

19



OFICINA ESPAÑOLA DE  
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 984 615**

51 Int. Cl.:

**H04N 19/577** (2014.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

86 Fecha de presentación y número de la solicitud internacional: **19.03.2020 PCT/CN2020/080156**

87 Fecha y número de publicación internacional: **24.09.2020 WO20187284**

96 Fecha de presentación y número de la solicitud europea: **19.03.2020 E 20773888 (1)**

97 Fecha y número de publicación de la concesión europea: **01.05.2024 EP 3942825**

54 Título: **Interpredicción de video basada en flujo óptico**

30 Prioridad:

**19.03.2019 IN 201931010751**

45 Fecha de publicación y mención en BOPI de la traducción de la patente:  
**30.10.2024**

73 Titular/es:

**HUAWEI TECHNOLOGIES CO., LTD. (100.0%)  
Huawei Administration Building, Bantian,  
Longgang District  
Shenzhen, Guangdong 518129, CN**

72 Inventor/es:

**SETHURAMAN, SRIRAM;  
A, JEEVA RAJ y  
KOTECHA, SAGAR**

74 Agente/Representante:

**PONS ARIÑO, Ángel**

**ES 2 984 615 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**

Interpredicción de video basada en flujo óptico

5 Campo técnico

Esta descripción se refiere a codificación y decodificación de video, especialmente se refiere a un método y aparato para interpredicción bi-predictiva que usa flujo óptico.

10 Antecedentes

La codificación de video (codificación y decodificación de video) se usa en una amplia gama de aplicaciones de video digital, por ejemplo, difusión de televisión digital, transmisión de video a través de Internet y redes móviles, aplicaciones conversacionales en tiempo real tal como chat de video, videoconferencias, DVD y discos Blu-ray, sistemas de adquisición y edición de contenido de video y videocámaras de aplicaciones de seguridad. La cantidad de datos de video necesarios para representar incluso un video relativamente corto puede ser sustancial, lo que puede dar por resultado dificultades cuando los datos se deben transmitir o comunicar de otra manera a través de una red de comunicaciones con capacidad de ancho de banda limitada. Por lo tanto, los datos de video generalmente se comprimen antes de comunicarse a través de redes de telecomunicaciones modernas. El tamaño de un video también podría ser un problema cuando el video se almacena en un dispositivo de almacenamiento porque los recursos de memoria pueden ser limitados.

Los dispositivos de compresión de video a menudo usan software y/o hardware en la fuente para codificar los datos de video antes de la transmisión o almacenamiento, disminuyendo así la cantidad de datos necesarios para representar imágenes de video digitales. Los datos comprimidos luego se reciben en el destino por un dispositivo de descompresión de video que decodifica los datos de video. Con recursos de red limitados y demandas cada vez mayores de mayor calidad de video, son deseables técnicas mejoradas de compresión y descompresión que mejoren la relación de compresión con poco o ningún sacrificio en la calidad de imagen.

En la compresión de video, la interpredicción es un proceso de uso de muestras reconstruidas de referencia previamente decodificadas mediante la especificación de vectores de movimiento con respecto a un bloque actual. Estos vectores de movimiento se pueden codificar como un residuo de predicción mediante el uso de predictores de vectores de movimiento espaciales o temporales. Los vectores de movimiento pueden tener una precisión de subpíxeles. Con el fin de derivar los valores de píxeles de precisión de subpíxeles en los cuadros de referencia de los valores enteros de posición reconstruidos, se aplica un filtro de interpolación.

La bi-predicción se refiere a un proceso donde la predicción para el bloque actual se deriva como una combinación ponderada de dos bloques de predicción derivados usando dos vectores de movimiento de dos áreas de imagen de referencia. En este caso, además de los vectores de movimiento, los índices de referencia para las imágenes de referencia a partir de las cuales se derivan los dos bloques de predicción también se necesitan codificar. Los vectores de movimiento para el bloque actual también se pueden derivar a través de un proceso de fusión donde los vectores de movimiento e índices de referencia de un vecino espacial se heredan sin codificar ningún residuo de vector de movimiento. Además de vecinos espaciales, los vectores de movimiento de cuadros de referencia previamente codificadas también se almacenan y usan como opciones de fusión temporal con escalado apropiado de los vectores de movimiento para cuidar la distancia a los cuadros de referencia con respecto a la distancia a los cuadros de referencia para el bloque actual.

El flujo óptico bidireccional (BPOF) es un refinamiento de movimiento por muestras que se realiza en la parte superior de la compensación de movimiento por bloques para bi-predicción. Dado que la estimación convencional del flujo óptico tiene problemas de complejidad o una brecha de eficiencia de compresión, existe una necesidad de dispositivos y métodos mejorados para interpredicción basada en flujo óptico.

LI (PANASONIC) J ET AL: "CE9: Simplification of BDOF's optical flow parameter derivation (Test 9.2.5)", REUNIÓN MPEG 126; 20190325 - 20190329; GINEBRA; (GRUPO DE EXPERTOS DE IMAGEN EN MOVIMIENTO O ISO/IEC JTC1/SC29/WG11), n.º M46856 13 de marzo de 2019 (2019-03-13), XP030209672, URL: [http://phenix.int-evry.fr/mpeg/doc\\_end\\_user/documents/126\\_Geneva/wg11/m46856-JVET-N0198-v2-N0198.zip](http://phenix.int-evry.fr/mpeg/doc_end_user/documents/126_Geneva/wg11/m46856-JVET-N0198-v2-N0198.zip) N0198/JVET-N0198-V1.docx, describe la simplificación de la derivación de parámetros de flujo óptico de BDOF.

60 Compendio

Los ejemplos del presente documento descripción proporcionan aparatos y métodos de interpredicción para codificación y decodificación de una imagen a través de interpredicción bi-predictiva que usa flujo óptico que puede mejorar la eficiencia de una señal de video. La descripción se desarrolla en los ejemplos y reivindicaciones contenidas en este archivo.

Según un primer aspecto, la descripción se refiere a un método para interpredicción basada en flujo óptico bidireccional (BDOF) para un bloque actual de una señal de video, que comprende:

5 determinar un desplazamiento de movimiento horizontal,  $v_x$ , y un desplazamiento de movimiento vertical,  $v_y$ , del bloque actual y en donde el desplazamiento de movimiento vertical se deriva con base en el desplazamiento de movimiento horizontal, una segunda variable  $s_2$ , una cuarta variable  $s_4$  y una quinta variable  $s_5$ , en donde:

10 la quinta variable  $s_5$  indica una suma de una pluralidad de primeros términos, en donde cada uno de la pluralidad de primeros términos se obtiene de un signo de un elemento de una segunda matriz y un elemento de una primera matriz, y el elemento de la primera matriz corresponde al elemento de la segunda matriz; cada elemento de la primera matriz se obtiene a partir de una suma de un primer gradiente de muestra predicho horizontal que corresponde a un primer cuadro de referencia del bloque actual y un segundo gradiente de muestra predicho horizontal que corresponde a un segundo cuadro de referencia del bloque actual, en donde el primer gradiente de muestra predicho horizontal y el segundo gradiente de muestra predicho horizontal corresponden al elemento de la primera matriz; y

15 cada elemento de la segunda matriz se obtiene a partir de una suma de un primer gradiente de muestra predicho vertical que corresponde al primer cuadro de referencia del bloque actual y un segundo gradiente de muestra predicho vertical que corresponde al segundo cuadro de referencia del bloque actual, en donde el primer gradiente de muestra predicho vertical y el segundo gradiente de muestra predicho vertical corresponden al elemento de la segunda matriz; y

20 en donde la segunda variable  $s_2$  indica una suma de valores absolutos de elementos de la segunda matriz y la cuarta variable  $s_4$  indica una suma de una pluralidad de segundos términos, en donde cada uno de la pluralidad de segundos términos se obtiene a partir de un signo de un elemento de la segunda matriz y un elemento de una tercera matriz, y el elemento de la tercera matriz corresponde al elemento de la segunda matriz, y cada elemento de la tercera matriz es una diferencia que se obtiene a partir de una primera muestra prevista del primer marco de referencia que corresponde al elemento de la tercera matriz y una segunda muestra prevista del segundo marco de referencia que corresponde al elemento de la tercera matriz; donde el desplazamiento horizontal se obtiene a partir de una primera variable  $s_1$  y una tercera variable  $s_3$ ; la primera variable  $s_1$  indica una suma de valores absolutos de elementos de la primera matriz; y la tercera variable  $s_3$  indica una suma de una pluralidad de terceros términos, donde cada uno de la pluralidad de terceros términos se obtiene a partir de un signo de un elemento de la primera matriz y un elemento de la tercera matriz, y el elemento de la tercera matriz corresponde al elemento de la primera matriz; y

35 determinar valores de muestra de predicción en el bloque actual usando valores de muestra de predicción correspondientes al primer fotograma de referencia, valores de muestra de predicción correspondientes al segundo fotograma de referencia, y el desplazamiento de movimiento horizontal y el desplazamiento de movimiento vertical.

40 Se observa que el bloque actual puede ser un bloque de cualquier tamaño, tal como un bloque de 4x4. El bloque actual puede ser un subbloque de un fotograma de la señal de video. Los píxeles del bloque actual se pueden referir usando posiciones absolutas de los píxeles con respecto a la esquina superior izquierda (por ejemplo, el píxel superior izquierdo) del cuadro, por ejemplo, (x,y), o posiciones relativas de los píxeles con respecto a la esquina superior izquierda (por ejemplo, el píxel superior izquierdo) del bloque, por ejemplo, (xBloque + i, yBloque+j). Aquí, (xBloque, yBloque) es la coordenada de la esquina superior izquierda (por ejemplo, el píxel superior izquierdo) del bloque con respecto a la esquina superior izquierda (por ejemplo, el píxel superior izquierdo) del fotograma.

50 Se puede entender que los términos "valor de píxel de predicción/valor de muestra de predicción", los términos "muestra/píxel" y "ubicación de muestra/ubicación de píxel" se pueden intercambiar entre sí en el presente documento descripción.

55 La primera matriz y la segunda matriz pueden ser cualquier matriz bidimensional que incluya filas y columnas y se puede hacer referencia a un elemento de la matriz usando (i,j) donde x es un índice horizontal/de fila e y es un índice vertical/de columna. El intervalo de i y j puede ser, por ejemplo,  $i = x\text{Bloque} - 1, \dots, x\text{Bloque} + 4$  y  $j = y\text{Bloque} - 1, \dots, y\text{Bloque} + 4$ . La primera y segunda matrices corresponden a, o se determinan para el bloque actual. En algunos ejemplos, el tamaño de la primera matriz es el mismo que el tamaño de la segunda matriz que puede ser mayor que el tamaño del bloque actual. Por ejemplo, el tamaño de la primera matriz y la segunda matriz puede ser 6x6, mientras que el bloque actual tiene un tamaño de 4x4.

60 Un elemento de la primera matriz (primer elemento) corresponde a un elemento de la segunda matriz (segundo elemento) si la posición del primer elemento en la primera matriz, (x,y) es la misma que la posición del segundo elemento en la segunda matriz, (p,q), es decir, (x,y)=(p,q). El primer gradiente de muestra predicho horizontal corresponde a un primer fotograma de referencia del bloque actual, lo que significa que el primer gradiente de

muestra predicho horizontal se genera con base en las muestras en el primer fotograma de referencia del bloque actual. El segundo gradiente de muestra predicho horizontal corresponde al segundo fotograma de referencia del bloque actual, lo que significa que el segundo gradiente de muestra predicho horizontal se genera con base en muestras en el segundo fotograma de referencia del bloque actual. El primer gradiente de muestra predicho horizontal corresponde al elemento de la primera matriz, lo que significa que el primer gradiente de muestra predicho horizontal se genera para la posición (x,y) del elemento en la primera matriz. Asimismo, el segundo gradiente de muestra predicho horizontal corresponde al elemento de la primera matriz, lo que significa que el segundo gradiente de muestra predicho horizontal se genera para la posición (x,y) del elemento en la primera matriz.

El primer gradiente de muestra predicho vertical corresponde al primer fotograma de referencia del bloque actual, lo que significa que el primer gradiente de muestra predicho vertical se genera con base en las muestras en el primer fotograma de referencia del bloque actual. El segundo gradiente de muestra predicho vertical corresponde al segundo fotograma de referencia del bloque actual, lo que significa que el segundo gradiente de muestra predicho vertical se genera con base en muestras en el segundo fotograma de referencia del bloque actual. El primer gradiente de muestra predicho vertical corresponde al elemento de la segunda matriz, lo que significa que el primer gradiente de muestra predicho vertical se genera para la posición (p,q) del elemento en la segunda matriz. Del mismo modo, el segundo gradiente de muestra predicho vertical corresponde al elemento de la segunda matriz, lo que significa que el segundo gradiente de muestra predicho vertical se genera para la posición (p,q) del elemento en la segunda matriz.

Cada elemento de una matriz se obtiene a partir de una suma de dos términos, lo que significa que el elemento se puede determinar como la suma de dos términos en sí o como un valor después de procesar la suma de dos términos. El procesamiento puede incluir desplazamiento a la izquierda, desplazamiento a la derecha, recorte, o una combinación de los mismos. De manera similar, un término se obtiene de un signo de un elemento de una segunda matriz y un elemento de una primera matriz, lo que significa que el término puede tomar el valor del elemento de la primera matriz por sí mismo o después de que el elemento de la primera matriz se procesa y aplicar el signo del elemento de la segunda matriz. El procesamiento del elemento de la primera matriz puede incluir desplazamiento hacia la izquierda, desplazamiento hacia la derecha, recorte, o una combinación de los mismos. El signo del elemento x se puede determinar como

$$\text{Signo}(x) = \begin{cases} 1 & ; & x > 0 \\ 0 & ; & x = 0 \\ -1 & ; & x < 0 \end{cases}$$

La tecnología presentada en el presente documento ajusta los valores de muestra bi-predichos del bloque actual con base en el desplazamiento de movimiento horizontal y el desplazamiento de movimiento vertical. El desplazamiento de movimiento vertical se calcula con base en una quinta variable  $s_5$  que solo implica la suma de los primeros términos obtenidos de un signo de un elemento de una segunda matriz y un elemento de una primera matriz. La aplicación de un signo de un elemento a otro elemento no implica operaciones de multiplicación. Del mismo modo, la suma tampoco implica multiplicaciones. Como resultado, la tecnología de interpredicción basada en BDOF presentada en el presente documento elimina las operaciones de multiplicación. En comparación con los enfoques tradicionales, la profundidad de bits de la quinta variable  $s_5$  se reduce porque las operaciones de multiplicación se reemplazan por la determinación de signo. Esto conduce a la reducción de la profundidad de bits del desplazamiento de movimiento horizontal,  $v_x$  y el desplazamiento de movimiento vertical,  $v_y$ , y también una reducción significativa en la complejidad de cálculo de la predicción y el tamaño del multiplicador.

Se debe señalar que un elemento de la tercera matriz (tercer elemento) corresponde a un elemento de la segunda matriz (segundo elemento) si la posición del tercer elemento en la tercera matriz, (k,l) es la misma que la posición del segundo elemento en la segunda matriz, (p,q), es decir, (k,l)=(p,q). La primera muestra predicha corresponde a un primer fotograma de referencia del bloque actual, lo que significa que la primera muestra predicha se encuentra en el primer fotograma de referencia del bloque actual. La segunda muestra predicha corresponde al segundo fotograma de referencia del bloque actual, lo que significa que la segunda muestra predicha se encuentra en el segundo fotograma de referencia del bloque actual. La primera muestra predicha corresponde al elemento de la tercera matriz, lo que significa que la primera muestra predicha se encuentra en la posición (k,l) del elemento en la tercera matriz. Del mismo modo, la segunda muestra predicha corresponde al elemento de la tercera matriz, lo que significa que la segunda muestra predicha se encuentra en la posición (k,l) del elemento en la tercera matriz.

Los valores adicionales, la segunda variable  $s_2$  y la cuarta variable  $s_4$  involucradas en el cálculo del desplazamiento de movimiento vertical del bloque actual tampoco implican operaciones de multiplicación. De manera similar a la quinta variable  $s_5$ , el cálculo de la cuarta variable  $s_4$  solo implica la suma de los segundos

términos obtenidos de un signo de un elemento de una segunda matriz y un elemento de una tercera matriz. La aplicación de un signo de un elemento a otro elemento no implica operaciones de multiplicación. Además, la suma tampoco implica multiplicaciones. En comparación con los enfoques tradicionales, la profundidad de bits de la cuarta variable  $s_4$  se reduce porque las operaciones de multiplicación se reemplazan por la determinación de signo. Del mismo modo, el cálculo de la segunda variable  $s_2$  implica tomar valores absolutos de elementos de la segunda matriz y no hay multiplicaciones involucradas. Como tal, la profundidad de bits de  $s_2$  también se reduce. En consecuencia, al calcular el desplazamiento de movimiento vertical, la profundidad de bits de los resultados también se reduce y la complejidad de cálculo se reduce significativamente.

El elemento de la tercera matriz (tercer elemento) corresponde al elemento de la primera matriz (primer elemento) significa que la posición del primer elemento en la primera matriz, (x,y) es la misma que la posición del tercer elemento en la tercera matriz, (k,l), es decir, (x,y)=(k,l).

Derivando el desplazamiento de movimiento horizontal basado en la primera variable  $s_1$  y una tercera variable  $s_3$ , se puede conseguir una reducción adicional de la complejidad de cálculo. Tanto el cálculo de la primera variable  $s_1$  como de una tercera variable  $s_3$  no implican multiplicaciones. En su lugar, sólo se toman valores absolutos, operaciones de signo y operaciones de suma.

En una forma de implementación posible del método según cualquier implementación anterior del primer aspecto o el primer aspecto como tal, el desplazamiento de movimiento horizontal se determina según:

$$v_x = -\frac{s_3}{s_1}$$

en donde  $v_x$  representa el desplazamiento de movimiento horizontal.

Esto muestra una posible forma de determinar  $v_x$  basándose en los términos de autocorrelación y correlación cruzada  $s_1$  y  $s_3$ .  $s_1$  y  $s_3$  se determinan sin operaciones de multiplicación, lo que reduce significativamente la complejidad de cálculo del proceso para determinar  $v_x$ . De este modo, calculando el desplazamiento horizontal de este modo, el desplazamiento horizontal puede determinarse eficientemente.

En una forma de implementación posible del método según cualquier implementación anterior del primer aspecto o el primer aspecto como tal, el desplazamiento de movimiento vertical  $v_y$  se determina según:

$$v_y = -\frac{(s_3 - v_x * s_2/2)}{s_2}$$

en donde  $v_x$  representa el desplazamiento de movimiento horizontal; y  $v_y$  representa el desplazamiento de movimiento vertical.

Esto muestra una posible manera de determinar  $v_y$  con base en términos de autocorrelación y correlación cruzada  $s_2$ ,  $s_4$  y  $s_5$ .  $s_2$ ,  $s_4$  y  $s_5$  se determinan sin operaciones de multiplicación, que reduce significativamente la complejidad de cálculo del proceso para determinar  $v_y$ . Como tal, al calcular el desplazamiento de movimiento vertical de esta manera, el desplazamiento de movimiento vertical se puede determinar de manera eficiente.

En una forma de implementación posible del método según cualquier implementación anterior del primer aspecto o el primer aspecto como tal,  $s_1$ ,  $s_2$ ,  $s_3$ ,  $s_4$  y  $s_5$  se determinan como:

$$s_1 = \sum_{i,j} abs(G_{x1} + G_{x0});$$

$$s_2 = \sum_{i,j} abs(G_{y1} + G_{y0});$$

$$s_3 = \sum_{i,j} signo(G_{x1} + G_{x0}) * (I^{(1)} - I^{(0)}).$$

$$s_3 = \sum_{i,j} \text{signo}(G_{y1} + G_{y0}) * (I^{(1)} - I^{(0)}),$$

$$s_5 = \sum_{i,j} \text{signo}(G_{y1} + G_{y0}) * (G_{x1} + G_{x0});$$

5 y en donde:

$I^{(0)}$  se obtiene a partir de los valores de muestra predichos que corresponden al primer fotograma de referencia, y  $I^{(1)}$  se obtiene a partir de los valores de muestra predichos que corresponden al segundo fotograma de referencia;

10

$G_{x0}$  y  $G_{x1}$  indica un conjunto de los gradientes de muestra predichos horizontales que corresponden al primer fotograma de referencia y al segundo fotograma de referencia, respectivamente;

15

$G_{y0}$  y  $G_{y1}$  indican un conjunto de los gradientes de muestra predichos verticales que corresponden al primer fotograma de referencia y al segundo fotograma de referencia, respectivamente; y

i y j son números enteros, en donde un valor de i varía de -1 a 4, y un valor de j varía de -1 a 4.

20

En algunos ejemplos,  $I^{(0)}$  es un bloque que incluye muestras predichas en el primer fotograma de referencia que se encuentran alrededor de un subbloque correspondiente al bloque actual. Por ejemplo, si el bloque actual es un bloque de 4x4,  $I^{(0)}$  puede ser un bloque de 6x6 en el primer fotograma de referencia que rodea el bloque de 4x4 que corresponde al bloque actual de 4x4. De manera similar,  $I^{(1)}$  es un bloque que incluye muestras predichas en el segundo fotograma de referencia que se encuentran alrededor de un subbloque correspondiente al bloque actual. Por ejemplo, si el bloque actual es un bloque de 4x4,  $I^{(1)}$  puede ser un bloque de 6x6 en el segundo fotograma de referencia que rodea el bloque de 4x4 que corresponde al bloque actual de 4x4.

25

Si el bloque actual es un bloque de 4x4 y el  $I^{(0)}$  y  $I^{(1)}$  son ambos bloques de 6x6, cada uno de  $G_{x0}$  y  $G_{x1}$  es un bloque de 6x6.

30

Se observa que al calcular los términos de autocorrelación y correlación cruzadas  $s_1, s_2, s_3, s_4$  y  $s_5$ , se puede aplicar desplazamiento para ajustar la precisión y/o profundidad de bits de  $s_1, s_2, s_3, s_4$  y  $s_5$ .

35

Se observa además que  $s_1, s_2, s_3, s_4$  y  $s_5$  se determinan sin operaciones de multiplicación, que reduce significativamente la complejidad de cálculo del proceso para determinar  $v_x$  y  $v_y$ .  $v_x$  y  $v_y$  se derivan con base en términos de autocorrelación y correlación cruzada como se muestra anteriormente con respecto a  $s_1, s_2, s_3, s_4$  y  $s_5$ .  $v_x$  y  $v_y$  se pueden determinar como  $v_x = \frac{s_3}{s_1}$  y  $v_y = \frac{(s_4 - v_x * s_5 / 2)}{s_2}$  si el término  $(I^{(1)} - I^{(0)})$  se cambia a  $(I^{(0)} - I^{(1)})$ .

40

En una forma de implementación posible del método según cualquier implementación anterior del primer aspecto o el primer aspecto como tal, el  $G_{x0}$  se determina como una diferencia obtenida de dos muestras predichas que corresponden al primer fotograma de referencia a lo largo de una dirección horizontal, y el  $G_{y0}$  se determina como una diferencia obtenida de dos muestras predichas que corresponden al primer fotograma de referencia a lo largo de una dirección vertical.

45

En algunos ejemplos, las dos muestras predichas que corresponden al primer fotograma de referencia a lo largo de una dirección horizontal tienen las mismas coordenadas verticales y diferentes coordenadas horizontales. Las dos muestras predichas que corresponden al primer fotograma de referencia a lo largo de una dirección vertical tienen las mismas coordenadas horizontales y diferentes coordenadas verticales. Cada una de las dos muestras predichas se puede procesar antes de calcular la diferencia, como desplazamiento a la derecha, desplazamiento a la izquierda o recorte.

50

En una forma de implementación posible del método según cualquier implementación anterior del primer aspecto o el primer aspecto como tal, el  $G_{x1}$  se determina como una diferencia obtenida de dos muestras predichas que corresponden al segundo fotograma de referencia a lo largo de una dirección horizontal, y el  $G_{y1}$  se determina como una diferencia obtenida de dos muestras predichas que corresponden al segundo fotograma de referencia a lo largo de una dirección vertical.

55

5 En algunos ejemplos, las dos muestras predichas que corresponden al segundo fotograma de referencia a lo largo de una dirección horizontal tienen las mismas coordenadas verticales y diferentes coordenadas horizontales. Las dos muestras predichas que corresponden al segundo fotograma de referencia a lo largo de una dirección vertical tienen las mismas coordenadas horizontales y diferentes coordenadas verticales. Cada una de las dos muestras predichas se puede procesar antes de calcular la diferencia, como desplazamiento a la derecha, desplazamiento a la izquierda o recorte.

10 En una forma de implementación posible del método según cualquier implementación anterior del primer aspecto o el primer aspecto como tal, los valores de muestra de predicción correspondientes al primer fotograma de referencia y los valores de muestra de predicción correspondientes al segundo fotograma de referencia se obtienen de los primeros y segundos fotogramas de referencia, respectivamente, usando un par de vectores de movimiento para el bloque actual con respecto a los primeros y segundos fotogramas de referencia.

15 En una forma de implementación posible del método según cualquier implementación anterior del primer aspecto o el primer aspecto como tal, los valores de muestra de predicción para el bloque actual son valores de muestra bi-predichos con base en la predicción de flujo óptico bidireccional (BDOF).

20 Según un segundo aspecto, la descripción se refiere a un dispositivo para codificación de datos de video, que comprende:

una memoria de datos de video; y

25 un codificador de video, en donde el codificador de video se configura para:

determinar un desplazamiento de movimiento horizontal,  $v_x$  y un desplazamiento de movimiento vertical,  $v_y$ , de un bloque actual de una señal de video, y en donde el desplazamiento de movimiento vertical se determina con base en el desplazamiento de movimiento horizontal y una quinta variable  $s_5$ , en donde:

30 la quinta variable  $s_5$  indica una suma de una pluralidad de términos, en donde cada uno de la pluralidad de términos se obtiene de un signo de un elemento de una segunda matriz y un elemento de una primera matriz, y el elemento de la primera matriz corresponde al elemento de la segunda matriz;

35 cada elemento de la primera matriz se obtiene a partir de una suma de un primer gradiente de muestra predicho horizontal que corresponde a un primer fotograma de referencia del bloque actual y un segundo gradiente de muestra predicho horizontal que corresponde a un segundo fotograma de referencia del bloque actual, en donde el primer gradiente de muestra predicho horizontal y el segundo gradiente de muestra predicho horizontal corresponden al elemento de la primera matriz; y

40 cada elemento de la segunda matriz se obtiene a partir de una suma de un primer gradiente de muestra predicho vertical que corresponde al primer fotograma de referencia del bloque actual y un segundo gradiente de muestra predicho vertical que corresponde al segundo fotograma de referencia del bloque actual, en donde el primer gradiente de muestra predicho vertical y el segundo gradiente de muestra predicho vertical corresponden al elemento de la segunda matriz; y

45 determinar valores de muestra de predicción en el bloque actual usando valores de muestra de predicción correspondientes al primer fotograma de referencia, valores de muestra de predicción correspondientes al segundo fotograma de referencia, y el desplazamiento de movimiento horizontal y el desplazamiento de movimiento vertical.

50 Se observa que el bloque actual puede ser un bloque de cualquier tamaño, tal como un bloque de 4x4. El bloque actual puede ser un subbloque de un fotograma de la señal de video. Los píxeles del bloque actual se pueden referir usando posiciones absolutas de los píxeles con respecto a la esquina superior izquierda del fotograma, por ejemplo, (x,y), o posiciones relativas de los píxeles con respecto a la esquina superior izquierda del bloque, por ejemplo, (xBloque + i, yBloque+j). Aquí, (xBloque, yBloque) es la coordenada de la esquina superior izquierda del bloque con respecto a la esquina superior izquierda del fotograma.

60 La primera matriz y la segunda matriz pueden ser cualquier matriz bidimensional que incluya filas y columnas y se puede hacer referencia a un elemento de la matriz usando (i,j) donde x es un índice horizontal/de fila e y es un índice vertical/de columna. El intervalo de i y j puede ser, por ejemplo,  $i = \text{xBloque} - 1, \dots, \text{xBloque} + 4$  y  $j = \text{yBloque} - 1, \dots, \text{yBloque} + 4$ . La primera y segunda matrices corresponden a, o se determinan para el bloque actual. En algunos ejemplos, el tamaño de la primera matriz es el mismo que el tamaño de la segunda matriz que puede ser mayor que el tamaño del bloque actual. Por ejemplo, el tamaño de la primera matriz y la segunda matriz puede ser 6x6, mientras que el bloque actual tiene un tamaño de 4x4.

65 Un elemento de la primera matriz (primer elemento) corresponde a un elemento de la segunda matriz (segundo

elemento) si la posición del primer elemento en la primera matriz, (x,y) es la misma que la posición del segundo elemento en la segunda matriz, (p,q), es decir, (x,y)=(p,q). El primer gradiente de muestra predicho horizontal corresponde a un primer fotograma de referencia del bloque actual, lo que significa que el primer gradiente de muestra predicho horizontal se genera con base en las muestras en el primer fotograma de referencia del bloque actual. El segundo gradiente de muestra predicho horizontal corresponde al segundo fotograma de referencia del bloque actual, lo que significa que el segundo gradiente de muestra predicho horizontal se genera con base en muestras en el segundo fotograma de referencia del bloque actual. El primer gradiente de muestra predicho horizontal corresponde al elemento de la primera matriz, lo que significa que el primer gradiente de muestra predicho horizontal se genera para la posición (x,y) del elemento en la primera matriz. Asimismo, el segundo gradiente de muestra predicho horizontal corresponde al elemento de la primera matriz, lo que significa que el segundo gradiente de muestra predicho horizontal se genera para la posición (x,y) del elemento en la primera matriz.

El primer gradiente de muestra predicho vertical corresponde al primer fotograma de referencia del bloque actual, lo que significa que el primer gradiente de muestra predicho vertical se genera con base en las muestras en el primer fotograma de referencia del bloque actual. El segundo gradiente de muestra predicho vertical corresponde al segundo fotograma de referencia del bloque actual, lo que significa que el segundo gradiente de muestra predicho vertical se genera con base en muestras en el segundo fotograma de referencia del bloque actual. El primer gradiente de muestra predicho vertical corresponde al elemento de la segunda matriz, lo que significa que el primer gradiente de muestra predicho vertical se genera para la posición (p,q) del elemento en la segunda matriz. Del mismo modo, el segundo gradiente de muestra predicho vertical corresponde al elemento de la segunda matriz, lo que significa que el segundo gradiente de muestra predicho vertical se genera para la posición (p,q) del elemento en la segunda matriz.

Cada elemento de una matriz se obtiene a partir de una suma de dos términos, lo que significa que el elemento se puede determinar como la suma de dos términos en sí o como un valor después de procesar la suma de dos términos. El procesamiento puede incluir desplazamiento a la izquierda, desplazamiento a la derecha, recorte, o una combinación de los mismos. De manera similar, un término se obtiene de un signo de un elemento de una segunda matriz y un elemento de una primera matriz, lo que significa que el término puede tomar el valor del elemento de la primera matriz por sí mismo o después de que el elemento de la primera matriz se procesa y aplicar el signo del elemento de la segunda matriz. El procesamiento del elemento de la primera matriz puede incluir desplazamiento hacia la izquierda, desplazamiento hacia la derecha, recorte, o una combinación de los mismos. El signo del elemento x se puede determinar como

$$\text{Signo}(x) = \begin{cases} 1 & ; & x > 0 \\ 0 & ; & x == 0 \\ -1 & ; & x < 0 \end{cases}$$

La tecnología presentada en el presente documento ajusta los valores de muestra bi-predichos del bloque actual con base en el desplazamiento de movimiento horizontal y el desplazamiento de movimiento vertical. El desplazamiento de movimiento vertical se calcula con base en una quinta variable  $s_5$  que solo implica la suma de términos obtenidos de un signo de un elemento de una segunda matriz y un elemento de una primera matriz. La aplicación de un signo de un elemento a otro elemento no implica operaciones de multiplicación. Del mismo modo, la suma tampoco implica multiplicaciones. Como resultado, la tecnología de interpredicción basada en BDOF presentada en el presente documento elimina las operaciones de multiplicación. En comparación con los enfoques tradicionales, la profundidad de bits de la quinta variable  $s_5$  se reduce porque las operaciones de multiplicación se reemplazan por la determinación de signo. Esto conduce a la reducción de la profundidad de bits del desplazamiento de movimiento horizontal,  $v_x$  y el desplazamiento de movimiento vertical,  $v_y$  y también una reducción significativa en la complejidad de cálculo de la predicción y el tamaño del multiplicador.

Según un tercer aspecto, la descripción se refiere a un dispositivo para decodificar datos de video, que comprende:

una memoria de datos de video; y

un decodificador de video, en donde el decodificador de video se configura para:

determinar un desplazamiento de movimiento horizontal,  $v_x$  y un desplazamiento de movimiento vertical,  $v_y$  de un bloque actual de una señal de video, y en donde el desplazamiento de movimiento vertical se determina con base en el desplazamiento de movimiento horizontal y una quinta variable  $s_5$ , en donde:

la quinta variable  $s_5$  indica una suma de una pluralidad de términos, en donde cada uno de la pluralidad de

## ES 2 984 615 T3

términos se obtiene de un signo de un elemento de una segunda matriz y un elemento de una primera matriz, y el elemento de la primera matriz corresponde al elemento de la segunda matriz;

5 cada elemento de la primera matriz se obtiene a partir de una suma de un primer gradiente de muestra predicho horizontal que corresponde a un primer fotograma de referencia del bloque actual y un segundo gradiente de muestra predicho horizontal que corresponde a un segundo fotograma de referencia del bloque actual, en donde el primer gradiente de muestra predicho horizontal y el segundo gradiente de muestra predicho horizontal corresponden al elemento de la primera matriz; y

10 cada elemento de la segunda matriz se obtiene a partir de una suma de un primer gradiente de muestra predicho vertical que corresponde al primer fotograma de referencia del bloque actual y un segundo gradiente de muestra predicho vertical que corresponde al segundo fotograma de referencia del bloque actual, en donde el primer gradiente de muestra predicho vertical y el segundo gradiente de muestra predicho vertical corresponden al elemento de la segunda matriz; y

15 determinar valores de muestra de predicción en el bloque actual usando valores de muestra de predicción correspondientes al primer fotograma de referencia, valores de muestra de predicción correspondientes al segundo fotograma de referencia, y el desplazamiento de movimiento horizontal y el desplazamiento de movimiento vertical.

20 Se observa que el bloque actual puede ser un bloque de cualquier tamaño, tal como un bloque de 4x4. El bloque actual puede ser un subbloque de un fotograma de la señal de video. Los píxeles del bloque actual se pueden referir usando posiciones absolutas de los píxeles con respecto a la esquina superior izquierda del fotograma, por ejemplo, (x,y), o posiciones relativas de los píxeles con respecto a la esquina superior izquierda del bloque, por ejemplo, (xBloque + i, yBloque+j). Aquí, (xBloque, yBloque) es la coordenada de la esquina superior izquierda del bloque con respecto a la esquina superior izquierda del fotograma.

