

19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 769 837**

51 Int. Cl.:

H04N 19/33 (2014.01)

H04N 19/51 (2014.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

86 Fecha de presentación y número de la solicitud internacional: **27.02.2014 PCT/US2014/019083**

87 Fecha y número de publicación internacional: **04.09.2014 WO14134334**

96 Fecha de presentación y número de la solicitud europea: **27.02.2014 E 14711366 (6)**

97 Fecha y número de publicación de la concesión europea: **30.10.2019 EP 2962463**

54 Título: **Escalamiento del vector de movimiento espacial para codificación escalable de vídeo**

30 Prioridad:

01.03.2013 US 201361771742 P

07.03.2013 US 201361774516 P

26.02.2014 US 201414191311

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

29.06.2020

73 Titular/es:

QUALCOMM INCORPORATED (100.0%)

5775 Morehouse Drive

San Diego, CA 92121-1714, US

72 Inventor/es:

SEREGIN, VADIM;

GUO, LIWEI;

CHEN, JIANLE y

RAPAKA, KRISHNAKANTH

74 Agente/Representante:

FORTEA LAGUNA, Juan José

ES 2 769 837 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

Escalamiento del vector de movimiento espacial para codificación escalable de vídeo

5 ANTECEDENTES

Campo

10 **[0001]** La presente divulgación se refiere al campo de la codificación y compresión de vídeo. En particular, se refiere a la codificación de vídeo de alta eficacia (HEVC) y sus extensiones, por ejemplo, codificación escalable de vídeo (SVC), vídeo de múltiples vistas y codificación 3D (MVC, 3DV), etc. En algunos modos de realización, la divulgación se refiere al escalamiento del vector de movimiento espacial para SVC.

Descripción de la técnica relacionada

15 **[0002]** Las capacidades del vídeo digital se pueden incorporar en una amplia gama de dispositivos, incluyendo televisores digitales, sistemas de radiodifusión digital directa, sistemas de radiodifusión inalámbrica, asistentes digitales personales (PDA), ordenadores portátiles o de sobremesa, ordenadores de tableta, lectores de libros electrónicos, cámaras digitales, dispositivos de grabación digital, reproductores de medios digitales, dispositivos de videojuegos, consolas de videojuegos, teléfonos móviles o de radio por satélite, los denominados "teléfonos inteligentes", dispositivos de videoconferencia, dispositivos de transmisión continua de vídeo 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 avanzada de vídeo (AVC), la norma de codificación de vídeo de alta eficacia (HEVC), actualmente en desarrollo, y las extensiones de dichas normas. Los dispositivos de vídeo pueden transmitir, recibir, codificar, descodificar y/o almacenar información de vídeo digital más eficazmente, implementando dichas técnicas de codificación de vídeo.

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) se puede dividir en bloques de vídeo, que también se pueden denominar bloques arbolados, unidades de codificación (CU) y/o nodos de codificación. Los bloques de vídeo en un fragmento intracodificado (I) de una imagen se codifican usando 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 predicción espacial con respecto a muestras de referencia en bloques contiguos de la misma imagen, o predicción temporal con respecto a muestras de referencia en otras imágenes de referencia. Las imágenes se pueden denominar tramas y las imágenes de referencia se pueden denominar tramas de referencia.

40 **[0004]** La predicción espacial o temporal da como resultado un bloque predictivo para un bloque que se va a codificar. Los datos residuales representan diferencias de píxeles entre el bloque original que se va a codificar y el bloque predictivo. Un bloque intercodificado se codifica de acuerdo con un vector de movimiento que apunta a un bloque de muestras de referencia que forman el bloque predictivo y los datos residuales que indican la diferencia entre el bloque codificado y el bloque predictivo. Un bloque intracodificado se codifica de acuerdo con un modo de intracodificación y los datos residuales. Para una compresión adicional, los datos residuales se pueden transformar 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, organizados inicialmente en una matriz bidimensional, se pueden explorar para producir un vector unidimensional de coeficientes de transformada, y se puede aplicar codificación por entropía para lograr aún más compresión.

50 **[0005]** En algunas situaciones es deseable escalar espacialmente vectores de movimiento de capas que tienen una escalabilidad espacial diferente a la capa que se está codificando o descodificando. En otras situaciones, es deseable asignar la posición de un bloque de vídeo de capa a una posición equivalente en la capa que se está codificando o descodificando. Sin embargo, el escalamiento espacial y la asignación de posición entre capas, en general, requieren costes de procesamiento incrementados debido a operaciones de división arbitrarias realizadas para determinar un parámetro de escalamiento apropiado.

60 **[0006]** V Seregin en "TE5: Cross-verification of the test 5.2.5 on base layer MV candidate" - JCT-VC-L0207 - informa sobre la predicción de vectores de movimiento entre capas, en los que se realiza un escalamiento de MV espacial reutilizando la función de escalamiento TMVP. Chuang TD *et al.* en "NonCE9: Division-free MV scaling - JCT-VC-G223 - divulgan un escalamiento del vector de movimiento sin división para la derivación de predictores del vector de movimiento espacial y temporal. Yang H *et al.* en "Description of scalable video coding technology proposed by Huawei Technologies" propone una solución de codificación escalable de vídeo para una compresión eficaz de la capa de mejoramiento. Misra K *et al.* en "Description of scalable video coding technology proposal by Sharp (proposal 1)" propone una extensión de codificación escalable de vídeo para HEVC.

65 SUMARIO DE LA DIVULGACIÓN

5 [0007] En general, la presente divulgación describe técnicas relacionadas con codificación escalable de vídeo (SVC). Las técnicas descritas a continuación proporcionan un mecanismo para indicar si se permite la predicción de sintaxis entre capas (incluyendo la información de movimiento) para códecs de capa base específicos (por ejemplo, HEVC, etc.).

[0008] De acuerdo con un aspecto de la invención, se proporciona un procedimiento de descodificación de información de vídeo como se define en la reivindicación 1.

10 [0009] De acuerdo con un aspecto de la invención, se proporciona un procedimiento de codificación de información de vídeo como se define en la reivindicación 2.

15 [0010] De acuerdo con un aspecto de la invención, se proporciona un medio legible por ordenador no transitorio que incluye instrucciones que, cuando se ejecutan, hacen que un aparato realice el procedimiento de las reivindicaciones 1-7.

[0011] De acuerdo con un aspecto de la invención, se proporciona un aparato configurado para descodificar información de vídeo como se define en la reivindicación 9.

20 [0012] De acuerdo con un aspecto de la invención, se proporciona un aparato de acuerdo con la reivindicación 10.

25 [0013] Los detalles de uno o más ejemplos se exponen en los dibujos adjuntos y la descripción a continuación, que no pretenden limitar el alcance completo de los conceptos inventivos descritos en el presente documento. Otros rasgos característicos, objetivos y ventajas resultarán evidentes a partir de la descripción y los dibujos, y a partir de las reivindicaciones.

BREVE DESCRIPCIÓN DE LOS DIBUJOS

30 [0014] A lo largo de los dibujos, los números de referencia se pueden volver a usar para indicar correspondencia entre los elementos mencionados. Los dibujos se proporcionan para ilustrar modos de realización de ejemplo descritos en el presente documento y no pretenden limitar el alcance de la divulgación.

35 La FIG. 1A es un diagrama de bloques que ilustra un sistema de codificación y descodificación de vídeo de ejemplo que puede utilizar técnicas de acuerdo con los aspectos descritos en la presente divulgación.

La FIG. 2 es un diagrama de bloques que ilustra un ejemplo de un codificador de vídeo que puede implementar técnicas de acuerdo con aspectos descritos en la presente divulgación.

40 La FIG. 3 es un diagrama de bloques que ilustra un ejemplo de un descodificador de vídeo que puede implementar técnicas de acuerdo con aspectos descritos en la presente divulgación.

45 La FIG. 4 es un diagrama de flujo que ilustra un modo de realización de un procedimiento de realización de escalamiento espacial usando una función de escalamiento temporal de acuerdo con aspectos de la presente divulgación.

DESCRIPCIÓN DETALLADA DE MODOS DE REALIZACIÓN ESPECÍFICOS

50 [0015] Las técnicas descritas en la presente divulgación se refieren en general a codificación escalable de vídeo (SVC) y codificación de vídeo de múltiples vistas/3D. Por ejemplo, las técnicas se pueden referir a, y usarse con o en, una extensión de codificación escalable de vídeo (SVC, a veces denominada SHVC) de codificación de vídeo de alta eficacia (HEVC). En una extensión de SVC, podría haber múltiples capas de información de vídeo. La capa en el nivel más bajo puede servir como capa base (BL), y la capa en la parte superior (o la capa más alta) puede servir como capa potenciada (EL). La "capa potenciada" a veces se denomina "capa de mejoramiento", y estos términos se pueden usar de manera intercambiable. La capa base a veces se denomina "capa de referencia", (RL) y estos términos también se pueden usar de manera intercambiable. Todas las capas entre la capa base y la capa superior pueden servir como cualquiera o ambas EL o capas de referencia (RL). Por ejemplo, una capa central puede ser una EL para las capas por debajo de la misma, tal como la capa base o cualquier capa de mejoramiento intermedia, y al mismo tiempo servir como una RL para las capas de mejoramiento por encima de la misma. Cada capa entre la capa base y la capa superior (o la capa más alta) se puede usar como referencia para la predicción entre capas por una capa más alta y puede usar una capa más baja como referencia para la predicción entre capas.

65 [0016] Para simplificar, los ejemplos se presentan en lo que respecta a únicamente dos capas: una BL y una EL; sin embargo, se debe entender bien que las ideas y modos de realización descritos a continuación son aplicables también a casos con múltiples capas. Además, para facilitar la explicación, a menudo se usan los términos "tramas" o "bloques". Sin embargo, estos términos no pretenden ser limitantes. Por ejemplo, las técnicas descritas a continuación se pueden

usar con cualquiera de una variedad de unidades de vídeo, incluyendo pero sin limitarse a píxeles, bloques (por ejemplo, CU, PU, TU, macrobloques, etc.), fragmentos, imágenes, etc.

Codificación de vídeo

[0017] Las normas de codificación de vídeo incluyen ITU-T H.261, ISO/IEC MPEG-1 Visual, ITU-T H.262 o ISO/IEC MPEG-2 Visual, ITU-T H.263, ISO/IEC MPEG-4 Visual e ITU-T H.264 (también conocida como ISO/IEC MPEG-4 AVC), incluyendo sus extensiones de codificación de vídeo ajustable a escala (SVC) y de codificación de vídeo de múltiples vistas (MVC). El último borrador conjunto de SVC y MVC se describe en "Advanced video coding for generic audiovisual services", recomendación ITU-T H.264, marzo de 2010.

[0018] Además, existe una nueva norma de codificación de vídeo, la codificación de vídeo de alta eficacia (HEVC), que se está desarrollando por el Equipo de Colaboración Conjunta sobre Codificación de Vídeo (JCT-VC) del Grupo de Expertos sobre Codificación de Vídeo (VCEG) de ITU-T y el Grupo de Expertos sobre Imágenes en Movimiento (MPEG) de ISO/IEC. Un borrador de memoria descriptiva textual de codificación de vídeo de alta eficacia (HEVC) reciente está disponible en http://phenix.int-evry.fr/jct/doc_end_user/documents/12_Geneva/wg11/JCTVC-L1003-v13.zip. Otro borrador de trabajo (WD) reciente de HEVC y denominado HEVC WD9, está disponible en http://phenix.int-evry.fr/jct/doc_end_user/documents/9_Geneva/wg11/JCTVC-L1003-v5.zip.

[0019] En SVC, se puede proporcionar la información de vídeo como múltiples capas. La capa en el nivel más bajo puede servir únicamente como una capa base (BL) y la capa en el nivel más alto puede servir como una capa de mejoramiento (EL). Todas las capas entre las capas superior e inferior pueden servir tanto como capas de mejoramiento como capas base. Por ejemplo, una capa central puede ser una EL para las capas por debajo de la misma, y al mismo tiempo como una BL para las capas por encima de la misma. Para simplificar la descripción, se puede suponer que existen dos capas, una BL y una EL, para ilustrar las técnicas descritas a continuación. Sin embargo, todas las técnicas descritas en el presente documento también son aplicables a casos con múltiples (más de dos) capas.

[0020] La codificación escalable de vídeo (SVC) se puede usar para proporcionar escalabilidad de calidad (también denominada relación de señal a ruido (SNR)), escalabilidad espacial y/o escalabilidad temporal. Por ejemplo, en un modo de realización, una capa de referencia (por ejemplo, una capa base) incluye información de vídeo suficiente para mostrar un vídeo en un primer nivel de calidad y la capa de mejoramiento incluye información de vídeo adicional con respecto a la capa de referencia de modo que la capa de referencia y la capa de mejoramiento incluyen conjuntamente información de vídeo suficiente para mostrar el vídeo en un segundo nivel de calidad superior al primer nivel (por ejemplo, menos ruido, mayor resolución, mejor velocidad de tramas, etc.). Una capa potenciada puede tener una resolución espacial diferente a una capa base. Por ejemplo, la proporción de aspecto espacial entre la EL y la BL puede ser 1,0, 1,5, 2,0 u otras proporciones diferentes. En otras palabras, el aspecto espacial de la EL puede ser igual a 1,0, 1,5 o 2,0 veces el aspecto espacial de la BL. En algunos ejemplos, el factor de escalamiento de la EL puede ser mayor que el de la BL. Por ejemplo, un tamaño de imágenes en la EL puede ser más grande que un tamaño de imágenes en la BL. De esta manera, puede ser posible, aunque no es una limitación, que la resolución espacial de la EL sea mayor que la resolución espacial de la BL.

