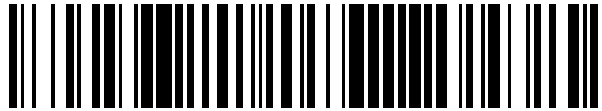


19



OFICINA ESPAÑOLA DE  
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 912 367**

51 Int. Cl.:

**H04N 19/105** (2014.01)

**H04N 19/52** (2014.01)

**H04N 19/176** (2014.01)

**H04N 19/51** (2014.01)

**H04N 19/157** (2014.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

86 Fecha de presentación y número de la solicitud internacional: **23.02.2012 PCT/US2012/026368**

87 Fecha y número de publicación internacional: **30.08.2012 WO12116212**

96 Fecha de presentación y número de la solicitud europea: **23.02.2012 E 12706428 (5)**

97 Fecha y número de publicación de la concesión europea: **23.03.2022 EP 2679011**

54 Título: **Jerarquía de bloques de vídeo de predicción de movimiento**

30 Prioridad:

**24.02.2011 US 201161446392 P**

**26.02.2011 US 201161447017 P**

**10.03.2011 US 201161451493 P**

**30.08.2011 US 201161529110 P**

**06.09.2011 US 201161531514 P**

**06.09.2011 US 201161531526 P**

**22.02.2012 US 201213402719**

45 Fecha de publicación y mención en BOPI de la traducción de la patente:  
**25.05.2022**

73 Titular/es:

**QUALCOMM INCORPORATED (100.0%)  
Attn: International IP Administration 5775  
Morehouse Drive  
San Diego, California 92121-1714, US**

72 Inventor/es:

**ZHENG, YUNFEI;  
CHIEN, WEI-JUNG y  
KARCZEWICZ, MARTA**

74 Agente/Representante:

**ISERN JARA, Jorge**

ES 2 912 367 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**

Jerarquía de bloques de vídeo de predicción de movimiento

5 Campo técnico

Esta divulgación se refiere al cifrado de vídeo.

Antecedentes

10 Las capacidades de vídeo digital se pueden incorporar en un amplio rango de dispositivos, incluidos televisores digitales, sistemas de emisión directa digital, sistemas de emisión inalámbricos, asistentes digitales personales (PDA), ordenadores portátiles o de escritorio, tabletas, lectores de libros electrónicos, cámaras digitales, dispositivos de grabación digital, reproductores de medios digitales, dispositivos de videojuegos, consolas de videojuegos, teléfonos celulares o de radio satelital, los llamados "teléfonos inteligentes", dispositivos de vídeo teleconferencia, dispositivos de transmisión de vídeo y similares. Los dispositivos de vídeo digital implementan técnicas de compresión de vídeo, tales como las descritas en los estándares definidos por MPEG-2, MPEG-4, ITU-T H.263, ITU-T H.264/MPEG-4, parte 10, cifrado de vídeo avanzada (AVC), el estándar de cifrado de vídeo de alta eficiencia (HEVC) actualmente en desarrollo, y extensiones de dichos estándares. Los dispositivos de vídeo pueden transmitir, recibir, codificar, decodificar y/o almacenar información de vídeo digital de manera más eficiente implementando dichas técnicas de compresión de vídeo.

25 La propuesta JCTVC-A125 titulada "BBC's Response to the Call for Proposals on Video Compression Technology" por T Davies, investigación y desarrollo de la BBC en la reunión de JCT-VC, 15-4-2010 en Dresden describe una estructura general para la representación de contenido de vídeo, predicción inter/intra, filtrado en bucle y cifrado de entropía. El documento US2009/304084 describe la predicción combinada del vector de movimiento y el índice de referencia para el cifrado de vídeo. El documento WO2011/001078 describe la predicción de un vector de movimiento de una partición de imagen actual que tiene una forma o tamaño geométrico diferente al de al menos una partición de imagen de referencia adyacente y la codificación y decodificación usando una de tales predicciones. El documento EP 2106146 describe la codificación y decodificación de información de vectores de movimiento.

Resumen

35 La invención se define en las reivindicaciones independientes adjuntas. Las características opcionales se definen en las reivindicaciones dependientes. La divulgación que permite la invención tal como se define en las reivindicaciones se encuentra en las realizaciones descritas en relación con las figuras 8 y 9 y los párrafos correspondientes. Las realizaciones restantes deben entenderse como ejemplos ilustrativos que no describen partes de la presente invención tal como se define en las reivindicaciones.

40 Los detalles de uno o más ejemplos se exponen en los dibujos adjuntos y la descripción a continuación. Otras características, objetos y ventajas serán evidentes a partir de la descripción y los dibujos, y a partir de las reivindicaciones.

Breve descripción de los dibujos

45 La figura 1 es un diagrama de bloques que ilustra un ejemplo de un sistema de codificación y decodificación de vídeo que puede implementar técnicas de esta divulgación.

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

La figura 3A es un diagrama conceptual que ilustra un bloque de vídeo actual y un conjunto de ejemplo de bloques de vídeo de predicción de movimiento.

55 La figura 3B es un diagrama conceptual que ilustra un bloque de vídeo actual y un conjunto de ejemplo de bloques de vídeo de predicción de movimiento.

60 La figura 3C es un diagrama conceptual que ilustra un bloque de vídeo actual y un conjunto de ejemplo de bloques de vídeo de predicción de movimiento.

La figura 4 es un diagrama conceptual que ilustra la relación temporal entre un cuadro de vídeo actual y cuadros de vídeo de referencia.

65 La figura 5 es un diagrama conceptual que ilustra la ordenación de un conjunto de bloques de vídeo de predicción de movimiento con base en una relación temporal con un bloque de vídeo actual.

La figura 6 es un diagrama conceptual que ilustra ejemplos de bloques de vídeo candidatos que pueden usarse para generar un conjunto de bloques de vídeo de predicción de movimiento.

5 La figura 7 es un diagrama conceptual que ilustra un ejemplo de un método para buscar bloques de vídeo candidatos con base en criterios para generar un conjunto de bloques de vídeo de predicción de movimiento.

La figura 8 es un diagrama conceptual que ilustra ejemplos de particiones de bloques de vídeo.

10 La figura 9A-9K son diagramas conceptuales que ilustran ejemplos que crean una jerarquía ordenada con base en una partición de bloque de vídeo actual para un conjunto de bloques de vídeo de predicción de movimiento.

La figura 10 es un diagrama de flujo que ilustra una técnica para codificar datos de vídeo.

15 La figura 11 es un diagrama de bloques que ilustra un ejemplo de una unidad decodificadora de vídeo que puede implementar técnicas de esta divulgación.

La figura 12 es un diagrama de flujo que ilustra una técnica para decodificar datos de vídeo.

20 Descripción detallada

Esta divulgación describe técnicas para generar conjuntos de bloques de vídeo de predicción de movimiento a partir de bloques de vídeo candidatos y crear una jerarquía ordenada de los bloques de vídeo de predicción de movimiento dentro de un conjunto. Un codificador de vídeo puede codificar información de movimiento para un bloque de vídeo actual usando la jerarquía ordenada y un decodificador de vídeo puede decodificar información de movimiento para un bloque de vídeo actual usando la jerarquía ordenada. Por ejemplo, un conjunto de bloques de vídeo candidatos puede incluir bloques de vídeo adyacentes a un bloque de vídeo actual. Un conjunto ordenado de bloques de vídeo de predicción de movimiento puede ser un subconjunto de los bloques de vídeo adyacentes. La información de movimiento para un bloque de vídeo actual se puede obtener utilizando el conjunto ordenado de bloques de vídeo de predicción de movimiento, incluido el uso de una de las siguientes técnicas: heredar un vector de movimiento de un bloque de vídeo de predicción de movimiento en el conjunto ordenado, calcular un vector de movimiento sumando o restando información de vector de movimiento residual con el vector de movimiento de un bloque de vídeo de predicción de movimiento en el conjunto ordenado, o calculando un vector de movimiento usando la información de vector de movimiento de uno o más bloques de vídeo de predicción de movimiento en el conjunto ordenado. El uso de una jerarquía ordenada puede permitir ahorrar bits.

35 En una técnica, se genera un conjunto de bloques de vídeo de predicción de movimiento analizando si los bloques de vídeo candidatos incluyen criterios específicos. Por ejemplo, se puede analizar un número de bloques de vídeo dentro de una distancia temporal o espacial específica a un bloque de vídeo actual para determinar si alguno de sus valores de identificación de referencia está dentro de un rango específico. En este ejemplo, los bloques de vídeo candidatos con un valor de identificación de referencia igual a un valor especificado pueden incluirse en un conjunto de bloques de vídeo de predicción de movimiento.

40 En una técnica descrita por esta divulgación, un conjunto de bloques de vídeo de predicción de movimiento se organiza en una jerarquía ordenada con base en una distancia temporal de un bloque de referencia asociado con cada bloque de predicción de movimiento y un bloque de referencia asociado con el bloque de vídeo actual que se está codificando. En otras palabras, los vectores de movimiento de bloques de predicción de movimiento que apuntan a bloques predictivos temporalmente más cercanos al bloque de vídeo actual pueden tener prioridad sobre los vectores de movimiento de bloques de predicción de movimiento que apuntan a bloques predictivos temporalmente más alejados del bloque de vídeo actual.

45 En una técnica descrita por esta divulgación, se crea una jerarquía ordenada de bloques de vídeo de predicción de movimiento con base en el tipo de partición de un bloque de vídeo actual. Se puede generar un subconjunto de tres bloques de predicción de movimiento a partir de un conjunto de cinco bloques de vídeo adyacentes con base en tipo de partición de un bloque de vídeo actual. Un tipo de partición puede corresponder a la forma de partición de una denominada PU consistente con el estándar emergente de cifrado de vídeo de alta eficiencia (HEVC).

50 La figura 1 es un diagrama de bloques que ilustra un ejemplo de un sistema 10 de codificación y decodificación de vídeo que puede implementar técnicas de esta divulgación. Como se muestra en la figura 1, el sistema 10 incluye un dispositivo 12 fuente que transmite vídeo codificado a un dispositivo 16 de destino a través de un canal 15 de comunicación. El dispositivo 12 fuente y el dispositivo 16 de destino pueden comprender cualquiera de un amplio rango de dispositivos. En algunos casos, el dispositivo 12 fuente y el dispositivo 16 de destino pueden comprender dispositivos de comunicación inalámbricos, tales como los llamados radioteléfonos celulares o satelitales. Las técnicas de esta divulgación, sin embargo, que se aplican generalmente a la codificación y decodificación pueden aplicarse a dispositivos no inalámbricos que incluyen capacidades de codificación y/o decodificación de vídeo. El dispositivo 12 fuente y el dispositivo 16 de destino son simplemente ejemplos de dispositivos de cifrado que pueden soportar las técnicas descritas en este documento.

En el ejemplo de la figura 1, el dispositivo 12 fuente puede incluir una fuente 20 de vídeo, un codificador 22 de vídeo, un modulador/demodulador (módem) 23 y un transmisor 24. El dispositivo 16 de destino puede incluir un receptor 26, un módem 27, un decodificador 28 de vídeo y un dispositivo 30 de visualización. Los elementos de sintaxis pueden generarse en el codificador 22 de vídeo como parte de un flujo de bits codificado, y los elementos de sintaxis pueden usarse por el decodificador 28 de vídeo para decodificar el flujo de bits.

La fuente 20 de vídeo puede comprender un dispositivo de captura de vídeo, tal como una cámara de vídeo, un archivo de vídeo que contiene el vídeo capturado previamente, una transmisión de vídeo de un proveedor de contenido de vídeo u otra fuente de vídeo. Como alternativa adicional, la fuente 20 de vídeo puede generar datos basados en gráficos de ordenador como la fuente de vídeo, o una combinación de vídeo en vivo, vídeo archivado y vídeo generado por ordenador. En algunos casos, si la fuente 20 de vídeo es una cámara de vídeo, el dispositivo 12 fuente y el dispositivo 16 de destino pueden formar los denominados teléfonos con cámara o vídeo teléfonos. En cada caso, el vídeo capturado, precapturado o generado por ordenador puede ser codificado por el codificador 22 de vídeo.

En algunos ejemplos (pero no en todos los casos), una vez que el codificador 22 de vídeo codifica los datos de vídeo, el módem 23 puede modular la información de vídeo codificada de acuerdo con un estándar de comunicación, por ejemplo, tal como acceso múltiple por división de código (CDMA), multiplexación por división de frecuencia ortogonal (OFDM) o cualquier otro estándar o técnica de comunicación. Los datos codificados y modulados pueden luego transmitirse al dispositivo 16 de destino a través del transmisor 24. El módem 23 puede incluir diversos mezcladores, filtros, amplificadores u otros componentes diseñados para la modulación de señales. El transmisor 24 puede incluir circuitos diseñados para transmitir datos, incluidos amplificadores, filtros y una o más antenas. El receptor 26 del dispositivo 16 de destino recibe información por el canal 15 y el módem 27 demodula la información. El proceso de decodificación de vídeo realizado por el decodificador 28 de vídeo puede incluir técnicas recíprocas a las técnicas de codificación realizadas por el codificador 22 de vídeo.

El canal 15 de comunicación puede comprender cualquier medio de comunicación inalámbrico o por cable, tal como un espectro de radiofrecuencia (RF) o una o más líneas de transmisión física, o cualquier combinación de medios inalámbricos y por cable. El canal 15 de comunicación puede formar parte de una red con base en paquetes, tal como una red de área local, una red de área amplia o una red global tal como internet. El canal 15 de comunicación generalmente representa cualquier medio de comunicación adecuado, o colección de diferentes medios de comunicación, para transmitir datos de vídeo desde el dispositivo 12 fuente al dispositivo 16 de destino. De nuevo, la figura 1 es meramente un ejemplo y las técnicas de esta divulgación pueden aplicarse a configuraciones de cifrado de vídeo (por ejemplo, codificación de vídeo o decodificación de vídeo) que no necesariamente incluyen ninguna comunicación de datos entre los dispositivos de codificación y decodificación. En otros ejemplos, los datos podrían recuperarse de una memoria local, transmitirse a través de una red o similar. Un dispositivo de codificación puede codificar y almacenar datos en la memoria, y/o un dispositivo de decodificación puede recuperar y decodificar datos de la memoria. En muchos casos, la codificación y decodificación la realizan dispositivos no relacionados que no se comunican entre sí, sino que simplemente codifican datos en la memoria y/o recuperan y decodifican datos de la memoria.

En algunos casos, el codificador 22 de vídeo y el decodificador 28 de vídeo pueden operar sustancialmente de acuerdo con un estándar de compresión de vídeo tal como el estándar emergente HEVC. Sin embargo, las técnicas de esta divulgación también pueden aplicarse en el contexto de una variedad de otros estándares de cifrado de vídeo, incluidos algunos estándares antiguos o estándares nuevos o emergentes. Aunque no se muestra en la figura 1, en algunos casos, el codificador 22 de vídeo y el decodificador 28 de vídeo pueden estar integrados cada uno con un codificador y decodificador de audio, y pueden incluir unidades MUX-DEMUX apropiadas, u otro hardware y software, para manejar la codificación de audio y vídeo en un flujo de datos común o flujos de datos separados. Si corresponde, las unidades MUX-DEMUX pueden cumplir con el protocolo multiplexor ITU H.223 u otros protocolos, tal como el protocolo de datagramas de usuario (UDP).