25 La primera matriz y la segunda matriz pueden ser cualquier matriz bidimensional que incluya filas y columnas y se puede hacer referencia a un elemento de la matriz usando (i,j) donde x es un índice horizontal/de fila e y es un índice vertical/de columna. El intervalo de i y j puede ser, por ejemplo,  $i = \text{xBloque} - 1, \dots, \text{xBloque} + 4$  y  $j = \text{yBloque} - 1, \dots, \text{yBloque} + 4$ . La primera y segunda matrices corresponden a, o se determinan para el bloque actual. En algunos ejemplos, el tamaño de la primera matriz es el mismo que el tamaño de la segunda matriz que puede ser mayor que el tamaño del bloque actual. Por ejemplo, el tamaño de la primera matriz y la segunda matriz puede ser 6x6, mientras que el bloque actual tiene un tamaño de 4x4.

30 Un elemento de la primera matriz (primer elemento) corresponde a un elemento de la segunda matriz (segundo elemento) si la posición del primer elemento en la primera matriz, (x,y) es la misma que la posición del segundo elemento en la segunda matriz, (p,q), es decir,  $(x,y)=(p,q)$ . El primer gradiente de muestra predicho horizontal corresponde a un primer fotograma de referencia del bloque actual, lo que significa que el primer gradiente de muestra predicho horizontal se genera con base en las muestras en el primer fotograma de referencia del bloque actual. El segundo gradiente de muestra predicho horizontal corresponde al segundo fotograma de referencia del bloque actual, lo que significa que el segundo gradiente de muestra predicho horizontal se genera con base en muestras en el segundo fotograma de referencia del bloque actual. El primer gradiente de muestra predicho horizontal corresponde al elemento de la primera matriz, lo que significa que el primer gradiente de muestra predicho horizontal se genera para la posición (x,y) del elemento en la primera matriz. Asimismo, el segundo gradiente de muestra predicho horizontal corresponde al elemento de la primera matriz, lo que significa que el segundo gradiente de muestra predicho horizontal se genera para la posición (x,y) del elemento en la primera matriz.

35 El primer gradiente de muestra predicho vertical corresponde al primer fotograma de referencia del bloque actual, lo que significa que el primer gradiente de muestra predicho vertical se genera con base en las muestras en el primer fotograma de referencia del bloque actual. El segundo gradiente de muestra predicho vertical corresponde al segundo fotograma de referencia del bloque actual, lo que significa que el segundo gradiente de muestra predicho vertical se genera con base en muestras en el segundo fotograma de referencia del bloque actual. El primer gradiente de muestra predicho vertical corresponde al elemento de la segunda matriz, lo que significa que el primer gradiente de muestra predicho vertical se genera para la posición (p,q) del elemento en la segunda matriz. Del mismo modo, el segundo gradiente de muestra predicho vertical corresponde al elemento de la segunda matriz, lo que significa que el segundo gradiente de muestra predicho vertical se genera para la posición (p,q) del elemento en la segunda matriz.

40 Cada elemento de una matriz se obtiene a partir de una suma de dos términos, lo que significa que el elemento se puede determinar como la suma de dos términos en sí o como un valor después de procesar la suma de dos términos. El procesamiento puede incluir desplazamiento a la izquierda, desplazamiento a la derecha, recorte, o una combinación de los mismos. De manera similar, un término se obtiene de un signo de un elemento de una segunda matriz y un elemento de una primera matriz, lo que significa que el término puede tomar el valor del elemento de la primera matriz por sí mismo o después de que el elemento de la primera

matriz se procesa y aplicar el signo del elemento de la segunda matriz. El procesamiento del elemento de la primera matriz puede incluir desplazamiento hacia la izquierda, desplazamiento hacia la derecha, recorte, o una combinación de los mismos. El signo del elemento  $x$  se puede determinar como

$$\text{Signo}(x) = \begin{cases} 1 & ; & x > 0 \\ 0 & ; & x = 0 \\ -1 & ; & x < 0 \end{cases}$$

5

La tecnología presentada en el presente documento ajusta los valores de muestra bi-predichos del bloque actual con base en el desplazamiento de movimiento horizontal y el desplazamiento de movimiento vertical. El desplazamiento de movimiento vertical se calcula con base en una quinta variable  $s_5$  que solo implica la suma de términos obtenidos de un signo de un elemento de una segunda matriz y un elemento de una primera matriz. La aplicación de un signo de un elemento a otro elemento no implica operaciones de multiplicación. Del mismo modo, la suma tampoco implica multiplicaciones. Como resultado, la tecnología de interpredicción basada en BDOF presentada en el presente documento elimina las operaciones de multiplicación. En comparación con los enfoques tradicionales, la profundidad de bits de la quinta variable  $s_5$  se reduce porque las operaciones de multiplicación se reemplazan por la determinación de signo. Esto conduce a la reducción de la profundidad de bits del desplazamiento de movimiento horizontal,  $v_x$  y el desplazamiento de movimiento vertical,  $v_y$ , y también una reducción significativa en la complejidad de cálculo de la predicción y el tamaño del multiplicador.

En una forma de implementación posible del dispositivo según cualquier implementación anterior del segundo y tercer aspecto o el segundo y tercer aspecto como tal, el desplazamiento de movimiento vertical se deriva con base en el desplazamiento de movimiento horizontal, una segunda variable  $s_2$ , una cuarta variable  $s_4$  y la quinta variable  $s_5$ , en donde:

la segunda variable  $s_2$  indica una suma de valores absolutos de elementos de la segunda matriz; y

la cuarta variable  $s_4$  indica una suma de una pluralidad de términos, en donde cada uno de la pluralidad de términos se obtiene de un signo de un elemento de la segunda matriz y un elemento de una tercera matriz, y el elemento de la tercera matriz corresponde al elemento de la segunda matriz, y cada elemento de la tercera matriz es una diferencia que se obtiene de una primera muestra predicha del primer fotograma de referencia que corresponde al elemento de la tercera matriz y una segunda muestra predicha del segundo fotograma de referencia que corresponde al elemento de la tercera matriz.

Se debe señalar que un elemento de la tercera matriz (primer elemento) corresponde a un elemento de la segunda matriz (segundo elemento) si la posición del primer elemento en la tercera matriz,  $(k,l)$  es la misma que la posición del segundo elemento en la segunda matriz,  $(p,q)$ , es decir,  $(k,l)=(p,q)$ . La primera muestra predicha corresponde a un primer fotograma de referencia del bloque actual, lo que significa que la primera muestra predicha se encuentra en el primer fotograma de referencia del bloque actual. La segunda muestra prevista corresponde al segundo fotograma de referencia del bloque actual, lo que significa que la segunda muestra prevista se encuentra en el segundo fotograma de referencia del bloque actual. La primera muestra predicha corresponde al elemento de la tercera matriz, lo que significa que la primera muestra predicha se encuentra en la posición  $(k,l)$  del elemento en la tercera matriz. Del mismo modo, la segunda muestra predicha corresponde al elemento de la tercera matriz, lo que significa que la segunda muestra predicha se encuentra en la posición  $(k,l)$  del elemento en la tercera matriz.

Los valores adicionales, la segunda variable  $s_2$  y la cuarta variable  $s_4$  involucradas en el cálculo del desplazamiento de movimiento vertical del bloque actual tampoco implican operaciones de multiplicación. De manera similar a la quinta variable  $s_5$ , el cálculo de la cuarta variable  $s_4$  solo implica la suma de términos obtenidos de un signo de un elemento de una segunda matriz y un elemento de una tercera matriz. La aplicación de un signo de un elemento a otro elemento no implica operaciones de multiplicación. Además, la suma tampoco implica multiplicaciones. En comparación con los enfoques tradicionales, la profundidad de bits de la cuarta variable  $s_4$  se reduce porque las operaciones de multiplicación se reemplazan por la determinación de signo. Del mismo modo, el cálculo de la segunda variable  $s_2$  implica tomar valores absolutos de elementos de la segunda matriz y no hay multiplicaciones involucradas. Como tal, la profundidad de bits de  $s_2$  también se reduce. En consecuencia, al calcular el desplazamiento de movimiento vertical, la profundidad de bits de los resultados también se reduce y la complejidad de cálculo se reduce significativamente.

En una forma de implementación posible del dispositivo según cualquier implementación anterior del segundo y tercer aspecto o el segundo y tercer aspecto como tal, el desplazamiento de movimiento horizontal se deriva con base en una primera variable  $s_1$  y una tercera variable  $s_3$ ;

60

la primera variable  $s_1$  indica una suma de valores absolutos de elementos de la primera matriz; y

la tercera variable  $s_3$  indica una suma de una pluralidad de términos, en donde cada uno de la pluralidad de términos se obtiene de un signo de un elemento de la primera matriz y un elemento de la tercera matriz, y el elemento de la tercera matriz corresponde al elemento de la primera matriz.

El elemento de la tercera matriz (tercer elemento) corresponde al elemento de la primera matriz (primer elemento) significa que la posición del primer elemento en la primera matriz, (x,y) es la misma que la posición del tercer elemento en la tercera matriz, (k,l), es decir, (x,y)=(k,l).

Al derivar el desplazamiento de movimiento horizontal con base en la primera variable  $s_1$  y una tercera variable  $s_3$ , se puede lograr una reducción de complejidad de cálculo adicional. Tanto el cálculo de la primera variable  $s_1$  como de una tercera variable  $s_3$  no implican multiplicaciones. Más bien, solo están involucradas operaciones de toma de valores absolutos, toma de signos y operaciones de suma.

En una forma de implementación posible del dispositivo según cualquier implementación anterior del segundo y tercer aspecto o el segundo y tercer aspecto como tal, el desplazamiento de movimiento horizontal se determina según

$$v_x = -\frac{s_3}{s_1}$$

en donde:

$v_x$  representa el desplazamiento de movimiento horizontal.

Esto muestra una posible manera de determinar  $v_x$  con base en términos de autocorrelación y correlación cruzada  $s_1$  y  $s_3$ .  $s_1$  y  $s_3$  se determinan sin operaciones de multiplicación, que reduce significativamente la complejidad de cálculo del proceso para determinar  $v_x$ . Como tal, al calcular el desplazamiento de movimiento horizontal de esta manera, el desplazamiento de movimiento horizontal se puede determinar de manera eficiente.

En una forma de implementación posible del dispositivo según cualquier implementación anterior del segundo y tercer aspecto o el segundo y tercer aspecto como tal, el desplazamiento de movimiento vertical  $v_y$  se determina según

$$v_y = -\frac{(s_4 - v_x * s_5/2)}{s_2},$$

en donde:

$v_x$  representa el desplazamiento de movimiento horizontal; y

$v_y$  representa el desplazamiento de movimiento vertical.

Esto muestra una posible manera de determinar  $v_y$  con base en términos de autocorrelación y correlación cruzada  $s_2$ ,  $s_4$  y  $s_5$ .  $s_2$ ,  $s_4$  y  $s_5$  se determinan sin operaciones de multiplicación, que reduce significativamente la complejidad de cálculo del proceso para determinar. Como tal, al calcular el desplazamiento de movimiento vertical de esta manera, el desplazamiento de movimiento vertical se puede determinar de manera eficiente.

En una forma de implementación posible del dispositivo según cualquier implementación anterior del segundo y tercer aspecto o el segundo y tercer aspecto como tal,  $s_1$ ,  $s_2$ ,  $s_3$ ,  $s_4$  y  $s_5$  se determinan como:

$$s_1 = \sum_{i,j} abs(G_{x1} + G_{x0});$$

$$s_2 = \sum_{i,j} abs(G_{y1} + G_{y0}).$$

55

$$s_3 = \sum_{i,j} \text{signo}(G_{x1} + G_{x0}) * (I^{(1)} - I^{(0)}),$$

$$s_4 = \sum_{i,j} \text{signo}(G_{y1} + G_{y0}) * (I^{(1)} - I^{(0)}),$$

$$s_5 = \sum_{i,j} \text{signo}(G_{y1} + G_{y0}) * (G_{x1} + G_{x0});$$

5

y en donde:

10  $I^{(0)}$  se obtiene a partir de los valores de muestra predichos que corresponden al primer fotograma de referencia, y  $I^{(1)}$  se obtiene a partir de los valores de muestra predichos que corresponden al segundo fotograma de referencia;

15  $G_{x0}$  y  $G_{x1}$  indica un conjunto de los gradientes de muestra predichos horizontales que corresponden al primer fotograma de referencia y al segundo fotograma de referencia, respectivamente;

20  $G_{y0}$  y  $G_{y1}$  indican un conjunto de los gradientes de muestra predichos verticales que corresponden al primer fotograma de referencia y al segundo fotograma de referencia, respectivamente; y

25  $i$  y  $j$  son números enteros, en donde un valor de  $i$  varía de -1 a 4, y un valor de  $j$  varía de -1 a 4.

30 En algunos ejemplos,  $I^{(0)}$  es un bloque que incluye muestras predichas en el primer fotograma de referencia que se encuentran alrededor de un subbloque correspondiente al bloque actual. Por ejemplo, si el bloque actual es un bloque de 4x4,  $I^{(0)}$  puede ser un bloque de 6x6 en el primer fotograma de referencia que rodea el bloque de 4x4 que corresponde al bloque actual de 4x4. De manera similar,  $I^{(1)}$  es un bloque que incluye muestras predichas en el segundo fotograma de referencia que se encuentran alrededor de un subbloque correspondiente al bloque actual. Por ejemplo, si el bloque actual es un bloque de 4x4,  $I^{(1)}$  puede ser un bloque de 6x6 en el segundo fotograma de referencia que rodea el bloque de 4x4 que corresponde al bloque actual de 4x4.

35 Si el bloque actual es un bloque de 4x4 y el  $I^{(0)}$  y  $I^{(1)}$  son ambos bloques de 6x6, cada uno de  $G_{x0}$  y  $G_{x1}$  es un bloque de 6x6.

40 Se observa que al calcular los términos de autocorrelación y correlación cruzadas  $s_1, s_2, s_3, s_4$  y  $s_5$ , se puede aplicar desplazamiento para ajustar la precisión y/o profundidad de bits de  $s_1, s_2, s_3, s_4$  y  $s_5$ .

45 Se observa además que  $s_1, s_2, s_3, s_4$  y  $s_5$  se determinan sin operaciones de multiplicación, que reduce significativamente la complejidad de cálculo del proceso para determinar  $v_x$  y  $v_y$ .  $v_x$  y  $v_y$  se derivan con base en términos de autocorrelación y correlación cruzada como se muestra anteriormente con respecto a  $s_1, s_2, s_3, s_4$  y  $s_5$ .  $v_x$  y  $v_y$  se pueden determinar como  $v_x = \frac{s_3}{s_1}$  y  $v_y = \frac{(s_4 - v_x * s_5 / 2)}{s_2}$  si el término  $(I^{(1)} - I^{(0)})$  se cambia a  $(I^{(0)} - I^{(1)})$ .

50 En una forma de implementación posible del dispositivo según cualquier implementación anterior del segundo y tercer aspecto o el segundo y tercer aspecto como tal, el  $G_{x0}$  se determina como una diferencia obtenida de dos muestras predichas que corresponden al primer fotograma de referencia a lo largo de una dirección horizontal, y el  $G_{y0}$  se determina como una diferencia obtenida de dos muestras predichas que corresponden al primer fotograma de referencia a lo largo de una dirección vertical.

En algunos ejemplos, las dos muestras predichas que corresponden al primer fotograma de referencia a lo largo de una dirección horizontal tienen las mismas coordenadas verticales y diferentes coordenadas horizontales. Las dos muestras predichas que corresponden al primer fotograma de referencia a lo largo de una dirección vertical tienen las mismas coordenadas horizontales y diferentes coordenadas verticales. Cada una de las dos muestras predichas se puede procesar antes de calcular la diferencia, como desplazamiento a la derecha, desplazamiento a la izquierda o recorte.

5 En una forma de implementación posible del dispositivo según cualquier implementación anterior del segundo y tercer aspecto o el segundo y tercer aspecto como tal, el  $G_{x1}$  se determina como una diferencia obtenida de dos muestras predichas que corresponden al segundo fotograma de referencia a lo largo de una dirección horizontal, y el  $G_{y1}$  se determina como una diferencia obtenida de dos muestras predichas que corresponden al segundo fotograma de referencia a lo largo de una dirección vertical.

10 En algunos ejemplos, las dos muestras predichas que corresponden al segundo fotograma de referencia a lo largo de una dirección horizontal tienen las mismas coordenadas verticales y diferentes coordenadas horizontales. Las dos muestras predichas que corresponden al segundo fotograma de referencia a lo largo de una dirección vertical tienen las mismas coordenadas horizontales y diferentes coordenadas verticales. Cada una de las dos muestras predichas se puede procesar antes de calcular la diferencia, como desplazamiento a la derecha, desplazamiento a la izquierda o recorte.

15 En una forma de implementación posible del dispositivo según cualquier implementación anterior del segundo y tercer aspecto o el segundo y tercer aspecto como tal, los valores de muestra de predicción correspondientes al primer fotograma de referencia y los valores de muestra de predicción correspondientes al segundo fotograma de referencia se obtienen de los primeros y segundos fotogramas de referencia, respectivamente, usando un par de vectores de movimiento para el bloque actual con respecto a los primeros y segundos fotogramas de referencia.

25 En una forma de implementación posible del dispositivo según cualquier implementación anterior del segundo y tercer aspecto o el segundo y tercer aspecto como tal, los valores de muestra de predicción para el bloque actual son valores de muestra bi-predichos con base en la predicción de flujo óptico bidireccional (BDOF).

Según un cuarto aspecto, la descripción se refiere a un aparato para realizar interpredicción basada en flujo óptico bidireccional (BDOF) para un bloque actual de una señal de video, que comprende:

30 una unidad de determinación configurada para determinar un desplazamiento de movimiento horizontal,  $v_x$  y un desplazamiento de movimiento vertical,  $v_y$ , del bloque actual, y en donde el desplazamiento de movimiento vertical se determina con base en el desplazamiento de movimiento horizontal y una quinta variable  $s_5$ , en donde:

35 la quinta variable  $s_5$  indica una suma de una pluralidad de términos, en donde cada uno de la pluralidad de términos se obtiene de un signo de un elemento de una segunda matriz y un elemento de una primera matriz, y el elemento de la primera matriz corresponde al elemento de la segunda matriz;

40 cada elemento de la primera matriz se obtiene a partir de una suma de un primer gradiente de muestra predicho horizontal que corresponde a un primer fotograma de referencia del bloque actual y un segundo gradiente de muestra predicho horizontal que corresponde a un segundo fotograma de referencia del bloque actual, en donde el primer gradiente de muestra predicho horizontal y el segundo gradiente de muestra predicho horizontal corresponden al elemento de la primera matriz; y

45 cada elemento de la segunda matriz se obtiene a partir de una suma de un primer gradiente de muestra predicho vertical que corresponde al primer fotograma de referencia del bloque actual y un segundo gradiente de muestra predicho vertical que corresponde al segundo fotograma de referencia del bloque actual, en donde el primer gradiente de muestra predicho vertical y el segundo gradiente de muestra predicho vertical corresponden al elemento de la segunda matriz; y

50 una unidad de procesamiento de predicción configurada para predecir valores de muestra de predicción en el bloque actual usando valores de muestra de predicción correspondientes al primer fotograma de referencia, valores de muestra de predicción correspondientes al segundo fotograma de referencia, y el desplazamiento de movimiento horizontal y el desplazamiento de movimiento vertical.

55 Se observa que el bloque actual puede ser un bloque de cualquier tamaño, tal como un bloque de 4x4. El bloque actual puede ser un subbloque de un fotograma de la señal de video. Los píxeles del bloque actual se pueden referir usando posiciones absolutas de los píxeles con respecto a la esquina superior izquierda del fotograma, por ejemplo, (x,y), o posiciones relativas de los píxeles con respecto a la esquina superior izquierda del bloque, por ejemplo, (xBloque + i, yBloque+j). Aquí, (xBloque, yBloque) es la coordenada de la esquina superior izquierda del bloque con respecto a la esquina superior izquierda del fotograma.

65 La primera matriz y la segunda matriz pueden ser cualquier matriz bidimensional que incluya filas y columnas y se puede hacer referencia a un elemento de la matriz usando (i,j) donde x es un índice horizontal/de fila e y es un índice vertical/de columna. El intervalo de i y j puede ser, por ejemplo,  $i = x\text{Bloque} - 1, \dots, x\text{Bloque} + 4$  y  $j = y\text{Bloque} - 1, \dots, y\text{Bloque} + 4$ . La primera y segunda matrices corresponden a, o se determinan para el bloque

actual. En algunos ejemplos, el tamaño de la primera matriz es el mismo que el tamaño de la segunda matriz que puede ser mayor que el tamaño del bloque actual. Por ejemplo, el tamaño de la primera matriz y la segunda matriz puede ser 6x6, mientras que el bloque actual tiene un tamaño de 4x4.

5 Un elemento de la primera matriz (primer elemento) corresponde a un elemento de la segunda matriz (segundo elemento) si la posición del primer elemento en la primera matriz, (x,y) es la misma que la posición del segundo elemento en la segunda matriz, (p,q), es decir, (x,y)=(p,q). El primer gradiente de muestra predicho horizontal corresponde a un primer fotograma de referencia del bloque actual, lo que significa que el primer gradiente de muestra predicho horizontal se genera con base en las muestras en el primer fotograma de referencia del bloque actual. El segundo gradiente de muestra predicho horizontal corresponde al segundo fotograma de referencia del bloque actual, lo que significa que el segundo gradiente de muestra predicho horizontal se genera con base en muestras en el segundo fotograma de referencia del bloque actual. El primer gradiente de muestra predicho horizontal corresponde al elemento de la primera matriz, lo que significa que el primer gradiente de muestra predicho horizontal se genera para la posición (x,y) del elemento en la primera matriz. Asimismo, el segundo gradiente de muestra predicho horizontal corresponde al elemento de la primera matriz, lo que significa que el segundo gradiente de muestra predicho horizontal se genera para la posición (x,y) del elemento en la primera matriz.

20 El primer gradiente de muestra predicho vertical corresponde al primer fotograma de referencia del bloque actual, lo que significa que el primer gradiente de muestra predicho vertical se genera con base en las muestras en el primer fotograma de referencia del bloque actual. El segundo gradiente de muestra predicho vertical corresponde al segundo fotograma de referencia del bloque actual, lo que significa que el segundo gradiente de muestra predicho vertical se genera con base en muestras en el segundo fotograma de referencia del bloque actual. El primer gradiente de muestra predicho vertical corresponde al elemento de la segunda matriz, lo que significa que el primer gradiente de muestra predicho vertical se genera para la posición (p,q) del elemento en la segunda matriz. Del mismo modo, el segundo gradiente de muestra predicho vertical corresponde al elemento de la segunda matriz, lo que significa que el segundo gradiente de muestra predicho vertical se genera para la posición (p,q) del elemento en la segunda matriz.

30 Cada elemento de una matriz se obtiene a partir de una suma de dos términos, lo que significa que el elemento se puede determinar como la suma de dos términos en sí o como un valor después de procesar la suma de dos términos. El procesamiento puede incluir desplazamiento a la izquierda, desplazamiento a la derecha, recorte, o una combinación de los mismos. De manera similar, un término se obtiene de un signo de un elemento de una segunda matriz y un elemento de una primera matriz, lo que significa que el término puede tomar el valor del elemento de la primera matriz por sí mismo o después de que el elemento de la primera matriz se procesa y aplicar el signo del elemento de la segunda matriz. El procesamiento del elemento de la primera matriz puede incluir desplazamiento hacia la izquierda, desplazamiento hacia la derecha, recorte, o una combinación de los mismos. El signo del elemento x se puede determinar como

$$\text{Signo}(x) = \begin{cases} 1 & ; & x > 0 \\ 0 & ; & x = 0 \\ -1 & ; & x < 0 \end{cases}$$

45 La tecnología presentada en el presente documento ajusta los valores de muestra bi-predichos del bloque actual con base en el desplazamiento de movimiento horizontal y el desplazamiento de movimiento vertical. El desplazamiento de movimiento vertical se calcula con base en una quinta variable  $s_5$  que solo implica la suma de términos obtenidos de un signo de un elemento de una segunda matriz y un elemento de una primera matriz. La aplicación de un signo de un elemento a otro elemento no implica operaciones de multiplicación. Del mismo modo, la suma tampoco implica multiplicaciones. Como resultado, la tecnología de interpredicción basada en BDOF presentada en el presente documento elimina las operaciones de multiplicación. En comparación con los enfoques tradicionales, la profundidad de bits de la quinta variable  $s_5$  se reduce porque las operaciones de multiplicación se reemplazan por la determinación de signo. Esto conduce a la reducción de la profundidad de bits del desplazamiento de movimiento horizontal,  $v_x$  y el desplazamiento de movimiento vertical,  $v_y$  y también una reducción significativa en la complejidad de cálculo de la predicción y el tamaño del multiplicador.

55 En una forma de implementación posible del dispositivo según cualquier implementación anterior del cuarto aspecto o el cuarto aspecto como tal, la unidad de determinación se configura para determinar el desplazamiento de movimiento vertical con base en el desplazamiento de movimiento horizontal, una segunda variable  $s_2$ , una cuarta variable  $s_4$  y la quinta variable  $s_5$ , en donde:

60 la segunda variable  $s_2$  indica una suma de valores absolutos de elementos de la segunda matriz; y

la cuarta variable  $s_4$  indica una suma de una pluralidad de términos, en donde cada uno de la pluralidad de

términos se obtiene de un signo de un elemento de la segunda matriz y un elemento de una tercera matriz, y el elemento de la tercera matriz corresponde al elemento de la segunda matriz, y cada elemento de la tercera matriz es una diferencia que se obtiene de una primera muestra predicha del primer fotograma de referencia que corresponde al elemento de la tercera matriz y una segunda muestra predicha del segundo fotograma de referencia que corresponde al elemento de la tercera matriz.

Se observa que un elemento de la tercera matriz (primer elemento) corresponde a un elemento de la segunda matriz (segundo elemento) si la posición del primer elemento en la tercera matriz,  $(k,l)$  es la misma que la posición del segundo elemento en la segunda matriz,  $(p,q)$ , es decir,  $(k,l)=(p,q)$ . La primera muestra predicha corresponde a un primer fotograma de referencia del bloque actual, lo que significa que la primera muestra predicha se encuentra en el primer fotograma de referencia del bloque actual. La segunda muestra predicha corresponde al segundo fotograma de referencia del bloque actual, lo que significa que la segunda muestra predicha se encuentra en el segundo fotograma de referencia del bloque actual. La primera muestra predicha corresponde al elemento de la tercera matriz, lo que significa que la primera muestra predicha se encuentra en la posición  $(k,l)$  del elemento en la tercera matriz. Del mismo modo, la segunda muestra predicha corresponde al elemento de la tercera matriz, lo que significa que la segunda muestra predicha se encuentra en la posición  $(k,l)$  del elemento en la tercera matriz.

Los valores adicionales, la segunda variable  $s_2$  y la cuarta variable  $s_4$  involucradas en el cálculo del desplazamiento de movimiento vertical del bloque actual tampoco implican operaciones de multiplicación. De manera similar a la quinta variable  $s_5$ , el cálculo de la cuarta variable  $s_4$  solo implica la suma de términos obtenidos de un signo de un elemento de una segunda matriz y un elemento de una tercera matriz. La aplicación de un signo de un elemento a otro elemento no implica operaciones de multiplicación. Además, la suma tampoco implica multiplicaciones. En comparación con los enfoques tradicionales, la profundidad de bits de la cuarta variable  $s_4$  se reduce porque las operaciones de multiplicación se reemplazan por la determinación de signo. Del mismo modo, el cálculo de la segunda variable  $s_2$  implica tomar valores absolutos de elementos de la segunda matriz y no hay multiplicaciones involucradas. Como tal, la profundidad de bits de  $s_2$  también se reduce. En consecuencia, al calcular el desplazamiento de movimiento vertical, la profundidad de bits de los resultados también se reduce y la complejidad de cálculo se reduce significativamente.

En una forma de implementación posible del dispositivo según cualquier implementación anterior del cuarto aspecto o el cuarto aspecto como tal, la unidad de determinación se configura para determinar el desplazamiento de movimiento horizontal con base en una primera variable  $s_1$  y una tercera variable  $s_3$ , en donde

la primera variable  $s_1$  indica una suma de valores absolutos de elementos de la primera matriz; y

la tercera variable  $s_3$  indica una suma de una pluralidad de términos, en donde cada uno de la pluralidad de términos se obtiene de un signo de un elemento de la primera matriz y un elemento de la tercera matriz, y el elemento de la tercera matriz corresponde al elemento de la primera matriz.

El elemento de la tercera matriz (tercer elemento) corresponde al elemento de la primera matriz (primer elemento) significa que la posición del primer elemento en la primera matriz,  $(x,y)$  es la misma que la posición del tercer elemento en la tercera matriz,  $(k,l)$ , es decir,  $(x,y)=(k,l)$ .

Al derivar el desplazamiento de movimiento horizontal con base en la primera variable  $s_1$  y una tercera variable  $s_3$ , se puede lograr una reducción de complejidad de cálculo adicional. Tanto el cálculo de la primera variable  $s_1$  como de una tercera variable  $s_3$  no implican multiplicaciones. Más bien, solo están involucradas operaciones de toma de valores absolutos, toma de signos y operaciones de suma.

En una forma de implementación posible del dispositivo según cualquier implementación anterior del cuarto aspecto o el cuarto aspecto como tal, la unidad de determinación se configura para determinar el desplazamiento de movimiento horizontal según

$$v_x = -\frac{s_3}{s_1}$$

en donde:

$v_x$  representa el desplazamiento de movimiento horizontal.

Esto muestra una posible manera de determinar  $v_x$  con base en términos de autocorrelación y correlación cruzada  $s_1$  y  $s_3$ .  $s_1$  y  $s_3$  se determinan sin operaciones de multiplicación, que reduce significativamente la complejidad de cálculo del proceso para determinar  $v_x$ . Como tal, al calcular el desplazamiento de movimiento horizontal de esta manera, el desplazamiento de movimiento horizontal se puede determinar de manera

eficiente.

En una forma de implementación posible del dispositivo según cualquier implementación anterior del cuarto aspecto o el cuarto aspecto como tal, la unidad de determinación se configura para determinar el desplazamiento de movimiento vertical según

$$v_y = -\frac{(s_4 - v_x * s_5/2)}{s_2},$$

en donde:

10

$v_x$  representa el desplazamiento de movimiento horizontal; y

$v_y$  representa el desplazamiento de movimiento vertical.

15 Esto muestra una posible manera de determinar  $v_y$  con base en términos de autocorrelación y correlación cruzada  $s_2$ ,  $s_4$  y  $s_5$ .  $s_2$ ,  $s_4$  y  $s_5$  se determinan sin operaciones de multiplicación, que reduce significativamente la complejidad de cálculo del proceso para determinar. Como tal, al calcular el desplazamiento de movimiento vertical de esta manera, el desplazamiento de movimiento vertical se puede determinar de manera eficiente.

20 En una forma de implementación posible del dispositivo según cualquier implementación anterior del cuarto aspecto o el cuarto aspecto como tal,  $s_1$ ,  $s_2$ ,  $s_3$ ,  $s_4$  y  $s_5$  se determinan como:

$$s_1 = \sum_{i,j} abs(G_{x1} + G_{x0});$$

$$s_2 = \sum_{i,j} abs(G_{y1} + G_{y0}),$$

25

$$s_3 = \sum_{i,j} signo(G_{x1} + G_{x0}) * (I^{(1)} - I^{(0)}).$$

$$s_4 = \sum_{i,j} signo(G_{y1} + G_{y0}) * (I^{(1)} - I^{(0)}),$$

30

$$s_5 = \sum_{i,j} signo(G_{y1} + G_{y0}) * (G_{x1} + G_{x0});$$

y en donde:

35  $I^{(0)}$  se obtiene a partir del valor de muestra predicho que corresponde al primer fotograma de referencia, y  $I^{(1)}$  se obtiene a partir del valor de muestra predicho que corresponde al segundo fotograma de referencia;

40  $G_{x0}$  y  $G_{x1}$  indica un conjunto de los gradientes de muestra predichos horizontales que corresponden al primer fotograma de referencia y al segundo fotograma de referencia, respectivamente;

40  $G_{y0}$  y  $G_{y1}$  indican un conjunto de los gradientes de muestra predichos verticales que corresponden al primer fotograma de referencia y al segundo fotograma de referencia, respectivamente; y

45  $i$  y  $j$  son números enteros, en donde un valor de  $i$  varía de -1 a 4, y un valor de  $j$  varía de -1 a 4.

En algunos ejemplos,  $I^{(0)}$  es un bloque que incluye muestras predichas en el primer fotograma de referencia

que se encuentran alrededor de un subbloque correspondiente al bloque actual. Por ejemplo, si el bloque actual es un bloque de 4x4,  $I^{(0)}$  puede ser un bloque de 6x6 en el primer fotograma de referencia que rodea el bloque de 4x4 que corresponde al bloque actual de 4x4. De manera similar,  $I^{(1)}$  es un bloque que incluye muestras predichas en el segundo fotograma de referencia que se encuentran alrededor de un subbloque correspondiente al bloque actual. Por ejemplo, si el bloque actual es un bloque de 4x4,  $I^{(1)}$  puede ser un bloque de 6x6 en el segundo fotograma de referencia que rodea el bloque de 4x4 que corresponde al bloque actual de 4x4.

Si el bloque actual es un bloque de 4x4 y el  $I^{(0)}$  y  $I^{(1)}$  son ambos bloques de 6x6, cada uno de  $G_{x0}$  y  $G_{x1}$  es un bloque de 6x6.

Se observa que al calcular los términos de autocorrelación y correlación cruzadas  $s_1, s_2, s_3, s_4$  y  $s_5$ , se puede aplicar desplazamiento para ajustar la precisión y/o profundidad de bits de  $s_1, s_2, s_3, s_4$  y  $s_5$ .

Se observa además que  $s_1, s_2, s_3, s_4$  y  $s_5$  se determinan sin operaciones de multiplicación, que reduce significativamente la complejidad de cálculo del proceso para determinar  $v_x$  y  $v_y$ .  $v_x$  y  $v_y$  se derivan con base en términos de autocorrelación y correlación cruzada como se muestra anteriormente con respecto a  $s_1, s_2, s_3, s_4$  y  $s_5$ .  $v_x$  y  $v_y$  se pueden determinar como  $v_x = \frac{s_3}{s_1}$  y  $v_y = \frac{(s_4 - v_x * s_5 / 2)}{s_2}$  si el término  $(I^{(1)} - I^{(0)})$  se cambia a  $(I^{(0)} - I^{(1)})$ .

En una forma de implementación posible del dispositivo según cualquier implementación anterior del cuarto aspecto o el cuarto aspecto como tal, el  $G_{x0}$  se determina como una diferencia obtenida de dos muestras predichas que corresponden al primer fotograma de referencia a lo largo de una dirección horizontal, y el  $G_{y0}$  se determina como una diferencia obtenida de dos muestras predichas que corresponden al primer fotograma de referencia a lo largo de una dirección vertical.

En algunos ejemplos, las dos muestras predichas que corresponden al primer fotograma de referencia a lo largo de una dirección horizontal tienen las mismas coordenadas verticales y diferentes coordenadas horizontales. Las dos muestras predichas que corresponden al primer fotograma de referencia a lo largo de una dirección vertical tienen las mismas coordenadas horizontales y diferentes coordenadas verticales. Cada una de las dos muestras predichas se puede procesar antes de calcular la diferencia, como desplazamiento a la derecha, desplazamiento a la izquierda o recorte.

En una forma de implementación posible del dispositivo según cualquier implementación anterior del cuarto aspecto o el cuarto aspecto como tal, el  $G_{x1}$  se determina como una diferencia obtenida de dos muestras predichas que corresponden al segundo fotograma de referencia a lo largo de una dirección horizontal, y el  $G_{y1}$  se determina como una diferencia obtenida de dos muestras predichas que corresponden al segundo fotograma de referencia a lo largo de una dirección vertical.