En la extensión de SVC para H.264, la predicción de un bloque actual se puede realizar usando las diferentes capas que se proporcionan para SVC. Dicha predicción se puede denominar predicción entre capas. Los procedimientos de predicción entre capas se pueden utilizar en SVC para reducir la redundancia entre capas. Algunos ejemplos de predicción entre capas pueden incluir la intrapredicción entre capas, la predicción de movimiento entre capas y la predicción residual entre capas. La intrapredicción entre capas usa la reconstrucción de bloques ubicados conjuntamente en la capa base para predecir el bloque actual en la capa de mejoramiento. La predicción de movimiento entre capas usa la información de movimiento (incluyendo los vectores de movimiento) de la capa base para predecir el movimiento en la capa de mejoramiento. La predicción residual entre capas usa el residuo de la capa base para predecir el residuo de la capa de mejoramiento.

[0021] En algunos modos de realización de predicción de movimiento entre capas, se pueden usar los datos de movimiento (incluyendo vectores de movimiento) de la capa base (por ejemplo, para el bloque ubicado conjuntamente) para predecir el bloque actual en la capa de mejoramiento. Por ejemplo, mientras codifican una unidad de vídeo en una capa de mejoramiento, los codificadores de vídeo pueden usar información de una capa de referencia para obtener datos de compensación de movimiento adicionales que se pueden usar para identificar hipótesis adicionales. Como estas hipótesis adicionales se derivan implícitamente de los datos ya existentes en un flujo de bits de vídeo, se puede obtener un rendimiento adicional en la codificación de vídeo con poco o ningún coste adicional en el tamaño del flujo de bits. En otro ejemplo, se puede usar la información de movimiento de unidades de vídeo espacialmente contiguas para ubicar hipótesis adicionales. A continuación, la hipótesis derivada se puede promediar o combinar de otro modo con una hipótesis codificada explícitamente para generar una mejor predicción de un valor de una unidad de vídeo. En determinadas situaciones, tales como cuando la resolución espacial de la capa base (o de referencia) es diferente a la resolución espacial de la capa del bloque actual, la información de movimiento de la capa base se escala espacialmente antes de usarse para codificar o descodificar el bloque actual. Del mismo modo, la posición de un

bloque en una capa base (o de referencia) se puede determinar mediante la asignación de posición entre capas, como se describe a continuación, cuando la capa tiene una resolución espacial diferente a la capa del bloque actual.

Escalamiento del vector de movimiento espacial

[0022] Un vector de movimiento (MV) de una capa base se puede usar como un candidato de MV en al menos modos AMVP o Merge. Debido a que la capa base y las capas de mejoramiento pueden tener potencialmente diferentes resoluciones, puede ser necesario que el MV de la capa base se escale de acuerdo con la diferencia de resolución antes de usarse como candidatos de predicción de MV entre capas. Dicho escalamiento se expresa por:

$$MV = \{MV_x, MV_y\},$$

$$\begin{cases} MV_x = blMV_x \times \frac{elW}{blW} \\ MV_y = blMV_y \times \frac{elH}{blH} \end{cases}$$

donde *el* indica la capa de mejoramiento, *bl* indica la capa base, *W* indica el ancho y *H* indica la altura.

[0023] Las fórmulas anteriores se proporcionan con propósitos ilustrativos, y se pueden usar fórmulas similares. Por ejemplo, se pueden usar fórmulas similares donde, por ejemplo, se añade desviación de redondeo.

Asignación de posición entre capas

[0024] La asignación de posición entre una imagen de capa base y una imagen de capa de mejoramiento se realiza para la predicción de textura y/o sintaxis entre capas. Por ejemplo, en la predicción de textura entre capas, para un píxel en la posición (x, y) de la capa de mejoramiento (que se puede denominar PEL(x,y)), se puede derivar una posición correspondiente (blx, bly). Por lo tanto, se puede usar un valor de píxel en (blx, bly) de la capa base, que se puede denominar PBL (blx, bly), para predecir el valor de PEL(x,y). La asignación de posición se puede expresar como:

$$\begin{cases} blx = x \times \frac{blW}{elW} \\ bly = y \times \frac{blH}{elH} \end{cases}$$

[0025] Las fórmulas anteriores se usan para propósito ilustrativo, se pueden usar fórmulas similares donde, por ejemplo, se puede añadir desviación de redondeo. Si la imagen de capa base y la imagen de capa de mejoramiento tienen un recorte diferente, los parámetros de recorte se pueden integrar en la fórmula. Sin embargo, una desventaja de usar las fórmulas anteriores es que se requiere una operación de división para proporciones de aspecto de escalabilidad de capa base y de mejoramiento arbitrarias. Una operación de división de este tipo es informáticamente costosa y requiere amplios recursos informáticos y/o mayores requisitos de memoria y ancho de banda, como se analiza en mayor detalle a continuación.

[0026] Se ha propuesto que la función de escalamiento del predictor del vector de movimiento temporal (TMVP) disponible en la norma HEVC se puede usar para el escalamiento de MV de la capa base. El parámetro de entrada para esta función es la variable escalada, calculada como $iScaleBase = (1 \ll 8) * \frac{elW}{blW}$, que es equivalente a $iScaleBase = 256 \times elW/blW$ (donde *elW* y *blW* corresponden a los anchos de la imagen de capa de mejoramiento y la imagen de capa base, respectivamente). La expresión $x \ll y$ se refiere a un desplazamiento a la izquierda de la representación binaria del número *x* en *y* bits, lo que equivale a dividir el número *x* por 2^y . De forma similar, la expresión $x \gg y$ se refiere a un desplazamiento a la derecha de la representación binaria del número *x* en *y* bits, lo que equivale a multiplicar el número *x* por 2^y . Para escalamiento nulo, por ejemplo, cuando los tamaños de imagen de capa base y de capa de mejoramiento son iguales, el parámetro de entrada *iScaleBase* es igual a 256. Para una escalabilidad de 2x (por ejemplo, cuando el ancho de la imagen de capa de mejoramiento es dos veces el ancho de la imagen de capa base), el parámetro de entrada *iScaleBase* es igual a 512, y para 1,5x el parámetro de entrada *iScaleBase* es igual a 384. Este parámetro de entrada se puede calcular una vez por secuencia. Sin embargo, como se menciona anteriormente, se requiere la división con un divisor arbitrario, que puede incrementar el coste y/o disminuir la eficacia del procedimiento de codificación y decodificación.

[0027] Diversos aspectos de los sistemas, aparatos y procedimientos novedosos se describen más adelante en el presente documento más completamente, con referencia a los dibujos adjuntos. Sin embargo, la presente divulgación se puede realizar de muchas formas diferentes y no se debería interpretar que está limitada a ninguna estructura o función específica presentada a lo largo de la presente divulgación. En cambio, estos aspectos se proporcionan de

modo que la presente divulgación sea exhaustiva y completa, y transmita por completo el alcance de la divulgación a los expertos en la técnica. En base a las enseñanzas del presente documento, un experto en la técnica debería apreciar que el alcance de la divulgación está concebido para cubrir cualquier aspecto de los sistemas, aparatos y procedimientos novedosos divulgados en el presente documento, ya sea implementados independientemente de, o en combinación con, cualquier otro aspecto de la invención. Por ejemplo, un aparato se puede implementar o un procedimiento se puede llevar a la práctica usando cualquier número de los aspectos expuestos en el presente documento. Además, el alcance de la invención está concebido para abarcar un aparato o procedimiento de este tipo, que se lleva a la práctica usando otra estructura, funcionalidad, o estructura y funcionalidad además de o distinta de los diversos aspectos de la invención expuestos en el presente documento. Se debe entender que cualquier aspecto divulgado en el presente documento se puede realizar mediante uno o más elementos de una reivindicación.

[0028] Aunque en el presente documento se describen aspectos particulares, muchas variaciones y permutaciones de estos aspectos se encuentran dentro del alcance de la divulgación. Aunque se mencionan algunos beneficios y ventajas de los aspectos preferentes, el alcance de la divulgación no está concebido para limitarse a beneficios, usos u objetivos particulares. En cambio, los aspectos de la divulgación están concebidos para ser ampliamente aplicables a diferentes tecnologías inalámbricas, configuraciones de sistema, redes y protocolos de transmisión, algunos de los que se ilustran a modo de ejemplo en las figuras y en la siguiente descripción de los aspectos preferentes. La descripción detallada y los dibujos son meramente ilustrativos de la divulgación, en lugar de limitantes, estando definido el alcance de la divulgación por las reivindicaciones adjuntas y equivalentes de las mismas.

Sistema de codificación de vídeo

[0029] La FIG. 1 es un diagrama de bloques que ilustra un sistema de codificación de vídeo 10 de ejemplo que puede utilizar técnicas de acuerdo con aspectos descritos en la presente divulgación. Como se usa descrito en el presente documento, el término "codificador de vídeo" se refiere, en general, tanto a codificadores de vídeo como a descodificadores de vídeo. En la presente divulgación, los términos "codificación de vídeo" o "codificación" se pueden referir genéricamente a la codificación de vídeo y a la descodificación de vídeo.

[0030] Como se muestra en la FIG. 1, el sistema de codificación de vídeo 10 incluye un dispositivo de origen 12 y un dispositivo de destino 14. El dispositivo de origen 12 genera datos de vídeo codificados. El dispositivo de destino 14 puede descodificar los datos de vídeo codificados generados por el dispositivo de origen 12. El dispositivo de origen 12 puede proporcionar los datos de vídeo al dispositivo de destino 14 por medio de un medio legible por ordenador 16. El dispositivo de origen 12 y el dispositivo de destino 14 pueden incluir una amplia gama de dispositivos, incluyendo ordenadores de sobremesa, ordenadores plegables (por ejemplo, portátiles), ordenadores de tableta, descodificadores, teléfonos portátiles, tales como los denominados teléfonos "inteligentes", los denominados paneles "inteligentes", televisores, cámaras, dispositivos de visualización, reproductores de medios digitales, consolas de videojuegos, ordenadores de coche, dispositivos de transmisión continua de vídeo o similares. El dispositivo de origen 12 y el dispositivo de destino 14 pueden estar equipados para la comunicación inalámbrica.

[0031] El dispositivo de destino 14 puede recibir los datos de vídeo codificados que se van a descodificar por medio del medio legible por ordenador 16. El medio legible por ordenador 16 puede comprender un tipo de medio o dispositivo que puede transportar los datos de vídeo codificados desde el dispositivo de origen 12 al dispositivo de destino 14. Por 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 transmitir al dispositivo de destino 14. El medio de comunicación puede comprender un medio de comunicación inalámbrica o cableada, tal como un espectro de radiofrecuencia (RF) o una o más líneas de transmisión físicas. 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 enrutadores, conmutadores, estaciones base u otro equipo que pueda ser útil para facilitar la comunicación desde el dispositivo de origen 12 al dispositivo de destino 14.

[0032] En algunos modos de realización, se pueden emitir datos codificados desde la interfaz de salida 22 a un dispositivo de almacenamiento. De forma similar, se puede acceder a los datos codificados desde el dispositivo de almacenamiento por una interfaz de entrada. El dispositivo de almacenamiento puede incluir cualquiera de una variedad de 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 u otros medios de almacenamiento digital para almacenar datos de vídeo. El dispositivo de almacenamiento puede corresponder a un servidor de archivos u otro dispositivo de almacenamiento intermedio que puede 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 por medio de transmisión continua o descarga. El servidor de archivos puede ser un tipo de servidor que pueda 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 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 una conexión de datos estándar, incluyendo una conexión a Internet. Esto puede incluir un canal inalámbrico (por ejemplo, una conexión wifi), una conexión cableada (por

ejemplo, DSL, módem de cable, etc.) o una combinación de ambas 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 las mismas.

5 **[0033]** Las técnicas de la presente divulgación pueden aplicar aplicaciones o configuraciones además de las aplicaciones o configuraciones inalámbricas. Las técnicas se pueden aplicar a la codificación de vídeo como soporte de una variedad de aplicaciones multimedia, tales como difusiones de televisión por el aire, transmisiones de televisión por cable, transmisiones de televisión por satélite, transmisiones de vídeo continuas por Internet, tales como la
10 transmisión continua adaptativa dinámica sobre HTTP (DASH), vídeo digital que se codifica en un medio de almacenamiento de datos, descodificación de vídeo digital almacenado en un medio de almacenamiento de datos u otras aplicaciones. En algunos modos de realización, el sistema 10 se puede configurar para admitir la transmisión de vídeo unidireccional o bidireccional para admitir aplicaciones tales como la transmisión continua de vídeo, la reproducción de vídeo, la radiodifusión de vídeo y/o la videotelefonía.

15 **[0034]** En la FIG. 1, el dispositivo de origen 12 incluye un origen de vídeo 18, un codificador de vídeo 20 y una interfaz de salida 22. El dispositivo de destino 14 incluye una interfaz de entrada 28, un descodificador de vídeo 30 y un dispositivo de visualización 32. El codificador de vídeo 20 del dispositivo de origen 12 se puede configurar para aplicar las técnicas para codificar un flujo de bits que incluye datos de vídeo que se ajustan a múltiples normas o extensiones
20 de normas. En otros modos de realización, 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 un origen de vídeo externo 18, tal 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.

25 **[0035]** El origen de vídeo 18 del dispositivo de origen 12 puede incluir un dispositivo de captación de vídeo, tal como una videocámara, un archivo de vídeo que contiene vídeo captado previamente y/o una interfaz de alimentación de vídeo para recibir vídeo desde un proveedor de contenidos de vídeo. El origen de vídeo 18 puede generar datos basados en gráficos de ordenador como vídeo de origen, o una combinación de vídeo en directo, vídeo archivado y vídeo generado por ordenador. En algunos modos de realización, si el origen de vídeo 18 es una videocámara, el
30 dispositivo de origen 12 y el dispositivo de destino 14 pueden formar los denominados teléfonos con cámara o videoteléfonos. El vídeo captado, precaptado o generado por ordenador se puede codificar por el codificador de vídeo 20. La información de vídeo codificada se puede emitir por la interfaz de salida 22 a un medio legible por ordenador 16.