El codificador 22 de vídeo y el decodificador 28 de vídeo pueden implementarse cada uno como uno o más microprocesadores, procesadores de señales digitales (DSP), circuitos integrados específicos de la aplicación (ASIC), arreglos de puertos programables en campo (FPGA), lógica discreta, software, hardware, firmware o combinaciones de los mismos. Cada codificador 22 de vídeo y decodificador 28 de vídeo puede incluirse en uno o más codificadores o decodificadores, cualquiera de los cuales puede integrarse como parte de un codificador/descodificador combinado (CODEC) en un dispositivo móvil respectivo, dispositivo de suscriptor, dispositivo de emisión, servidor, o similar. En esta divulgación, la expresión cifrador se refiere a un codificador, decodificador o CODEC, y las expresiones cifrador, codificador, decodificador y CODEC se refieren a máquinas específicas diseñadas para el cifrado (codificación y/o decodificación) de datos de vídeo consistente con esta divulgación.

En algunos casos, los dispositivos 12, 16 pueden operar de manera sustancialmente simétrica. Por ejemplo, cada uno de los dispositivos 12, 16 puede incluir componentes de codificación y decodificación de vídeo. Por lo tanto, el

sistema 10 puede soportar la transmisión de vídeo unidireccional o bidireccional entre los dispositivos 12, 16, de vídeo por ejemplo, para transmisión de vídeo, reproducción de vídeo, emisión de vídeo o vídeo de telefonía.

5 Durante el proceso de codificación, el codificador 22 de vídeo puede ejecutar un número de técnicas u operaciones de cifrado. En general, el codificador 22 de vídeo opera en bloques de datos de vídeo consistentes con el estándar HEVC. HEVC se refiere a unidades de cifrado (CU), que se pueden dividir de acuerdo con un esquema de división de cuatro quadrees. Una "LCU" se refiere a la unidad de cifrado de mayor tamaño (por ejemplo, la "unidad de cifrado más grande") soportada en una situación determinada. El tamaño de la LCU puede señalizarse como parte del flujo de bits, por ejemplo, como sintaxis de nivel de secuencia. La LCU se puede dividir en CU más pequeñas. 10 Las CU pueden dividirse en unidades de predicción (PU) con fines de predicción. Las PU pueden tener formas cuadradas o rectangulares. Las transformaciones no son fijas en el estándar HEVC emergente, sino que se definen de acuerdo con los tamaños de las unidades de transformación (TU), que pueden ser del mismo tamaño que una CU dada, o posiblemente más pequeñas. Las muestras residuales correspondientes a una CU se pueden subdividir en unidades más pequeñas utilizando una estructura quadtree conocida como "quadtree residual" (RQT). Los nodos hoja del RQT pueden denominarse unidades de transformación (TU). Las TU se pueden transformar y cuantificar. 15 Los elementos de sintaxis se pueden definir en el nivel LCU, el nivel CU, el nivel PU y el nivel TU. Los elementos denominados "banderas divididas" pueden incluirse como sintaxis de nivel de CU para indicar si una CU determinada se subdivide en cuatro CU más. Por ejemplo, CU0 puede referirse a la LCU, y CU1 a CU4 pueden comprender sub-CU de LCU.

20 Consistente con HEVC, los bloques de vídeo se denominan unidades de cifrado (CU) y existen muchas CU dentro de cuadros de vídeo individuales (u otras unidades de vídeo definidas independientemente, tales como segmentos). Los cuadros, segmentos, porciones de cuadros, grupos de imágenes u otras estructuras de datos pueden definirse como unidades de información de vídeo que incluyen una pluralidad de CU. Las CU pueden tener diferentes 25 tamaños consistente con el estándar HEVC, y el flujo de bits puede definir las unidades de cifrado más grandes (LCU) como el tamaño más grande de CU. Con el estándar HEVC, las LCU se pueden dividir en CU cada vez más pequeñas de acuerdo con un esquema de partición de quadtree, y las diferentes CU que se definen en el esquema se pueden dividir en las denominadas unidades de predicción (PU). Las LCU, CU y PU son bloques de vídeo en el sentido de esta divulgación.

30 El codificador 22 de vídeo puede realizar un cifrado predictivo en el que un bloque de vídeo que se codifica (por ejemplo, una PU de una CU dentro de una LCU) se compara con uno o más candidatos predictivos para identificar un bloque predictivo. Este proceso de cifrado predictivo puede ser intra (en cuyo caso, los datos predictivos se generan con base en los datos intra vecinos dentro del mismo cuadro o segmento de vídeo) o inter (en cuyo caso, 35 los datos predictivos se generan con base en los datos de vídeo en cuadros o segmentos anteriores o posteriores). Se pueden soportar muchos modos de cifrado diferentes, y el codificador 22 de vídeo puede seleccionar un modo de cifrado de vídeo deseable. De acuerdo con esta divulgación, al menos algunos bloques de vídeo pueden codificarse usando los procesos descritos en este documento. Las técnicas de compresión de vídeo incluyen predicción espacial (intra-imagen) y/o predicción temporal (inter-imagen) para reducir o eliminar la redundancia inherente a las 40 secuencias de vídeo. Para el cifrado de vídeo basado en bloques, un segmento de vídeo (por ejemplo, un cuadro de vídeo o una porción de un cuadro de vídeo) puede dividirse en bloques de vídeo, que también pueden denominarse bloques de árboles, unidades de cifrado (CU) y/o nodos de cifrado. Los bloques de vídeo en un segmento intracodificado (I) de una imagen se codifican usando predicción espacial con respecto a muestras de referencia en bloques vecinos en la misma imagen. Los bloques de vídeo en un segmento intercodificado (P o B) de una imagen 45 pueden usar predicción espacial con respecto a muestras de referencia en bloques vecinos en la misma imagen o predicción temporal con respecto a muestras de referencia en otras imágenes de referencia. Las imágenes pueden denominarse cuadros, y las imágenes de referencia pueden denominarse cuadros de referencia.

50 Cuando el codificador 22 de vídeo usa estimación de movimiento y compensación de movimiento para reducir la redundancia temporal en una secuencia de vídeo, se puede generar un vector de movimiento para identificar un bloque predictivo de datos de vídeo. El decodificador 28 de vídeo puede usar el vector de movimiento para predecir los valores del bloque de vídeo actual que se está codificando. Por ejemplo, los valores del bloque de vídeo predictivo pueden restarse de los valores del bloque de vídeo actual para producir un bloque de datos residuales. El vector de movimiento, junto con los datos residuales, se puede comunicar desde el codificador 22 de vídeo al 55 decodificador de vídeo a través del canal 15 de comunicación. El decodificador 28 de vídeo puede ubicar el mismo bloque predictivo (con base en el vector de movimiento) y reconstruir el bloque de vídeo codificado al combinar los datos residuales con los datos del bloque predictivo.

60 El codificador 22 de vídeo puede usar el modo de fusión para codificar la información de movimiento de un bloque de vídeo actual. El modo de fusión es un modo de cifrado de vídeo en el que la información de movimiento (tales como vectores de movimiento, índices de cuadros de referencia, direcciones de predicción u otra información) de un bloque de vídeo vecino se hereda para un bloque de vídeo actual que se está codificando. Se puede usar un valor de índice para identificar al vecino del que el bloque de vídeo actual hereda su información de movimiento (por ejemplo, de arriba, a la derecha, izquierda, abajo a la izquierda o coubicado desde un cuadro temporalmente 65 adyacente). Otro caso en el que se usa el vector de movimiento de un bloque de vídeo vecino en el cifrado de un bloque de vídeo actual es la llamada predicción de vector de movimiento. En este caso, se aplica el cifrado

5 predictivo de vectores de movimiento para reducir la cantidad de datos necesarios para comunicar el vector de movimiento. Por ejemplo, en lugar de codificar y comunicar el propio vector de movimiento, el codificador 22 de vídeo puede codificar y comunicar una diferencia de vector de movimiento (MVD) con respecto a un vector de movimiento conocido (o cognoscible). En H.264/AVC, el vector de movimiento conocido, que se puede usar con el MVD para definir el vector de movimiento actual, se puede definir mediante el llamado predictor de vector de movimiento (MVP), que se deriva como la mediana de los vectores de movimiento asociados con los bloques vecinos.

10 El codificador 22 de vídeo puede usar predicción de vector de movimiento adaptable (AMVP) para codificar la información de movimiento de un bloque de vídeo actual. La AMVP crea un conjunto de candidatos de vector de movimiento al incluir varios bloques vecinos en direcciones espaciales y temporales como candidatos para el MVP. En AMVP, el codificador 22 de vídeo selecciona el bloque predictor más preciso del conjunto de candidatos con base en el análisis de la tasa de codificación y la distorsión (por ejemplo, utilizando el denominado análisis de coste de distorsión de tasa). Se puede transmitir un índice predictor de vector de movimiento (mvp\_idx) a un decodificador 15 28 de vídeo para informar al decodificador 28 de vídeo dónde ubicar el MVP. También se puede transmitir un MVD al decodificador 28 de vídeo. El decodificador 28 de vídeo puede combinar el MVD con el MVP (definido por el índice predictor de vector de movimiento) para generar el vector de movimiento para el bloque de vídeo actual.

20 Después de generar el bloque predictivo, las diferencias entre el bloque de vídeo actual que se codifica y el bloque predictivo se codifican como un bloque residual, y la sintaxis de predicción (tal como un vector de movimiento en el caso del cifrado interno, o un modo predictivo en el caso del intra cifrado) se usa para identificar el bloque predictivo. Además, con AMVP o el modo de fusión, el bloque vecino utilizado para identificar el bloque de predicción puede codificarse, por ejemplo, a través de un valor de índice que identifica a un vecino particular de acuerdo con una jerarquía ordenada como se describe en este documento.

25 En algunos casos, el bloque residual puede transformarse y cuantificarse. Las técnicas de transformación pueden comprender un proceso DCT o un proceso conceptualmente similar, transformaciones de números enteros, transformaciones de ondículas u otros tipos de transformaciones. En un proceso DCT, por ejemplo, el proceso de transformación convierte un conjunto de valores de píxeles (por ejemplo, valores de píxeles residuales) en coeficientes de transformación, que pueden representar la energía de los valores de píxeles en el dominio de la frecuencia. Los coeficientes de transformación pueden cuantificarse. En particular, la cuantificación se puede aplicar a los coeficientes de transformación y generalmente implica un proceso que limita el número de bits asociados con cualquier coeficiente de transformación dado. Más específicamente, la cuantificación puede aplicarse de acuerdo con un parámetro de cuantificación (QP) definido en el nivel de LCU. En consecuencia, se puede aplicar el mismo nivel de cuantificación a todos los coeficientes de transformación en las TU asociadas con diferentes PU de CU 30 dentro de una LCU. Sin embargo, en lugar de señalar el QP en sí mismo, se puede señalar un cambio o diferencia (es decir, un delta) en el QP con la LCU para indicar el cambio en QP en relación con el de una LCU anterior.

40 Después de la transformación y la cuantificación, el cifrado de entropía se puede realizar en los bloques de vídeo residuales transformados y cuantificados. Los elementos de sintaxis también pueden incluirse en el flujo de bits codificado por entropía. En general, el cifrado entrópico comprende uno o más procesos que comprimen colectivamente una secuencia de coeficientes de transformación cuantificados y/u otra información de sintaxis. Pueden realizarse técnicas de exploración sobre los coeficientes de transformación cuantificados para definir uno o más vectores de coeficientes unidimensionales serializados a partir de bloques de vídeo bidimensionales. A 45 continuación, los coeficientes escaneados se codifican en entropía junto con cualquier información de sintaxis, por ejemplo, mediante cifrado de longitud variable adaptable al contenido (CAVLC), cifrado aritmético binario adaptable al contexto (CABAC) u otro proceso de cifrado de entropía.

50 Como parte del proceso de codificación, los bloques de vídeo codificados pueden decodificarse para generar los datos de vídeo que se usan para el cifrado posterior con base en la predicción de los bloques de vídeo posteriores. Esto a menudo se denomina bucle de decodificación del proceso de codificación y, en general, imita la decodificación que realiza un dispositivo decodificador. En el bucle de decodificación de un codificador o un decodificador, se pueden usar técnicas de filtrado para mejorar la calidad del vídeo y, por ejemplo, suavizar los límites de los píxeles y posiblemente eliminar los artefactos del vídeo decodificado. Este filtrado puede ser en bucle 55 o post-bucle. Con el filtrado en bucle, el filtrado de datos de vídeo reconstruidos ocurre en el bucle de cifrado, lo que significa que los datos filtrados son almacenados por un codificador o decodificador para su uso posterior en la predicción de datos de imágenes posteriores. Por el contrario, con el filtrado posterior al bucle, el filtrado de datos de vídeo reconstruidos se produce fuera del bucle de cifrado, lo que significa que un codificador o decodificador almacena versiones sin filtrar de los datos para su uso posterior en la predicción de datos de imagen posteriores. El filtrado de bucle a menudo sigue un proceso de filtrado de desbloqueo separado, que generalmente aplica el filtrado 60 a los píxeles que están en o cerca de los límites de los bloques de vídeo adyacentes para eliminar los artefactos de bloqueo que se manifiestan en los límites de los bloques de vídeo.

65 La figura 2 es un diagrama de bloques que ilustra un codificador 50 de vídeo consistente con esta divulgación. El codificador 50 de vídeo puede corresponder al codificador 22 de vídeo del dispositivo 12, o un codificador de vídeo de un dispositivo diferente. Como se muestra en la figura 2, el codificador 50 de vídeo incluye una unidad 31 de

partición quadtree, una unidad 32 de codificación de predicción, una memoria 34, un módulo 38 de transformación, una unidad 40 de cuantificación, una unidad 42 de cuantificación inversa, un módulo 44 de transformación inversa, una unidad 46 de cifrado de entropía, una unidad 47 de filtro, que puede incluir filtros de desbloqueo y filtros posteriores al bucle y/o en bucle, agregador 48 y agregador 51. Los datos de vídeo codificados y la información de sintaxis que define la manera de la codificación pueden comunicarse a la unidad 46 de codificación de entropía, que realiza la codificación de entropía en el flujo de bits.

Como se muestra en la figura 2, la unidad 32 de codificación de predicción puede soportar una pluralidad de diferentes modos 35 de cifrado en la codificación de bloques de vídeo. Los modos 35 pueden incluir modos de inter cifrado que definen datos predictivos de diferentes cuadros (o segmentos) de vídeo. Los modos de inter cifrado pueden ser bipredictivos, lo que significa que se utilizan dos listas diferentes (por ejemplo, lista 0 y lista 1) de datos predictivos (y típicamente dos vectores de movimiento diferentes) para identificar los datos predictivos. Los modos de inter cifrado pueden ser alternativamente unipredictivos, lo que significa que se usa una lista (por ejemplo, la lista 0) de datos predictivos (y típicamente un vector de movimiento) para identificar los datos predictivos. Pueden realizarse interpolaciones, compensaciones u otras técnicas junto con la generación de datos predictivos. También se pueden soportar los llamados modos SKIP y modos DIRECT, que heredan información de movimiento asociada con un bloque coubicado de otro cuadro (o segmento). Los bloques de modo SKIP no incluyen ninguna información residual, mientras que los bloques de modo DIRECTO incluyen información residual.

Además, los modos 35 pueden incluir modos de inter cifrado, que definen datos predictivos con base en datos dentro del mismo cuadro (o segmento) de vídeo que el que está siendo codificado. Los modos de intra cifrado pueden incluir modos direccionales que definen datos predictivos con base en datos en una dirección particular dentro del mismo cuadro, así como modos DC y/o planos que definen datos predictivos con base en el promedio o promedio ponderado de datos vecinos. La unidad 32 de codificación de predicción puede seleccionar el modo para un bloque dado con base en algunos criterios, tal como con base en un análisis de distorsión de tasa o algunas características del bloque, tal como el tamaño del bloque, la textura u otras características.

Según esta divulgación, la unidad 32 de codificación de predicción soporta uno o más modos que realizan predicción de vector de movimiento adaptativo (AMVP) como se describió anteriormente o un modo de fusión como se describió anteriormente. En estos u otros casos, la información de movimiento se puede heredar de un bloque de la manera descrita en este documento y la señalización del bloque del que se produce dicha herencia se puede realizar de la manera descrita en este documento.

