

19



OFICINA ESPAÑOLA DE  
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 757 703**

51 Int. Cl.:

**H04N 19/577** (2014.01)

**H04N 19/70** (2014.01)

**H04N 19/147** (2014.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

86 Fecha de presentación y número de la solicitud internacional: **18.02.2011 PCT/US2011/025506**

87 Fecha y número de publicación internacional: **25.08.2011 WO11103482**

96 Fecha de presentación y número de la solicitud europea: **18.02.2011 E 11706121 (8)**

97 Fecha y número de publicación de la concesión europea: **21.08.2019 EP 2537344**

54 Título: **Señalización de tipo de bloques en la codificación de vídeo**

30 Prioridad:

**19.02.2010 US 30626110 P**  
**17.02.2011 US 201113029967**

45 Fecha de publicación y mención en BOPI de la traducción de la patente:  
**29.04.2020**

73 Titular/es:

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

72 Inventor/es:

**PANCHAL, RAHUL P. y**  
**KARCZEWICZ, MARTA**

74 Agente/Representante:

**FORTEA LAGUNA, Juan José**

ES 2 757 703 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

Señalización de tipo de bloques en la codificación de vídeo

- 5 **[0001]** Esta solicitud reivindica el beneficio de la solicitud provisional estadounidense número 61/306.261 presentada el 19 de febrero de 2010.

## CAMPO TÉCNICO

- 10 **[0002]** La divulgación se refiere a la codificación de vídeo y, más en particular, a técnicas de codificación de vídeo que usan la predicción bidireccional.

## ANTECEDENTES

- 15 **[0003]** Funcionalidades de multimedia digital se pueden incorporar en una amplia gama de dispositivos, incluyendo televisiones digitales, sistemas de radiodifusión directa digitales, sistemas de radiodifusión inalámbricos, asistentes digitales personales (PDA), ordenadores portátiles o de escritorio, ordenadores de tableta, cámaras digitales, dispositivos de grabación digitales, dispositivos de videojuegos, consolas de videojuegos, teléfonos celulares o de radio por satélite, y similares. Los dispositivos de multimedia digital pueden implementar técnicas de codificación de vídeo, tales como las de las normas MPEG-2, ITU-H.263, MPEG-4, o ITU-H.264/MPEG-4, parte 10, codificación avanzada de vídeo (AVC), para transmitir y recibir o almacenar y recuperar datos de vídeo digital de manera más eficaz. Las técnicas de codificación de vídeo pueden realizar la compresión de vídeo por medio de la predicción espacial y temporal para reducir o retirar la redundancia inherente a las secuencias de vídeo. Continúan emergiendo y evolucionando nuevas normas de vídeo, como la norma de codificación de vídeo de alta eficacia (HEVC), que está siendo desarrollado por un equipo colaborativo conjunto de codificación de vídeo (JCT-VC) establecido por el grupo de expertos en imágenes en movimiento (MPEG) de la ISO/IEC y el grupo de expertos de codificación de vídeo (VCEG) de ITU-T.

- 30 **[0004]** En la codificación de vídeo, la compresión a menudo incluye predicción espacial, estimación de movimiento y compensación de movimiento. La intracodificación se basa en la predicción espacial y la codificación de transformada, tal como la transformada de coseno discreta (DCT), para reducir o retirar la redundancia espacial entre bloques de vídeo dentro de una trama de vídeo dada. La intercodificación se basa en la predicción temporal codificación de transformada para reducir o retirar la redundancia temporal entre bloques de vídeo de tramas de vídeo sucesivos de una secuencia de vídeo. Las tramas intracodificadas ("tramas I") se usan a menudo como puntos de acceso aleatorio, así como referencias para la intercodificación de otras tramas. Sin embargo, las tramas I típicamente presentan menos compresión que otras tramas.

- 40 **[0005]** Para la intercodificación, un codificador de vídeo realiza una estimación de movimiento para rastrear el movimiento de bloques de vídeo coincidentes entre dos o más tramas contiguas u otras unidades de información de vídeo que incluyen bloques de vídeo codificados, tales como cortes de tramas. Las tramas intercodificadas pueden incluir tramas predictivas ("tramas P"), que pueden incluir bloques de vídeo predichos a partir de una trama previa, y tramas predictivas bidireccionales ("tramas B"), que pueden incluir bloques de vídeo predichos a partir de una trama previa y una trama posterior de una secuencia de vídeo. Los términos tramas P y tramas B son algo históricos en el sentido de que las primeras técnicas de codificación limitaron la predicción en direcciones específicas. Es posible que los formatos y normas de codificación más nuevos no limiten la dirección de predicción de tramas P o tramas B. Por tanto, el término "bidireccional" ahora se refiere a la predicción basada en dos o más listas de datos de referencia, independientemente de la relación temporal de dichos datos de referencia en relación con los datos que se codifican.

- 50 **[0006]** De acuerdo con las normas de vídeo más nuevas, como ITU H.264, por ejemplo, la predicción bidireccional se puede basar en dos listas diferentes que no necesariamente necesitan tener datos que residan temporalmente antes y después del bloque de vídeo actual. En otras palabras, los bloques de vídeo B se pueden predecir a partir de dos listas de datos, que pueden corresponder a datos de dos tramas previas, dos tramas posteriores, o una trama previa y una trama posterior. Por el contrario, los bloques de vídeo P se predicen en base a una lista, es decir, una estructura de datos, que puede corresponder a una trama predictiva, por ejemplo, una trama previa o una trama posterior.

- 60 **[0007]** Para los bloques de vídeo P y B, la estimación de movimiento genera vectores de movimiento, que indican el desplazamiento de los bloques de vídeo en relación con los bloques de vídeo de predicción correspondientes en la(s) trama(s) de referencia predictivas u otra(s) unidad(es) de referencia. La compensación de movimiento usa los vectores de movimiento para generar bloques de vídeo de predicción a partir de la(s) trama(s) de referencia o unidad(es) de referencia. Después de la compensación de movimiento, un bloque de vídeo residual se forma restando el bloque de vídeo de predicción del bloque de vídeo original que se va a codificar. El codificador de vídeo normalmente aplica procesos de codificación de transformada, cuantificación y por entropía para reducir aún más la velocidad de bits asociada con la comunicación del bloque residual. Las unidades I y P se usan comúnmente para definir bloques de referencia para la intercodificación de las unidades P y B.

[0008] Flierl, M *et al*, en "Generalized B pictures and the draft H.264/AVC video-compression standard", IEEE TRANSACTIONS ON CIRCUITS AND SYSTEMS FOR VIDEO TECHNOLOGY, IEEE SERVICE CENTER, PISCATAWAY, NJ, US, (20030701), vol. 13, no. 7, páginas 587 - 597, revisan el uso de imágenes B en el contexto del proyecto de norma H.264/AVC de compresión de vídeo, enfocándose en la selección de imágenes de referencia y señales de predicción de compensación de movimiento combinadas linealmente.

[0009] Gary Sullivan *et al*, en "Technical Description of ITU-T VCEG Draft H.26L Algorithm in Response to Video and DCinema CfPs", ITU STUDY GROUP 16 - VIDEO CODING EXPERTS GROUP -ISO/IEC MPEG & ITU-T VCEG (ISO/IEC JTC1/SC29/WG11 e ITU-T SG16 Q6), XX, XX, (20010710), no. M7512, introducen una definición de imágenes B y analizan su uso.

[0010] Jeon, en "New syntax for Bi-dir pred. mode in MH pictures", ITU STUDY GROUP 16 - VIDEO CODING EXPERTS GROUP -ISO/IEC MPEG & ITU-T VCEG (ISO/IEC JTC1/SC29/WG11 e ITU-T SG16 Q6), XX, XX, (20020510), no. JVT-C121, analiza el uso de dos parámetros de índice de referencia de imagen como los índices relativos para un búfer de múltiples imágenes.

[0011] Cabe señalar que los documentos mencionados anteriormente especifican un elemento sintáctico (es decir, MB\_type) para señalar tanto el tamaño de división como la dirección de predicción para los bloques de vídeo bidireccionales.

[0012] Adachi *et al*, en "Core Exp. Results Improved MB Pred Modes", ITU STUDY GROUP 16 - VIDEO CODING EXPERTS GROUP -ISO/IEC MPEG & ITU-T VCEG (ISO/IEC JTC1/SC29/WG11 e ITU-T SG16 Q6), XX, XX, (20020201), no. JVT-B090, analizan la sintaxis de trama B.

[0013] Cabe señalar que el último documento especifica dos elementos sintácticos para bloques de vídeo bidireccionales, es decir, MB\_type y Ref\_frame, con lo que MB\_type es indicativo del modo de división y el elemento sintáctico Ref\_frame es indicativo tanto de una dirección de predicción como una trama de referencia.

[0014] "Draft ITU-T Recommendation and Final Draft International Standard of Joint Video Specification (ITU-T Rec. H.264 | ISO/IEC 14496-10 AVC)" ITU STUDY GROUP 16 - VIDEO CODING EXPERTS GROUP -ISO/IEC MPEG & ITU-T VCEG(ISO/IEC JTC1/SC29/WG11 e ITU-T SG16 Q6), XX, XX, (20030314) no. JVT-G050r1, describe la semántica de la capa de macrobloque y la semántica de la predicción de macrobloque.

## SUMARIO

[0015] Esta divulgación describe técnicas de codificación y descodificación de vídeo aplicables a la predicción bidireccional. En la predicción bidireccional, un bloque de vídeo se codifica y descodifica de manera predictiva en base a dos listas diferentes de datos de referencia predictivos. En un aspecto de esta divulgación, se describen técnicas para señalar desde un codificador a un descodificador un tipo de bloque para un bloque de vídeo, donde la señalización del tipo de bloque incluye señalar por separado un tamaño de división para el bloque de vídeo, una primera dirección de predicción para el bloque de vídeo y, en algunos casos, una segunda dirección de predicción para el bloque de vídeo. Las técnicas de la presente divulgación pueden usar elementos sintácticos separados para señalar el tamaño de división, la primera dirección de predicción y la segunda dirección de predicción. Adicionalmente, se describen técnicas para señalar un tipo de subbloque desde un codificador a un descodificador, donde la señalización del tipo de subbloque también puede incluir la señalización por separado de un tamaño de división y una dirección de predicción usando elementos sintácticos separados. Los elementos sintácticos separados se pueden codificar y transmitir individualmente desde el codificador al descodificador.

[0016] La invención se define en las reivindicaciones independientes adjuntas. Las reivindicaciones dependientes definen rasgos característicos opcionales de la invención.

[0017] Los detalles de uno o más aspectos de la divulgación se exponen en los dibujos adjuntos y en la descripción siguiente. Otros rasgos característicos, objetivos y ventajas de las técnicas descritas en la presente divulgación resultarán evidentes a partir de la descripción y de los dibujos, y a partir de las reivindicaciones.

## BREVE DESCRIPCIÓN DE LOS DIBUJOS

### [0018]

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

La FIG. 2 es un diagrama de bloques que ilustra un ejemplo de un codificador de vídeo que puede realizar las técnicas acordes con esta divulgación.

La FIG. 3 es un diagrama de bloques que ilustra un ejemplo de una unidad de compensación de movimiento con mayor detalle.

5 La FIG. 4 es un diagrama de bloques que ilustra un ejemplo de un codificador de vídeo que puede realizar técnicas acordes con esta divulgación.

La FIG. 5 es un diagrama de flujo que ilustra un ejemplo de un proceso realizado por un codificador de vídeo acorde con esta divulgación.

10 La FIG. 6 es un diagrama de flujo que ilustra un ejemplo de un proceso realizado por un descodificador de vídeo acorde con esta divulgación.

## DESCRIPCIÓN DETALLADA

15 **[0019]** Esta divulgación describe técnicas de codificación y descodificación de vídeo aplicables a la predicción bidireccional. En la predicción bidireccional, un bloque de vídeo se codifica y descodifica de manera predictiva en base a dos listas diferentes de datos de referencia predictivos. En un aspecto de esta divulgación, se describen técnicas para señalar desde un codificador a un descodificador un tipo de bloque para un bloque de vídeo, donde la señalización del tipo de bloque incluye señalar por separado un tamaño de división para el bloque de vídeo, una primera dirección de predicción para el bloque de vídeo y, en algunos casos, una segunda dirección de predicción para el bloque de vídeo. Las técnicas de la presente divulgación pueden usar elementos sintácticos separados para señalar el tamaño de división, la primera dirección de predicción y la segunda dirección de predicción. Adicionalmente, se describen técnicas para señalar desde un codificador a un descodificador un tipo de subbloque, donde la señalización del tipo de subbloque también puede incluir señalar por separado un tamaño de división y una dirección de predicción usando elementos sintácticos separados. Los elementos sintácticos separados se pueden codificar y transmitir individualmente desde el codificador al descodificador.