35 **[0036]** 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 (por ejemplo, medios de almacenamiento no transitorio), 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. Un servidor de red (no mostrado) puede recibir datos de vídeo codificados desde el dispositivo de origen 12 y proporcionar los datos de vídeo codificados al dispositivo de destino
40 14 (por ejemplo, por medio de transmisión por red). Un dispositivo informático de una instalación de producción de un medio, tal como una instalación 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.

45 **[0037]** La interfaz de entrada 28 del dispositivo de destino 14 puede recibir información desde el medio legible por ordenador 16. La información del medio legible por ordenador 16 puede incluir información sintáctica definida por el codificador de vídeo 20, que se puede usar por el descodificador de vídeo 30, que incluye elementos sintácticos que describen características y/o procesamiento de bloques y otras unidades codificadas, por ejemplo, los GOP. El
50 dispositivo de visualización 32 muestra los datos de vídeo descodificados a un usuario, y puede incluir cualquiera de una variedad de 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.

55 **[0038]** El codificador de vídeo 20 y el descodificador de vídeo 30 pueden funcionar de acuerdo con una norma de codificación de vídeo, tal como la norma de codificación de vídeo de alta eficacia (HEVC), actualmente en fase de desarrollo, y se pueden ajustar al modelo de prueba de la HEVC (HM). De forma alternativa, el codificador de vídeo 20 y el descodificador de vídeo 30 pueden funcionar de acuerdo con otras normas privadas o industriales, tales como la norma ITU-T H.264, denominada de forma alternativa MPEG-4, parte 10, codificación avanzada de vídeo (AVC), o extensiones de dichas normas. Sin embargo, las técnicas de la presente divulgación no están limitadas a ninguna
60 norma de codificación particular. Otros ejemplos de normas de codificación de vídeo incluyen MPEG-2 e ITU-T H.263. Aunque no se muestra en la FIG. 1, en algunos aspectos, el codificador de vídeo 20 y el descodificador de vídeo 30 pueden estar cada uno integrados con un codificador y descodificador de audio, y pueden incluir unidades MUX-DEMUX adecuadas, u otro hardware y software, para manejar la codificación tanto de audio como de vídeo en un flujo de datos común o en flujos de datos separados. Si fuera aplicable, las unidades MUX-DEMUX se pueden ajustar al
65 protocolo multiplexor ITU H.223 o a otros protocolos tales como el protocolo de datagramas de usuario (UDP).

5 [0039] El codificador de vídeo 20 y el descodificador de vídeo 30 se pueden implementar cada uno como cualquiera de una variedad de circuitos codificadores adecuados, tales como uno o más microprocesadores, procesadores de señales digitales (DSP), circuitos integrados específicos de la aplicación (ASIC), matrices de puertas programables *in situ* (FPGA), lógica discreta, software, hardware, firmware o cualquier combinación de los mismos. Cuando las técnicas se implementan parcialmente en software, un dispositivo puede almacenar instrucciones para el software en un medio legible por ordenador no transitorio y ejecutar las instrucciones en hardware usando uno o más procesadores para realizar las técnicas de la presente divulgación. El codificador de vídeo 20 y el descodificador de vídeo 30 pueden estar incluidos cada uno en uno o más codificadores o descodificadores, cualquiera de los que puede estar integrado como parte de un codificador/descodificador (CÓDEC) combinado en un dispositivo respectivo. Un dispositivo que incluye el codificador de vídeo 20 y/o el descodificador de vídeo 30 puede comprender un circuito integrado, un microprocesador y/o un dispositivo de comunicación inalámbrica, tal como un teléfono móvil.

15 [0040] El JCT-VC está trabajando en el desarrollo de la norma HEVC. Los esfuerzos de normalización de la HEVC se basan en un modelo en evolución de un dispositivo de codificación de vídeo denominado modelo de prueba de la HEVC (HM). El HM supone varias capacidades adicionales de los dispositivos de codificación de vídeo con respecto a los dispositivos existentes de acuerdo, por ejemplo, con ITU-T H.264/AVC. Por ejemplo, mientras que H.264 proporciona nueve modos de codificación por intrapredicción, el HM puede proporcionar hasta treinta y tres modos de codificación por intrapredicción.

20 [0041] En general, el modelo de funcionamiento del HM describe que una trama o imagen de vídeo se puede dividir en una secuencia de bloques arbolados o unidades de codificación de máximo tamaño (LCU), que incluyen muestras tanto de luma como de croma. Los datos sintácticos dentro de un flujo de bits pueden definir un tamaño para la LCU, que es la unidad de codificación más grande en lo que respecta al número de píxeles. Un fragmento incluye un número de bloques arbolados consecutivos en orden de codificación. Una trama o imagen de vídeo se puede dividir en uno o más fragmentos. Cada bloque arbolado se puede separar en unidades de codificación (CU) de acuerdo con un árbol cuaternario. En general, una estructura de datos de árbol cuaternario incluye un nodo por CU, con un nodo raíz correspondiente al bloque arbolado. Si una CU se separa en cuatro sub-CU, el nodo correspondiente a la CU incluye cuatro nodos hoja, de los que cada uno corresponde a una de las sub-CU.

30 [0042] Cada nodo de la estructura de datos de árbol cuaternario puede proporcionar datos sintácticos para la CU correspondiente. Por ejemplo, un nodo en el árbol cuaternario puede incluir un indicador de separación, que indica si la CU correspondiente al nodo se separa en sub-CU. Los elementos sintácticos para una CU se pueden definir de manera recursiva y pueden depender de si la CU se separa en sub-CU. Si una CU no se separa adicionalmente, se denomina CU hoja. En la presente divulgación, cuatro sub-CU de una CU hoja también se denominarán CU hojas, incluso si no existe ninguna separación explícita de la CU hoja original. Por ejemplo, si una CU de tamaño 16x16 no se separa adicionalmente, las cuatro sub-CU de 8x8 también se denominarán CU hojas aunque la CU de 16x16 nunca se separara.

40 [0043] Una CU tiene un propósito similar a un macrobloque de la norma H.264, excepto que una CU no tiene una distinción de tamaño. Por ejemplo, un bloque arbolado se puede separar en cuatro nodos hijos (también denominados sub-CU) y cada nodo hijo puede a su vez ser un nodo padre y separarse en otros cuatro nodos hijos. Un nodo hijo final, no separado, denominado nodo hoja del árbol cuaternario, comprende un nodo de codificación, también denominado una CU hoja. Los datos sintácticos asociados a un flujo de bits codificado pueden definir un número máximo de veces en que se puede separar un bloque arbolado, denominado profundidad de CU máxima, y también pueden definir un tamaño mínimo de los nodos de codificación. En consecuencia, un flujo de bits también puede definir una unidad de codificación más pequeña (SCU). La presente divulgación usa el término "bloque" para referirse a cualquiera de una CU, PU o TU, en el contexto de la HEVC, o a estructuras de datos similares en el contexto de otras normas (por ejemplo, macrobloques y subbloques de los mismos en la H.264/AVC).

50 [0044] Una CU incluye un nodo de codificación y unidades de predicción (PU) y unidades de transformada (TU) asociadas al nodo de codificación. Un tamaño de la CU corresponde a un tamaño del nodo de codificación y debe tener conformación cuadrada. El tamaño de la CU puede variar desde 8x8 píxeles hasta el tamaño del bloque arbolado, con un máximo de 64x64 píxeles o mayor. Cada CU puede contener una o más PU y una o más TU. Los datos sintácticos asociados a una CU pueden describir, por ejemplo, la división de la CU en una o más PU. Los modos de división pueden diferir entre si la CU está codificada en modo de salto o directo, codificada en modo de intrapredicción o codificada en modo de interpretación. Las PU se pueden dividir para tener una conformación no cuadrada. Los datos sintácticos asociados a una CU también pueden describir, por ejemplo, la división de la CU en una o más TU de acuerdo con un árbol cuaternario. Una TU puede tener una conformación cuadrada o no cuadrada (por ejemplo, rectangular).

60 [0045] La norma HEVC permite transformaciones de acuerdo con las TU, que pueden ser diferentes para diferentes CU. El tamaño de las TU se basa típicamente en el tamaño de las PU dentro de una CU dada definida para una LCU dividida, aunque puede que no sea siempre así. Las TU son típicamente del mismo tamaño o más pequeñas que las PU. En algunos ejemplos, las muestras residuales correspondientes a una CU se pueden subdividir en unidades más pequeñas usando una estructura de árbol cuaternario conocida como "árbol cuaternario residual" (RQT). Los nodos

hoja del RQT se pueden denominar unidades de transformada (TU). Los valores de diferencias de píxeles asociados a las TU se pueden transformar para producir coeficientes de transformada, que se pueden cuantificar.

[0046] Una CU hoja puede incluir una o más unidades de predicción (PU). En general, una PU representa una zona espacial correspondiente a la totalidad, o a una parte, de la CU correspondiente, y puede incluir datos para recuperar una muestra de referencia para la PU. Además, una PU incluye datos relacionados con la predicción. Por ejemplo, cuando la PU se codifica mediante intramodo se pueden incluir datos para la PU en un árbol cuaternario residual (RQT), que puede incluir datos que describen un modo de intrapredicción para una TU correspondiente a la PU. Como otro ejemplo, cuando la PU se codifica mediante intermodo, la PU puede incluir datos que definen uno o más vectores de movimiento para la PU. Los datos que definen el vector de movimiento para una PU pueden describir, por ejemplo, una componente horizontal del vector de movimiento, una componente vertical del vector de movimiento, una resolución para el vector de movimiento (por ejemplo, precisión de un cuarto de píxel o precisión de un octavo de píxel), una imagen de referencia a la que apunta el vector de movimiento y/o una lista de imágenes de referencia (por ejemplo, lista 0, lista 1 o lista C) para el vector de movimiento.

[0047] Una CU hoja que tiene una o más PU también puede incluir una o más unidades de transformada (TU). Las unidades de transformada se pueden especificar usando un RQT (también denominado estructura de árbol cuaternario de TU), como se analiza anteriormente. Por ejemplo, un indicador de separación puede indicar si una CU hoja está separada en cuatro unidades de transformada. A continuación, cada unidad de transformada se puede separar adicionalmente en sub-TU adicionales. Cuando una TU no se separa adicionalmente, se puede denominar una TU hoja. En general, para la intracodificación, todas las TU hojas que pertenecen a una CU hoja comparten el mismo modo de intrapredicción. Es decir, el mismo modo de intrapredicción se aplica en general para calcular valores predichos para todas las TU de una CU hoja. Para la intracodificación, un codificador de vídeo puede calcular un valor residual para cada TU hoja usando el modo de intrapredicción, como una diferencia entre la parte de la CU correspondiente a la TU y el bloque original. Una TU no se limita necesariamente al tamaño de una PU. Por tanto, las TU pueden ser más grandes o más pequeñas que una PU. Para la intracodificación, una PU se puede ubicar conjuntamente con una TU hoja correspondiente para la misma CU. En algunos ejemplos, el tamaño máximo de una TU hoja puede corresponder al tamaño de la CU hoja correspondiente.

[0048] Además, las TU de las CU hojas también se pueden asociar a las respectivas estructuras de datos de árbol cuaternario, denominadas árboles cuaternarios residuales (RQT). Es decir, una CU hoja puede incluir un árbol cuaternario que indica cómo se divide la CU hoja en las TU. El nodo raíz de un árbol cuaternario de TU corresponde en general a una CU hoja, mientras que el nodo raíz de un árbol cuaternario de CU corresponde en general a un bloque arbolado (o LCU). Las TU del RQT que no se separan se denominan TU hojas. En general, la presente divulgación usa los términos CU y TU para hacer referencia a una CU hoja y a una TU hoja, respectivamente, a menos que se indique de otro modo.

[0049] Una secuencia de vídeo incluye típicamente una serie de tramas o imágenes de vídeo. Un grupo de imágenes (GOP) comprende, en general, una serie de una o más de las imágenes de vídeo. Un GOP puede incluir datos sintácticos en una cabecera del GOP, una cabecera de una o más de las imágenes, o en otras ubicaciones, que describen un número de imágenes incluidas en el GOP. Cada fragmento de una imagen puede incluir datos sintácticos de fragmento que describen un modo de codificación para el fragmento respectivo. El codificador de vídeo 20 funciona típicamente sobre bloques de vídeo dentro de fragmentos de vídeo individuales para codificar los datos de vídeo. Un bloque de vídeo puede corresponder a un nodo de codificación dentro de una CU. Los bloques de vídeo pueden tener tamaños fijos o variados y pueden diferir en tamaño de acuerdo con una norma de codificación especificada.

[0050] Como ejemplo, el HM admite la predicción en diversos tamaños de PU. Suponiendo que el tamaño de una CU particular sea $2N \times 2N$, el HM admite la intrapredicción en tamaños de PU de $2N \times 2N$ o $N \times N$, y la interpredicción en tamaños de PU simétricos de $2N \times 2N$, $2N \times N$, $N \times 2N$ o $N \times N$. El HM también admite la división asimétrica para la interpredicción en tamaños de PU de $2N \times nU$, $2N \times nD$, $nL \times 2N$ y $nR \times 2N$. En la división asimétrica, una dirección de una CU no se divide, mientras que la otra dirección se divide en un 25 % y un 75 %. La parte de la CU correspondiente a la división de un 25 % se indica por una "n" seguida de una indicación de "arriba", "abajo", "izquierda" o "derecha". Por tanto, por ejemplo, " $2N \times nU$ " se refiere a una CU de $2N \times 2N$ que se divide horizontalmente con una PU de $2N \times 0,5N$ en la parte superior y una PU de $2N \times 1,5N$ en la parte inferior.