En algunos ejemplos, las dos muestras predichas que corresponden al segundo fotograma de referencia a lo largo de una dirección horizontal tienen las mismas coordenadas verticales y diferentes coordenadas horizontales. Las dos muestras predichas que corresponden al segundo fotograma de referencia a lo largo de una dirección vertical tienen las mismas coordenadas horizontales y diferentes coordenadas verticales. Cada una de las dos muestras predichas se puede procesar antes de calcular la diferencia, como desplazamiento a la derecha, desplazamiento a la izquierda o recorte.

En una forma de implementación posible del dispositivo según cualquier implementación anterior del cuarto aspecto o el cuarto aspecto como tal, los valores de muestra de predicción correspondientes al primer fotograma de referencia y los valores de muestra de predicción correspondientes al segundo fotograma de referencia se obtienen de los primeros y segundos fotogramas de referencia, respectivamente, usando un par de vectores de movimiento para el bloque actual con respecto a los primeros y segundos fotogramas de referencia.

En una forma de implementación posible del dispositivo según cualquier implementación anterior del cuarto aspecto o el cuarto aspecto como tal, los valores de muestra de predicción para el bloque actual son valores de muestra bi-predichos con base en la predicción de flujo óptico bidireccional (BDOF).

El método según algún aspecto de la invención se puede realizar por el aparato según algún aspecto de la invención. Otras características y formas de implementación del método según algún aspecto de la invención resultan directamente de la funcionalidad del aparato según algún aspecto de la invención y sus diferentes formas de implementación.

Se observa que un dispositivo de codificación puede ser un dispositivo de codificación o dispositivo de decodificación.

Según otro aspecto, la invención se refiere a un aparato para decodificar una secuencia de video que incluye un procesador y una memoria. La memoria almacena instrucciones que hacen que el procesador realice el método mostrado anteriormente.

5

Según otro aspecto, la invención se refiere a un aparato para codificación de una secuencia de video que incluye un procesador y una memoria. La memoria almacena instrucciones que hacen que el procesador realice el método mostrado anteriormente.

10

Según otro aspecto, se propone un medio de almacenamiento legible por ordenador que tiene almacenadas instrucciones que, cuando se ejecutan, provocan que uno o más procesadores se configuren para codificar datos de video. Las instrucciones hacen que el uno o más procesadores realicen un método mostrado anteriormente.

15

Según otro aspecto, se proporciona un producto de programa de ordenador con un código de programa para realizar el método mostrado anteriormente cuando el programa de ordenador se ejecuta en un ordenador.

Los detalles de una o más modalidades se establecen en las figuras anexas y la descripción más adelante. Otras características, objetos y ventajas serán evidentes a partir de la descripción, dibujos y reivindicaciones.

20

Por claridad, cualquiera de las modalidades anteriores se puede combinar con cualquiera de una o más de las otras realizaciones anteriores para crear una nueva realización dentro del alcance del presente documento descripción.

25

Estas y otras características se entenderán más claramente a partir de la siguiente descripción detallada tomada en conjunto con las figuras y reivindicaciones anexas.

Breve descripción de los dibujos

30

Otras realizaciones de la invención se describirán con respecto a las siguientes figuras, en donde:

La Fig. 1A muestra un diagrama de bloques que muestra un ejemplo de un sistema de codificación de video configurado para implementar las realizaciones presentadas en el presente documento;

35

La Fig. 1B muestra un diagrama de bloques que muestra otro ejemplo de un sistema de codificación de video configurado para implementar realizaciones;

la Fig. 2 muestra un diagrama de bloques que muestra un ejemplo de un codificador de video configurado para implementar realizaciones presentadas en el presente documento.

40

la Fig. 3 es un diagrama de bloques que muestra una estructura de ejemplo de un decodificador de video configurado para implementar realizaciones presentadas en el presente documento.

45

la Fig. 4 es un diagrama de bloques que ilustra un ejemplo de un aparato de codificación o un aparato de decodificación;

la Fig. 5 es un diagrama de bloques que ilustra otro ejemplo de un aparato de codificación o un aparato de decodificación.

50

la Fig. 6 es un diagrama que ilustra la relación entre una ventana de 6x6 para calcular la autocorrelación y correlación cruzada de gradientes y un subbloque de 4x4.

la Fig. 7 es un diagrama que muestra un ejemplo de flujo óptico bi-predictivo.

55

la Fig. 8 es un diagrama de flujo que ilustra un ejemplo de un proceso para interpredicción basada en flujo óptico según una modalidad.

la Fig. 9 es un diagrama de flujo que ilustra un ejemplo de un método para interpredicción basada en flujo óptico bidireccional (BDOF) para un bloque actual de una señal de video.

60

la Fig. 10 es un diagrama de bloques que muestra una estructura de ejemplo de un aparato para interpredicción basada en flujo óptico bidireccional (BDOF) para un bloque actual de una señal de video.

65

la Fig. 11 es un diagrama de bloques que muestra una estructura de ejemplo de un sistema de suministro de contenido que proporciona un servicio de suministro de contenido.

la Fig. 12 es un diagrama de bloques que muestra una estructura de un ejemplo de un dispositivo terminal.

En las diversas figuras, se usarán signos de referencia idénticos para características idénticas o funcionalmente equivalentes.

5

Descripción detallada de las realizaciones

10 En la siguiente descripción, se hace referencia a las figuras anexas, que forman parte de la descripción, y en las que se muestran, a manera de ilustración, aspectos específicos en los que se puede colocar el presente documento invención. Se entiende que se pueden utilizar otros aspectos y se pueden realizar cambios estructurales o lógicos sin desviarse del alcance del presente documento invención. Por lo tanto, la siguiente descripción detallada no se debe tomar en un sentido limitante, ya que el alcance del presente documento invención se define por las reivindicaciones anexas.

15 Por ejemplo, se entiende que una descripción en conexión con un método descrito también se puede mantener verdadera para un dispositivo o sistema correspondiente configurado para realizar el método y viceversa. Por ejemplo, si se describen uno o una pluralidad de pasos de método específicos, un dispositivo correspondiente puede incluir una o una pluralidad de unidades, por ejemplo, unidades funcionales, para realizar el paso de método descrito (por ejemplo, una unidad para realizar el paso o pluralidad de pasos, o una pluralidad de unidades para realizar cada uno o más de la pluralidad de pasos), incluso si esta unidad o más no se describen o ilustran explícitamente en las figuras. Por otro lado, por ejemplo, si un aparato específico se describe con base en una o una pluralidad de unidades, por ejemplo, unidades funcionales, un método correspondiente puede incluir un paso para realizar la funcionalidad de la una o pluralidad de unidades (por ejemplo, un paso para realizar la funcionalidad de la una o pluralidad de unidades, o una pluralidad de pasos para realizar cada una la funcionalidad de una o más de la pluralidad de unidades), incluso si este paso o pluralidad de pasos no se describen o ilustran explícitamente en las figuras. Además, se entiende que las características de las diversas realizaciones y/o aspectos de ejemplo descritos a continuación se pueden combinar entre sí, a menos que se indique específicamente lo contrario.

30 El presente documento descripción proporciona una tecnología que calcula un segundo componente de flujo óptico con base en un primer componente de flujo óptico calculado y sin requerir alguna multiplicación costosa. La tecnología se puede emplear tanto en los extremos de codificación como de decodificación cuando se activa BPOF. Proporciona dispositivos y métodos mejorados para interpredicción basada en flujo óptico, de modo que se mejora la eficiencia de compresión en comparación con la BPOF convencional sin aumentar la complejidad de cálculo de la BPOF.

35

Para describir la invención en detalle, se utilizarán los siguientes términos, abreviaturas y notaciones:

40 POC Recuento de Orden de Imagen en orden de visualización

40

MV Vector de Movimiento

MCP Predicción Compensada de Movimiento

45 HEVC Codificación de Video de Alta Eficiencia

45

BPOF Corrección de lado de decodificador basada en flujo Óptico Bi-predictivo para MCP

50 BDOF Flujo Óptico Bidireccional

50

Como se usa en el presente documento, una señal de video o secuencia de video es un conjunto de fotogramas posteriores que presentan una imagen en movimiento. En otras palabras, una señal de video o secuencia de video consiste en una pluralidad de fotogramas (también denominados fotografías o imágenes).

55 Como se usa en el presente documento, unidad de árbol de codificación (CTU) denota la raíz de una estructura de codificación de la secuencia de video de un tamaño predefinido, que contiene una parte de un fotograma (por ejemplo, 64x64 píxeles). Una CTU se puede particionar en varias CU.

60 Como se usa en el presente documento, la unidad de codificación (CU) denota una estructura de codificación básica de la secuencia de video de un tamaño predefinido, que contiene una parte de un fotograma, que pertenece a una CTU. Una CU se puede particionar en otras CU.

60

Como se usa en el presente documento, unidad de predicción (PU) denota una estructura de codificación, que es el resultado de partición de una CU.

65

Como se usa en el presente documento, el término co-ubicado en el presente documento denota un bloque o

región en un segundo fotograma, es decir, un fotograma de referencia, correspondiente a un bloque o región real en un primer fotograma, es decir, un fotograma actual.

5 La codificación de video normalmente se refiere al procesamiento de una secuencia de imágenes, que forman el video o secuencia de video. En lugar del término "fotografía", el término "fotograma" o "imagen" se pueden usar como sinónimos en el campo de codificación de video. La codificación de video (o codificación en general) consta de dos partes: codificación de video y decodificación de video. La codificación de video se realiza en el lado fuente, que comprende típicamente procesar (por ejemplo mediante compresión) las imágenes de video originales para reducir la cantidad de datos necesarios para representar las imágenes de video (para almacenamiento y/o transmisión más eficiente). La decodificación de video se realiza en el lado de destino y generalmente comprende el procesamiento inverso en comparación con el codificador para reconstruir las imágenes de video. Se entenderá que las realizaciones que se refieren a "codificación" de imágenes de video (o imágenes en general) se refieren a "codificación" o "decodificación" de imágenes de video o secuencias de video respectivas. La combinación de la parte de codificación y la parte de decodificación también se conoce como códec (codificación y decodificación).

20 En caso de codificación de video sin pérdidas, las imágenes de video originales se pueden reconstruir, es decir, las imágenes de video reconstruidas tienen la misma calidad que las imágenes de video originales (asumiendo que no hay pérdida de transmisión u otra pérdida de datos durante el almacenamiento o transmisión). En caso de codificación de video con pérdidas, se realiza compresión adicional, por ejemplo, mediante cuantificación, para reducir la cantidad de datos que representan las imágenes de video, que no se pueden reconstruir completamente en el decodificador, es decir, la calidad de las imágenes de video reconstruidas es menor o peor en comparación con la calidad de las imágenes de video originales.

25 Varias normas de codificación de video pertenecen al grupo de "códecs de video híbridos con pérdidas" (es decir, combinan predicción espacial y temporal en la codificación de transformada bidimensional para aplicar cuantificación en el dominio de transformada). Cada imagen de una secuencia de video se particiona típicamente en un conjunto de bloques no superpuestos y la codificación se realiza típicamente a nivel de bloque. En otras palabras, en el codificador el video se procesa normalmente, es decir, se codifica, a nivel de bloque (bloque de video), por ejemplo, al usar predicción espacial (intraimagen) y predicción temporal (interimagen) para generar un bloque de predicción, restar el bloque de predicción del bloque actual (bloque actualmente procesado/que se va a procesar) para obtener un bloque residual, transformar el bloque residual y cuantificar el bloque residual en el dominio de transformada para reducir la cantidad de datos que se van a transmitir (compresión). En el decodificador, el procesamiento inverso en comparación con el codificador se aplica al bloque codificado o comprimido para reconstruir el bloque actual para su representación. Además, el codificador duplica el bucle de procesamiento de decodificador de modo que ambos generarán predicciones idénticas (por ejemplo, intra e interpredicciones) y/o reconstrucciones para procesamiento, es decir, codificación, los bloques posteriores.

40 Las realizaciones de un sistema 10 de codificación de video, un codificador 20 de video y un decodificador 30 de video se describen con respecto a las Fig. 1A-3.

45 La Fig. 1A es un diagrama de bloques esquemático que ilustra un sistema de codificación de ejemplo 10, por ejemplo, un sistema 10 de codificación de video (o sistema de codificación 10 en resumen) que puede implementar técnicas presentadas en el presente documento. El codificador 20 de video (o codificador 20 para abreviar) y el decodificador 30 de video (o decodificador 30 para abreviar) del sistema 10 de codificación de video representan ejemplos de dispositivos que se pueden configurar para realizar técnicas según varios ejemplos descritos en el presente documento.

50 Como se muestra en la Fig. 1A, el sistema 10 de codificación comprende un dispositivo 12 fuente configurado para proporcionar datos 21 de imagen codificados, por ejemplo, a un dispositivo destino 14 para decodificar los datos de imagen codificados 21. El dispositivo 12 fuente incluye un codificador 20 y puede incluir adicionalmente una fuente 16 de imagen, un preprocesador (o unidad de preprocesamiento) 18, tal como un preprocesador 18 de imágenes y una interfaz de comunicación o unidad 22 de comunicación.

55 La fuente 16 de imágenes puede incluir cualquier tipo de dispositivo de captura de imágenes, tal como una cámara para capturar una imagen del mundo real, cualquier tipo de dispositivo generador de imágenes, tal como un procesador de gráficos de ordenador para generar una imagen animada por ordenador. La fuente 16 de imágenes también puede incluir cualquier tipo de otro dispositivo para obtener y/o proporcionar una imagen del mundo real, una imagen generada por ordenador (por ejemplo, un contenido de pantalla, una imagen de realidad virtual (VR)) y/o cualquier combinación de los mismos (por ejemplo, una imagen de realidad aumentada (AR)). La fuente de imagen puede ser cualquier tipo de memoria o almacenamiento que almacene cualquiera de las imágenes mencionadas anteriormente.

65 A diferencia del preprocesador 18 y el procesamiento realizado por la unidad 18 de preprocesamiento, la imagen o datos 17 de imagen también se pueden referir como datos 17 de imagen sin procesar o imagen sin

procesar. El preprocesador 18 se configura para recibir los datos 17 de imagen (sin procesar) y para realizar preprocesamiento en los datos 17 de imagen para obtener una imagen 19 preprocesada o datos 19 de imagen preprocesados. El preprocesamiento realizado por el preprocesador 18 puede incluir recorte, conversión de formato de color (por ejemplo de RGB a YCbCr), corrección de color o eliminación de ruido.

- 5 El codificador 20 de video se configura para recibir los datos 19 de imagen preprocesados y proporcionar datos 21 de imagen codificados (se describirán detalles adicionales más adelante, por ejemplo, con respecto a la Fig. 2). La interfaz 22 de comunicación del dispositivo 12 fuente se puede configurar para recibir los datos 21 de imagen codificados y transmitir los datos 21 de imagen codificados (o cualquier versión procesada adicional de los mimos) a través del canal 13 de comunicación a otro dispositivo, por ejemplo, el dispositivo 14 destino o cualquier otro dispositivo, para almacenamiento o reconstrucción directa. El dispositivo 14 destino comprende un decodificador 30 (por ejemplo, un decodificador 30 de video) y puede incluir adicionalmente, una interfaz de comunicación o unidad de comunicación 28, un postprocesador 32 (o unidad de postprocesador 32 ) y un dispositivo de visualización 34. La interfaz 28 de comunicación del dispositivo 14 destino se configura para recibir los datos 21 de imagen codificados (o cualquier versión procesada adicional de los mismos), por ejemplo, directamente desde el dispositivo 12 fuente o desde cualquier otra fuente, por ejemplo, un dispositivo de almacenamiento, por ejemplo, un dispositivo de almacenamiento de datos de imagen codificados, y proporcionar los datos 21 de imagen codificados al decodificador 30.
- 10
- 15
- 20 La interfaz 22 de comunicación y la interfaz 28 de comunicación se pueden configurar para transmitir o recibir los datos 21 de imagen codificados o datos 21 codificados mediante un enlace de comunicación directa entre el dispositivo 12 fuente y el dispositivo 14 destino, tal como una conexión inalámbrica o por cable directa, o mediante cualquier tipo de red, por ejemplo, una red inalámbrica o por cable o cualquier combinación de las mismas, o cualquier tipo de red pública y privada, o cualquier tipo de combinación de las mismas.
- 25
- 30 La interfaz 22 de comunicación se puede configurar, por ejemplo, para empaquetar los datos 21 de imagen codificados en un formato adecuado, por ejemplo, paquetes, y/o procesar los datos de imagen codificados usando cualquier tipo de codificación o procesamiento de transmisión para transmisión a través de un enlace de comunicación o red de comunicación. La interfaz 28 de comunicación, que forma la contraparte de la interfaz 22 de comunicación, se puede configurar, por ejemplo, para recibir los datos transmitidos y procesar los datos de transmisión usando cualquier tipo de decodificación o procesamiento de transmisión correspondiente y/o desempaqueto para obtener los datos 21 de imagen codificados.
- 35 Tanto la interfaz 22 de comunicación como la interfaz 28 de comunicación se pueden configurar como interfaces de comunicación unidireccionales como se indica por la flecha para el canal de comunicación 13 en la Fig. 1A que apuntan desde el dispositivo 12 fuente al dispositivo 14 destino, o interfaces de comunicación bidireccionales, y se pueden configurar, por ejemplo, para enviar y recibir mensajes, por ejemplo, para configurar una conexión, para reconocer e intercambiar cualquier otra información relacionada con el enlace de comunicación y/o transmisión de datos, por ejemplo, transmisión de datos de imagen codificados.
- 40
- 45 El decodificador 30 del dispositivo 14 destino se configura para recibir los datos 21 de imagen codificados y proporcionar datos de imagen 31 decodificados o una imagen 31 decodificada (se describirán más detalles a continuación, por ejemplo, con respecto a la Fig. 3 o Fig. 5). El postprocesador 32 del dispositivo 14 destino se configura para postprocesar los datos de imagen decodificados 31 (también llamados datos de imagen reconstruidos), por ejemplo, la imagen decodificada 31, para obtener datos 33 de imagen postprocesados, por ejemplo, una imagen 33 postprocesada. El postprocesamiento realizado por la unidad 32 de postprocesamiento puede comprender, por ejemplo, conversión de formato de color (por ejemplo, de YCbCr a RGB), corrección de color, recorte o remuestreo, o cualquier otro procesamiento, por ejemplo, para preparar los datos 31 de imagen decodificados para visualización, por ejemplo, por el dispositivo 34 de visualización.
- 50
- 55 El dispositivo 34 de visualización del dispositivo 14 destino se configura para recibir los datos 33 de imagen postprocesados para visualizar la imagen, por ejemplo, a un usuario o espectador. El dispositivo 34 de visualización puede ser o comprender cualquier tipo de elemento de visualización para representar la imagen reconstruida, por ejemplo, un elemento de visualización o monitor integrado o externo. Los elementos de visualización pueden, por ejemplo, comprender elementos de visualización de cristal líquido (LCD), elementos de visualización de diodos emisores de luz orgánicos (OLED), elementos de visualización de plasma, proyectores, elementos de visualización micro-led, cristal líquido sobre silicio (LCoS), procesador de luz digital (DLP) o cualquier tipo de elemento de visualización.
- 60 Aunque la Fig. 1A representa el dispositivo 12 fuente y el dispositivo 14 destino como dispositivos separados, realizaciones de dispositivos también pueden comprender ambas funcionalidades, el dispositivo 12 fuente o funcionalidad correspondiente y el dispositivo 14 destino o funcionalidad correspondiente. En estas realizaciones, el dispositivo 12 fuente o funcionalidad correspondiente y el dispositivo 14 destino o funcionalidad correspondiente se pueden implementar usando el mismo hardware y/o software o mediante hardware y/o software separado o cualquier combinación de los mismos.
- 65

Como será evidente para el experto con base en la descripción, la existencia y (exacta) división de funcionalidades de las diferentes unidades o funcionalidades dentro del dispositivo 12 fuente y/o dispositivo 14 destino como se muestra en la Fig. 1A puede variar dependiendo del dispositivo y aplicación reales.

5 El codificador 20 (por ejemplo, un codificador 20 de video) o decodificador 30 (por ejemplo, un decodificador 30 de video) o ambos codificador 20 y decodificador 30 se puede implementar mediante circuitería de procesamiento como se muestra en la Fig. 1B, tal como uno o más microprocesadores, procesadores de señales digitales (DSP), circuitos integrados de aplicación específica (ASIC), matrices de puertas programables en campo (FPGA), lógica discreta, hardware, codificación de video dedicada o cualquier combinación de los mismos. El codificador 20 se puede implementar a través de la circuitería 46 de procesamiento para incorporar los diversos módulos como se analiza con respecto al codificador 20 de la Fig. 2 y/o cualquier otro sistema o subsistema codificador descrito en el presente documento. El decodificador 30 se puede implementar a través de la circuitería 46 de procesamiento para incorporar los diversos módulos como se analiza con respecto al decodificador 30 de la Fig. 3 y/o cualquier otro sistema o subsistema decodificador descrito en el presente documento. La circuitería de procesamiento se puede configurar para realizar las diversas operaciones descritas más adelante. Como se mostrará en la Fig. 5, si las técnicas se implementan parcialmente en software, un dispositivo puede almacenar instrucciones para el software en un almacenamiento medio de almacenamiento legible por ordenador adecuado y no transitorio y puede ejecutar las instrucciones en hardware usando uno o más procesadores para realizar las técnicas de esta descripción. Cualquiera del codificador 20 de video y el decodificador 30 de video se pueden integrar como parte de un codificador/decodificador combinado (códec) en un solo dispositivo, por ejemplo, como se muestra en la Fig. 1B.

25 El dispositivo 12 fuente y el dispositivo 14 destino pueden comprender cualquiera de una amplia gama de dispositivos, que incluyen cualquier tipo de dispositivo portátil o estacionario, por ejemplo, miniordenadores u ordenadores portátiles, teléfonos móviles, teléfonos inteligentes, tabletas u ordenadores de tableta, cámaras, ordenadores de escritorio, decodificadores, televisiones, dispositivos de visualización, reproductores de medios digitales, consolas de videojuegos, dispositivos de transmisión en tiempo de video (tal como servidores de servicios de contenido o servidores de entrega de contenido), dispositivo receptor de difusión, dispositivo transmisor de difusión o similares y pueden usar ningún o cualquier tipo de sistema operativo. En algunos casos, el dispositivo 12 fuente y el dispositivo 14 destino pueden estar equipados para comunicación inalámbrica. Por lo tanto, el dispositivo 12 fuente y el dispositivo 14 destino pueden ser dispositivos de comunicación inalámbrica.

35 El sistema 10 de codificación de video ilustrado en la Fig. 1A es meramente un ejemplo y las técnicas presentadas en el presente documento se pueden aplicar a configuraciones de codificación de video (por ejemplo, codificación de video o decodificación de video) que no incluyen necesariamente alguna comunicación de datos entre los dispositivos de codificación y decodificación. En otros ejemplos, los datos se recuperan de una memoria local, se transmiten en tiempo real a través de una red o similares. Un dispositivo de codificación de video puede codificar y almacenar datos en memoria, y/o un dispositivo de decodificación de video puede recuperar y decodificar datos de memoria. En algunos ejemplos, la codificación y decodificación se realizan mediante dispositivos que no se comunican entre sí, sino que simplemente codifican datos para memoria y/o recuperan y decodifican datos de memoria.

45 Para mayor comodidad de descripción, se describen realizaciones en el presente documento, por ejemplo, con referencia a Codificación de Video de Alta Eficiencia (HEVC) o al software de referencia de Codificación de Video Versátil (VVC), la norma de codificación de Video de próxima generación desarrollada por el Equipo de Colaboración Conjunta en Codificación de Video (JCT-VC) del Grupo de Expertos en Codificación de Video ITU-T (VCEG) y el Grupo de Expertos en Imágenes en Movimiento ISO/IEC (MPEG). Un experto en la técnica entenderá que las realizaciones de la invención no se limitan a HEVC o VVC.

#### Codificador y método de codificación

55 La Fig. 2 muestra un diagrama de bloques esquemático de un codificador 20 de video de ejemplo que se configura para implementar las técnicas presentadas en el presente documento. En el ejemplo de la Fig. 2, el codificador 20 de video comprende una entrada 201 (o interfaz 201 de entrada), una unidad 204 de cálculo residual, una unidad 206 de procesamiento de transformada, una unidad 208 de cuantificación, una unidad 210 de cuantificación inversa y una unidad 212 de procesamiento de transformada inversa, una unidad 214 de reconstrucción, una unidad 220 de filtro de bucle, una memoria intermedia 230 de imágenes decodificadas (DPB), una unidad 260 de selección de modo, una unidad 270 de codificación entrópica y una salida 272 (o interfaz 272 de salida). La unidad 260 de selección de modo puede incluir una unidad 244 de interpredicción, una unidad 254 de intrapredicción y una unidad 262 de partición. La unidad 244 de intrapredicción puede incluir una unidad de estimación de movimiento y una unidad de compensación de movimiento (no se muestra). Un codificador 20 de video como se muestra en la Fig. 2 también se puede referir como codificador de video híbrido o un codificador de video según un códec de video híbrido.

La unidad 204 de cálculo residual, la unidad 206 de procesamiento de transformada, la unidad 208 de cuantificación, la unidad 260 de selección de modo se pueden referir como que forman una ruta de señal directa del codificador 20, mientras que la unidad 210 de cuantificación inversa, la unidad 212 de procesamiento de transformada inversa, la unidad 214 de reconstrucción, la memoria 216 intermedia, el filtro 220 de bucle, la memoria intermedia 230 de imágenes decodificadas (DPB), la unidad 244 de predicción inversa y la unidad 254 de intrapredicción se pueden referir como que forman una ruta de señal inversa del codificador 20 de video, en donde la ruta de señal inversa del codificador 20 de video corresponde a la ruta de señal del decodificador (ver el decodificador de video 30 en la Fig. 3). También se hace referencia a la unidad 210 de cuantificación inversa, la unidad 212 de procesamiento de transformada inversa, la unidad de reconstrucción 214, el filtro de bucle 220, la memoria intermedia 230 de imágenes decodificadas (DPB), la unidad 244 de interpretación y la unidad 254 de intrapredicción que forman el “decodificador integrado” del codificador 20 de video.

#### Imágenes y partición de imágenes (imágenes y bloques)

El codificador 20 se puede configurar para recibir, por ejemplo, mediante la entrada 201, una imagen 17 (o datos 17 de imagen), por ejemplo, imagen de una secuencia de imágenes que forman un video o secuencia de video. La imagen o datos de imagen recibidos también pueden ser una imagen 19 preprocesada (o datos 19 de imagen preprocesados). En aras de simplicidad, la siguiente descripción se refiere a la imagen 17. La imagen 17 también se puede referir como imagen actual o imagen que se va a codificar (en particular, en la codificación de video para distinguir la imagen actual de otras imágenes, por ejemplo, imágenes previamente codificadas y/o decodificadas de la misma secuencia de video, es decir, la secuencia de video que también comprende la imagen actual).

Una imagen (digital) se puede considerar como una matriz bidimensional o matriz de muestras con valores de intensidad. Una muestra en la matriz también se puede referir como a píxel (forma corta de elemento de imagen) o un pel. El número de muestras en dirección horizontal y vertical (o eje) de la matriz o imagen define el tamaño y/o resolución de la imagen. Para la representación del color, típicamente se emplean tres componentes de color, es decir, la imagen puede estar representada o incluir tres matrices de muestras. En formato RGB o espacio de color, una imagen comprende una matriz de muestras roja, verde y azul correspondiente. Sin embargo, en la codificación de video, cada píxel se representa típicamente en un formato de luminancia y crominancia o espacio de color, por ejemplo, YCbCr, que comprende un componente de luminancia indicado por Y (a veces también se usa L en su lugar) y dos componentes de crominancia indicados por Cb y Cr. El componente de luminancia (o luma corta) Y representa el brillo o la intensidad del nivel de gris (por ejemplo, como en una imagen a escala de grises), mientras que los dos componentes de crominancia (o croma corta) Cb y Cr representan los componentes de cromaticidad o información de color. Por consiguiente, una imagen en formato YCbCr comprende una matriz de muestras de luminancia de valores (Y) de muestra de luminancia y dos matrices de muestras de crominancia de valores (Cb y Cr) de crominancia. Las imágenes en formato RGB se pueden convertir o transformar en formato YCbCr y viceversa, el proceso también se conoce como transformación o conversión de color. Si una imagen es monocromática, la imagen puede comprender solo una matriz de muestras de luminancia. Por consiguiente, una imagen puede ser, por ejemplo, una matriz de muestras de luma en formato monocromático o una matriz de muestras de luma y dos matrices correspondientes de muestras de croma en formato de color 4:2:0, 4:2:2 y 4:4:4.

Las realizaciones del codificador 20 de video pueden comprender una unidad de partición de imagen (no representada en la Fig. 2) configurada para dividir la imagen 17 en una pluralidad de bloques 203 de imagen (típicamente no superpuestos). Estos bloques también se pueden referir como bloques raíz, macrobloques (H.264/AVC) o bloques de árbol de codificación (CTB) o unidades de árbol de codificación (CTU) (H.265/HEVC y VVC). La unidad de partición de imagen se puede configurar para usar el mismo tamaño de bloque para todas las imágenes de una secuencia de video y la cuadrícula correspondiente que define el tamaño de bloque, o para cambiar el tamaño de bloque entre imágenes o subconjuntos o grupos de imágenes, y particionar cada imagen en los bloques correspondientes.

En realizaciones adicionales, el codificador de video se puede configurar para recibir directamente un bloque 203 de la imagen 17, por ejemplo, uno, varios o todos los bloques que forman la imagen 17. El bloque 203 de imagen también se puede referir como bloque de imagen actual o bloque de imagen que se va a codificar.

Al igual que la imagen 17, el bloque 203 de imagen también se puede considerar como una matriz bidimensional o matriz de muestras con valores de intensidad (valores de muestra), aunque de dimensión más pequeña que la imagen 17. En otras palabras, el bloque 203 puede comprender, por ejemplo, una matriz de muestras (por ejemplo, una matriz de luma en el caso de una imagen 17 monocromática, o una matriz de luma o croma en el caso de una imagen de color) o tres matrices de muestras (por ejemplo, una matriz de luma y dos matrices de croma en el caso de una imagen de color 17) o cualquier otro número y/o tipo de matrices dependiendo del formato de color aplicado. El número de muestras en dirección horizontal y vertical (o eje) del bloque 203 define el tamaño del bloque 203. Por consiguiente, un bloque puede, por ejemplo, una matriz de muestras MxN (columna M por fila N) o una matriz MxN de coeficientes de transformada.

Se pueden configurar realizaciones del codificador 20 de video como se muestra en la Fig. 2 para codificar el bloque de imagen 17 por bloque, por ejemplo, la codificación y predicción se realiza por bloque 203. Las realizaciones del codificador 20 de video como se muestra en la Fig. 2 pueden configurarse además para dividir y/o codificar la imagen mediante el uso de segmentos (también denominados segmentos de video), en donde una imagen se puede particionar o codificar mediante el uso de uno o más segmentos (típicamente no superpuestos), y cada segmento puede comprender uno o más bloques (por ejemplo, CTU).

Las realizaciones del codificador 20 de video como se muestra en la Fig. 2 se pueden configurar adicionalmente para particionar y/o codificar la imagen mediante el uso de grupos de mosaicos (también denominados grupos de mosaicos de video) y/o mosaicos (también denominados mosaicos de video), en donde una imagen se puede particionar o codificar mediante el uso de uno o más grupos de mosaicos (típicamente no superpuestos), y cada grupo de mosaicos puede comprender, por ejemplo uno o más bloques (por ejemplo, CTU) o una o más mosaicos, en donde cada mosaico, por ejemplo puede tener forma rectangular y puede comprender uno o más bloques (por ejemplo, CTU), por ejemplo, bloques completos o fraccionarios.

#### 15 Cálculo residual

La unidad 204 de cálculo residual se puede configurar para calcular un bloque 205 residual (también denominado residual 205) con base en el bloque 203 de imagen y un bloque 265 de predicción (los detalles adicionales sobre el bloque 265 de predicción se proporcionan más adelante), por ejemplo al restar valores de muestra del bloque 265 de predicción de valores de muestra del bloque 203 de imagen, muestra por muestra (pixel por pixel) para obtener el bloque 205 residual en el dominio de muestra.

#### 25 Transformada

La unidad 206 de procesamiento de transformada se puede configurar para aplicar una transformada, por ejemplo, una transformada discreta de coseno (DCT) o una transformada discreta de seno (DST), a los valores de muestra del bloque residual 205 para obtener coeficientes 207 de transformada en un dominio de transformada. Los coeficientes 207 de transformada también se pueden referir como coeficientes residuales de transformada y representan el bloque 205 residual en el dominio de transformada.

La unidad 206 de procesamiento de transformada se puede configurar para aplicar aproximaciones enteras de DCT/DST, tal como las transformadas especificadas para H.265/HEVC. En comparación con una transformada DCT ortogonal, estas aproximaciones enteras generalmente se modifican en escala por un factor determinado. Para preservar la norma del bloque residual que se procesa mediante transformadas directas e inversas, se aplican factores de modificación de escala adicionales como parte del proceso de transformada. Los factores de escala se eligen típicamente con base en ciertas limitaciones, como factores de escala que son una potencia de dos para operaciones de desplazamiento, profundidad de bits de los coeficientes de transformada, compensación entre precisión y costos de implementación, etc. Los factores de modificación de escala específicos se especifican, por ejemplo, para la transformada inversa, por ejemplo, mediante la unidad 212 de procesamiento de transformada inversa (y la transformada inversa correspondiente, por ejemplo, mediante la unidad 312 de procesamiento de transformada inversa en el decodificador 30 de video) y los factores de modificación de escala correspondientes para la transformada directa, por ejemplo, mediante la unidad 206 de procesamiento de transformada, en un codificador 20 se pueden especificar en consecuencia.

Se pueden configurar realizaciones del codificador 20 de video (respectivamente unidad 206 de procesamiento de transformada) para emitir parámetros de transformada, por ejemplo, un tipo de transformada o transformada, por ejemplo, directamente o codificada o comprimida a través de la unidad 270 de codificación entrópica, de modo que, por ejemplo, el decodificador 30 de video pueda recibir y usar los parámetros de transformada para decodificación.

#### 50 Cuantificación

La unidad 208 de cuantificación se puede configurar para cuantificar los coeficientes 207 de transformada para obtener coeficientes 209 cuantificados, por ejemplo, mediante la aplicación de cuantificación escalar o cuantificación vectorial. Los coeficientes 209 cuantificados también se pueden referir como coeficientes 209 de transformada cuantificados o coeficientes 209 residuales cuantificados.

El proceso de cuantificación puede reducir la profundidad de bits asociada con algunos o todos los coeficientes 207 de transformada. Por ejemplo, un coeficiente de transformada de n-bits se puede redondear hacia abajo a un coeficiente de transformada de m-bits durante la cuantificación, donde n es mayor que m. El grado de cuantificación se puede modificar ajustando un parámetro de cuantificación (QP). Por ejemplo, para cuantificación escalar, se puede aplicar modificación de escala diferente para lograr cuantificación más fina o más gruesa. Los tamaños de paso de cuantificación más pequeños corresponden a una cuantificación más fina, mientras que los tamaños de paso de cuantificación más grandes corresponden a una cuantificación más gruesa. El tamaño de paso de cuantificación aplicable se puede indicar mediante un parámetro de

cuantificación (QP). El parámetro de cuantificación puede ser, por ejemplo, un índice de un conjunto predefinido de tamaños de paso de cuantificación aplicables. Por ejemplo, parámetros de cuantificación pequeños pueden corresponder a una cuantificación fina (tamaños de paso de cuantificación pequeños) y parámetros de cuantificación grandes pueden corresponder a una cuantificación gruesa (tamaños de paso de cuantificación grandes) o viceversa.

