación intrapredictiva del bloque de vídeo recibido con respecto a uno o más bloques vecinos en la misma trama o fragmento que el bloque a codificar para proporcionar predicción espacial. El codificador de vídeo 20 puede llevar a cabo múltiples pasadas de codificación, por ejemplo, para seleccionar un modo de codificación adecuado para cada bloque de datos de vídeo.

[0102] Además, la unidad de partición 48 puede dividir bloques de datos de vídeo en subbloques, basándose en la evaluación de los esquemas de partición anteriores en las pasadas de codificación anteriores. Por ejemplo, la unidad de partición 48 puede dividir inicialmente una trama o un fragmento en varias LCU, y dividir cada una de las LCU en varias sub-CU, basándose en un análisis de velocidad-distorsión (por ejemplo, una optimización de velocidad-distorsión). La unidad de selección de modo 40 puede producir además una estructura de datos de árbol cuádruple que indica la división de una LCU en varias sub-CU. Las CU de nodos hojas del árbol cuádruple pueden incluir una o más PU y una o más TU.

[0103] La unidad de selección de modo 40 puede seleccionar uno de los modos de codificación, intra o inter, por ejemplo, basándose en los resultados de errores, y proporciona el bloque intracodificado o intercodificado resultante al sumador 50 para generar datos de bloques residuales, y al sumador 62 para reconstruir el bloque codificado para su uso como una trama de referencia. La unidad de selección de modo 40 también proporciona elementos de sintaxis, tales como vectores de movimiento, indicadores de intramodo, información de división y otra información de sintaxis de este tipo, a la unidad de codificación por entropía 56.

[0104] La unidad de estimación de movimiento 42 y la unidad de compensación de movimiento 44 pueden estar altamente integradas, pero se ilustran por separado con fines conceptuales. La estimación de movimiento, realizada por la unidad de estimación de movimiento 42, es el proceso de generación de vectores de movimiento, que estiman el movimiento para los bloques de vídeo. Un vector de movimiento, por ejemplo, puede indicar el desplazamiento de una PU de un bloque de vídeo dentro de una trama o imagen de vídeo actual, con respecto a un bloque predictivo dentro de una imagen de referencia (u otra unidad codificada) con respecto al bloque actual que se está codificando dentro del fragmento actual (u otra unidad codificada). Un bloque predictivo es un bloque que se revela como estrechamente coincidente con el bloque a codificar, en lo que respecta a la diferencia de píxeles, lo cual puede determinarse mediante la suma de diferencias absolutas (SAD), la suma de las diferencias al cuadrado (SSD) u otras métricas de diferencia. En algunos ejemplos, el codificador de vídeo 20 puede calcular valores para posiciones de píxel subentero de las imágenes de referencia almacenadas en la memoria intermedia de imágenes decodificadas 64. Por ejemplo, el codificador de vídeo 20 puede interpolar valores de posiciones de un cuarto de píxel, posiciones de un octavo de píxel u otras posiciones de píxel fraccionado de la imagen de referencia. Por lo tanto, la unidad de estimación de movimiento 42 puede realizar una búsqueda de movimiento con respecto a las posiciones de píxel completo y a las posiciones de píxel fraccionado, y emitir un vector de movimiento con una precisión de píxel fraccionado.

[0105] La unidad de estimación de movimiento 42 calcula un vector de movimiento para una PU de un bloque de vídeo de un fragmento sometido a intercodificación, comparando la posición de la PU con la posición de un bloque predictivo de una imagen de referencia. La imagen de referencia puede seleccionarse entre una primera lista de imágenes de referencia (lista 0) o una segunda lista de imágenes de referencia (lista 1), cada una de las cuales identifica una o más imágenes de referencia almacenadas en la memoria intermedia de imágenes decodificadas 64. La unidad de estimación de movimiento 42 envía el vector de movimiento calculado a la unidad de codificación por entropía 56 y a la unidad de compensación de movimiento 44.

[0106] La compensación de movimiento, realizada por la unidad de compensación de movimiento 44, puede implicar extraer o generar el bloque predictivo basándose en el vector de movimiento determinado por la unidad de estimación de movimiento 42. De nuevo, la unidad de estimación de movimiento 42 y la unidad de compensación de movimiento 44 pueden integrarse funcionalmente, en algunos ejemplos. Tras recibir el vector de movimiento para la PU del bloque de vídeo actual, la unidad de compensación de movimiento 44 puede localizar el bloque predictivo al que apunta el vector de movimiento en una de las listas de imágenes de referencia. El sumador 50 forma un bloque de vídeo residual restando los valores de píxel del bloque predictivo a los valores de píxel del bloque de vídeo actual que se está codificando, generando valores de diferencias de píxel, como se analiza posteriormente. En general, la unidad de estimación de movimiento 42 realiza la estimación de movimiento con respecto a los componentes de luma, y la unidad de compensación de movimiento 44 utiliza los vectores de movimiento calculados basándose en los componentes de luma, tanto para los componentes de croma como para los componentes de luma. La unidad de selección de modo 40 también puede generar elementos de sintaxis asociados a los bloques de vídeo y al fragmento de vídeo para su uso por parte del decodificador de vídeo 30 a la hora de decodificar los bloques de vídeo del fragmento de vídeo.

[0107] La unidad de intrapredicción 46 puede intrapredicir un bloque actual, como alternativa a la interpredicción llevada a cabo por la unidad de estimación de movimiento 42 y la unidad de compensación de movimiento 44, como se ha descrito anteriormente. En particular, la unidad de intrapredicción 46 puede determinar un modo de intrapredicción a usar para codificar un bloque actual. En algunos ejemplos, la unidad de intrapredicción 46 puede codificar un bloque actual usando diversos modos de intrapredicción, por ejemplo, durante diferentes pasadas de codificación, y la unidad de intrapredicción 46 (o la unidad de selección de modo 40, en algunos ejemplos) puede seleccionar un modo de intrapredicción adecuado a usar, entre los modos probados.

[0108] Por ejemplo, la unidad de intrapredicción 46 puede calcular valores de velocidad-distorsión usando un análisis de velocidad-distorsión para los diversos modos de intrapredicción probados, y seleccionar el modo de intrapredicción que tenga las mejores características de velocidad-distorsión entre los modos probados. El análisis de velocidad-distorsión determina, en general, una cantidad de distorsión (o error) entre un bloque codificado y un bloque original no codificado que fue codificado para producir el bloque codificado, así como una tasa de bits (es decir, un número de bits) usada para producir el bloque codificado. La unidad de intrapredicción 46 puede calcular relaciones a partir de las distorsiones y velocidades de los diversos bloques codificados, para determinar qué modo de intrapredicción presenta el mejor valor de velocidad-distorsión para el bloque.

[0109] Después de seleccionar un modo de intrapredicción para un bloque, la unidad de intrapredicción 46 puede proporcionar información, indicativa del modo de intrapredicción seleccionado para el bloque, a la unidad de codificación por entropía 56. La unidad de codificación por entropía 56 puede codificar la información que indica el modo de intrapredicción seleccionado. El codificador de vídeo 20 puede incluir, en el flujo de bits transmitido, datos de configuración, que pueden incluir una pluralidad de tablas de índices de modo de intrapredicción y una pluralidad de tablas modificadas de índices de modo de intrapredicción (también denominadas tablas de asignación de palabras código), definiciones de contextos de codificación para diversos bloques e indicaciones de un modo de intrapredicción más probable, una tabla de índices de modo de intrapredicción y una tabla modificada de índices de modo de intrapredicción a usar para cada uno de los contextos.

[0110] El codificador de vídeo 20 forma un bloque de vídeo residual restando los datos de predicción de la unidad de selección de modo 40 del bloque de vídeo original que se está codificando. El sumador 50 representa el componente o los componentes que realizan esta operación de resta. La unidad de procesamiento de transformada 52 aplica una transformada, tal como una transformada discreta del coseno (DCT) o una transformada conceptualmente similar, al bloque residual, produciendo un bloque de vídeo que comprende valores residuales de coeficientes de transformada. La unidad de procesamiento de transformada 52 puede llevar a cabo otras transformadas que son conceptualmente similares a la DCT. También se podrían usar transformaciones de ondículas, transformaciones de números enteros, transformaciones de subbandas u otros tipos de transformaciones. En cualquier caso, la unidad de procesamiento de transformada 52 aplica la transformada al bloque residual, produciendo un bloque de coeficientes de transformada residuales. La transformada puede convertir la información residual desde un dominio de valores de píxel a un dominio de transformada, tal como un dominio de la frecuencia. La unidad de procesamiento de transformada 52 puede enviar los coeficientes de transformada resultantes a la unidad de cuantificación 54. La unidad de cuantificación 54 cuantifica los coeficientes de transformada para reducir aún más la tasa de bits. El proceso de cuantificación puede reducir la profundidad de bits asociada a algunos de, o a todos, los coeficientes. El grado de cuantificación se puede modificar ajustando un parámetro de cuantificación. En algunos ejemplos, la unidad de cuantificación 54 puede realizar, a continuación, una exploración de la matriz que incluye los coeficientes de transformada cuantificados. De forma alternativa, la unidad de codificación por entropía 56 puede realizar la exploración.

[0111] Tras la cuantificación, la unidad de codificación por entropía 56 codifica por entropía los coeficientes de transformada cuantificados. Por ejemplo, la unidad de codificación por entropía 56 puede llevar a cabo la codificación de longitud variable adaptativa al contexto (CAVLC), la codificación aritmética binaria adaptativa el contexto (CABAC), la codificación aritmética binaria adaptativa el contexto basada en sintaxis (SBAC), la codificación por entropía por división en intervalos de probabilidad (PIPE) u otra técnica de codificación por entropía. En el caso de la codificación por entropía basada en el contexto, el contexto se puede basar en bloques vecinos. Tras la codificación por entropía por parte de la unidad de codificación por entropía 56, el flujo de bits codificado se puede transmitir a otro dispositivo (por ejemplo, el decodificador de vídeo 30) o archivar para su posterior transmisión o recuperación.

[0112] La unidad de cuantificación inversa 58 y la unidad de procesamiento de transformada inversa 60 aplican la cuantificación inversa y la transformación inversa, respectivamente, para reconstruir el bloque residual en el dominio de píxel, por ejemplo, para su posterior uso como bloque de referencia. La unidad de compensación de movimiento 44 puede calcular un bloque de referencia añadiendo el bloque residual a un bloque predictivo de una de las tramas de la memoria intermedia de imágenes de referencia 64. La unidad de compensación de movimiento 44 también puede aplicar uno o más filtros de interpolación al bloque residual reconstruido para calcular valores de píxel subentero para su uso en la estimación de movimiento. El sumador 62 añade el bloque residual reconstruido al bloque de predicción compensado por movimiento, producido por la unidad de compensación de movimiento 44, para producir un bloque de vídeo reconstruido para su almacenamiento en la memoria intermedia de imágenes decodificadas 64. El bloque de vídeo reconstruido puede ser utilizado por la unidad de estimación de movimiento 42 y la unidad de compensación de movimiento 44, como bloque de referencia para intercodificar un bloque en una trama de vídeo posterior.

[0113] De acuerdo con las técnicas descritas en esta divulgación, el codificador de vídeo 20 está configurado para realizar la tabla de consulta 3D basada en la escalabilidad de la gama de colores cuando se codifican datos de vídeo multicapa. El codificador de vídeo 20 puede predecir y codificar datos de vídeo multicapa de acuerdo con cualquiera de la extensión SHVC, la extensión MV-HEVC y la extensión 3D-HEVC, u otras extensiones de codificación de vídeo multicapa. Específicamente, la unidad de procesamiento de predicción de color 66 del codificador de vídeo 20 puede generar imágenes de referencia entre capas usadas para predecir bloques de vídeo en una imagen de una capa superior de los datos de vídeo cuando una gama de colores para la capa superior de los datos de vídeo sea diferente de una gama de colores para una capa inferior de datos de vídeo.