En general, durante el proceso de codificación, el codificador 50 de vídeo recibe datos de vídeo de entrada. La unidad 32 de codificación de predicción realiza técnicas de cifrado predictiva en bloques de vídeo (por ejemplo, CU y PU). La unidad 31 de partición quadtree puede dividir una LCU en CU y PU más pequeñas de acuerdo con la partición HEVC. Para el inter cifrado, la unidad 32 de codificación de predicción compara CU o PU con diversos candidatos predictivos en uno o más cuadros o segmentos de referencia de vídeo (por ejemplo, una o más "listas" de datos de referencia) para definir un bloque predictivo. Para el intra cifrado, la unidad 32 de codificación de predicción genera un bloque predictivo con base en datos vecinos dentro del mismo cuadro o segmento de vídeo. La unidad 32 de codificación de predicción genere el bloque de predicción y el agregador 48 resta el bloque de predicción de la CU o PU que se está codificando para generar un bloque residual. Nuevamente, al menos algunos bloques de vídeo pueden codificarse usando la AMVP descrita en este documento.

En algunos casos, la unidad 32 de codificación de predicción puede incluir una unidad de distorsión de tasa (R-D) que compara los resultados de cifrado de bloques de vídeo (por ejemplo, CU o PU) en diferentes modos. En este caso, la unidad 32 de codificación de predicción también puede incluir una unidad de selección de modo para analizar los resultados de cifrado en términos de tasa de cifrado (es decir, bits de cifrado necesarios para el bloque) y distorsión (por ejemplo, representación de la calidad de vídeo del bloque codificado en relación con el bloque original) para hacer selecciones de modo para bloques de vídeo. De esta forma, la unidad R-D puede proporcionar análisis de los resultados de diferentes modos para permitir que la unidad de selección de modo seleccione el modo deseado para diferentes bloques de vídeo. Consistente con esta divulgación, se puede seleccionar un modo que realiza AMVP cuando una unidad R-D lo identifica como el modo deseado para un bloque de vídeo dado, por ejemplo, debido a ganancias de cifrado o eficiencia de cifrado. Alternativamente, consistente con esta divulgación, se puede seleccionar un modo de fusión en el que la información de movimiento se hereda de un bloque vecino. En estos u otros ejemplos, se puede definir y utilizar un conjunto ordenado de vecinos en el cifrado consistente con esta divulgación.

Haciendo referencia de nuevo a la figura 2, después de que la unidad 32 de codificación de predicción genere el bloque de predicción, y después de que el agregador 48 reste el bloque de predicción del bloque de vídeo que se está codificando para generar un bloque residual de valores de píxeles residuales, la unidad 38 de transformación aplica una transformación al bloque residual. La transformación puede comprender una transformación de coseno discreta (DCT) o una transformación conceptualmente similar tal como la definida por el estándar ITU H.264 o el estándar HEVC. Se pueden definir las denominadas estructuras de "mariposa" para realizar las transformaciones, o también se podría usar la multiplicación basada en matrices. En algunos ejemplos, consistente con el estándar HEVC, el tamaño de la transformación puede variar para diferentes CU, por ejemplo, según el nivel de partición que

se produzca con respecto a una LCU determinada. Las unidades de transformación (TU) pueden definirse para establecer el tamaño de transformación aplicado por la unidad 38 de transformación. También podrían usarse transformaciones de ondas, transformaciones de enteros, transformaciones de subbanda u otros tipos de transformaciones. En cualquier caso, la unidad de transformación aplica la transformación al bloque residual, produciendo un bloque de coeficientes de transformación residual. La transformación, en general, puede convertir la información residual de un dominio de píxeles a un dominio de frecuencia.

A continuación, la unidad 40 de cuantificación cuantifica los coeficientes de transformación residuales para reducir aún más la tasa de bits. La unidad 40 de cuantificación, por ejemplo, puede limitar el número de bits utilizados para codificar cada uno de los coeficientes. En particular, la unidad 40 de cuantificación puede aplicar el QP delta definido para la LCU para definir el nivel de cuantificación a aplicar (tal como combinando el QP delta con el QP de la LCU anterior o algún otro QP conocido). Después de realizar la cuantificación en las muestras residuales, la unidad 46 de cifrado de entropía puede escanear y codificar los datos por entropía.

CAVLC es un tipo de técnica de cifrado de entropía soportada por el estándar ITU H.264 y el estándar HEVC emergente, que puede aplicarse de forma vectorial mediante la unidad 46 de cifrado de entropía. CAVLC utiliza tablas de cifrado de longitud variable (VLC) de una manera que comprime de manera efectiva "ejecuciones" serializadas de coeficientes y/o elementos de sintaxis. CABAC es otro tipo de técnica de cifrado de entropía soportada por el estándar ITU H.264 o el estándar HEVC, que se puede aplicar de forma vectorizada por la unidad 46 de cifrado de entropía. CABAC puede implicar varias etapas, incluida la binarización, la selección del modelo de contexto y el cifrado aritmético binario. En este caso, la unidad 46 de cifrado de entropía codifica coeficientes y elementos de sintaxis de acuerdo con CABAC. También existen muchos otros tipos de técnicas de cifrado de entropía, y es probable que surjan nuevas técnicas de cifrado de entropía en el futuro. Esta divulgación no se limita a ninguna técnica de cifrado de entropía específica.

Después del cifrado de entropía por parte de la unidad 46 de codificación de entropía, el vídeo codificado puede transmitirse a otro dispositivo o archivarse para su posterior transmisión o recuperación. El vídeo codificado puede comprender los vectores codificados por entropía y diversa información de sintaxis. El decodificador puede utilizar dicha información para configurar correctamente el proceso de decodificación. La unidad 42 de cuantificación inversa y la unidad 44 de transformación inversa aplican cuantificación inversa y transformación inversa, respectivamente, para reconstruir el bloque residual en el dominio de píxeles. El sumador 51 agrega el bloque residual reconstruido al bloque de predicción producido por la unidad 32 de codificación de predicción para producir un bloque de vídeo reconstruido para almacenamiento en la memoria 34. Sin embargo, antes de dicho almacenamiento, la unidad 47 de filtro puede aplicar filtrado al bloque de vídeo para mejorar la calidad del vídeo. El filtrado aplicado por la unidad 47 de filtro puede reducir los artefactos y suavizar los límites de los píxeles. Además, el filtrado puede mejorar la compresión al generar bloques de vídeo predictivos que comprenden coincidencias cercanas con los bloques de vídeo que se están codificando.

Según las técnicas descritas en este documento, la unidad 32 de codificación de predicción puede usar jerarquías ordenadas para identificar un bloque predictivo para codificar un bloque actual y/o puede generar un valor de índice que identifica un bloque predictivo particular de acuerdo con una jerarquía ordenada. La figura 3A es un diagrama conceptual que ilustra un bloque de vídeo actual y bloques de vídeo de movimiento predictivo candidatos (es decir, arriba (T), arriba a la derecha (TR), izquierda (L), abajo a la izquierda (BL) o cúbicados desde un cuadro temporalmente adyacente (Temp)) del cual un bloque de vídeo actual puede derivar información de movimiento. La figura 3B es un diagrama conceptual que ilustra un bloque de vídeo actual y uno de una pluralidad de conjuntos de bloques de vídeo de movimiento predictivo que pueden derivarse del conjunto de bloques de vídeo de movimiento predictivo candidatos en la figura 3A. La figura 3C es un diagrama conceptual que ilustra un bloque de vídeo actual y bloques de vídeo de movimiento predictivo candidatos (es decir, arriba (T), arriba a la izquierda (TL) arriba a la derecha (TR), izquierda (L) o abajo a la izquierda (BL) de la que un bloque de vídeo actual puede derivar información de movimiento.

Las figuras 4 y 5 son diagramas conceptuales que ilustran el uso de una jerarquía ordenada de bloques de vídeo de predicción de movimiento para identificar un bloque de vídeo predictivo para codificar un bloque de vídeo actual. En el ejemplo, ilustrado en las figuras 4 y 5, la distancia temporal entre un bloque de vídeo actual y cada uno de los bloques de vídeo de predicción de movimiento se usa para crear una jerarquía ordenada. La jerarquía ordenada puede ser creada por el codificador 50 de vídeo con base en los datos de vídeo de entrada o puede crearse de forma avanzada y almacenarse en la memoria 34. La creación de una jerarquía ordenada con base en la distancia temporal puede explotar el hecho de que los bloques de vídeo de predicción de movimiento tienen distancias temporales más cortas en relación con un bloque de vídeo actual pueden ser mejores predictores que los bloques de vídeo que tienen distancias temporales más largas. En el ejemplo ilustrado en las figuras 4 y 5, el conjunto de bloques de predicción de movimiento incluye los cinco bloques mostrados en la figura 3A. En otros ejemplos, el conjunto de bloques de vídeo de predicción de movimiento puede incluir más o menos bloques de vídeo de predicción de movimiento. El tamaño del conjunto y los bloques de vídeo de predicción de movimiento incluidos en el conjunto pueden variar para cada bloque de vídeo actual. Por ejemplo, se puede generar un conjunto de tres bloques de vídeo de predicción de movimiento usando los cinco bloques de vídeo que se muestran en la figura 5.

Se puede usar un conteo de órdenes de imagen (POC) asociado con la información de movimiento de un bloque de vídeo de predicción de movimiento para definir la distancia temporal entre cada uno de los bloques de vídeo de predicción de movimiento y el bloque de vídeo actual. En el ejemplo ilustrado en las figuras 4 y 5, el bloque de vídeo actual que se está codificando se ubica en el cuadro 5 (POC=5). La información de movimiento de los bloques de vídeo de predicción de movimiento apunta al cuadro 0 para el bloque L, cuadro 1 para el bloque BL, cuadro 2 para el bloque T, cuadro 3 para el bloque Temp y cuadro 4 para el bloque TR. En consecuencia, la jerarquía de los bloques de predicción de movimiento puede definirse como: bloque TR seguido del bloque Temp, seguido del bloque T, seguido del bloque BL, seguido del bloque L.

Como se describió anteriormente, la unidad 32 de codificación de predicción puede utilizar la jerarquía ordenada de ejemplo ilustrada en la figura 5 para codificar la información de movimiento para un bloque de vídeo actual. En un ejemplo, una jerarquía ordenada puede programarse por adelantado y almacenarse en la memoria 34. En otro ejemplo, el codificador 50 de vídeo puede generar jerarquías de forma adaptativa mediante el análisis de datos de vídeo. Una vez que se determina una jerarquía, a cada uno de los bloques de vídeo de predicción de movimiento se le pueden asignar palabras de código variables como valores de índice. Al bloque de vídeo de predicción de movimiento que tiene la probabilidad más alta de ser el bloque de vídeo de predicción de movimiento mejor clasificado para un bloque de vídeo actual determinado se le puede asignar la palabra de código más corta. En el ejemplo, mostrado en la figura 5 el bloque de vídeo TR puede tener la palabra de código más corta. Al asignar valores de índice de manera diferente, dependiendo de la jerarquía ordenada (por ejemplo, la distancia temporal de la información de movimiento), se pueden lograr ahorros de bits. En algunos casos, se pueden usar códigos de longitud variable para asignar códigos más cortos a bloques de vídeo de predicción de movimiento con una mejor correlación (por ejemplo, en términos de la distancia temporal de la información de movimiento). En otros casos, se pueden usar códigos fijos, pero se pueden excluir algunos bloques de vídeo de predicción de movimiento, lo que permite códigos fijos más cortos debido al uso de menos bloques de vídeo de predicción de movimiento.

La unidad 32 de codificación de predicción puede comparar la información de movimiento del bloque de vídeo actual con la información de movimiento de los bloques de predicción de movimiento en un conjunto y seleccionar un valor de índice para el vídeo actual, donde el valor de índice identifica uno de los bloques de vídeo de predicción de movimiento. Con base en el modo de cifrado, la información de movimiento del bloque de vídeo actual puede generarse usando el valor de índice para: heredar un vector de movimiento del bloque de vídeo de predicción de movimiento identificado o calcular un vector de movimiento sumando o restando información de vector de movimiento residual con el vector de movimiento del bloque de vídeo de predicción de movimiento identificado.

El método de ejemplo ilustrado en las figuras 4 y 5 pueden basarse en un escenario en el que el bloque de vídeo actual y el bloque de vídeo de predicción de movimiento usan un modo direccional unipredictivo. Sin embargo, el método de la figura 4 y 5 también pueden extenderse a escenarios bipredictivos, donde cada bloque de vídeo tiene dos vectores de movimiento, considerando la distancia combinada de los dos bloques predictivos diferentes de bloques de vídeo de predicción de movimiento codificados en modo bipredictivo, en relación con el bloque de vídeo actual. En algunos ejemplos, si cualquiera de los bloques de vídeo de predicción de movimiento tiene el mismo POC, entonces se puede usar un orden predeterminado o se pueden usar otros criterios para ordenar los bloques de vídeo de predicción de movimiento. En un ejemplo, el orden predeterminado puede ser el bloque T, seguido del bloque L, seguido del bloque Temp, seguido del bloque TR, seguido del bloque BL. Para un conjunto de cinco bloques, cualquiera de los 120 órdenes posibles se puede utilizar como orden predeterminado. Otros criterios que pueden usarse para determinar el orden pueden incluir: lista de referencia, índice de referencia, dirección de predicción, tamaño de bloque, tamaño de unidad de predicción, tipo de partición de predicción, índice de transformación, tamaño de transformación u otra información relacionada con bloques de vídeo. Por ejemplo, el orden puede basarse en si el tamaño o la forma del bloque de vídeo que se codifica es el tamaño o la forma de los bloques de vídeo de predicción de movimiento. Si uno o más bloques de vídeo de predicción de movimiento no se pueden ordenar únicamente según una característica temporal específica (por ejemplo, cada bloque de vídeo de predicción de movimiento hace referencia al mismo bloque predictivo), se puede usar un segundo criterio para realizar un análisis adicional o definir un orden predeterminado. El segundo criterio puede basarse en otras técnicas de pedido descritas en este documento. En la invención reivindicada, un conjunto de bloques de vídeo de predicción de movimiento se organiza en una jerarquía ordenada con base en una forma de partición del bloque de vídeo actual.

Según la divulgación, la unidad 32 de codificación de predicción puede usar otras técnicas para codificar la información de movimiento de un bloque de vídeo actual. La figura 6 es un diagrama conceptual que muestra un ejemplo de posibles bloques de vídeo que pueden analizarse con el fin de determinar cómo codificar la información de movimiento de un bloque de vídeo actual. En la figura 6, los bloques de vídeo se ubican en el mismo cuadro que el bloque de vídeo actual. En otro ejemplo, los bloques de vídeo candidatos también pueden ubicarse en cuadros diferentes (ya codificados/descodificados) que el bloque de vídeo actual. Por ejemplo, un bloque ubicado en el bloque de vídeo actual de uno o más cuadros previamente codificados también puede ser un bloque de vídeo candidato. La unidad 32 de codificación de predicción puede analizar la información de movimiento de los bloques de vídeo candidatos ilustrados en la figura 6.

La figura 7 es un diagrama conceptual que ilustra un ejemplo de un método para analizar bloques de vídeo con base en criterios para codificar información de movimiento para un bloque de vídeo actual. Esencialmente, el ejemplo ilustrado en la figura 7 ilustra una forma eficiente de comparar información de movimiento de un bloque de vídeo actual con bloques de vídeo que pueden usarse para codificar la información de movimiento de un bloque de vídeo actual. El ejemplo descrito según la figura 7 puede utilizarse para buscar conjuntos de diversos tamaños de bloques de vídeo de predicción de movimiento.

