

19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 711 449**

51 Int. Cl.:

H04N 19/52	(2014.01)
H04N 19/577	(2014.01)
H04N 19/105	(2014.01)
H04N 19/147	(2014.01)
H04N 19/176	(2014.01)
H04N 19/46	(2014.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

- 86 Fecha de presentación y número de la solicitud internacional: **29.02.2012 PCT/US2012/027136**
- 87 Fecha y número de publicación internacional: **27.09.2012 WO12128903**
- 96 Fecha de presentación y número de la solicitud europea: **29.02.2012 E 12707694 (1)**
- 97 Fecha y número de publicación de la concesión europea: **14.11.2018 EP 2689582**

54 Título: **Modo de fusión bipredictivo basado en vecinos unipredictivos y bipredictivos en la codificación de vídeo**

30 Prioridad:

21.03.2011 US 201161454862 P
29.06.2011 US 201161502703 P
23.12.2011 US 201113336799

45 Fecha de publicación y mención en BOPI de la traducción de la patente:
03.05.2019

73 Titular/es:

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

72 Inventor/es:

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

74 Agente/Representante:

FORTEA LAGUNA, Juan José

ES 2 711 449 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

Modo de fusión bipredictivo basado en vecinos unipredictivos y bipredictivos en la codificación de vídeo

5 CAMPO TÉCNICO

[0001] Esta divulgación se refiere a técnicas de codificación de vídeo usadas para comprimir datos de vídeo y, más particularmente, a modos de codificación de vídeo usados en la compresión de vídeo.

10 ANTECEDENTES

[0002] Las capacidades de vídeo digital pueden incorporarse a una amplia gama de dispositivos de vídeo, incluidos televisores digitales, sistemas de radiodifusión directa digital, dispositivos de comunicación inalámbrica, tales como auriculares de teléfonos inalámbricos, sistemas de radiodifusión inalámbrica, asistentes digitales personales (PDA), ordenadores portátiles o de escritorio, tablets, cámaras digitales, dispositivos de grabación digital, dispositivos de videojuegos, consolas de videojuegos, reproductores multimedia personales y similares. Dichos dispositivos de vídeo pueden implementar técnicas de compresión de vídeo, tales como las descritas en MPEG-2, MPEG-4 o ITU-T H.264/MPEG-4, parte 10, Codificación avanzada de vídeo (AVC), con el fin de comprimir datos de vídeo. Las técnicas de compresión de vídeo realizan una predicción espacial y/o temporal para reducir o eliminar la redundancia inherente a las secuencias de vídeo. El "Equipo de Colaboración Conjunta en Codificación de Vídeo" (JCTVC), que es una colaboración entre MPEG e ITU-T, está desarrollando normas de vídeo emergentes, tales como la norma de Codificación de Vídeo de Alta Eficiencia (HEVC), que siguen surgiendo y evolucionando. La norma HVEC emergente se denomina algunas veces H.265.

[0003] Estas y otras normas y técnicas de codificación de vídeo usan la codificación de vídeo basada en bloques. Las técnicas de codificación de vídeo basadas en bloques dividen los datos de vídeo de una trama de vídeo (o una porción de los mismos) en bloques de vídeo y luego codifican los bloques de vídeo usando técnicas de compresión predictiva basada en bloques. Los bloques de vídeo se pueden dividir además en particiones de bloques de vídeo. Los bloques de vídeo (o las particiones de los mismos) pueden denominarse unidades de codificación (CU) y pueden codificarse usando una o más técnicas de codificación específicas de vídeo, así como técnicas generales de compresión de datos. Se pueden seleccionar diferentes modos y usarse para codificar los bloques de vídeo.

[0004] Con la norma HEVC emergente, las unidades de codificación más grandes (LCU) se pueden dividir en CU cada vez más pequeñas de acuerdo con un sistema de partición en árbol cuaternario. Las CU se pueden predecir basándose en las llamadas unidades de predicción (PU), que pueden tener tamaños de partición correspondientes al tamaño de las CU o más pequeños que el tamaño de las CU, de modo que se pueden usar múltiples PU para predecir una CU dada.