[0114] La unidad de procesamiento de predicción de color 66 del codificador de vídeo 20 puede realizar la predicción de color usando una tabla de consulta 3D para la escalabilidad de la gama de colores para convertir datos de color de una imagen de referencia en una primera gama de colores para la capa inferior de los datos de vídeo a una segunda gama de colores para la capa superior de los datos de vídeo. En algunos ejemplos, la

unidad de procesamiento de predicción de color 66 puede generar una tabla de consulta 3D separada para cada uno de los componentes de color, es decir, un componente de luma, un primer componente de croma y un segundo componente de croma. Cada una de las tablas de consulta 3D incluye una dimensión de luma, una primera dimensión de croma y una segunda dimensión de croma, y se indexa utilizando los tres componentes de color independientes.

[0115] De acuerdo con las técnicas descritas en esta descripción, la unidad de procesamiento de predicción de color 66 del codificador de vídeo 20 genera al menos una tabla de consulta 3D que tiene un tamaño que es diferente para el componente de luma que para cada uno del primer componente de croma y el segundo componente de croma. La unidad de procesamiento de predicción de color 66 puede generar esta tabla de consulta 3D asimétrica usando un número de segmentos diferente para la dimensión de luma de la tabla de consulta 3D. En otro ejemplo, la unidad de procesamiento de predicción de color 66 puede generar la tabla de consulta 3D para que tenga un tamaño más grande, es decir, más segmentos para el componente de luma que para cada uno de los primero y segundo componentes de croma. De esta manera, el tamaño total de la tabla puede reducirse mientras se mantiene un buen rendimiento de la codificación con una resolución más alta para el componente de luma que para los primero y segundo componentes de croma.

[0116] En algunos ejemplos, la unidad de procesamiento de predicción de color 66 puede generar cada una de las tablas de consulta 3D para que tengan el mismo tamaño, en el que el tamaño de la dimensión de luma de la tabla de consulta 3D sea diferente de cada tamaño de la primera dimensión de croma de la tabla de consulta 3D y un tamaño de la segunda dimensión de croma de la tabla de consulta 3D. En otros ejemplos, la unidad de procesamiento de predicción de color 66 puede generar una tabla de consulta 3D de componente de luma para que tenga un primer tamaño que sea diferente de cada uno de un segundo tamaño de una tabla de consulta 3D del primer componente de croma y un tercer tamaño de una tabla de consulta 3D del segundo componente de croma. En otro ejemplo, la unidad de procesamiento de predicción de color 66 puede generar una tabla de consulta 3D de componente de luma para que tenga un valor de precisión diferente que cada una de la tabla de consulta 3D del primer componente de croma y la tabla de consulta 3D del segundo componente de croma. En algunos casos, para reducir aún más la complejidad, la unidad de procesamiento de predicción de color 66 solo puede generar una tabla de consulta 3D del componente de luma, realizar la predicción del componente de luma usando la tabla de consulta 3D del componente de luma y realizar la predicción del primer y segundo componente de croma usando una transformación lineal 1D o una transformación lineal por tramos.

[0117] Para mencionar otro ejemplo, de acuerdo con las técnicas de esta descripción, la unidad de procesamiento de predicción de color 66 puede generar una tabla de consulta 3D desequilibrada utilizando más segmentos para la dimensión de la tabla de consulta 3D asociada con el componente de color utilizado como índice de tabla para la tabla de consulta 3D. Por ejemplo, la tabla de consulta 3D del componente de luma puede tener una dimensión de luma más grande que cada una de una primera dimensión de croma y una segunda dimensión de croma basándose en el componente de luma que se utiliza como un índice de tabla para la tabla de consulta 3D del componente de luma. Los tamaños de las tablas de consulta 3D del primer y segundo componente de croma se pueden determinar de manera similar basándose en el primer o segundo componente de croma respectivo que se utiliza en el índice de tabla. De esta manera, el tamaño total de cada una de las tablas puede reducirse mientras se mantiene un buen rendimiento de la codificación con una resolución más alta para el componente de color utilizado como el índice de tabla.

[0118] Tras generar la tabla de consulta 3D, la unidad de procesamiento de predicción de color 66 realiza la predicción de color de una imagen de referencia para la capa inferior de los datos de vídeo usando la tabla de consulta 3D, y genera una imagen de referencia entre capas para la capa superior de la capa superior de los datos de vídeo basándose en la imagen de referencia predictiva de color. Tras generar la imagen de referencia entre capas, la unidad de compensación de movimiento 44 del codificador de vídeo 20 puede funcionar como se describe anteriormente para predecir bloques de vídeo en una imagen de la capa superior de los datos de vídeo basándose en las imágenes de referencia entre capas generadas usando la tabla de consulta 3D. El codificador de vídeo 20 puede codificar entonces datos residuales de los bloques de vídeo predictivos en un flujo de bits para su transmisión al decodificador de vídeo 30. En algunos ejemplos, el codificador de vídeo 20 también puede codificar uno o más elementos de sintaxis que indican el tamaño de la tabla de consulta 3D en el flujo de bits, donde el tamaño es diferente para el componente de luma que para los primero y segundo componentes de croma.

[0119] La FIG. 12 es un diagrama de bloques que ilustra un ejemplo de decodificador de vídeo 30 que puede implementar técnicas para determinar el uso de una tabla de consulta 3D basada en la escalabilidad de la gama de colores en la codificación de vídeo multicapa. En el ejemplo de la FIG. 12, el decodificador de vídeo 30 incluye una unidad de decodificación por entropía 70, una memoria de datos de vídeo 71, una unidad de compensación de movimiento 72, una unidad de procesamiento de intrapredicción 74, una unidad de procesamiento de predicción de color 86, una unidad de cuantificación inversa 76, una unidad de procesamiento de transformada inversa 78, una memoria intermedia de imágenes decodificadas 82 y un sumador 80. En algunos ejemplos, el decodificador de vídeo 30 puede realizar una pasada de decodificación, en general, recíproca a la pasada de codificación descrita con respecto al codificador de vídeo 20 (FIG. 11). La unidad de compensación de

movimiento 72 puede generar datos de predicción basándose en vectores de movimiento recibidos desde la unidad de decodificación por entropía 70, mientras que la unidad de intrapredicción 74 puede generar datos de predicción basándose en indicadores de modo de intrapredicción recibidos desde la unidad de decodificación por entropía 70.

5

[0120] La memoria de datos de vídeo 71 puede almacenar datos de vídeo, tales como un flujo de bits de vídeo codificado, a decodificar por parte de los componentes del decodificador de vídeo 30. Los datos de vídeo almacenados en la memoria de datos de vídeo 71 pueden obtenerse, por ejemplo, a partir del medio legible por ordenador 16, por ejemplo, desde una fuente de vídeo local, tal como una cámara, mediante comunicación de datos de vídeo por red cableada o inalámbrica, o accediendo a medios de almacenamiento físico de datos. La memoria de datos de vídeo 71 puede formar una memoria intermedia de imágenes codificadas (CPB) que almacene datos de vídeo codificados a partir de un flujo de bits de vídeo codificado. La memoria intermedia de imágenes decodificadas 82 puede ser una memoria de imágenes de referencia que almacena datos de vídeo de referencia para su uso en la decodificación de datos de vídeo mediante el decodificador de vídeo 30, por ejemplo, en los modos de intracodificación o intercodificación. La memoria de datos de vídeo 71 y la memoria intermedia de imágenes decodificadas 82 pueden estar formadas por cualquiera de una variedad de dispositivos de memoria, tales como memoria dinámica de acceso aleatorio (DRAM), incluyendo DRAM síncrona (SDRAM), RAM magnetorresistiva (MRAM), RAM resistiva (RRAM) u otros tipos de dispositivos de memoria. La memoria de datos de vídeo 71 y la memoria intermedia de imágenes decodificadas 82 pueden proporcionarse mediante el mismo dispositivo de memoria o dispositivos de memoria independientes. En diversos ejemplos, la memoria de datos de vídeo 71 puede estar en un chip con otros componentes del decodificador de vídeo 30, o fuera del chip con respecto a esos componentes.

10

15

20

[0121] Durante el proceso de decodificación, el decodificador de vídeo 30 recibe un flujo de bits de vídeo codificado, que representa bloques de vídeo de un fragmento de vídeo codificado y elementos de sintaxis asociados, desde el codificador de vídeo 20. La unidad de decodificación por entropía 70 del decodificador de vídeo 30 decodifica por entropía el flujo de bits para generar coeficientes cuantificados, vectores de movimiento o indicadores de modo de intrapredicción y otros elementos de sintaxis. La unidad de decodificación por entropía 70 reenvía los vectores de movimiento y otros elementos de sintaxis a la unidad de compensación de movimiento 72. El decodificador de vídeo 30 puede recibir los elementos de sintaxis en el nivel de fragmento de vídeo y/o el nivel de bloque de vídeo.

25

30

[0122] Cuando el fragmento de vídeo se codifica como un fragmento intracodificado (I), la unidad de procesamiento de intrapredicción 74 puede generar datos de predicción para un bloque de vídeo del fragmento de vídeo actual, basándose en un modo de intrapredicción señalado y datos de bloques previamente decodificados de la trama o imagen actual. Cuando la trama de vídeo se codifica como un fragmento intercodificado (es decir, B o P), la unidad de compensación de movimiento 72 produce bloques predictivos para un bloque de vídeo del fragmento de vídeo actual basándose en los vectores de movimiento y en otros elementos de sintaxis recibidos desde la unidad de decodificación por entropía 70. Los bloques predictivos se pueden producir a partir de una de las imágenes de referencia dentro de una de las listas de imágenes de referencia. El decodificador de vídeo 30 puede construir las listas de imágenes de referencia, la Lista 0 y la Lista 1, usando técnicas de construcción por defecto basándose en las imágenes de referencia almacenadas en la memoria intermedia de imágenes decodificadas 82. La unidad de compensación de movimiento 72 determina la información de predicción para un bloque de vídeo del fragmento de vídeo actual analizando los vectores de movimiento y otros elementos de sintaxis, y usa la información de predicción para producir los bloques predictivos para el bloque de vídeo actual que se está decodificando. Por ejemplo, la unidad de compensación de movimiento 72 usa algunos de los elementos de sintaxis recibidos para determinar un modo de predicción (por ejemplo, intrapredicción o interpredicción), usado para codificar los bloques de vídeo del fragmento de vídeo, un tipo de fragmento de interpredicción (por ejemplo, fragmento B o fragmento P), información de construcción para una o más de las listas de imágenes de referencia para el fragmento, vectores de movimiento para cada bloque de vídeo intercodificado del fragmento, el estado de interpredicción para cada bloque de vídeo intercodificado del fragmento y otra información para decodificar los bloques de vídeo en el fragmento de vídeo actual.

35

40

45

50

[0123] La unidad de compensación de movimiento 72 también puede realizar la interpolación basándose en filtros de interpolación. La unidad de compensación de movimiento 72 puede usar filtros de interpolación como los usados por el codificador de vídeo 20 durante la codificación de los bloques de vídeo para calcular valores interpolados de píxeles subenteros de los bloques de referencia. En este caso, la unidad de compensación de movimiento 72 puede determinar los filtros de interpolación usados por el codificador de vídeo 20 a partir de los elementos de sintaxis recibidos y usar los filtros de interpolación para producir bloques predictivos.