20 **[0020]** Los elementos sintácticos separados, como se usan en esta divulgación, en general se refieren a una pluralidad de elementos sintácticos para señalar, desde un codificador a un descodificador, una pluralidad de partes de información, tales como tamaños de división y direcciones de predicción. Por el contrario, un elemento sintáctico conjunto se refiere a un único elemento sintáctico que se usa para señalar una pluralidad de partes de información desde un codificador a un descodificador. Por ejemplo, a diferencia de los esquemas de codificación existentes en los que un elemento sintáctico conjunto se podría usar para señalar un tamaño de división y una dirección de predicción, las técnicas de esta divulgación incluyen el uso de un primer elemento sintáctico para señalar el tamaño de división y un segundo elemento sintáctico separado para señalar la dirección de la predicción. En algunos casos, un elemento sintáctico conjunto también podría ser un elemento sintáctico separado en relación con otro elemento sintáctico. Por ejemplo, un elemento sintáctico conjunto usado para señalar una primera dirección de predicción y una segunda dirección de predicción podría ser un elemento sintáctico separado en relación con un elemento sintáctico que se usa para señalar un tamaño de división. Como se analizará con más detalle a continuación, los elementos sintácticos separados se pueden codificar y procesar independientemente unos de otros. Por ejemplo, dos elementos sintácticos separados pueden experimentar dos procesos de codificación por entropía diferentes antes de ser transmitidos.

30 **[0021]** Un esquema de codificación que usa elementos sintácticos separados para señalar el tamaño de división y la dirección de predicción se puede extender más fácilmente para su uso con tamaños de bloque de vídeo más grandes en comparación con un esquema de codificación que usa elementos sintácticos conjuntos. Por ejemplo, la introducción de un nuevo tamaño de bloque de vídeo o un nuevo tamaño de división en un esquema de codificación que utiliza elementos sintácticos conjuntos podría requerir la introducción de más elementos sintácticos nuevos en comparación con un esquema de codificación que señala por separado el tamaño de división y la dirección de predicción, como se describe en esta divulgación. Adicionalmente, el uso de elementos sintácticos separados en comparación con un elemento sintáctico conjunto puede resultar en ganancias de eficacia de codificación debido a la codificación por entropía mejorada en algunos casos.

35 **[0022]** La FIG. 1 es un diagrama de bloques que ilustra un sistema de codificación y descodificación de vídeo ejemplar 10 que puede implementar las técnicas de esta divulgación. Como se muestra en la FIG. 1, el sistema 10 incluye un dispositivo de origen 12 que transmite el vídeo codificado a un dispositivo de destino 16 por medio de un canal de comunicación 15. El dispositivo de origen 12 y el dispositivo de destino 16 pueden comprender uno cualquiera de una amplia gama de dispositivos. En algunos casos, el dispositivo de origen 12 y el dispositivo de destino 16 comprenden dispositivos de comunicación inalámbrica, tales como microteléfonos inalámbricos, los denominados radioteléfonos celulares o vía satélite o cualquier dispositivo inalámbrico que pueda comunicar información de vídeo a través de un canal de comunicación 15, que puede o no ser inalámbrico. Sin embargo, las técnicas de esta descripción, que implican tipos de bloques y tipos de subbloques de señalización para unidades B desde un codificador a un descodificador, no se limitan necesariamente a aplicaciones o configuraciones inalámbricas.

60 **[0023]** En el ejemplo de la FIG. 1, el dispositivo de origen 12 puede incluir una fuente de vídeo 20, un codificador de vídeo 22, un modulador/desmodulador (módem) 23 y un transmisor 24. El dispositivo de destino 16 puede incluir un

receptor 26, un módem 27, un descodificador de vídeo 28 y un dispositivo de visualización 30. De acuerdo con esta divulgación, el codificador de vídeo 22 del dispositivo de origen 12 se puede configurar para generar elementos sintácticos separados para indicar un tamaño de división, una primera dirección de predicción y una segunda dirección para un bloque de vídeo, y adicionalmente, para generar elementos sintácticos separados para indicar un tamaño de división y una dirección de predicción para un subbloque. Los elementos sintácticos separados se pueden codificar y transmitir individualmente desde el codificador de vídeo 22 al descodificador de vídeo 28. El descodificador de vídeo 28 puede recibir los elementos sintácticos separados que indican la selección. En consecuencia, el descodificador de vídeo 28 puede realizar la descodificación de vídeo apropiada en base a los elementos sintácticos recibidos.

**[0024]** El sistema 10 ilustrado de la FIG. 1 es simplemente un ejemplo. Las técnicas de señalización de esta divulgación se pueden realizar mediante cualquier dispositivo de codificación que admita la predicción compensada por movimiento bidireccional. El dispositivo de origen 12 y el dispositivo de destino 16 son simplemente ejemplos de dichos dispositivos de codificación en los que el dispositivo de origen 12 genera datos de vídeo codificados para su transmisión al dispositivo de destino 16. En algunos casos, los dispositivos 12, 16 pueden funcionar de manera sustancialmente simétrica, de modo que cada uno de los dispositivos 12, 16 incluya componentes de codificación y de descodificación de vídeo. De ahí que el sistema 10 pueda admitir la transmisión de vídeo unidireccional o bidireccional entre los dispositivos de vídeo 12, 16, por ejemplo, para la transmisión continua de vídeo, la reproducción de vídeo, la radiodifusión de vídeo o la videotelefonía.

**[0025]** La fuente de vídeo 20 del dispositivo de origen 12 puede incluir un dispositivo de captura de vídeo, tal como una cámara de vídeo, un archivo de vídeo que contenga vídeo capturado previamente o un suministro de vídeo de un proveedor de contenido de vídeo. Como otra alternativa, la fuente de vídeo 20 puede generar datos basados en gráficos de ordenador como la fuente de vídeo, o una combinación de vídeo en directo, vídeo archivado y vídeo generado por ordenador. En algunos casos, si la fuente de vídeo 20 es una cámara de vídeo, el dispositivo de origen 12 y el dispositivo de destino 16 pueden formar los llamados teléfonos con cámara o videotelefonos. En cada caso, el vídeo capturado, precapturado o generado por ordenador se puede codificar mediante el codificador de vídeo 22. La información de vídeo codificada se puede modular a continuación por el módem 23 de acuerdo con una norma de comunicación, por ejemplo, tal como el acceso múltiple por división de código (CDMA), la multiplexación por división de frecuencias ortogonales (OFDM) u otra técnica o norma de comunicación, y transmitirse al dispositivo de destino 16 por medio 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, incluyendo amplificadores, filtros y una o más antenas.

**[0026]** El receptor 26 del dispositivo de destino 16 recibe información a través del canal 15, y el módem 27 desmodula la información. La información comunicada a través del canal 15 puede incluir información definida por el codificador de vídeo 22, que también puede ser usada por el descodificador de vídeo 28 de acuerdo con esta divulgación. El dispositivo de visualización 30 muestra los datos de vídeo descodificados a un usuario y puede comprender cualquiera de una variedad de dispositivos de visualización, tales como un tubo de rayos catódicos, una pantalla de cristal líquido (LCD), una pantalla de plasma, una pantalla de diodos orgánicos emisores de luz (OLED) u otro tipo de dispositivo de visualización.

**[0027]** En el ejemplo de la FIG 1, el canal 15 de comunicación puede comprender cualquier medio de comunicación inalámbrica o alámbrica, tal como un espectro de radiofrecuencia (RF) o una o más líneas de transmisión físicas, o cualquier combinación de medios inalámbricos y alámbricos. En consecuencia, el módem 23 y el transmisor 24 pueden admitir muchos posibles protocolos inalámbricos, protocolos alámbricos o protocolos alámbricos e inalámbricos. El canal de comunicación 15 puede formar parte de una red basada en paquetes, tal como una red de área local (LAN), una red de área amplia (WAN) o una red global, tal como Internet, que comprenda una interconexión de una o más redes. El canal de comunicación 15 representa en general cualquier medio de comunicación adecuado o una colección de diferentes medios de comunicación, para transmitir datos de vídeo desde el dispositivo de origen 12 hasta el dispositivo de destino 16. El canal de comunicación 15 puede incluir enrutadores, conmutadores, estaciones base o cualquier otro equipo que pueda ser útil para facilitar la comunicación desde el dispositivo de origen 12 hasta el dispositivo de destino 16.

**[0028]** El codificador de vídeo 22 y el descodificador de vídeo 28 pueden funcionar de acuerdo con una norma de compresión de vídeo, tal como la norma ITU-T H.265 de nueva aparición. Sin embargo, las técnicas de esta divulgación no están limitadas a ninguna norma de codificación particular. Aunque no se muestra en la FIG. 1, en algunos aspectos, tanto el codificador de vídeo 22 como el descodificador de vídeo 28 pueden estar cada uno integrados con un codificador y descodificador de audio, y pueden incluir unidades MUX-DEMUX adecuadas, u otro tipo de hardware y programas informáticos, para manejar la codificación tanto de audio como de vídeo en un flujo de datos común o en flujos de datos separados. Si fuera aplicable, las unidades MUX-DEMUX pueden ser conformes al protocolo multiplexor ITU H.223 o a otros protocolos tales como el protocolo de datagramas de usuario (UDP).

**[0029]** El codificador de vídeo 22 y el descodificador de vídeo 28 pueden cada uno implementarse como uno o más microprocesadores, procesadores de señales digitales (DSP), circuitos integrados específicos de la aplicación (ASIC), matrices de puertas programables por campo (FPGA), lógica discreta, programas informáticos, hardware, firmware o cualquier combinación de los mismos. Cada uno del codificador de vídeo 22 y del descodificador de vídeo 28 se

pueden incluir en uno o más codificadores o descodificadores, pudiendo cualquiera de los mismos estar integrado como parte de un codificador/descodificador combinado (CODEC) en un respectivo dispositivo móvil, dispositivo de abonado, dispositivo de radiodifusión, servidor o similares.

5 **[0030]** Una secuencia de vídeo incluye típicamente una serie de tramas de vídeo. Un codificador de vídeo 20 funciona sobre bloques de vídeo de tramas de vídeo individuales a fin de codificar los datos de vídeo. En un ejemplo, un bloque de vídeo puede corresponder a un macrobloque o a una división de un macrobloque. Los macrobloques son un tipo de videobloque definido por la norma ITU H.264 y otras normas. Los macrobloques generalmente se refieren a bloques de datos de 16x16, aunque el término también se usa comúnmente para referirse genéricamente a cualquier  
10 bloque de vídeo de tamaño NxN. Como un ejemplo, la norma ITU-T H.264 admite la intrapredicción para varios tamaños de bloque, tales como 16x16, 8x8 o 4x4 para componentes de luma, y 8x8 para componentes de crominancia, así como la interpredicción para varios tamaños de bloque, tales como 16x16, 16x8, 8x16, 8x8, 8x4, 4x8 y 4x4 para componentes de luma y tamaños ajustados a escala correspondientes para componentes de crominancia. En esta divulgación, "NxN" se refiere a las dimensiones de píxel del bloque en lo que respecta a las dimensiones vertical y  
15 horizontal, por ejemplo 16x16 píxeles. En general, un bloque de 16x16 tendrá 16 píxeles en una dirección vertical y 16 píxeles en una dirección horizontal. Asimismo, un bloque de NxN tiene, en general, N píxeles en una dirección vertical y N píxeles en una dirección horizontal, donde N representa un valor entero positivo. Los píxeles en un bloque se pueden disponer en filas y columnas.

20 **[0031]** La norma ITU H.265 de nueva aparición define nuevos términos para bloques de vídeo. En particular, con la ITU H.265, los bloques de vídeo (o sus divisiones) se pueden denominar "unidades codificadas". Con la norma ITU-T H.265, las unidades codificadas más grandes (LCU) se pueden dividir en unidades codificadas (CU) cada vez más pequeñas de acuerdo con un esquema de división de árbol cuaternario, y las diferentes CU que se definen en el esquema se pueden dividir además en las llamadas unidades de predicción (PU). Las LCU, las CU y las PU son todas  
25 bloques de vídeo en el sentido de esta divulgación. También se pueden usar otros tipos de bloques de vídeo, de acuerdo con la norma ITU H.265 u otras normas de codificación de vídeo. Por tanto, la frase "bloques de vídeo" se refiere a cualquier tamaño de bloque de vídeo.