[0005] Se pueden usar diferentes modos para codificar las CU. Por ejemplo, se pueden usar diferentes modos de intracodificación para codificar las CU basándose en los datos predictivos dentro de la misma trama o segmento a fin de aprovechar la redundancia espacial dentro de una trama de vídeo. De forma alternativa, los modos de intercodificación se pueden usar para codificar CU basadas en datos predictivos a partir de otra trama o segmento, a fin de aprovechar la redundancia temporal a través de tramas de una secuencia de vídeo. Después de que la codificación predictiva se realice de acuerdo con un modo seleccionado, se puede realizar entonces la codificación de transformada, tal como las transformadas de coseno discretas (DCT), las transformadas de enteros o similares. Con la HEVC, la codificación de transformada puede producirse con respecto a las unidades de transformada (TU), que también pueden tener diferentes tamaños de transformadas en la norma HEVC. También se puede realizar la cuantización de los coeficientes de transformada, el escaneo de los coeficientes de transformada cuantizados y la codificación por entropía. La información sintáctica se señala con datos de vídeo codificados, por ejemplo, en un encabezado de segmento de vídeo o en un encabezado de bloque de vídeo, con el fin de informar al decodificador sobre cómo decodificar los datos de vídeo. Entre otras cosas, la información sintáctica puede identificar el modo que se usó en la codificación de vídeo de diferentes bloques de vídeo.

[0006] El modo de fusión es un modo de intercodificación específica usado en la compresión de vídeo. Con el modo de fusión, el vector de movimiento de un bloque de vídeo vecino se hereda para un bloque de vídeo actual que se esté codificando. En algunos casos, el modo de fusión causa que un bloque de vídeo actual herede el vector de movimiento de un vecino predefinido y, en otros casos, se puede usar un valor de índice para identificar el vecino específico del que el bloque de vídeo actual hereda su vector de movimiento (por ejemplo, arriba, arriba a la derecha, a la izquierda, abajo a la izquierda o coubicados desde una trama temporalmente adyacente).

60

5 [0007] "Extended merging scheme using motion-hypothesis prediction [Sistema de fusión extendido usando predicción de hipótesis de movimiento]", no. JCTVC-B023, 23 de julio de 2013, LIM J et al. describe la predicción de hipótesis de movimiento en la que se predice una partición actual basándose en una suma ponderada de dos particiones, una de las cuales se ubica usando un parámetro de movimiento de una partición vecina anterior y otra de las cuales se ubica usando un parámetro de movimiento de una partición vecina izquierda.

10 [0008] "Video coding technology proposal by RWTH Aachen University [Propuesta de tecnología de codificación de vídeo de la RWTH Aachen University]", no. JCTVC-A112, 17 de abril de 2010, KAMP S et al. divulga la Derivación de Vectores de Movimiento del lado del Decodificador que produce una señal de predicción basada en una dirección de predicción. Para la predicción unidireccional se usan dos candidatos de movimiento con el coste más bajo para obtener dos señales de predicción que entonces se suman y se dividen por 2. Para la predicción bidireccional de dos candidatos de movimiento con el menor coste se seleccionan las señales sumadas y divididas por 2.

15 **SUMARIO**

20 [0009] Esta divulgación describe un modo de fusión bipredictivo en el que un bloque de vídeo codificado en el modo de fusión bipredictivo puede heredar su información de movimiento de dos bloques vecinos diferentes, en el que los dos bloques vecinos diferentes se codificaron cada uno en un modo unipredictivo. La codificación bipredictiva puede mejorar la capacidad para lograr la compresión o mejorar la calidad del vídeo en un nivel de compresión dado. Sin embargo, en algunos casos, puede que no haya (o que haya pocos) vecinos que se codificaron en un modo bipredictivo, haciendo por tanto que la bipredicción no esté disponible (o sea limitada) con respecto a la codificación del modo de fusión. El modo de fusión bipredictivo descrito puede aumentar el número de candidatos bipredictivos que pueden usarse en el contexto de la codificación del modo de fusión al permitir que se usen dos vecinos unipredictivos independientes para definir la información de movimiento bipredictivo para un bloque de vídeo.