55

60

[0124] La unidad de cuantificación inversa 76 cuantifica inversamente, es decir, descuantifica, los coeficientes de transformada cuantificados proporcionados en el flujo de bits y decodificados por la unidad de decodificación por entropía 70. El proceso de cuantificación inversa puede incluir el uso de un parámetro de cuantificación QP_Y , calculado por el decodificador de vídeo 30 para cada bloque de vídeo en el fragmento de vídeo, para determinar un grado de cuantificación y, asimismo, un grado de cuantificación inversa que debería aplicarse. La unidad de procesamiento de transformada inversa 78 aplica una transformada inversa, por ejemplo una DCT inversa, una

65

transformada inversa de números enteros o un proceso de transformada inversa conceptualmente similar, a los coeficientes de transformada con el fin de producir bloques residuales en el dominio de píxel.

5 **[0125]** Una vez que la unidad de compensación de movimiento 72 ha generado el bloque predictivo para el bloque de vídeo actual basándose en los vectores de movimiento y en otros elementos de sintaxis, el decodificador de vídeo 30 forma un bloque de vídeo decodificado sumando los bloques residuales procedentes de la unidad de procesamiento de transformada inversa 78 a los bloques predictivos correspondientes generados por la unidad de compensación de movimiento 72. El sumador 80 representa el componente o los componentes que realizan esta operación de suma. Si se desea, también se puede aplicar un filtro de eliminación de bloques para filtrar los bloques decodificados a fin de eliminar distorsiones de efecto pixelado. También se pueden usar otros filtros de bucle (en el bucle de codificación o bien después del bucle de codificación) para suavizar las transiciones de píxeles o mejorar de otro modo la calidad de vídeo. Los bloques de vídeo decodificados en una trama o imagen determinada se almacenan a continuación en la memoria intermedia de imágenes decodificadas 82, que almacena imágenes de referencia usadas para una posterior compensación de movimiento. La memoria intermedia de imágenes decodificadas 82 almacena también vídeo decodificado para su presentación posterior en un dispositivo de visualización, tal como el dispositivo de visualización 32 de la FIG. 1.

20 **[0126]** De acuerdo con las técnicas descritas en esta divulgación, el decodificador de vídeo 30 está configurado para realizar la tabla de consulta 3D basada en la escalabilidad de la gama de colores cuando se decodifican datos de vídeo multicapa. El decodificador de vídeo 30 puede decodificar y reconstruir los datos de vídeo multicapa predictivos de acuerdo con cualquiera de la extensión SHVC, la extensión MV-HEVC, la extensión 3D-HEVC u otras extensiones de codificación de vídeo multicapa a HEVC. Específicamente, la unidad de procesamiento de predicción de color 86 del decodificador de vídeo 30 puede generar imágenes de referencia entre capas usadas para predecir bloques de vídeo en una imagen de una capa superior de los datos de vídeo cuando una gama de colores para la capa superior de los datos de vídeo sea diferente a una gama de colores para una capa inferior de datos de vídeo.

30 **[0127]** La unidad de procesamiento de predicción de color 86 del decodificador de vídeo 30 puede realizar la predicción de color usando una tabla de consulta 3D para la escalabilidad de la gama de colores para convertir los datos de color de una imagen de referencia en una primera gama de colores para la capa inferior de los datos de vídeo a una segunda gama de colores para la capa superior de los datos de vídeo. En algunos ejemplos, la unidad de procesamiento de predicción de color 86 puede generar una tabla de consulta 3D independiente para cada uno de los componentes de color, es decir, un componente de luma, un primer componente de croma y un segundo componente de croma. Cada una de las tablas de consulta 3D incluye una dimensión de luma, una dimensión de primer croma y una dimensión de segundo croma, y se indexa usando los tres componentes de color independientes.

40 **[0128]** De acuerdo con las técnicas descritas en esta descripción, la unidad de procesamiento de predicción de color 86 del decodificador de vídeo 30 genera al menos una tabla de consulta 3D que tiene un tamaño que es diferente para el componente de luma que para cada uno del primer componente de croma y el segundo componente de croma. En algunos ejemplos, el decodificador de vídeo 30 puede decodificar uno o más elementos de sintaxis que indican el tamaño de la tabla de consulta 3D en el flujo de bits, donde el tamaño es diferente para el componente de luma que para los primero y segundo componentes de croma. La unidad de procesamiento de predicción de color 86 puede generar esta tabla de consulta 3D asimétrica de acuerdo con el tamaño indicado usando un número diferente de segmentos para la dimensión de luma de la tabla de consulta 3D. En otro ejemplo, la unidad de procesamiento de predicción de color 86 puede generar la tabla de consulta 3D para que tenga un tamaño más grande, es decir, más segmentos para el componente de luma que para cada uno de los primero y segundo componentes de croma. De esta manera, el tamaño total de la tabla puede reducirse mientras se mantiene un buen rendimiento de la codificación con una resolución más alta para el componente de luma que para los primero y segundo componentes de croma.

55 **[0129]** En algunos ejemplos, la unidad de procesamiento de predicción de color 86 puede generar cada una de las tablas de consulta 3D para que tengan el mismo tamaño, en el que el tamaño de la dimensión de luma de la tabla de consulta 3D sea diferente de cada tamaño de la primera dimensión de croma de la tabla de consulta 3D y un tamaño de la segunda dimensión de croma de la tabla de consulta 3D. En otros ejemplos, la unidad de procesamiento de predicción de color 86 puede generar una tabla de consulta 3D de componente de luma para que tenga un primer tamaño que sea diferente de cada uno de un segundo tamaño de una primera tabla de consulta 3D de componente de croma y un tercer tamaño de una segunda búsqueda 3D de componente de croma mesa. En otro ejemplo, la unidad de procesamiento de predicción de color 86 puede generar una tabla de consulta 3D de componente de luma para que tenga un valor de precisión diferente que cada una de la primera tabla de consulta 3D de componente de croma y la segunda tabla de consulta 3D de componente de croma. En algunos casos, para reducir aún más la complejidad, la unidad de procesamiento de predicción de color 86 solo puede generar una tabla de consulta 3D del componente de luma, realizar la predicción del componente de luma usando la tabla de consulta 3D del componente de luma y realizar la predicción del primer y segundo componente de croma usando una transformación lineal 1D o una transformación lineal por tramos.

[0130] Para mencionar otro ejemplo, de acuerdo con las técnicas de esta descripción, la unidad de procesamiento de predicción de color 86 puede generar una tabla de consulta 3D desequilibrada utilizando más segmentos para la dimensión de la tabla de consulta 3D asociada con el componente de color utilizado como índice de tabla para la tabla de consulta 3D. Por ejemplo, la tabla de consulta 3D del componente de luma puede tener una dimensión de luma más grande que cada una de una primera dimensión de croma y una segunda dimensión de croma basándose en el componente de luma que se utiliza como un índice de tabla para la tabla de consulta 3D del componente de luma. Los tamaños de las tablas de consulta 3D del primer y segundo componente de croma se pueden determinar de manera similar basándose en el primer o segundo componente de croma respectivo que se utiliza en el índice de tabla. De esta manera, el tamaño total de cada una de las tablas puede reducirse mientras se mantiene un buen rendimiento de la codificación con una resolución más alta para el componente de color utilizado como el índice de tabla.

[0131] Tras generar la tabla de consulta 3D, la unidad de procesamiento de predicción de color 86 realiza la predicción de color de una imagen de referencia para la capa inferior de los datos de vídeo usando la tabla de consulta 3D, y genera una imagen de referencia entre capas para la capa superior de la capa superior de los datos de vídeo basándose en la imagen de referencia predictiva de color. Tras generar las imágenes de referencia entre capas, la unidad de compensación de movimiento 72 del decodificador de vídeo 30 puede funcionar como se describe anteriormente para reconstruir bloques de vídeo predictivos en una imagen de la capa superior de los datos de vídeo basándose en datos residuales decodificados y en las imágenes de referencia entre capas generadas usando la tabla de consulta 3D.

[0132] La FIG. 13 es un diagrama de flujo que ilustra una operación de ejemplo de generación de imágenes de referencia entre capas que utilizan una tabla de consulta 3D basada en la escalabilidad de la gama de colores. La operación de ejemplo de la FIG. 13 se describe en el presente documento como realizado por la unidad de procesamiento de predicción de color 86 del decodificador de vídeo 30 de la FIG. 12. En otros ejemplos, la operación puede realizarse mediante la unidad de procesamiento de predicción de color 66 del codificador de vídeo 20 de la FIG. 11 o la unidad de procesamiento de predicción de color 144 de la FIG. 8.

[0133] Convencionalmente, en la codificación de vídeo multicapa, una capa inferior de datos de vídeo (por ejemplo, una capa base) y una capa superior de los datos de vídeo (por ejemplo, una capa de mejora) incluyen datos de color en la misma gama de colores, por ejemplo, gama de colores BT.709 HD. En este caso, un codificador de vídeo y/o un decodificador de vídeo pueden generar imágenes de referencia entre capas para la capa superior de los datos de vídeo como versiones muestreadas en sentido ascendente de imágenes de referencia cúbicadas para la capa inferior de los datos de vídeo. En algunos ejemplos, sin embargo, una capa inferior de los datos de vídeo pueden incluir datos de color en una primera gama de colores, por ejemplo, BT.709, y una capa superior de los datos de vídeo pueden incluir datos de color a una segunda gama de colores diferente, por ejemplo, una gama de colores BT.2020 UHD. En este ejemplo, para generar imágenes de referencia entre capas para la capa superior de los datos de vídeo, un codificador de vídeo 30 utiliza la unidad de procesamiento de predicción de color 86 para realizar la predicción de color y convertir los datos de color de una imagen de referencia en la primera gama de colores para la capa inferior de los datos de vídeo a la segunda gama de colores para la capa superior de los datos de vídeo. La unidad de procesamiento de predicción de color 86 del codificador de vídeo 30 puede realizar la predicción de color usando una tabla de consulta 3D para la escalabilidad de la gama de colores.

[0134] De acuerdo con las técnicas de esta descripción, la unidad de procesamiento de predicción de color 86 genera al menos una tabla de consulta 3D para la escalabilidad de la gama de colores que tiene un tamaño diferente para un componente de luma que para cada uno de un primer componente de croma y un segundo componente de croma (180). La unidad de decodificación de entropía 70 puede decodificar uno o más elementos de sintaxis que indican el tamaño de la, al menos una, tabla de consulta 3D, donde el tamaño es diferente para el componente de luma que para cada uno de los primero y segundo componentes de croma. La unidad de procesamiento de predicción de color 86 puede generar esta tabla de consulta 3D asimétrica de acuerdo con el tamaño indicado usando un número diferente de segmentos para la dimensión de luma de la tabla de consulta 3D que para los primero y segundo componentes de croma de la tabla de consulta 3D.

[0135] Por ejemplo, la unidad de procesamiento de predicción de color 86 puede generar la tabla de consulta 3D para que tenga un tamaño mayor, es decir, más segmentos, para el componente de luma que para cada uno de los primero y segundo componentes de croma. En un ejemplo, la tabla de consulta 3D puede tener un tamaño de hasta 8x2x2. De esta manera, el tamaño total de la tabla puede reducirse mientras se mantiene un buen rendimiento de la codificación con una resolución más alta para el componente de luma que para los primero y segundo componentes de croma.

[0136] En algunos ejemplos, la unidad de procesamiento de predicción de color 86 puede generar una tabla de consulta 3D independiente para cada uno de los componentes de color, es decir, un componente de luma (Y), un primer componente de croma (U) y un segundo componente de croma (V). Cada una de las tablas de consulta 3D incluye una dimensión de luma (Y), una primera dimensión de croma (U) y una segunda dimensión de croma (V), y se indexa usando los tres componentes de color independientes (Y, U, V). En algunos ejemplos, la unidad

de procesamiento de predicción de color 86 puede generar cada una de las tablas de consulta 3D para que tenga el mismo tamaño $M \times N \times K$, en el que el tamaño (M) de la dimensión de luma de la tabla de consulta 3D es diferente de cada tamaño (N) de la primera dimensión de croma de la tabla de consulta 3D y el tamaño (K) de la segunda dimensión de croma de la tabla de consulta 3D. Por ejemplo, el tamaño de la dimensión de luma puede ser mayor que los tamaños de las dimensiones de croma, $M > N$ y $M > K$, y los tamaños de las dimensiones de croma pueden ser los mismos, $N=K$. En un ejemplo, cada una de las tablas de consulta 3D puede tener el mismo tamaño de $8 \times 2 \times 2$.