[0051] En la presente divulgación, " $N \times N$ " y "N por N" se pueden usar de manera intercambiable para hacer referencia a las dimensiones de píxeles de un bloque de vídeo en lo que respecta a dimensiones verticales y horizontales, por ejemplo, 16×16 píxeles o 16 por 16 píxeles. En general, un bloque de 16×16 tendrá 16 píxeles en una dirección vertical ($y = 16$) y 16 píxeles en una dirección horizontal ($x = 16$). Del mismo modo, un bloque de $N \times N$ tiene, en general, N píxeles en una dirección vertical y N píxeles en una dirección horizontal, donde N representa un valor entero no negativo. Los píxeles en un bloque se pueden organizar en filas y columnas. Además, no es necesario que los bloques tengan necesariamente el mismo número de píxeles en la dirección horizontal que en la dirección vertical. Por ejemplo, los bloques pueden comprender $N \times M$ píxeles, donde M no es necesariamente igual a N.

[0052] Tras la codificación intrapredictiva o interpredictiva usando las PU de una CU, el codificador de vídeo 20 puede calcular los datos residuales para las TU de la CU. Las PU pueden comprender datos sintácticos que describen un

procedimiento o modo de generación de datos de píxeles predictivos en el dominio espacial (también denominado el dominio de píxel) y las TU pueden comprender coeficientes en el dominio de transformada, tras la aplicación de una transformada, por ejemplo, una transformada de seno discreta (DST), una transformada de coseno discreta (DCT), una transformada entera, una transformada de tren de ondas o una transformada conceptualmente similar, en los datos de vídeo residuales. Los datos residuales pueden corresponder a diferencias de píxeles entre píxeles de la imagen no codificada y valores de predicción correspondientes a las PU. El codificador de vídeo 20 puede formar las TU, incluyendo los datos residuales para la CU y, a continuación, transformar las TU para producir coeficientes de transformada para la CU.

[0053] Tras cualquier transformada para producir coeficientes de transformada, el codificador de vídeo 20 puede realizar la cuantificación de los coeficientes de transformada. La cuantificación es un término amplio que pretende tener su significado ordinario más amplio. En un modo de realización, la cuantificación se refiere a un procedimiento en el que unos coeficientes de transformada se cuantifican para reducir posiblemente la cantidad de datos usados para representar los coeficientes, proporcionando compresión adicional. El procedimiento de cuantificación puede reducir la profundidad de bits asociada a algunos o a todos los coeficientes. Por ejemplo, un valor de n bits se puede redondear a la baja hasta un valor de m bits durante la cuantificación, donde n es mayor que m .

[0054] Tras la cuantificación, el codificador de vídeo puede explorar los coeficientes de transformada, produciendo un vector unidimensional a partir de la matriz bidimensional que incluye los coeficientes de transformada cuantificados. La exploración se puede diseñar para colocar los coeficientes de energía más alta (y, por lo tanto, de frecuencia más baja) en la parte delantera de la matriz y para colocar los coeficientes de energía más baja (y, por lo tanto, de frecuencia más alta) en la parte trasera de la matriz. En algunos ejemplos, el codificador de vídeo 20 puede utilizar un orden de exploración predefinido para explorar los coeficientes de transformada cuantificados, para producir un vector en serie que se pueda codificar por entropía. En otros ejemplos, el codificador de vídeo 20 puede realizar una exploración adaptativa. Después de explorar los coeficientes de transformada cuantificados para formar un vector unidimensional, el codificador de vídeo 20 puede codificar por entropía el vector unidimensional, por ejemplo, de acuerdo con la codificación de longitud variable adaptativa al contexto (CAVLC), la codificación aritmética binaria adaptativa al contexto (CABAC), la codificación aritmética binaria adaptativa al contexto basada en sintaxis (SBAC), la codificación por entropía por división en intervalos de probabilidad (PIPE) o con otra metodología de codificación por entropía. El codificador de vídeo 20 también puede codificar por entropía los elementos sintácticos asociados a los datos de vídeo codificados, para su uso por el descodificador de vídeo 30 en la descodificación de los datos de vídeo.

[0055] Para realizar la CABAC, el codificador de vídeo 20 puede asignar un contexto dentro de un modelo contextual a un símbolo que se va a transmitir. El contexto se puede referir, por ejemplo, a si los valores contiguos del símbolo son distintos de cero o no. Para realizar la CAVLC, el codificador de vídeo 20 puede seleccionar un código de longitud variable para un símbolo que se va a transmitir. Las palabras de código en la VLC se pueden construir de modo que los códigos relativamente más cortos correspondan a símbolos más probables, mientras que los códigos más largos correspondan a símbolos menos probables. De esta manera, el uso de la VLC puede lograr un ahorro de bits con respecto, por ejemplo, al uso de palabras de código de igual longitud para cada símbolo que se va a transmitir. La determinación de la probabilidad se puede basar en un contexto asignado al símbolo.

[0056] Además, el codificador de vídeo 20 puede enviar datos sintácticos, tales como datos sintácticos basados en bloques, datos sintácticos basados en tramas y datos sintácticos basados en el GOP, al descodificador de vídeo 30, por ejemplo, en una cabecera de trama, una cabecera de bloque, una cabecera de fragmento o una cabecera de GOP. Los datos sintácticos de GOP pueden describir un número de tramas en el GOP respectivo, y los datos sintácticos de trama pueden indicar un modo de codificación/predicción usado para codificar la trama correspondiente.

Codificador de vídeo

[0057] La FIG. 2 es un diagrama de bloques que ilustra un ejemplo de un codificador de vídeo que puede implementar técnicas de acuerdo con aspectos descritos en la presente divulgación. El codificador de vídeo 20 se puede configurar para realizar cualquiera o todas las técnicas de la presente divulgación, incluyendo pero sin limitarse a, los procedimientos de realización de escalamiento espacial usando una función de escalamiento temporal descrita en mayor detalle a continuación con respecto a la FIG. 4. Como ejemplo, la unidad de predicción entre capas 66 (cuando se proporcionan) se puede configurar para realizar cualquiera o todas las técnicas descritas en la presente divulgación. Sin embargo, los aspectos de la presente divulgación no se limitan a ello. En algunos ejemplos, las técnicas descritas en la presente divulgación se pueden compartir entre los diversos componentes del codificador de vídeo 20. En algunos ejemplos, además de o en lugar de, un procesador (no mostrado) se puede configurar para realizar cualquiera o todas las técnicas descritas en la presente divulgación.

[0058] El codificador de vídeo 20 puede realizar predicciones intra, inter y entre capas (a veces denominadas codificación intra, inter o entre capas) 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. La codificación entre capas se basa en la predicción basada en vídeo dentro de una capa(s) diferente(s) dentro de la misma secuencia de codificación de

vídeo. El intramodo (modo I) se puede referir a cualquiera de varios modos de codificación de base espacial. Los intermodos, tales como la predicción unidireccional (modo P) o la bipredicción (modo B), se pueden referir a cualquiera de varios modos de codificación de base temporal.

5 **[0059]** Como se muestra en la FIG. 2, 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. 2, el codificador de vídeo 20 incluye una unidad de selección de modo 40, una memoria de tramas 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 la unidad de compensación de movimiento 44, la unidad de estimación de movimiento 42, la unidad de intrapredicción 46, la unidad de predicción entre capas 66 y la unidad de división 48.

15 **[0060]** Para la reconstrucción de bloques de vídeo, el codificador de vídeo 20 incluye también una unidad de cuantificación inversa 58, una unidad de transformada inversa 60 y un sumador 62. También se puede incluir un filtro de eliminación de bloques (no mostrado en la FIG. 2) 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 se pueden usar filtros adicionales (en el bucle o posteriormente al 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).

20 **[0061]** Durante el procedimiento de codificación, el codificador de vídeo 20 recibe una trama o un fragmento de vídeo que se va a codificar. La trama o el fragmento se pueden dividir en múltiples bloques de vídeo. La unidad de estimación de movimiento 42 y la unidad de compensación de movimiento 44 realizan 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 realizar la codificación intrapredictiva del bloque de vídeo recibido con respecto a uno o más bloques contiguos en la misma trama o fragmento que el bloque que se va a codificar para proporcionar predicción espacial. El codificador de vídeo 20 puede realizar múltiples pases de codificación, por ejemplo, para seleccionar un modo de codificación apropiado para cada bloque de datos de vídeo.

30 **[0062]** Además, la unidad de división 48 puede dividir bloques de datos de vídeo en subbloques, en base a la evaluación de esquemas de división previos en pases de codificación previos. Por ejemplo, la unidad de división 48 puede dividir inicialmente una trama o un fragmento en LCU, y dividir cada una de las LCU en sub-CU, en base a un análisis de tasa-distorsión (por ejemplo, una optimización de tasa-distorsión, etc.). La unidad de selección de modo 40 puede producir además una estructura de datos de árbol cuaternario, indicativa de la división de una LCU en las sub-CU. Las CU de nodos hojas del árbol cuaternario pueden incluir una o más PU y una o más TU.

35 **[0063]** La unidad de selección de modo 40 puede seleccionar uno de los modos de codificación, modo de intrapredicción, interpredicción o predicción entre capas, por ejemplo, en base a los resultados erróneos, y proporcionar el bloque intracodificado, intercodificado o codificado entre capas 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 sintácticos, tales como vectores de movimiento, indicadores de intramodo, información de división y otra información sintáctica de este tipo, a la unidad de codificación por entropía 56.

45 **[0064]** 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 propósitos conceptuales. La estimación del movimiento, realizada por la unidad de estimación de movimiento 42, es el procedimiento 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 trama de referencia (u otra unidad codificada), con respecto al bloque actual que se está codificando dentro de la trama actual (u otra unidad codificada). Un bloque predictivo es un bloque que se encuentra como estrechamente coincidente con el bloque que se va a codificar, en lo que respecta a la diferencia de píxeles, lo que se puede determinar por la suma de diferencia absoluta (SAD), la suma de diferencia al cuadrado (SSD) u otras métricas de diferencia. En algunos ejemplos, el codificador de vídeo 20 puede calcular valores para posiciones de píxel fraccionario de imágenes de referencia almacenadas en la memoria de tramas de referencia 64. Por ejemplo, el codificador de vídeo 20 puede interpolar valores de posiciones de un cuarto de píxel, posiciones de un octavo de píxel u otras posiciones de píxel fraccionario 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 las posiciones de píxel fraccionario y emitir un vector de movimiento con una precisión de píxel fraccionario.

60 **[0065]** La unidad de estimación de movimiento 42 calcula un vector de movimiento para una PU de un bloque de vídeo en un fragmento intercodificado, comparando la posición de la PU con la posición de un bloque predictivo de una imagen de referencia. La imagen de referencia se puede seleccionar de una primera lista de imágenes de referencia (lista 0) o una segunda lista de imágenes de referencia (lista 1), de las que cada una identifica una o más imágenes de referencia almacenadas en la memoria de tramas de referencia 64. La unidad de estimación de movimiento 42 envía el vector de movimiento calculado a la unidad de codificación por entropía 56 y a la unidad de compensación de movimiento 44.

5 [0066] La compensación de movimiento, realizada por la unidad de compensación de movimiento 44, puede implicar capturar o generar el bloque predictivo en base al vector de movimiento determinado por la unidad de estimación de movimiento 42. La unidad de estimación de movimiento 42 y la unidad de compensación de movimiento 44 se pueden integrar 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 ubicar 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, formando valores de diferencias de píxel, como se analiza a continuación. En algunos modos de realización, la unidad de estimación de movimiento 42 puede realizar la estimación de movimiento con respecto a los componentes de luma, y la unidad de compensación de movimiento 44 puede usar los vectores de movimiento calculados en base a los componentes de luma, tanto para los componentes de croma como para los componentes de luma. La unidad de selección de modo 40 puede generar elementos sintácticos asociados a los bloques de vídeo y al fragmento de vídeo para su uso por el descodificador de vídeo 30 en la descodificación de los bloques de vídeo del fragmento de vídeo.

15 [0067] La unidad de intrapredicción 46 puede intrapredecir o calcular un bloque actual, como alternativa a la interpredicción realizada por la unidad de estimación de movimiento 42 y la unidad de compensación de movimiento 44, como se describe 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 pases de codificación separados, y la unidad de intrapredicción 46 (o la unidad de selección de modo 40, en algunos ejemplos) puede seleccionar un modo de intrapredicción apropiada a usar de los modos sometidos a prueba.

20 [0068] Por ejemplo, la unidad de intrapredicción 46 puede calcular valores de tasa-distorsión usando un análisis de tasa-distorsión para los diversos modos de intrapredicción sometidos a prueba, y seleccionar el modo de intrapredicción que tenga las mejores características de tasa-distorsión entre los modos sometidos a prueba. El análisis de tasa-distorsión determina, en general, una cantidad de distorsión (o de errores) entre un bloque codificado y un bloque original, no codificado, que se codificó 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 proporciones a partir de las distorsiones y tasas para los diversos bloques codificados, para determinar qué modo de intrapredicción presenta el mejor valor de tasa-distorsión para el bloque.

25 [0069] 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 del modo de intrapredicción y una pluralidad de tablas de índices del modo de intrapredicción modificadas (también denominadas tablas de asignación de palabras de código), definiciones de contextos de codificación para diversos bloques e indicaciones de un modo de intrapredicción más probable, una tabla de índices del modo de intrapredicción y una tabla de índices del modo de intrapredicción modificada para usar para cada uno de los contextos.