En el ejemplo mostrado en la figura 7, hay once bloques de vídeo de predicción de movimiento. Cada uno de los bloques de vídeo incluye un valor de modo de dirección (es decir, predicción única o predicción doble), un valor de lista de referencia y un valor de índice de referencia. Antes de comparar la información de movimiento de un bloque de vídeo actual con la información de movimiento de cada uno de los once bloques de vídeo de predicción de movimiento ilustrados en la figura 7, puede ocurrir una primera comparación del valor del modo de dirección, el valor de la lista de referencia y el valor del índice de referencia. Esto puede provocar que se produzcan menos comparaciones de información de movimiento. De esta manera, la unidad 32 de codificación de predicción puede buscar eficientemente un bloque de vídeo de predicción de movimiento para un bloque de vídeo actual. De acuerdo con el ejemplo mostrado en la figura 7, el valor de la lista de referencia y el valor del índice de referencia de un bloque de vídeo de predicción de movimiento pueden compararse con el bloque de vídeo actual. En el ejemplo ilustrado en la figura 7, se puede determinar si el vector de movimiento de un bloque de vídeo de predicción de movimiento predice a partir de la misma lista de referencia y el mismo índice de referencia que el vector de movimiento de los bloques de vídeo actuales. Como se ilustra en el ejemplo mostrado en la figura 7, asumiendo que un bloque de vídeo actual está siendo codificado en modo de predicción bidireccional, y los dos vectores de movimiento usados en este modo de predicción bidireccional apuntan a la lista de referencia L1 y al índice 0 de referencia. Se puede realizar una búsqueda de los bloques de vídeo de predicción de movimiento para encontrar bloques codificados en modo de bipredicción, y los dos vectores de movimiento usados en este modo de bipredicción apuntan a la lista de referencia L1 y al índice 0 de referencia.

En el método de búsqueda de ejemplo ilustrado en la figura 7, la búsqueda comienza desde la izquierda a lo largo de la dirección de búsqueda izquierda (de los bloques 0 a 4 de vídeo), si se encuentra una coincidencia (en este ejemplo, el bloque 2 de vídeo candidato es una coincidencia), se puede detener la búsqueda izquierda y se inicia una búsqueda de arriba desde los bloques 5 a 10 de vídeo candidatos a lo largo de la dirección de búsqueda de arriba. Una vez que se encuentra la primera coincidencia de los principales candidatos del bloque de vídeo (en este ejemplo, el bloque 6 de vídeo es una coincidencia), se puede detener la búsqueda principal. La información de movimiento del bloque de vídeo actual se puede comparar con la del bloque 2 de vídeo y el bloque 6 de vídeo. Este proceso se puede repetir según sea necesario. La información de movimiento unitario de un bloque de vídeo predictivo se encuentra dentro de un umbral de información de movimiento de un bloque de vídeo actual.

Cabe señalar que en el ejemplo que se muestra en la figura 7, si no se tuvo en cuenta la dirección de predicción, el conjunto de bloques de predicción de movimiento puede incluir el bloque 0 de vídeo (primera coincidencia en la búsqueda izquierda) y el bloque 6 de vídeo (primera coincidencia en la búsqueda de arriba). El bloque 0 de vídeo candidato puede, en última instancia, no ser útil para predecir la información del vector de movimiento del bloque de vídeo actual, porque está codificado en modo de predicción única.

Consistente con los ejemplos de esta divulgación, se pueden agregar criterios adicionales para analizar los bloques de vídeo candidatos. Además de la lista de referencia, el índice de referencia y la dirección de predicción, los criterios adicionales pueden incluir uno o más de tamaño de bloque, tamaño de unidad de predicción, tipo de partición de predicción, índice de transformación, tamaño de transformación u otra información relacionada con el bloque de vídeo.

La unidad 32 de codificación de predicción puede generar un valor de índice para informar a un decodificador dónde ubicar un bloque de predicción de movimiento (por ejemplo, arriba, a la izquierda o ubicado en el mismo lugar). Un decodificador puede realizar un proceso de búsqueda recíproca para determinar un bloque de vídeo de predicción de movimiento. De esta manera, un decodificador puede generar un vector de movimiento para un bloque de vídeo actual realizando una búsqueda de un subconjunto de bloques de vídeo. Con referencia a la figura 7, un valor de índice puede indicar un subconjunto de bloques de vídeo de predicción de movimiento (es decir, bloques 0 a 4 de vídeo o bloques 5 a 10 de vídeo) de un conjunto conocido. Usando el valor de índice, un decodificador puede comparar información tal como el valor del modo de dirección, un valor de lista de referencia y un valor de índice de referencia de los bloques de vídeo de predicción de movimiento con el bloque de vídeo actual. Si hay una "coincidencia", la información de movimiento para un bloque de vídeo actual puede generarse a partir del primer bloque de vídeo predictivo que produce una "coincidencia". De esta manera, la información de movimiento para un bloque de vídeo actual puede codificarse usando un valor de índice que identifica ese subconjunto. Esto puede lograr un ahorro significativo de bits en comparación con la generación de un valor de índice que identifica un bloque de vídeo predictivo.

Según la divulgación, la unidad 32 de codificación de predicción codifica la información de movimiento de un bloque de vídeo actual usando una jerarquía ordenada donde una jerarquía ordenada se basa en información de partición. La figura 8 es un diagrama conceptual que ilustra ejemplos de particiones de bloques de vídeo. Las formas de

partición de predicción que se muestran en la figura 8 son algunos ejemplos de formas de partición de predicción, que pueden definirse mediante una unidad de predicción (PU) y un índice de PU consistente con el estándar emergente de cifrado de vídeo de alta eficiencia (HEVC). Al codificar la información de movimiento de un bloque de vídeo actual utilizando la forma de partición (que puede definirse mediante una forma de PU y un índice que define el tamaño de la PU), los cinco bloques candidatos pueden ordenarse con mayor probabilidad a menor probabilidad. Donde la probabilidad corresponde a la probabilidad de que el vector de movimiento de uno de los cinco bloques de vídeo candidatos "coincida" con el vector de movimiento del bloque de vídeo actual. Los conjuntos ordenados se pueden programar de antemano y almacenar tanto en un codificador como en un decodificador.

Las figuras 9A - 9K son diagramas conceptuales que ilustran ejemplos de creación de una jerarquía ordenada con base en un tipo de partición para un conjunto de bloques de vídeo de predicción de movimiento. En los ejemplos ilustrados en las figuras 9A - 9K, el conjunto de bloques de predicción de movimiento incluye los cinco bloques que se muestran en la figura 1A. En otros ejemplos, el conjunto de bloques de vídeo de predicción de movimiento puede incluir más o menos bloques de vídeo de predicción de movimiento. Por ejemplo, un conjunto puede incluir tres bloques de vídeo de predicción de movimiento. Incluir menos bloques de vídeo de predicción de movimiento en un conjunto puede reducir la complejidad del cifrado.

En los ejemplos mostrados en las figuras 9A - 9K, los números dentro de los bloques de vídeo de predicción de movimiento representan el orden jerárquico de los bloques de vídeo de predicción de movimiento. Para cada forma de partición de predicción (que puede estar definida por una forma de PU y un índice que define un tamaño de PU), la ordenación puede programarse por adelantado y almacenarse tanto en el codificador como en el decodificador.

Por ejemplo, en la figura 9A una jerarquía de bloques de vídeo de predicción de movimiento cuando la forma de partición es  $2N \times 2N$  puede definirse como: bloque L, bloque T, bloque BL, bloque TR, Temp.

Por ejemplo, en la figura 9B una jerarquía de bloques de vídeo de predicción de movimiento cuando la forma de partición es  $2N \times N_0$  puede definirse como: bloque T, bloque L, bloque TR, bloque BL, Temp.

Por ejemplo, en la figura 9C una jerarquía de bloques de vídeo de predicción de movimiento cuando la forma de partición es  $N \times 2N_1$  puede definirse como: bloque L, bloque BL, bloque Temp, bloque TR, bloque T.

Por ejemplo, en la figura 9D una jerarquía de bloques de vídeo de predicción de movimiento cuando la forma de partición es  $N \times 2N_0$  puede definirse como: bloque L, bloque T, bloque BL, Temp, bloque TR.

Por ejemplo, en la figura 9E una jerarquía de bloques de vídeo de predicción de movimiento cuando la forma de partición es  $N \times 2N_1$  puede definirse como: bloque T, bloque TR, Temp, bloque BL, bloque L.

Por ejemplo, en la figura 9F una jerarquía de bloques de vídeo de predicción de movimiento cuando la forma de partición es  $N \times N_0$  puede definirse como: bloque L, bloque T, bloque BL, bloque TR, Temp.

Por ejemplo, en la figura 9G una jerarquía de bloques de vídeo de predicción de movimiento cuando la forma de partición es  $N \times N_2$  puede definirse como: bloque L, bloque BL, bloque T, bloque TR, Temp.

Por ejemplo, en la figura 9H una jerarquía de bloques de vídeo de predicción de movimiento cuando la forma de partición es  $N \times N_1$  puede definirse como: bloque T, bloque TR, bloque L, Temp, bloque BL.

Por ejemplo, en la figura 9I una jerarquía de bloques de vídeo de predicción de movimiento cuando la forma de partición es  $N \times N_3$  puede definirse como: bloque L, bloque T, Temp, bloque TR, bloque BL.

Por ejemplo, en la figura 9J una jerarquía de bloques de vídeo de predicción de movimiento cuando la forma de partición es  $N \times 2N_0$  puede definirse como: bloque TL, bloque T, bloque BL, Temp, bloque TR, bloque L.

Por ejemplo, en la figura 9K una jerarquía de bloques de vídeo de predicción de movimiento cuando la forma de partición es  $N \times 2N_1$  puede definirse como: bloque T, bloque TR, Temp, bloque BL, bloque TL, bloque L.

La unidad 32 de codificación de predicción puede codificar información de movimiento de un bloque de vídeo actual utilizando el orden jerárquico de ejemplo ilustrado en las figuras 9A-9I. La ordenación jerárquica de ejemplo se puede almacenar en la memoria 34. La figura 10 es un ejemplo de un diagrama de flujo que ilustra una técnica para codificar datos de vídeo usando el ejemplo de ordenación jerárquica en las figuras 9A-9I. Cabe señalar que aunque la figura 10 se describe junto con el codificador 50 de vídeo, los pasos descritos en la figura 10 puede ser realizado por otros dispositivos y componentes. En el paso 250, la unidad 32 de codificación de predicción obtiene un vector de movimiento para un bloque de vídeo actual. Como se describió anteriormente, un vector de movimiento indica un bloque de vídeo predictivo que puede usarse para codificar un bloque de vídeo actual. En el paso 252, la unidad 32 de codificación de predicción obtiene un tipo de partición para un bloque de vídeo actual. La unidad 32 de codificación de predicción puede recibir un valor de tipo de partición desde la unidad 31 de partición quadtree. En un ejemplo, un tipo de partición corresponde a uno de los tipos de partición descritos en las figuras 9A-9I.

En el paso 254, la unidad 32 de codificación de predicción selecciona uno de una pluralidad de conjuntos definidos de bloques de vídeo predictivo candidatos ordenados con base en el tipo de partición. Por ejemplo, si el tipo de partición es Nx2N\_1, el conjunto de bloques de vídeo predictivo candidatos ordenados puede definirse como: bloque T, bloque TR, Temp, bloque BL, bloque L. En el paso 256, la unidad 32 de codificación de predicción selecciona un bloque de vídeo predictivo del seleccionado de la pluralidad de conjuntos definidos de bloques de vídeo predictivos candidatos ordenados con base en el vector de movimiento del bloque de vídeo actual. En el paso 256, el vector de movimiento del bloque de vídeo actual puede compararse con cada uno de los vectores de movimiento de los bloques de vídeo candidatos dentro del conjunto. La comparación se puede realizar de forma similar a la búsqueda descrita de acuerdo con la figura 7, donde se realiza el número mínimo de comparaciones de vectores de movimiento y si se encuentra un vector de movimiento dentro de un umbral del vector de movimiento del bloque de vídeo actual, la búsqueda se completa. En el paso 258, la unidad 32 de codificación de predicción genera un valor de índice que identifica el bloque de vídeo predictivo seleccionado. Los valores de índice para cada uno de los bloques de vídeo predictivo pueden almacenarse en la memoria 34.

Las palabras de código de longitud variable se pueden usar como valores de índice para cada uno de los bloques de vídeo de predicción de movimiento. Al bloque de vídeo de predicción de movimiento que tiene la probabilidad más alta de ser el bloque de vídeo de predicción de movimiento mejor clasificado para un bloque de vídeo actual determinado se le puede asignar la palabra de código más corta. Mediante la asignación de valores de índice de longitud variable, se pueden lograr ahorros de bits. Un decodificador puede programarse para conocer la misma jerarquía y, por lo tanto, puede interpretar correctamente la palabra clave recibida para realizar la misma selección que un codificador. En un ejemplo, el bloque de vídeo predictivo de mayor clasificación en cada uno de los conjuntos definidos en las figuras 9A-9I puede tener asignado un valor de índice de un bit. Solo se considera un subconjunto (por ejemplo, los 3 principales de 5) en cualquier escenario dado, lo que reduce la complejidad del cifrado. De esta manera, si se codifican varios bloques de vídeo en un grupo usando solo el subconjunto de los bloques de vídeo, el número de valores de índice usados para codificar el grupo se reduce aún más. En este caso, un codificador 50 de vídeo señala un conjunto reducido de valores de índice para un grupo de bloques de vídeo codificados.

Los ejemplos de creación de una jerarquía ordenada de bloques de vídeo de predicción de movimiento con base en la forma de partición descritos en este documento se pueden utilizar junto con métodos para generar información de movimiento para un bloque de vídeo actual. Por ejemplo, un codificador puede generar un bloque de vídeo actual usando cualquiera de las siguientes técnicas: heredar un vector de movimiento del bloque de vídeo de predicción de movimiento identificado, calcular un vector de movimiento sumando o restando información de vector de movimiento residual con un vector de movimiento de un bloque de vídeo de predicción de movimiento identificado, o calcular un vector de movimiento usando la información de vector de movimiento de uno o más bloques de vídeo de predicción de movimiento de clasificación más alta seleccionando un vector de movimiento mediano o promediando vectores de movimiento.

La figura 11 es un diagrama de bloques que ilustra un ejemplo de un decodificador 60 de vídeo, que decodifica una secuencia de vídeo que está codificada de la manera descrita en este documento. Las técnicas de esta divulgación pueden ser realizadas por el decodificador 60 de vídeo en algunos ejemplos. En particular, el decodificador 60 de vídeo puede realizar una o más de las técnicas para determinar la información de movimiento para un bloque de vídeo actual descrito en este documento como parte de un proceso de decodificación.

El decodificador 60 de vídeo incluye una unidad 52 de decodificación de entropía, que realiza la función de decodificación recíproca de la codificación realizada por la unidad 46 de codificación de entropía de la figura 2. En particular, la unidad 52 de decodificación de entropía puede realizar la decodificación CAVLC o CABAC, o cualquier otro tipo de decodificación de entropía utilizada por el codificador 50 de vídeo. El decodificador 60 de vídeo también incluye una unidad 54 de decodificación de predicción, una unidad 56 de cuantificación inversa, una unidad 58 de transformación inversa, una memoria 62 y un sumador 64. En particular, al igual que el codificador 50 de vídeo, el decodificador 60 de vídeo incluye una unidad 54 de decodificación de predicción y una unidad 57 de filtro. La unidad 54 de decodificación de predicción del decodificador 60 de vídeo puede incluir una unidad 86 de compensación de movimiento, que decodifica bloques intercodificados y posiblemente incluye uno o más filtros de interpolación para la interpolación de subpíxeles en el proceso de compensación de movimiento. La unidad 54 de decodificación de predicción también puede incluir una unidad de predicción intra para decodificar modos intra. La unidad 54 de decodificación de predicción puede soportar una pluralidad de modos 35 que incluyen uno o más modos que soportan AMVP y/o uno o más modos de combinación. La unidad 57 de filtro puede filtrar la salida del sumador 64, y puede recibir información de filtro decodificada por entropía para definir los coeficientes de filtro aplicados en el filtrado de bucle.