[0137] En otros ejemplos, la unidad de procesamiento de predicción de color 86 puede generar una tabla de consulta 3D de componentes de luma (LUT_Y) para que tenga un primer tamaño, por ejemplo, $M \times M \times M$ o $M \times N \times K$, que sea diferente de cada uno de un segundo tamaño de una tabla de consulta 3D del primer componente de croma (LUT_U), por ejemplo, $N \times N \times N$, y un tercer tamaño de una tabla de consulta 3D del segundo componente de croma (LUT_V), por ejemplo, $K \times K \times K$. Por ejemplo, el tamaño de la tabla de consulta 3D del componente de luma puede ser mayor que los tamaños de las tablas de consulta 3D del componente de croma. Las tablas de consulta 3D del primer y segundo componente de croma pueden ser del mismo tamaño o de tamaños diferentes. En un ejemplo, la LUT_Y puede tener el tamaño de $8 \times 2 \times 2$, y cada una de la LUT_U y la LUT_V puede tener el tamaño de $2 \times 2 \times 2$.

[0138] En otro ejemplo, la unidad de procesamiento de predicción de color 86 puede generar la tabla de consulta 3D del componente de luma con un valor de precisión diferente que el valor de precisión de cada una de la primera tabla de consulta 3D del componente de croma y la segunda tabla de consulta 3D del componente de croma. Por ejemplo, el valor de precisión de la tabla de consulta 3D del componente de luma puede ser mayor que el valor de precisión de las tablas de consulta 3D del componente de croma. En algunos casos, para reducir aún más la complejidad, la unidad de procesamiento de predicción de color 86 solo puede generar una tabla de consulta 3D del componente de luma.

[0139] Para mencionar otro ejemplo, de acuerdo con las técnicas de esta descripción, la unidad de procesamiento de predicción de color 86 del decodificador de vídeo 30 puede generar la, al menos una, tabla de consulta 3D como una tabla de consulta 3D desequilibrada utilizando más segmentos para la dimensión de la tabla de consulta 3D asociada con el componente de color utilizado como un índice de tabla para la tabla de consulta 3D. Por ejemplo, la tabla de consulta 3D del componente de luma puede tener una dimensión de luma más grande que una primera dimensión de croma y una segunda dimensión de croma basándose en el componente de luma que se utiliza como un índice de tabla para la tabla de consulta 3D del componente de luma. En este ejemplo, la LUT_Y puede tener el tamaño $M \times N \times N$, donde $M > N$. Los tamaños de las tablas de consulta 3D del primer y segundo componente de croma se pueden determinar de manera similar. Por ejemplo, la LUT_U puede tener el tamaño $N \times M \times N$ basándose en el primer componente de croma que se utiliza en el índice de tabla, y la LUT_V puede tener el tamaño $N \times N \times M$ basándose en el segundo componente de croma que se utiliza en el índice de tabla. De esta manera, el tamaño total de cada una de las tablas puede reducirse mientras se mantiene un buen rendimiento de la codificación con una resolución más alta para el componente de color utilizado como el índice de tabla.

[0140] Después de generar la, al menos una, tabla de consulta 3D, la unidad de procesamiento de predicción de color 86 realiza la predicción de color utilizando la, al menos una, tabla de consulta 3D para convertir los datos de color de la imagen de referencia en la primera gama de colores para la capa inferior de los datos de vídeo a la segunda gama de colores para la capa superior de los datos de vídeo (182). A continuación la unidad de procesamiento de predicción de color 86 genera al menos una imagen de referencia entre capas para la capa superior de los datos de vídeo basándose en los datos de color convertidos (184). En los casos en que solo se genera una tabla de consulta 3D del componente de luma, la unidad de procesamiento de predicción de color 86 puede realizar la predicción del componente de luma usando la tabla de consulta 3D del componente de luma, y realizar la predicción del primer y segundo componentes de croma usando una transformación lineal 1D o una transformación lineal por tramos para generar un imagen de referencia entre capas. El decodificador de vídeo 30 puede entonces reconstruir bloques de vídeo predictivos en una imagen de la capa superior de los datos de vídeo basándose en los datos residuales decodificados y la, al menos una, imagen de referencia entre capas generada usando la tabla de consulta 3D.

[0141] Ciertos aspectos de esta divulgación se han descrito con respecto a las extensiones de la norma de HEVC con fines ilustrativos. Sin embargo, las técnicas descritas en esta divulgación pueden ser útiles para otros procesos de codificación de vídeo, que incluyen otros procesos de codificación de vídeo, estándar o patentados, aún no desarrollados.

[0142] Un codificador de vídeo, como se describe en esta divulgación, puede referirse a un codificador de vídeo o a un decodificador de vídeo. De manera similar, una unidad de codificación de vídeo puede referirse a un codificador de vídeo o a un decodificador de vídeo. Del mismo modo, la codificación de vídeo puede referirse a la codificación de vídeo o a la decodificación de vídeo, según corresponda.

[0143] Ha de reconocerse que, según el ejemplo, ciertas acciones o sucesos de cualquiera de las técnicas descritas en el presente documento pueden realizarse en una secuencia distinta, pueden añadirse, combinarse u omitirse por completo (por ejemplo, no todas las acciones o sucesos descritos son necesarios para la puesta en práctica de las técnicas). Asimismo, en determinados ejemplos, las acciones o sucesos se pueden realizar simultáneamente, por ejemplo, mediante procesamiento de múltiples hilos, procesamiento de interrupciones o múltiples procesadores, en lugar de secuencialmente.

[0144] En uno o más ejemplos, las funciones descritas pueden implementarse en hardware, software, firmware o cualquier combinación de los mismos. Si se implementan en software, las funciones se pueden almacenar en, o transmitir como, una o más instrucciones o código en un medio legible por ordenador y ejecutarse mediante una unidad de procesamiento basada en hardware. Los medios legibles por ordenador pueden incluir medios de almacenamiento legibles por ordenador, que corresponden a un medio tangible tal como medios de almacenamiento de datos, o medios de comunicación que incluyen cualquier medio que facilite la transferencia de un programa informático desde un lugar a otro, por ejemplo, de acuerdo con un protocolo de comunicación. De esta manera, los medios legibles por ordenador pueden corresponder, en general, a (1) medios de almacenamiento tangibles legibles por ordenador que sean no transitorios o (2) un medio de comunicación tal como una señal o una onda portadora. Los medios de almacenamiento de datos pueden ser cualquier medio disponible al que se pueda acceder desde uno o más ordenadores o uno o más procesadores para recuperar instrucciones, código y/o estructuras de datos para la implementación de las técnicas descritas en esta divulgación. Un producto de programa informático puede incluir un medio legible por ordenador.

[0145] A modo de ejemplo, y no de limitación, dichos medios de almacenamiento legibles por ordenador pueden comprender RAM, ROM, EEPROM, CD-ROM u otro almacenamiento en disco óptico, almacenamiento en disco magnético, u otros dispositivos de almacenamiento magnético, memoria flash o cualquier otro medio que pueda usarse para almacenar código de programa deseado en forma de instrucciones o estructuras de datos y al que pueda accederse mediante un ordenador. Además, cualquier conexión recibe adecuadamente la denominación de medio legible por ordenador. Por ejemplo, si las instrucciones se transmiten desde un sitio web, un servidor u otro origen remoto usando un cable coaxial, un cable de fibra óptica, un par trenzado, una línea de abonado digital (DSL) o tecnologías inalámbricas tales como infrarrojos, radio y microondas, entonces el cable coaxial, el cable de fibra óptica, el par trenzado, la DSL o las tecnologías inalámbricas tales como infrarrojos, radio y microondas se incluyen en la definición de medio. Sin embargo, debería entenderse que los medios de almacenamiento legibles por ordenador y los medios de almacenamiento de datos no incluyen conexiones, ondas portadoras, señales u otros medios transitorios, sino que, en cambio, están orientados a medios de almacenamiento tangibles no transitorios. El término disco, como se usa en el presente documento, incluye el disco compacto (CD), el disco láser, el disco óptico, el disco versátil digital (DVD), el disco flexible y el disco Blu-ray, donde algunos discos normalmente reproducen datos de manera magnética, mientras que otros discos reproducen datos ópticamente con láseres. Las combinaciones de lo anterior se deberían incluir también dentro del alcance de los medios legibles por ordenador.

[0146] Las instrucciones pueden ser ejecutadas por uno o más procesadores, tales como uno o más procesadores de señales digitales (DSP), microprocesadores de propósito general, circuitos integrados específicos de la aplicación (ASIC), matrices lógicas programables in situ (FPGA) u otros circuitos lógicos integrados o discretos equivalentes. Por consiguiente, el término "procesador", como se usa en el presente documento, puede referirse a cualquier estructura anterior o a cualquier otra estructura adecuada para la implementación de las técnicas descritas en el presente documento. Además, en algunos aspectos, la funcionalidad descrita en el presente documento se puede proporcionar en módulos de hardware y/o software dedicados configurados para la codificación y la decodificación, o incorporados en un códec combinado. Además, las técnicas se podrían implementar totalmente en uno o más circuitos o elementos lógicos.

[0147] Las técnicas de esta divulgación se pueden implementar en una amplia variedad de dispositivos o aparatos, incluidos un microteléfono inalámbrico, un circuito integrado (IC) o un conjunto de IC (por ejemplo, un conjunto de chips). Diversos componentes, módulos o unidades se describen en esta divulgación para enfatizar aspectos funcionales de dispositivos configurados para realizar las técnicas divulgadas, pero no requieren necesariamente su realización mediante diferentes unidades de hardware. En cambio, como se ha descrito anteriormente, diversas unidades se pueden combinar en una unidad de hardware de códec o proporcionarse mediante un grupo de unidades de hardware interoperativas, incluidos uno o más procesadores, como se ha descrito anteriormente, conjuntamente con software y/o firmware adecuados.

[0148] Se han descrito diversos ejemplos. No todos los ejemplos están dentro del alcance de las siguientes reivindicaciones.