25 [0010] En un ejemplo, esta divulgación describe un procedimiento para decodificar datos de vídeo. El procedimiento comprende recibir uno o más elementos sintácticos para un bloque de vídeo actual, en el que el bloque de vídeo actual se codifica de acuerdo con un modo de fusión bipredictivo y, basándose en uno o más elementos sintácticos, identificar uno de una pluralidad de conjuntos de dos bloques de vídeo vecinos diferentes codificados en los modos unipredictivos o de un bloque vecino del conjunto de bloques vecinos candidatos que se codifique en el modo bipredictivo, en el que los bloques de vídeo vecinos son de un conjunto de bloques vecinos candidatos, comprendiendo el conjunto de bloques vecinos candidatos un vecino superior, un vecino superior derecho, un vecino izquierdo, un vecino inferior izquierdo y un vecino temporal coubicado desde otra trama de vídeo. El procedimiento también comprende, donde el elemento sintáctico identifica uno o una pluralidad de conjuntos de dos bloques de vídeo vecinos diferentes codificados en el modo unipredictivo, usar información de movimiento de los dos bloques de vídeo vecinos diferentes para decodificar un bloque de vídeo actual de acuerdo con el modo de fusión bipredictivo y, donde el elemento sintáctico identifica bloques de vídeo vecinos codificados en el modo bipredictivo, usar los dos vectores de movimiento bipredictivos asociados con el bloque de vídeo vecino para decodificar un bloque de vídeo actual de acuerdo con el modo de fusión bipredictivo.

30 [0011] En otro ejemplo, esta divulgación describe un procedimiento para codificar datos de vídeo. El procedimiento comprende seleccionar un modo de fusión bipredictivo para codificar un bloque de vídeo actual, identificar un primer conjunto de dos bloques de vídeo vecinos diferentes codificados en los modos unipredictivos de un conjunto de bloques vecinos candidatos, comprendiendo el conjunto de bloques vecinos candidatos un vecino superior, un vecino superior derecho, un vecino izquierdo, un vecino inferior izquierdo y un vecino temporal coubicado desde otra trama de vídeo, identificar un segundo conjunto de dos bloques de vídeo vecinos diferentes codificados en los modos unipredictivos a partir del conjunto de bloques vecinos candidatos, y usar dos vectores de movimiento unipredictivos diferentes, uno asociado con cada uno de los dos bloques de vídeo vecinos diferentes de dicho segundo conjunto, como un segundo conjunto candidato de vectores de movimiento para codificar el bloque de vídeo actual de acuerdo con el modo de fusión bipredictivo, identificar al menos uno de los bloques vecinos del conjunto de bloques vecinos candidatos que se codifique en el modo bipredictivo, y usar los dos vectores de movimiento bipredictivo de dicho al menos uno de los bloques vecinos como un tercer conjunto candidato de vectores de movimiento para codificar el bloque de vídeo actual de acuerdo con el modo de fusión bipredictivo, seleccionar uno de los conjuntos de vectores de movimiento candidatos para codificar el conjunto candidato seleccionado de bloque de vídeo actual de vectores de movimiento.