30 [0070] El codificador de vídeo 20 puede incluir una unidad de predicción entre capas 66. La unidad de predicción entre capas 66 se configura para predecir un bloque actual (por ejemplo, un bloque actual en la EL) usando una o más capas diferentes que están disponibles en SVC (por ejemplo, una capa base o de referencia). Dicha predicción se puede denominar predicción entre capas. La unidad de predicción entre capas 66 utiliza procedimientos de predicción para reducir la redundancia entre capas, mejorando de este modo la eficacia de codificación y reduciendo los requisitos de recursos informáticos. Algunos ejemplos de predicción entre capas incluyen la intrapredicción entre capas, la predicción de movimiento entre capas y la predicción residual entre capas. La intrapredicción entre capas usa la reconstrucción de bloques ubicados conjuntamente en la capa base para predecir el bloque actual en la capa de mejoramiento. La predicción de movimiento entre capas usa la información de movimiento de la capa base para predecir el movimiento en la capa de mejoramiento. La predicción residual entre capas usa el residuo de la capa base para predecir el residuo de la capa de mejoramiento. Cuando las capas base y de mejoramiento tienen diferentes resoluciones espaciales, se puede realizar el escalamiento de vectores de movimiento espacial y/o la asignación de posición entre capas por la unidad de predicción entre capas 66 usando una función de escalamiento temporal, como se describe con mayor detalle a continuación.

35 [0071] 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 de coseno discreta (DCT) o una transformada conceptualmente similar, al bloque residual, produciendo un bloque de vídeo que comprende valores de coeficientes de transformada residuales. La unidad de procesamiento de transformada 52 puede realizar otras transformadas que son conceptualmente similares a la DCT. Por ejemplo, también se pueden usar transformadas de seno discretas (DST), transformadas de tren de ondas, transformadas de enteros, transformadas de subbanda u otros tipos de transformadas.

[0072] La unidad de procesamiento de transformada 52 puede aplicar 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 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 adicionalmente la tasa de bits. El procedimiento de cuantificación puede reducir la profundidad de bits asociada a algunos 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.

[0073] 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 realizar 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 contiguos. Tras la codificación por entropía por la unidad de codificación por entropía 56, el flujo de bits codificado se puede transmitir a otro dispositivo (por ejemplo, el descodificador de vídeo 30) o archivar para su posterior transmisión o recuperación.

[0074] La unidad de cuantificación inversa 58 y la unidad 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 un 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 de tramas 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 píxeles fraccionarios, 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 de tramas de referencia 64. El bloque de vídeo reconstruido se puede usar 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.

Descodificador de vídeo

[0075] La FIG. 3 es un diagrama de bloques que ilustra un ejemplo de un descodificador de vídeo que puede implementar técnicas de acuerdo con aspectos descritos en la presente divulgación. El descodificador de vídeo 30 se puede configurar para realizar cualquiera o todas las técnicas de la presente divulgación, incluyendo pero sin limitarse a, los procedimientos de realización de escalamiento espacial usando una función de escalamiento temporal descrita con mayor detalle a continuación con respecto a la FIG. 4. Como ejemplo, la unidad de predicción entre capas 75 se puede configurar para realizar cualquiera o todas las técnicas descritas en la presente divulgación. Sin embargo, los aspectos de la presente divulgación no se limitan a ello. En algunos ejemplos, las técnicas descritas en la presente divulgación se pueden compartir entre los diversos componentes del descodificador de vídeo 30. En algunos ejemplos, además de o en lugar de, un procesador (no mostrado) se puede configurar para realizar cualquiera o todas las técnicas descritas en la presente divulgación.

[0076] En el ejemplo de la FIG. 3A, el descodificador de vídeo 30 incluye una unidad de descodificación por entropía 70, una unidad de compensación de movimiento 72, una unidad de intrapredicción 74, una unidad de predicción entre capas 75, una unidad de cuantificación inversa 76, una unidad de transformación inversa 78, una memoria de tramas de referencia 82 y un sumador 80. En algunos modos de realización, la unidad de compensación de movimiento 72 y/o la unidad de intrapredicción 74 se pueden configurar para realizar predicción entre capas, caso en el que la unidad de predicción entre capas 75 se puede omitir. En algunos ejemplos, el descodificador de vídeo 30 puede realizar un pase de descodificación, en general, recíproco al pase de codificación descrito con respecto al codificador de vídeo 20 (FIG. 2). La unidad de compensación de movimiento 72 puede generar datos de predicción en base a vectores de movimiento recibidos desde la unidad de descodificación por entropía 70, mientras que la unidad de intrapredicción 74 puede generar datos de predicción en base a indicadores de modo de intrapredicción recibidos desde la unidad de descodificación por entropía 70.

[0077] Durante el procedimiento de descodificación, el descodificador de vídeo 30 recibe un flujo de bits de vídeo codificado que representa los bloques de vídeo de un fragmento de vídeo codificado y elementos sintácticos asociados desde el codificador de vídeo 20. La unidad de descodificación por entropía 70 del descodificador de vídeo 30 descodifica por entropía el flujo de bits para generar coeficientes cuantificados, vectores de movimiento o indicadores de modo de intrapredicción y otros elementos sintácticos. La unidad de descodificación por entropía 70 reenvía los vectores de movimiento y otros elementos sintácticos a la unidad de compensación de movimiento 72. El descodificador de vídeo 30 puede recibir los elementos sintácticos en el nivel de fragmento de vídeo y/o el nivel de bloque de vídeo.

[0078] Cuando el fragmento de vídeo se codifica como un fragmento intracodificado (I), la unidad de intrapredicción 74 puede generar datos de predicción para un bloque de vídeo del fragmento de vídeo actual en base a un modo de intrapredicción señalado y datos de bloques previamente descodificados de la trama o imagen actual. Cuando la trama de vídeo se codifica como un fragmento intercodificado (por ejemplo, B, P o GPB), la unidad de compensación de movimiento 72 produce bloques predictivos para un bloque de vídeo del fragmento de vídeo actual en base a los vectores de movimiento y otros elementos sintácticos recibidos desde la unidad de descodificación 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 descodificador de vídeo 30 puede construir las listas de tramas de referencia, la lista 0 y la lista 1, usando técnicas de construcción por defecto en base a las imágenes de referencia almacenadas en la memoria de tramas de referencia 92. La unidad de compensación de movimiento 72 determina la información de predicción para un bloque de vídeo del fragmento de vídeo actual, analizando sintácticamente los vectores de movimiento y otros elementos sintácticos, y usa la información de predicción para producir los bloques predictivos para el bloque de vídeo actual que se está descodificando. Por ejemplo, la unidad de compensación de movimiento 72 usa algunos de los elementos sintácticos recibidos para determinar un modo de predicción (por ejemplo, intra o interpredicción) usado para codificar los bloques de vídeo del fragmento de vídeo, un tipo de fragmento de interpredicción (por ejemplo, un fragmento B, un fragmento P o un fragmento GPB), información de construcción para una o más de las listas de imágenes de referencia para el fragmento, vectores de movimiento para cada bloque de vídeo intercodificado del fragmento, el estado de interpredicción para cada bloque de vídeo intercodificado del fragmento y otra información para descodificar los bloques de vídeo en el fragmento de vídeo actual.

[0079] La unidad de compensación de movimiento 72 también puede realizar la interpolación en base a filtros de interpolación. La unidad de compensación de movimiento 72 puede usar filtros de interpolación como se usan por el codificador de vídeo 20 durante la codificación de los bloques de vídeo, para calcular valores interpolados para píxeles fraccionarios 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 sintácticos recibidos y usar los filtros de interpolación para producir bloques predictivos.

[0080] El descodificador de vídeo 30 también puede incluir una unidad de predicción entre capas 75. La unidad de predicción entre capas 75 se configura para predecir un bloque actual (por ejemplo, un bloque actual en la EL) usando una o más capas diferentes que están disponibles en SVC (por ejemplo, una capa base o de referencia). Dicha predicción se puede denominar predicción entre capas. La unidad de predicción entre capas 75 utiliza procedimientos de predicción para reducir la redundancia entre capas, mejorando de este modo la eficacia de codificación y reduciendo los requisitos de recursos informáticos. Algunos ejemplos de predicción entre capas incluyen la intrapredicción entre capas, la predicción de movimiento entre capas y la predicción residual entre capas. La intrapredicción entre capas usa la reconstrucción de bloques ubicados conjuntamente en la capa base para predecir el bloque actual en la capa de mejoramiento. La predicción de movimiento entre capas usa la información de movimiento de la capa base para predecir el movimiento en la capa de mejoramiento. La predicción residual entre capas usa el residuo de la capa base para predecir el residuo de la capa de mejoramiento. Cuando las capas base y de mejoramiento tienen diferentes resoluciones espaciales, se puede realizar el escalamiento del vector de movimiento espacial y/o la asignación de posición entre capas por la unidad de predicción entre capas 75 usando una función de escalamiento temporal, como se describe con mayor detalle a continuación.

[0081] La unidad de cuantificación inversa 76 cuantifica de manera inversa, por ejemplo, descuantifica, los coeficientes de transformada cuantificados proporcionados en el flujo de bits y descodificados por la unidad de descodificación por entropía 70. El procedimiento de cuantificación inversa puede incluir el uso de un parámetro de cuantificación QPY, calculado por el descodificador de vídeo 30 para cada bloque de vídeo en el fragmento de vídeo para determinar un grado de cuantificación y, del mismo modo, un grado de cuantificación inversa que se debería aplicar.

[0082] La unidad de transformada inversa 78 aplica una transformada inversa, por ejemplo, una DCT inversa, una DST inversa, una transformada entera inversa, o un procedimiento de transformada inversa conceptualmente similar, a los coeficientes de transformada para producir bloques residuales en el dominio de píxel.

[0083] Después de que la unidad de compensación de movimiento 72 genera el bloque predictivo para el bloque de vídeo actual en base a los vectores de movimiento y otros elementos sintácticos, el descodificador de vídeo 30 forma un bloque de vídeo descodificado sumando los bloques residuales de la unidad de transformada inversa 78 a los correspondientes bloques predictivos generados por la unidad de compensación de movimiento 72. El sumador 90 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 descodificados para 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 del vídeo. Los bloques de vídeo descodificados en una trama o imagen dada se almacenan a continuación en la memoria de imágenes de referencia 92, que almacena imágenes de referencia usadas para una compensación de movimiento posterior. La memoria de tramas de referencia 82 almacena también vídeo descodificado para su presentación posterior en un dispositivo de visualización, tal como el dispositivo de visualización 32 de la FIG. 1.

Escalamiento del vector de movimiento

[0084] Como se analiza anteriormente, un vector de movimiento se escala como parte de un procedimiento de codificación o decodificación. Por ejemplo, los vectores de movimiento de una capa que tiene una resolución espacial diferente a la capa del bloque actual se pueden escalar antes de usar para codificar o decodificar el bloque actual. Un procesador se configura para implementar un procedimiento de escalamiento del vector de movimiento. La primera etapa del procedimiento puede ser para cambiar el parámetro de entrada *iScaleBase*, analizado anteriormente, en un ejemplo que no se encuentra dentro del alcance de la invención, a:

$$iScaleBase = \frac{e_{lH}}{b_{lH}} \times 256.$$

Un *iScaleBase* determinado usando una proporción de los valores de altura de la capa base y de mejoramiento proporciona ventajas sobre el uso de una proporción de los valores de ancho de la capa base y de mejoramiento porque el valor de altura es normalmente más pequeño que el ancho.

[0085] De acuerdo con la invención, para el cálculo de parámetros de entrada, se añade una desviación a la ecuación como sigue:

$$iScaleBase = \frac{e_{lW} \times 256 + desvia}{b_{lW}},$$

donde la desviación puede ser $b_{lW}/2$, que corresponde a un valor de desviación de redondeo de 0,5. En otro modo de realización que proporciona una desviación de redondeo, el parámetro de entrada puede ser:

$$\left(\frac{b_{lW}}{2} - 1\right).$$

[0086] Como se analiza anteriormente, estos parámetros de entrada se usan como la entrada a una función de escalamiento temporal TMVP que ya existe bajo la norma HEVC para realizar escalamiento espacial. El uso de la función de escalamiento temporal TMVP existente para realizar escalamiento espacial proporciona la ventaja de permitir el escalamiento espacial sin introducir una nueva función de escalamiento específica a la norma HEVC o sus extensiones (tal como la extensión SHVC).

Proporción de escalabilidad para las direcciones horizontal y vertical

[0087] Si se desea que la proporción de escalabilidad no sea la misma en las direcciones horizontal y vertical, a continuación los procedimientos descritos anteriormente se pueden extender introduciendo dos parámetros de escala *iScaleBaseX* e *iScaleBaseY* para las direcciones horizontal y vertical, respectivamente. En un ejemplo:

$$iScaleBaseX = \frac{e_{lW} * N}{b_{lW}},$$

$$iScaleBaseY = \frac{e_{lH} * N}{b_{lH}},$$

donde el factor escalado, N, puede ser un número entero, tal como 256. En otro ejemplo, se puede añadir una desviación al numerador de cada parámetro de escala.

[0088] En este modo de realización, la función de escalamiento TMVP tendrá dos parámetros de entrada, *iScaleBaseX* e *iScaleBaseY*, y se puede aplicar el procedimiento de escalamiento TMVP para cada componente del vector de movimiento independientemente. Por ejemplo, la función TMVP se puede llamar dos veces; la primera vez usando *iScaleBaseX* como entrada, y la segunda vez usando *iScaleBaseY* como entrada.