La cuantificación puede incluir división por un tamaño de paso de cuantificación y una correspondiente y/o la cuantificación inversa, por ejemplo, por unidad 210 de cuantificación inversa, puede incluir multiplicación por el tamaño de paso de cuantificación. Las realizaciones según algunas normas, por ejemplo, HEVC, se pueden configurar para usar un parámetro de cuantificación para determinar el tamaño de paso de cuantificación. En general, el tamaño de paso de cuantificación se puede calcular con base en un parámetro de cuantificación usando de una aproximación de punto fijo de una ecuación que incluye división. Se pueden introducir factores de modificación de escala adicionales para cuantificación y descuantificación para restaurar la norma del bloque residual, que se puede modificar debido a la modificación escala utilizada en la aproximación de punto fijo de la ecuación para el tamaño de paso de cuantificación y el parámetro de cuantificación. En una implementación de ejemplo, se puede combinar transformada inversa y descuantificación. De manera alternativa, tablas de cuantificación personalizadas se pueden utilizar y señalar desde un codificador a un decodificador, por ejemplo, en un flujo de bits. La cuantificación es una operación con pérdidas, en donde la pérdida aumenta con el incremento de tamaños de paso de cuantificación.

Se pueden configurar realizaciones del codificador 20 de video (respectivamente unidad 208 de cuantificación) para emitir parámetros de cuantificación (QP), por ejemplo, directamente o codificados a través de la unidad 270 de codificación entrópica, de modo que, por ejemplo, el decodificador 30 de video pueda recibir y aplicar los parámetros de cuantificación para decodificación.

#### 25 Cuantificación inversa

La unidad 210 de cuantificación inversa se configura para aplicar la cuantificación inversa de la unidad 208 de cuantificación en los coeficientes cuantificados para obtener coeficientes 211 descuantificados, por ejemplo, mediante la aplicación del inverso del esquema de cuantificación aplicado por la unidad 208 de cuantificación con base en o utilizando el mismo tamaño de paso de cuantificación que la unidad 208 de cuantificación. Los coeficientes 211 descuantificados también se pueden referir como coeficientes 211 residuales descuantificados y corresponden - aunque generalmente no son idénticos a los coeficientes de transformada debido a la pérdida por cuantificación - a los coeficientes 207 de transformada.

#### 35 Transformada inversa

La unidad 212 de procesamiento de transformada inversa se configura para aplicar la transformada inversa de la transformada aplicada por la unidad 206 de procesamiento de transformada, por ejemplo, una transformada inversa discreta de coseno (DCT) o transformada inversa discreta de seno (DST) u otras transformadas inversas, para obtener un bloque 213 residual reconstruido (o coeficientes 213 correspondientes descuantificados) en el dominio de muestra. El bloque 213 residual reconstruido también se puede referir como bloque 213 de transformada.

#### 45 Reconstrucción

La unidad 214 de reconstrucción (por ejemplo, añadidor o sumador 214) se configura para sumar el bloque 213 de transformada (es decir, el bloque 213 residual reconstruido) al bloque 265 de predicción para obtener un bloque 215 reconstruido en el dominio de muestra, por ejemplo, sumando - muestra por muestra - valores de muestra del bloque 213 residual reconstruido y valores de muestra del bloque 265 de predicción.

#### Filtrado

La unidad 220 de filtro de bucle (o "filtro de bucle" 220 corto), se configura para filtrar el bloque 215 reconstruido para obtener un bloque 221 filtrado, o en general, para filtrar muestras reconstruidas para obtener muestras filtradas. La unidad de filtro de bucle se configura, por ejemplo, para suavizar transiciones de píxeles o mejorar de otro modo la calidad de video. La unidad 220 de filtro de bucle puede comprender uno o más filtros de bucle tal como un filtro de desbloqueo, un filtro de desplazamiento adaptativo de muestra (SAO) o uno o más filtros adicionales, por ejemplo, un filtro bilateral, un filtro de bucle adaptativo (ALF), una definición, filtros de suavizado o filtros colaborativos o cualquier combinación de los mismos. Aunque la unidad 220 de filtro de bucle se muestra en la Fig. 2 como un filtro en bucle, en otras configuraciones, la unidad 220 de filtro de bucle se puede implementar como un filtro postbucle. El bloque 221 filtrado también se puede referir como bloque 221 reconstruido filtrado.

Se pueden configurar realizaciones del codificador 20 de video (respectivamente unidad 220 de filtro de bucle) para emitir parámetros de filtro de bucle (tal como información de desplazamiento adaptativa de muestra), por

ejemplo, directamente o codificada a través de la unidad 270 de codificación entrópica, de modo que, por ejemplo, un decodificador 30 puede recibir y aplicar los mismos parámetros de filtro de bucle o filtros de bucle respectivos para decodificación.

5 Memoria intermedia de imágenes decodificadas

La memoria intermedia 230 de imágenes decodificadas (DPB) puede ser una memoria que almacena imágenes de referencia, o en datos de imágenes de referencia generales, para codificar datos de video mediante el codificador 20 de video. La DPB 230 puede estar formada por cualquiera de una variedad de dispositivos de memoria, tal como memoria de acceso aleatorio dinámica (DRAM), que incluye DRAM síncrona (SDRAM), RAM magneto-resistiva (MRAM), RAM resistiva (RRAM) u otros tipos de dispositivos de memoria. Se puede configurar la memoria intermedia 230 de imágenes decodificadas (DPB) para almacenar uno o más bloques 221 filtrados. La memoria intermedia 230 de imágenes decodificadas se puede configurar además para almacenar otros bloques previamente filtrados, por ejemplo, bloques 221 previamente reconstruidos y filtrados, de la misma imagen actual o de imágenes diferentes, por ejemplo, imágenes previamente reconstruidas, y puede proporcionar imágenes completas previamente reconstruidas, es decir, decodificadas, imágenes (y bloques y muestras de referencia correspondientes) y/o una imagen actual parcialmente reconstruida (y bloques y muestras de referencia correspondientes), por ejemplo para interpredicción. La memoria intermedia 230 de imágenes decodificadas (DPB) también se puede configurar para almacenar uno o más bloques 215 reconstruidos no filtrados, o en general muestras reconstruidas no filtradas, por ejemplo, si el bloque 215 reconstruido no se filtra mediante la unidad 220 de filtro de bucle, o cualquier otra versión procesada adicional de los bloques o muestras reconstruidas.

Selección de modo (partición y predicción)

La unidad 260 de selección de modo comprende la unidad 262 de partición, la unidad 244 de interpredicción y la unidad 254 de intrapredicción, y se configura para recibir u obtener datos de imagen originales, por ejemplo, un bloque 203 original (bloque 203 actual de la imagen 17 actual), y datos de imagen reconstruidos, por ejemplo, muestras reconstruidas filtradas y/o no filtradas o bloques de la misma imagen (actual) y/o de una o una pluralidad de imágenes previamente decodificadas, por ejemplo, de la memoria intermedia 230 de imágenes decodificadas u otras memorias intermedias (por ejemplo, memoria intermedia lineal, no mostrada). Los datos de imagen reconstruidos se utilizan como datos de imagen de referencia para predicción, por ejemplo, interpredicción o intrapredicción, para obtener un bloque 265 de predicción o predictor 265.

La unidad 260 de selección de modo se puede configurar para determinar o seleccionar una partición para un modo de predicción de bloque actual (que incluye sin partición) y un modo de predicción (por ejemplo, un modo de intra- o interpredicción) y generar un bloque 265 de predicción correspondiente, que se utiliza para el cálculo del bloque 205 residual y para la reconstrucción del bloque 215 reconstruido.

Las realizaciones de la unidad 260 de selección de modo se pueden configurar para seleccionar la partición y el modo de predicción (por ejemplo de aquellos soportados por o disponibles para la unidad 260 de selección de modo), que proporcionan la mejor coincidencia o, en otras palabras, el residuo mínimo (residuo mínimo significa mejor compresión para transmisión o almacenamiento) o una sobrecarga de señalización mínima (sobrecarga de señalización mínima significa mejor compresión para transmisión o almacenamiento), o que considera o equilibra ambos. La unidad 260 de selección de modo se puede configurar para determinar el modo de partición y predicción con base en la optimización de distorsión de tasa (RDO), es decir, seleccionar el modo de predicción que proporciona una distorsión de tasa mínima. Términos como "mejor", "mínimo", "óptimo", etc. en este contexto no se refieren necesariamente a un "mejor", "mínimo", "óptimo" general etc., pero también pueden hacer referencia al cumplimiento de un criterio de terminación o selección como un valor que excede o cae por debajo de un umbral u otras restricciones que conducen potencialmente a una "selección subóptima", pero que reduce la complejidad y el tiempo de procesamiento.

En otras palabras, la unidad 262 de partición se puede configurar para particionar el bloque 203 en particiones o subbloques de bloques más pequeños (que forman nuevamente bloques), por ejemplo, iterativamente usando particiones de árbol cuádruple (QT), particiones binarias (BT) o particiones de árbol triple (TT) o cualquier combinación de las mismas, y para realizar, por ejemplo, la predicción para cada una de las particiones o subbloques de bloques, en donde la selección de modo comprende la selección de la estructura de árbol del bloque 203 particionado y los modos de predicción se aplican a cada una de las particiones de bloques o subbloques.

En lo siguiente, la partición (por ejemplo, mediante la unidad 260 de partición) y el procesamiento de predicción (por la unidad 244 de interpredicción y la unidad 254 de intrapredicción) realizados por un codificador de video de ejemplo 20 se explicarán en más detalle.

65 Partición

La unidad 262 de partición puede particionar (o dividir) un bloque 203 actual en particiones más pequeñas, por ejemplo, bloques más pequeños de tamaño cuadrado o rectangular. Estos bloques más pequeños (que también se pueden referir como subbloques) se pueden particionar además en particiones aún más pequeñas. Esto también se refiere a la partición de árboles o partición jerárquica de árboles, en donde un bloque raíz, por ejemplo, en el nivel de árbol raíz 0 (nivel de jerarquía 0, profundidad 0), se puede particionar recursivamente, por ejemplo, particionar en dos o más bloques de un siguiente nivel de árbol inferior, por ejemplo, nodos en nivel de árbol 1 (nivel de jerarquía 1, profundidad 1), en donde estos bloques se pueden particionar nuevamente en dos o más bloques de un siguiente nivel inferior, por ejemplo, nivel de árbol 2 (nivel de jerarquía 2, profundidad 2), etc. hasta que se termina la partición, por ejemplo, porque se cumple un criterio de terminación, por ejemplo, se alcanza una profundidad de árbol máxima o un tamaño de bloque mínimo. Los bloques que no se particionan adicionalmente también se conocen como bloques de hoja o nodos hoja del árbol. Un árbol que usa particiones en dos particiones se conoce como árbol binario (BT), un árbol que usa particiones en tres particiones se conoce como árbol ternario (TT), y un árbol que usa particiones en cuatro particiones se conoce como árbol cuádruple (QT).

Como se mencionó anteriormente, el término "bloque", tal como se usa en el presente documento, puede ser una parte, en particular una parte cuadrada o rectangular, de una imagen. Con referencia, por ejemplo, a HEVC y VVC, el bloque puede ser o corresponder a una unidad de árbol de codificación (CTU), una unidad de codificación (CU), una unidad de predicción (PU) y una unidad de transformada (TU) y/o a los bloques correspondientes, por ejemplo, un bloque de árbol de codificación (CTB), un bloque de codificación (CB), un bloque de transformada (TB) o un bloque de predicción (PB).

Por ejemplo, una unidad de árbol de codificación (CTU) puede ser o comprender un CTB de muestras de luma, dos CTB correspondientes de muestras de croma de una imagen que tiene tres matrices de muestras o un CTB de muestras de una imagen monocromática o una imagen que se codifica usando tres planos de color separados y estructuras de sintaxis utilizadas para codificar las muestras. De manera correspondiente, un bloque de árbol de codificación (CTB) puede ser un bloque de muestras de  $N \times N$  para algún valor de  $N$  de modo que la división de un componente en CTB sea una partición. Una unidad de codificación (CU) puede ser o comprender un bloque de codificación de muestras de luma, dos bloques de codificación correspondientes de muestras de croma de una imagen que tiene tres matrices de muestras, o un bloque de codificación de muestras de una imagen monocromática o una imagen que se codifica usando tres planos de color separados y estructuras de sintaxis utilizadas para codificar las muestras. En consecuencia, un bloque de codificación (CB) puede ser un bloque de muestras de  $M \times N$  para algunos valores de  $M$  y  $N$  de modo que la división de un CTB en bloques de codificación sea una partición.

En realizaciones, por ejemplo, según HEVC, una unidad de árbol de codificación (CTU) se puede dividir en CU mediante el uso de una estructura de árbol cuádruple denotada como árbol de codificación. La decisión de si se debe codificar un área de imagen utilizando predicción interimagen (temporal) o intrainimagen (espacial) se toma al nivel de CU. Cada CU se puede dividir además en una, dos o cuatro PU según el tipo de división PU. Dentro de una PU, se aplica el mismo proceso de predicción y la información relevante se transmite al decodificador de una forma por PU. Después de obtener el bloque residual mediante la aplicación del proceso de predicción con base en el tipo de división PU, una CU se puede particionar en unidades de transformada (TU) según otra estructura de árbol cuádruple similar al árbol de codificación para la CU.

En realizaciones, por ejemplo, según la norma de codificación de video más reciente actualmente en desarrollo, a la que se hace referencia como codificación de video versátil (VVC), se usa, por ejemplo, una partición de árbol cuádruple y árbol binario (QTBT) combinada para particionar un bloque de codificación. En la estructura de bloque QTBT, una CU puede tener una forma cuadrada o rectangular. Por ejemplo, una unidad de árbol de codificación (CTU) se particiona primero mediante una estructura de árbol cuádruple. Los nodos hoja de árbol cuádruple se particionan adicionalmente mediante un árbol binario o estructura de árbol ternario (o triple). Los nodos hoja de árbol de partición se denominan unidades de codificación (CU), y esa segmentación se utiliza para predicción y procesamiento de transformada sin ninguna partición adicional. Esto significa que la CU, PU y TU tienen el mismo tamaño de bloque en la estructura de bloque de codificación QTBT. En paralelo, la partición múltiple, por ejemplo, la partición de árbol triple se puede usar junto con la estructura de bloque QTBT.

En algunas realizaciones, tal como en el borrador de norma VVC, con el fin de facilitar una segmentación de procesamiento en hardware con memoria interna limitada en comparación con la CTU, se define una unidad de datos de segmentación virtual (VPDU). Una VPDU es una partición virtual de una CTU en subbloques uniformes de muestras de luminancia y muestras de crominancia correspondientes con un orden de procesamiento específico a través de las particiones dentro de la CTU de tal manera que el procesamiento de una VPDU dada no depende del procesamiento de cualquier otra VPDU futura en el orden de procesamiento. Sin embargo, ciertos elementos de la sintaxis todavía se pueden señalar en el flujo de bits a nivel de CTU y se aplicarán a todas las VPDU en esa CTU. Se pueden imponer ciertas restricciones a la partición para asegurar que una unidad de codificación abarque completamente una o más VPDU, pero no puede cubrir parcialmente una VPDU. En un ejemplo, la unidad 260 de selección de modo del codificador 20 de video se puede configurar para realizar cualquier combinación de las técnicas de partición descritas en el presente

documento.

5 Como se describió anteriormente, el codificador 20 de video se configura para determinar o seleccionar el mejor modo de predicción o un modo de predicción óptimo de un conjunto de modos de predicción (por ejemplo, predeterminados). El conjunto de modos de predicción puede comprender, por ejemplo, modos de intrapredicción y/o modos de interpredicción.

#### Intrapredicción

10 El conjunto de modos de intrapredicción puede comprender 35 modos de intrapredicción diferentes, por ejemplo, modos no direccionales como modo DC (o medio) y modo plano, o modos direccionales, por ejemplo, como se define en HEVC, o puede comprender 67 modos de intrapredicción diferentes, por ejemplo, modos no direccionales como modo DC (o medio) y modo plano, o modos direccionales, por ejemplo, como se define para VVC.

15 La unidad 254 de intrapredicción se configura para usar muestras reconstruidas de bloques vecinos de la misma imagen actual para generar un bloque 265 de intrapredicción según un modo de intrapredicción del conjunto de modos de intrapredicción.

20 La unidad 254 de intrapredicción (o en general la unidad 260 de selección de modo) se configura además para emitir parámetros de intrapredicción (o en información general indicativa del modo de intrapredicción seleccionado para el bloque) a la unidad 270 de codificación entrópica en forma de elementos de sintaxis 266 para inclusión en los datos 21 de imagen codificados, de modo que, por ejemplo, el decodificador 30 de video puede recibir y usar los parámetros de predicción para decodificación.

25

#### Interpredicción

30 El conjunto de (o posibles) modos de interpredicción depende de las imágenes de referencia disponibles (es decir, imágenes anteriores al menos parcialmente decodificadas, por ejemplo, almacenadas en la DPB 230) y otros parámetros de interpredicción, por ejemplo, si la imagen de referencia completa o solo una parte, por ejemplo, un área de ventana de búsqueda alrededor del área del bloque actual, de la imagen de referencia se utiliza para buscar un bloque de referencia que coincida mejor y/o por ejemplo, si se aplica interpolación de píxeles, por ejemplo, interpolación de medio/semipíxel y/o cuarto de píxel, o no. Además de los modos de predicción anteriores, se puede aplicar un modo de salto y/o un modo directo.

35

40 La unidad 244 de interpredicción puede incluir una unidad de estimación de movimiento (ME) y una unidad de compensación de movimiento (MC) (ambas no se muestran en la Fig. 2). La unidad de estimación de movimiento se puede configurar para recibir u obtener el bloque 203 de imagen (bloque 203 de imagen actual de la imagen 17 actual) y una imagen 231 decodificada, o al menos uno o una pluralidad de bloques previamente reconstruidos, por ejemplo, bloques reconstruidos de una o una pluralidad de otras/diferentes imágenes 231 previamente decodificadas, para estimación de movimiento. Por ejemplo, una secuencia de video puede comprender la imagen actual y las imágenes 231 previamente decodificadas, o, en otras palabras, la imagen actual y las imágenes 231 previamente decodificadas pueden ser parte de o formar una secuencia de imágenes que forman una secuencia de video.

45

50 El codificador 20 se puede configurar, por ejemplo, para seleccionar un bloque de referencia de una pluralidad de bloques de referencia de la misma o diferentes imágenes de la pluralidad de otras imágenes y proporcionar una imagen de referencia (o índice de imagen de referencia) y/o un desplazamiento (desplazamiento espacial) entre la posición (coordenadas x, y) del bloque de referencia y la posición del bloque actual como parámetros de interpredicción a la unidad de estimación de movimiento. Este desplazamiento también se conoce como vector de movimiento (MV).

55 La unidad de compensación de movimiento se configura para obtener, por ejemplo, recibir, un parámetro de interpredicción y para realizar interpredicción con base en o mediante el uso del parámetro de interpredicción para obtener un bloque 265 de interpredicción. La compensación de movimiento, realizada por la unidad de compensación de movimiento, puede implicar recuperar o generar el bloque de predicción con base en el vector de movimiento/bloque determinado por estimación de movimiento, posiblemente realizando interpolaciones a precisión de subpíxeles. El filtrado de interpolación puede generar muestras de píxeles adicionales a partir de muestras de píxeles conocidas, aumentando así la cantidad de bloques de predicción candidatos que se pueden usar para codificar un bloque de imagen. Después de recibir el vector de movimiento para la PU del bloque de imagen actual, la unidad de compensación de movimiento puede ubicar el bloque de predicción al que apunta el vector de movimiento en una de las listas de imágenes de referencia.

60

65 La unidad de compensación movimiento también puede generar elementos de sintaxis asociados con los bloques y segmentos de video para su uso por el decodificador 30 de video en la decodificación de los bloques de imagen del segmento de video. Además o como una alternativa a los segmentos y los elementos de sintaxis

respectivos, se pueden generar o usar grupos de mosaicos y/o mosaicos y elementos de sintaxis respectivos.

5 Como se analizará en detalle más adelante, las realizaciones presentadas en el presente documento proporcionan una mejora a la unidad 244 de interpredicción al proporcionar una predicción de vector de movimiento más precisa usada por la unidad de interpredicción cuando se realiza la interpredicción, por ejemplo, la interpredicción basada en flujo óptico bidireccional (BDOF).

#### Codificación entrópica

10 La unidad 270 de codificación entrópica se configura para aplicar, por ejemplo, un algoritmo o esquema de codificación entrópica (por ejemplo, un esquema de codificación de longitud variable (VLC), un esquema VLC adaptativo al contexto (CAVLC), un esquema de codificación aritmética binaria adaptativa a la probabilidad (CABAC), una codificación aritmética binaria adaptativa al contexto basada en sintaxis (SBAC), codificación entrópica de partición de intervalo de probabilidad (PIPE) u otra metodología o técnica de codificación entrópica) o derivación (sin compresión) en los coeficientes 209 cuantificados, parámetros de interpredicción, parámetros de intrapredicción, parámetros de filtro de bucle y/u otros elementos de sintaxis para obtener datos 21 de imagen codificados que se pueden generar a través de la salida 272, por ejemplo, en forma de un flujo de bits codificado 21, de modo que, por ejemplo, el decodificador 30 de video puede recibir y usar los parámetros para decodificación. El flujo de bits 21 codificado se puede transmitir al decodificador 30 de video, 20 o almacenarse en una memoria para su posterior transmisión o recuperación mediante el decodificador 30 de video.

25 Otras variaciones estructurales del codificador 20 de video se pueden usar para codificar el flujo de video. Por ejemplo, un codificador no basado en transformada 20 puede cuantificar la señal residual directamente sin la unidad 206 de procesamiento de transformada para determinados bloques o fotogramas. En otra implementación, un codificador 20 puede tener la unidad 208 de cuantificación y la unidad 210 de cuantificación inversa combinadas en una sola unidad.

#### Decodificador y método de decodificación

30 La Fig. 3 muestra un decodificador 30 de video de ejemplo que se configura para implementar las técnicas del presente documento solicitud. El decodificador 30 de video se configura para recibir datos 21 de imagen codificados (por ejemplo, flujo de bits 21 codificado), por ejemplo, codificado por el codificador 20, para obtener una imagen 331 decodificada. Los datos de imagen codificados comprenden información para decodificar los 35 datos de imagen codificados, por ejemplo, datos que representan bloques de imagen de un segmento de video codificado (y/o grupos de mosaicos o mosaicos) y elementos de sintaxis asociados.

40 En el ejemplo de la Fig. 3, el decodificador 30 comprende una unidad 304 de decodificación entrópica, una unidad 310 de cuantificación inversa, una unidad 312 de procesamiento de transformada inversa, una unidad 314 de reconstrucción (por ejemplo, un sumador 314), un filtro 320 de bucle, una memoria intermedia 330 de imágenes decodificadas (DBP), una unidad 360 de aplicación de modo, una unidad 344 de interpredicción y una unidad 354 de intrapredicción. La unidad 344 de intrapredicción puede ser o incluir una unidad de compensación de movimiento. El decodificador de video 30 puede, en algunos ejemplos, realizar una pasada de decodificación generalmente recíproca a la pasada de codificación descrita con respecto al codificador 100 45 de video de la Fig. 2.

50 Tal como se explica con respecto al codificador 20, también se hace referencia a la unidad 210 de cuantificación inversa, la unidad 212 de procesamiento de transformada inversa, la unidad 214 de reconstrucción, el filtro 220 de bucle, la memoria intermedia 230 de imágenes decodificadas (DPB), la unidad 344 de interpredicción y la unidad 354 de intrapredicción para formar el "decodificador integrado" del codificador 20 de video. Por consiguiente, la unidad 310 de cuantificación inversa puede ser idéntica en función a la unidad 110 de cuantificación inversa, la unidad 312 de procesamiento de transformada inversa puede ser idéntica en función a la unidad de procesamiento de transformada inversa 212, la unidad de reconstrucción 314 puede ser idéntica en función a la unidad de reconstrucción 214, el filtro de bucle 320 puede ser idéntico en función al filtro 220 55 de bucle y la memoria intermedia 330 de imágenes decodificadas puede ser idéntica en función a la memoria intermedia 230 de imágenes decodificadas. Por lo tanto, las explicaciones proporcionadas para las respectivas unidades y funciones del codificador 20 de video se aplican de manera correspondiente a las respectivas unidades y funciones del decodificador 30 de video.

#### Decodificación entrópica

60 La unidad 304 de decodificación entrópica se configura para analizar el flujo de bits 21 (o en general datos 21 de imagen codificados) y realizar, por ejemplo, decodificación entrópica a los datos 21 de imagen codificados para obtener, por ejemplo, coeficientes 309 cuantificados y/o parámetros de codificación decodificados (no mostrados en la Fig. 3), por ejemplo, cualquiera o todos los parámetros de interpredicción (por ejemplo, índice de imagen de referencia y vector de movimiento), parámetro intrapredicción (por ejemplo, modo o índice 65

intrapredicción), parámetros de transformada, parámetros de cuantificación, parámetros de filtro de bucle y/u otros elementos de sintaxis. La unidad 304 de decodificación entrópica se puede configurar para aplicar los algoritmos o esquemas de decodificación correspondientes a los esquemas de codificación como se describe con respecto a la unidad 270 de codificación entrópica del codificador 20. La unidad 304 de decodificación entrópica se puede configurar además para proporcionar parámetros de interpredicción, parámetros de intrapredicción y/u otros elementos de sintaxis a la unidad 360 de aplicación de modo y otros parámetros a otras unidades del decodificador 30. El decodificador 30 de video puede recibir los elementos de sintaxis al nivel de segmento de video y/o al nivel de bloque de video. Además o como una alternativa a los segmentos y los elementos de sintaxis respectivos, se pueden recibir y/o utilizar grupos de mosaicos y/o mosaicos y elementos de sintaxis respectivos.

#### Cuantificación inversa

La unidad 310 de cuantificación inversa se puede configurar para recibir parámetros de cuantificación (QP) (o en información general relacionada con la cuantificación inversa) y coeficientes cuantificados a partir de los datos 21 de imagen codificados (por ejemplo, mediante análisis y/o decodificación, por ejemplo, mediante unidad 304 de decodificación entrópica) y para aplicar con base en los parámetros de cuantificación una cuantificación inversa en los coeficientes 309 cuantificados decodificados para obtener coeficientes 311 descuantificados, que también se pueden referir como coeficientes 311 de transformada. El proceso de cuantificación inversa puede incluir el uso de un parámetro de cuantificación determinado mediante el codificador 20 de video para cada bloque de video en el segmento de video (o grupo de mosaicos o mosaico) para determinar un grado de cuantificación y, de manera similar, un grado de cuantificación inversa que se debe aplicar.

#### Transformada inversa

La unidad 312 de procesamiento de transformada inversa se puede configurar para recibir coeficientes 311 descuantificados, también denominados coeficientes 311 de transformada, y para aplicar una transformada a los coeficientes 311 descuantificados para obtener bloques 213 residuales reconstruidos en el dominio de muestra. Los bloques 213 residuales reconstruidos también se puede referir como bloques 313 de transformada. La transformada puede ser una transformada inversa, por ejemplo, una DCT inversa, una DST inversa, una transformada de enteros inversa o un proceso de transformada inversa conceptualmente similar. La unidad 312 de procesamiento de transformada inversa se puede configurar adicionalmente para recibir parámetros de transformada o información correspondiente a partir de los datos 21 de imagen codificados (por ejemplo, mediante análisis y/o decodificación, por ejemplo, mediante unidad 304 de decodificación entrópica) para determinar la transformada que se va a aplicar a los coeficientes 311 descuantificados.

#### Reconstrucción

La unidad 314 de reconstrucción (por ejemplo, añadidor o sumador 314) se puede configurar para sumar el bloque 313 residual reconstruido al bloque 365 de predicción para obtener un bloque 315 reconstruido en el dominio de muestra, por ejemplo, sumando los valores de muestra del bloque 313 residual reconstruido y los valores de muestra del bloque 365 de predicción.

#### Filtrado

La unidad 320 de filtro de bucle (ya sea en el bucle de codificación o después del bucle de codificación) se configura para filtrar el bloque 315 reconstruido para obtener un bloque 321 filtrado, por ejemplo, para suavizar transiciones de píxeles o mejorar de otro modo la calidad de video. La unidad 320 de filtro de bucle puede comprender uno o más filtros de bucle tal como un filtro de desbloqueo, un filtro de desplazamiento adaptativo de muestra (SAO) o uno o más filtros adicionales, por ejemplo, un filtro bilateral, un filtro de bucle adaptativo (ALF), una definición, filtros de suavizado o filtros colaborativos o cualquier combinación de los mismos. Aunque la unidad de filtro 320 de bucle se muestra en la Fig. 3 como un filtro en bucle, en otras configuraciones, la unidad 320 de filtro de bucle se puede implementar como un filtro post-bucle.

#### Memoria intermedia de imágenes decodificadas

Los bloques 321 de video decodificados de una imagen se almacenan en la memoria intermedia de imágenes 330 decodificadas, que almacena las imágenes 331 decodificadas como imágenes de referencia para compensación de movimiento posterior para otras imágenes y/o para visualización de salida, respectivamente. El decodificador 30 se configura para emitir la imagen 311 decodificada, por ejemplo, a través de la salida 312, para su presentación o visualización a un usuario.

#### Predicción

La unidad 344 de interpredicción puede ser idéntica a la unidad 244 de interpredicción (en particular a la unidad

de compensación de movimiento) y la unidad 354 de intrapredicción puede ser idéntica a la unidad 254 de intrapredicción en función, y toma decisiones de división o partición y predicción con base en los parámetros de partición y/o predicción o información respectiva recibida a partir de los datos 21 de imagen codificados (por ejemplo, mediante análisis y/o decodificación, por ejemplo, mediante la unidad 304 de decodificación entrópica). La unidad 360 de aplicación de modo se puede configurar para realizar la predicción (intra o interpredicción) por bloque con base en imágenes reconstruidas, bloques o muestras respectivas (filtradas o no filtradas) para obtener el bloque 365 de predicción.

10 Cuando el segmento de video se codifica como un segmento intracodificado (I), la unidad 354 de intrapredicción de la unidad 360 de aplicación de modo se configura para generar el bloque 365 de predicción para un bloque de imagen del segmento de video actual con base en un modo de intrapredicción señalado y datos de bloques previamente decodificados de la imagen actual. Cuando la imagen de video se codifica como un segmento intercodificado (es decir, B o P), la unidad 344 de intercodificación (por ejemplo, unidad de compensación de movimiento) de la unidad 360 de aplicación de modo se configura para producir bloques 365 de predicción para un bloque de video del segmento de video actual con base en los vectores de movimiento y otros elementos de sintaxis recibidos de la unidad 304 de decodificación entrópica. Para interpredicción, los bloques de predicción se pueden producir a partir de una de las imágenes de referencia dentro de una de las listas de imágenes de referencia. El decodificador 30 de video puede construir las listas de fotogramas de referencia, Lista 0 y Lista 1, usando técnicas de construcción predeterminadas basadas en imágenes de referencia almacenadas en DPB 330. Lo mismo o similar se puede solicitar o mediante realizaciones mediante el uso de grupos de mosaicos (por ejemplo, grupos de mosaicos de video) y/o mosaicos (por ejemplo, mosaicos de video) además o alternativamente a segmentos (por ejemplo, segmentos de video), por ejemplo, un video se puede codificar mediante el uso de grupos de mosaicos y/o mosaicos I, P o B.

25 Como se analizará en detalle más adelante, las realizaciones presentadas en el presente documento proporcionan una mejora a la unidad 344 de interpredicción al proporcionar una predicción de vector de movimiento más precisa usada por la unidad de interpredicción cuando se realiza la interpredicción, por ejemplo, la interpredicción basada en flujo óptico bidireccional (BDOF).

30 La unidad 360 de aplicación de modo se configura para determinar la información de predicción para un bloque de video del segmento de video actual al analizar los vectores de movimiento o información relacionada y otros elementos de sintaxis y usa la información de predicción para producir los bloques de predicción para el bloque de video actual que se decodifica. Por ejemplo, la unidad 360 de aplicación de modo usa algunos de los elementos de sintaxis recibidos para determinar un modo de predicción (por ejemplo, intra o interpredicción) utilizado para codificar los bloques de video del segmento de video, un tipo de segmento de interpredicción (por ejemplo, segmento B, segmento P o segmento GPB), información de construcción para una o más de las listas de imágenes de referencia para el segmento, vectores de movimiento para cada bloque de video intercodificado del segmento, estado de interpredicción para cada bloque de video intercodificado del segmento y otra información para decodificar los bloques de video en el segmento de video actual. Lo mismo o similar se puede solicitar o mediante realizaciones mediante el uso de grupos de mosaicos (por ejemplo, grupos de mosaicos de video) y/o mosaicos (por ejemplo, mosaicos de video) además o alternativamente a segmentos (por ejemplo, segmentos de video), por ejemplo, un video se puede codificar mediante el uso de grupos de mosaicos y/o mosaicos I, P o B.

45 Las realizaciones del decodificador de video 30 como se muestran en la Fig. 3 se pueden configurar para particionar y/o decodificar la imagen mediante el uso de segmentos (también denominados segmentos de video), en donde una imagen se puede particionar o decodificar mediante el uso de uno o más segmentos (típicamente no superpuestos), y cada segmento puede comprender uno o más bloques (por ejemplo, CTU).

50 Las realizaciones del decodificador 30 de video como se muestra en la Fig. 3 se pueden configurar para particionar y/o decodificar la imagen mediante el uso de grupos de mosaicos (también denominados grupos de mosaicos de video) y/o mosaicos (también denominados mosaicos de video), en donde una imagen se puede particionar o decodificar mediante el uso de uno o más grupos de mosaicos (típicamente no superpuestos), y cada grupo de mosaicos puede comprender, por ejemplo uno o más bloques (por ejemplo, CTU) o una o más mosaicos, en donde cada mosaico, por ejemplo puede tener forma rectangular y puede comprender uno o más bloques (por ejemplo, CTU), por ejemplo, bloques completos o fraccionarios.

60 Otras variaciones del decodificador 30 de video se pueden usar para decodificar los datos 21 de imagen codificados. Por ejemplo, el decodificador 30 puede producir el flujo de video de salida sin la unidad 320 de filtrado de bucle. Por ejemplo, un decodificador 30 no basado en transformada puede cuantificar inversamente la señal residual directamente sin la unidad 312 de procesamiento de transformada inversa para determinados bloques o fotogramas. En otra implementación, el decodificador de video 30 puede tener la unidad 310 de cuantificación inversa y la unidad 312 de procesamiento de transformada inversa combinadas en una sola unidad.

65 Se debe entender que, en el codificador 20 y el decodificador 30, un resultado de procesamiento de un paso

actual se puede procesar adicionalmente y luego enviar al siguiente paso. Por ejemplo, después de filtrado de interpolación, derivación de vector de movimiento o filtrado de bucle, una operación adicional, tal como recorte o desplazamiento, se puede realizar en el resultado de procesamiento del filtrado de interpolación, derivación de vector de movimiento o filtrado de bucle.