Al recibir datos de vídeo codificados, la unidad 52 de decodificación de entropía realiza una decodificación recíproca a la codificación realizada por la unidad 46 de codificación de entropía (del codificador 50 en la figura 2). En el decodificador, la unidad 52 de decodificación de entropía analiza el flujo de bits para determinar las LCU y la partición correspondiente asociada con las LCU. En algunos ejemplos, una LCU o las CU de la LCU pueden definir modos de cifrado que se utilizaron, y estos modos de cifrado pueden incluir el modo de fusión bipredictivo. En

consecuencia, la unidad 52 de decodificación de entropía puede enviar la información de sintaxis a la unidad de predicción que identifica el modo de fusión bipredictivo.

La figura 12 es un ejemplo de un diagrama de flujo que ilustra una técnica para decodificar datos de vídeo usando el ejemplo de ordenación jerárquica en las figuras 9A-9I. Cabe señalar que aunque la figura 12 se describe junto con el codificador 60 de vídeo, los pasos descritos en la figura 12 puede ser realizado por otros dispositivos y componentes. En el paso 350, la unidad 54 de decodificación predictiva obtiene un valor de índice para un bloque de vídeo actual. Como se ha descrito anteriormente según la figura 10, un valor de índice indica un bloque de vídeo predictivo que puede usarse para generar un vector de movimiento para un bloque de vídeo actual. En el paso 352, la unidad 54 de decodificación predictiva obtiene un tipo de partición para un bloque de vídeo actual. En un ejemplo, un tipo de partición corresponde a uno de los tipos de partición descritos en las figuras 9A-9I. En el paso 354, la unidad 54 de decodificación predictiva selecciona uno de una pluralidad de conjuntos definidos de bloques de vídeo predictivo candidatos ordenados con base en tipo de partición. Por ejemplo, si el tipo de partición es  $Nx2N_1$ , el conjunto de bloques de vídeo predictivo candidatos ordenados puede definirse como: bloque T, bloque TR, Temp, bloque BL, bloque L. En el paso 356, la unidad 54 de decodificación predictiva selecciona un bloque de vídeo predictivo del seleccionado de la pluralidad de conjuntos definidos de bloques de vídeo predictivo candidatos ordenados con base en el valor de índice. En el paso 358, la unidad 54 de decodificación predictiva genera un vector de movimiento. Por ejemplo, la unidad 54 de decodificación predictiva puede generar un vector de movimiento usando cualquiera de las siguientes técnicas: heredar un vector de movimiento del bloque de vídeo de predicción de movimiento identificado, calcular un vector de movimiento sumando o restando información de vector de movimiento residual con un vector de movimiento de un bloque de vídeo de predicción de movimiento identificado, o calcular un vector de movimiento usando la información de vector de movimiento de uno o más bloques de vídeo de predicción de movimiento de clasificación más alta seleccionando un vector de movimiento mediano o promediando vectores de movimiento.

En el ejemplo descrito en la figura 12, el decodificador 60 puede programarse para conocer las jerarquías de forma de partición descritas en las figuras 9A-9I y, por lo tanto, puede interpretar adecuadamente el valor de índice recibido para hacer la misma selección de bloque de vídeo de predicción que el codificador 50. En un ejemplo, el bloque de vídeo predictivo de clasificación más alta en cada uno de los conjuntos definidos en las figuras 9A-9I puede tener asignado un valor de índice de un bit. En otros ejemplos, solo se puede considerar un subconjunto (por ejemplo, los 3 principales de 5) en cualquier escenario dado, lo que puede reducir la complejidad del cifrado. Por ejemplo, para una partición  $Nx2N_1$ , al conjunto de bloques de vídeo predictivo candidatos ordenados, bloque T, bloque TR y Temp, se le pueden asignar los siguientes valores de índice: 1, 01 y 00. De esta manera, si se codifican varios bloques de vídeo usando solo los valores de índice de los 3 bloques de vídeo principales pueden lograr más ahorros de bits. En este caso, un codificador 50 de vídeo puede señalar el número de bloques de vídeo de movimiento predictivo para un grupo de bloques de vídeo codificados. Esto se puede lograr de manera similar a como se señala el conjunto de bloques de la izquierda en el ejemplo descrito según la figura 7.

Debe reconocerse que, según el ejemplo, ciertos actos o eventos de cualquiera de las técnicas descritas en este documento se pueden realizar en una secuencia diferente, se pueden agregar, fusionar u omitir por completo (por ejemplo, no todos los actos o eventos descritos son necesarios para la práctica de las técnicas). Además, en ciertos ejemplos, los actos o eventos se pueden realizar simultáneamente, por ejemplo, a través del procesamiento de subprocesos múltiples, el procesamiento de interrupciones o múltiples procesadores, en lugar de secuencialmente.

En uno o más ejemplos, las funciones descritas pueden implementarse en hardware, software, firmware o cualquier combinación de los mismos. Si se implementan en software, las funciones pueden almacenarse o transmitirse como una o más instrucciones o código en un medio legible por ordenador y ser ejecutadas por una unidad de procesamiento con base en hardware. Los medios legibles por ordenador pueden incluir medios de almacenamiento legibles por ordenador, que corresponden a un medio tangible tales como medios de almacenamiento de datos o medios de comunicación que incluyen cualquier medio que facilite la transferencia de un programa de ordenador de un lugar a otro, por ejemplo, de acuerdo con un protocolo de comunicación. De esta manera, los medios legibles por ordenador generalmente pueden corresponder a (1) un medio de almacenamiento tangible legible por ordenador que no es transitorio o (2) un medio de comunicación tal como una señal o una onda portadora. Los medios de almacenamiento de datos pueden ser cualquier medio disponible al que puedan acceder uno o más ordenadores o uno o más procesadores para recuperar instrucciones, códigos y/o estructuras de datos para la implementación de las técnicas descritas en esta divulgación. Un producto de programa informático puede incluir un medio legible por ordenador.

A modo de ejemplo, y sin limitación, tales medios de almacenamiento legibles por ordenador pueden comprender RAM, ROM, EEPROM, CD-ROM u otro almacenamiento en disco óptico, almacenamiento en disco magnético u otros dispositivos de almacenamiento magnético, memoria flash o cualquier otro medio que puede usarse para almacenar el código de programa deseado en forma de instrucciones o estructuras de datos y al que puede acceder un ordenador. Además, cualquier conexión se denomina correctamente un medio legible por ordenador. Por ejemplo, si las instrucciones se transmiten desde un sitio web, servidor u otra fuente remota mediante un cable coaxial, cable de fibra óptica, par trenzado, línea de suscriptor digital (DSL) o tecnologías inalámbricas tales como infrarrojos, radio y microondas, entonces el cable coaxial, el cable de fibra óptica, el par trenzado, el DSL o las

5 tecnologías inalámbricas tales como infrarrojos, radio y microondas se incluyen en la definición de medio. Debe entenderse, 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 están dirigidos a medios de almacenamiento tangibles no transitorios. Disco y disco, como se usa en este 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 reproducen datos ópticamente con láser.

10 Las instrucciones pueden ser ejecutadas por uno o más procesadores, tales como uno o más procesadores de señales digitales (DSP), microprocesadores de uso general, circuitos integrados específicos de la aplicación (ASIC), arreglos lógicos programables en campo (FPGA) u otra circuitería lógica integrada o discreta equivalente. En consecuencia, la expresión "procesador", como se usa en este documento, puede referirse a cualquiera de las estructuras anteriores o a cualquier otra estructura adecuada para la implementación de las técnicas descritas en este documento. Además, en algunos aspectos, la funcionalidad descrita en este documento puede proporcionarse dentro de módulos de software y/o hardware dedicados configurados para codificar y decodificar, o incorporarse en un códec combinado. Además, las técnicas podrían implementarse completamente en uno o más circuitos o elementos lógicos.

20 Las técnicas de esta divulgación pueden implementarse en una amplia variedad de dispositivos o aparatos, incluidos un teléfono inalámbrico, un circuito integrado (IC) o un conjunto de IC (por ejemplo, un conjunto de chips). En esta divulgación se describen diversos componentes, módulos o unidades para enfatizar los aspectos funcionales de los dispositivos configurados para realizar las técnicas divulgadas, pero no necesariamente requieren la realización por diferentes unidades de hardware. Más bien, como se describe anteriormente, diversas unidades pueden combinarse en una unidad de hardware de códec o proporcionarse mediante una colección de unidades de hardware interoperativas, que incluyen uno o más procesadores como se describió anteriormente, junto con software y/o firmware adecuados.

25 Se han descrito diversos ejemplos.

**REIVINDICACIONES**

1. Un método para decodificar datos de vídeo de acuerdo con un modo de fusión, el método que comprende:

5 obtener (350) un valor de índice para un bloque de vídeo actual codificado en el modo de fusión, identificando el valor de índice, en un subconjunto de bloques de vídeo predictivo candidato, un bloque de vídeo predictivo seleccionado;  
 obtener una forma de unidad de predicción y un índice de unidad de predicción para el bloque de vídeo actual, en donde el índice de unidad de predicción especifica la ubicación relativa de la unidad de predicción dentro del bloque  
 10 de vídeo actual;  
 generar un conjunto ordenado de bloques de vídeo predictivos candidatos con base en vecinos espaciales y temporales del bloque de vídeo actual, en donde el orden de los bloques de vídeo predictivos candidatos se basa en la forma de la unidad de predicción y el índice de la unidad de predicción;  
 limitar el conjunto ordenado de bloques predictivos candidatos a un subconjunto de los bloques de vídeo predictivos  
 15 candidatos;  
 seleccionar un bloque de vídeo predictivo del subconjunto con base en el valor de índice; y  
 generar (358) información de movimiento para el bloque de vídeo actual de acuerdo con el modo de fusión con base en información de movimiento del bloque de vídeo predictivo, en donde generar la información de movimiento para el bloque de vídeo actual comprende heredar información de movimiento del bloque de vídeo predictivo.

20 2. El método de la reivindicación 1, en donde el subconjunto incluye tres bloques predictivos candidatos seleccionados de cinco bloques de vídeo posibles.

25 3. El método de la reivindicación 1 o 2, en donde el conjunto de bloques de vídeo predictivos candidatos incluye: un bloque de vídeo izquierdo adyacente al bloque de vídeo actual (L), un bloque de vídeo superior adyacente al bloque de vídeo actual (T), un bloque de vídeo superior derecho adyacente al bloque de vídeo actual (TR), un bloque de vídeo inferior izquierdo adyacente al bloque de vídeo actual (BL) y un bloque de vídeo temporal adyacente al bloque de vídeo actual (Temp).

30 4. El método de la reivindicación 1, en donde heredar la información de movimiento incluye heredar un vector de movimiento y un índice de cuadro de referencia del bloque de vídeo predictivo.

35 5. El método de la reivindicación 1, en donde el conjunto incluye: un bloque de vídeo de arriba izquierdo adyacente al bloque de vídeo actual (TL), un bloque de vídeo de arriba adyacente al bloque de vídeo actual (T), un bloque de vídeo de arriba derecho adyacente al bloque de vídeo actual (TR), un bloque de vídeo izquierdo adyacente al bloque de vídeo actual (L), un bloque de vídeo de abajo izquierdo adyacente al bloque de vídeo actual (BL), y un bloque de vídeo temporal adyacente al bloque de vídeo actual (Temp), en donde el bloque de vídeo de arriba adyacente (T) está ubicado a la izquierda del bloque de vídeo de arriba derecho adyacente (TR) y en donde el bloque de vídeo izquierdo adyacente (L) está ubicado en la parte de arriba del bloque de vídeo de abajo izquierdo adyacente (BL).

40 6. Un dispositivo para decodificar datos de vídeo de acuerdo con un modo de fusión, el dispositivo que comprende:

medios para obtener un valor de índice para un bloque de vídeo actual, identificando el valor de índice, en un subconjunto de bloques de vídeo predictivos candidatos, un bloque de vídeo predictivo seleccionado;  
 45 medios para obtener una forma de unidad de predicción y un índice de unidad de predicción para el bloque de vídeo actual, en donde el índice de unidad de predicción especifica la ubicación relativa de la unidad de predicción dentro del bloque de vídeo actual;  
 medios para generar un conjunto ordenado de bloques de vídeo predictivos candidatos con base en vecinos espaciales y temporales del bloque de vídeo actual, en donde el orden de los bloques de vídeo predictivos candidatos se basa en la forma de la unidad de predicción y el índice de la unidad de predicción;  
 50 medios configurados para limitar el conjunto de bloques de vídeo predictivos candidatos a un subconjunto de los bloques de vídeo predictivos candidatos;  
 medios para seleccionar un bloque de vídeo predictivo del subconjunto con base en el valor de índice;  
 medios para generar información de movimiento para el bloque de vídeo actual de acuerdo con el modo de fusión  
 55 con base en información de movimiento del bloque de vídeo predictivo, en donde generar información de movimiento para el bloque de vídeo actual comprende heredar información de movimiento del bloque de vídeo predictivo.