[0089] Además, todos los procedimientos y técnicas descritos a continuación en lo que respecta a una proporción de escalabilidad se pueden aplicar de forma similar a las proporciones de escalabilidad horizontal y vertical independientemente. Por ejemplo, se pueden realizar procedimientos similares para la asignación de posición descrita a continuación, donde se pueden introducir dos parámetros de escalamiento (por ejemplo, uno para la dirección horizontal y otro para la dirección vertical). Para el cálculo del parámetro horizontal, se usan *e_{lW}* y *b_{lW}*, y para el cálculo del parámetro vertical, se usan *e_{lH}* y *b_{lH}*.

Reducción del intervalo de datos para la división

[0090] En algunos ejemplos para reducir los costes informáticos, de ancho de banda y/o de memoria, la operación de división del parámetro *iScaleBase* se limita restringiendo el divisor dentro de un determinado (por ejemplo, predeterminado) intervalo. Al hacerlo, la operación de división se puede implementar con una tabla de consulta de tamaño más pequeño.

[0091] Los valores de ancho o altura de la capa base y las capas de mejoramiento se reducen de modo que el ancho de la capa base estará en un intervalo de (0, blMax), donde blMax es el valor de ancho de la capa base máximo posible y se puede relacionar con el tamaño de la tabla de consulta. Cuanto más pequeño sea este valor, más pequeña será la tabla de consulta que se puede usar. Se puede encontrar un umbral óptimo entre el tamaño de la tabla de consulta y la exactitud cambiando el valor de blMax.

[0092] Para un número dado blMax, se calcula el número de desplazamientos (o el divisor de potencia de 2) para el ancho o la altura de la capa base, para que el valor esté en el intervalo de (0, blMax). En un ejemplo, el número de desplazamientos se denomina N, y el divisor correspondiente de la capa base es 2^N . Se puede utilizar la potencia de dos para simplificar la implementación, ya que se pueden realizar las operaciones de división por dos por desplazamiento de bits. Sin embargo, en otros ejemplos, el valor puede ser cualquier número más pequeño que el ancho o la altura de la capa base.

[0093] Seguidamente, la capa base y la capa de mejoramiento se desplazan a la derecha en N para mantener la misma proporción de escalabilidad. El cálculo del parámetro de entrada final se puede expresar como:

$$iScaleBase = ((eIW \gg N) \ll 8) / (blW \gg N).$$

[0094] Se puede usar una fórmula variante para reducir el error de redondeo:

$$iScaleBase = ((eIW \ll 8) \gg N) / (blW \gg N),$$

donde (blW >> N) está en un intervalo (0, blMax), que tiene un límite superior, y el desplazamiento a la izquierda en 8 se usa para normalizar el parámetro de entrada de la función de escalamiento TMVP (por ejemplo, la escala de 1 corresponde a 256).

[0095] Además, se puede añadir la desviación de redondeo a la fórmula anterior. Por ejemplo, se puede añadir la desviación como:

$$iScaleBase = (((eIW \gg N) + desvia) \ll 8) / (blW \gg N),$$

donde la desviación puede ser (blW >> (N+1)), que corresponde a un valor de desviación de redondeo de 0,5. Otra variante para la desviación de redondeo puede ser (blW >> (N+1))-1, en la que se usa el redondeo hacia cero.

[0096] La siguiente variante se puede usar para reducir el error de redondeo para un intervalo de datos reducidos de eIW:

$$iScaleBase = (((eIW \ll 8) \gg N) + desvia) / (blW \gg N).$$

De forma similar, la desviación se puede establecer en (blW >> (N+1)), correspondiente a un valor de desviación de redondeo de 0,5, o (blW >> (N+1))-1, correspondiente al redondeo hacia cero. La variable blW y eIW en todas las ecuaciones anteriores se puede reemplazar por blH y eIH.

Implementación de C++

[0097] Un ejemplo de una implementación de C++ es como sigue:

```

Int blMax = 128;
Int N = 0;
Int blWidth = iBWidth;
while (blWidth > blMax)
{
    blWidth >>= 1;
    N++;
}
assert((iBWidth >> N) < blMax);
iScaleBase = (iEWidth >> N) << 8 / (iBWidth >> N);

```

donde *iBWidth* es el ancho de la capa base, *iEWidth* es el ancho de la capa de mejoramiento. Aunque este ejemplo omite la desviación de redondeo, se puede incluir la desviación de redondeo, como se analiza anteriormente.

Evitando la operación de división

[0098] Para facilitar la conversión de una operación de división a una función de acceso a la tabla de consulta, el cálculo del parámetro de entrada se puede implementar como sigue:

$$iScaleBase = elW * \frac{2^K}{blW} \times 256 \gg K$$

[0099] En otro ejemplo, se añade una desviación:

5
$$iScaleBase = elW * \frac{2^K + desviación}{blW} \times 256 \gg K$$

donde la desviación se puede establecer en $blW/2$ o $\left(\frac{blW}{2} - 1\right)$, como se analiza anteriormente. En otro ejemplo:

10
$$iScaleBase = elW * \frac{2^K + desviación}{blW} \gg (K - 8)$$

donde K es el valor constante predefinido (por ejemplo, un valor de ejemplo de K es 16). Con esta expresión, la división $\frac{2^K + desviación}{blW}$ se puede reemplazar fácilmente por una tabla de consulta.

15 [0100] Los procedimientos descritos anteriormente también se pueden aplicar a los procedimientos de esta sección "Evitando la operación de división". Por ejemplo, en un ejemplo, blW se desplaza a la derecha en N bits, de modo que el valor desplazado de $(blW \gg N)$ está en un intervalo predefinido (0, blMax). El valor de $(2^K / (blW \gg N))$ se almacena en una tabla de consulta para $(blW \gg N)$ en el intervalo de (0, blMax).

20 [0101] El cálculo se puede realizar como:

$$iScaleBase = (elW \gg N) * \frac{2^K + desviación}{blW \gg N} \gg (K - 8)$$

En otro ejemplo:

25
$$iScaleBase = elW * \frac{2^K + desviación}{blW \gg N} \gg (K + N - 8)$$

donde la desviación se puede establecer en $(blW \gg (N + 1))$ o $((blW \gg (N + 1)) - 1)$ como se menciona anteriormente.

30 [0102] La variable blW y elW en todas las ecuaciones anteriores se puede reemplazar por blH y elH, como se menciona anteriormente.

Mejora de la asignación de posición entre capas

35 [0103] La simplificación de la operación de división se puede aplicar a otros elementos donde se usa la división de acuerdo con la proporción de aspecto de escalabilidad. Por ejemplo, también podría ser necesario dividir por el ancho o la altura de la capa base para ubicar la posición del píxel ubicado conjuntamente en la imagen de la capa base. En este caso, se pueden aplicar la misma técnica o procedimientos analizados anteriormente y usados para escalar un vector de movimiento, para calcular la posición de un píxel ubicado conjuntamente o un bloque ubicado conjuntamente en la capa base.

40 [0104] Como se menciona anteriormente, las coordenadas x e y de un píxel (o bloque) en una capa base (blx, bly) se pueden expresar como:

45
$$\begin{cases} blx = x \times \frac{blW}{elW} \\ bly = y \times \frac{blH}{elH} \end{cases}$$

Donde x e y corresponden a las coordenadas del píxel en la capa de mejoramiento, y blW, elW, blH y elH corresponden a los anchos y las alturas de la capa base y de mejoramiento, respectivamente. La ecuación anterior se puede modificar como sigue:

50
$$blx = (x * iScaleBase' + desvia) \gg M$$

$$bly = (y * iScaleBase' + desvia) \gg M$$

donde, en un ejemplo que no se encuentra dentro del alcance de las reivindicaciones, iScaleBase' se puede calcular

$$\text{como } iScaleBase' = \frac{blW}{elW} * (1 \ll M)$$

5 **[0105]** Todos los procedimientos mencionados anteriormente se pueden aplicar a la ecuación anterior para reducir el coste de la operación de división por elW. Además, se puede aplicar una cualquiera de las desviaciones, descritas anteriormente, para definir la desviación 1, anteriormente.

10 **Implementación de la tabla de consulta**

[0106] En otro modo de realización, la operación de división se elimina reemplazándola con una tabla de consulta que se basa en valores reducidos de intervalo de los anchos y/o las alturas de la capa base y la capa de mejoramiento.

15 **[0107]** En un ejemplo, los valores del ancho y/o la altura de la capa base y de la capa de mejoramiento se reducen para garantizar que se encontrarán dentro de un intervalo de (0, blMax) y (0, elMax), respectivamente. blMax es el valor del ancho y/o la altura de la capa base máximo posible y elMax es el valor del ancho y/o la altura de la capa de mejoramiento máximo posible. Tanto blMax como elMax pueden estar conectados o relacionados con el tamaño de la tabla de consulta. Cuanto más pequeños sean estos valores, más pequeña será la tabla de consulta que se puede usar.

20 **[0108]** Este enfoque proporciona una compensación entre la complejidad informática y la exactitud del escalamiento. Se puede encontrar un umbral óptimo entre el tamaño de la tabla de consulta y la exactitud cambiando los valores de blMax y elMax.

25 **[0109]** Para un número dado blMax, se puede calcular el número de desplazamientos (o el divisor de potencia de 2) para el ancho o la altura de la capa base y/o de mejoramiento para que el valor esté en el intervalo de (0, blMax) y/o (0, elMax) respectivamente. Por ejemplo, el número de desplazamientos se puede representar como Nb, y el divisor correspondiente de la capa base será 2^{Nb}. También se pueden aplicar las mismas técnicas y procedimientos a una capa de mejoramiento. La potencia de 2 se puede elegir debido a la fácil operación, puesto que la división se puede realizar por desplazamientos; sin embargo, en otros modos de realización puede ser cualquier número más pequeño que el ancho o la altura de la capa base.

35 **[0110]** La capa base y las capas de mejoramiento se desplazan a la derecha posteriormente en Nb y Ne para mantener la misma proporción de escalabilidad. El cálculo del parámetro de entrada final se puede expresar como:

$$iScaleBase = \text{LookUpTable}[(iEWidth \gg N_e)][(iBWidth \gg N_b)] .$$

40 **[0111]** La función LookUpTable se puede implementar de muchas maneras en base a la exactitud del redondeo, el coste de distorsión de la tasa, etc. Además, las variables iEWidth e iBWidth en todas las ecuaciones anteriores se pueden reemplazar por iEHeight e iBHeight.

45 **[0112]** La FIG. 4 ilustra un modo de realización de un procedimiento de realización de escalamiento espacial usando una función de escalamiento temporal que se puede realizar por el codificador de vídeo 20 de la FIG. 2 o el descodificador de vídeo 30 de la FIG. 3. El procedimiento 400 se puede realizar por una cualquiera o más de la unidad de estimación de movimiento 42, la unidad de compensación de movimiento 44, la unidad de intrapredicción 46 y la unidad de predicción entre capas 66 del codificador de vídeo 20 de la FIG. 2. En otro modo de realización, el procedimiento 400 se puede realizar por una cualquiera o más de la unidad de compensación de movimiento 72, la unidad de intrapredicción 74 y la unidad de predicción entre capas 75 del descodificador de la FIG. 3.

50 **[0113]** El procedimiento 400 comienza en el bloque 410. En el bloque 420, se determinan los valores de dimensión espacial de la capa base (BL) y la capa de mejoramiento (EL). Por ejemplo, el ancho de la imagen de la capa base y el ancho de la imagen de la capa de mejoramiento se pueden determinar en el bloque 420, y la altura de la imagen de la capa base y la altura de la imagen de la capa de mejoramiento se determinan en el bloque 420. En un modo de realización, la dimensión es la misma dimensión tanto en BL como en EL.

55 **[0114]** En el bloque 430, se determina un factor de escalamiento en base a los valores de dimensión espacial de BL y EL, de modo que el factor de escalamiento se restringe dentro de un intervalo predeterminado. Por ejemplo, el valor de la dimensión de BL se puede restringir para que se encuentre dentro del intervalo de (0, blMax), como se analiza anteriormente, y el valor de la dimensión de EL se puede restringir para que se encuentre dentro del intervalo de (0, elMax), como se analiza anteriormente. En un modo de realización, el factor de escalamiento se restringe multiplicando una proporción de los valores de la dimensión de EL y BL por un primer valor y dividiendo el producto por un segundo valor. En un modo de realización, el segundo valor es una potencia de 2, de modo que la división se puede realizar por una operación de desplazamiento a la derecha. El factor de escalamiento se restringe de modo que se puede usar con un procedimiento o función de escalamiento del vector de movimiento preexistente, tal como el procedimiento de

escalamiento TMVP proporcionado por la especificación HEVC. En un modo de realización, el factor de escalamiento se restringe usando una función de recorte.

5 **[0115]** En el bloque 440 un elemento asociado a la EL o bien la BL se escala espacialmente usando una función de escalamiento temporal y el factor de escalamiento. Por ejemplo, en un modo de realización, el elemento corresponde a un vector de movimiento asociado a un bloque de vídeo en la BL. En otro modo de realización, el elemento corresponde a la ubicación espacial (o posición) de un píxel o bloque de vídeo en la BL. El procedimiento 400 continúa al bloque 450 y termina.

10 **[0116]** Si bien la divulgación anterior ha descrito modos de realización particulares, son posibles muchas variaciones. Por ejemplo, como se mencionó anteriormente, las técnicas anteriores se pueden aplicar a la codificación de vídeo 3D. En algunos modos de realización de vídeo 3D, una capa de referencia (por ejemplo, una capa base) incluye información de vídeo suficiente para mostrar una primera vista de un vídeo y la capa de mejoramiento incluye información de vídeo adicional con respecto a la capa de referencia de modo que la capa de referencia y la capa de mejoramiento incluyen conjuntamente información de vídeo suficiente para mostrar una segunda vista del vídeo. Estas dos vistas se pueden usar para generar una imagen estereoscópica. Como se analiza anteriormente, la información de movimiento de la capa de referencia se puede usar para identificar hipótesis implícitas adicionales cuando se codifica o descodifica una unidad de vídeo en la capa de mejoramiento, de acuerdo con aspectos de la divulgación. Esto puede proporcionar una mayor eficacia de codificación para un flujo de bits de vídeo 3D.