5

Se debe observar que se pueden aplicar operaciones adicionales a los vectores de movimiento derivados del bloque actual (que incluyen, pero no se limitan a, vectores de movimiento de punto de control del modo afín, vectores de movimiento de subbloque en modos afín, plano, ATMVP, vectores de movimiento temporal, etc.). Por ejemplo, el valor del vector de movimiento se restringe a un intervalo predefinido según su bit representativo. Si el bit representativo del vector de movimiento es Profundidadbit, entonces el intervalo es  $-2^{(Profundidadbit - 1)} \sim 2^{(Profundidadbit - 1) - 1}$ , donde "A" significa exponenciación. Por ejemplo, si Profundidadbit se establece igual a 16, el intervalo es  $-32768 \sim 32767$ ; si Profundidadbit se establece igual a 18, el intervalo es  $-131072 \sim 131071$ . Por ejemplo, el valor del vector de movimiento derivado (por ejemplo, los MV de cuatro subbloques de 4x4 dentro de un bloque de 8x8) se restringe de modo que la diferencia máxima entre las partes enteras de los cuatro MV de subbloque de 4x4 no es más de N píxeles, tal como no más de 1 píxel. Aquí se proporcionan dos métodos para restringir el vector de movimiento según la profundidad de bits.

10

15

Método 1: eliminar el MSB de desbordamiento (bit más significativo) mediante operaciones de flujo

20

$$ux = (mvx + 2^{Profundidadbit}) \% 2^{Profundidadbit} \quad (1)$$

$$mvx = (ux \geq 2^{Profundidadbit - 1}) ? (ux - 2^{Profundidadbit}) : ux \quad (2)$$

25

$$uy = (mvy + 2^{Profundidadbit}) \% 2^{Profundidadbit} \quad (3)$$

$$mvy = (uy \geq 2^{Profundidadbit - 1}) ? (uy - 2^{Profundidadbit}) : uy \quad (4)$$

30

donde mvx es un componente horizontal de un vector de movimiento de un bloque de imagen o un subbloque, mvy es un componente vertical de un vector de movimiento de un bloque de imagen o un subbloque y ux y uy indican un valor intermedio.

35

Por ejemplo, si el valor de mvx es  $-32769$ , después de aplicar las fórmulas (1) y (2), el valor resultante es  $32767$ . En un sistema informático, los números decimales se almacenan como complementos de dos. El complemento de dos de  $-32769$  es  $1,0111,1111,1111,1111$  (17 bits), luego el MSB se descarta, por lo que el complemento de dos resultante es  $0111,1111,1111,1111$  (el número decimal es  $32767$ ), que es el mismo que el producto mediante la aplicación de las fórmulas (1) y (2).

$$ux = (mvp_x + mvd_x + 2^{Profundidadbit}) \% 2^{Profundidadbit} \quad (5)$$

40

$$mvx = (ux \geq 2^{Profundidadbit - 1}) ? (ux - 2^{Profundidadbit}) : ux \quad (6)$$

$$uy = (mvp_y + mvd_y + 2^{Profundidadbit}) \% 2^{Profundidadbit} \quad (7)$$

45

$$mvy = (uy \geq 2^{Profundidadbit - 1}) ? (uy - 2^{Profundidadbit}) : uy \quad (8)$$

Las operaciones se pueden aplicar durante la suma de mvp y mvd, como se muestra en la fórmula (5) a (8).

Método 2: eliminar el MSB de desbordamiento recortando el valor

50

$$vx = \text{Clip3}(-2^{Profundidadbit - 1}, 2^{Profundidadbit - 1} - 1, vx) \quad (9)$$

$$vy = \text{Clip3}(-2^{Profundidadbit - 1}, 2^{Profundidadbit - 1} - 1, vy) \quad (10)$$

55

donde vx es un componente horizontal de un vector de movimiento de un bloque de imagen o un subbloque, vy es un componente vertical de un vector de movimiento de un bloque de imagen o un subbloque; x, y y z corresponden respectivamente a tres valores de entrada del proceso de recorte MV y la definición de función Clip3 es como sigue:

$$\text{Clip3}(x, y, z) = \begin{cases} x & ; & z < x \\ y & ; & z > y \\ z & ; & \text{de otra forma} \end{cases} \quad (11)$$

60

La Fig. 4 es un diagrama esquemático de un dispositivo de codificación de video 400 según una modalidad. El dispositivo 400 de codificación de video es adecuado para implementar las realizaciones descritas como se describe en el presente documento. En una modalidad, el dispositivo de codificación de video 400 puede ser un decodificador tal como decodificador de video 30 de la Fig. 1A o un codificador tal como codificador 20 de video de la Fig. 1A.

El dispositivo 400 de codificación de video comprende puertos 410 de ingreso (o puertos 410 de entrada) y unidades 420 receptoras (Rx) para recibir datos; un procesador, unidad lógica o unidad 430 de procesamiento central (CPU) para procesar los datos; unidades 440 transmisoras (Tx) y puertos 450 de egreso (o puertos 450 de salida) para transmitir los datos; y una memoria 460 para almacenar los datos. El dispositivo 400 de codificación de video también puede comprender componentes óptico a eléctrico (OE) y componentes eléctrico a óptico (EO) acoplados a los puertos 410 de ingreso, las unidades 420 receptoras, las unidades 440 transmisoras y los puertos 450 de egreso para salida o entrada de señales ópticas o eléctricas.

El procesador 430 se implementa por hardware y software. El procesador 430 se puede implementar como uno o más chips de CPU, núcleos (por ejemplo, como un procesador de múltiples núcleos), FPGA, ASIC y DSP. El procesador 430 está en comunicación con los puertos de ingreso 410, unidades 420 receptoras, unidades 440 transmisoras, puertos 450 de egreso y memoria 460. El procesador 430 comprende un módulo 470 de codificación. El módulo 470 de codificación implementa las realizaciones descritas anteriormente. Por ejemplo, el módulo 470 de codificación implementa, procesa, prepara o proporciona las diversas operaciones de codificación. Por lo tanto, la inclusión del módulo 470 de codificación proporciona una mejora sustancial a la funcionalidad del dispositivo 400 de codificación de video y efectúa una transformación del dispositivo 400 de codificación de video a un estado diferente. Alternativamente, el módulo 470 de codificación se implementa como instrucciones almacenadas en la memoria 460 y ejecutadas por el procesador 430.

La memoria 460 puede comprender uno o más discos, unidades de cinta y unidades de estado sólido y se puede usar como un dispositivo de almacenamiento de datos de desbordamiento, para almacenar programas cuando estos programas se seleccionan para ejecución, y para almacenar instrucciones y datos que se leen durante la ejecución de programa. La memoria 460 puede ser, por ejemplo, volátil y/o no volátil y puede ser memoria de solo A lectura (ROM), memoria de acceso aleatorio (RAM), memoria ternaria direccionable por contenido (TCAM) y/o memoria estática de acceso aleatorio (SRAM).

La Fig. 5 es un diagrama de bloques simplificado de un aparato 500 que se puede usar como cualquiera o ambos del dispositivo 12 fuente y el dispositivo 14 destino de la Fig. 1A según una modalidad de ejemplo. Un procesador 502 en el aparato 500 puede ser una unidad de procesamiento central. Alternativamente, el procesador 502 puede ser cualquier otro tipo de dispositivo o múltiples dispositivos capaces de manipular o procesar información ahora existente o desarrollada posteriormente. Aunque las implementaciones descritas se pueden poner en práctica con un único procesador como se muestra, por ejemplo, el procesador 502, se pueden lograr ventajas en velocidad y eficiencia usando más de un procesador.

Una memoria 504 en el aparato 500 puede ser un dispositivo de memoria de solo lectura (ROM) o un dispositivo de memoria de acceso aleatorio (RAM) en una implementación. Se puede usar cualquier otro tipo adecuado de dispositivo de almacenamiento como la memoria 504. La memoria 504 puede incluir código y datos 506 a los que se accede por el procesador 502 usando un bus 512. La memoria 504 puede incluir además un sistema operativo 508 y programas de aplicación 510, los programas 510 de aplicación que incluyen al menos un programa que permite que el procesador 502 realice los métodos descritos en el presente documento. Por ejemplo, los programas 510 de solicitud pueden incluir aplicaciones 1 a N, que incluyen además una aplicación de codificación de video que realiza los métodos descritos en el presente documento.

El aparato 500 también puede incluir uno o más dispositivos de salida, tal como una pantalla 518. El elemento de visualización 518 puede ser, en un ejemplo, un elemento de visualización sensible al tacto que combina un elemento de visualización con un elemento sensible al tacto que se puede operar para detectar entradas táctiles. El elemento de visualización 518 se puede acoplar al procesador 502 mediante el bus 512.

Aunque se representa aquí como un solo bus, el bus 512 del aparato 500 puede estar compuesto de múltiples buses. Además, el almacenamiento secundario 514 se puede acoplar directamente a los otros componentes del aparato 500 o se puede acceder a través de una red y puede comprender una sola unidad integrada tal como una tarjeta de memoria o múltiples unidades tal como múltiples tarjetas de memoria. Por lo tanto, el aparato 500 se puede implementar en una amplia variedad de configuraciones.

#### Refinamiento de vector de movimiento (MVR)

Los vectores de movimiento generalmente se determinan al menos parcialmente en el lado codificador y se señalan al decodificador dentro del flujo de bits codificado. Sin embargo, los vectores de movimiento también se pueden refinar en el decodificador (y también en el codificador) a partir de los vectores de movimiento iniciales indicados en el flujo de bits. En este caso, por ejemplo, se puede usar la similitud entre los parches de píxeles ya decodificados señalados por los vectores de movimiento iniciales para mejorar la precisión de los vectores de movimiento iniciales. Este refinamiento de movimiento proporciona una ventaja de reducir la sobrecarga de señalización: la precisión de los vectores de movimiento inicial se mejora de la misma manera tanto en el codificador como en el decodificador y, por lo tanto, no se necesita señalización adicional para el refinamiento.

Se observa que los vectores de movimiento iniciales antes del refinamiento podrían no ser los mejores vectores de movimiento que dan por resultado la mejor predicción. Dado que los vectores de movimiento iniciales se señalizan en el flujo de bits, podría no ser posible representar el vector de movimiento inicial con una precisión muy alta (que incrementaría la velocidad de bits), por lo tanto, el proceso de refinamiento de vector de movimiento se usa para mejorar la precisión del vector de movimiento inicial. Los vectores de movimiento iniciales podrían, por ejemplo, ser los vectores de movimiento que se utilizan en la predicción de un bloque vecino de un bloque actual. En este caso, es suficiente señalar una indicación en el flujo de bits, indicando vectores de movimiento de qué bloque vecino se usa por el bloque actual. Este mecanismo de predicción es muy eficiente para reducir el número de bits para representar los vectores de movimiento iniciales. Sin embargo, la precisión de los vectores de movimiento iniciales podría ser baja, ya que, en general, no se espera que los vectores de movimiento de dos bloques vecinos sean idénticos.

Con el fin de mejorar adicionalmente la precisión de los vectores de movimiento sin un incremento adicional en la sobrecarga de señalización, puede ser beneficioso refinar adicionalmente los vectores de movimiento derivados en el lado de codificador y proporcionados (señalizados) en el flujo de bits. El refinamiento del vector de movimiento se puede realizar en el decodificador sin asistencia del codificador. El codificador en su bucle de decodificador puede emplear el mismo refinamiento para obtener vectores de movimiento refinados correspondientes que estarían disponibles en el decodificador. El refinamiento para un bloque actual que se está reconstruyendo en una imagen actual se realiza al determinar una plantilla de muestras reconstruidas, determinar un espacio de búsqueda alrededor de la información de movimiento inicial para el bloque actual y encontrar en el espacio de búsqueda una porción de imagen de referencia que mejor coincida con la plantilla. La mejor porción coincidente determina los vectores de movimiento refinados para el bloque actual que luego se utilizan para obtener las muestras interpredichas para el bloque actual, es decir, el bloque actual que se reconstruye. El refinamiento de vector de movimiento es una parte de la Unidad (244) de Interpredicción en la Fig. 2 y 344 en la Fig. 3.

El refinamiento de vector de movimiento se puede realizar según los siguientes pasos. Típicamente, se pueden determinar vectores de movimiento iniciales con base en una indicación en el flujo de bits. Por ejemplo, un índice se puede señalar en el flujo de bits que indica una posición en una lista de vectores de movimiento candidatos. En otro ejemplo, se pueden señalar un índice de predictor de vector de movimiento y un valor de diferencia de vector de movimiento en el flujo de bits. Los vectores de movimiento que se determinan con base en una indicación en el flujo de bits se definen como vectores de movimiento iniciales. En el caso de bi-predicción, la interpredicción para el bloque actual se obtiene como una combinación ponderada del bloque predicho de muestras determinado según dos vectores de movimiento MV0 y MV1. Aquí, MV0 es el vector de movimiento inicial en una primera imagen de referencia en la lista L0 y MV1 es el vector de movimiento inicial en la segunda imagen de referencia en la lista L1.

Usando los vectores de movimiento inicial, se determinan los pares de vector de movimiento candidato de refinamiento (MV). Al menos, se deben determinar dos pares de candidatos de refinamiento. Típicamente, los pares de vectores de movimiento candidatos de refinamiento se determinan con base en el par de vectores de movimiento iniciales (MV0, MV1). Además, los pares de MV candidatos se determinan al añadir pequeñas diferencias de vector de movimiento a MV0 y MV1. Por ejemplo, los pares de MV candidatos pueden incluir los siguientes:

- (MV0, MV1)
- (MV0 + (0,1), MV1 + (0,-1))
- (MV0 + (1,0), MV1 + (-1,0))
- (MV0 + (0,-1), MV1 + (0,1))
- (MV0 + (-1,0), MV1 + (1,0))
- ...

Aquí, (1,-1) denota un vector que tiene un desplazamiento de 1 en la dirección horizontal (o x) y un desplazamiento de -1 en la dirección vertical (o y). Se observa que la lista anterior de pares de candidatos son solo ejemplos de explicación y la invención no se limita a una lista específica de candidatos. En algunos ejemplos, el espacio de búsqueda del proceso de refinamiento de vector de movimiento incluye los pares de vector de movimiento (MV) candidato de refinamiento.

En una bi-predicción de bloque actual, dos bloques de predicción obtenidos usando el primer vector de movimiento respectivo para la lista L0 y el segundo vector de movimiento para la lista L1, se combinan en una única señal de predicción, que puede proporcionar una mejor adaptación a la señal original que la

unipredicción, lo que resulta en menos información residual y posiblemente una compresión más eficiente.

En el refinamiento de vector de movimiento, los dos bloques de predicción obtenidos con el primer vector de movimiento respectivo y el segundo vector de movimiento de un par de MV candidatos se comparan con base en una métrica de similitud para cada uno de los pares de MV candidatos de refinamiento. Un par de MV candidatos que da por resultado la métrica de similitud más alta se selecciona como los vectores de movimiento refinados. El vector de movimiento refinado para una primera imagen de referencia en la lista L0 y el vector de movimiento refinado para una segunda imagen de referencia en la lista L1 se indican MV0' y MV1', respectivamente. En otras palabras, se obtienen predicciones correspondientes a vector de movimiento de lista L0 y vector de movimiento de lista L1 del par de vectores de movimiento candidatos, que luego se comparan con base en una métrica de similitud. El par de vectores de movimiento candidatos que tiene la mayor similitud asociada se selecciona como par de MV refinado.

Por lo general, la salida del proceso de refinamiento son MV refinados. Los MV refinados pueden ser los mismos que los MV iniciales o pueden ser diferentes con los MV iniciales, dependiendo de qué par de MV candidatos logre la mayor similitud, el par de MV candidatos formado por MV iniciales también se encuentra entre los candidatos de par de MV. En otras palabras, si el par de MV candidatos más altos que logran la mayor similitud está formado por los MV iniciales, los MV refinados y los MV iniciales son iguales entre sí.

En lugar de seleccionar la posición que maximiza una métrica de similitud, otro método es seleccionar una posición que minimice una métrica de disimilitud. La medida de comparación de disimilitud podría ser SAD (suma de diferencias absolutas), MRSAD (suma de diferencias absolutas de media eliminada), SSE (suma de errores cuadrados), etc. El SAD entre dos bloques de predicción se puede obtener usando un par de MV candidatos (CMV0, CMV1), el SAD se puede calcular de la siguiente manera:

$$SAD(CMV0, CMV1) = \sum_{x=0}^{nCbw-1} \sum_{y=0}^{nCbH-1} \text{abs}(\text{predMuestrasL0}[x][y] - \text{predMuestrasL1}[x][y]) \quad (12)$$

donde nCbH y nCbW son la altura y el ancho de los bloques de predicción; la función abs(a) especifica el valor absoluto del argumento a; y predMuestrasL0 y predSMuestrasL1 son muestras de bloques de predicción obtenidas según un par de MV candidatos que se denota por (CMV0, CMV1).

Alternativamente, las mediciones de comparación de disimilitud se pueden obtener evaluando solo un subconjunto de muestras en un bloque de predicción, con el fin de reducir la cantidad de cálculos. A continuación se presenta un ejemplo, donde las filas de muestras se incluyen alternativamente en el cálculo SAD (se evalúa cada segunda fila).

$$SAD(CMV0, CMV1) = \sum_{x=0}^{nCbw-1} \sum_{y=0}^{nCbH/2-1} \text{abs}(\text{predMuestrasL0}[x][2 * y] - \text{predMuestrasL1}[x][2 * y]) \quad (13)$$

Un ejemplo de refinamiento de vector de movimiento se explica en el documento JVET-M1001-v3, "Versatile Video Coding (Draft 4)" de JVET (de ITU-T SG 16 WP 3 e ISO/IEC JTC 1/SC 29/WG 11) que está disponible públicamente en <http://phenix.it-sudparis.eu/jvet/>. La sección "8.4.3 Decoder side motion vector refinement process" en el documento ejemplifica el refinamiento de vector de movimiento.

Con el fin de reducir los requisitos de memoria interna para el refinamiento, en algunas realizaciones, el proceso de refinamiento de vector de movimiento se puede realizar independientemente en bloques de muestras de luma. Los bloques de muestras de luma se pueden obtener dividiendo el bloque codificado de muestras que exceden un cierto ancho predeterminado o altura predeterminada en muestras de luma en subbloques de muestras que son menores o iguales al ancho predeterminado y la altura predeterminada en la muestra de luma. El par de MV refinados para cada subbloque dentro de un bloque codificado dividido puede ser diferente. Luego se realiza la interpredicción tanto para luma como para croma para cada subbloque utilizando el par de MV refinados de ese subbloque.

Denota el ancho y la altura máximos permitidos del subbloque como max\_sb\_anchura y max\_sb\_altura, respectivamente. Una unidad de codificación actual de tamaño cbAnchura x cbAltura que es elegible para aplicar MVR se particiona típicamente en una cantidad de subbloques numSbs cada uno de tamaño sbAnchura x sbAltura de la siguiente manera:

$$\text{numSbs} = \text{numSbX} * \text{numSbY} \quad (14)$$

$$\text{numSbX} = (\text{cbAnchura} > \text{max\_sb\_anchura}) ? (\text{cbAnchura} / \text{max\_sb\_anchura}) : 1 \quad (15)$$

$$\text{numSbY} = (\text{cbAltura} > \text{max\_sb\_altura}) ? (\text{cbAltura} / \text{max\_sb\_altura}) : 1 \quad (16)$$

$$\text{sbAnchura} = (\text{cbAnchura} > \text{max\_sb\_anchura}) ? \text{max\_sb\_anchura} : \text{cbAnchura} \quad (17)$$

$$\text{sbAltura} = (\text{cbAltura} > \text{max\_sb\_altura}) ? \text{max\_sb\_altura} : \text{cbAltura} \quad (18)$$

5 donde la expresión  $(x>y)? a : b$  devuelve valor  $a$  si  $x>y$  es verdadero y devuelve  $b$  si  $x>y$  es falso. Cada MV del par MV inicial puede tener una precisión de píxeles fraccionarios. En otras palabras, el MV puede indicar un desplazamiento entre un bloque actual de muestras y una región de referencia remuestreada. Este desplazamiento puede apuntar a una posición fraccionaria en las direcciones horizontal y vertical desde la cuadrícula entera de muestras de referencia reconstruidas.

10 Normalmente, se realiza una interpolación bidimensional de los valores de cuadrícula de muestra entera de referencia reconstruidos para obtener los valores de muestra en la ubicación de desplazamiento de muestra fraccionada. El proceso de obtención de muestras predichas de las imágenes de referencia reconstruidas usando un par de MV candidatos puede ser a través de uno de los siguientes métodos:

- Redondear la parte fraccionaria del par MV inicial a la ubicación entera más cercana y obtener los valores de cuadrícula entera de las imágenes de referencia reconstruidas.
- 15 • Realizar una interpolación bilineal separable de 2 derivaciones (por ejemplo, bilineal) para obtener los valores de muestra predichos con la precisión de píxeles fraccionados indicada por el par de MV iniciales.
- Realizar una interpolación separable de mayor derivaciones (por ejemplo, 8 o 6 derivaciones) para obtener los valores de muestra predichos con la precisión de píxeles fraccionados indicada por el par de MV iniciales .

20 En tanto que los pares de MV candidatos pueden tener un desplazamiento de subpíxeles arbitrario con respecto al par de MV iniciales, en algunas realizaciones, por simplicidad de búsqueda, los pares de MV candidatos se eligen con una distancia de píxeles enteros con respecto al par de MV iniciales. En estos casos, las muestras predichas a través de los pares de MV candidatos se pueden obtener mediante la realización de una predicción para un bloque de muestras alrededor del par de MV iniciales para cubrir todas las posiciones de refinamiento alrededor del par de MV iniciales.

25 En algunas realizaciones, después de que se han evaluado los valores de coste de disimilitud para los pares de MV candidatos que se encuentran a una distancia entera del par de MV iniciales, se pueden agregar y evaluar pares de MV candidatos adicionales en desplazamientos de subpíxeles desde la mejor posición de valor de coste. Las muestras predichas se obtienen para cada una de estas posiciones utilizando uno de los métodos descritos anteriormente y los costos de disimilitud se evalúan y comparan para obtener la posición de disimilitud más baja. En otras realizaciones, para evitar este proceso de predicción costoso en cálculo para cada posición de distancia de subpíxeles alrededor de la mejor posición de distancia entera de coste, se recuerdan los valores de costo de distancia entera evaluados y se encaja una superficie de errores paramétricos en las proximidades de la mejor posición de distancia entera. El mínimo de esta superficie de errores se calcula analíticamente y se utiliza como la posición con la mínima disimilitud. En estos casos, el valor de costo de disimilitud se deriva de los valores de coste de distancia entera calculados.

30 La aplicación de refinamiento de vector de movimiento para un bloque codificado de muestras dado se puede condicionar a ciertas propiedades de codificación del bloque codificado de muestras. Algunos ejemplos de estas propiedades de codificación incluyen la distancia en el número de imágenes (cuando se muestrean a una velocidad de fotograma uniforme) desde la imagen actual a las dos imágenes de referencia usadas para bi-predicción del bloque codificado de muestras son iguales y caen en lados opuestos de la imagen actual. Las propiedades de codificación también pueden incluir la disimilitud inicial entre los dos bloques predichos obtenidos usando el par de MV iniciales es menor que un umbral predeterminado por muestra.

35 En algunas implementaciones, se aplica BPOF a bloques bi-predichos cuando las dos predicciones son de imágenes de referencia diferentes. No se aplica BPOF a casos de modo de fusión temporal avanzada afín, bi-predictiva ponderada compensada en movimiento y basada en subbloques.

#### Refinamiento de flujo óptico bi-predictivo

40 El refinamiento de flujo óptico bi-predictivo es un proceso para mejorar la precisión de bi-predicción de un bloque sin proporcionar explícitamente señales adicionales en el flujo de bits que no sean las señales para bi-predicción. El refinamiento de flujo óptico bi-predictivo se puede implementar en la unidad de interpredicción 244 en la Fig. 2 y la unidad de interpredicción 344 en la Fig. 3. La entrada del proceso de refinamiento de flujo óptico son las muestras de predicción de dos imágenes de referencia y la salida del refinamiento de flujo óptico es la predicción combinada (predBIO) que se calcula según la ecuación de flujo óptico.

45 En la bi-predicción, dos interpredicciones se obtienen de dos fotogramas de referencia según dos vectores de movimiento, tal como el par de vectores de movimiento MV0 y MV1 o el par de vectores de movimiento refinado analizado anteriormente. Las dos predicciones se pueden combinar, tal como a través de promedios

ponderados. La predicción combinada puede dar como resultado una energía residual reducida debido a que el ruido de cuantificación en las dos predicciones se cancela, proporcionando así más eficiencia de codificación en comparación con la unipredicción (es decir, la predicción usando un vector de movimiento). En un ejemplo, la combinación ponderada en bi-predicción se puede realizar de la siguiente manera:

5  
 Bi-predicción = Predicción1 \* W1 + Predicción2 \* W2 + K, (19)

donde W1 y W2 son factores de ponderación que se pueden señalar en el flujo de bits o se pueden predefinir. K es un factor aditivo que también se puede señalar o predefinir. Como un ejemplo, la bi-predicción se podría obtener a través de

10  
 Bi-predicción = (Predicción1 + Predicción2)/2, (20)

donde W1 y W2 se establecen a ½ y K se establece a 0.

15  
 La precisión de la bi-predicción se puede mejorar mediante refinamiento de flujo óptico. Un flujo óptico es el patrón de movimiento aparente de objetos de imagen entre dos fotogramas producido por el movimiento del objeto o cámara. El proceso de refinamiento de flujo óptico mejora la precisión de la bi-predicción mediante la determinación del flujo óptico entre los dos fotogramas de referencia y el ajuste de la bi-predicción con base en el flujo óptico determinado.

20  
 Considerar un píxel  $I(x,y,t)$  en un primer fotograma con  $x$  e  $y$  correspondientes a coordenadas espaciales y  $t$  correspondiente a la dimensión de tiempo. Se mueve por una distancia  $(dx,dy)$  en el siguiente fotograma tomado después del tiempo  $dt$ . Suponiendo que los píxeles en los dos fotogramas son los mismos y la intensidad no cambia dentro del tiempo  $dt$ , la ecuación de flujo óptico se puede formular como:

25  
 $I(x,y,t) = I(x+ v_x, y+ v_y, t+dt)$  (21)

30  
 $I(x,y,t)$  especifica la intensidad (es decir, el valor de muestra) de un píxel en las coordenadas de  $(x,y,t)$ . Con base en una suposición de que el movimiento o desplazamiento del píxel es pequeño y otras suposiciones tales como términos de mayor orden en una expansión de serie de Taylor se pueden ignorar, las ecuaciones de flujo óptico se pueden escribir como:

$$\frac{\partial I}{\partial t} + v_x \frac{\partial I}{\partial x} + v_y \frac{\partial I}{\partial y} = 0, \quad (22)$$

donde  $\frac{\partial I}{\partial x}$  y  $\frac{\partial I}{\partial y}$  son los gradientes de muestra espaciales horizontales y verticales en la posición  $(x,y)$  y  $\frac{\partial I}{\partial t}$  es la derivada temporal parcial en  $(x,y)$ . En algunos ejemplos, los gradientes de muestra se pueden obtener mediante:

35  
 $\partial I(x, y, t) / \partial x = I(x + 1, y, t) - I(x - 1, y, t),$

40  
 $\partial I(x, y, t) / \partial y = I(x, y + 1, t) - I(x, y - 1, t).$

El refinamiento de flujo óptico utiliza el principio que se muestra en la Ecuación. (22) con el fin de mejorar la calidad de bi-predicción. En algunas implementaciones, el refinamiento de flujo óptico se realiza calculando gradientes de muestra  $\frac{\partial I}{\partial x}$  y  $\frac{\partial I}{\partial y}$ , calculando la diferencia entre la primera predicción y la segunda predicción ( $I^{(0)} - I^{(1)}$ ), y calculando el desplazamiento de píxeles o grupo de píxeles  $(v_x, v_y)$ . El desplazamiento se calcula de modo que se minimiza el error  $\Delta$  entre muestras en los dos fotogramas de referencia obtenidos utilizando la ecuación de flujo óptico. El error  $\Delta$  se define como:

45  
 $\Delta = (I^{(0)} - I^{(1)}) + v_x(\tau_0 \frac{\partial I^{(0)}}{\partial x} + \tau_1 \frac{\partial I^{(1)}}{\partial x}) + v_y(\tau_0 \frac{\partial I^{(0)}}{\partial y} + \tau_1 \frac{\partial I^{(1)}}{\partial y})$  (23)

donde  $I^{(0)}$  representa un valor de muestra en la primera predicción (por ejemplo, muestras de predicción en el primer fotograma de referencia en L0), e  $I^{(1)}$  es el valor de muestra en la segunda predicción (por ejemplo, muestras de predicción en el segundo fotograma de referencia en L1) que corresponde a  $I^{(0)}$ .  $v_x$  y  $v_y$  son los desplazamientos calculados en dirección  $-x$  y  $-y$ , respectivamente. Los gradientes de la muestra en el primer fotograma de referencia en direcciones  $-x$  e  $-y$ , respectivamente, son  $\partial I^{(0)} / \partial x$  y  $\partial I^{(0)} / \partial y$ .  $\partial I^{(1)} / \partial x$  y  $\partial I^{(1)} / \partial y$  son los gradientes de muestra en el segundo fotograma de referencia en direcciones  $-x$  y  $-y$ , respectivamente.  $\tau_1$  y  $\tau_0$  denotan las distancias del fotograma actual a los primeros y segundos fotogramas de referencia, respectivamente. La Fig. 7 ilustra la relación entre las diversas variables involucradas en la ecuación. (23).

55  
 Para determinar el desplazamiento  $(v_x, v_y)$  en la Ecuación. (23), se utiliza un parche de muestras alrededor de una posición dada  $(x,y)$  para resolver el problema de minimización mencionado anteriormente. Algunos enfoques minimizan la suma de errores cuadrados para diferentes píxeles en un parche de los fotogramas de referencia. Otros enfoques minimizan la suma de errores absolutos. Después de determinar el desplazamiento

$(v_x, v_y)$ , la predicción combinada en la posición dada  $(x,y)$  se puede determinar como:

$$pred_{BIO} = \frac{1}{2} \left( (I^{(0)} + I^{(1)} + \frac{v_x}{2} (\tau_1 \frac{\partial I^{(1)}}{\partial x} - \tau_0 \frac{\partial I^{(0)}}{\partial x}) + \frac{v_y}{2} (\tau_1 \frac{\partial I^{(1)}}{\partial y} - \tau_0 \frac{\partial I^{(0)}}{\partial y}) \right) \quad (24)$$

donde  $pred_{BIO}$  es la predicción modificada en la posición  $(x,y)$  que es la salida del proceso de refinamiento de flujo óptico.

- 5 A partir de esta ecuación, se puede determinar que el desplazamiento determinado con base en el BDOF es, asumiendo  $\tau_0$  e  $\tau_1$  igual a 1.

$$BDOF_{desplazamiento} = v_x \left( \frac{\partial I^{(1)}}{\partial x} - \frac{\partial I^{(0)}}{\partial x} \right) / 2 + v_y \left( \frac{\partial I^{(1)}}{\partial y} - \frac{\partial I^{(0)}}{\partial y} \right) / 2.$$

- 10 En algunas realizaciones, con el fin simplificar la complejidad de estimar el desplazamiento para cada píxel, el desplazamiento se estima para un grupo de píxeles. Por ejemplo, el desplazamiento se puede estimar para un bloque de 4x4 píxeles, tal como muestras de luma de 4x4, en lugar de un píxel individual. En estos ejemplos, para calcular la bi-predicción mejorada para un bloque de muestras de luma de 4x4, los desplazamientos se estiman usando valores de muestra cerca del bloque de muestras de luma de 4x4, tal como un bloque de muestras de luma de 8x8 con el bloque de 4x4 de muestras en su centro. Una unidad de codificación se divide en subbloques cuando el ancho o la altura de la unidad de codificación excede 16. En el límite de un subbloque, los gradientes de muestra se calculan usando valores de muestra de referencia de cuadrícula entera sin ninguna interpolación compensada en movimiento separable 2-D. Después de eso, las muestras extendidas y los gradientes de muestra fuera de las posiciones de subbloque se obtienen extendiendo los valores de gradiente de muestra y muestra más cercana desde el límite del subbloque.

La entrada del proceso de refinamiento de flujo óptico son las muestras de predicción de dos imágenes de referencia y la salida del refinamiento de flujo óptico es la predicción combinada ( $pred_{BIO}$ ) que se calcula según la ecuación de flujo óptico.

- 25 En una versión adoptada actualmente de BDOF, las siguientes ecuaciones se usan para calcular el flujo óptico  $(v_x, v_y)$  basado en los gradientes de muestra horizontales y verticales de un bloque de 6x6 de muestras centrado en un bloque actual de 4x4 de muestras.

$$s_1 = \sum_{i,j} \left( \frac{\partial I^{(0)}}{\partial x} + \frac{\partial I^{(1)}}{\partial x} \right) * \left( \frac{\partial I^{(0)}}{\partial x} + \frac{\partial I^{(1)}}{\partial x} \right) \quad (25)$$

$$s_2 = \sum_{i,j} \left( \frac{\partial I^{(0)}}{\partial y} + \frac{\partial I^{(1)}}{\partial y} \right) * \left( \frac{\partial I^{(0)}}{\partial y} + \frac{\partial I^{(1)}}{\partial y} \right) \quad (26)$$

$$s_3 = \sum_{i,j} \left( \frac{\partial I^{(0)}}{\partial x} + \frac{\partial I^{(1)}}{\partial x} \right) * (I^{(1)} - I^{(0)}) \quad (27)$$

$$s_4 = \sum_{i,j} \left( \frac{\partial I^{(0)}}{\partial y} + \frac{\partial I^{(1)}}{\partial y} \right) * (I^{(1)} - I^{(0)}) \quad (28)$$

$$s_5 = \sum_{i,j} \left( \frac{\partial I^{(0)}}{\partial x} + \frac{\partial I^{(1)}}{\partial x} \right) * \left( \frac{\partial I^{(0)}}{\partial y} + \frac{\partial I^{(1)}}{\partial y} \right) \quad (29)$$

$$v_x = -\frac{s_3}{s_1} \quad (30)$$

$$v_y = -\frac{(s_4 - (v_x/2) * s_5)}{s_2} \quad (31)$$

- 30 El desplazamiento de flujo óptico  $(v_x, v_y)$  también se conoce en el presente documento como “flujo óptico  $(v_x, v_y)$ ”. La operación de división requerida para calcular  $v_x$  y  $v_y$  se simplifica a expensas de la precisión al desplazar hacia la derecha el numerador con solo la posición del bit más significativo en el denominador. En ciertas otras técnicas anteriores, la división se reemplaza por una tabla de búsqueda de N bits que contiene los valores recíprocos con un desplazamiento variable que representa la posición del bit más significativo para

mejorar la precisión. Sin embargo, la tabla de búsqueda da como resultado una mayor memoria en el chip. Una tabla de búsqueda de N bits con una precisión de M bits para el recíproco requiere N\*M bits de SRAM.

5 Un ejemplo de refinamiento de vector de movimiento se explica en la sección 8.4.7.4 "Proceso de predicción de flujo óptico bidireccional" del documento JVET-M1001, Codificación de video versátil (Borrador 4).