**[0032]** Los bloques de vídeo pueden tener tamaños fijos o variados y pueden diferir en tamaño de acuerdo con una  
30 norma de codificación especificada. Cada trama de vídeo puede incluir una pluralidad de cortes. Cada corte puede incluir una pluralidad de macrobloques, que se pueden disponer en divisiones, también denominadas subbloques. De acuerdo con el esquema de división de árbol cuaternario mencionado anteriormente, una primera CU de  $N/2 \times N/2$  puede ser un subbloque de una LCU de NxN, una segunda CU de  $N/4 \times N/4$  también puede ser un subbloque de la primera CU. Una PU de  $N/8 \times N/8$  podría ser un subbloque de la segunda CU. De forma similar, como otro ejemplo, los tamaños de bloque que son menores de 16x16 se pueden denominar divisiones de un bloque de vídeo de 16x16 o subbloques del bloque de vídeo de 16x16. Del mismo modo, para un bloque de NxN, los tamaños de bloque inferiores a NxN se pueden denominar divisiones o subbloques del bloque de NxN. Los bloques de vídeo pueden comprender bloques de datos de píxeles en el dominio de píxel, o bloques de coeficientes de transformada en el dominio de transformada, por ejemplo, tras la aplicación de una transformada, tal como una transformada de coseno discreta (DCT), una transformada de números enteros, una transformada de ondículas o una transformada conceptualmente similar a los datos de bloques de vídeo residuales que representan diferencias de píxeles entre bloques de vídeo codificados y bloques de vídeo predictivos. En algunos casos, un bloque de vídeo puede comprender bloques de coeficientes de transformada cuantificados en el dominio de transformada.

45 **[0033]** Los bloques de vídeo más pequeños pueden proporcionar una mejor resolución y se pueden usar para localizaciones de una trama de vídeo que incluyen altos niveles de detalle. Los bloques de vídeo más grandes pueden proporcionar una mayor eficacia de codificación y se pueden usar para localizaciones de una trama de vídeo que incluye un bajo nivel de detalle. Un corte se puede considerar una pluralidad de bloques y/o subbloques de vídeo. Cada corte puede ser una serie independientemente descodificable de bloques de vídeo de una trama de vídeo. De forma alternativa, las propias tramas pueden ser series descodificables de bloques de vídeo, u otras partes de una trama se pueden definir como series descodificables de bloques de vídeo. El término «series de bloques de vídeo» se puede referir a cualquier parte independientemente descodificable de una trama de vídeo, tal como una trama completa, un corte de una trama, un grupo de imágenes (GOP), denominado también secuencia, u otra unidad independientemente descodificable definida de acuerdo con las técnicas de codificación aplicables.

55 **[0034]** Después de la codificación predictiva entre bases, y después de cualquier transformación (como una transformación entera de 4x4 u 8x8 o una transformada de coseno discreta o DCT), se puede realizar la cuantización. La cuantificación se refiere en general a un proceso en el que los coeficientes se cuantifican para reducir posiblemente la cantidad de datos usados para representar los coeficientes. El proceso de cuantificación puede reducir la profundidad de bits asociada a algunos de, o a todos, los coeficientes. Tras la cuantificación, se puede realizar la codificación por entropía, por ejemplo de acuerdo con la codificación de longitud variable adaptativa al contenido (CAVLC), con la codificación aritmética binaria adaptativa al contexto (CABAC) o con otra metodología de codificación por entropía.

65 **[0035]** Las técnicas de esta descripción son específicamente aplicables a los bloques de vídeo B que hacen uso de la predicción bidireccional, que podría incluir tanto la predicción bidireccional ponderada como la predicción

bidireccional no ponderada. En esta divulgación, el término "unidades B" se usará para referirse en general a cualquier tipo de unidades B, que pueden incluir tramas B, cortes B o posiblemente otras unidades de vídeo que incluyan al menos algunos bloques de vídeo B.

5 **[0036]** Como se mencionó anteriormente, la predicción bidireccional es la predicción de los llamados "bloques de vídeo B" en base a dos listas diferentes de datos. Los bloques de vídeo B usan dos listas de imágenes de referencia previamente codificadas, la lista 0 y la lista 1. Estas dos listas pueden cada una contener imágenes codificadas pasadas y/o futuras en orden temporal. Los bloques de vídeo B se pueden predecir de una entre varias maneras: predicción compensada por movimiento de una imagen de referencia de lista 0, predicción compensada por movimiento de una imagen de referencia de lista 1 o predicción compensada por movimiento de la combinación de imágenes de referencia de ambas listas, lista 0 y lista 1. Para obtener la combinación de las imágenes de referencia, tanto de la lista 0 como de la lista 1, se obtienen dos áreas de referencia compensadas por movimiento de imágenes de referencia de la lista 0 y de la lista 1, respectivamente.

15 **[0037]** Los bloques de vídeo B se pueden predecir a partir de dos listas de datos de dos tramas previas, dos listas de datos de tramas posteriores o una lista de datos de una trama previa y una de una trama posterior. Aunque la lista 0 y la lista 1 no se correlacionan necesariamente con una dirección temporal particular, la selección de si realizar una compensación de movimiento mediante la lista 0, la lista 1 o ambas listas 0 y 1 se conoce comúnmente como dirección de predicción. A diferencia de los bloques de vídeo B, los bloques de vídeo P se predicen en base a una lista, que puede corresponder a una trama predictiva, por ejemplo, una trama previa o una trama posterior. Las tramas B y las tramas P se pueden denominar más en general unidades P y unidades B. Las unidades P y las unidades B también se pueden realizar en unidades más pequeñas, tales como cortes de tramas o partes de tramas. Las unidades B pueden incluir bloques de vídeo B, bloques de vídeo P o bloques de vídeo I. Las unidades P pueden incluir bloques de vídeo P o bloques de vídeo I. Las unidades I pueden incluir solo bloques de vídeo I.

25 **[0038]** Los bloques de vídeo B que utilizan predicción bidireccional (es decir, tanto la lista 0 como la lista 1) se pueden codificar usando bien la predicción bidireccional ponderada o la predicción bidireccional no ponderada. La predicción bidireccional ponderada se refiere a la predicción bidireccional que permite asignar factores de peso a las dos listas diferentes. Cada lista puede comprender un conjunto de datos asociados con una trama predictiva u otra serie de bloques de vídeo. En la predicción bidireccional ponderada, una lista se puede ponderar con mayor peso en la generación de datos predictivos. Si una de las listas tiene datos que son más similares al bloque de vídeo que se está codificando, por ejemplo, esa lista se puede ponderar con mayor peso que la otra lista.

35 **[0039]** Para diferentes tipos de predicción bidireccional ponderada de acuerdo con ITU-T H.264, por ejemplo, el codificador de vídeo 22 y el descodificador de vídeo 28 en general pueden admitir tres tipos diferentes de modos de predicción. Un primer modo de predicción, denominado "predicción ponderada predeterminada", se refiere a la predicción ponderada en la que los factores de peso asociados con dos o más listas diferentes están predefinidos por alguna configuración predeterminada. La predicción ponderada predeterminada, en algunos casos, puede asignar ponderaciones iguales a cada una de las listas, aunque la ponderación desigual también se podría predefinir para un modo de predicción ponderada predeterminada.

45 **[0040]** Un segundo modo de predicción, denominado "predicción ponderada implícita", se refiere a la predicción ponderada en la que los factores de peso asociados con dos o más listas diferentes se definen en base a algunos factores implícitos asociados con los datos. Por ejemplo, los factores de peso implícitos se pueden definir por las posiciones temporales relativas de los datos en las dos listas diferentes en relación con los datos que se codifican de forma predictiva. Tanto en la predicción ponderada predeterminada como en la predicción ponderada implícita, los factores de peso no se incluyen en el flujo de bits. En cambio, el descodificador de vídeo 28 se puede programar para conocer los factores de peso (para la predeterminada) o se puede programar para saber cómo derivar los factores de peso (para la implícita).

50 **[0041]** Un tercer modo de predicción, denominado "predicción ponderada explícita", se refiere a la predicción ponderada en la que los factores de peso se definen dinámicamente como parte del proceso de codificación y se codifican en el flujo de bits. La predicción ponderada explícita es diferente de la predicción ponderada predeterminada y la predicción ponderada implícita a este respecto, por ejemplo, la predicción ponderada explícita da como resultado factores de ponderación que se codifican como parte del flujo de bits.

55 **[0042]** Estos tres modos de predicción bidireccional ponderada se presentan simplemente para proporcionar contexto para las técnicas de señalización descritas en la presente divulgación. Sin embargo, se contempla que las técnicas de la presente divulgación también se pueden implementar junto con la predicción bidireccional no ponderada o modos de predicción bidireccional ponderada distintos de los descritos. Por ejemplo, se contempla que las técnicas de la presente divulgación se pueden implementar usando todos los diversos modos de predicción bidireccional incluidos en la norma H.265.

60 **[0043]** El codificador de vídeo 22 puede seleccionar un tamaño de bloque de vídeo y un tamaño de división que proporcione métricas de distorsión de velocidad favorables. De acuerdo con un aspecto de esta divulgación, el codificador de vídeo 22 puede determinar una dirección de predicción para una división del bloque de vídeo. Si se

elige la predicción bidireccional, el codificador de vídeo 22 puede determinar además uno de los modos de predicción bidireccional ponderada anteriores para la división del bloque de vídeo. Para seleccionar tamaños de bloques de vídeo que proporcionen métricas de distorsión de velocidad favorables, las métricas de distorsión de velocidad se pueden analizar tanto para bloques de vídeo grandes (por ejemplo, 32x32, 64x64 o mayores) como para bloques de vídeo pequeños (por ejemplo, 16x16 o más pequeños) con diferentes esquemas de división (es decir, diferentes tamaños de división). Como se ilustrará a continuación, un bloque de vídeo de NxN se puede dividir en una sola división de NxN, dos divisiones de N/2xN, dos divisiones de NxN/2 o cuatro divisiones de N/2xN/2. En algunos casos, una división se puede dividir aún más.

**[0044]** Un codificador puede comparar métricas de distorsión de velocidad entre bloques de vídeo de diferentes tamaños y esquemas de división, tales como bloques de vídeo de 16x16 con múltiples esquemas de división, bloques de vídeo de 32x32 con múltiples esquemas de división y bloques de vídeo de 64x64 con múltiples esquemas de división, para una serie codificada de bloques de vídeo, tal como una trama o un corte. El codificador puede seleccionar el tamaño del bloque de vídeo y el esquema de división que da como resultado la mejor distorsión de velocidad y codificar el bloque de vídeo usando el tamaño de bloque de vídeo seleccionado, es decir, el tamaño del bloque de vídeo con la mejor distorsión de velocidad.

**[0045]** La selección se puede basar en la codificación de la trama o del corte en tres o más pasadas, por ejemplo, una primera pasada usando bloques de vídeo de 16x16 píxeles, una segunda pasada usando bloques de vídeo de 32x32 píxeles y una tercera pasada usando bloques de vídeo de 64x64 píxeles, y comparando las métricas de distorsión de velocidad para cada pasada. De esta manera, un codificador puede mejorar la distorsión de velocidad variando el tamaño del bloque de vídeo y el esquema de división y seleccionando el tamaño de bloque de vídeo y el esquema de división que da como resultado la distorsión de velocidad mejor u óptima para una serie dada de bloques de vídeo, como un corte o una trama. El codificador puede transmitir además información sintáctica para la serie de bloques de vídeo, por ejemplo, como parte de un encabezado de trama o un encabezado de corte, que identifica el tamaño de los bloques de vídeo usados en la serie de bloques de vídeo. En un encabezado de bloque, el codificador puede transmitir además elementos sintácticos separados que identifican el esquema de división usado y la dirección de predicción para cada división del bloque. Si un bloque se divide en 4 subbloques, el codificador puede transmitir en un encabezado de subbloque elementos sintácticos separados que identifican el esquema de división para el subbloque y una dirección de predicción.

**[0046]** Extendiendo estos conceptos a la terminología emergente para H.265, la selección se puede basar en la codificación de la trama o el corte en tres o más pasadas, por ejemplo, una primera pasada con LCU de 16x16, una segunda pasada con LCU de 32x32 y una tercera pasada con LCU de 64x64, y comparando métricas de distorsión de velocidad para cada pasada. De esta manera, un codificador puede mejorar la distorsión de la velocidad variando el tamaño de LCU y el esquema de división y seleccionando el tamaño de LCU y el esquema de división que da como resultado la mejor u óptima distorsión de velocidad para una serie dada de bloques de vídeo, como un corte o una trama. El codificador puede transmitir además información sintáctica para la serie de bloques de vídeo, por ejemplo, como parte de un encabezado de trama o un encabezado de corte, que identifica el tamaño de LCU usado en la serie de bloques de vídeo. En un encabezado de LCU, el codificador puede transmitir además elementos sintácticos separados que identifican el esquema de división usado para la LCU y la dirección de predicción para cada división de la LCU. Si una LCU se divide en 4 CU, el codificador puede transmitir en un encabezado de CU elementos sintácticos separados que identifican el esquema de división para la CU y una dirección de predicción.