REIVINDICACIONES

- 5 1. Un procedimiento de procesamiento de datos de vídeo multicapa, el procedimiento que comprende:
- 10 generar una pluralidad de tablas de consulta tridimensionales, 3D, para la escalabilidad de la gama de colores, en el que la pluralidad de tablas de consulta 3D comprende una primera tabla de consulta para un componente de luma y una segunda tabla de consulta para un primer componente de croma y una tercera tabla de consulta para un segundo componente de croma en el que cada una de la primera tabla de consulta y la segunda tabla de consulta y la tercera tabla de consulta comprenden 3 dimensiones que comprenden una dimensión de luma y una primera dimensión de croma y una segunda dimensión de croma y en el que el tamaño de la tabla de consulta 3D para el componente de luma es mayor que el tamaño de la tabla de consulta 3D para el primer componente de croma y en el que el tamaño de la tabla de consulta 3D para el primer componente de croma es el mismo que el tamaño de la tabla de consulta 3D para el segundo componente de croma y en el que el tamaño de dimensión de luma de cada una de las tablas de consulta 3D es mayor que el tamaño de cada una de la primera dimensión de croma y la segunda dimensión de croma;
- 15 codificar una pluralidad de elementos de sintaxis que indican los tamaños para las tablas de consulta 3D;
- 20 realizar la predicción de color usando la, al menos una, tabla de consulta 3D para convertir los datos de color de una imagen de referencia en una primera gama de colores para una capa inferior de los datos de vídeo a una segunda gama de colores para una capa superior de los datos de vídeo basándose en los elementos de sintaxis codificados; y
- 25 generar al menos una imagen de referencia entre capas para la capa superior de los datos de vídeo basándose en los datos de color convertidos.
- 30 2. El procedimiento de la reivindicación 1, en el que la generación de la pluralidad de tablas de consulta 3D comprende además generar una tabla de consulta 3D del componente de luma que tiene un primer valor de precisión, generar una tabla de consulta 3D del primer componente de croma que tiene un segundo valor de precisión y generar una tabla de consulta 3D del segundo componente de croma que tiene el segundo valor de precisión, en el que el primer valor de precisión es diferente del segundo valor de precisión.
- 35 3. El procedimiento de la reivindicación 1, que comprende además:
- 40 predecir bloques de vídeo en una imagen de la capa superior de los datos de vídeo basándose en la, al menos una, imagen de referencia entre capas generada usando la, al menos una, tabla de consulta 3D;
- 45 codificar datos residuales de los bloques de vídeo predictivos en un flujo de bits; y
- codificar uno o más elementos de sintaxis que indican el tamaño de la, al menos una, tabla de consulta 3D en el flujo de bits, en el que los uno o más elementos de sintaxis indican que el tamaño es diferente para el componente de luma que para cada uno de los primero y segundo componentes de croma.
- 50 4. El procedimiento de la reivindicación 1, que comprende además:
- 55 decodificar uno o más elementos de sintaxis que indican el tamaño de la, al menos una, tabla de consulta 3D, en el que los uno o más elementos de sintaxis indican que el tamaño es diferente para el componente de luma que para cada uno de los primero y segundo componentes de croma;
- 60 decodificar datos residuales de bloques de vídeo predictivos; y
- reconstruir los bloques de vídeo predictivos en una imagen de la capa superior de los datos de vídeo basándose en los datos residuales decodificados y la, al menos una, imagen de referencia entre capas generada usando la, al menos una, tabla de consulta 3D.
- 65 5. Un dispositivo de procesamiento de vídeo para procesar datos de vídeo multicapa, el dispositivo que comprende:
- medios para generar una pluralidad de tablas de consulta tridimensionales, 3D, para la escalabilidad de la gama de colores, en el que la pluralidad de tablas de consulta 3D comprende una primera tabla

- 5 de consulta para un componente de luma y una segunda tabla de consulta para un primer componente de croma y una tercera tabla de consulta para un segundo componente de croma en el que cada una de la primera tabla de consulta y la segunda tabla de consulta y la tercera tabla de consulta comprenden 3 dimensiones que comprenden una dimensión de luma y una primera dimensión de croma y una segunda dimensión de croma y en el que el tamaño de la tabla de consulta 3D para el componente de luma es mayor que el tamaño de la tabla de consulta 3D para el primer componente de croma y en el que el tamaño de la tabla de consulta 3D para el primer componente de croma es el mismo que el tamaño de la tabla de consulta 3D para el segundo componente de croma y en el que el tamaño de dimensión de luma de cada una de las tablas de consulta 3D es mayor que el tamaño de cada una de la primera dimensión de croma y la segunda dimensión de croma;
- 10
- medios para codificar una pluralidad de elementos de sintaxis que indican los tamaños para las tablas de consulta 3D;
- 15
- medios para realizar la predicción de color usando la, al menos una, tabla de consulta 3D para convertir los datos de color de una imagen de referencia en una primera gama de colores para una capa inferior de los datos de vídeo a una segunda gama de colores para una capa superior de los datos de vídeo basándose en los elementos de sintaxis codificados; y
- 20
- medios para generar al menos una imagen de referencia entre capas para la capa superior de los datos de vídeo basándose en los datos de color convertidos.
6. Un medio de almacenamiento legible por ordenador que almacena instrucciones para procesar datos de vídeo multicapa que, cuando son ejecutadas, hacen que uno o más procesadores realicen el procedimiento de cualquiera de las reivindicaciones 1 a 4.
- 25