10 Como se analizó anteriormente, un flujo óptico contiene dos componentes:  $v_x$  en la dirección horizontal y  $v_y$  en la dirección vertical. En comparación con el método mostrado en las ecuaciones (25)-(31), el cálculo presentado en el presente documento para los dos componentes:  $v_x$  en la dirección horizontal y  $v_y$  en la dirección vertical elimina las operaciones de multiplicación y reduce la profundidad de bits de los términos.

En particular, el flujo óptico se puede estimar como sigue:

$$s_1 = \sum_{i,j} \text{abs}(G_{x1}(i,j) + G_{x0}(i,j)) \quad (32)$$

$$s_2 = \sum_{i,j} \text{abs}(G_{y1}(i,j) + G_{y0}(i,j)) \quad (33)$$

$$s_3 = \sum_{i,j} (\text{signo}(G_{x1}(i,j) + G_{x0}(i,j))) * (I^{(1)}(i,j) - I^{(0)}(i,j)) \quad (34)$$

$$s_4 = \sum_{i,j} (\text{signo}(G_{y1}(i,j) + G_{y0}(i,j))) * (I^{(1)}(i,j) - I^{(0)}(i,j)) \quad (35)$$

$$s_5 = \sum_{i,j} (\text{signo}(G_{y1}(i,j) + G_{y0}(i,j))) * (G_{x1}(i,j) + G_{x0}(i,j)) \quad (36)$$

$$v_x = -\frac{s_2}{s_1} \quad (37)$$

$$v_y = -\frac{(s_4 - v_x * s_5/2)}{s_2} \quad (38)$$

15

donde  $G_{x0}(i,j) = \frac{\partial I^{(0)}}{\partial x}$  y  $G_{x1}(i,j) = \frac{\partial I^{(1)}}{\partial x}$  son gradientes de muestra predichos horizontales en el píxel  $(i,j)$  en el primer y segundo fotogramas de referencia, respectivamente, y  $G_{y0}(i,j) = \frac{\partial I^{(0)}}{\partial y}$  y  $G_{y1}(i,j) = \frac{\partial I^{(1)}}{\partial y}$  son gradientes de muestra predichos verticales en el píxel  $(i,j)$  en el primer y segundo fotogramas de referencia, respectivamente. Aquí,  $i$  y  $j$  son enteros y se extienden sobre un conjunto de posiciones de muestra centradas en un bloque actual de posiciones de muestra. En una modalidad, para un bloque de 4x4, se emplea un bloque de 6x6 de posiciones de muestra centradas en el bloque de 4x4. En un ejemplo, el valor de  $i$  varía de -1 a 4, y el valor de  $j$  varía de -1 a 4.

20

Se debe entender que durante el cálculo de los términos de autocorrelación y correlación cruzada  $s_1$ - $s_5$  mostrados anteriormente, uno o más términos se pueden desplazar para ajustar la precisión y profundidad de bits de los valores.

25

Se debe observar además que las ecuaciones (32)-(38) enumeradas anteriormente son solo para fines ilustrativos y no se deben interpretar como limitantes. Varios términos en estas ecuaciones se pueden preprocesar antes de combinarse con otros términos en estas ecuaciones. Por ejemplo, el término  $(G_{y1}(i,j) + G_{y0}(i,j))$  o  $G_{x1}(i,j) + G_{x0}(i,j)$  se puede desplazar, voltear al cambiar su signo o procesarse de otro modo antes de usarse como se muestra en las ecuaciones anteriores para calcular  $s_1$ - $s_5$ . De manera similar, el término  $(I^{(1)} - I^{(0)})$  también se puede preprocesar antes de combinarse con otros términos en las ecuaciones anteriores. Del mismo modo, varios valores determinados en las ecuaciones anteriores también se pueden postprocesar antes de usarse para calcular el valor de  $v_x$  y  $v_y$ . Por ejemplo,  $s_k$  ( $k=1$  a 5) determinado anteriormente se puede postprocesar al añadir los bits inferiores de su valor a los bits superiores de su valor para determinar la versión final de  $s_k$ . Esta versión final luego se puede utilizar como se muestra arriba para determinar  $v_x$  y  $v_y$ .

30

35

Como se puede ver de esta implementación, se determina una nueva cantidad  $s_5$  para facilitar el cálculo del segundo componente del flujo óptico,  $v_y$ , en este ejemplo.  $s_5$  se determina con base en la suma del producto del signo de la suma de los gradientes de muestra verticales predichos en los dos fotogramas de referencia y la suma de los gradientes de muestra horizontales predichos en los dos fotogramas de referencia. El cálculo de  $s_5$  se puede realizar sin operaciones de multiplicación. Por ejemplo, el cálculo se puede realizar al sumar o restar condicionalmente la suma de gradientes de muestra predichos horizontales en una posición de muestra

40

45

dada con base en el signo de la suma de los gradientes de muestra predichos verticales en esa posición de muestra. Luego el componente vertical de flujo óptico,  $v_y$ , se modifica con base en  $v_x, s_5$ , y  $s_2$ . En algunas implementaciones, la división por  $s_1$  o  $s_2$  en las ecuaciones (37)-(38) se puede simplificar mediante la aplicación de un desplazamiento a la derecha igual a la posición de la posición de bit más significativa en  $s_1$  y  $s_2$  respectivamente. Como resultado, el cálculo de los componentes  $s_1$ - $s_5$ , especialmente  $s_5$ , para el flujo óptico como se especifica en las ecuaciones (32)-(44) ha reducido la profundidad de bits porque el cálculo se basa en la operación de signo. En particular, la profundidad de bits de  $v_x$  se reduce porque el cálculo implica solo operación absoluta y operación de signo, de modo que en el cálculo de  $v_y$ , se puede reducir la profundidad de bits del multiplicador para  $v_x \times s_5$ . Por lo tanto, la complejidad del cálculo para  $v_x$  y  $v_y$  se reduce significativamente.

Como se puede ver en lo anterior, la BDOF requiere mucho menos cálculo, especialmente en términos de número de multiplicaciones y el tamaño del multiplicador. En algunos ejemplos, la BDOF se usa para refinar la señal de bi-predicción de una CU a nivel de subbloque de 4x4 y solo se aplica al componente de luma. El modo BDOF se basa en el concepto de flujo óptico, que asume que el movimiento de un objeto es suave. Para cada subbloque de 4 x4, se calcula un refinamiento de movimiento o desplazamiento de movimiento ( $v_x, v_y$ ) minimizando la diferencia entre las muestras de predicción L0 y L1. El refinamiento de movimiento se usa luego para ajustar los valores de muestra bi-predichos en el subbloque 4x4.

Como se analizó anteriormente, los gradientes horizontales y verticales,  $\frac{\partial I^{(k)}}{\partial x}(i, j)$  y  $\frac{\partial I^{(k)}}{\partial y}(i, j)$ ,  $k = 0, 1$ , se pueden calcular calculando la diferencia entre dos muestras vecinas en los fotogramas de referencia correspondientes. Las muestras se pueden desplazar con base en la profundidad de bits de luma antes de calcular la diferencia. La autocorrelación y correlación cruzada de los gradientes,  $s_1, s_2, s_3, s_4$  y  $s_5$ , se calculan para una ventana de 6x6 alrededor del subbloque de 4x4. La Fig. 6 muestra la relación entre la ventana de 6x6 y el subbloque de 4x4. Como se puede ver, con el fin de derivar los valores de gradiente utilizados en  $s_1, s_2, s_3, s_4$  y  $s_5$ , se deben generar algunas muestras de predicción  $I^{(k)}(i, j)$  en la lista  $k$  ( $k = 0, 1$ ) fuera de los límites de CU actuales (posiciones grises). En el ejemplo mostrado en la Fig. 6, la BDOF usa una fila/columna extendida alrededor de los límites de la CU. Estos valores de muestra extendida se utilizan solo en el cálculo de gradiente. Para los pasos restantes en el proceso BDOF, si se necesitan valores de muestra y gradiente fuera de los límites de la CU, se rellenan (es decir, se repiten) desde sus vecinos más cercanos.

El refinamiento de movimiento ( $v_x, v_y$ ) luego se deriva usando los términos de correlación cruzada y autocorrelación usando lo siguiente:

$$v_x = S_1 > 0? \text{clip3} \left( -th'_{BIO}, th'_{BIO}, -((S_3 \cdot 2^{n_b - n_a}) \gg \lfloor \log_2 S_1 \rfloor) \right) : 0 \quad (3-28)$$

$$v_y = S_5 > 0? \text{clip3} \left( -th'_{BIO}, th'_{BIO}, -\left( (S_6 \cdot 2^{n_b - n_a} - ((v_x S_{2,m}) \ll n_{S_2} + v_x S_{2,s}) / 2 \right) \gg \lfloor \log_2 S_5 \rfloor \right) : 0$$

donde  $S_{2,m} = S_2 \gg n_{S_2}, S_{2,s} = S_2 \& (2^{n_{S_2}} - 1), th'_{BIO} = 2^{\max(5, BD-7)}$ .  $\lfloor \cdot \rfloor$  es la función de piso, y  $n_{S_2} = 12$ .

Con base en el refinamiento de movimiento y los gradientes, se calcula el siguiente ajuste para cada muestra en el subbloque de 4 x4:

$$b(x, y) = \text{rnd} \left( \left( v_x \left( \frac{\partial I^{(1)}(x, y)}{\partial x} - \frac{\partial I^{(0)}(x, y)}{\partial x} \right) + v_y \left( \frac{\partial I^{(1)}(x, y)}{\partial y} - \frac{\partial I^{(0)}(x, y)}{\partial y} \right) + 1 \right) / 2 \right) \quad (3-29)$$

Finalmente, las muestras BDOF de la CU se calculan ajustando las muestras de bi-predicción de la siguiente manera:

$$\text{pred}_{BDOF}(x, y) = (I^{(0)}(x, y) + I^{(1)}(x, y) + b(x, y) + o_{\text{desplazamiento}}) \gg \text{desplazamiento} \quad (3-30)$$

La Fig. 8 ilustra un ejemplo de un proceso 800 para realizar refinamiento de flujo óptico bi-predictivo con base en el cálculo de flujo óptico presentado en el presente documento. Uno o más dispositivos informáticos (por ejemplo, el aparato de codificación 200 o el aparato de decodificación 300) implementan operaciones representadas en la Fig. 8 mediante la ejecución de código de programa adecuado.

El bloque 810 corresponde al primer paso descrito anteriormente. En este bloque, se obtienen dos vectores de movimiento como entrada. Los vectores de movimiento iniciales se pueden determinar con base en una información de indicación en el flujo de bits. Por ejemplo, un índice se puede señalar en el flujo de bits, el índice indica una posición en una lista de vectores de movimiento candidatos. En otro ejemplo, se pueden señalar un índice de predictor de vector de movimiento y un valor de diferencia de vector de movimiento en el flujo de bits. En otro ejemplo, estos vectores de movimiento se pueden derivar como un vector de movimiento de refinamiento mediante el uso de refinamiento de vector de movimiento a partir de un par inicial de vectores de movimiento que se indican en el flujo de bits. En otro ejemplo, las indicaciones de fotograma de referencia

se pueden obtener del flujo de bits para indicar el fotograma de referencia con el que se asocia un vector de movimiento dado en el par de vectores de movimiento obtenido. Por ejemplo, la indicación puede especificar que un fotograma de una primera lista L0 de fotogramas de referencia está asociado con el vector de movimiento MV0 en el par de vectores de movimiento y un fotograma de una segunda lista L1 de fotogramas de referencia está asociado con el vector de movimiento MV1 en el par de vectores de movimiento.

El bloque 820 corresponde al segundo paso descrito anteriormente. En este bloque, se puede obtener una predicción unidireccional en cada uno de los dos fotogramas de referencia (es decir, muestras de luma reconstruidas) según el par de vectores de movimiento obtenido y un filtro de interpolación de K derivaciones. Por ejemplo, la predicción obtiene valores de muestra de referencia reconstruidos cuando el vector de movimiento corresponde a una posición de muestra entera. Si el vector de movimiento tiene un componente horizontal distinto de cero, pero un componente vertical cero, realiza una interpolación de K derivaciones horizontal para obtener los valores de muestra predichos. Si el vector de movimiento tiene un componente vertical distinto de cero, pero un componente horizontal cero, realiza una interpolación de K derivaciones vertical para obtener los valores de muestra predichos. Si el vector de movimiento tiene valores distintos de cero para los componentes horizontales y verticales, se realiza una interpolación de K derivaciones separable 2-D con la interpolación horizontal realizada primero seguida de la interpolación vertical para obtener los valores de muestra predichos. De esta manera, se genera una primera predicción usando MV0 en un fotograma de referencia de la primera lista de fotogramas de referencia L0 y se genera una segunda predicción usando MV1 en un fotograma de referencia de la segunda lista L1 de fotogramas de referencia.

El bloque 830 corresponde al tercer paso descrito anteriormente. En este bloque, se estima un flujo óptico para cada subbloque en una unidad de codificación actual dada usando las predicciones obtenidas en el segundo paso en los fotogramas de referencia. Según las anotaciones analizadas anteriormente, las muestras de predicción en la predicción obtenida en el primer fotograma de referencia se denotan como  $I^{(0)}$  y las muestras de predicción en la predicción obtenida en el segundo fotograma de referencia se denotan como  $I^{(1)}$ . El gradiente de muestra horizontal en una posición  $(i, j)$  se puede calcular al tomar la diferencia entre el valor de muestra predicho a la derecha de esta posición y el valor de muestra a la izquierda de esta posición, es decir,  $\partial I / \partial x = I(x + 1, y) - I(x - 1, y)$ . El gradiente de muestra vertical en una posición  $(i, j)$  se puede calcular al tomar la diferencia entre el valor de muestra predicho por debajo de esta posición y el valor de muestra por encima de esta posición, es decir  $\partial I / \partial y = I(x, y+1) - I(x, y-1)$ . Se señala que para una imagen o un fotograma, la dirección horizontal apunta de izquierda a derecha, y la dirección vertical apunta de arriba a abajo. En algunos ejemplos,  $\partial I^{(0)} / \partial x$ ,  $\partial I^{(0)} / \partial y$  y  $\partial I^{(1)} / \partial x$ ,  $\partial I^{(1)} / \partial y$  se calculan para un conjunto de posiciones dentro del subbloque de codificación actual. Con base en los gradientes de muestra determinados, el flujo óptico se puede determinar usando el método descrito anteriormente con las ecuaciones respectivas (31)-(40) o el método de estimación de flujo óptico iterativo descrito anteriormente con las ecuaciones respectivas (38)-(46) o las ecuaciones (47)-(52).

El bloque 840 corresponde al cuarto paso descrito anteriormente. En este bloque, las muestras finales inter-bipredichas para el bloque de codificación actual se pueden calcular según la ecuación (24) que teniendo en cuenta los valores de muestra predichos, los gradientes de muestra determinados y el flujo óptico estimado.

La Fig. 9 es un diagrama de flujo de un método de ejemplo 900 para interpredicción basada en flujo óptico bidireccional (BDOF) para un bloque actual de una señal de video.

En el paso 910, el método 900 implica determinar un desplazamiento de movimiento horizontal,  $v_x$  y un desplazamiento de movimiento vertical,  $v_y$ , del bloque actual, y en donde el desplazamiento de movimiento vertical se determina con base en el desplazamiento de movimiento horizontal y una quinta variable  $s_5$ . La quinta variable  $s_5$  indica una suma de una pluralidad de términos. Cada uno de la pluralidad de términos se obtiene de un signo de un elemento de una segunda matriz y un elemento de una primera matriz. El elemento de la primera matriz corresponde al elemento de la segunda matriz.

Cada elemento de la primera matriz se obtiene a partir de una suma de un primer gradiente de muestra predicho horizontal que corresponde a un primer fotograma de referencia del bloque actual y un segundo gradiente de muestra predicho horizontal que corresponde a un segundo fotograma de referencia del bloque actual. El primer gradiente de muestra predicho horizontal y el segundo gradiente de muestra predicho horizontal corresponden al elemento de la primera matriz. Cada elemento de la segunda matriz se obtiene a partir de una suma de un primer gradiente de muestra predicho vertical que corresponde al primer fotograma de referencia del bloque actual y un segundo gradiente de muestra predicho vertical que corresponde al segundo fotograma de referencia del bloque actual. El primer gradiente de muestra predicho vertical y el segundo gradiente de muestra predicho vertical corresponden al elemento de la segunda matriz.

En el paso 920, el proceso 900 implica determinar valores de muestra de predicción en el bloque actual usando valores de muestra de predicción correspondientes al primer fotograma de referencia, valores de muestra de predicción correspondientes al segundo fotograma de referencia, y el desplazamiento de movimiento horizontal

y el desplazamiento de movimiento vertical.

5 Se observa que el bloque actual puede ser un bloque de cualquier tamaño, tal como un bloque de 4x4. El bloque actual puede ser un subbloque de un fotograma de la señal de video. Los píxeles del bloque actual se pueden referir usando posiciones absolutas de los píxeles con respecto a la esquina superior izquierda del fotograma, por ejemplo, (x,y), o posiciones relativas de los píxeles con respecto a la esquina superior izquierda del bloque, por ejemplo, (xBloque + i, yBloque+j). Aquí, (xBloque, yBloque) es la coordenada de la esquina superior izquierda del bloque con respecto a la esquina superior izquierda del fotograma.

10 La primera matriz y la segunda matriz pueden ser cualquier matriz bidimensional que incluya filas y columnas y se puede hacer referencia a un elemento de la matriz usando (i,j) donde x es un índice horizontal/de fila e y es un índice vertical/de columna. El intervalo de i y j puede ser, por ejemplo, i = xBloque -1, ..., xBloque + 4 y j = yBloque -1, ..., yBloque + 4. La primera y segunda matrices corresponden a, o se determinan para el bloque actual. En algunos ejemplos, el tamaño de la primera matriz es el mismo que el tamaño de la segunda matriz  
15 que puede ser mayor que el tamaño del bloque actual. Por ejemplo, el tamaño de la primera matriz y la segunda matriz puede ser 6x6, mientras que el bloque actual tiene un tamaño de 4x4.

20 Un elemento de la primera matriz (primer elemento) corresponde a un elemento de la segunda matriz (segundo elemento) si la posición del primer elemento en la primera matriz, (x,y) es la misma que la posición del segundo elemento en la segunda matriz, (p,q), es decir, (x,y)=(p,q). El primer gradiente de muestra predicho horizontal corresponde a un primer fotograma de referencia del bloque actual, lo que significa que el primer gradiente de muestra predicho horizontal se genera con base en las muestras en el primer fotograma de referencia del bloque actual. El segundo gradiente de muestra predicho horizontal corresponde al segundo fotograma de referencia del bloque actual, lo que significa que el segundo gradiente de muestra predicho horizontal se genera con base en muestras en el segundo fotograma de referencia del bloque actual. El primer gradiente de muestra predicho horizontal corresponde al elemento de la primera matriz, lo que significa que el primer gradiente de muestra predicho horizontal se genera para la posición (x,y) del elemento en la primera matriz. Asimismo, el segundo gradiente de muestra predicho horizontal corresponde al elemento de la primera matriz, lo que significa que el segundo gradiente de muestra predicho horizontal se genera para la posición (x,y) del elemento  
25 en la primera matriz.  
30

35 El primer gradiente de muestra predicho vertical corresponde al primer fotograma de referencia del bloque actual, lo que significa que el primer gradiente de muestra predicho vertical se genera con base en las muestras en el primer fotograma de referencia del bloque actual. El segundo gradiente de muestra predicho vertical corresponde al segundo fotograma de referencia del bloque actual, lo que significa que el segundo gradiente de muestra predicho vertical se genera con base en muestras en el segundo fotograma de referencia del bloque actual. El primer gradiente de muestra predicho vertical corresponde al elemento de la segunda matriz, lo que significa que el primer gradiente de muestra predicho vertical se genera para la posición (p,q) del elemento en la segunda matriz. Del mismo modo, el segundo gradiente de muestra predicho vertical corresponde al elemento de la segunda matriz, lo que significa que el segundo gradiente de muestra predicho vertical se genera para la posición (p,q) del elemento en la segunda matriz.  
40

45 Cada elemento de una matriz se obtiene a partir de una suma de dos términos, lo que significa que el elemento se puede determinar como la suma de dos términos en sí o como un valor después de procesar la suma de dos términos. El procesamiento puede incluir desplazamiento a la izquierda, desplazamiento a la derecha, recorte, o una combinación de los mismos. De manera similar, un término se obtiene de un signo de un elemento de una segunda matriz y un elemento de una primera matriz, lo que significa que el término puede tomar el valor del elemento de la primera matriz por sí mismo o después de que el elemento de la primera matriz se procesa y aplicar el signo del elemento de la segunda matriz. El procesamiento del elemento de la primera matriz puede incluir desplazamiento hacia la izquierda, desplazamiento hacia la derecha, recorte, o una combinación de los mismos. El signo del elemento x se puede determinar como  
50

$$\text{Signo}(x) = \begin{cases} 1 & ; & x > 0 \\ 0 & ; & x == 0 \\ -1 & ; & x < 0 \end{cases}$$

55 La tecnología presentada en el presente documento ajusta los valores de muestra bi-predichos del bloque actual con base en el desplazamiento de movimiento horizontal y el desplazamiento de movimiento vertical. El desplazamiento de movimiento vertical se calcula con base en una quinta variable  $s_5$  que solo implica la suma de términos obtenidos de un signo de un elemento de una segunda matriz y un elemento de una primera matriz. La aplicación de un signo de un elemento a otro elemento no implica operaciones de multiplicación. Del mismo modo, la suma tampoco implica multiplicaciones. Como resultado, la tecnología de interpredicción basada en BDOF presentada en el presente documento elimina las operaciones de multiplicación. En comparación con  
60

los enfoques tradicionales, la profundidad de bits de la quinta variable  $s_5$  se reduce porque las operaciones de multiplicación se reemplazan por la determinación de signo. Esto conduce a la reducción de la profundidad de bits del desplazamiento de movimiento horizontal,  $v_x$  y el desplazamiento de movimiento vertical,  $v_y$ , y también una reducción significativa en la complejidad de cálculo de la predicción y el tamaño del multiplicador.

5

La Fig. 10 ilustra un dispositivo 1000 para interpredicción basada en flujo óptico bidireccional (BDOF) para un bloque actual de una señal de video. El dispositivo 1000 incluye:

10

una unidad 1001 de determinación configurada para determinar un desplazamiento de movimiento horizontal,  $v_x$  y un desplazamiento de movimiento vertical,  $v_y$  del bloque actual, y en donde el desplazamiento de movimiento vertical se determina con base en el desplazamiento de movimiento horizontal y una quinta variable  $s_5$ . La quinta variable  $s_5$  indica una suma de una pluralidad de términos, en donde cada uno de la pluralidad de términos se obtiene de un signo de un elemento de una segunda matriz y un elemento de una primera matriz, y el elemento de la primera matriz corresponde al elemento de la segunda matriz; y

15

una unidad 1003 de procesamiento de predicción configurada para predecir valores de muestra de predicción en el bloque actual usando valores de muestra de predicción correspondientes al primer fotograma de referencia, valores de muestra de predicción correspondientes al segundo fotograma de referencia, y el desplazamiento de movimiento horizontal y el desplazamiento de movimiento vertical.

20

La unidad 1001 de determinación se configura además para obtener cada elemento de la primera matriz a partir de una suma de un primer gradiente de muestra predicho horizontal que corresponde a un primer fotograma de referencia del bloque actual y un segundo gradiente de muestra predicho horizontal que corresponde a un segundo fotograma de referencia del bloque actual. El primer gradiente de muestra predicho horizontal y el segundo gradiente de muestra predicho horizontal corresponden al elemento de la primera matriz.

25

La unidad 1001 de determinación se configura además para obtener cada elemento de la segunda matriz a partir de una suma de un primer gradiente de muestra predicho vertical que corresponde al primer fotograma de referencia del bloque actual y un segundo gradiente de muestra predicho vertical que corresponde al segundo fotograma de referencia del bloque actual. El primer gradiente de muestra predicho vertical y el segundo gradiente de muestra predicho vertical corresponden al elemento de la segunda matriz.

30

Correspondientemente, en un ejemplo, una estructura de ejemplo del dispositivo 1000 puede corresponder al codificador 200 en la Fig. 2. En otro ejemplo, una estructura de ejemplo del dispositivo 1000 puede corresponder al decodificador 300 en la Fig. 3. En otro ejemplo, una estructura de ejemplo del dispositivo 1000 puede corresponder a la unidad 244 de interpredicción en la Fig. 2. En otro ejemplo, una estructura de ejemplo del dispositivo 1000 puede corresponder a la unidad 344 de interpredicción en la Fig. 3.

35

La tecnología presentada en el presente documento para calcular el flujo óptico y las muestras bi-predichas mejora la eficiencia de codificación mediante el cálculo de un segundo componente del flujo óptico con base en un primer componente del flujo óptico calculado independientemente. La complejidad de cálculo se mantiene baja porque el cálculo dependiente tampoco requiere ninguna operación de multiplicación. La suma del producto del signo de la suma de gradientes en la segunda dirección y la suma de gradientes en la primera dirección se puede realizar sin ninguna multiplicación al sumar o restar condicionalmente la suma de gradientes en la primera dirección con base en el signo de la suma de gradientes en la segunda dirección. La tecnología presentada en el presente documento también logra una eficiencia de compresión similar a los métodos que usan operaciones de multiplicación.

45

50

El presente documento descripción proporciona los siguientes aspectos adicionales.

Según un primer aspecto, un método para interpredicción basada en flujo óptico, comprende:

55

- determinar un flujo óptico para un bloque de codificación actual, en donde un segundo componente del flujo óptico se determina o deriva con base en un primer componente del flujo óptico (tal como,  $v_y$  se basa en  $v_x$  o  $v_x$  se basa en  $v_y$  en inter-bipredicción basada en flujo óptico bi-predictivo) mediante una primera formulación; y

60

- obtener o derivar valores de muestra de predicción (tal como valores de muestra bi-predichos) para el subbloque actual usando el flujo óptico determinado para el bloque de codificación actual.

En una forma de implementación posible del método según cualquier implementación anterior del primer aspecto o el primer aspecto como tal, el paso de determinar un flujo óptico para un bloque de codificación actual comprende:

65

## ES 2 984 615 T3

calcular un flujo óptico para un bloque de codificación actual, en donde un segundo componente de flujo óptico se calcula usando

- 5 - un primer componente de flujo óptico calculado,
- signo y valor absoluto de una suma de gradientes de muestra predichos correspondientes (tales como posiciones de muestra correspondientes en los dos bloques predichos) a través de dos fotogramas de referencia en una dirección correspondiente al segundo componente, y
- 10 - la suma de los gradientes de muestra predichos correspondientes en los dos fotogramas de referencia en una dirección correspondiente al primer componente.

En una forma de implementación posible del método según cualquier implementación anterior del primer aspecto o el primer aspecto como tal, el paso de obtener o derivar valores de muestra de predicción (tal como valores de muestra bi-predichos) para el subbloque actual usando el flujo óptico determinado para el bloque de codificación actual, comprende:

20 obtener valores de muestra bi-predichos para el bloque de codificación actual usando un conjunto de valores de muestra predichos, el flujo óptico calculado y gradientes de muestra horizontales y verticales en dos fotogramas de referencia, en donde el conjunto de valores de muestra predichos se obtiene en cada uno de los dos fotogramas de referencia usando un par de vectores de movimiento para el bloque de codificación actual con respecto a los dos fotogramas de referencia.

25 En una forma de implementación posible del método según cualquier implementación anterior del primer aspecto o el primer aspecto como tal, el método comprende además:

calcular un flujo óptico para un bloque de codificación actual, en donde un segundo componente de flujo óptico se calcula usando

- 30 - un primer componente de flujo óptico calculado,
- signo y valor absoluto de una suma de gradientes de muestra predichos correspondientes a través de dos fotogramas de referencia en una dirección correspondiente al segundo componente, y
- 35 - la suma de los gradientes de muestra predichos correspondientes en los dos fotogramas de referencia en una dirección correspondiente al primer componente;

40 - obtener valores de muestra bi-predichos para el bloque de codificación actual usando un conjunto de valores de muestra predichos, el flujo óptico calculado y gradientes de muestra horizontales y verticales en los dos fotogramas de referencia, en donde el conjunto de valores de muestra predichos se obtiene en cada uno de los dos fotogramas de referencia usando un par de vectores de movimiento para el bloque de codificación actual con respecto a los dos fotogramas de referencia.

45 Según un segundo aspecto, un método para interpredicción basada en flujo óptico comprende:

- obtener un par de vectores de movimiento para un bloque de codificación actual con respecto a dos fotogramas de referencia;
- 50 - obtener un conjunto de muestras predichas en cada fotograma de referencia usando el par de vectores de movimiento obtenido y los valores de muestra de luma reconstruidos de los dos fotogramas de referencia;

55 - calcular un flujo óptico para el bloque de codificación actual usando las primeras diferencias de muestra predichas entre las muestras correspondientes en los dos fotogramas de referencia, y la suma de gradientes de muestra horizontales (sGx) y gradientes de muestra verticales (sGy) correspondientes en los dos fotogramas de referencia; en donde un segundo componente de flujo óptico se calcula usando

- un primer componente de flujo óptico calculado,
- 60 - signo y valor absoluto de suma de gradientes de muestra predichos correspondientes a través de las dos referencias en la dirección correspondiente al segundo componente, y
- la suma de gradientes de muestra predichos correspondientes en las dos referencias en la dirección correspondiente al primer componente;

65 - obtener los valores de muestra bi-predichos para el bloque de codificación actual usando los primeros valores de muestra predichos, el flujo óptico calculado y los gradientes de muestra horizontales y verticales en las dos

referencias con base en la ecuación de predicción para BDOF (flujo óptico bidireccional).

5 En una forma de implementación posible del método según cualquier implementación anterior del segundo aspecto o el segundo aspecto como tal, los valores de muestra de luma de referencia reconstruidos de los dos fotogramas de referencia comprenden los valores de muestra de luma vecinos reconstruidos de los dos fotogramas de referencia.

10 En una forma de implementación posible del método según cualquier implementación anterior del segundo aspecto o el segundo aspecto como tal, se calcula un flujo óptico según una ecuación de flujo óptico, de la siguiente manera:

$$\Delta = (I^{(0)} - I^{(1)}) + v_x(\tau_0 \frac{\partial I^{(0)}}{\partial x} + \tau_1 \frac{\partial I^{(1)}}{\partial x}) + v_y(\tau_0 \frac{\partial I^{(0)}}{\partial y} + \tau_1 \frac{\partial I^{(1)}}{\partial y})$$

15 donde  $I^{(0)}$  corresponde al valor de la muestra en una primera predicción,  $I^{(1)}$  es el valor de la muestra en una segunda predicción,  $v_x$  e  $v_y$  son los desplazamientos calculados en dirección  $-x$  y  $-y$ , y  $\partial I^{(0)}/\partial x$  y  $\partial I^{(0)}/\partial y$  son los gradientes en direcciones  $x$  e  $y$ ,  $\tau_1$  y  $\tau_0$  denotan las distancias a las imágenes de referencia, donde se obtienen la primera predicción y la segunda predicción.

20 En una forma de implementación posible del método según cualquier implementación anterior del segundo aspecto o el segundo aspecto como tal, el método se usa para predicción bidireccional;

correspondientemente, el par de vectores de movimiento incluye un primer vector de movimiento que corresponde a una primera lista de fotogramas de referencia y un segundo vector de movimiento que

25 corresponde a una segunda lista de fotogramas de referencia; el conjunto obtenido de muestras predichas comprende un primer conjunto de muestras predichas obtenidas según el primer vector de movimiento y un segundo conjunto de muestras predichas obtenidas según el segundo vector de movimiento;

30 los gradientes de muestra horizontales y verticales comprenden un primer conjunto de gradientes de muestra horizontales y verticales calculados usando el primer conjunto de muestras predichas, y un segundo conjunto de gradientes de muestra horizontales y verticales calculados usando el segundo conjunto de muestras predichas;

35 se obtiene un desplazamiento de movimiento con base en el primer y segundo conjunto de gradientes horizontales y verticales, y el primer y segundo conjunto de muestras predichas; y los valores de muestra de predicción para el subbloque actual se obtienen usando el desplazamiento de movimiento.

40 En una forma de implementación posible del método según cualquier implementación anterior del segundo aspecto o el segundo aspecto como tal, se determina o se deriva un segundo componente del flujo óptico (tal como  $v_y$ ) con base en un primer componente (tal como  $v_x$ ) del flujo óptico y uno o más de un primer valor, un segundo valor, un tercer valor, un cuarto valor y un quinto valor;

45 el primer componente (tal como  $v_x$ ) del flujo óptico se determina o deriva con base en uno o más del primer valor, el segundo valor, el tercer valor, el cuarto valor y el quinto valor.

50 En una forma de implementación posible del método según cualquier implementación anterior del segundo aspecto o el segundo aspecto como tal,  $v_y$  se basa en  $v_x$  en inter-bipredicción basada en flujo óptico bi-predictivo, de la siguiente manera:

$$v_x = -\frac{s_3}{s_1}$$

$$v_y = -\frac{(s_4 - v_x * s_5/2)}{s_2}$$

55 En una forma de implementación posible del método según cualquier implementación anterior del segundo aspecto o el segundo aspecto como tal,

$$s_1 = \sum_{i,j} abs(Gx1 + Gx0)$$

$$s_2 = \sum_{i,j} \text{abs}(Gy1 + Gy0)$$

$$s_3 = \sum_{i,j} f(Gx1 + Gx0) * (I^{(1)} - I^{(0)})$$

5

$$s_4 = \sum_{i,j} \text{signo}(Gy1 + Gy0) * (I^{(1)} - I^{(0)})$$

$$s_5 = \sum_{i,j} \text{signo}(Gy1 + Gy0) * (Gx1 + Gx0)$$

10 Según un cuarto aspecto, un codificador (20) comprende circuitería de procesamiento para llevar a cabo el método según cualquiera del primer y segundo aspectos.

Según un quinto aspecto, un decodificador (30) comprende circuitería de procesamiento para llevar a cabo el método según cualquiera del primer y segundo aspectos.

15

Según un sexto aspecto, un producto de programa de ordenador comprende un código de programa para realizar el método según cualquiera del primer y segundo aspectos.

20 Según un séptimo aspecto, un medio legible por ordenador no transitorio que porta un código de programa que, cuando se ejecuta por un dispositivo informático, hace que el dispositivo informático realice el método de uno cualquiera del primer y segundo aspectos.

Según un octavo aspecto, un decodificador comprende:

25 uno o más procesadores; y

un medio de almacenamiento legible por ordenador no transitorio acoplado a los procesadores y que almacena programación para ejecución por los procesadores, en donde la programación, cuando se ejecuta por los procesadores, configura el decodificador para llevar a cabo el método según cualquiera del primer y segundo aspectos.

30

Según un noveno aspecto, un codificador, comprende:

uno o más procesadores; y

35

un medio de almacenamiento legible por ordenador no transitorio acoplado a los procesadores y que almacena programación para ejecución por los procesadores, en donde la programación, cuando se ejecuta por los procesadores, configura el codificador para llevar a cabo el método según uno cualquiera del primer y el segundo aspectos.