35 [0012] En otro ejemplo, esta divulgación describe un dispositivo para decodificar datos de vídeo, comprendiendo el dispositivo medios para recibir uno o más elementos sintácticos para un bloque de vídeo actual, en el que el bloque de vídeo actual se codifica de acuerdo con un modo de fusión bipredictivo, medios para identificar uno de una pluralidad de conjuntos de dos bloques de vídeo vecinos diferentes codificados en los modos unipredictivos o un bloque vecino del conjunto de bloques vecinos candidatos que se codifique en el modo bipredictivo, basándose en el

uno o más elementos sintácticos, en el que los bloques de vídeo vecinos proceden de un conjunto de bloques vecinos candidatos, comprendiendo el conjunto de bloques vecinos candidatos un vecino superior, un vecino superior derecho, un vecino izquierdo, un vecino inferior izquierdo y un vecino temporal coubicado desde otra trama de vídeo, y medios para usar la información de movimiento de los dos bloques de vídeo vecinos diferentes para decodificar un bloque de vídeo actual de acuerdo con el modo de fusión bipredictivo donde el elemento sintáctico identifica uno o una pluralidad de conjuntos de dos bloques de vídeo vecinos diferentes codificados en el modo unipredictivo o para usar los dos vectores de movimiento bipredictivo asociados con el bloque de vídeo vecino para decodificar un bloque de vídeo actual de acuerdo con el modo de fusión bipredictivo donde el elemento sintáctico identifica bloques de vídeo vecinos codificados en el modo bipredictivo.

[0013] En otro ejemplo, esta divulgación describe un dispositivo para codificar datos de vídeo, comprendiendo el dispositivo medios para seleccionar un modo de fusión bipredictivo para codificar un bloque de vídeo actual, medios para identificar un primer conjunto de dos bloques de vídeo vecinos diferentes codificados en los modos unipredictivos a partir de un conjunto de bloques vecinos candidatos, comprendiendo el conjunto de bloques vecinos candidatos un vecino superior, un vecino superior derecho, un vecino izquierdo, un vecino inferior izquierdo y un vecino temporal coubicado desde otra trama de vídeo, medios para usar dos vectores de movimiento unipredictivos diferentes, uno asociado con cada uno de los dos bloques de vídeo vecinos diferentes de dicho primer conjunto, como un primer conjunto candidato de vectores de movimiento para codificar el bloque de vídeo actual de acuerdo con el modo de fusión bipredictivo, medios para identificar un segundo conjunto de dos bloques de vídeo vecinos diferentes codificados en los modos unipredictivos a partir del conjunto de bloques vecinos candidatos, y usar dos vectores de movimiento unipredictivos diferentes, uno asociado con cada uno de los dos bloques de vídeo vecinos diferentes de dicho segundo conjunto, como un segundo conjunto candidato de vectores de movimiento para codificar el bloque de vídeo actual de acuerdo con el modo de fusión bipredictivo, medios para identificar al menos uno de los bloques vecinos del conjunto de bloques vecinos candidatos que se codifique en el modo bipredictivo, y usar los dos vectores de movimiento bipredictivos de dicho al menos uno de los bloques vecinos como tercer conjunto candidato de vectores de movimiento para codificar el bloque de vídeo actual de acuerdo con el modo de fusión bipredictivo, medios para seleccionar uno de los conjuntos de vectores de movimiento candidatos para codificar el bloque de vídeo actual de acuerdo con el modo de fusión bipredictivo y medios para generar uno o más elementos sintácticos para identificar el conjunto candidato seleccionado de vectores de movimiento.

[0014] Las técnicas descritas en esta divulgación pueden implementarse en hardware, software, firmware o en cualquier combinación de los mismos. Por ejemplo, diversas técnicas pueden implementarse o ejecutarse por uno o más procesadores. Como se usa en el presente documento, un procesador puede referirse a un microprocesador, a un circuito integrado específico de la aplicación (ASIC), a una matriz de puerta programable por campo (FPGA), a un procesador de señales digitales (DSP), o a otro circuito lógico integrado o discreto equivalente. El software puede ejecutarse por uno o más procesadores. El software que comprende instrucciones para ejecutar las técnicas se puede almacenar inicialmente en un medio legible por ordenador y cargarse y ejecutarse mediante un procesador.