**[0047]** Para algunas/os tramas o cortes de vídeo, los bloques de vídeo grandes pueden presentar ahorros sustanciales de velocidad de bits y, de este modo, producir los mejores resultados de distorsión de velocidad, dada una distorsión relativamente baja. Sin embargo, para otras/os tramas o cortes de vídeo, los bloques de vídeo más pequeños pueden presentar menos distorsión, superando la velocidad de bits en el análisis de costes de distorsión de velocidad. De ahí que, en diferentes casos, 64x64, 32x32 o 16x16 pueden ser apropiados para diferentes tramas o cortes de vídeo, por ejemplo, dependiendo del contenido y la complejidad del vídeo. De forma similar, diferentes esquemas de división también pueden ser apropiados para diferentes bloques de vídeo dependiendo del contenido y la complejidad del vídeo.

**[0048]** Se pueden generar y codificar dos o más elementos sintácticos separados en el flujo de bits para identificar un tamaño de división para el bloque de vídeo y una o más direcciones de predicción para las divisiones. El dispositivo 12 transmite los dos o más elementos sintácticos separados al descodificador 28. El descodificador 28 puede descodificar e interpretar el/los elemento(s) sintáctico(s), y en base a el/los elemento(s) sintáctico(s), el descodificador 28 puede reconstruir el bloque de vídeo.

**[0049]** La FIG. 2 es un diagrama de bloques que ilustra un ejemplo de un codificador de vídeo 50 que puede realizar las técnicas acordes con esta divulgación. El codificador de vídeo 50 puede corresponder al codificador de vídeo 22 del dispositivo de origen 12, o a un codificador de vídeo de un dispositivo diferente. El codificador de vídeo 50 puede realizar intracodificación e intercodificación de bloques dentro de tramas de vídeo, aunque los componentes de intracodificación no se muestran en la FIG. 2 para facilitar la ilustración. La intracodificación se basa en la predicción espacial para reducir o retirar la redundancia espacial en el vídeo dentro de una trama de vídeo dada. La intercodificación se basa en la predicción temporal para reducir o retirar la redundancia temporal en el vídeo dentro de

tramas contiguas de una secuencia de vídeo. El intramodo (modo I) se puede referir al modo de compresión con base espacial, y los intermodos, tales como de predicción (modo P) o bidireccional (modo B), se pueden referir a los modos de compresión con base temporal.

5 **[0050]** Como se muestra en la FIG. 2, el codificador de vídeo 50 recibe un bloque de vídeo actual dentro de una trama o un corte de vídeo que se va a codificar. En el ejemplo de la FIG. 2, el codificador de vídeo 50 incluye una unidad de predicción 31, que incluye una unidad de estimación de movimiento 32 y una unidad de compensación de movimiento 35, una memoria 34, un sumador 48, una unidad de transformada 38, una unidad de cuantificación 40 y una unidad de codificación por entropía 46. Para la reconstrucción de bloques de vídeo, el codificador de vídeo 50 incluye también una unidad de cuantificación inversa 42, una unidad de transformada inversa 44 y un sumador 51. El descodificador de vídeo 50 puede incluir también un filtro de desbloqueo (no mostrado) para filtrar los delimitadores de bloque para retirar distorsiones de efecto pixelado del vídeo reconstruido. Si se desea, el filtro de desbloqueo filtraría típicamente la salida del sumador 51.

15 **[0051]** Durante el proceso de codificación, el codificador de vídeo 50 recibe un bloque de vídeo que se va a codificar, y la unidad de estimación de movimiento 32 y la unidad de compensación de movimiento 35 realizan una codificación interpredictiva. La unidad de estimación de movimiento 32 y la unidad de compensación de movimiento 35 pueden estar altamente integradas dentro de la unidad de predicción 31, pero se muestran por separado con propósitos ilustrativos. Se considera típicamente que la estimación de movimiento es el proceso de generar vectores de movimiento, que estiman un movimiento para bloques de vídeo. Un vector de movimiento, por ejemplo, puede indicar el desplazamiento de un bloque predictivo dentro de una trama predictiva (u otra serie de bloques de vídeo) con respecto al bloque actual que se codifica dentro de la trama actual (u otra serie de bloques de vídeo). La compensación de movimiento se considera típicamente el proceso de extraer o generar el bloque predictivo en base al vector de movimiento determinado por la estimación de movimiento. De nuevo, la unidad de estimación de movimiento 32 y la unidad de compensación de movimiento 35 se pueden integrar funcionalmente. Con propósitos demostrativos, las técnicas descritas en esta divulgación se describen como realizadas por la unidad de compensación de movimiento 35, pero más en general, se podrían realizar por cualquier parte de la unidad de predicción 31.

30 **[0052]** La unidad de estimación de movimiento 32 selecciona el vector de movimiento apropiado para que se codifique el bloque de vídeo comparando el bloque de vídeo con los bloques de vídeo de una o más series predictivas de bloques de vídeo (por ejemplo, una trama previa y/o futura en términos de tiempo o temporalmente). La unidad de estimación de movimiento 32 puede, como ejemplo, seleccionar un vector de movimiento para una trama B de varias formas. De una forma, la unidad de estimación de movimiento 32 puede seleccionar una trama previa o futura de un primer conjunto de tramas (denominado lista 0) y determinar un vector de movimiento usando solamente esta trama previa o futura a partir de la lista 0. De forma alternativa, la unidad de estimación de movimiento 32 puede seleccionar una trama previa o futura de un segundo conjunto de tramas (denominado lista 1) y determinar un vector de movimiento usando solamente esta trama previa o futura a partir de la lista 1. Aún en otro ejemplo, la unidad de estimación de movimiento 32 puede seleccionar una primera trama a partir de la lista 0 y una segunda trama a partir de la lista 1 y seleccionar uno o más vectores de movimiento de la primera trama de la lista 0 y de la segunda trama de la lista 1. Esta tercera forma de predicción se puede denominar estimación de movimiento bipredictiva. El vector de movimiento seleccionado para cualquier lista dada puede apuntar a un bloque de vídeo predictivo que es más similar al bloque de vídeo que se está codificando, por ejemplo, como se define por una métrica tal como la suma de la diferencia absoluta (SAD) o la suma de la diferencia cuadrática (SSD) de los valores de píxeles del bloque predictivo en relación con los valores de píxeles del bloque que se codifica.

45 **[0053]** En un ejemplo, si la unidad de estimación de movimiento 32 selecciona la estimación de movimiento bidireccional para un bloque de vídeo, se pueden usar tres algoritmos o modos bipredictivos compensados por movimiento para predecir una trama B o partes de la misma, tales como bloques de vídeo, macrobloques, LCU, CU y PU o cualquier otra parte discreta y/o contigua de una trama B. Un primer algoritmo o modo bipredictivo compensado por movimiento, que comúnmente se denomina predicción ponderada predeterminada, puede implicar la aplicación de pesos predeterminados a cada bloque de vídeo identificado de la primera trama de la lista 0 y de la segunda trama de la lista 1. Los pesos predeterminados se pueden programar de acuerdo con la norma, y a menudo se seleccionan para que sean iguales para la predicción ponderada predeterminada. Los bloques ponderados de las primera y segunda tramas se añaden y se dividen a continuación por el número total de tramas usadas para predecir la trama B, por ejemplo, dos en este caso.

60 **[0054]** Las unidades B usan dos listas de imágenes de referencia previamente codificadas, la lista 0 y la lista 1. Estas dos listas pueden cada una contener imágenes codificadas pasadas y/o futuras en orden temporal. Los bloques en una unidad B se pueden predecir de una entre varias maneras: predicción compensada por movimiento de una imagen de referencia de lista 0, predicción compensada por movimiento de una imagen de referencia de lista 1 o predicción compensada por movimiento de la combinación de imágenes de referencia de ambas listas, lista 0 y lista 1. Para obtener la combinación de las imágenes de referencia, tanto de la lista 0 como de la lista 1, se obtienen dos áreas de referencia compensadas por movimiento de imágenes de referencia de la lista 0 y de la lista 1, respectivamente. Su combinación se usará para predecir el bloque actual.

65

**[0055]** Como se mencionó, las unidades B pueden permitir tres tipos de predicción ponderada. Para simplificar, solo se muestra a continuación la predicción hacia delante en la predicción unidireccional, aunque también se podría usar la predicción hacia atrás. La predicción ponderada predeterminada se puede definir mediante las siguientes ecuaciones para la predicción unidireccional y la predicción bidireccional, respectivamente.

5

Predicción unidireccional:  $\text{pred}(i,j) = \text{pred0}(i,j)$

Predicción bidireccional:  $\text{pred}(i,j) = (\text{pred0}(i,j) + \text{pred1}(i,j)+1) \gg 1$

10 donde  $\text{pred0}(i,j)$  y  $\text{pred1}(i,j)$  son datos de predicción de la lista 0 y la lista 1.

**[0056]** La predicción ponderada implícita se puede definir mediante las siguientes ecuaciones para la predicción unidireccional y la predicción bidireccional, respectivamente.

15 Predicción unidireccional:  $\text{pred}(i,j) = \text{pred0}(i,j)$

Predicción bidireccional:  $\text{pred}(i,j) = (\text{pred0}(i,j)*w0+\text{pred1}(i,j)*w1+32)\gg 6$

20 En este caso, cada predicción se ajusta a escala mediante un factor de ponderación  $w0$  o  $w1$ , donde  $w0$  y  $w1$  se calculan en base a la posición temporal relativa de las imágenes de referencia de la lista 0 y la lista 1.

**[0057]** La predicción ponderada explícita se puede definir mediante las siguientes ecuaciones para la predicción unidireccional y la predicción bidireccional, respectivamente.

25 Predicción unidireccional:  $\text{pred}(i,j) = (\text{pred0}(i,j)*w0+2^{r-1})\gg r + o1$

Predicción bidireccional:

30

$$\text{pred}(i,j)=(\text{pred0}(i,j)*w0+\text{pred1}(i,j)*w1+2^r)\gg(r+1)+((o1+o2+1)\gg 1)$$

En este caso, los factores de ponderación se determinan por el codificador y se transmiten en el encabezado del corte, y  $o1$  y  $o2$  son desplazamientos de imagen para las imágenes de referencia de la lista 0 y la lista 1, respectivamente.

**[0058]** La tabla 1 a continuación muestra 23 tipos de bloques diferentes (marcados 0-22) para codificar un bloque de vídeo de 16x16 de una unidad B, aunque el mismo concepto se puede extender a un bloque de vídeo de NxN. La columna marcada "Tipo de bloque" muestra una notación abreviada separada del Número de tipo de bloque para identificar el tipo de bloque. La columna marcada "tamaño de división" identifica cómo se divide el bloque de vídeo (D-16x16, 16x16, 8x16, 16x8 u 8x8). Para los tipos de bloque que usan el modo directo (D-16x16), ningún vector de movimiento y, por lo tanto, ninguna dirección de predicción, se señala desde el codificador al descodificador porque el vector de movimiento se determina en el descodificador en base a los bloques de vídeo vecinos. Para los tipos de bloque que incluyen una división (es decir, tamaños de división de 16x16), la columna marcada "Dir. de predicción 0" identifica la dirección de predicción para la única división. Para los tipos de bloque que incluyen dos divisiones (es decir, tamaños de división de 8x16 y 16x8), la columna marcada "Dir. de predicción 0" identifica una dirección de predicción para la primera división, y la columna marcada "Dir. de predicción 1" identifica una dirección de predicción para la segunda división. Como se analiza anteriormente, las direcciones de predicción incluyen usar solo la lista 0 (L0), solo la lista 1 (L1) y la estimación de movimiento bipredictiva (Bi) usar tanto la lista 0 como la lista 1.

Tabla 1

Tipo de bloque #	Tipo de bloque	Tamaño de la división	Dir. de predicción 0	Dir. de predicción 1
0	B_Direct_16x16	D-16x16	-	-
1	B_L0_16x16	16x16	L0	-
2	B_L1_16x16	16x16	L1	-
3	B_Bi_16x16	16x16	Bi	-
4	B_L0_L016x8	16x8	L0	L0
5	B_L0_L08x16	8x16	L0	L0
6	B_L1_L1_16x8	16x8	L1	L1
7	B_L1_L1_8x16	8x16	L1	L1
8	B_L0_L1_16x8	16x8	L0	L1

Tipo de bloque #	Tipo de bloque	Tamaño de la división	Dir. de predicción 0	Dir. de predicción 1
9	B_L0_L1_8x16	8x16	L0	L1
10	B_L1_L0_16x8	16x8	L1	L0
11	B_L1_L0_8x16	8x16	L1	L0
12	B_L0_Bi_16x8	16x8	L0	Bi
13	B_L0_Bi_8x16	8x16	L0	Bi
14	B_L1_Bi_16x8	16x8	L1	Bi
15	B_L1_Bi_8x16	8x16	L1	Bi
16	B_Bi_L0_16x8	16x8	Bi	L0
17	B_Bi_L0_8x16	8x16	Bi	L0
18	B_Bi_L1_16x8	16x8	Bi	L1
19	B_Bi_L1_8x16	8x16	Bi	L1
20	B_Bi_Bi_16x8	16x8	Bi	Bi
21	B_Bi_Bi_8x16	8x16	Bi	Bi
22	B_8x8	8x8	-	-

[0059] En lugar de que cada uno de los 23 tipos de bloques anteriores tenga un solo elemento sintáctico único (es decir, 23 elementos sintácticos diferentes), las técnicas de la presente divulgación incluyen usar elementos sintácticos separados para cada tamaño de división, dirección de predicción 0 y dirección de predicción 1. En consecuencia, los 23 tipos de bloques anteriores se pueden representar como una combinación de 5 tamaños de división (D-16x16, 16x16, 8x16, 16x8 y 8x8) y 3 direcciones de predicción (L0, L1 y Bi). Por tanto, usando la opción 12 como ejemplo, para un bloque de vídeo con una división de 16x8 donde la primera división se predice usando la lista 0 y la segunda división se predice usando tanto la lista 0 como la lista 1, se podrían generar tres elementos sintácticos separados. El primer elemento sintáctico podría identificar divisiones de 16x8; el segundo elemento sintáctico podría identificar la lista 0 (L0), y el tercer elemento sintáctico podría identificar la predicción bidireccional (Bi).