7. Un método para codificar datos de vídeo de acuerdo con un modo de fusión, el método que comprende:

60 obtener (250) un vector de movimiento para un bloque de vídeo actual;  
 obtener una forma de unidad de predicción y un índice de unidad de predicción para el bloque de vídeo actual, en donde el índice de unidad de predicción especifica la ubicación relativa de la unidad de predicción dentro del bloque de vídeo actual;  
 generar un conjunto ordenado de bloques de vídeo predictivos candidatos con base en vecinos espaciales y temporales del bloque de vídeo actual, en donde el orden de los bloques de vídeo predictivos candidatos se basa en la forma de la unidad de predicción y el índice de la unidad de predicción;

- limitar el conjunto ordenado de bloques de vídeo predictivos candidatos a un subconjunto de los bloques de vídeo predictivos candidatos;  
 seleccionar un bloque de vídeo predictivo del subconjunto con base en el vector de movimiento; y  
 generar (258) un valor de índice que identifique, en el subconjunto de bloques de vídeo predictivo candidatos, el  
 5 bloque de vídeo predictivo seleccionado.
8. El método de la reivindicación 7, en donde el subconjunto incluye tres bloques predictivos candidatos seleccionados de cinco bloques de vídeo posibles.
- 10 9. El método de la reivindicación 7 u 8, en donde los bloques de vídeo predictivos candidatos incluyen: un bloque de vídeo izquierdo adyacente al bloque de vídeo actual (L), un bloque de vídeo superior adyacente al bloque de vídeo actual (T), un bloque de vídeo superior derecho adyacente al bloque de vídeo actual (TR), un bloque de vídeo inferior izquierdo adyacente al bloque de vídeo actual (BL) y un bloque de vídeo temporal adyacente al bloque de vídeo actual (Temp).
- 15 10. El método de la reivindicación 7, en donde el conjunto incluye: un bloque de vídeo de arriba izquierdo adyacente al bloque de vídeo actual (TL), un bloque de vídeo de arriba adyacente al bloque de vídeo actual (T), un bloque de vídeo de arriba derecho adyacente al bloque de vídeo actual (TR), un bloque de vídeo izquierdo adyacente al bloque de vídeo actual (L), un bloque de vídeo de abajo izquierdo adyacente al bloque de vídeo actual (BL) y un bloque de vídeo temporal adyacente al bloque de vídeo actual (Temp), en donde el bloque de vídeo de arriba adyacente (T) está ubicado a la izquierda del bloque de vídeo de arriba derecho adyacente (TR) y en donde el bloque de vídeo izquierdo adyacente (L) está ubicado en la parte de arriba del bloque de vídeo de abajo izquierdo adyacente (BL).
- 20 11. Un dispositivo para codificar datos de vídeo de acuerdo con un modo de fusión, el dispositivo que comprende:
- 25 medios para obtener un vector de movimiento para un bloque de vídeo actual;  
 medios para obtener una forma de unidad de predicción y un índice de unidad de predicción para el bloque de vídeo actual, en donde el índice de unidad de predicción especifica la ubicación relativa de la unidad de predicción dentro del bloque de vídeo actual;
- 30 medios para generar un conjunto de bloques de vídeo predictivos candidatos con base en vecinos espaciales y temporales del bloque de vídeo actual, en donde el orden de los bloques de vídeo predictivos candidatos se basa en la forma de la unidad de predicción y el índice de la unidad de predicción;  
 medios configurados para limitar el conjunto ordenado de bloques de vídeo predictivos candidatos a un subconjunto de los bloques de vídeo predictivos candidatos;
- 35 medios para seleccionar un bloque de vídeo predictivo del subconjunto con base en el vector de movimiento; y  
 medios para generar un valor de índice que identifica, en el subconjunto de bloques de vídeo predictivo candidatos, el bloque de vídeo predictivo seleccionado.
- 40 12. El dispositivo de la reivindicación 11, en donde el subconjunto incluye tres bloques predictivos candidatos seleccionados de cinco bloques de vídeo posibles.
- 45 13. El dispositivo de la reivindicación 11 o 12, en donde los bloques de vídeo predictivos candidatos incluyen: un bloque de vídeo izquierdo adyacente al bloque de vídeo actual (L), un bloque de vídeo superior adyacente al bloque de vídeo actual (T), un bloque de vídeo superior derecho adyacente al bloque de vídeo actual (TR), un bloque de vídeo inferior izquierdo adyacente al bloque de vídeo actual (BL) y un bloque de vídeo temporal adyacente al bloque de vídeo actual (Temp).
- 50 14. El dispositivo de la reivindicación 11, en donde el conjunto incluye: un bloque de vídeo de arriba izquierdo adyacente al bloque de vídeo actual (TL), un bloque de vídeo de arriba adyacente al bloque de vídeo actual (T), un bloque de vídeo de arriba derecho adyacente al bloque de vídeo actual (TR), un bloque de vídeo izquierdo adyacente al bloque de vídeo actual (L), un bloque de vídeo de abajo izquierdo adyacente al bloque de vídeo actual (BL), y un bloque de vídeo temporal adyacente al bloque de vídeo actual (Temp), en donde el bloque de vídeo de arriba adyacente (T) está ubicado a la izquierda del bloque de vídeo de arriba derecho adyacente (TR) y en donde el bloque de vídeo izquierdo adyacente (L) está ubicado en la parte de arriba del bloque de vídeo de abajo izquierdo adyacente (BL).
- 55 15. Un producto de programa de ordenador que comprende un medio de almacenamiento legible por ordenador que tiene almacenadas instrucciones en el mismo que, cuando son ejecutadas por un procesador, hacen que el procesador de un dispositivo lleve a cabo el método de cualquiera de las reivindicaciones 1 a 5 o 7 a 10.
- 60

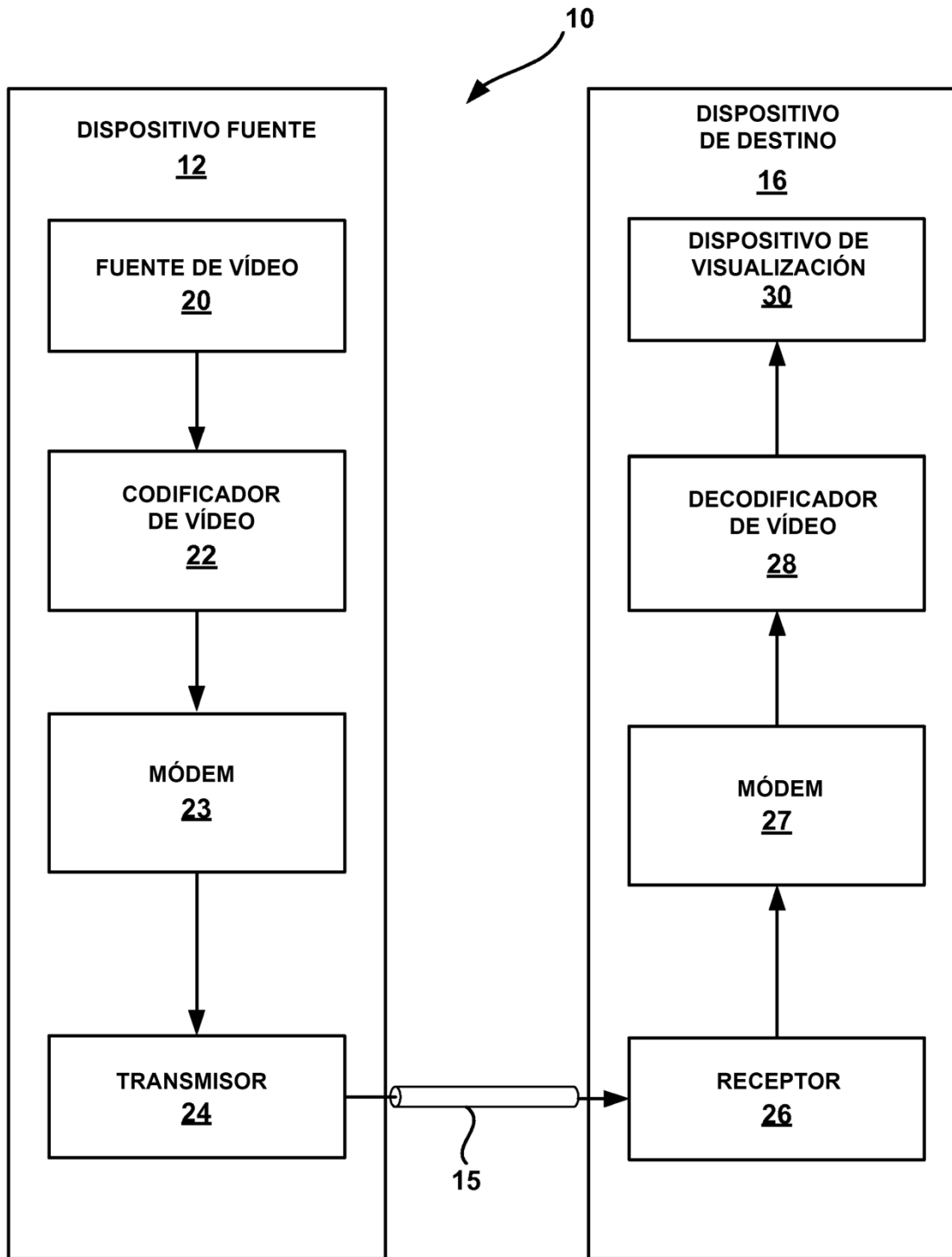


FIG. 1

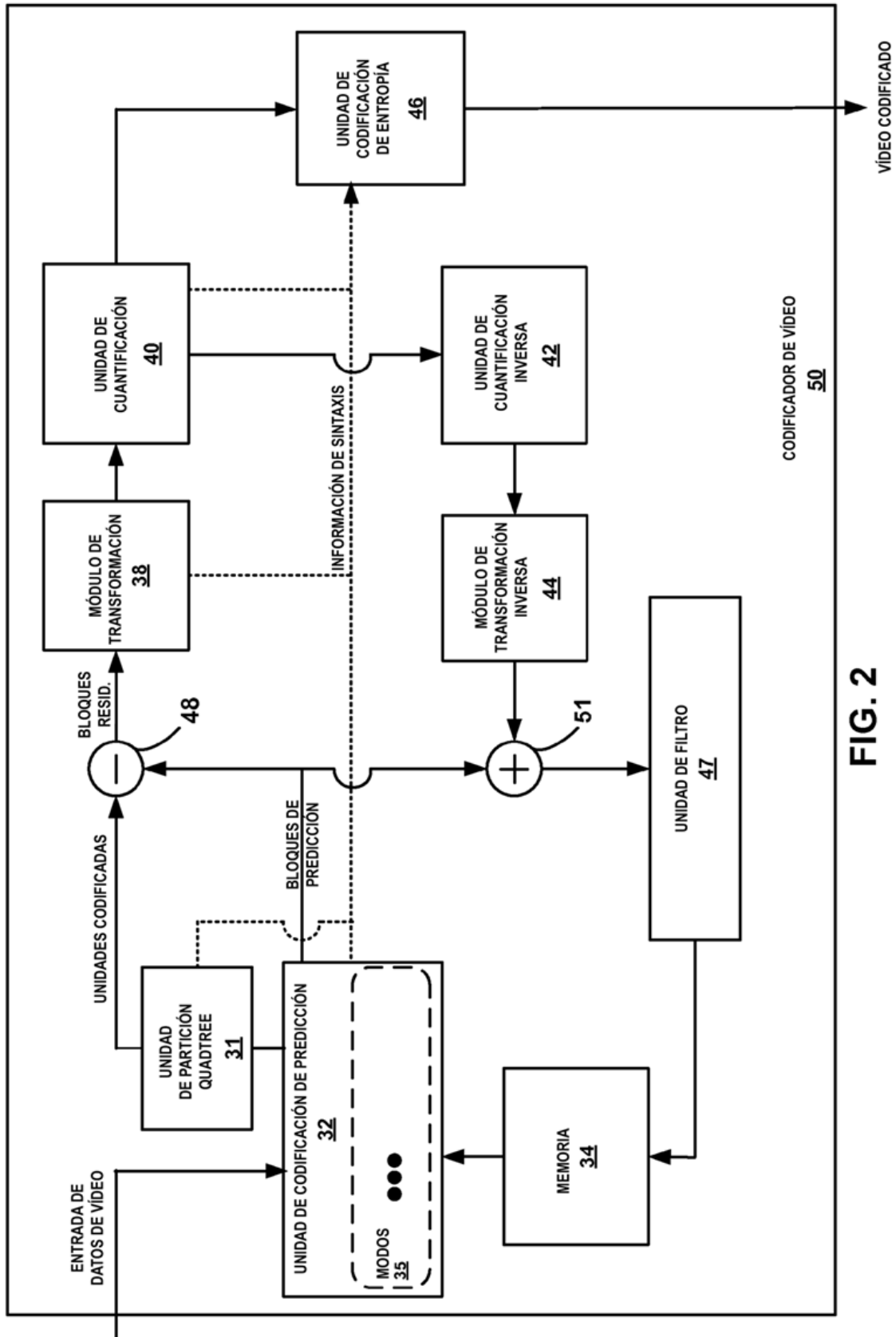
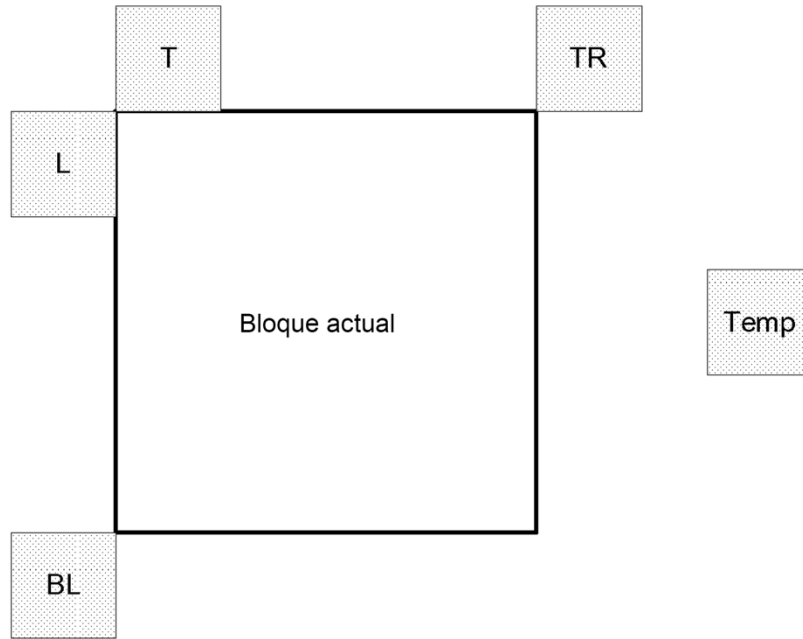
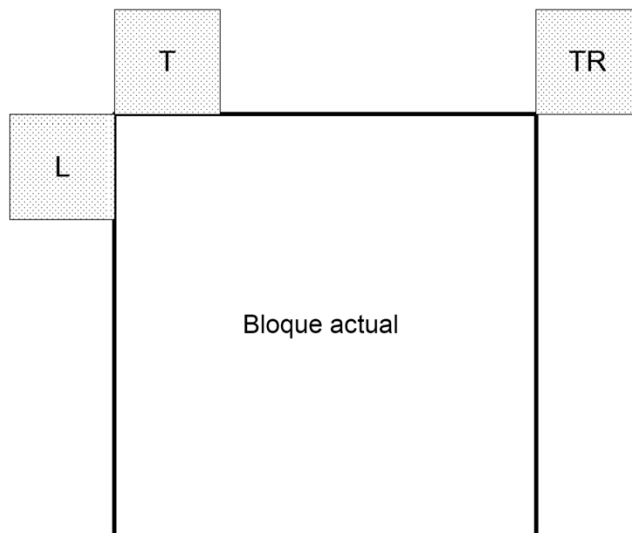