[0015] Por consiguiente, esta divulgación también contempla medios de almacenamiento legibles por ordenador que comprenden instrucciones para causar que un procesador realice cualquiera de una variedad de técnicas descritas en esta divulgación. En algunos casos, el medio de almacenamiento legible por ordenador puede formar parte de un producto de almacenamiento de programa informático, que puede venderse a los fabricantes y/o usarse en un dispositivo. El producto de programa informático puede incluir el medio legible por ordenador y, en algunos casos, también puede incluir materiales de embalaje.

[0016] Los detalles de uno o más aspectos de la divulgación se exponen en los dibujos adjuntos y en la descripción siguiente. Otras características, objetivos y ventajas de las técnicas descritas en esta divulgación resultarán evidentes a partir de la descripción y de los dibujos, y a partir de las reivindicaciones.

BREVE DESCRIPCIÓN DE LOS DIBUJOS

[0017]

La FIG. 1 es un diagrama de bloques que ilustra un sistema de codificación y decodificación de vídeo a modo de ejemplo que puede implementar una o más de las técnicas de esta divulgación.

La FIG. 2 es un diagrama conceptual que ilustra la partición de las unidades de codificación (CU) compatibles con las técnicas de esta divulgación.

La FIG. 3 es un diagrama conceptual que ilustra algunas posibles relaciones entre las CU, las unidades de predicción (PU) y las unidades de transformada (TU) compatibles con las técnicas de esta divulgación.

La FIG. 4 es un diagrama de bloques que ilustra un codificador de vídeo que puede implementar las técnicas de esta divulgación.

5 La FIG. 5 es un diagrama de bloques que ilustra una unidad de predicción a modo de ejemplo de un codificador, compatible con uno o más ejemplos de esta divulgación.

La FIG. 6 es un diagrama de bloques que ilustra un decodificador de vídeo que puede implementar las técnicas de esta divulgación.

10 La FIG. 7 es un diagrama conceptual que ilustra la ubicación de diferentes bloques de vídeo vecinos en relación con el bloque de vídeo actual, de modo que el bloque de vídeo actual puede usar información de uno o más bloques de vídeo vecinos diferentes en un modo de fusión bipredictivo, compatible con esta divulgación.

15 Las FIGS. 8 y 9 son diagramas de flujo que ilustran técnicas compatibles con esta divulgación.

DESCRIPCIÓN DETALLADA

20 **[0018]** En la mayoría de los sistemas de codificación de vídeo, la estimación de movimiento y la compensación de movimiento se usan para reducir la redundancia temporal en una secuencia de vídeo, con el fin de lograr la compresión de datos. En este caso, puede generarse un vector de movimiento a fin de identificar un bloque predictivo de datos de vídeo, por ejemplo, a partir de otra trama o segmento de vídeo, que se puede usar para predecir los valores del bloque de vídeo actual que se esté codificando. Los valores del bloque de vídeo predictivo se restan de los valores del bloque de vídeo actual para producir un bloque de datos residuales. El vector de movimiento se comunica desde el codificador al decodificador, junto con los datos residuales. El decodificador puede ubicar el mismo bloque predictivo (basándose en el vector de movimiento) y reconstruir el bloque de vídeo codificado mediante la fusión de los datos residuales con los datos del bloque predictivo. También se pueden usar muchas otras técnicas de compresión, tales como las transformadas y la codificación por entropía, para mejorar aún más la compresión de vídeo.

30 **[0019]** El proceso de estimación de movimiento se realiza usualmente en el codificador. La información de movimiento (tales como los vectores de movimiento, los índices de vectores de movimiento, las direcciones de predicción u otra información) se puede codificar y transmitir desde el codificador al decodificador para que el decodificador pueda identificar el mismo bloque predictivo que se usó para codificar un bloque de vídeo dado. Se pueden usar muchos modos de codificación diferentes para permitir diferentes tipos de predicción temporal entre dos tramas diferentes o tipos diferentes de predicción espacial dentro de una trama dada.