[0060] El tipo de bloque 22 de la tabla 1 corresponde al bloque de vídeo de 16x16 que se divide en 4 subbloques de 8x8, en cuyo caso cada subbloque podría usar su propio esquema de división. Por lo tanto, en respuesta a la recepción de un elemento sintáctico que indica una división de 8x8 para un bloque de vídeo de 16x16, un codificador de vídeo sabrá buscar sintaxis adicional en un encabezado de subbloque para indicar un tipo de subbloque. Como con los ejemplos de tipo de bloque anteriores, el tipo de subbloque se puede señalar usando dos o más elementos sintácticos separados. Para un subbloque de 8x8, que indica un tamaño de división y una dirección de predicción para las divisiones del subbloque, un primer elemento sintáctico puede identificar el tamaño de la división y un segundo elemento sintáctico puede identificar la dirección de predicción. La tabla 2 a continuación muestra un ejemplo de los tipos de subbloques (marcados 0-13) que se podrían usar para un subbloque de 8x8. La columna marcada "Tipo de subbloque #" proporciona un número para cada tipo de subbloque, mientras que la columna marcada "Tipo de subbloque" proporciona una notación abreviada para el tipo de subbloque. La columna marcada "Tamaño de división" identifica un esquema de división para el subbloque, y la columna marcada "Dir. de predicción" identifica una dirección de predicción para las divisiones del subbloque.

Tabla 2

Tipo de subbloque #	Tipo de subbloque	Tamaño de la división	Dir. de predicción
0	B_Direct_8x8	D-8x8	-
1	B_L0_8x8	8x8	L0
2	B_L1_8x8	8x8	L1
3	B_Bi_8x8	8x8	Bi
4	B_L0_8x4	8x4	L0
5	B_L0_4x8	4x8	L0
6	B_L1_8x4	8x4	L1
7	B_L1_4x8	4x8	L1
8	B_Bi_8x4	8x4	Bi

Tipo de subbloque #	Tipo de subbloque	Tamaño de la división	Dir. de predicción
9	B_Bi_4x8	4x8	Bi
10	B_LO_4x4	4x4	L0
11	B_L1_4x4	4x4	L1
12	B_Bi_4x4	4x4	Bi

**[0061]** En algunas implementaciones, como en el ejemplo de la tabla 2, ambas divisiones de 8x4 del subbloque pueden usar la misma dirección de predicción. Por ejemplo, ambas divisiones del subbloque tipo 4 de la tabla 2 tienen la misma dirección de predicción (L0). Sin embargo, en otras implementaciones, diferentes divisiones de un subbloque pueden tener cada una dirección de predicción única, similar a la tabla 1 pero con divisiones más pequeñas.

**[0062]** Un bloque de vídeo puede tener múltiples niveles de división. Por ejemplo, un corte o una trama puede contener un elemento sintáctico en el encabezado de corte o el encabezado de trama que indica que el corte o la trama tiene un tamaño de LCU de 64x64. Cada LCU, señalizada en un encabezado de LCU, puede tener elementos sintácticos separados que indican un tamaño de división y una o dos direcciones de predicción para divisiones de la LCU, o el encabezado de LCU puede tener un elemento sintáctico que indica que la LCU está dividida en 4 CU de 32x32. Cada CU de 32x32, señalizada en un encabezado de CU, puede tener elementos sintácticos separados que indican un tamaño de división y una o dos direcciones de predicción de las divisiones, o puede tener un elemento sintáctico que indica que la CU de 32x32 se divide además en 4 CU de 16x16. Del mismo modo, cada encabezado de CU de 16x16 puede tener elementos sintácticos separados que indiquen un tamaño de división y una o dos direcciones de predicción o un elemento sintáctico que indique una división adicional.

**[0063]** Como con los tipos de bloque analizados en la tabla 1, los tipos de subbloques de la tabla 2 se pueden señalar usando elementos sintácticos separados. En lugar de que cada uno de los 13 tipos de bloques anteriores tenga un solo elemento sintáctico único (es decir, 13 elementos sintácticos diferentes), las técnicas de la presente divulgación incluyen el uso de elementos sintácticos separados para cada tamaño de división y dirección de predicción. En consecuencia, los 13 tipos de bloques anteriores se pueden representar como una combinación de 5 tamaños de división (D-8x8, 8x8, 4x8, 8x4 y 4x4) y 3 direcciones de predicción (L0, L1 y Bi). Por tanto, usando la opción 7 como un ejemplo, para un subbloque de vídeo con una división de 4x8 donde las divisiones se deben predecir usando la lista 1, se podrían generar dos elementos sintácticos separados. El primer elemento sintáctico podría identificar la división de 4x8, y el segundo elemento sintáctico podría identificar la lista 1 (L1).

**[0064]** Una vez que la unidad de compensación de movimiento 35 identifica los datos de predicción deseados, como se describe en el presente documento, el codificador de vídeo 50 forma un bloque de vídeo residual restando los datos de predicción del bloque de vídeo original que se está codificando. El sumador 48 representa el componente o los componentes que realizan esta operación de resta. La unidad de transformada 38 aplica una transformada, tal como una transformada de coseno discreta (DCT) o una transformada conceptualmente similar, al bloque residual, produciendo un bloque de vídeo que comprende coeficientes de bloque de transformada residual. La unidad de transformada 38, por ejemplo, puede realizar otras transformadas, tales como las definidas por una norma de codificación, que sean conceptualmente similares a la DCT. También se podrían usar transformadas de ondículas, transformadas de números enteros, transformadas de subbandas u otros tipos de transformadas. En cualquier caso, la unidad de transformada 38 aplica la transformada al bloque residual, produciendo un bloque de coeficientes de transformada residuales. La transformada puede convertir la información residual de un dominio de píxel en un dominio de frecuencia.

**[0065]** La unidad de cuantificación 40 cuantifica los coeficientes de transformada residuales para reducir más la velocidad de bits. Tras la cuantificación, la unidad de codificación por entropía 46 codifica por entropía los coeficientes de transformada cuantificados. La unidad de codificación por entropía 46 codifica además los elementos sintácticos generados por la unidad de compensación de movimiento 35. Por ejemplo, la unidad de codificación por entropía 46 puede realizar la codificación de longitud variable adaptativa al contenido (CAVLC), la codificación aritmética binaria adaptativa al contexto (CABAC) u otra técnica de codificación por entropía. Como los elementos sintácticos generados por la unidad de compensación de movimiento 35 son elementos sintácticos separados, cada elemento sintáctico puede experimentar un proceso de codificación por entropía diferente. Por ejemplo, la unidad de codificación por entropía 46 puede mantener estadísticas para aplicar CAVLC a elementos sintácticos para tamaños de división y estadísticas separadas para aplicar CAVLC a elementos sintácticos para direcciones de predicción. Tras la codificación por entropía por la unidad de codificación por entropía 46, el vídeo codificado se puede transmitir a otro dispositivo o archivar para su transmisión o recuperación posterior. El flujo de bits codificado puede incluir bloques residuales codificados por entropía, vectores de movimiento para dichos bloques y otra sintaxis tal como la sintaxis descrita en el presente documento.

**[0066]** La unidad de cuantificación inversa 42 y la unidad de transformada inversa 44 aplican la cuantificación inversa y la transformada inversa, respectivamente, para reconstruir el bloque residual en el dominio de píxel, por ejemplo, para su uso posterior como bloque de referencia de la manera descrita anteriormente. El sumador 51 suma el bloque

residual reconstruido al bloque de predicción compensado por movimiento, producido por la unidad de compensación de movimiento 35, para producir un bloque de vídeo reconstruido para su almacenamiento en la memoria 34. El bloque de vídeo reconstruido puede usarse por la unidad de estimación de movimiento 32 y por la unidad de compensación de movimiento 35 como bloque de referencia para intercodificar un bloque en una trama de vídeo posterior.

**[0067]** La FIG. 3 es un diagrama de bloques que ilustra con mayor detalle un ejemplo de la unidad de compensación de movimiento 35 de la FIG. 2. Como se analiza anteriormente, la unidad de compensación de movimiento 35 podría muy probablemente estar altamente integrada con la unidad de estimación de movimiento 32 dentro de la unidad de predicción 31. Por lo tanto, se debe entender que la división de funcionalidad entre la unidad de estimación de movimiento 32 y la unidad de compensación de movimiento 35 en esta divulgación es solo para propósitos de explicación. Como se muestra en el ejemplo de la FIG. 3, la unidad de compensación de movimiento 35 se acopla a la memoria 34, que almacena el primer y el segundo conjuntos de unidades de referencia o tramas de referencia como la lista 0 52A y la lista 1 52B. Además, la memoria 34 puede almacenar los datos de vídeo actuales 53 codificados. La memoria 34 puede comprender una estructura de memoria compartida, o posiblemente varias memorias diferentes, unidades de almacenamiento, memorias intermedias u otros tipos de almacenamiento que faciliten el almacenamiento de cualesquiera datos analizados en el presente documento. La lista 0 52A y la lista 1 52B son datos asociados con dos unidades predictivas diferentes, por ejemplo, datos de dos tramas o cortes o macrobloques diferentes, de acuerdo con la predicción bidireccional. Una vez más, la predicción bidireccional no se limita necesariamente a ninguna dirección de predicción y, por tanto, la lista 0 52A y la lista 1 52B pueden almacenar datos de dos tramas o cortes previas/os, dos tramas o cortes posteriores, o una trama previa o un corte previo y una trama posterior o un corte posterior. Además, en algunos casos, la lista 0 52A y/o la lista 1 52B podrían cada una incluir datos asociados con múltiples tramas, cortes o bloques de vídeo. La lista 0 52A y/o la lista 1 52B son simplemente dos conjuntos diferentes de datos predictivos posibles, y cada lista puede incluir una trama o un corte, o varios tramas, cortes o bloques de vídeo en cualquier dirección en relación con el bloque de vídeo actual que se codifica.

**[0068]** Como se muestra en la FIG. 3, la unidad de compensación de movimiento 35 incluye una unidad de predicción ponderada predeterminada 54, una unidad de predicción ponderada implícita 56 y una unidad de predicción ponderada explícita 58 para determinar un tipo de predicción bidireccional que se va a usar para codificar una división de un bloque de vídeo. La unidad de compensación de movimiento 35 incluye además una unidad de predicción de lista 0 60 y una unidad de predicción de lista 1 62 para determinar si la predicción hacia delante (lista 0, por ejemplo) o la predicción hacia atrás (lista 1, por ejemplo) se debe usar para codificar un bloque de vídeo. Como se mencionó previamente, la predicción bidireccional puede no estar limitada a direcciones temporales específicas. Por lo tanto, aunque los términos hacia delante y hacia atrás cuando se analiza la predicción todavía se usan comúnmente para referirse a dos listas diferentes, las dos listas no necesariamente tienen ninguna dirección temporal específica. Las unidades 54, 56 y 58 respectivamente pueden realizar predicciones ponderadas predeterminadas, predicciones ponderadas implícitas y predicciones ponderadas explícitas como se describe en el presente documento. La unidad de análisis de distorsión de velocidad (R-D) 64 puede seleccionar uno de los datos de predicción ponderados entre estas posibilidades o puede seleccionar uno de los datos de predicción de lista 0 o de predicción de lista 1, y puede implementar técnicas de esta divulgación para facilitar el proceso de selección.