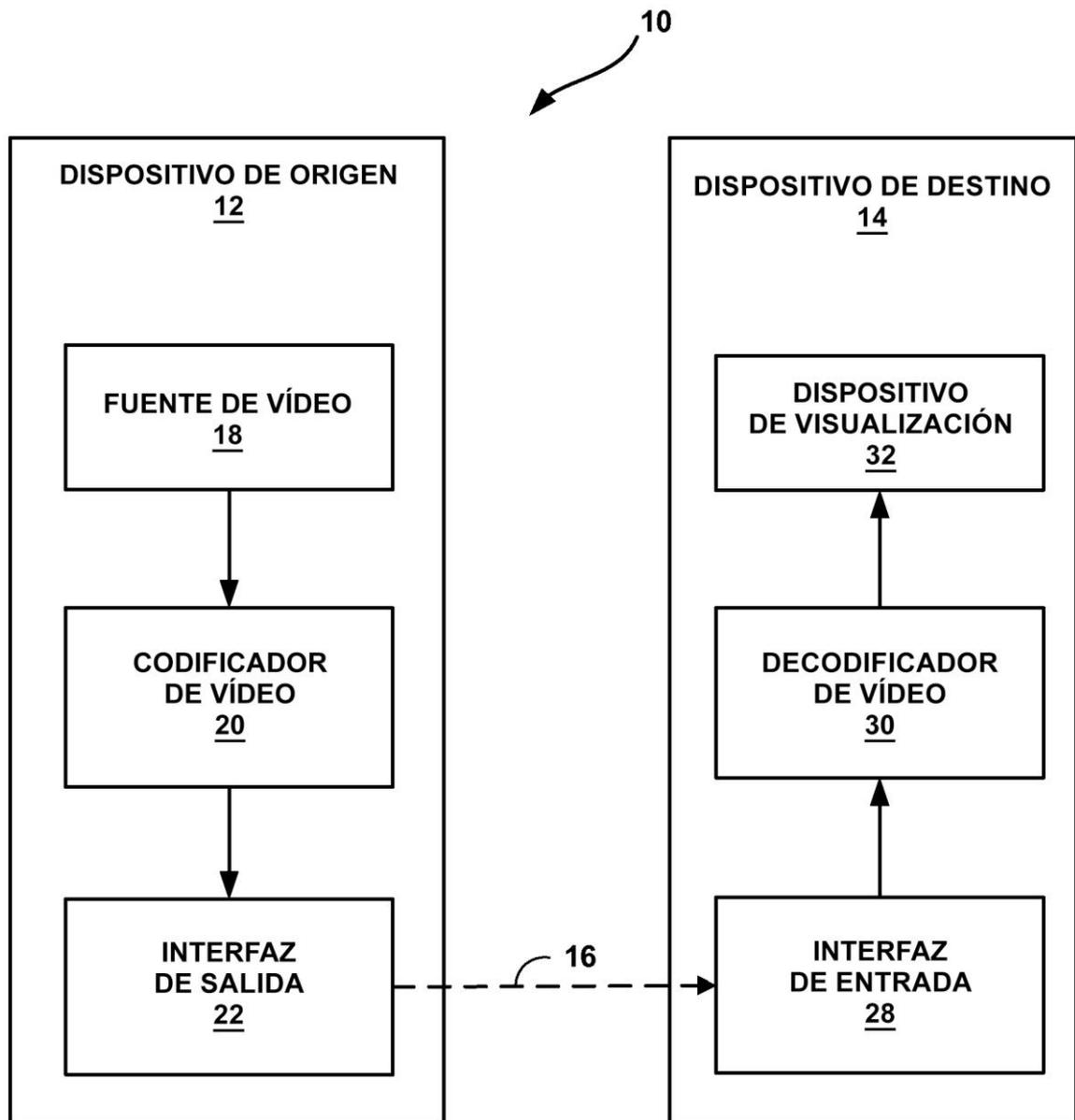


FIG. 1

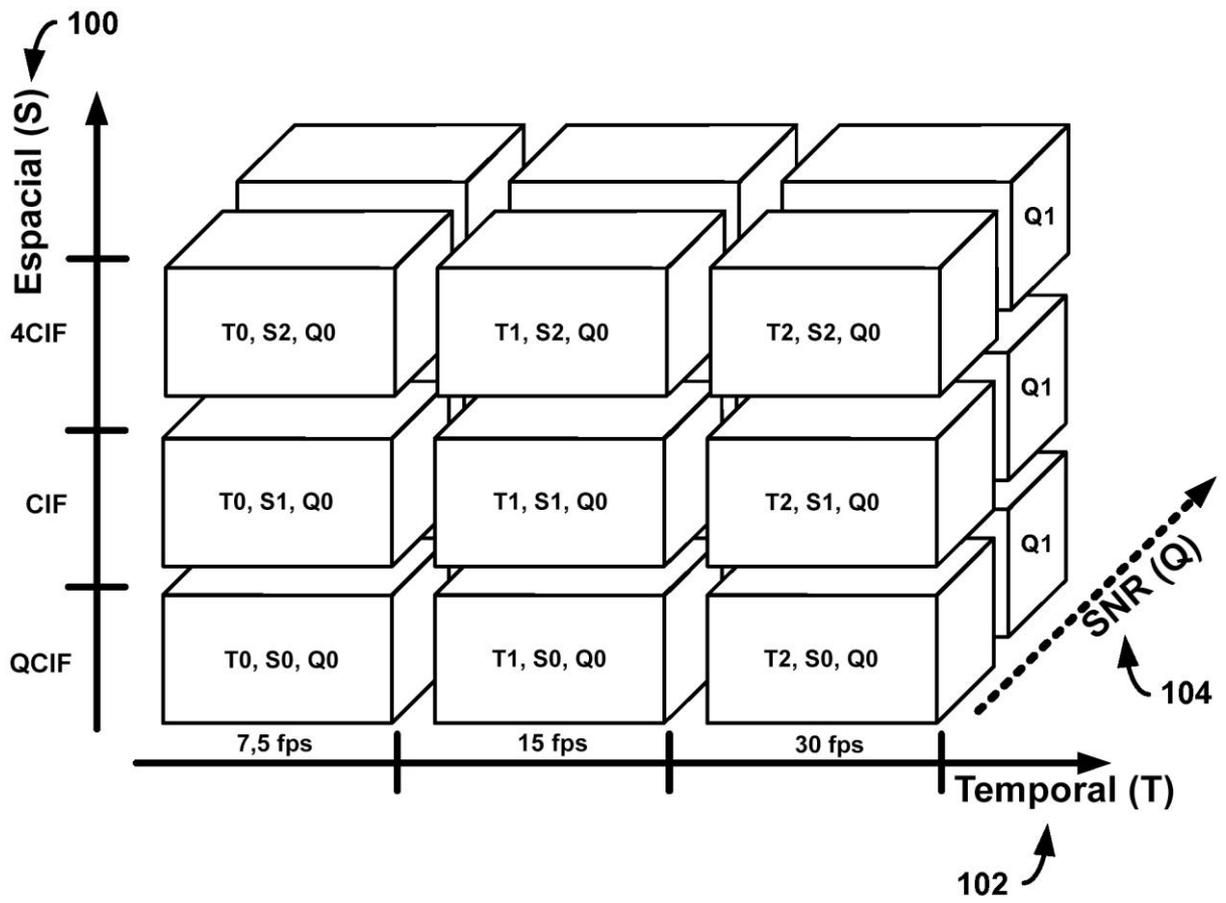


FIG. 2

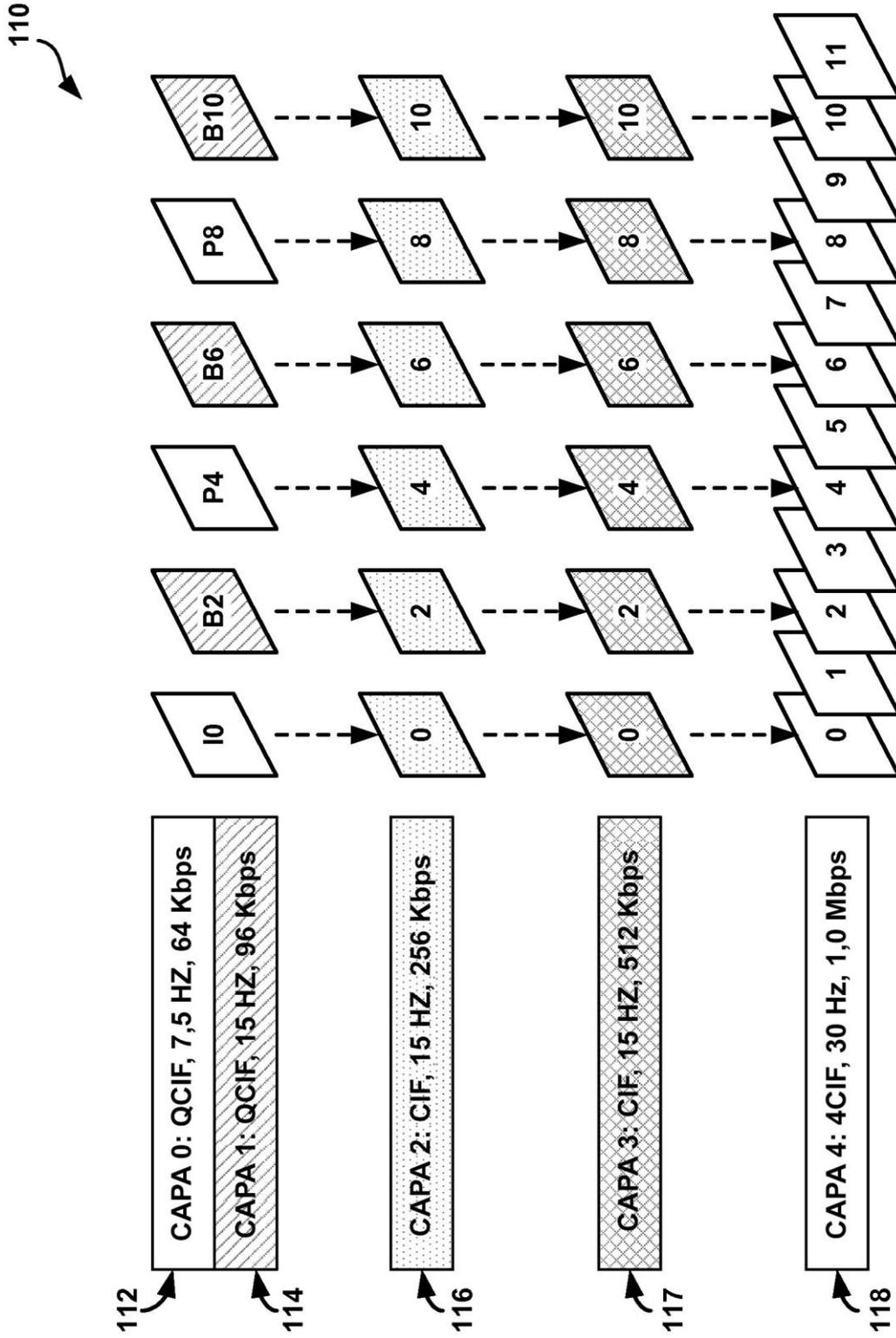


FIG. 3

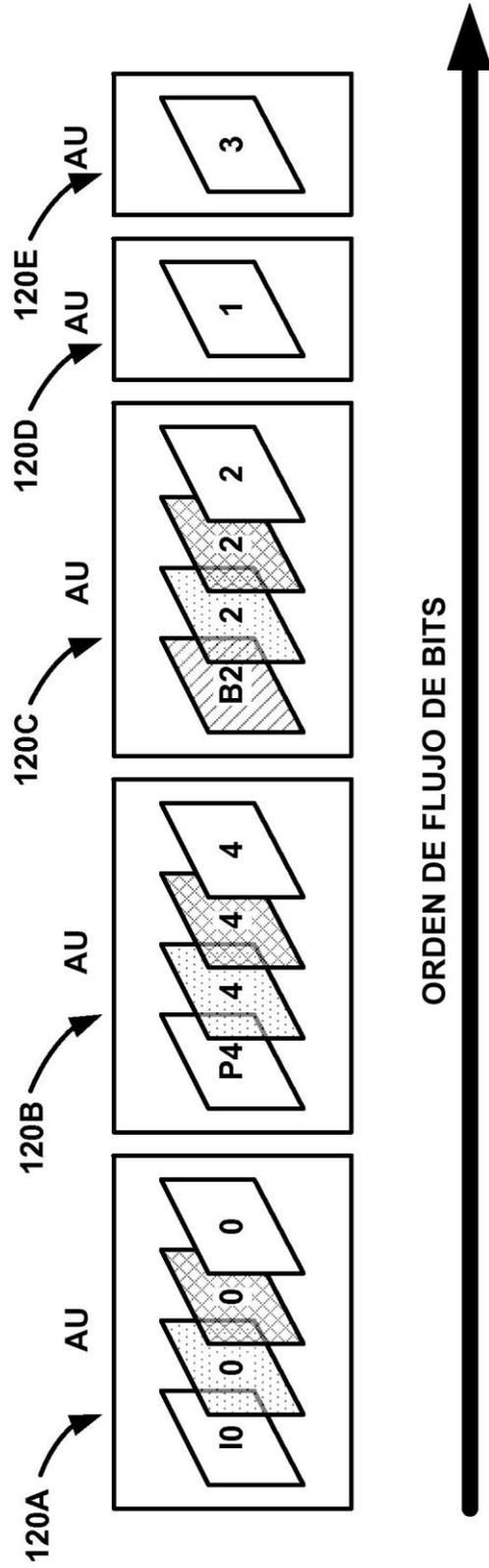


FIG. 4