40

Según un décimo aspecto, un aparato para interpredicción basada en flujo óptico, que comprende:

- una unidad de determinación, configurada para determinar un flujo óptico para un bloque de codificación actual, en donde un segundo componente del flujo óptico se determina o deriva con base en un primer componente del flujo óptico (tal como, en inter-bipredicción basada en flujo óptico bi-predictivo); y

45

- una unidad de obtención, configurada para obtener o derivar valores de muestra de predicción (tal como valores de muestra bi-predichos) para el subbloque actual usando el flujo óptico determinado para el bloque de codificación actual.

50

Lo siguiente es una explicación de las aplicaciones del método de codificación, así como el método de

decodificación tal como se muestra en las realizaciones mencionadas anteriormente, y un sistema que las usa.

La Fig. 11 es un diagrama de bloques que muestra un sistema de distribución de contenido 3100 para realizar el servicio de distribución de contenido. Este sistema de distribución de contenido 3100 incluye el dispositivo 3102 de captura, el dispositivo 3106 terminal y opcionalmente incluye el elemento de visualización 3126. El dispositivo 3102 de captura se comunica con el dispositivo 3106 terminal a través del enlace 3104 de comunicación. El enlace de comunicación puede incluir el canal 13 de comunicación descrito anteriormente. El enlace 3104 de comunicación incluye, pero no se limita a WIFI, Ethernet, Cable, inalámbrico (3G/4G/5G), USB, o cualquier tipo de combinación de los mismos, o similares.

El dispositivo 3102 de captura genera datos y puede codificar los datos mediante el método de codificación como se muestra en las realizaciones anteriores. De manera alternativa, el dispositivo 3102 de captura puede distribuir los datos a un servidor de transmisión en tiempo real (no mostrado en las Figuras), y el servidor codifica los datos y transmite los datos codificados al dispositivo terminal 3106. El dispositivo 3102 de captura incluye, pero no se limita a, cámara, teléfono inteligente o almohadilla, ordenador u ordenador portátil, sistema de videoconferencia, PDA, dispositivo montado en vehículo, o una combinación de cualquiera de ellos, o similares. Por ejemplo, el dispositivo 3102 de captura puede incluir el dispositivo 12 fuente como se describió anteriormente. Cuando los datos incluyen video, el codificador 20 de video incluido en el dispositivo 3102 de captura puede realizar realmente procesamiento de codificación de video. Cuando los datos incluyen audio (es decir, voz), un codificador de audio incluido en el dispositivo 3102 de captura puede realizar realmente el procesamiento de codificación de audio. Para algunos escenarios prácticos, el dispositivo 3102 de captura distribuye los datos codificados de video y audio al multiplexarlos conjuntamente. Para otros escenarios prácticos, por ejemplo en el sistema de videoconferencia, los datos de audio codificados y los datos de video codificados no se multiplexan. El dispositivo de captura 3102 distribuye los datos de audio codificados y los datos de video codificados al dispositivo 3106 terminal por separado.

En el sistema 3100 de distribución de contenido, el dispositivo 310 terminal recibe y reproduce los datos codificados. El dispositivo 3106 terminal podría ser un dispositivo con capacidad de recepción y recuperación de datos, tal como teléfono inteligente o almohadilla 3108, ordenador u ordenador portátil 3110, grabadora de video de red (NVR)/grabadora de video digital (DVR) 3112, TV 3114, decodificador (STB) 3116, sistema de videoconferencia 3118, sistema de videovigilancia 3120, asistente 3122 digital personal (PDA), dispositivo montado en vehículo 3124, o una combinación de cualquiera de los mismos, o similares capaces de decodificar los datos codificados mencionados anteriormente. Por ejemplo, el dispositivo 3106 terminal puede incluir el dispositivo 14 destino como se describió anteriormente. Cuando los datos codificados incluyen video, se prioriza el decodificador de video 30 incluido en el dispositivo terminal para realizar decodificación de video. Cuando los datos codificados incluyen audio, se prioriza un decodificador de audio incluido en el dispositivo terminal para realizar procesamiento de decodificación de audio.

Para un dispositivo terminal con su elemento de visualización, por ejemplo, teléfono inteligente o almohadilla 3108, ordenador 3110 portátil, grabadora de video de red (NVR)/grabadora de video digital (DVR) 3112, TV 3114, asistente 3122 digital personal (PDA), o dispositivo 3124 montado en vehículo, el dispositivo terminal puede alimentar los datos decodificados a su pantalla. Para un dispositivo terminal equipado sin pantalla, por ejemplo, STB 3116, sistema 3118 de videoconferencia o sistema 3120 de vigilancia de video, se pone en contacto un elemento de visualización 3126 externo en el mismo para recibir y mostrar los datos decodificados.

Cuando cada dispositivo en este sistema realiza codificación o decodificación, se puede usar el dispositivo de codificación de imágenes o el dispositivo de decodificación de imágenes, como se muestra en las realizaciones mencionadas anteriormente.

La Fig. 12 es un diagrama que muestra una estructura de un ejemplo del dispositivo 3106 terminal. Después de que el dispositivo 3106 terminal recibe flujo del dispositivo 3102 de captura, la unidad de procedimiento de protocolo 3202 analiza el protocolo de transmisión del flujo. El protocolo incluye, pero no se limita a, Protocolo de Transmisión en Tiempo Real (RTSP), Protocolo de Transferencia de Hipertexto (HTTP), Protocolo de Transmisión en tiempo real en Vivo HTTP (HLS), MPEG-DASH, Protocolo de Transporte en Tiempo Real (RTP), Protocolo de Mensajería en Tiempo Real (RTMP), o cualquier tipo de combinación de los mismos, o similares.

Después de que la unidad 3202 de procedimiento de protocolo procesa la secuencia, se genera el archivo de secuencia. El archivo se emite a una unidad 3204 de demultiplexación. La unidad 3204 de demultiplexación puede separar los datos multiplexados en los datos de audio codificados y los datos de video codificados. Como se describió anteriormente, para algunos escenarios prácticos, por ejemplo en el sistema de videoconferencia, los datos de audio codificados y los datos de video codificados no se multiplexan. En esta situación, los datos codificados se transmiten al decodificador 3206 de video y al decodificador 3208 de audio sin a través de la unidad 3204 de desmultiplexación.

Mediante el procesamiento de desmultiplexación, se generan el flujo elemental de video (ES), ES de audio y,

opcionalmente, subtítulos. El decodificador 3206 de video, que incluye el decodificador 30 de video como se explica en las realizaciones mencionadas anteriormente, decodifica el ES de video mediante el método de decodificación como se muestra en las realizaciones mencionadas anteriormente para generar fotogramas de video, y alimenta estos datos a la unidad 3212 síncrona. El decodificador 3208 de audio, decodifica el ES de audio para generar fotograma de audio y alimenta estos datos a la unidad 3212 síncrona. De manera alternativa, el fotograma de video se puede almacenar en una memoria intermedia (no mostrada en la Fig. Y) antes de alimentarlo a la unidad 3212 síncrona. De forma similar, el fotograma de audio se puede almacenar en una memoria intermedia (no mostrada en la Fig. Y) antes de alimentarlo a la unidad 3212 síncrona.

10 La unidad 3212 síncrona sincroniza el fotograma de video y el fotograma de audio y suministra el video/audio a un elemento de visualización 3214 de video/audio. Por ejemplo, la unidad 3212 síncrona sincroniza la presentación de la información de video y audio. La información se puede codificar en la sintaxis utilizando marcas de tiempo relativas a la presentación de datos audiovisuales codificados y marcas de tiempo relativas a la distribución del propio flujo de datos.

15 Si se incluye los subtítulos en el flujo, el decodificador 3210 de subtítulos decodifica el subtítulo y lo sincroniza con el fotograma de video y el fotograma de audio y suministra el video/audio/subtítulos a una elemento de visualización 3216 de video/audio/subtítulo.

20 El presente documento invención no se limita al sistema mencionado anteriormente, y ya sea el dispositivo de codificación de imágenes o el dispositivo de decodificación de imágenes en las realizaciones mencionadas anteriormente se pueden incorporar en otro sistema, por ejemplo, un sistema de automóvil.

#### Operadores matemáticos

25 Los operadores matemáticos usados en esta solicitud son similares a aquellos usados en el lenguaje de programación C. Sin embargo, los resultados de las operaciones de división de números enteros y cambio aritmético se definen con mayor precisión, y se definen operaciones adicionales, tal como exponenciación y división con valores reales. Las convenciones de numeración y conteo generalmente comienzan desde 0, por ejemplo, "el primero" es equivalente al 0-ésimo, "el segundo" es equivalente al 1-ero, etc.

#### Operadores aritméticos

Los siguientes operadores aritméticos se definen como sigue:

|                     |  |
|---------------------|--|
| +                   | Suma   |
| -                   | Resta (como operador de dos argumentos) o negación (como operador de prefijo unario)   |
| *                   | Multiplicación, que incluye multiplicación de matrices   |
| $x^y$               | Exponenciación. Especifica x a la potencia de y. En otros contextos, esta notación se utiliza para superíndices que no pretende que se interprete como exponenciación. |
| /                   | División entera con truncamiento del resultado hacia cero. Por ejemplo, $7 / 4$ y $-7 / -4$ se truncan a 1 y $-7 / 4$ y $7 / -4$ se truncan a -1.                      |
| +                   | Se usa para denotar división en ecuaciones matemáticas donde no se pretende truncar o redondear.   |
| $\frac{x}{y}$       | Se usa para denotar división en ecuaciones matemáticas donde no se pretende truncar o redondear.   |
| $\sum_{i=x}^y f(i)$ | La suma de f(i) con i que toma todos los valores enteros de x hasta y que incluye y.   |
| $x \% y$            | Módulo. Resto de x dividido por y, definido solo para enteros x e y con $x \geq 0$ e $y > 0$ .   |

#### Operadores lógicos

Los siguientes operadores lógicos se definen como sigue:

|    |             |  |
|----|-------------|--|
| 40 | $x \&\& y$  | Booleano lógico "y" de x y y   |
|    | $x \ \  y$  | Booleano lógico "o" de x e y   |
|    | !           | Booleano lógico "no"   |
| 45 | $x ? y : z$ | Si x es VERDADERO o no igual a 0, se evalúa al valor de y; de lo contrario, se evalúa al valor de z. |

#### Operadores relacionales

50 Los siguientes operadores relacionales se definen como sigue:

|    |     |                   |
|----|-----|-------------------|
|    | >   | Mayor que         |
|    | >=  | Mayor o igual que |
| 5  | <   | Menor que         |
|    | <=  | Menor o igual que |
|    | = = | Igual a           |
| 10 | !=  | No igual a        |

15 Cuando un operador relacional se aplica a un elemento o variable de sintaxis al que se le ha asignado el valor "na" (no aplicable), el valor "na" se trata como un valor distinto para el elemento o variable de sintaxis. Se considera que el valor "na" no es igual a ningún otro valor.

#### Operadores de nivel de bit

20 Los siguientes operadores de nivel de bit se definen como sigue:

& "y" a nivel de bit. Cuando se opera en argumentos enteros, opera en una representación de complemento de dos del valor entero. Cuando se opera en un argumento binario que contiene menos bits que otro argumento, se extiende el argumento más corto agregando bits más significativos iguales a 0.

25 | "o" a nivel de bit. Cuando se opera en argumentos enteros, opera en una representación de complemento de dos del valor entero. Cuando se opera en un argumento binario que contiene menos bits que otro argumento, se extiende el argumento más corto agregando bits más significativos iguales a 0.

30 ^ "exclusivo o" a nivel de bit. Cuando se opera en argumentos enteros, opera en una representación de complemento de dos del valor entero. Cuando se opera en un argumento binario que contiene menos bits que otro argumento, se extiende el argumento más corto agregando bits más significativos iguales a 0.

35  $x \gg y$  Desplazamiento aritmético derecho de una representación entera de complemento de dos de  $x$  por  $y$  dígitos binarios. Esta función se define solo para valores enteros no negativos de  $y$ . Los bits trasladados a los bits más significativos (MSB) como resultado del desplazamiento a la derecha tienen un valor igual al MSB de  $x$  antes de la operación de desplazamiento.

40  $x \ll y$  Desplazamiento aritmético izquierdo de una representación entera de complemento de dos de  $x$  por  $y$  dígitos binarios. Esta función se define solo para valores enteros no negativos de  $y$ . Los bits trasladados a los bits menos significativos (LSB) como resultado del desplazamiento a la izquierda tienen un valor igual a 0.

#### Operadores de asignación

45 Los siguientes operadores aritméticos se definen como sigue:

= Operador de asignación

50 + + Incremento, es decir,  $x + +$  es equivalente a  $x = x + 1$ ; cuando se utiliza en un índice de matriz, se evalúa al valor de la variable antes de la operación de incremento.

-- Decremento, es decir,  $x - -$  es equivalente a  $x = x - 1$ ; cuando se utiliza en un índice de matriz, se evalúa el valor de la variable antes de la operación de disminución.

55 += Incremento por cantidad especificada, es decir,  $x += 3$  es equivalente a  $x = x + 3$ , y  $x += (-3)$  es equivalente a  $x = x + (-3)$ .

-= Decremento por cantidad especificada, es decir,  $x -= 3$  es equivalente a  $x = x - 3$ , y  $x -= (-3)$  es equivalente a  $x = x - (-3)$ .

60 Notación de intervalo

La siguiente notación se usa para especificar un intervalo de valores:

65  $x = y..z$   $x$  supone valores enteros desde  $y$  a  $z$ , inclusive, con  $x$ ,  $y$  y  $z$  que son números enteros y  $z$  mayores que  $y$ .

Funciones matemáticas

Se definen las siguientes funciones matemáticas:

$$5 \quad \text{Abs}(x) = \begin{cases} x & ; \quad x \geq 0 \\ -x & ; \quad x < 0 \end{cases}$$

Asin(x) la función seno inversa trigonométrica, que opera en un argumento x que está en el intervalo de -1.0 a 1.0, inclusive, con un valor de salida en el intervalo de  $-\pi/2$  a  $\pi/2$ , inclusive, en unidades de radianes

10 Atan(x) la función tangente inversa trigonométrica, que opera en un argumento x, con un valor de salida en el intervalo de  $-\pi/2$  a  $\pi/2$ , inclusive, en unidades de radianes

$$\text{Atan2}(y, x) = \begin{cases} \text{Atan}\left(\frac{y}{x}\right) & ; \quad x > 0 \\ \text{Atan}\left(\frac{y}{x}\right) + \pi & ; \quad x < 0 \ \&\& \ y \geq 0 \\ \text{Atan}\left(\frac{y}{x}\right) - \pi & ; \quad x < 0 \ \&\& \ y < 0 \\ +\frac{\pi}{2} & ; \quad x = 0 \ \&\& \ y \geq 0 \\ -\frac{\pi}{2} & ; \quad \text{de otra forma} \end{cases}$$

15 Techo(x) el número entero más pequeño mayor o igual a x.

$$\text{Clip1v}(x) = \text{Clip3}(0, (1 \ll \text{ProfundidadBitv}) - 1, x)$$

$$\text{Clip1c}(x) = \text{Clip3}(0, (1 \ll \text{ProfundidadBitc}) - 1, x)$$

20

$$\text{Clip3}(x, y, z) = \begin{cases} x & ; \quad z < x \\ y & ; \quad z > y \\ z & ; \quad \text{de otra forma} \end{cases}$$

Cos(x) la función coseno trigonométrica que opera en un argumento x en unidades de radianes.

25 Suelo(x) el número entero más grande menor o igual a x.

$$\text{GetCurrMsb}(a, b, c, d) = \begin{cases} c + d & ; \quad b - a \geq d / 2 \\ c - d & ; \quad a - b > d / 2 \\ c & ; \quad \text{de otra forma} \end{cases}$$

30 Ln(x) el logaritmo natural de x (el logaritmo base-e, donde e es la constante de base del logaritmo natural 2.718 281 828...).

Log2(x) el logaritmo base-2 de x.

Log10(x) el logaritmo base-10 de x.

35

$$\text{Min}(x, y) = \begin{cases} x & ; \quad x \leq y \\ y & ; \quad x > y \end{cases}$$

$$\text{Max}(x, y) = \begin{cases} x & ; \quad x \geq y \\ y & ; \quad x < y \end{cases}$$

40 Redondear(x) = Signo(x) \* Suelo(Abs(x) + 0,5)

$$\text{Signo}(x) = \begin{cases} 1 & ; \quad x > 0 \\ 0 & ; \quad x = 0 \\ -1 & ; \quad x < 0 \end{cases}$$

Sin(x) la función seno trigonométrica que opera en un argumento x en unidades de radianes

45

$$\text{Sqrt}(x) = \sqrt{x}$$

$$\text{Cambiar}(x, y) = (y, x)$$

50 Tan(x) la función tangente trigonométrica que opera en un argumento x en unidades de radianes

La invención se ha descrito junto con diversas realizaciones en el presente documento. Sin embargo, se pueden comprender y efectuar otras variaciones a las realizaciones descritas por aquellos expertos en la técnica en la práctica de la invención reivindicada, a partir de un estudio de los dibujos, la descripción y las reivindicaciones anexas. En las reivindicaciones, la palabra "comprendiendo" no excluye otros elementos o pasos y el artículo indefinido "un" o "una" no excluye una pluralidad. Un solo procesador u otra unidad puede cumplir con las funciones de varios artículos mencionados en las reivindicaciones. El simple hecho de que determinadas mediciones se mencionen en reivindicaciones dependientes generalmente diferentes no indica que una combinación de estas mediciones no se pueda usar con ventaja. Un programa informático se puede almacenar/distribuir en un medio adecuado, tal como un medio de almacenamiento óptico o un medio de estado sólido suministrado junto con o como parte de otro hardware, pero también se puede distribuir en otras formas, tal como a través de Internet u otros sistemas de comunicación alámbricos o inalámbricos.

El experto en la técnica entenderá que los "bloques" ("unidades") de las diversas figuras (método y aparato) representan o describen funcionalidades de realizaciones de la invención (en lugar de necesariamente "unidades" individuales en hardware o software) y por lo tanto describen igualmente funciones o características de realizaciones de aparatos así como realizaciones de métodos (unidad = paso).

La terminología de "unidades" se utiliza simplemente con fines ilustrativos de la funcionalidad de realizaciones del codificador/decodificador y no se pretende que limite la descripción.

En las diversas realizaciones proporcionadas en el presente documento solicitud, se debe entender que el sistema, aparato y método descritos se pueden implementar de otras maneras. Por ejemplo, la modalidad de aparato descrita es meramente de ejemplo. Por ejemplo, la división de unidades es simplemente una división de funciones lógicas y puede ser otra división en la implementación real. Por ejemplo, una pluralidad de unidades o componentes se pueden combinar o integrar en otro sistema, o algunas características se pueden ignorar o no realizar. Además, los acoplamientos mutuos o directos o conexiones de comunicación mostradas o analizadas se pueden implementar al utilizar algunas interfaces. Los acoplamientos indirectos o las conexiones de comunicación entre los aparatos o unidades se pueden implementar de forma electrónica, mecánica u otras.

Las unidades descritas como partes separadas pueden o no estar separadas físicamente, y las partes mostradas como unidades pueden o no ser unidades físicas, pueden estar ubicadas en una posición, o pueden distribuirse en una pluralidad de unidades de red. Algunos o todas las unidades se pueden seleccionar según las necesidades reales para lograr los objetivos de las soluciones de las realizaciones.

Además, las unidades funcionales en las realizaciones del presente documento invención se pueden integrar en una unidad de procesamiento, o cada una de las unidades puede existir físicamente por separado, o dos o más unidades se integran en una unidad.

Las realizaciones de la invención pueden comprender además un aparato, por ejemplo, codificador y/o decodificador, que comprende una circuitería de procesamiento configurada para realizar cualquiera de los métodos y/o procesos descritos en el presente documento.

Aunque las realizaciones de la invención se han descrito principalmente con base en codificación de video, se debe señalar que las realizaciones del sistema de codificación 10, codificador 20 y decodificador 30 (y, en consecuencia, el sistema 10) y las otras realizaciones descritas en el presente documento también se pueden configurar para el procesamiento o codificación de imágenes fijas, es decir, el procesamiento o codificación de una imagen individual independiente de cualquier imagen anterior o consecutiva como en la codificación de video. En general, solo las unidades 244 de interpretación (codificador) y 344 (decodificador) pueden no estar disponibles en caso de que la codificación de procesamiento de imágenes se limite a una sola imagen 17. Todas las demás funcionalidades (también denominadas herramientas o tecnologías) del codificador 20 de video y un decodificador 30 de video se pueden usar igualmente para procesamiento de imágenes fijas, por ejemplo, cálculo residual 204/304, transformada 206, cuantificación 208, cuantificación inversa 210/310, transformada (inversa) 212/312, partición 262/362, intrapredicción 254/354 y/o filtrado de bucle 220, 320 y codificación entrópica 270 y decodificación entrópica 304.

Las realizaciones, por ejemplo, del codificador 20 y el decodificador 30, y las funciones descritas en el presente documento, por ejemplo, con referencia al codificador 20 y el decodificador 30, se pueden implementar en hardware, software, firmware, o cualquier combinación de los mismos. Si se implementan en software, las funciones se pueden almacenar en un medio legible por ordenador o transmitir sobre medios de comunicación como una o más instrucciones o código y ejecutarse por una unidad de procesamiento basada en hardware. Los medios legibles por ordenador pueden incluir medios de almacenamiento legibles por ordenador, que corresponden a un medio tangible tal como medios de almacenamiento de datos o medios de comunicación que incluyen cualquier medio que facilite transferencia de un programa informático de un lugar a otro, por ejemplo, según un protocolo de comunicación. De esta manera, los medios legibles por ordenador

5 generalmente pueden corresponder a (1) medios de almacenamiento legibles por ordenador tangibles que no son transitorios o (2) un medio de comunicación tal como una señal u onda portadora. Los medios de almacenamiento de datos pueden ser cualquier medio disponible al que se puede acceder por uno o más ordenadores o uno o más procesadores para recuperar instrucciones, código y/o estructuras de datos para implementación de las técnicas descritas en esta descripción. Un producto de programa informático puede incluir un medio legible por ordenador.

10 A manera de ejemplo y no de limitación, estos medios de almacenamiento legibles por ordenador pueden comprender RAM, ROM, EEPROM, CD-ROM u otro almacenamiento de disco óptico, almacenamiento de disco magnético, u otros dispositivos de almacenamiento magnético, memoria flash o cualquier otro medio que se pueda utilizar para almacenar código de programa deseado en forma de instrucciones o estructuras de datos y al que se puede acceder mediante un ordenador. Además, cualquier conexión se denomina apropiadamente un medio legible por ordenador. Por ejemplo, si las instrucciones se transmiten desde un sitio web, servidor u otra fuente remota usando un cable coaxial, cable de fibra óptica, par trenzado, línea de abonado digital (DSL) o tecnologías inalámbricas tal como infrarrojo, radiofrecuencia y microondas, entonces el cable coaxial, cable de fibra óptica, par trenzado, DSL o tecnologías inalámbricas tal como infrarrojo, radiofrecuencia y microondas se incluyen en la definición de medio. Se debe entender, sin embargo, que los medios de almacenamiento legibles por ordenador y los medios de almacenamiento de datos no incluyen conexiones, ondas portadoras, señales u otros medios transitorios, sino que en su lugar se dirigen a medios de almacenamiento tangibles no transitorios. El disco digital y disco, como se usa en el presente documento, incluye disco compacto (CD), disco láser, disco óptico, disco versátil digital (DVD), disquete y disco Blu-ray, donde los discos generalmente reproducen datos magnéticamente, mientras que los discos digitales reproducen datos ópticamente con láseres. Las combinaciones de los anteriores también se deben incluir dentro del alcance de los medios legibles por ordenador.

25 Se puede ejecutar instrucciones de programa por uno o más procesadores, tal como uno o más procesadores de señales digitales (DSP), microprocesadores de propósito general, circuitos integrados de aplicación específica (ASIC), matrices lógicas programables en campo (FPGA) u otra circuitería lógica integrada o discreta equivalentes. Por consiguiente, el término "procesador", como se usa en el presente documento, se puede referir a cualquiera de la estructura anterior o cualquier otra estructura adecuada para 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 y/o hardware dedicados configurados para codificación y decodificación, o se incorporan en un códec combinado. Además, las técnicas se podrían implementar completamente en uno o más circuitos o elementos lógicos.

35 Las técnicas de esta descripción se pueden implementar en una amplia variedad de dispositivos o aparatos, que incluyen un aparato inalámbrico, un circuito integrado (IC) o un conjunto de IC (por ejemplo, un conjunto de chips). Varios componentes, módulos o unidades se describen en esta descripción para enfatizar aspectos funcionales de dispositivos configurados para realizar las técnicas descritas, pero no requieren necesariamente la realización por diferentes unidades de hardware. Más bien, como se describió anteriormente, se pueden combinar varias unidades en una unidad de hardware de códec o proporcionadas por una recolección de unidades de hardware interoperativas, que incluyen uno o más procesadores como se describió anteriormente, junto con software y/o firmware adecuado.

**REIVINDICACIONES**

1. Un método para interpredicción basada en flujo óptico bidireccional (BDOF) para un bloque actual de una señal de video, que comprende:

5 determinar un desplazamiento de movimiento horizontal  $v_x$  y un desplazamiento de movimiento vertical del bloque actual  $v_y$ , y en donde el desplazamiento de movimiento vertical se deriva con base en el desplazamiento de movimiento horizontal, una segunda variable  $s_2$ , una cuarta variable  $s_4$ , y una quinta variable  $s_5$ , en donde:

10 la quinta variable  $s_5$  indica una suma de una pluralidad de primeros términos, en donde cada uno de la pluralidad de primeros términos se obtiene de un signo de un elemento de una segunda matriz y un elemento de una primera matriz, y el elemento de la primera matriz corresponde al elemento de la segunda matriz;

15 en donde cada elemento de la primera matriz se obtiene a partir de una suma de un primer gradiente de muestra predicho horizontal que corresponde a un primer fotograma de referencia del bloque actual y un segundo gradiente de muestra predicho horizontal que corresponde a un segundo fotograma de referencia del bloque actual, en donde el primer gradiente de muestra predicho horizontal y el segundo gradiente de muestra predicho horizontal corresponden al elemento de la primera matriz; y

20 cada elemento de la segunda matriz se obtiene a partir de una suma de un primer gradiente de muestra predicho vertical que corresponde al primer fotograma de referencia del bloque actual y un segundo gradiente de muestra predicho vertical que corresponde al segundo fotograma de referencia del bloque actual, en donde el primer gradiente de muestra predicho vertical y el segundo gradiente de muestra predicho vertical corresponden al elemento de la segunda matriz; y

25 en donde

la segunda variable indica una suma de valores absolutos de elementos de la segunda matriz; y

30 la cuarta variable  $s_4$  indica una suma de una pluralidad de segundos términos, en donde cada uno de la pluralidad de segundos términos se obtiene de un signo de un elemento de la segunda matriz y un elemento de una tercera matriz, y el elemento de la tercera matriz corresponde al elemento de la segunda matriz, y cada elemento de la tercera matriz es una diferencia que se obtiene de una primera muestra predicha del primer fotograma de referencia que corresponde al elemento de la tercera matriz y una segunda muestra predicha del

35 segundo fotograma de referencia que corresponde al elemento de la tercera matriz;

en donde:

40 el desplazamiento de movimiento horizontal se deriva con base en una primera variable  $s_1$  y una tercera variable  $s_3$ ;

la primera variable  $s_1$  indica una suma de valores absolutos de elementos de la primera matriz; y

45 la tercera variable  $s_3$  indica una suma de una pluralidad de terceros términos, en donde cada uno de la pluralidad de terceros términos se obtiene de un signo de un elemento de la primera matriz y un elemento de la tercera matriz, y el elemento de la tercera matriz corresponde al elemento de la primera matriz, y

50 en donde el método además comprende determinar valores de muestra de predicción en el bloque actual usando valores de muestra de predicción correspondientes al primer cuadro de referencia, valores de muestra de predicción correspondientes al segundo cuadro de referencia, el desplazamiento de movimiento horizontal y el desplazamiento de movimiento vertical.

2. El método de la reivindicación 1, en donde el desplazamiento de movimiento horizontal se determina según

55 
$$v_x = -\frac{s_3}{s_1}$$

en donde:

60  $v_x$  representa el desplazamiento de movimiento horizontal.

3. El método de la reivindicación 1, en donde el desplazamiento de movimiento vertical  $v_y$ , se determina según

$$v_y = -\frac{(s_4 - v_x * s_5/2)}{s_2},$$

en donde:

$v_x$  representa el desplazamiento de movimiento horizontal; y

$v_y$  representa el desplazamiento de movimiento vertical.

4. El método de la reivindicación 2 o 3, en donde  $s_1$ ,  $s_2$ ,  $s_3$ ,  $s_4$  y  $s_5$  se determinan como:

$$s_1 = \sum_{i,j} \text{abs}(G_{x1} + G_{x0});$$

$$s_2 = \sum_{i,j} \text{abs}(G_{y1} + G_{y0}),$$

$$s_3 = \sum_{i,j} \text{signo}(G_{x1} + G_{x0}) * (I^{(1)} - I^{(0)}),$$

$$s_4 = \sum_{i,j} \text{signo}(G_{y1} + G_{y0}) * (I^{(1)} - I^{(0)}),$$

$$s_5 = \sum_{i,j} \text{signo}(G_{y1} + G_{y0}) * (G_{x1} + G_{x0});$$

y en donde:

$I^{(0)}$  se obtiene a partir de los valores de muestra predichos que corresponden al primer fotograma de referencia, y  $I^{(1)}$  se obtiene a partir de los valores de muestra predichos que corresponden al segundo fotograma de referencia;

$G_{x0}$  y  $G_{x1}$  indica un conjunto de los gradientes de muestra predichos horizontales que corresponden al primer fotograma de referencia y al segundo fotograma de referencia, respectivamente;

$G_{y0}$  y  $G_{y1}$  indican un conjunto de los gradientes de muestra predichos verticales que corresponden al primer fotograma de referencia y al segundo fotograma de referencia, respectivamente;

$i$  y  $j$  son números enteros, en donde un valor de  $i$  varía de -1 a 4, y un valor de  $j$  varía de -1 a 4.

5. El método de la reivindicación 4, en donde el  $G_{x0}$  se determina como una diferencia obtenida de dos muestras predichas que corresponden al primer fotograma de referencia a lo largo de una dirección horizontal, y el  $G_{y0}$  se determina como una diferencia obtenida de dos muestras predichas que corresponden al primer fotograma de referencia a lo largo de una dirección vertical.

6. El método de la reivindicación 4, en donde el  $G_{x1}$  se determina como una diferencia obtenida de dos muestras predichas que corresponden al segundo fotograma de referencia a lo largo de una dirección horizontal, y el  $G_{y1}$  se determina como una diferencia obtenida de dos muestras predichas que corresponden al segundo fotograma de referencia a lo largo de una dirección vertical.

7. El método de una cualquiera de las reivindicaciones 1 a 6, en donde los valores de muestra de predicción correspondientes al primer fotograma de referencia y los valores de muestra de predicción correspondientes al segundo fotograma de referencia se obtienen de los primeros y segundos fotogramas de referencia, respectivamente, usando un par de vectores de movimiento para el bloque actual con respecto a los primeros

y segundos fotogramas de referencia.

- 5 8. El método de una cualquiera de las reivindicaciones 1 a 7, en donde los valores de muestra de predicción para el bloque actual son valores de muestra bi-predichos con base en la predicción de flujo óptico bidireccional (BDOF).
9. Un codificador (20) que comprende circuitería de procesamiento para llevar a cabo el método según una cualquiera de las reivindicaciones 1 a 8.
- 10 10. Un decodificador (30) que comprende circuitería de procesamiento para llevar a cabo el método según cualquiera de las reivindicaciones 1 a 8.
11. Un producto de programa informático que comprende un código de programa para realizar el método según una cualquiera de las reivindicaciones 1 a 8.
- 15 12. Un medio legible por ordenador no transitorio que porta un código de programa que, cuando se ejecuta por un dispositivo informático, hace que el dispositivo informático realice el método de una cualquiera de las reivindicaciones 1 a 8.