**[0069]** La unidad de compensación de movimiento 35 también puede incluir una unidad de redondeo que provoca que una o más de las unidades 54, 56, 58, 60 y 62 generen versiones tanto redondeadas como no redondeadas de los datos de predicción ponderados respectivos y una unidad de cálculo de desplazamiento, que calcula el desplazamiento como una diferencia entre un promedio de valores de bloque de vídeo de un bloque que se codifica y un promedio de valores de bloque de vídeo del bloque de predicción. Por simplicidad, la unidad de redondeo y la unidad de desplazamiento no se muestran en la FIG. 3.

**[0070]** La unidad de análisis R-D 64 puede analizar los diferentes datos predictivos de las unidades 54, 56, 58, 60 y 62, y puede seleccionar los datos predictivos que generan los mejores resultados en términos de calidad, o en términos de velocidad y distorsión. Dependiendo de varias preferencias de implementación, la unidad de análisis R-D puede configurarse para equilibrar los intereses en competencia de la velocidad de codificación (es decir, el número de bits) y el nivel de calidad de imagen. La unidad de análisis R-D 64 emite los datos predictivos seleccionados, que se pueden restar del bloque de vídeo que se codifica por medio del sumador 48 (FIG. 2). Como se analiza anteriormente, además de analizar qué tipo de datos predichos genera los mejores resultados, la unidad de análisis R-D 64 también puede analizar qué tamaño de unidad codificada y qué tipo de división para la unidad codificada genera los mejores resultados.

**[0071]** Los elementos sintácticos se pueden usar para informar a un descodificador de la manera o el procedimiento que usó la unidad de compensación de movimiento 35 para codificar los datos y, por lo tanto, qué manera y procedimiento debería usar el descodificador para generar los datos predictivos ponderados. De acuerdo con la presente divulgación, los elementos sintácticos pueden incluir elementos sintácticos separados para señalar un tamaño de división y una dirección de predicción. Si se selecciona la predicción bidireccional, los elementos sintácticos, por ejemplo, pueden indicar además si se debe usar la predicción ponderada predeterminada, implícita o explícita. Si se debe usar la predicción ponderada explícita, los elementos sintácticos pueden identificar además los factores de peso y el desplazamiento, que nuevamente pueden ser factores de peso y desplazamiento asociados con la predicción ponderada explícita, o pueden ser factores de peso que se definieron realmente por la unidad de predicción ponderada

predeterminada 54 o la unidad de predicción ponderada implícita 56 con la adición del desplazamiento de la unidad de cálculo de desplazamiento 62.

**[0072]** La FIG. 4 es un diagrama de bloques que ilustra un descodificador de vídeo 70 ejemplar, que puede realizar las técnicas de descodificación recíproca a las técnicas de codificación descritas anteriormente. El descodificador de vídeo 70 también incluye una unidad de descodificación por entropía 72, una unidad de predicción 75, una unidad de cuantificación inversa 76, una unidad de transformada inversa 78, una memoria 74 y un sumador 79. La unidad de predicción 75 puede incluir una unidad de compensación de movimiento (MC) 88, así como componentes de predicción espacial, que no se muestran por simplicidad y facilidad de ilustración.

**[0073]** El descodificador de vídeo 70 puede recibir datos de vídeo codificados y dos o más elementos sintácticos separados que indican un tamaño de división y una dirección de predicción. En base a los dos o más elementos sintácticos separados, la unidad de MC 86 de la unidad de predicción 75 puede generar datos de predicción ponderados que dependen de dos o más listas de datos, como se describe en el presente documento. El descodificador de vídeo 70 puede descodificar los datos de vídeo usando los datos de predicción ponderados, por ejemplo, invocando al sumador 79 para añadir los datos de predicción ponderados (por ejemplo, un bloque de predicción) a los datos residuales (por ejemplo, un bloque residual).

**[0074]** En general, la unidad de descodificación por entropía 72 recibe un flujo de bits codificado y descodifica por entropía el flujo de bits para generar coeficientes cuantificados, vectores de movimiento y otros elementos sintácticos. De acuerdo con las técnicas de esta divulgación, los otros elementos sintácticos pueden incluir dos o más elementos sintácticos separados que señalizan un tamaño de división y una dirección de predicción. La información de movimiento (por ejemplo, vectores de movimiento) y otra sintaxis se remiten a la unidad de predicción 75 para su uso en la generación de los datos predictivos. La unidad de predicción 75 realiza predicciones bidireccionales consecuentes con esta divulgación, y posiblemente implementando la predicción ponderada predeterminada, implícita o explícita de acuerdo con los elementos sintácticos recibidos. Los elementos sintácticos pueden identificar el tipo de predicción ponderada que se va a usar, los coeficientes y el desplazamiento si se va a usar una predicción ponderada explícita, y como se analizó anteriormente, pueden incluir dos o más elementos sintácticos separados que identifiquen un tamaño de división y una dirección de predicción.

**[0075]** Los coeficientes cuantificados se envían desde la unidad de descodificación por entropía 72 a la unidad de cuantificación inversa 76, que realiza la cuantificación inversa. La unidad de transformada inversa 78 a continuación transforma inversamente los coeficientes descuantificados de regreso al dominio de pixel para generar un bloque residual. El sumador 79 combina los datos de predicción (por ejemplo, un bloque de predicción) generados por la unidad de predicción 75 con el bloque residual de la unidad de transformada inversa 78 para crear un bloque de vídeo reconstruido, que se puede almacenar en la memoria 74 y/o emitir del descodificador de vídeo 70 como salida de vídeo descodificada.

**[0076]** La FIG. 5 es un diagrama de flujo que ilustra un proceso de ejemplo un codificador de vídeo consecuente con esta divulgación. La FIG. 5 se describirá desde la perspectiva del codificador de vídeo 50 de la FIG. 2. Como se muestra en la FIG. 5, la unidad de compensación de movimiento 35 genera los primeros datos de predicción ponderados (501), y genera los segundos datos de predicción ponderados (502). La unidad de compensación de movimiento 35 selecciona a continuación los datos de predicción de los primeros y segundos datos de predicción ponderados en base a un análisis de distorsión de velocidad (503). En particular, la unidad de compensación de movimiento 35 puede determinar métricas de coste para los primeros y segundos datos de predicción ponderados que cuantifican y equilibran la tasa de codificación y la calidad de codificación asociada con los primeros y segundos datos de predicción ponderados, y puede seleccionar los datos de predicción con el coste más bajo en términos de velocidad y distorsión. El codificador de vídeo 50 puede a continuación codificar datos de vídeo en base a los datos de predicción seleccionados (504). Por ejemplo, el codificador de vídeo 50 puede invocar al sumador 48 para restar los datos de predicción seleccionados de los datos de vídeo que se están codificando, y a continuación invocar la unidad de transformada 38 para la transformación, la unidad de cuantificación 40 para la cuantificación y la unidad de codificación por entropía 46 para la codificación por entropía de los coeficientes residuales cuantificados y transformados. En este caso, la unidad de compensación de movimiento 35 puede generar dos o más elementos sintácticos separados para indicar un tamaño de división y una dirección de predicción para los datos de predicción, y puede remitir dichos elementos sintácticos a la unidad de codificación por entropía 46 para su inclusión en el flujo de bits codificado (505). El proceso de la FIG. 5 se puede realizar para una pluralidad de bloques de vídeo B dentro de un corte o una trama u otra serie de bloques de vídeo, y también se puede realizar para subbloques de bloques de vídeo B. Por ejemplo, el proceso de la FIG. 5 se podría realizar para cualquiera de una LCU, CU o PU.

**[0077]** La FIG. 6 es un diagrama de flujo que ilustra un proceso de ejemplo realizado por un descodificador de vídeo consecuente con esta divulgación. La FIG. 6 se describirá desde la perspectiva del descodificador de vídeo 70 de la FIG. 4. Como se muestra en la FIG. 6, el descodificador de vídeo recibe datos de vídeo codificados (601) y recibe dos o más elementos sintácticos separados que indican un tamaño de división y una dirección de predicción que se usó para codificar los datos de vídeo (602). En particular, la unidad de descodificación por entropía 72 puede recibir un flujo de bits codificado que incluye los datos de vídeo y los dos o más elementos sintácticos separados. Después de la descodificación por entropía, la unidad de descodificación por entropía 72 puede emitir los datos de vídeo como

coeficientes de transformada cuantificados, que se cuantifican inversamente por la unidad 76 y se transforman inversamente por la unidad 78. La unidad de descodificación por entropía 72 puede emitir elementos sintácticos a la unidad de predicción 75, que incluye los dos o más elementos sintácticos separados que indican un tamaño de división y una dirección de predicción que se usó para codificar los datos de vídeo, vectores de movimiento y posiblemente otra sintaxis.

**[0078]** La unidad de predicción 75 invoca la unidad de compensación de movimiento 86 para la descodificación predictiva basada en bloques. Al hacerlo, la unidad de compensación de movimiento 86 genera datos de predicción ponderados en base a los dos o más elementos sintácticos separados que indican el tamaño de división y la dirección de predicción (603). El descodificador de vídeo 70 puede a continuación descodificar los datos de vídeo usando los datos de predicción ponderados (604). En particular, el descodificador de vídeo 70 puede invocar el sumador 79 para combinar datos de predicción ponderados (por ejemplo, un bloque de predicción) con datos de vídeo residuales (por ejemplo, un bloque residual) para generar una reconstrucción de los datos de vídeo (por ejemplo, un bloque de vídeo reconstruido). El proceso de la FIG. 6 se puede realizar para una pluralidad de bloques de vídeo B dentro de un corte o trama u otra serie de bloques de vídeo, y también se puede realizar para subbloques de bloques de vídeo B. Por ejemplo, el proceso de la FIG. 5 se podría realizar para cualquiera de una LCU, CU o PU.

**[0079]** Las técnicas de esta divulgación se pueden implementar en una amplia variedad de dispositivos o aparatos, incluyendo un microteléfono inalámbrico, un circuito integrado (IC) o un conjunto de IC (es decir, un conjunto de chips). Se ha descrito cualquier componente, módulo o unidad proporcionado para enfatizar aspectos funcionales y no necesariamente requieren su realización por diferentes unidades de hardware. Las técnicas descritas en el presente documento se pueden implementar en hardware, programas informáticos, firmware o en cualquier combinación de los mismos. Cualesquiera rasgos característicos descritos como módulos, unidades o componentes se pueden implementar juntos en un dispositivo lógico integrado o por separado, como dispositivos lógicos discretos pero interoperables. En algunos casos, diversos rasgos característicos se pueden implementar como un dispositivo de circuito integrado, tal como un chip de circuito integrado o conjunto de chips.

**[0080]** Si se implementan en programas informáticos, las técnicas se pueden realizar, al menos en parte, mediante un medio legible por ordenador que comprende instrucciones que, al ejecutarse en un procesador, realizan uno o más de los procedimientos descritos anteriormente. El medio legible por ordenador puede comprender un medio de almacenamiento legible por ordenador y puede formar parte de un producto de programa informático, que puede incluir materiales de empaquetado. El medio de almacenamiento legible por ordenador puede comprender memoria de acceso aleatorio (RAM) tal como memoria de acceso aleatorio dinámica síncrona (SDRAM), memoria de solo lectura (ROM), memoria de acceso aleatorio no volátil (NVRAM), memoria de solo lectura programable y borrable eléctricamente (EEPROM), memoria flash, medios de almacenamiento de datos magnéticos u ópticos, y similares. Las técnicas se pueden realizar adicionalmente, o de forma alternativa, al menos en parte por un medio de comunicación legible por ordenador que lleve o comunique un código en forma de instrucciones o estructuras de datos y al que se pueda acceder, y que se pueda leer y/o ejecutar por un ordenador.

**[0081]** El código puede ser ejecutado por uno o más procesadores, tales como uno o más procesadores de señales digitales (DSP), microprocesadores de propósito general, circuitos integrados de aplicación específica (ASIC), matrices lógicas programables por campo (FPGA), u otra circuitería lógica integrada o discreta equivalente. En consecuencia, el término "procesador", como se usa en el presente documento, se puede referir a cualquiera de las estructuras anteriores o a cualquier otra estructura adecuada para la implementación de las técnicas descritas en el presente documento. Además, en algunos aspectos, la funcionalidad descrita en el presente documento se puede proporcionar dentro de módulos de programas informáticos y/o módulos de hardware dedicados configurados para la codificación y la descodificación, o incorporarse en un códec de vídeo combinado. Además, las técnicas se podrían implementar totalmente en uno o más circuitos o elementos lógicos.

**[0082]** La divulgación también contempla cualquiera de una variedad de dispositivos de circuitos integrados que incluyan circuitería para implementar una o más de las técnicas descritas en esta divulgación. Dicha circuitería se puede proporcionar en un único chip de circuito integrado o en múltiples chips de circuitos integrados interoperables en un denominado conjunto de chips. Dichos dispositivos de circuitos integrados pueden usarse en una variedad de aplicaciones, algunas de las cuales pueden incluir el uso en dispositivos de comunicación inalámbrica, tales como teléfonos móviles.