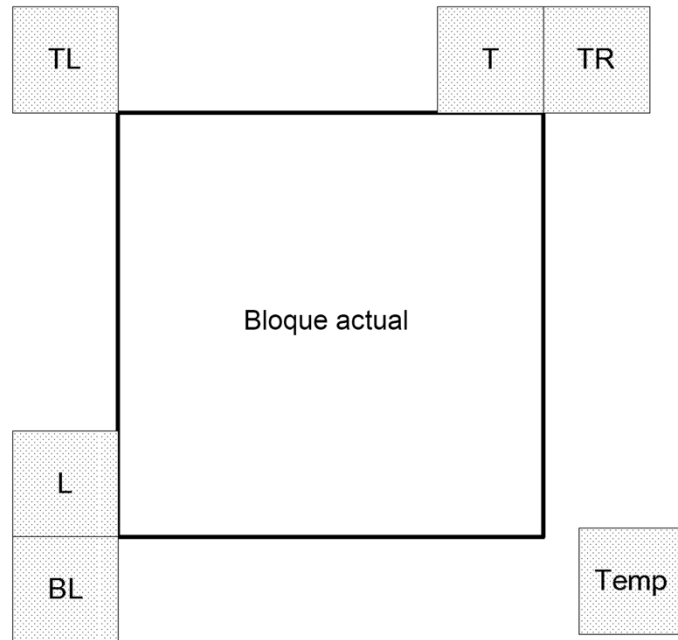
FIG. 2



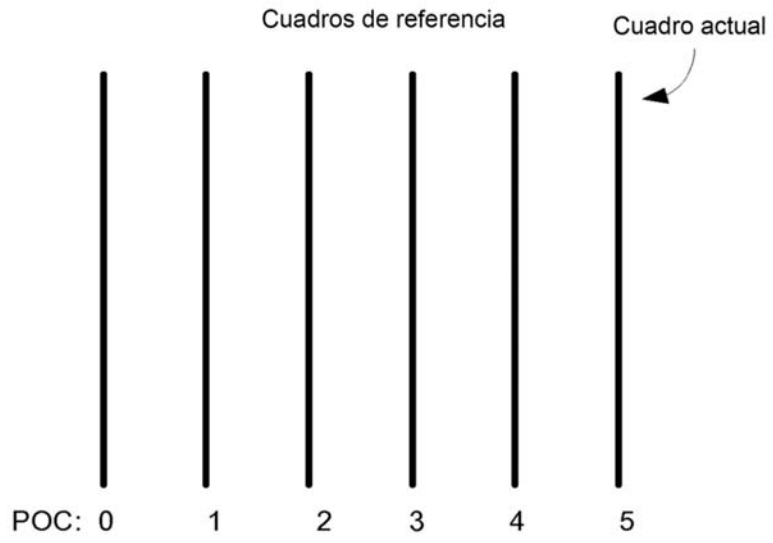
**FIG. 3A**



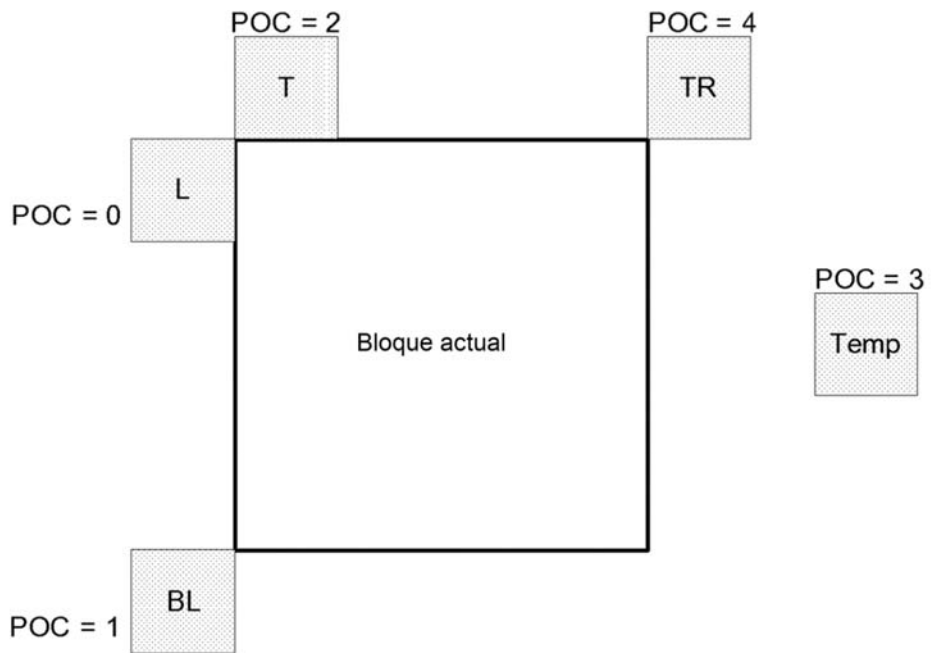
**FIG. 3B**



**FIG. 3C**



**FIG. 4**



**FIG. 5**

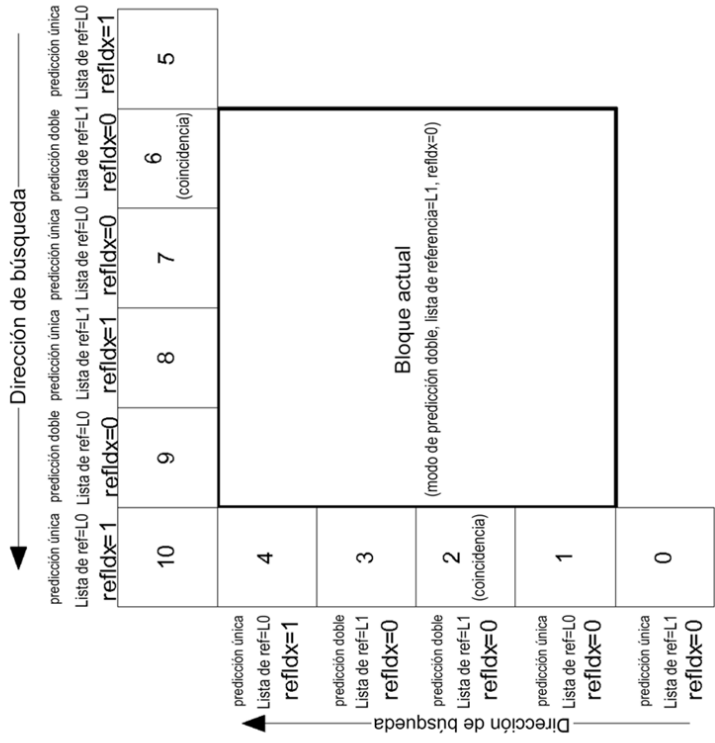


FIG. 7

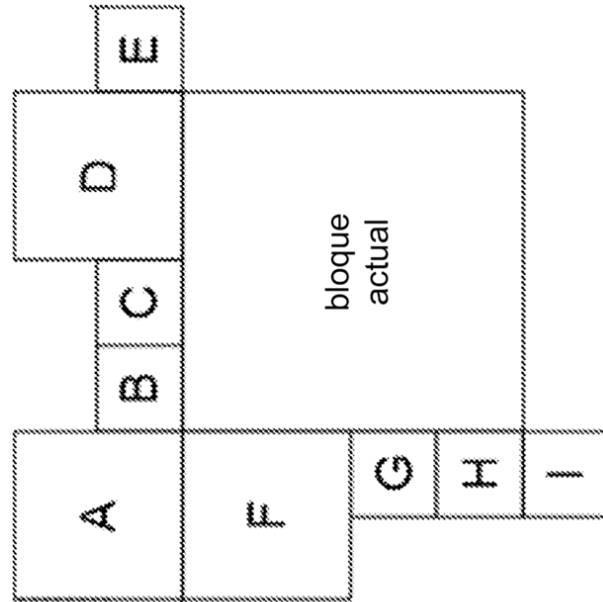
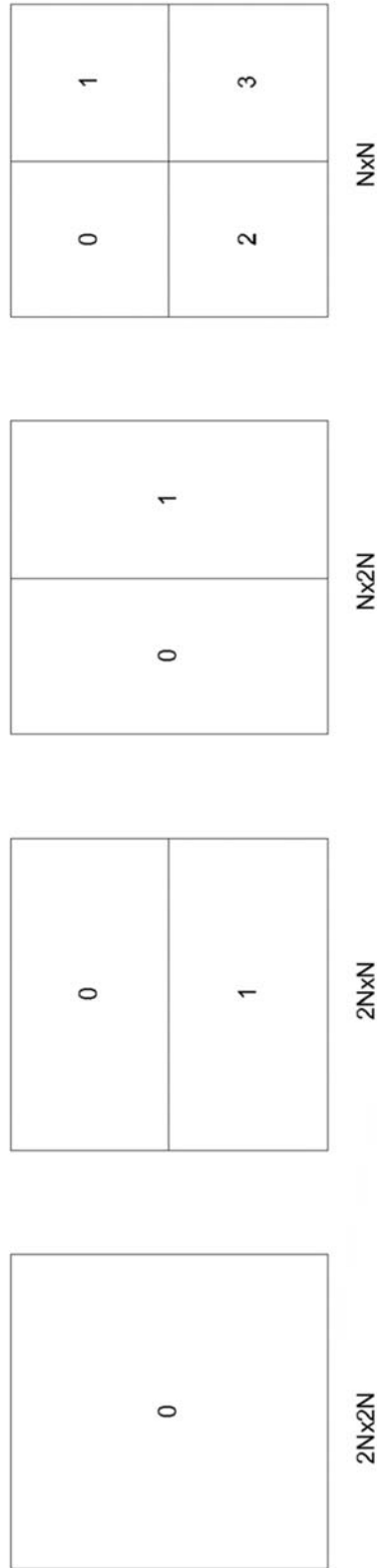
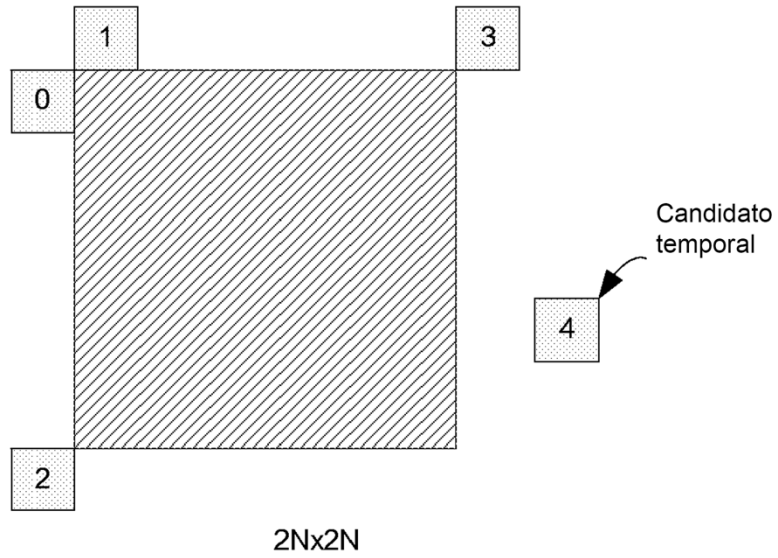