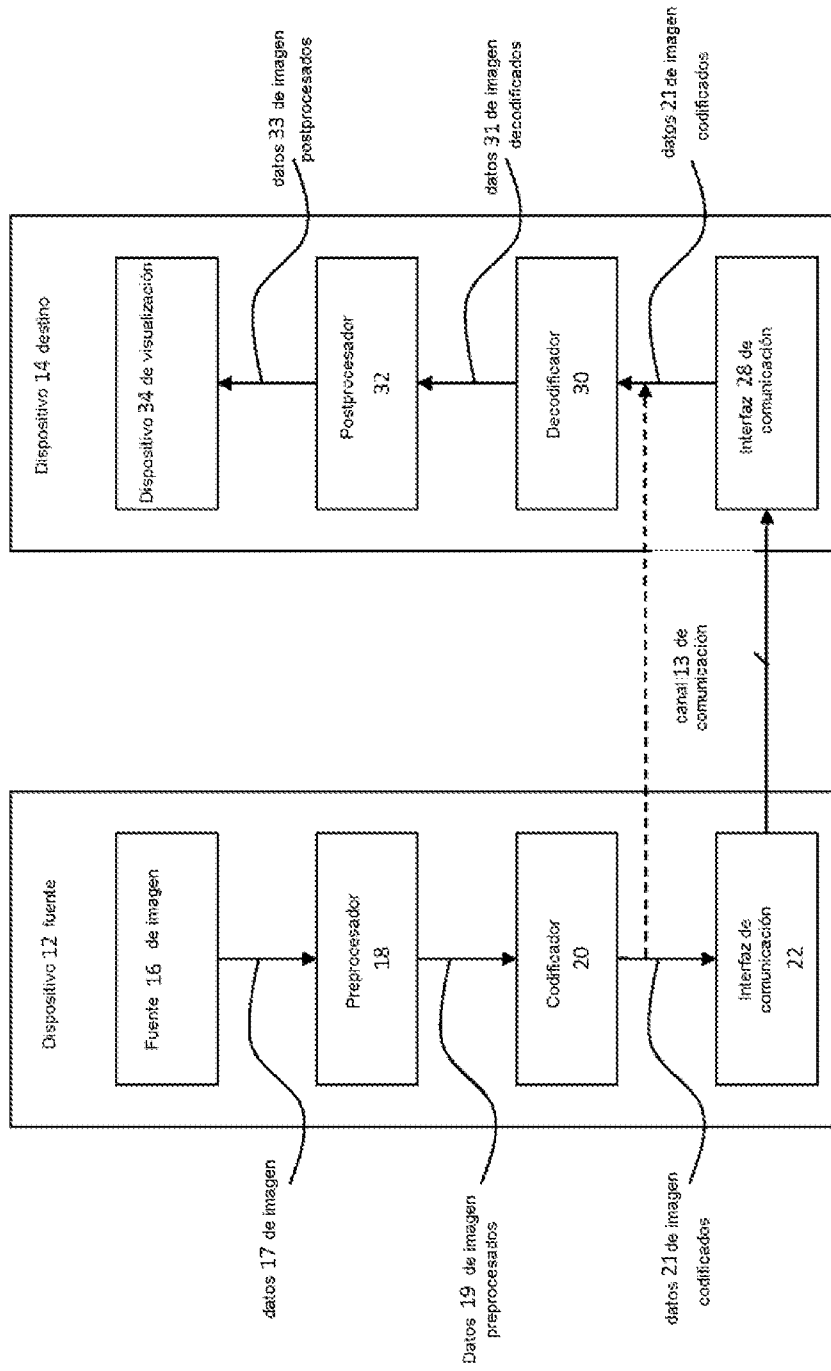


FIG.1A

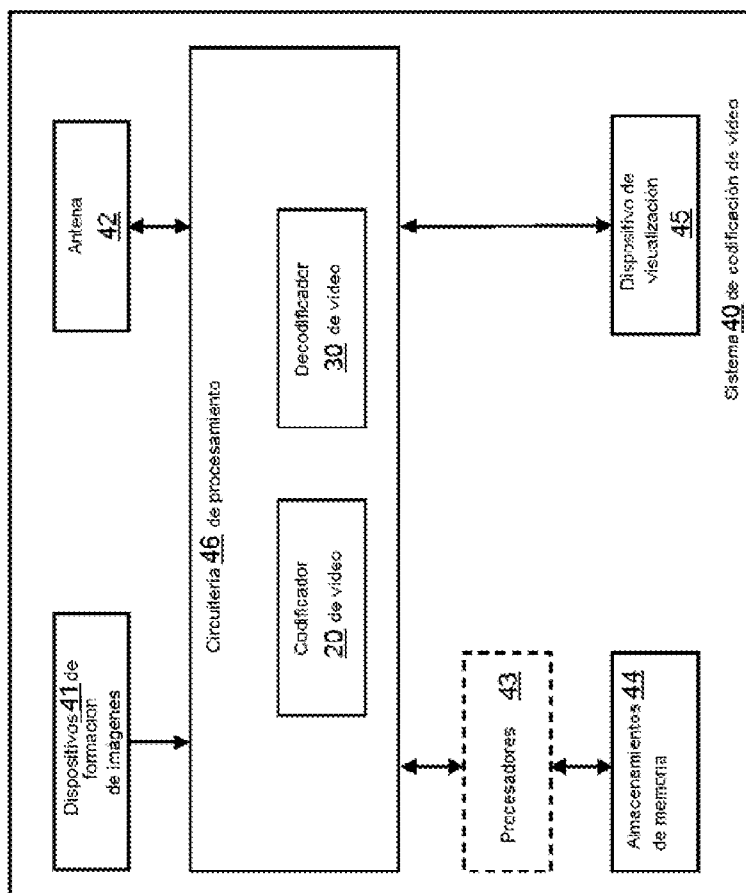


FIG.1B

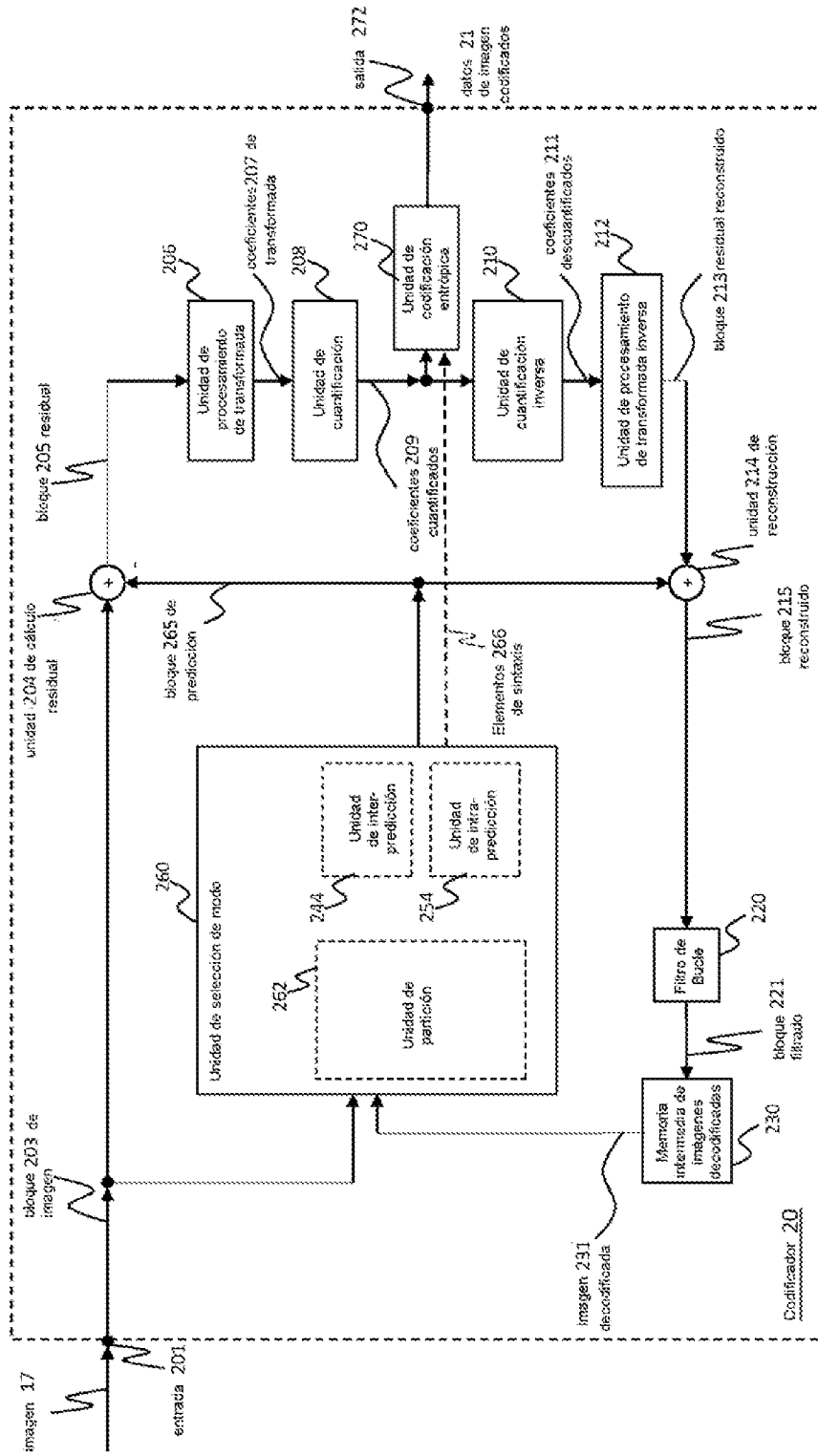


FIG. 2

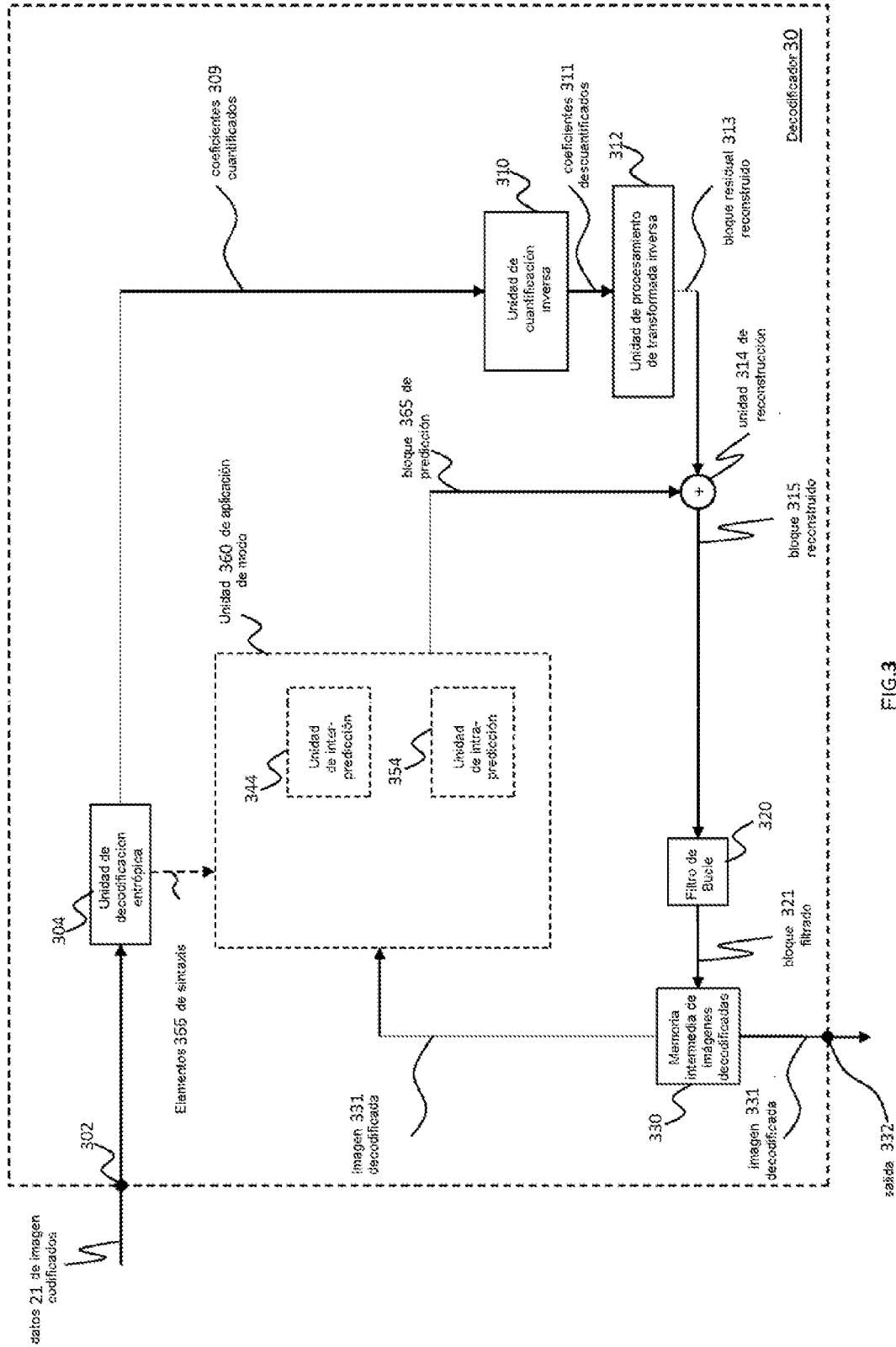


FIG. 3

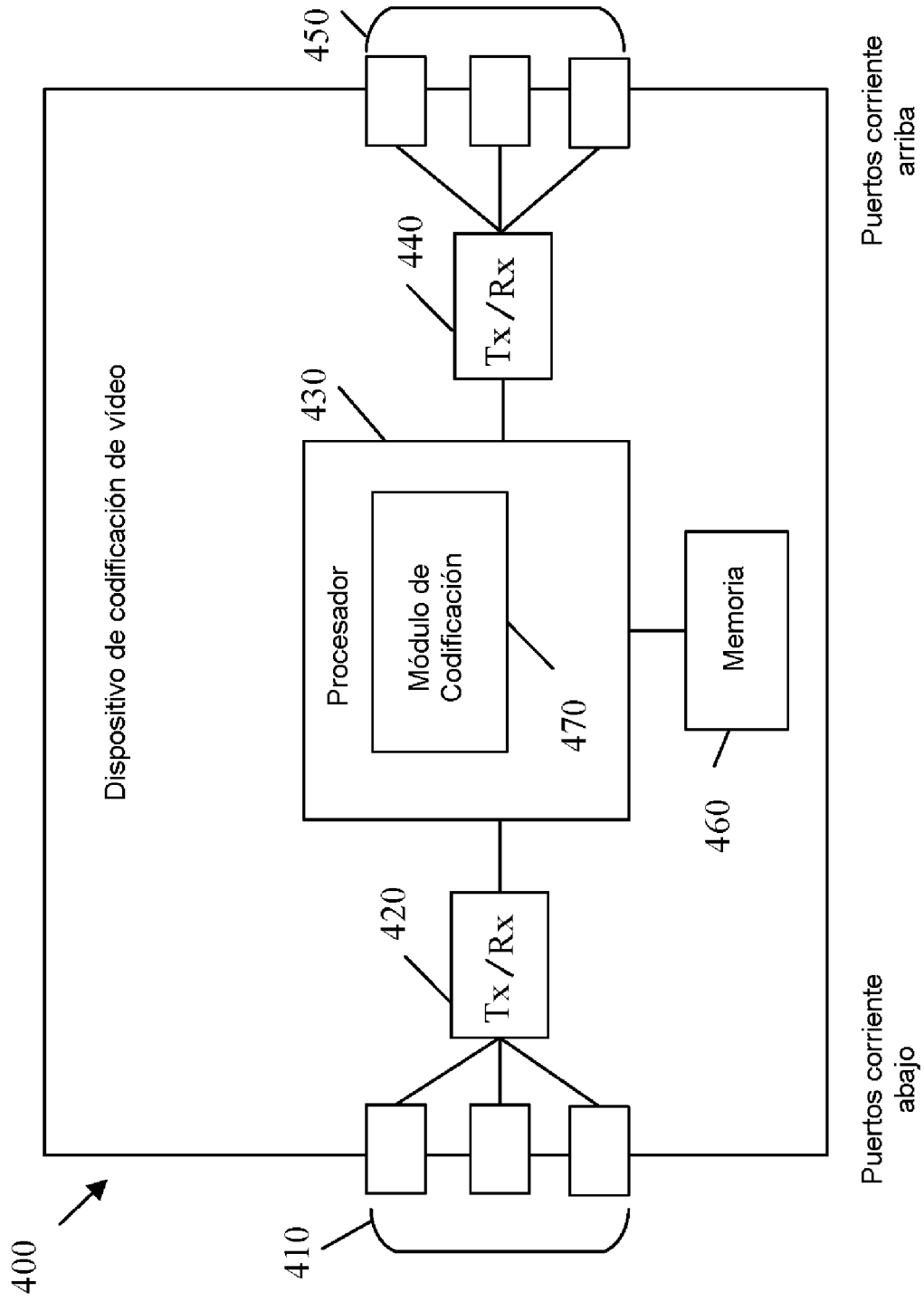


FIG.4

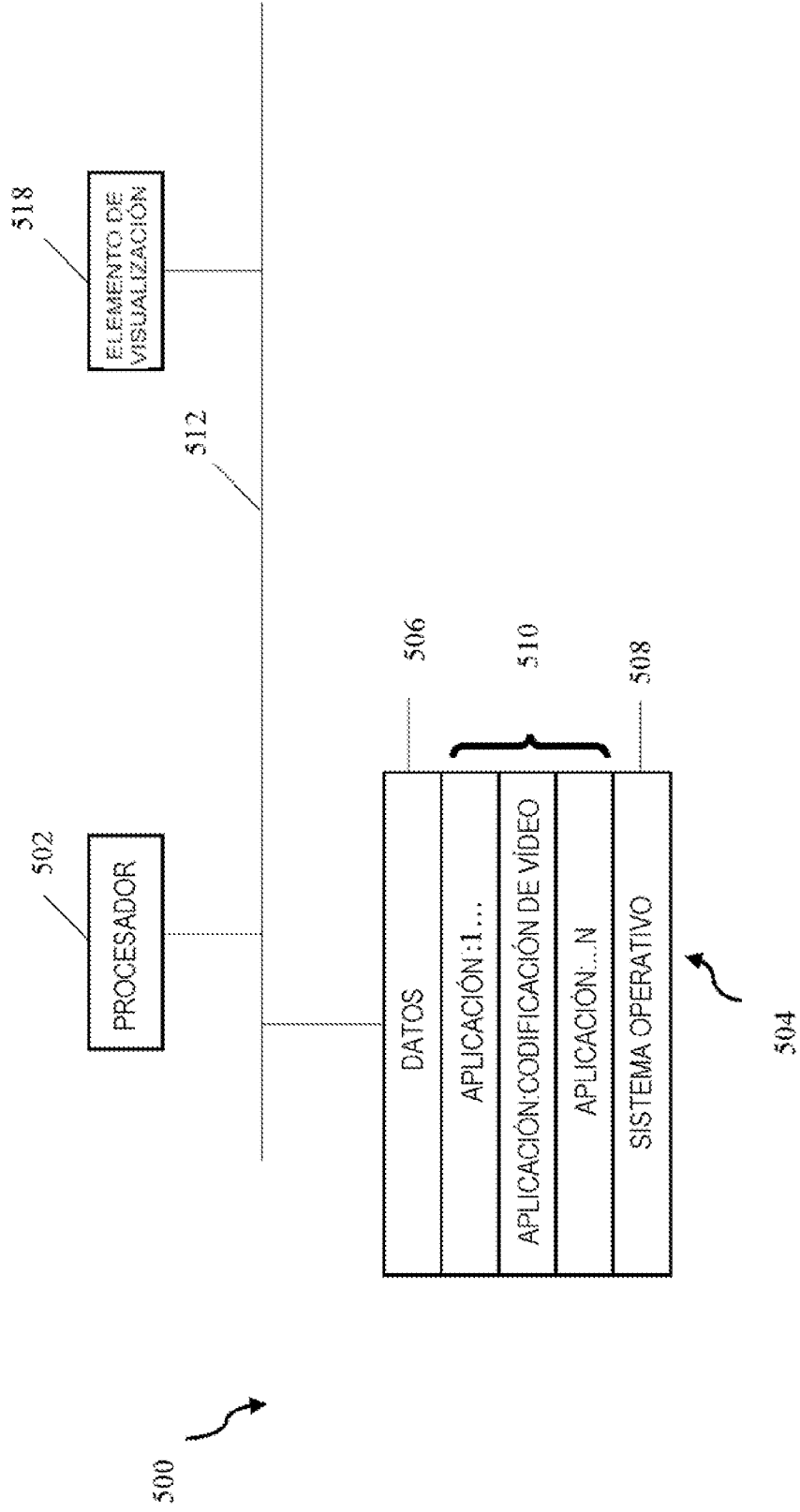


FIG. 5

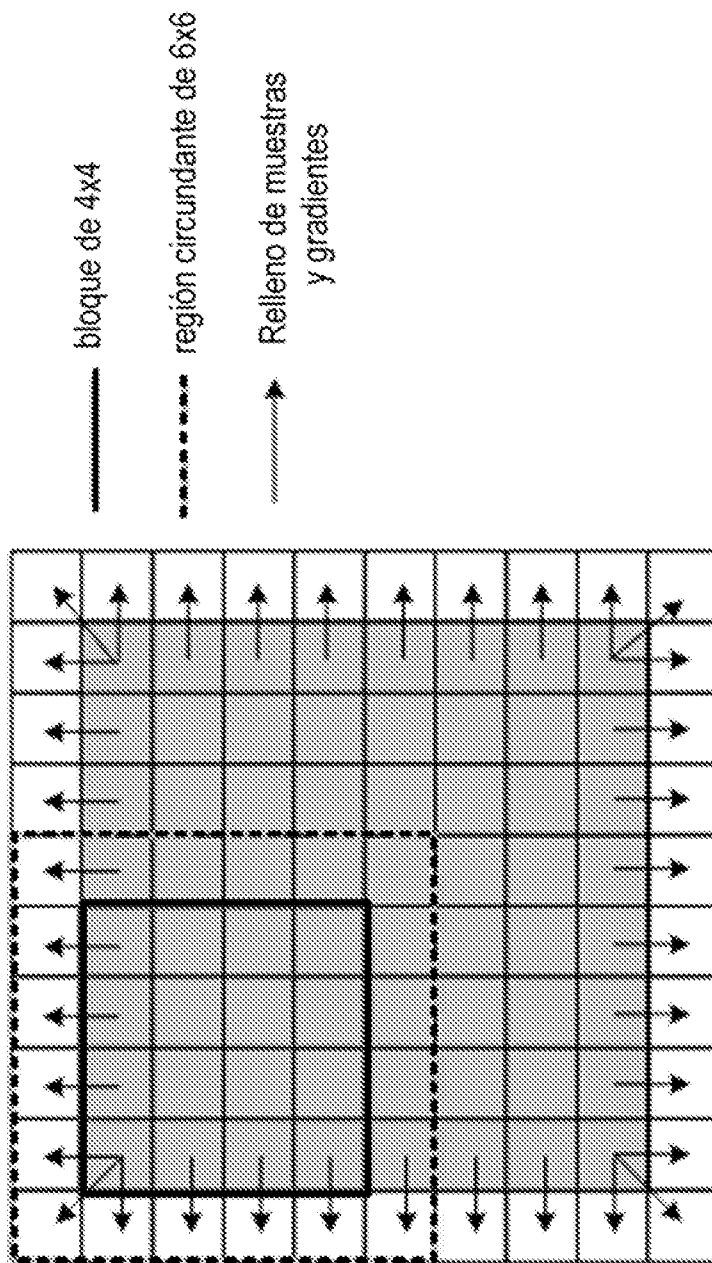


Fig. 6

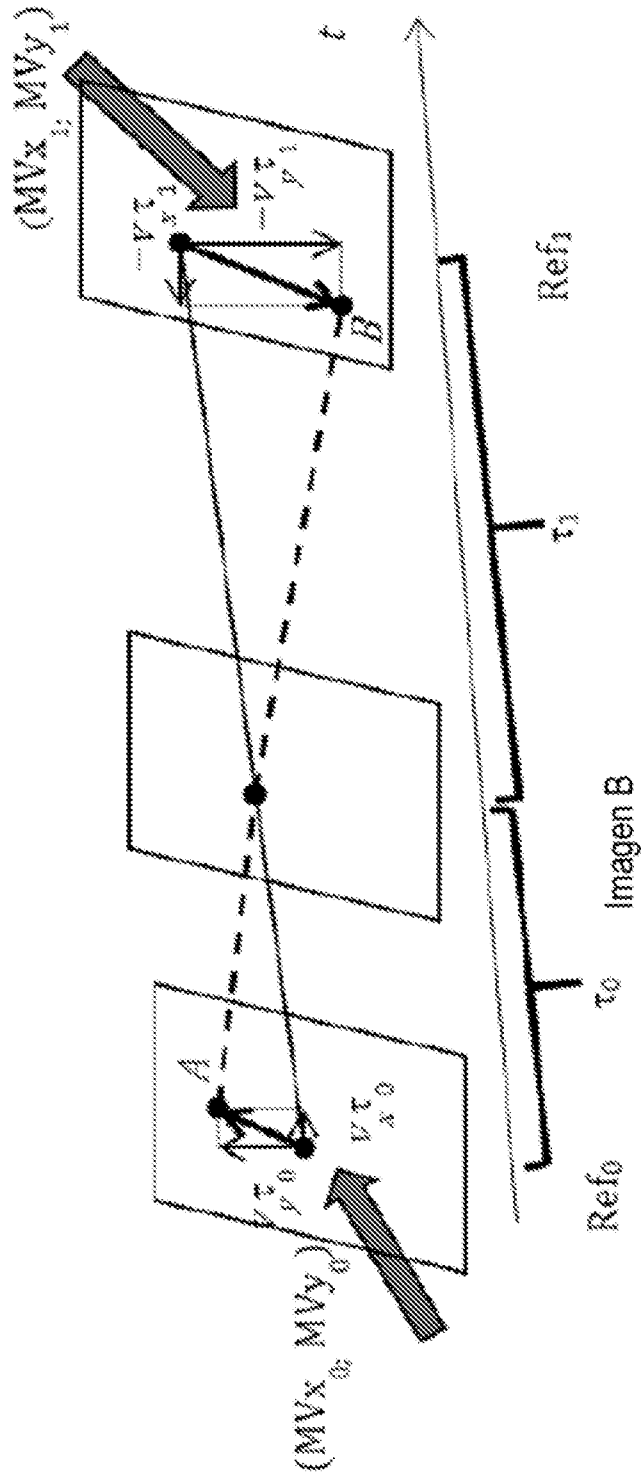


FIG 7: Flujo Óptico Bi-predictivo

800

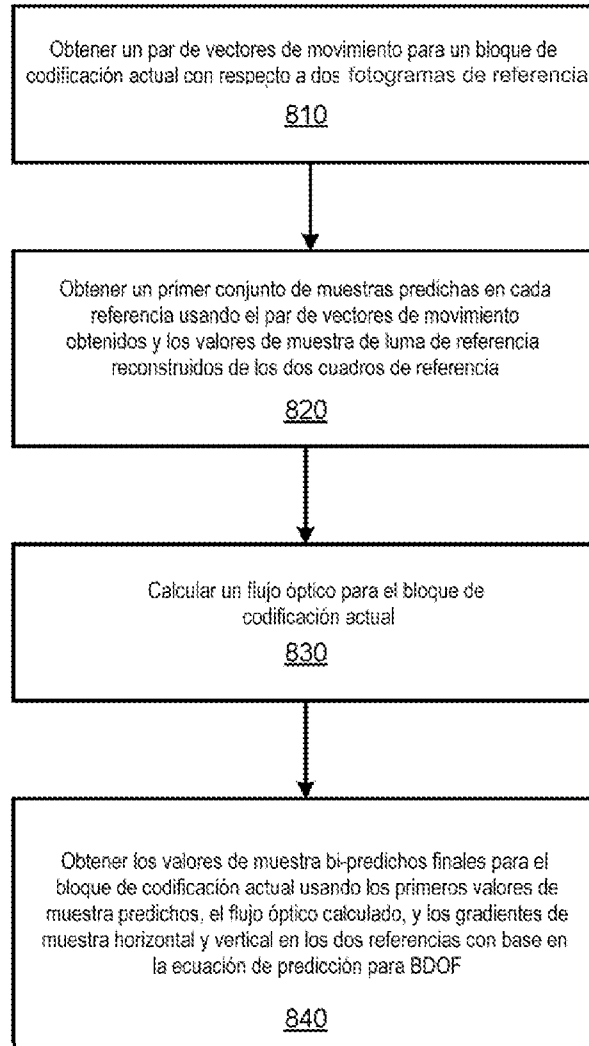


Fig. 8

900

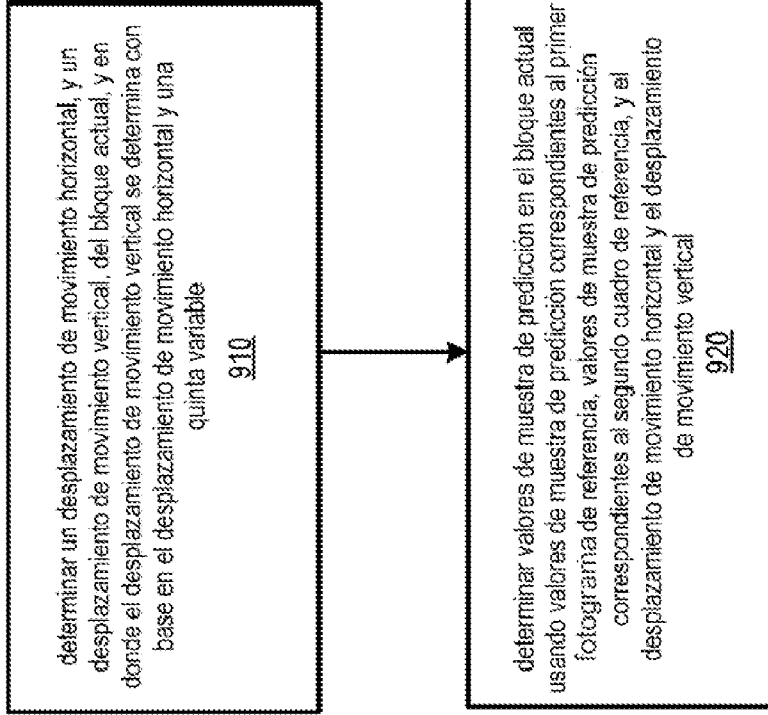
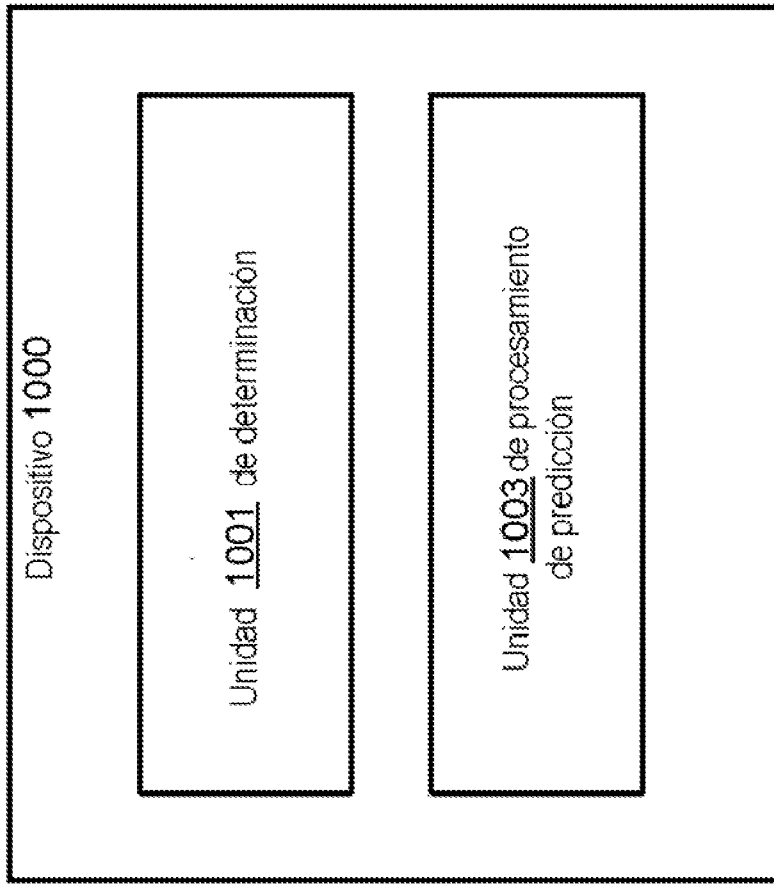


FIG. 9



**FIG. 10**

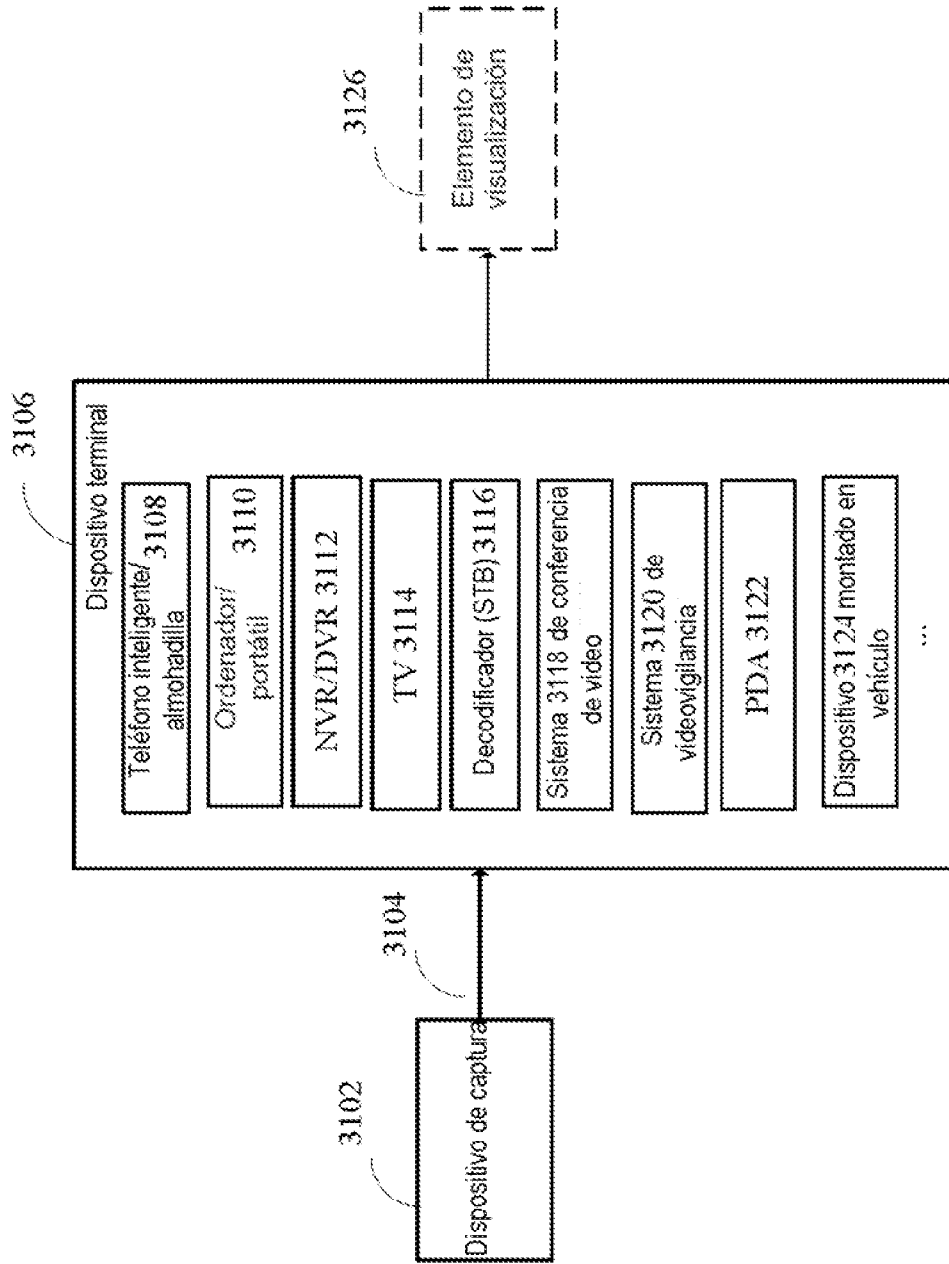


FIG. 11

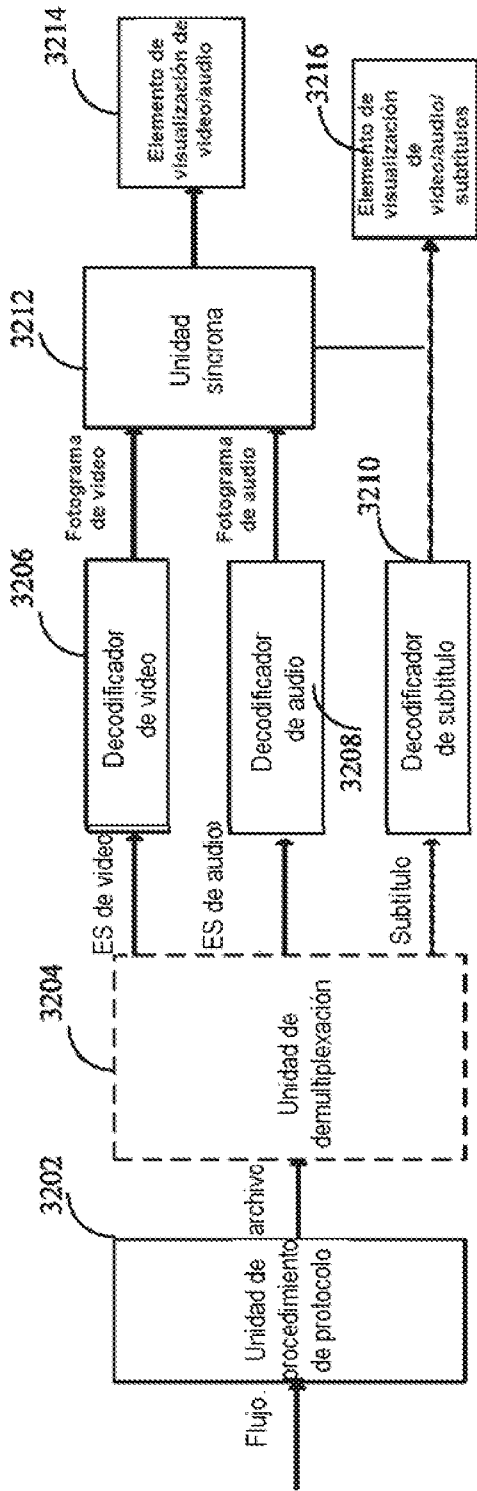


FIG. 12