20 **[0117]** Se debe reconocer que dependiendo del ejemplo, determinados actos o acontecimientos de cualquiera de las técnicas descritas en el presente documento se pueden realizar en una secuencia diferente, se pueden añadir, fusionar u omitir por completo (por ejemplo, no todos los actos o acontecimientos descritos son necesarios para la puesta en práctica de las técnicas). Además, en determinados ejemplos, los actos o acontecimientos se pueden realizar simultáneamente, por ejemplo, a través de procesamiento de múltiples subprocedimientos, procesamiento de interrupciones o múltiples procesadores, en lugar de secuencialmente.

25 **[0118]** La información y las señales divulgadas en el presente documento se pueden representar usando cualquiera de una variedad de tecnologías y técnicas diferentes. Por ejemplo, los datos, las instrucciones, los comandos, la información, las señales, los bits, los símbolos y los chips que se puedan haber mencionado por toda la descripción anterior se pueden representar mediante tensiones, corrientes, ondas electromagnéticas, campos o partículas magnéticas, campos o partículas ópticos o cualquier combinación de los mismos.

30 **[0119]** Los diversos bloques lógicos, módulos, circuitos y etapas de algoritmo ilustrativos descritos en relación con los modos de realización divulgados en el presente documento se pueden implementar como hardware electrónico, software informático o combinaciones de ambos. Para ilustrar claramente esta intercambiabilidad de hardware y software, anteriormente se han descrito diversos componentes, bloques, módulos, circuitos y etapas ilustrativas, en general, en lo que respecta a su funcionalidad. Que dicha funcionalidad se implemente como hardware o software depende de la aplicación y las restricciones de diseño particulares impuestas en el sistema global. Los expertos en la técnica pueden implementar la funcionalidad descrita de formas variadas para cada aplicación particular, pero no se debería interpretar que dichas decisiones de implementación suponen apartarse del alcance de la presente invención.

35 **[0120]** Las técnicas descritas en el presente documento se pueden implementar en hardware, software, firmware o en cualquier combinación de los mismos. Dichas técnicas se pueden implementar en cualquiera de una variedad de dispositivos tales como ordenadores de propósito general, equipos manuales de dispositivos de comunicación inalámbrica o dispositivos de circuitos integrados que tienen múltiples usos, incluyendo su aplicación en equipos manuales de dispositivos de comunicación inalámbrica y otros dispositivos. Todos los rasgos característicos descritos como módulos o componentes se pueden implementar conjuntamente en un dispositivo lógico integrado o por separado, como dispositivos lógicos discretos pero interoperables. Si se implementan en software, las técnicas se pueden realizar, al menos en parte, por un medio de almacenamiento de datos legible por ordenador que comprenda código de programa que incluye instrucciones que, cuando se ejecutan, realizan uno o más de los procedimientos descritos anteriormente. El medio de almacenamiento de datos legible por ordenador puede formar parte de un producto de programa informático, que puede incluir materiales de embalaje. El medio legible por ordenador puede comprender memoria o medios de almacenamiento de datos, tales como memoria de acceso aleatorio (RAM), tal como memoria de acceso aleatorio dinámica síncrona (SDRAM), memoria de solo lectura (ROM), memoria de acceso aleatorio no volátil (NVRAM), memoria de solo lectura programable y borrable eléctricamente (EEPROM), memoria FLASH, medios de almacenamiento de datos magnéticos u ópticos y similares. Las técnicas se pueden realizar adicionalmente, o de forma alternativa, al menos en parte, por un medio de comunicación legible por ordenador que transporta o comunica código de programa en forma de instrucciones o estructuras de datos y a las que se puede acceder, leer y/o ejecutar por un ordenador, tales como señales u ondas propagadas.

40 **[0121]** El código de programa se puede ejecutar por un procesador, que puede incluir 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. Un procesador de este tipo se puede configurar para realizar cualquiera de las técnicas descritas en la presente divulgación. Un procesador de propósito general puede ser un microprocesador,

5 pero, de forma alternativa, el procesador puede ser cualquier procesador, controlador, microcontrolador o máquina de
estados convencional. Un procesador también se puede implementar como una combinación de dispositivos
informáticos, por ejemplo, una combinación de un DSP y un microprocesador, una pluralidad de microprocesadores,
10 uno o más microprocesadores junto con un núcleo de DSP o cualquier otra configuración de este tipo. En
consecuencia, el término "procesador", como se usa en el presente documento, se puede referir a cualquiera de las
estructuras anteriores, cualquier combinación de la estructura anterior, o cualquier otra estructura o aparato adecuado
para la implementación de las técnicas descritas en el presente documento. Además, en algunos aspectos, la
funcionalidad descrita en el presente documento se puede proporcionar dentro de módulos de software o módulos de
15 hardware dedicados configurados para codificar y descodificar, o incorporados en un codificador-descodificador de
vídeo combinado (CÓDEC).

[0122] Se han descrito diversos modos de realización de la invención. Estos y otros modos de realización están dentro
del alcance de las siguientes reivindicaciones.

15

REIVINDICACIONES

1. Un procedimiento de descodificación de información de vídeo, comprendiendo el procedimiento:

5 con un procesador acoplado a una unidad de memoria configurada para almacenar información de vídeo asociada a una capa base, BL, y una capa de mejoramiento, EL, en el que la BL comprende una o más imágenes de BL que tienen un valor de ancho de BL y un valor de altura de BL, y la EL comprende una o más imágenes de EL que tienen un valor de ancho de EL y un valor de altura de EL,

10 determinar un primer factor de escalamiento en base a una proporción de una suma de a) el valor de ancho de EL multiplicado por 256 y b) un valor de desviación, al valor de ancho de BL;

determinar un segundo factor de escalamiento en base a una proporción de una suma de a) el valor de altura de EL multiplicado por 256 y b) un valor de desviación, al valor de altura de BL, en el que el primer factor de escalamiento es diferente del segundo factor de escalamiento;

15 escalar espacialmente una componente horizontal de un elemento asociado a la capa base usando el primer factor de escalamiento determinado como una entrada a una función de escalamiento, en el que la función de escalamiento se configura para escalar temporalmente vectores de movimiento para generar predictores del vector de movimiento temporales;

20 escalar espacialmente una componente vertical del elemento asociado a la BL usando el segundo factor de escalamiento determinado como una entrada a la función de escalamiento; y

25 descodificar la información de vídeo usando las componentes horizontal y vertical escaladas espacialmente del elemento.

2. Un procedimiento de codificación de información de vídeo, comprendiendo el procedimiento:

30 con un procesador acoplado a una unidad de memoria configurada para almacenar información de vídeo asociada a una capa base, BL, una capa de mejoramiento, EL, en el que la BL comprende una o más imágenes de BL que tienen un valor de ancho de BL y un valor de altura de BL, y la EL comprende una o más imágenes de EL que tienen un valor de ancho de EL y un valor de altura de EL,

35 determinar un primer factor de escalamiento en base a una proporción de una suma de a) el valor de ancho de EL multiplicado por 256 y b) un valor de desviación, al valor de ancho de BL;

determinar un segundo factor de escalamiento en base a una proporción de una suma de a) el valor de altura de EL multiplicado por 256 y b) un valor de desviación, al valor de altura de BL, en el que el primer factor de escalamiento es diferente del segundo factor de escalamiento;

40 escalar espacialmente una componente horizontal de un elemento asociado a la capa base usando el primer factor de escalamiento determinado como una entrada a una función de escalamiento, en el que la función de escalamiento se configura para escalar temporalmente vectores de movimiento para generar predictores del vector de movimiento temporales;

45 escalar espacialmente una componente vertical del elemento asociado a la BL usando el segundo factor de escalamiento determinado como una entrada a la función de escalamiento; y

50 codificar la información de vídeo usando las componentes horizontal y vertical escaladas espacialmente del elemento.

3. El procedimiento de la reivindicación 1 o la reivindicación 2, en el que el valor de desviación es la mitad del valor de ancho de BL; o en el que el valor de desviación es uno menos de la mitad del valor de ancho de BL.

4. El procedimiento de la reivindicación 1 o la reivindicación 2, en el que el valor de ancho de EL y el valor de ancho de BL están dentro de un intervalo de 0 a un valor de dimensión de ancho máximo de la BL.

5. El procedimiento de la reivindicación 1 o la reivindicación 2, en el que el elemento comprende un vector de movimiento; o

en el que el elemento comprende una posición espacial de un píxel o bloque de vídeo.

6. El procedimiento de la reivindicación 1 o la reivindicación 2, en el que determinar los primer y segundo factores de escalamiento comprende determinar los primer y segundo factores de escalamiento sin realizar una operación de división; y/o

en el que determinar los primer y segundo factores de escalamiento comprende determinar los primer y segundo factores de escalamiento usando una tabla de consulta para cada operación de división asociada a los primer y segundo factores de escalamiento.

5 7. El procedimiento de la reivindicación 1 o la reivindicación 2, en el que el valor de altura de BL y el valor de altura de EL están dentro de un intervalo de 0 a un valor de dimensión de ancho máximo de la BL.

10 8. Un medio legible por ordenador no transitorio que comprende instrucciones que, cuando se ejecutan, hacen que un aparato realice el procedimiento de cualquiera de las reivindicaciones 1- 7.

9. Un aparato configurado para descodificar información de vídeo, comprendiendo el aparato:

15 medios para determinar un primer factor de escalamiento en base a una proporción de una suma de a) un valor de ancho de la capa de mejoramiento, EL, multiplicado por 256 y b) un valor de desviación, a un valor de ancho de la capa base, BL;

20 medios para determinar un segundo factor de escalamiento en base a una proporción de una suma de a) un valor de altura de EL multiplicado por 256 y b) un valor de desviación, a un valor de altura de BL, en el que el primer factor de escalamiento es diferente del segundo factor de escalamiento;

25 medios para escalar espacialmente una componente horizontal de un elemento asociado a la capa base usando el primer factor de escalamiento determinado como una entrada para una función de escalamiento, en el que la función de escalamiento se configura para escalar temporalmente vectores de movimiento para generar predictores del vector de movimiento temporales;

medios para escalar espacialmente una componente vertical del elemento asociado a la BL usando el segundo factor de escalamiento determinado como una entrada a la función de escalamiento; y

30 medios para descodificar la información de vídeo usando las componentes horizontal y vertical escaladas espacialmente del elemento.

10. Un aparato configurado para codificar información de vídeo, comprendiendo el aparato:

35 medios para determinar un primer factor de escalamiento en base a una proporción de una suma de a) un valor de ancho de la capa de mejoramiento, EL, multiplicado por 256 y b) un valor de desviación, a un valor de ancho de la capa base, BL;

40 medios para determinar un segundo factor de escalamiento en base a una proporción de una suma de a) un valor de altura de EL multiplicado por 256 y b) un valor de desviación, a un valor de altura de BL, en el que el primer factor de escalamiento es diferente del segundo factor de escalamiento;

45 medios para escalar espacialmente una componente horizontal de un elemento asociado a la capa base usando el primer factor de escalamiento determinado como una entrada para una función de escalamiento, en el que la función de escalamiento se configura para escalar temporalmente vectores de movimiento para generar predictores del vector de movimiento temporales;

50 medios para escalar espacialmente una componente vertical del elemento asociado a la BL usando el segundo factor de escalamiento determinado como una entrada a la función de escalamiento; y

medios para codificar la información de vídeo usando las componentes horizontal y vertical escaladas espacialmente del elemento.

55 11. El aparato de la reivindicación 9 o la reivindicación 10, en el que el aparato es un dispositivo seleccionado del grupo que consiste en un televisor digital, un sistema de radiodifusión digital directa, un sistema de radiodifusión inalámbrica, un asistente digital personal, PDA, un ordenador portátil, un ordenador de sobremesa, un ordenador de tableta, un lector de libros electrónicos, una cámara digital, un dispositivo de grabación digital, un reproductor de medios digitales, un dispositivo de videojuegos, una consola de videojuegos, un teléfono móvil, un teléfono de radio por satélite, un teléfono inteligente, un dispositivo de videoconferencia y un dispositivo de transmisión
60 continua de vídeo.