**[0083]** Se han descrito diversos modos de realización de la invención. Estos y otros modos de realización están dentro del alcance de las siguientes reivindicaciones.

**REIVINDICACIONES**

1. Un procedimiento para codificar un bloque de vídeo bidireccional, bloque de vídeo B, comprendiendo el procedimiento:
- 5           usar la predicción bidireccional para generar (501, 502) primeros datos de predicción y segundos datos de predicción para el bloque de vídeo B;
- 10           seleccionar (503) uno de los primeros datos de predicción y los segundos datos de predicción en base a una o más métricas de distorsión de velocidad;
- 15           en base a los datos de predicción seleccionados, generar (505) un primer elemento sintáctico que es indicativo de un tamaño de división para el bloque de vídeo B, y que no es indicativo de una dirección de predicción;
- 20           en base a los datos de predicción seleccionados, generar (505) un segundo elemento sintáctico, separado del primer elemento sintáctico, e indicativo de una dirección de predicción para la o cada división del bloque de vídeo B, en el que el segundo elemento sintáctico no es indicativo de una trama de referencia, o, generar una pluralidad de segundos elementos sintácticos, separados del primer elemento sintáctico, y respectivamente indicativos de una dirección de predicción respectiva para divisiones respectivas del bloque de vídeo B en el que la pluralidad de segundos elementos sintácticos no son indicativos de una trama de referencia, en el que la dirección de predicción se selecciona de un grupo de direcciones de predicción que comprende una primera lista, una segunda lista y tanto la primera lista como la segunda lista; y,
- 25           emitir el primer y segundo elementos sintácticos.
2. El procedimiento de la reivindicación 1, en el que seleccionar uno de los primeros datos de predicción y los segundos datos de predicción comprende evaluar la una o más métricas de distorsión de velocidad para cada uno de los primeros datos de predicción y los segundos datos de predicción, y en el que la una o más métricas de distorsión de velocidad se basan al menos en parte en una velocidad de codificación y una calidad de codificación asociada con cada uno de los primeros datos de predicción y los segundos datos de predicción.
3. El procedimiento de la reivindicación 1, en el que generar los primeros datos de predicción comprende determinar un primer vector de movimiento, y generar los segundos datos de predicción comprende determinar un segundo vector de movimiento, en el que generar el segundo elemento sintáctico se basa preferentemente al menos en parte en cuál del primer vector de movimientos o del segundo vector de movimiento corresponde a los datos de predicción seleccionados.
4. El procedimiento de la reivindicación 1, en el que generar los primeros datos de predicción comprende determinar un primer tamaño de división, y generar los segundos datos de predicción comprende determinar un segundo tamaño de división, en el que generar el primer elemento sintáctico se basa preferentemente al menos en parte en cuál del primer tamaño de división o segundo tamaño de la división corresponde a los datos de predicción seleccionados.
5. El procedimiento de la reivindicación 1, en el que emitir el primer y el segundo elemento sintáctico comprende:
- emitir una primera pluralidad de bits indicativos del primer elemento sintáctico; y
- emitir una segunda pluralidad de bits indicativos del segundo elemento sintáctico, siendo dicha primera y segunda pluralidad de bits independientes entre sí.
6. El procedimiento de la reivindicación 1, en el que el bloque de vídeo B comprende una división o dos divisiones, y en el que la pluralidad de segundos elementos sintácticos comprende un segundo elemento sintáctico y un tercer elemento sintáctico, siendo el segundo elemento sintáctico indicativo de una primera dirección de predicción para una primera división del bloque de vídeo B, y siendo el tercer elemento sintáctico indicativo de una segunda dirección de predicción para una segunda división del bloque de vídeo B.
7. El procedimiento de la reivindicación 1, que comprende además: realizar un proceso de codificación por entropía en el primer elemento sintáctico, y preferentemente realizar un segundo proceso de codificación por entropía en el segundo elemento sintáctico.
8. Un dispositivo para codificar un bloque de vídeo bidireccional, bloque de vídeo B, comprendiendo el dispositivo:
- 65           medios para generar, usando la predicción bidireccional, primeros datos de predicción y segundos datos de predicción para el bloque de vídeo B

medios para seleccionar uno de los primeros datos de predicción y los segundos datos de predicción en base a una o más métricas de distorsión de velocidad;

5 medios para generar, en base a los datos de predicción seleccionados, un primer elemento sintáctico que es indicativo de un tamaño de división para el bloque de vídeo B, y que no es indicativo de una dirección de predicción;

10 medios para generar un segundo elemento sintáctico, separado del primer elemento sintáctico, e indicativo de una dirección de predicción para la o cada división del bloque de vídeo B, en base a los datos de predicción seleccionados, en el que el segundo elemento sintáctico no es indicativo de una trama de referencia, o, generar una pluralidad de segundos elementos sintácticos, separados del primer elemento sintáctico, y respectivamente indicativos de una dirección de predicción respectiva para divisiones respectivas del bloque de vídeo B en el que la pluralidad de segundos elementos sintácticos no son indicativos de una trama de referencia, en el que la dirección de predicción se selecciona de un grupo de direcciones de predicción que comprende una primera lista, una segunda lista y tanto la primera lista como la segunda lista; y,

20 medios para emitir el primer y segundo elementos sintácticos.

9. Un procedimiento para descodificar un bloque de vídeo bidireccional, bloque de vídeo B, comprendiendo el procedimiento:

25 recibir (601) datos de vídeo codificados;

recibir (602), en los datos de vídeo codificados, un primer elemento sintáctico que es indicativo de un tamaño de división usado para codificar el bloque de vídeo B, y que no es indicativo de una dirección de predicción;

30 recibir (602), en los datos de vídeo codificados:

un segundo elemento sintáctico, separado del primer elemento sintáctico, e indicativo de una dirección de predicción usada para codificar la o cada división del bloque de vídeo B, en el que el segundo elemento sintáctico no es indicativo de una trama de referencia, o

35 una pluralidad de segundos elementos sintácticos, separados del primer elemento sintáctico, y respectivamente indicativos de una dirección de predicción respectiva para las divisiones respectivas del bloque de vídeo B, en el que la pluralidad de segundos elementos sintácticos no son indicativos de una trama de referencia,

40 en el que el primer elemento sintáctico está separado del segundo elemento sintáctico, en el que la dirección de predicción se selecciona de un grupo de direcciones de predicción que comprende una primera lista, una segunda lista, y tanto la primera lista como la segunda lista; y,

45 descodificar (604) el bloque de vídeo B basado al menos en parte en el primer elemento sintáctico recibido y el/los segundo(s) elemento(s) sintáctico(s).

10. El procedimiento de la reivindicación 9, en el que descodificar el bloque de vídeo B comprende generar datos de predicción basados al menos en parte en el primer elemento sintáctico y el segundo elemento sintáctico.

50 11. El procedimiento de la reivindicación 9, en el que recibir el primer elemento sintáctico y el segundo elemento sintáctico comprende:

recibir una primera pluralidad de bits indicativos del primer elemento sintáctico; y

55 recibir una segunda pluralidad de bits indicativos del segundo elemento sintáctico, siendo dicha primera y segunda pluralidad de bits independientes entre sí.

60 12. El procedimiento de la reivindicación 9, en el que el bloque de vídeo B comprende una división o dos divisiones, y en el que la pluralidad de segundos elementos sintácticos comprende un segundo elemento sintáctico y un tercer elemento sintáctico, siendo el segundo elemento sintáctico indicativo de una primera dirección de predicción para una primera división del bloque de vídeo B, y siendo el tercer elemento sintáctico indicativo de una segunda dirección de predicción usada para codificar una segunda división del bloque de vídeo B.

65 13. El procedimiento de la reivindicación 9, que comprende, además:

realizar un proceso de descodificación por entropía en el primer elemento sintáctico, y preferentemente realizar un segundo proceso de descodificación por entropía en el segundo elemento sintáctico.

- 5      **14.** Un dispositivo para descodificar un bloque de vídeo bidireccional, bloque de vídeo B, comprendiendo el dispositivo:
- medios para recibir los datos de vídeo codificados;
- 10           medios para recibir un primer elemento sintáctico que es indicativo de un tamaño de división usado para codificar el bloque de vídeo B, y que no es indicativo de una dirección de predicción;
- medios para recibir los datos de vídeo codificados:
- 15           un segundo elemento sintáctico, separado del primer elemento sintáctico, e indicativo de una dirección de predicción usada para codificar la o cada división del vídeo B en el que el segundo elemento sintáctico no es indicativo de una trama de referencia, o
- 20           una pluralidad de segundos elementos sintácticos, separados del primer elemento sintáctico, y respectivamente indicativos de una dirección de predicción respectiva para las divisiones respectivas del bloque de vídeo B, en el que la pluralidad de segundos elementos sintácticos no son indicativos de una trama de referencia,
- 25           en el que el primer elemento sintáctico está separado del segundo elemento sintáctico, en el que la dirección de predicción se selecciona de un grupo de direcciones de predicción que comprende una primera lista, una segunda lista, y tanto una primera lista como la segunda lista; y,
- medios para descodificar el bloque de vídeo B basado al menos en parte en el primer elemento sintáctico recibido y el/los segundo(s) elemento(s) sintáctico(s).
- 30      **15.** Un medio de almacenamiento legible por ordenador que comprende instrucciones que al ejecutarse en un procesador provocan que el procesador realice la codificación o descodificación de un bloque de vídeo bidireccional, bloque de vídeo B, de acuerdo con el procedimiento de una cualquiera de las reivindicaciones 1 a 7 o el procedimiento de una cualquiera de las reivindicaciones 9 a 13, respectivamente.