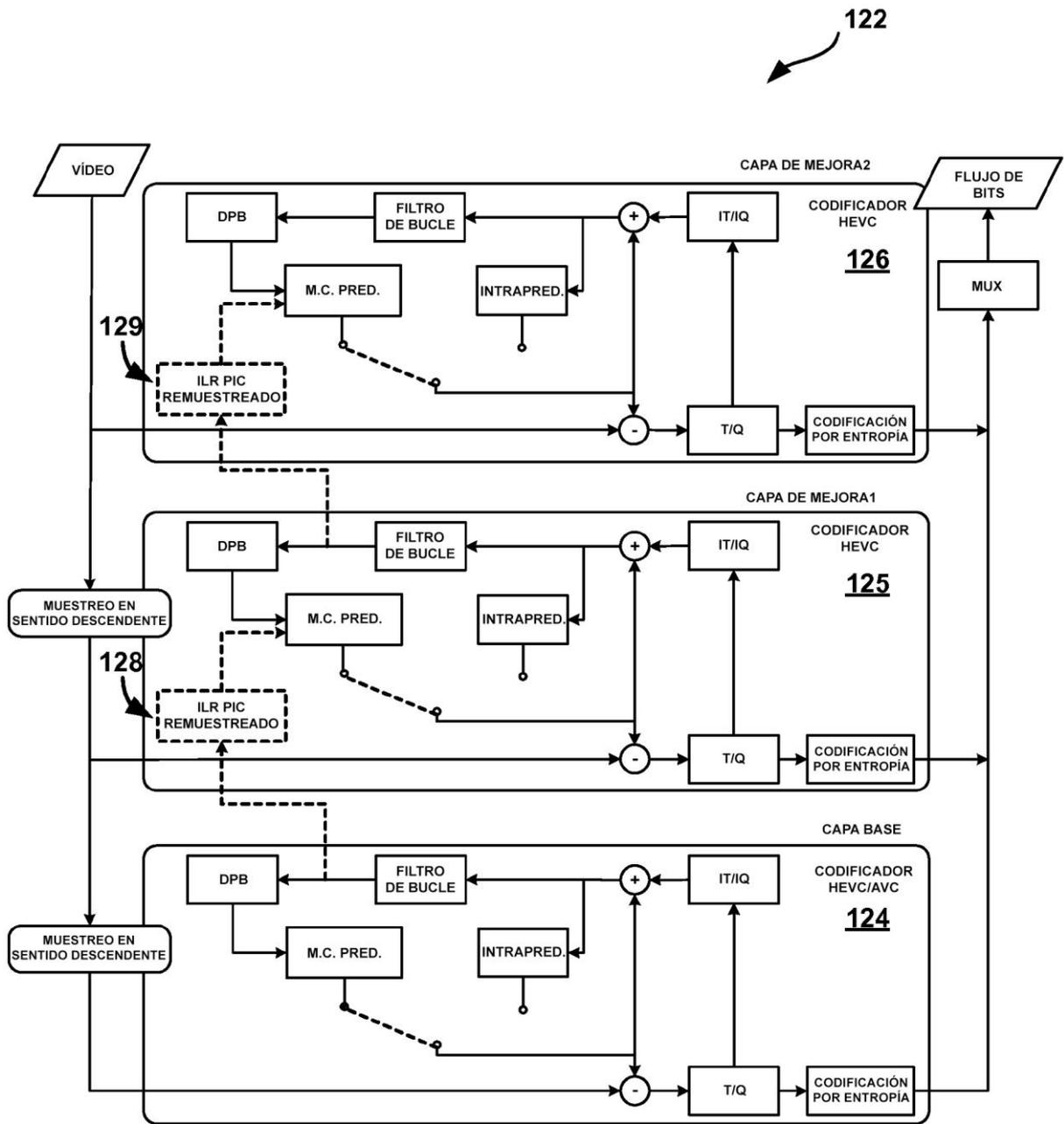


FIG. 5

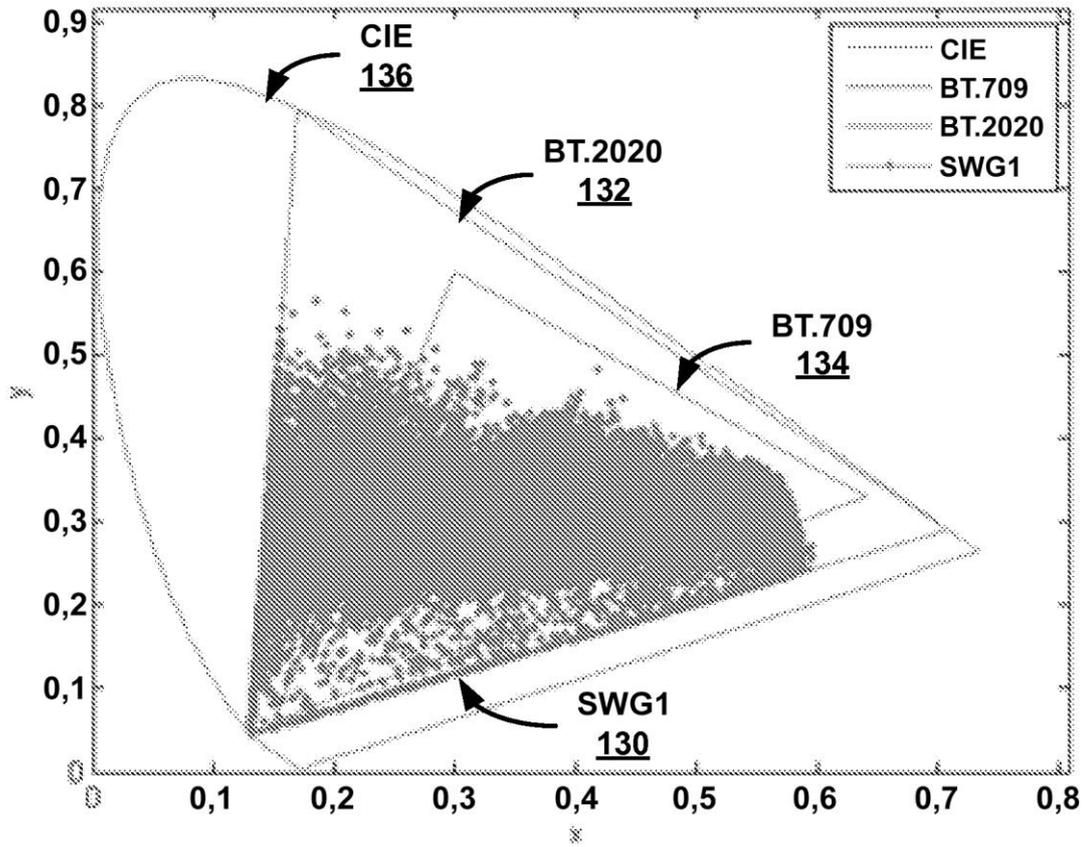


FIG. 6

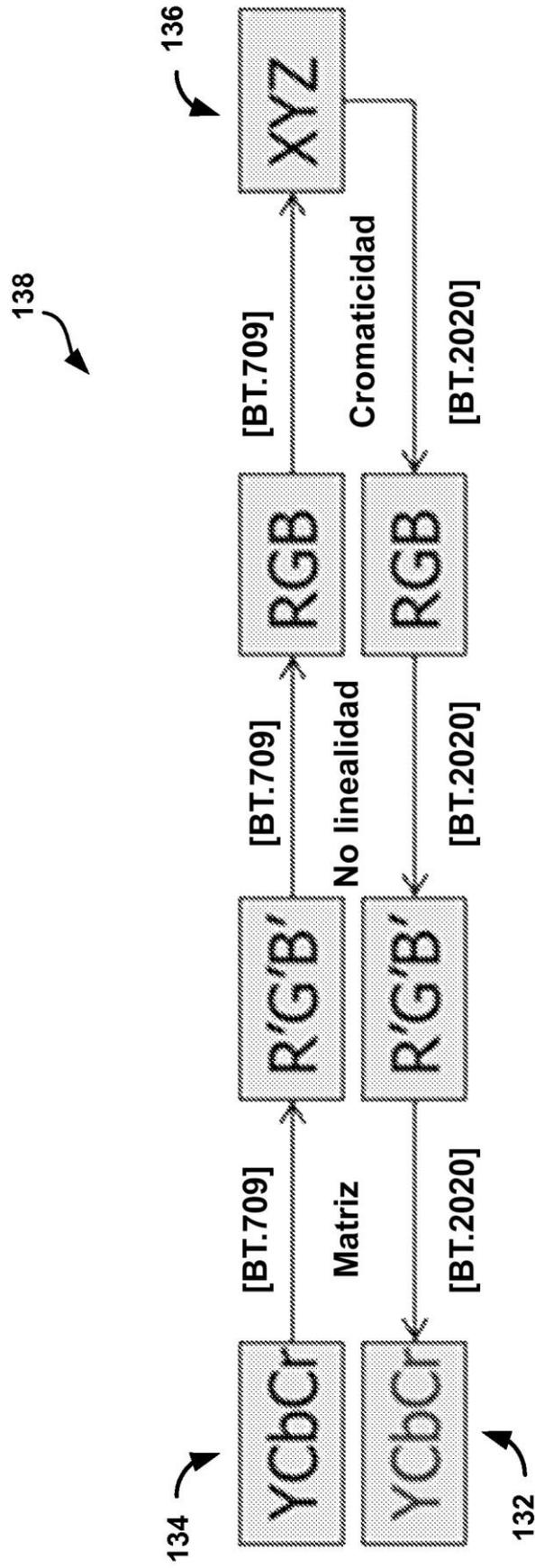


FIG. 7

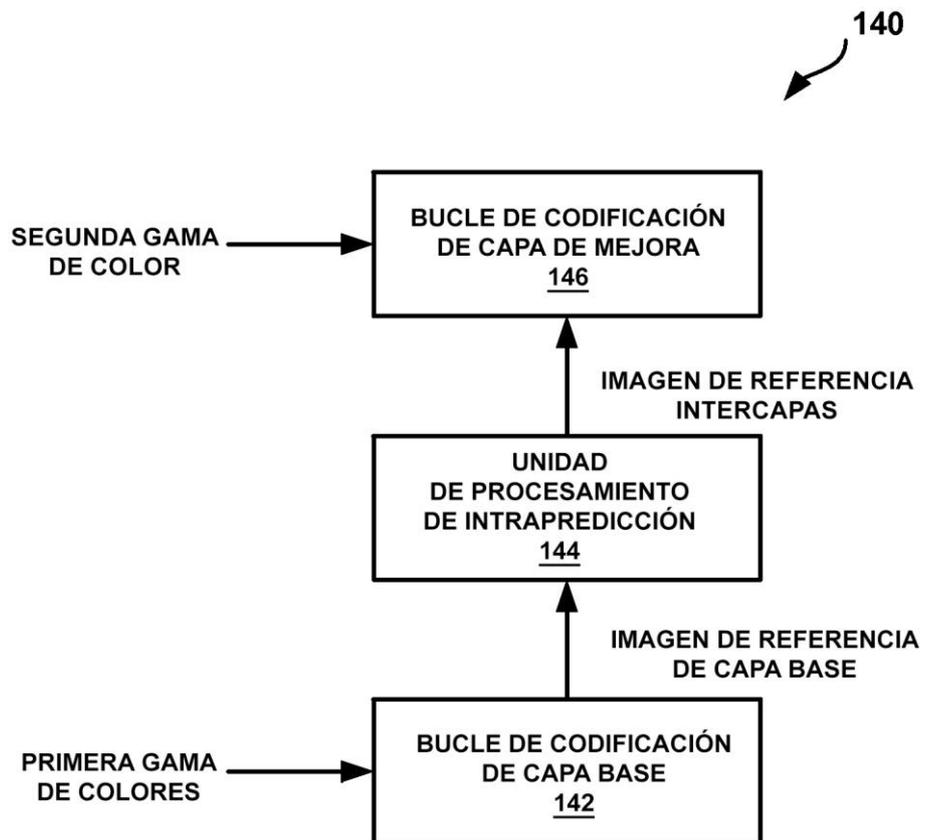


FIG. 8

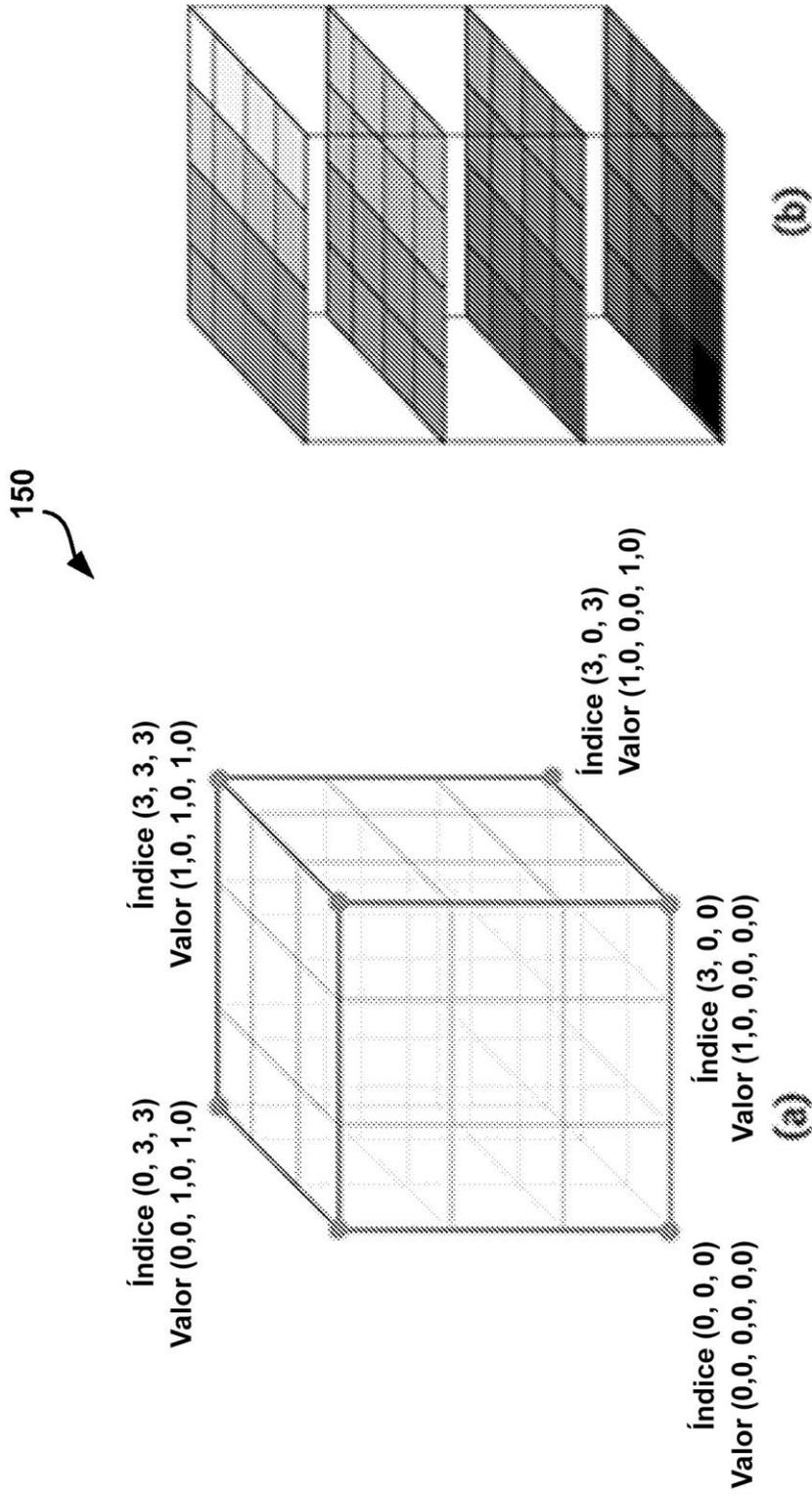