FIG. 6

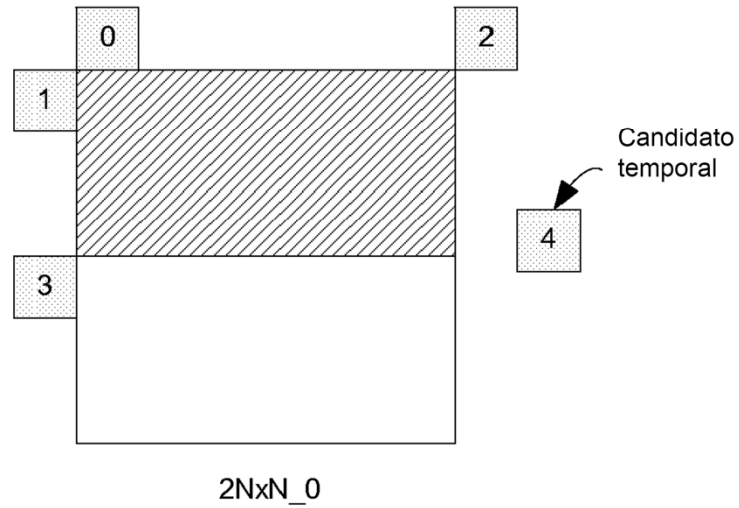


**FIG. 8**



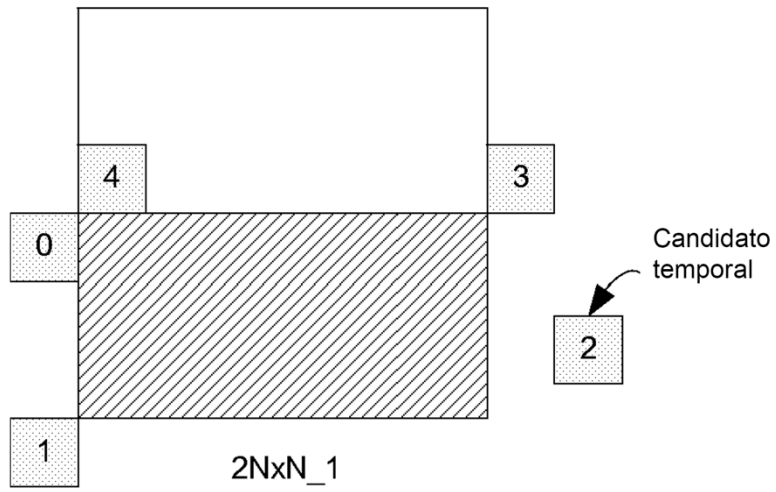
$2N \times 2N$

**FIG. 9A**

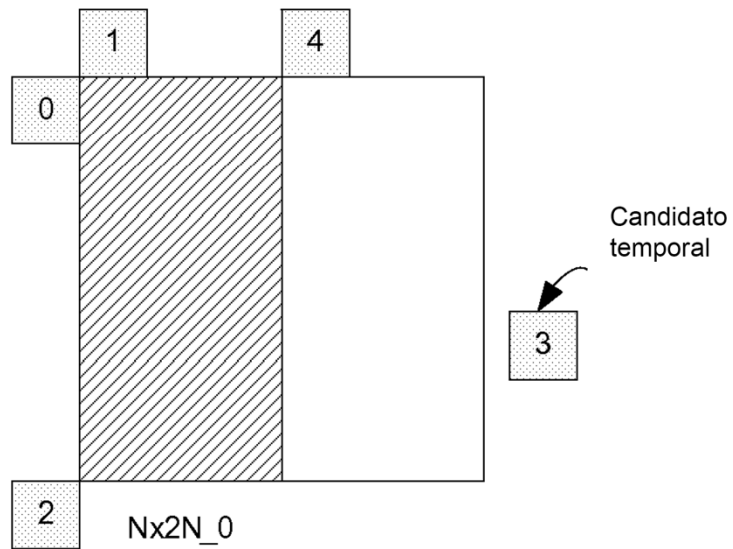


$2N \times N_0$

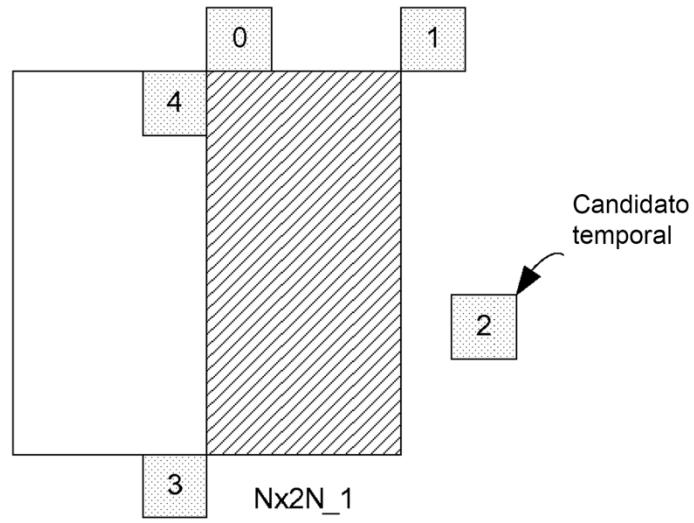
**FIG. 9B**



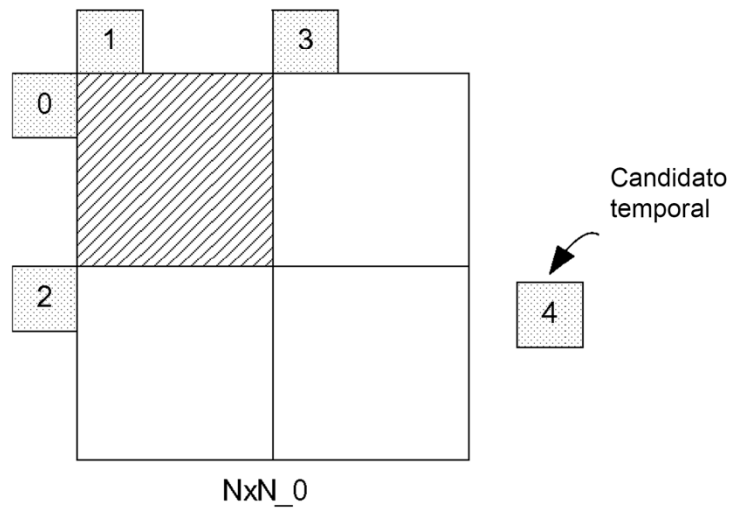
**FIG. 9C**



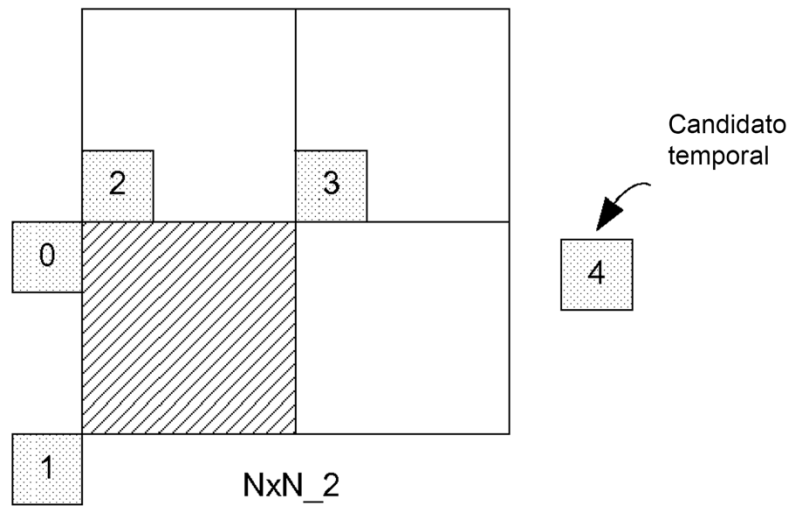
**FIG. 9D**



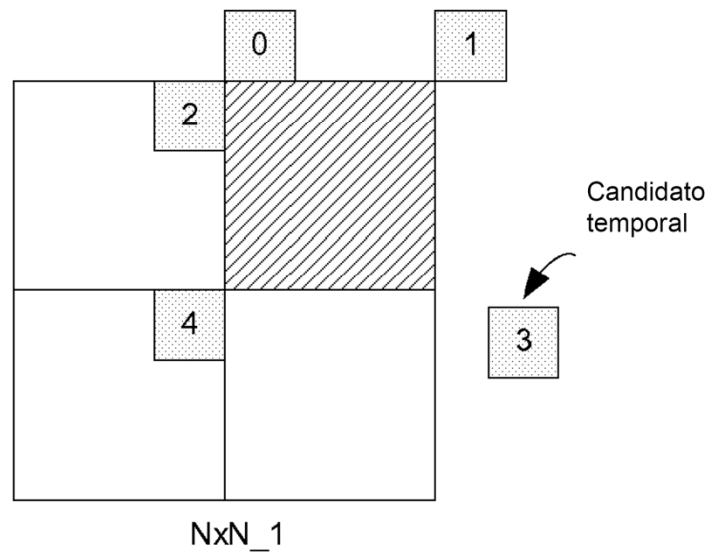
**FIG. 9E**



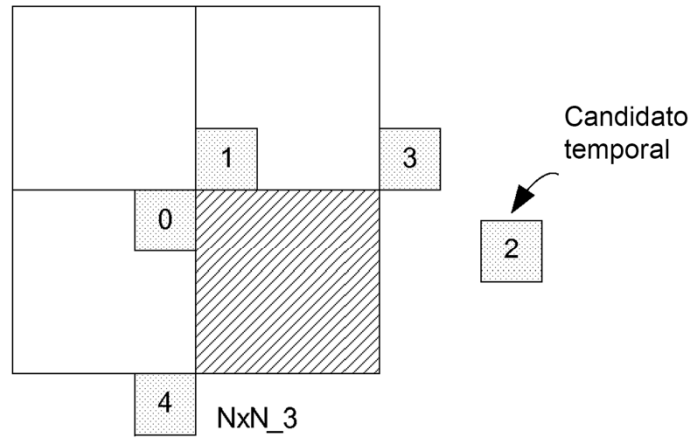
**FIG. 9F**



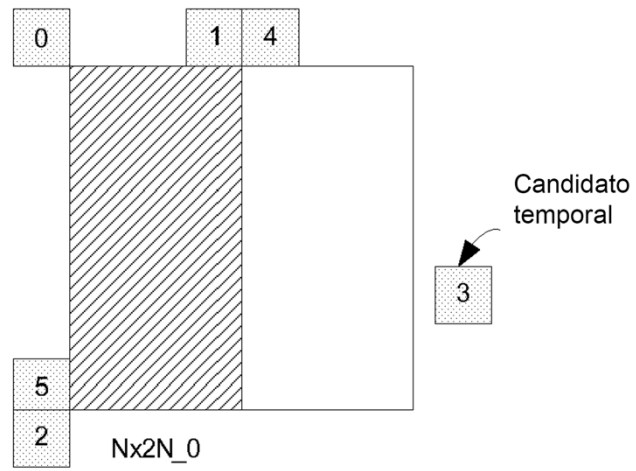
**FIG. 9G**



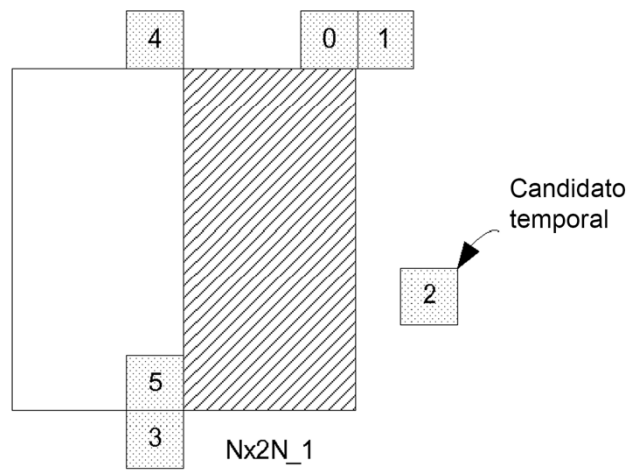
**FIG. 9H**



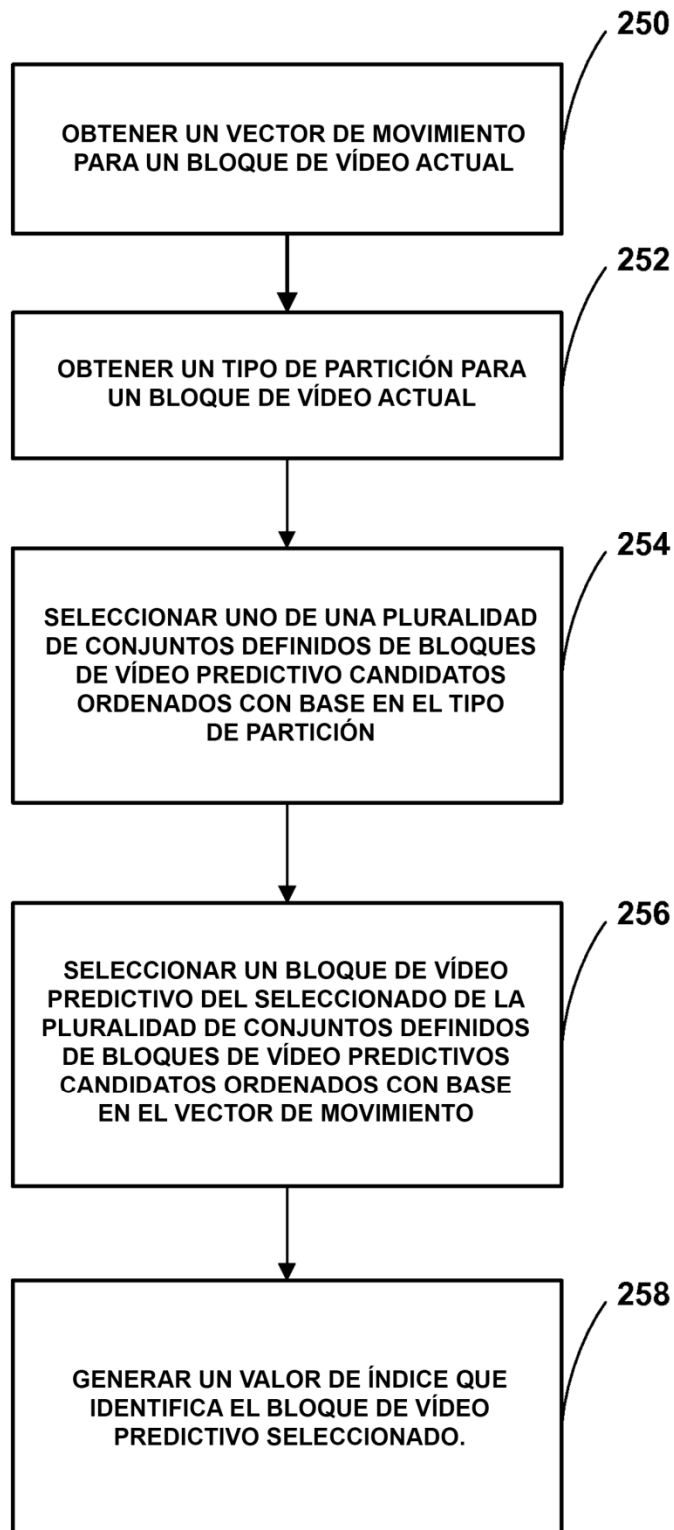
**FIG. 9I**



**FIG. 9J**



**FIG. 9K**



**FIG. 10**

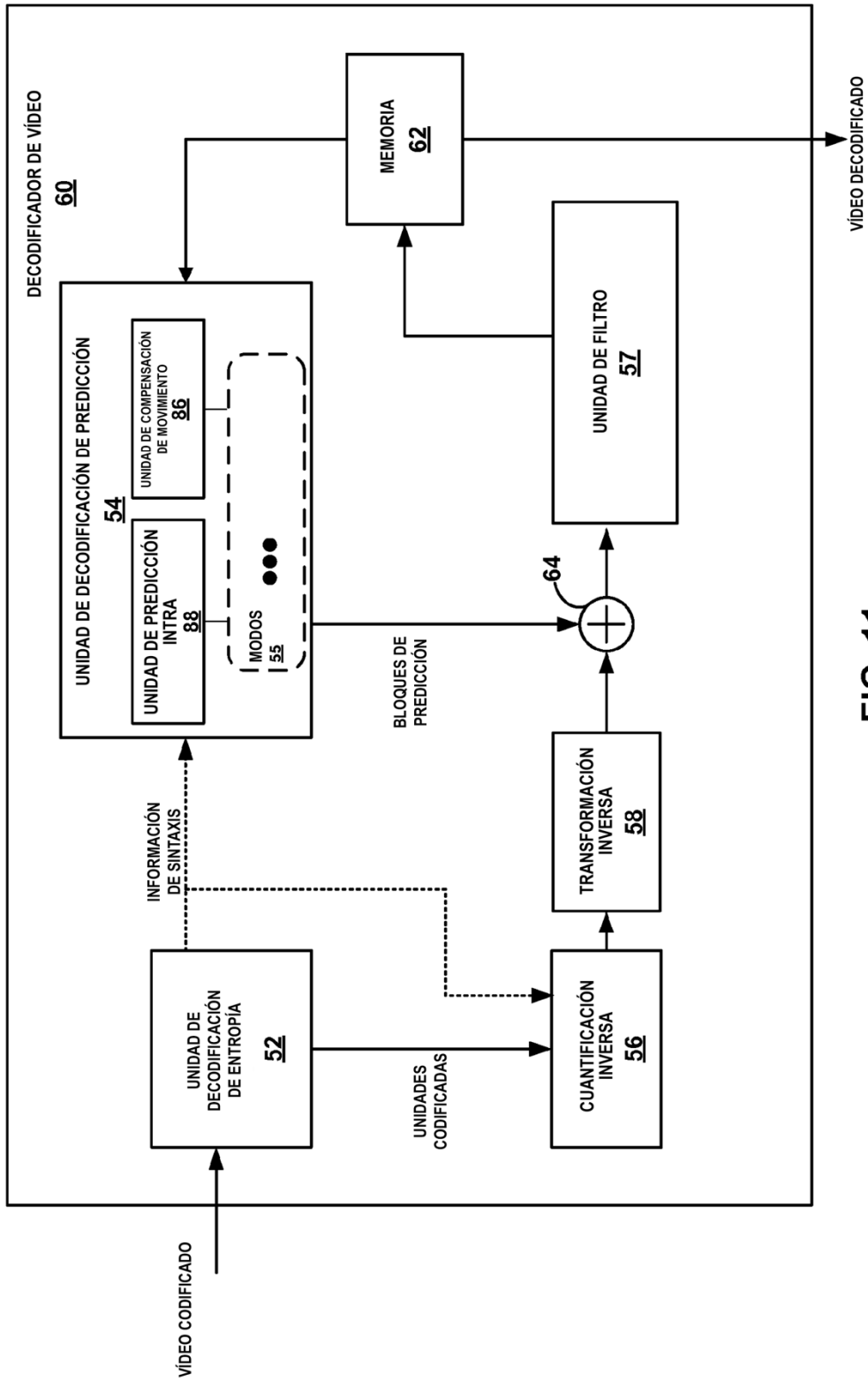
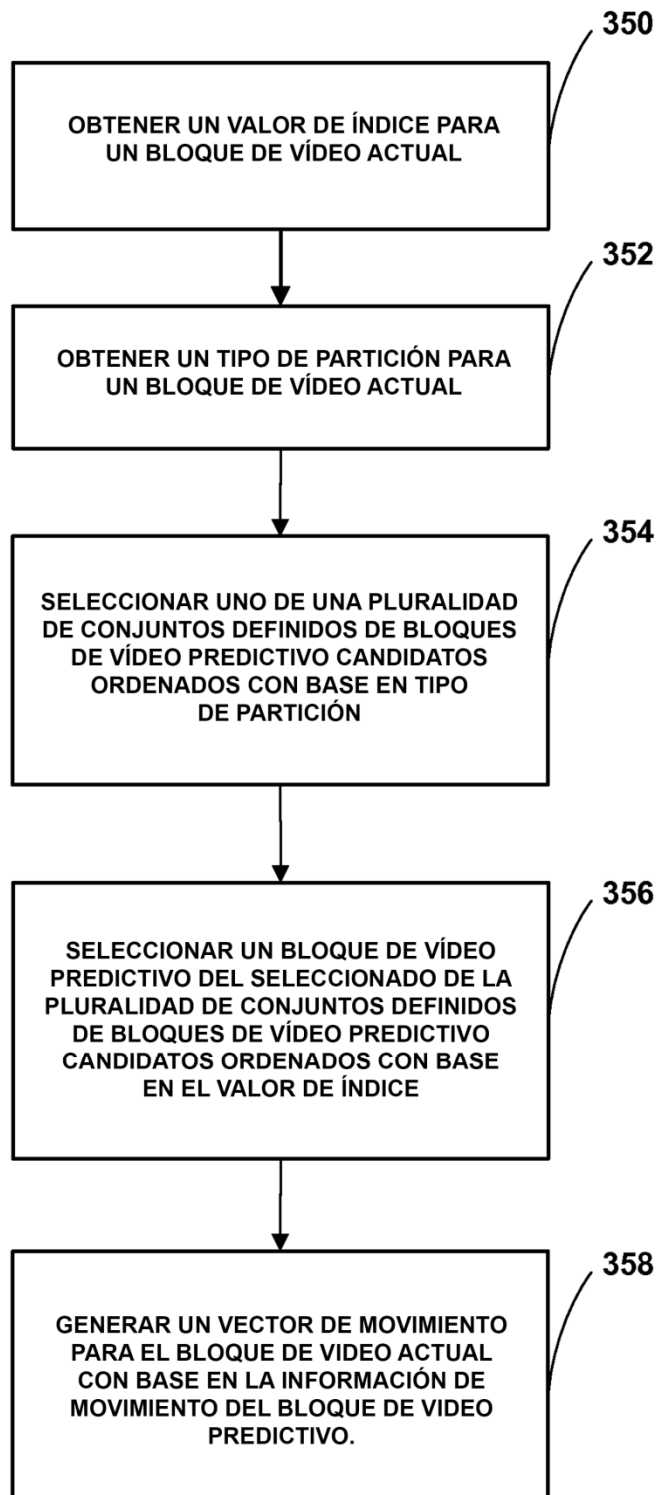


FIG. 11



**FIG. 12**