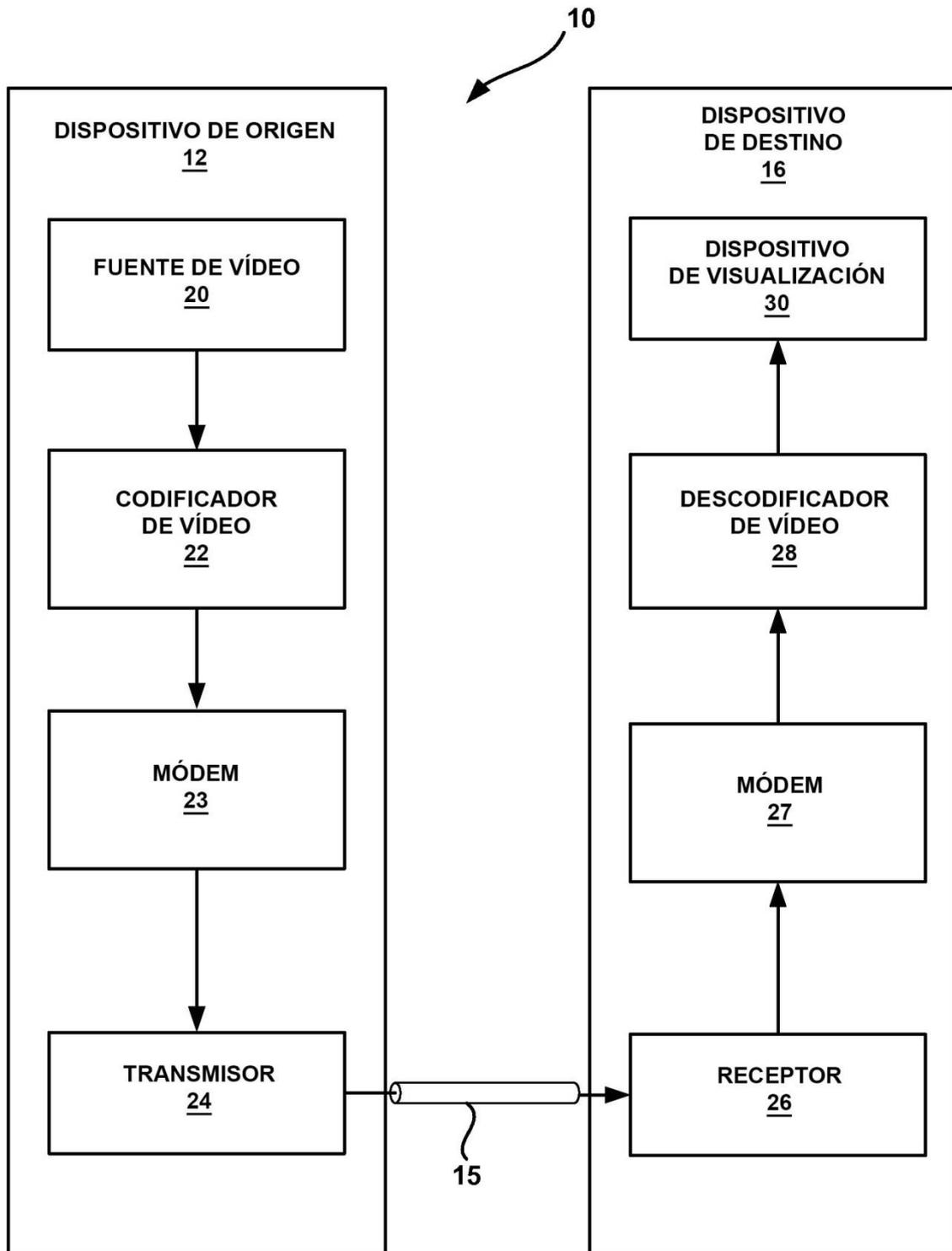


FIG. 1



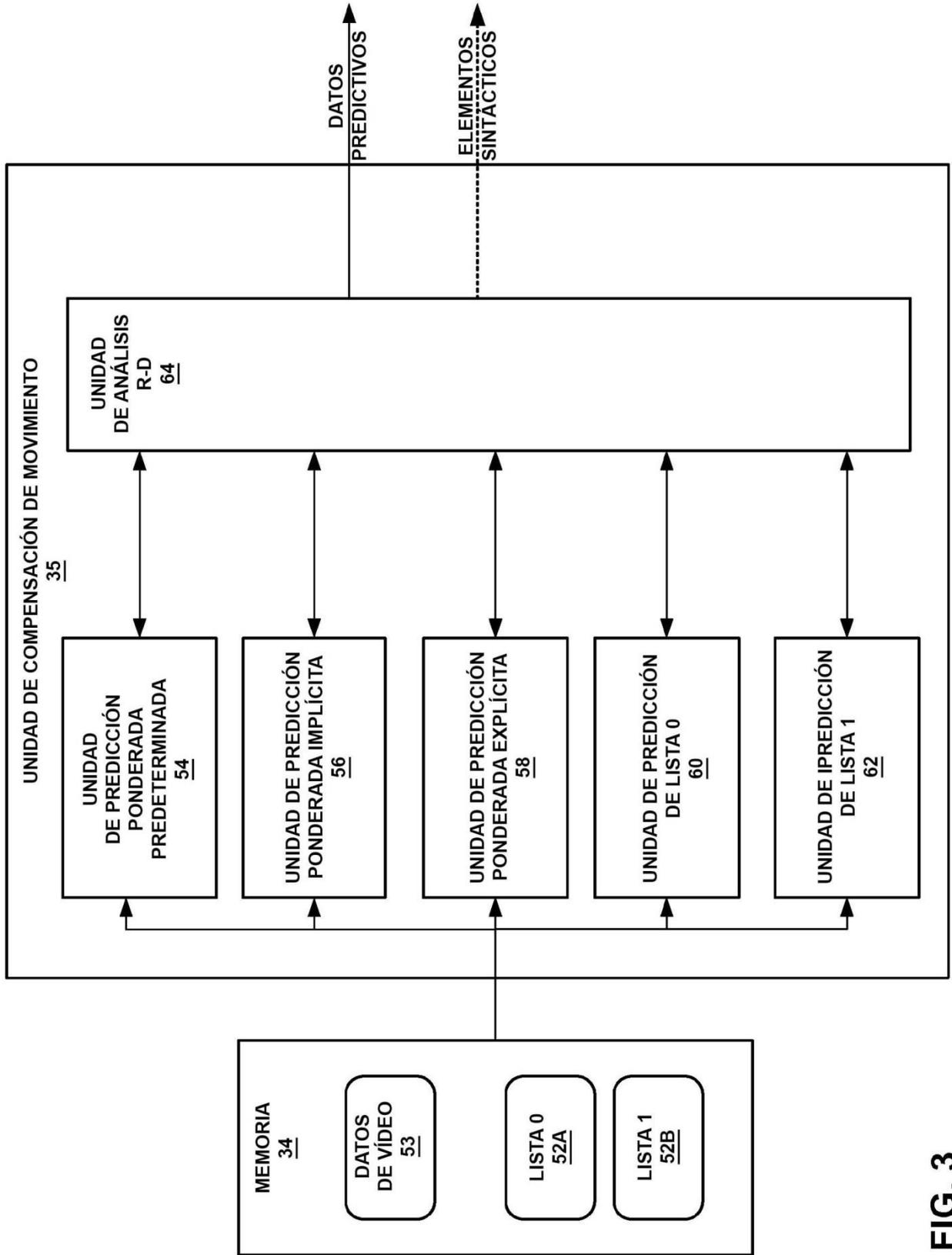


FIG. 3

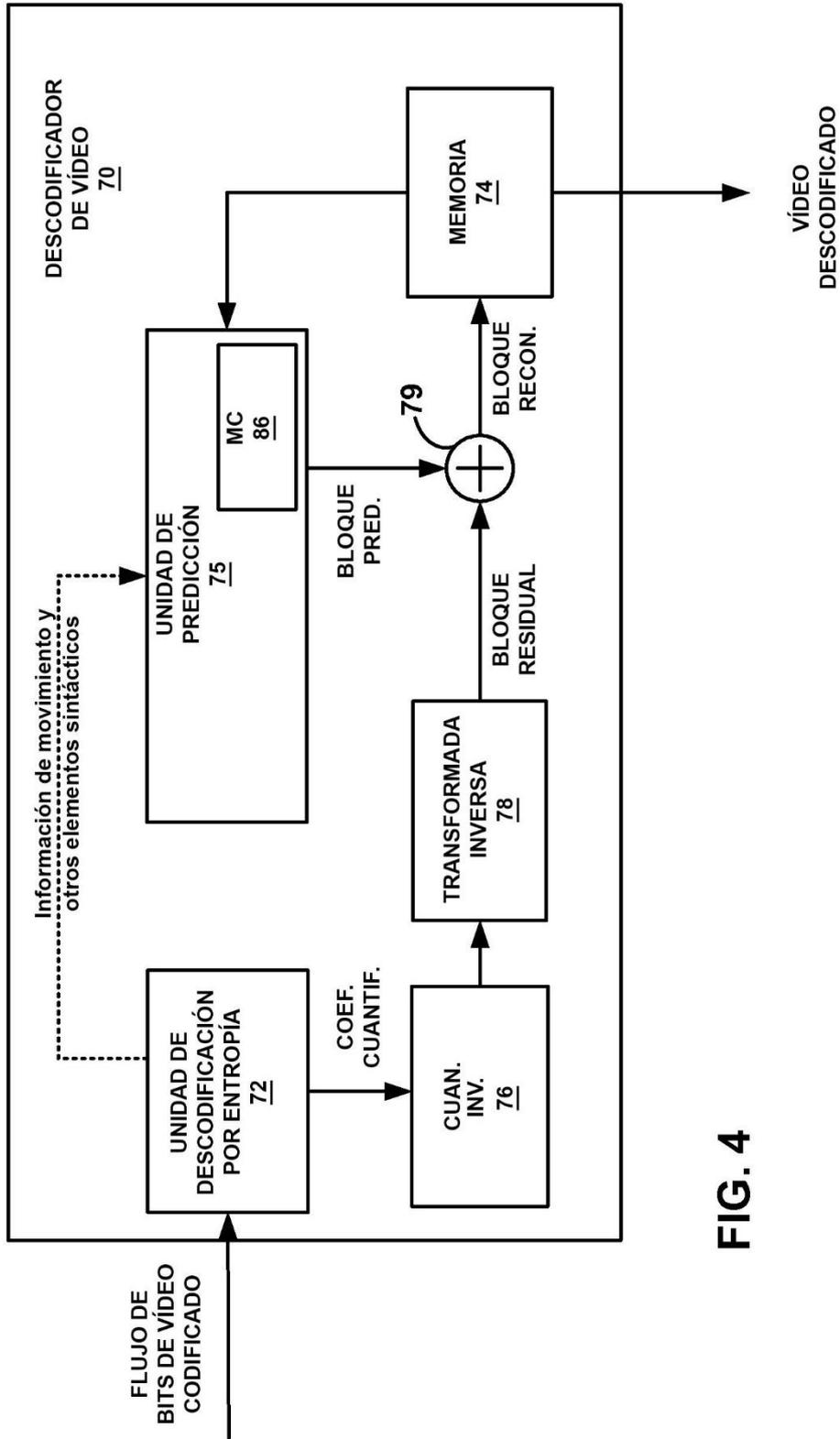
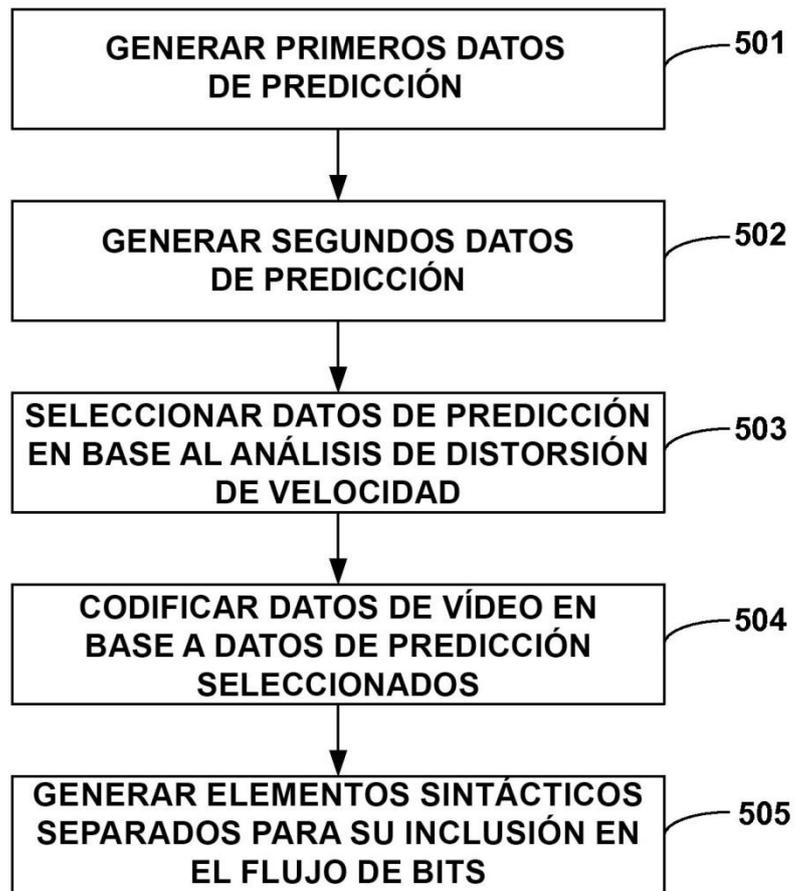
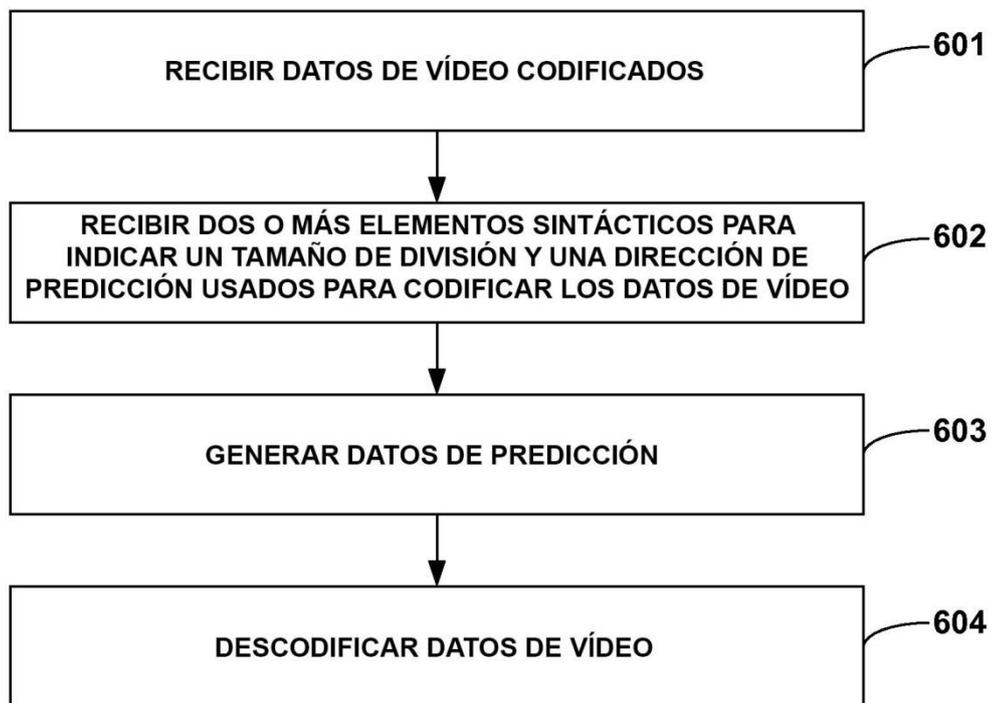


FIG. 4



**FIG. 5**



**FIG. 6**