FIG. 9

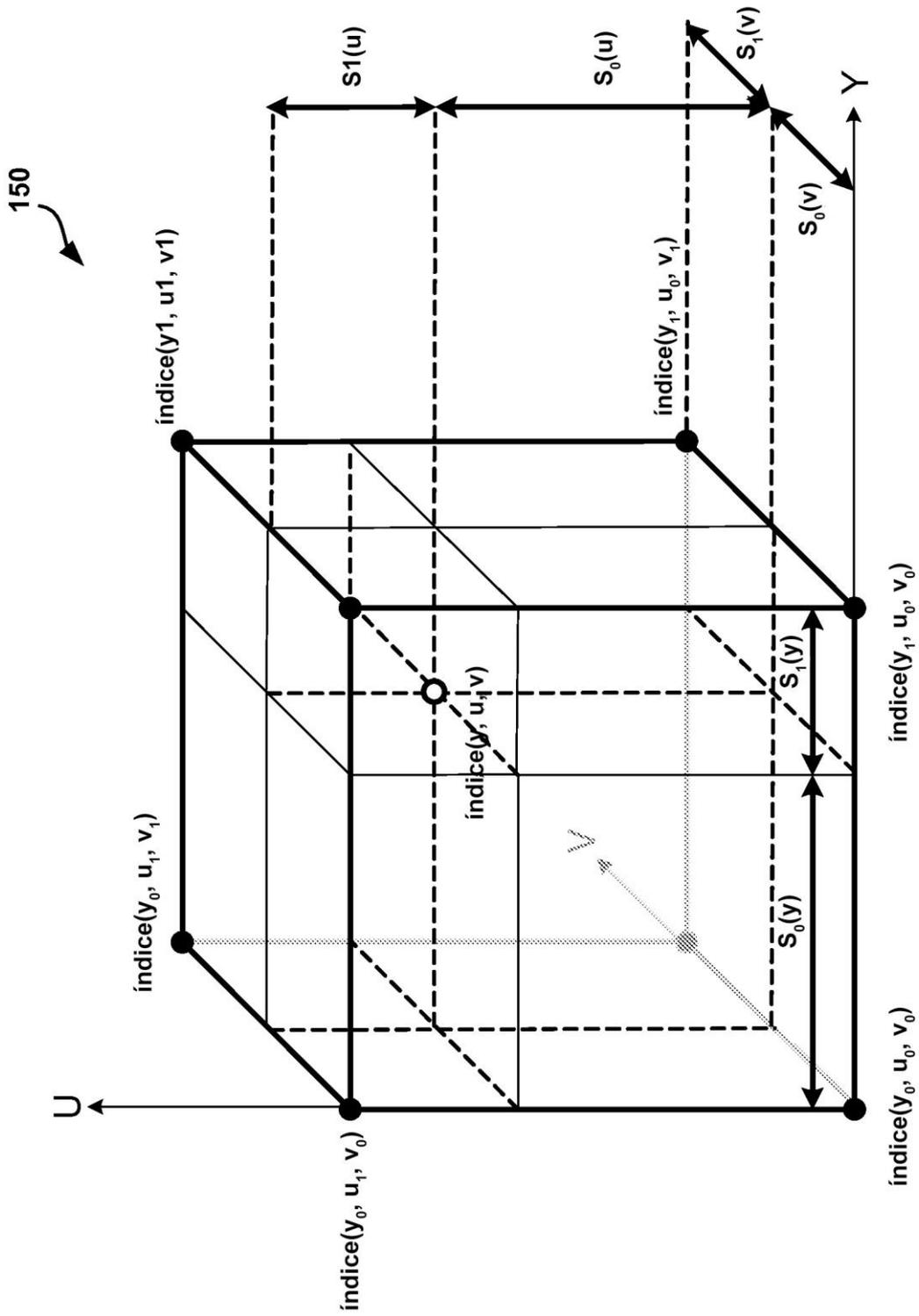


FIG. 10

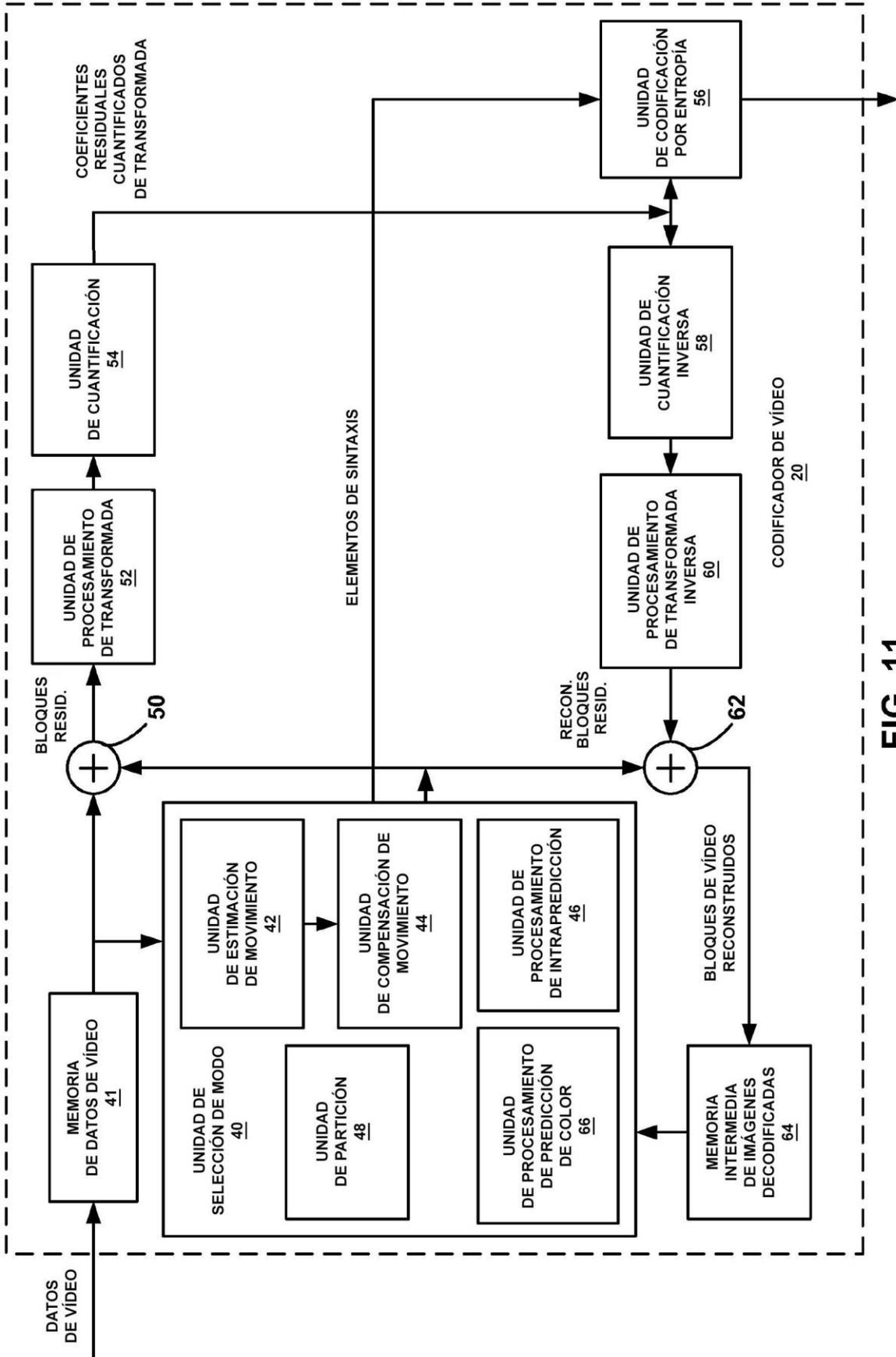


FIG. 11

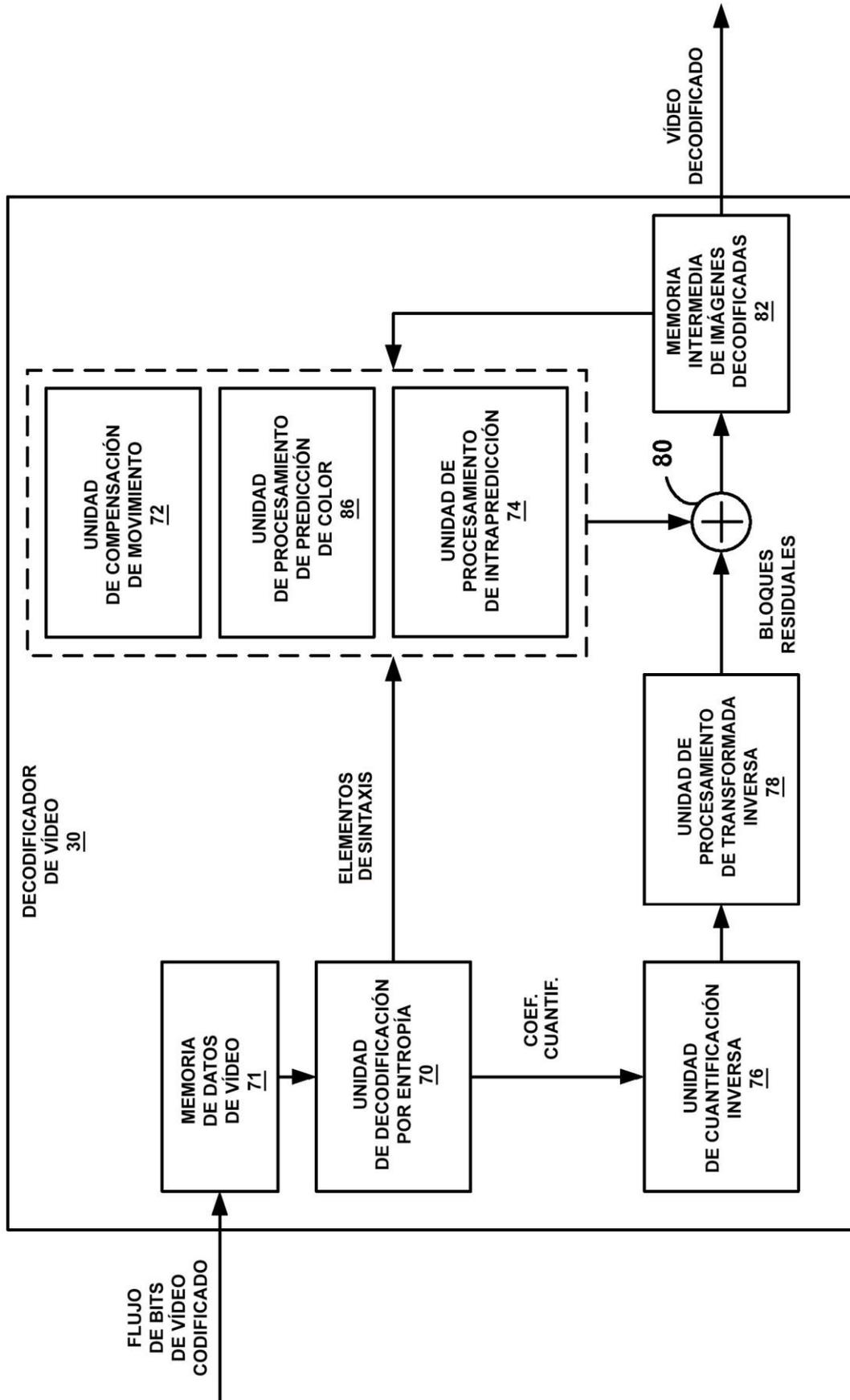


FIG. 12

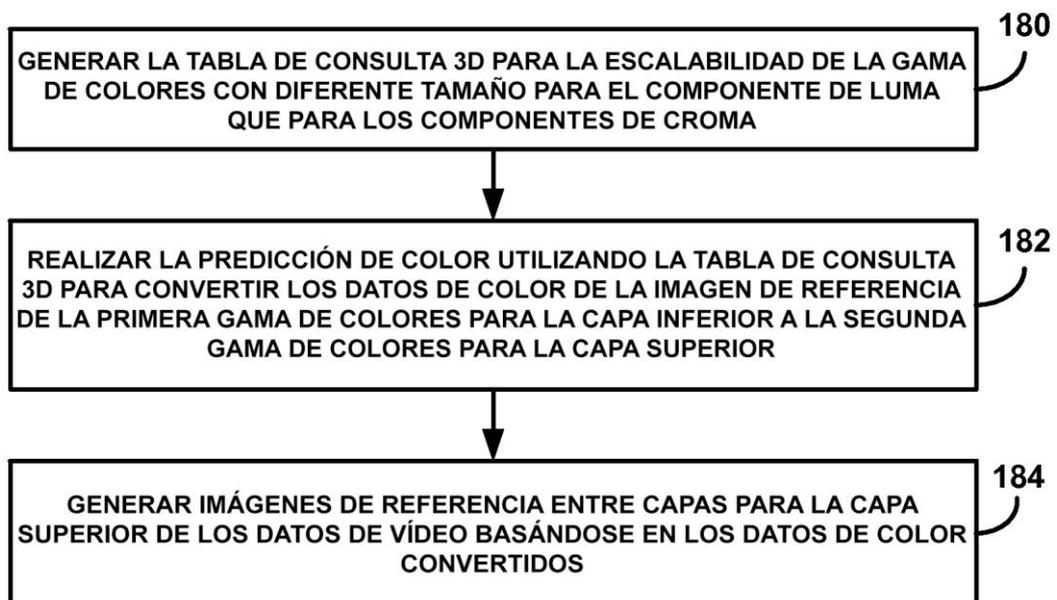


FIG. 13