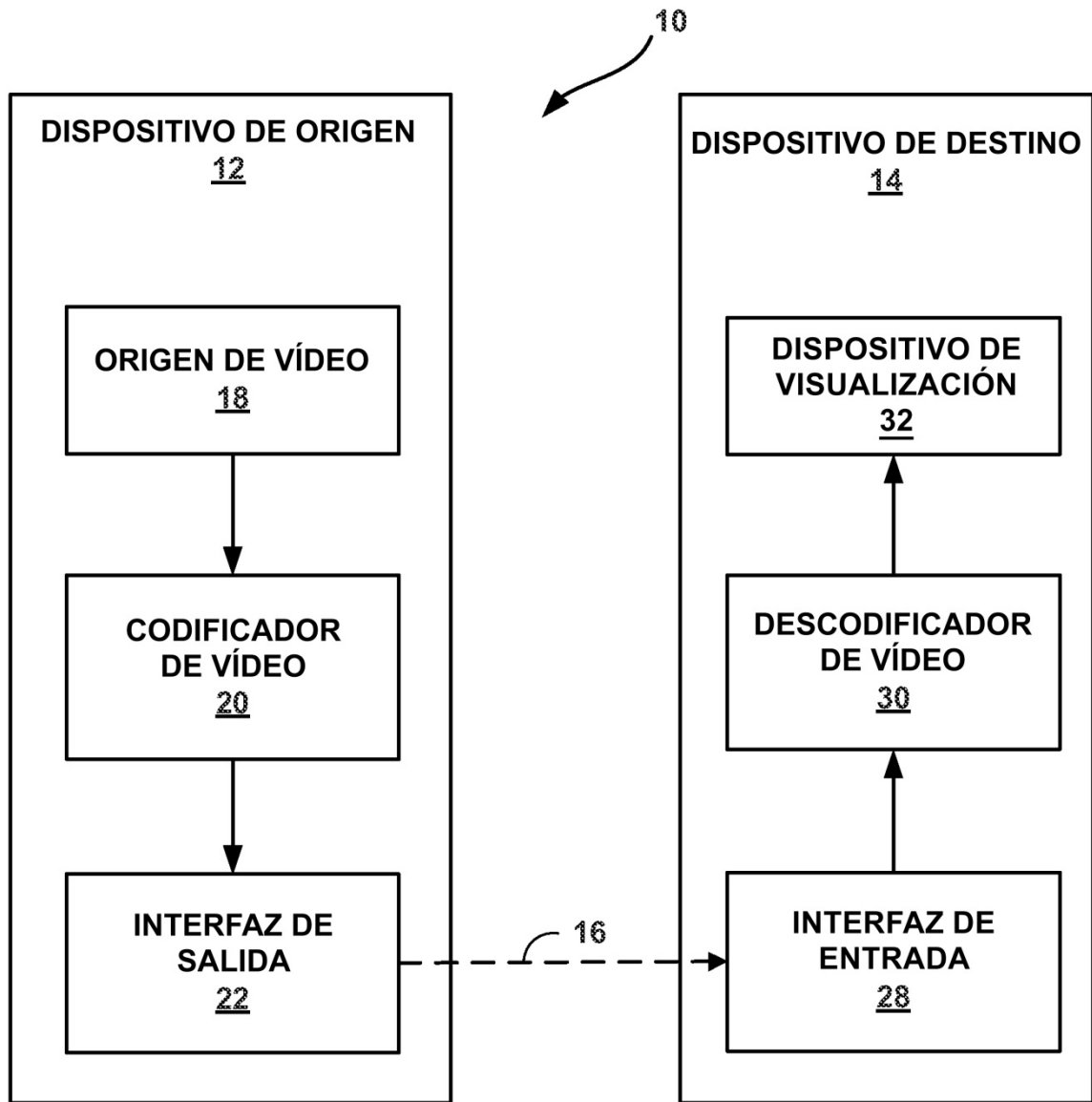


FIG. 1

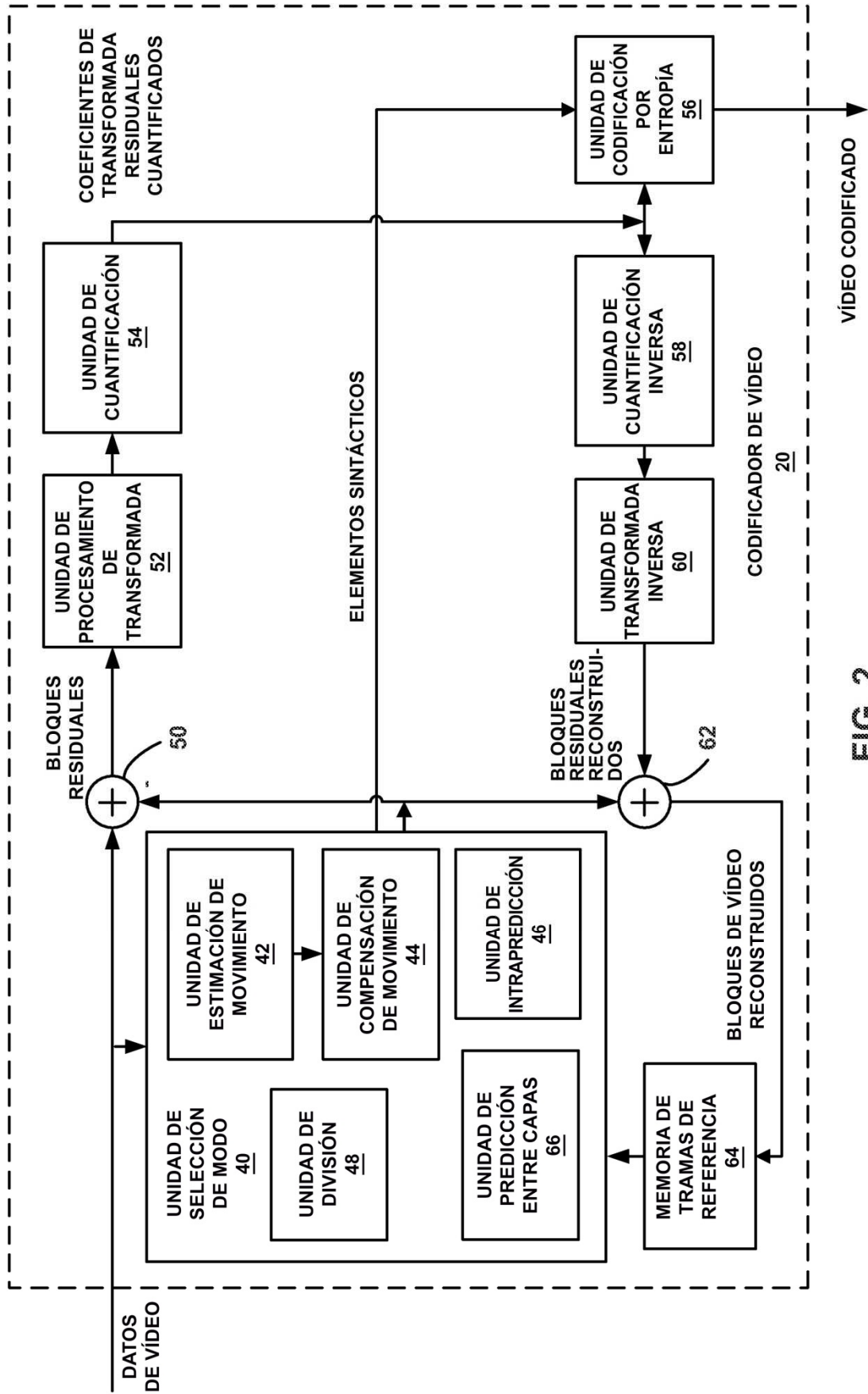


FIG. 2

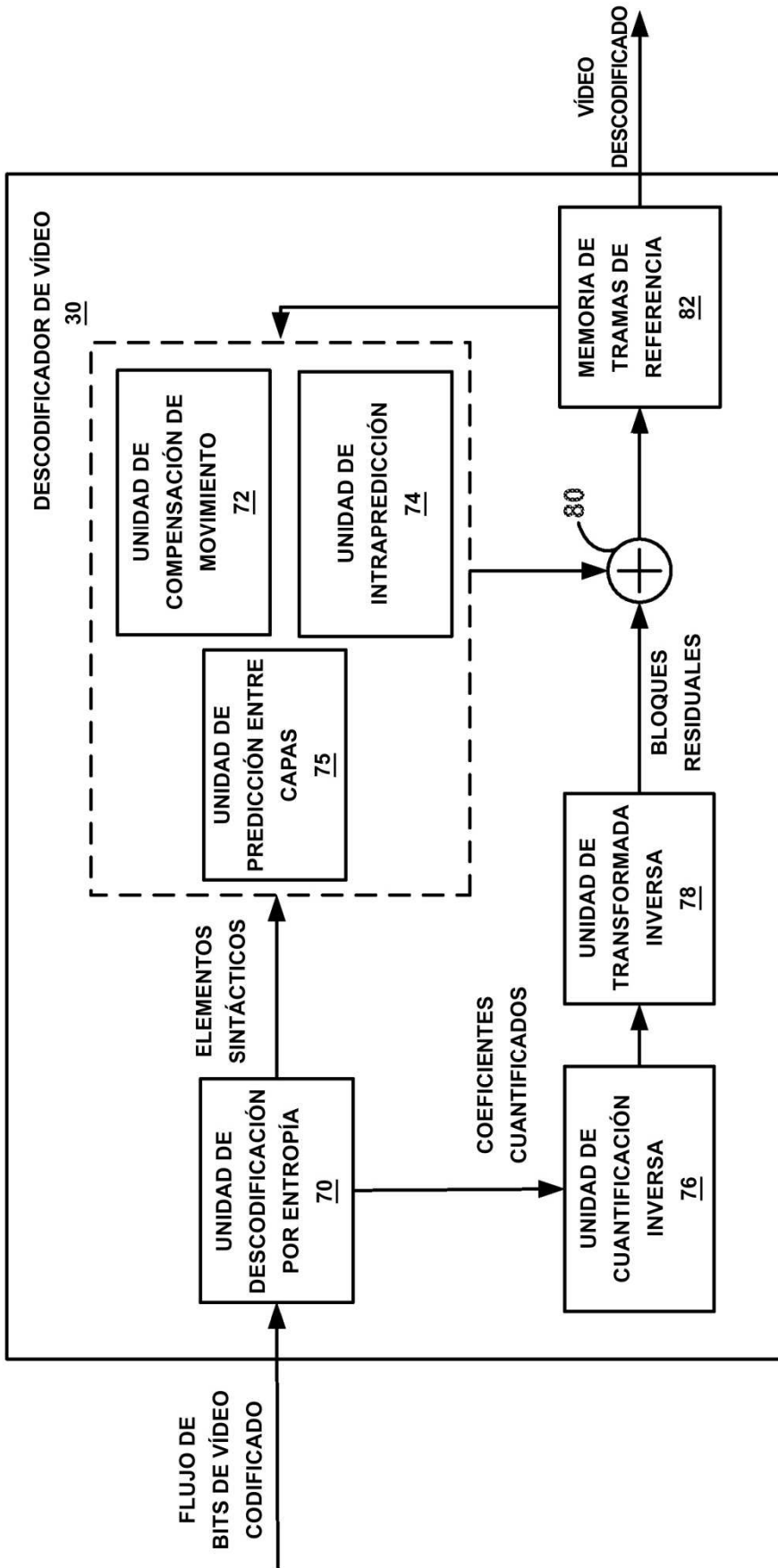


FIG. 3

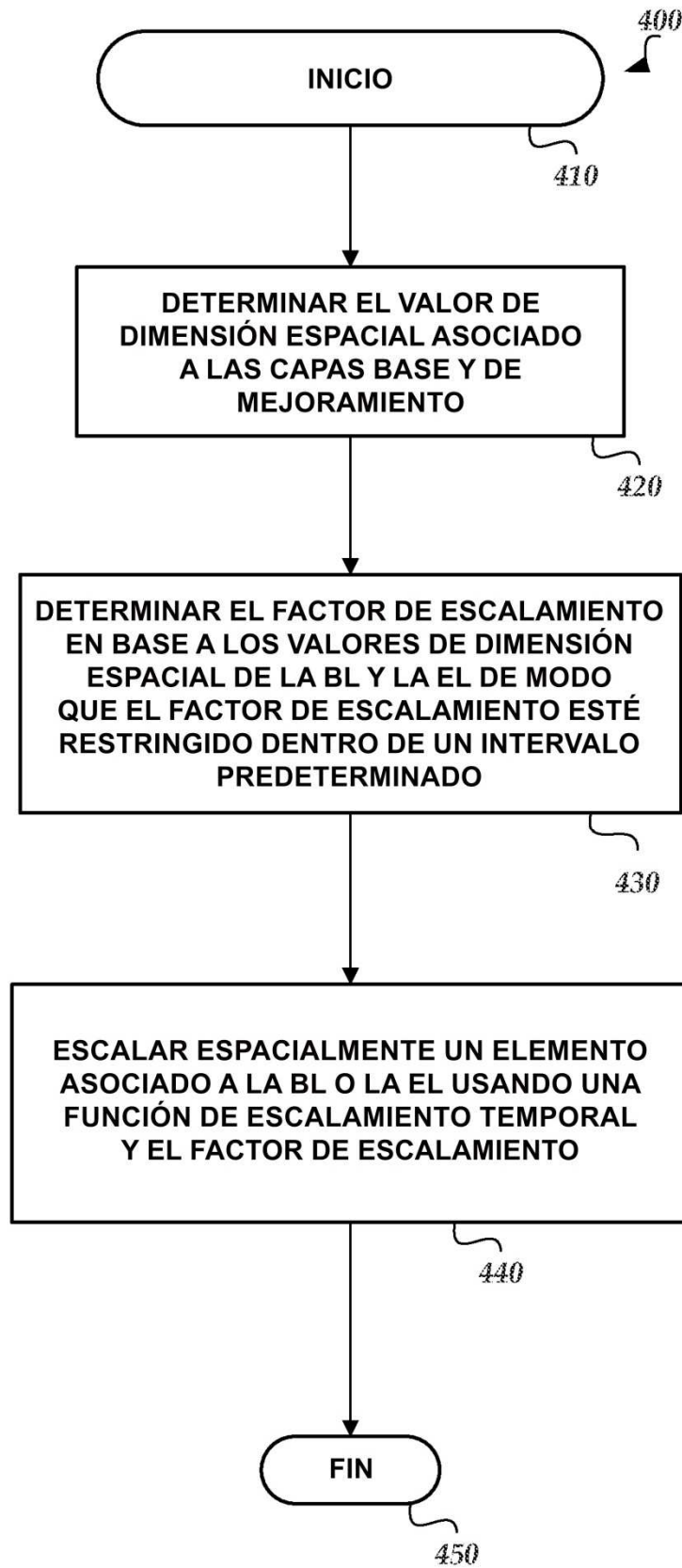


FIG. 4