40 **[0020]** Con el llamado modo de fusión, la información de movimiento de un bloque de vídeo vecino se hereda para el bloque de vídeo actual que se esté codificando. En este caso, el propio vector de movimiento no se transmite para un bloque de vídeo codificado en el modo de fusión. Más bien, se puede usar un valor de índice para identificar al vecino del que el bloque de vídeo actual hereda su vector de movimiento (y posiblemente otra información de movimiento). Por ejemplo, la información de movimiento se puede heredar de un vecino superior, un vecino superior derecho, un vecino izquierdo, un vecino inferior izquierdo o un vecino temporal coubicado desde una trama temporalmente adyacente.

45 **[0021]** Con la mayoría de los modos de fusión, si el vecino se codifica en un modo unipredictivo, entonces el bloque de vídeo actual hereda un vector de movimiento. Si el vecino se codifica en un modo bipredictivo, entonces el bloque de vídeo actual hereda dos vectores de movimiento. En dichos ejemplos, un bloque que se codifica en el modo de fusión está limitado por la información de movimiento de sus vecinos. La unipredicción y la bipredicción se denominan algunas veces predicción unidireccional (P) y predicción bidireccional (B), pero el término "direccional" está en general fuera de lugar porque, con las normas modernas de codificación de vídeo, la bipredicción se basa simplemente en dos listas diferentes de datos predictivos y la dirección no es obligatoria. En otras palabras, los datos en las dos listas diferentes para la bipredicción pueden proceder de tramas anteriores o posteriores, y no es necesario que sean bidireccionales a partir de las tramas anteriores y posteriores, respectivamente. Por esta razón, esta divulgación usa los términos unipredicción y bipredicción en lugar de los términos predicción unidireccional y predicción bidireccional.

55 **[0022]** La codificación bipredictiva puede mejorar la capacidad para lograr la compresión o mejorar la calidad del vídeo en un nivel de compresión dado. Sin embargo, en algunos casos, puede que no haya (o que haya pocos) vecinos que se codificaron en un modo bipredictivo, lo que hace por tanto que la bipredicción no esté disponible (o limitada) en la codificación del modo de fusión. Por ejemplo, con el modo de fusión convencional, si no hay un modo bipredictivo en ninguno de los bloques vecinos, el bloque actual puede perder la oportunidad de aprovechar los beneficios que puedan surgir de la bipredicción.

5 [0023] Esta divulgación describe un modo de fusión bipredictivo como una extensión o adición a las técnicas del modo de fusión. Más específicamente, esta divulgación describe un modo de fusión bipredictivo que puede heredar información de movimiento de dos bloques vecinos diferentes, en el que los dos bloques vecinos diferentes se codificaron cada uno en un modo unipredictivo. El modo de fusión bipredictivo descrito puede aumentar el número de candidatos bipredictivos que pueden usarse en el contexto de la codificación del modo de fusión.

10 [0024] La FIG. 1 es un diagrama de bloques que ilustra un sistema de codificación y decodificación de vídeo 10 a modo de ejemplo 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 a través 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 pueden comprender auriculares de dispositivos de comunicación inalámbrica, tales como los llamados radioteléfonos móviles o satelitales. Sin embargo, las técnicas de esta divulgación, que se aplican en general a la codificación y a la decodificación de bloques de vídeo en el modo de fusión bipredictivo, se pueden aplicar a dispositivos no inalámbricos que incluyan capacidades de codificación y/o decodificación de vídeo. El dispositivo de origen 12 y el dispositivo de destino 16 son meramente ejemplos de dispositivos de codificación que pueden soportar las técnicas descritas en el presente documento.

20 [0025] En el ejemplo de la FIG. 1, el dispositivo de origen 12 incluye una fuente de vídeo 20, un codificador de vídeo 22, un modulador/demodulador (módem) 23 y un transmisor 24. El dispositivo de destino 16 puede incluir un receptor 26, un módem 27, un decodificador de vídeo 28 y un dispositivo de visualización 30. De acuerdo con esta divulgación, el codificador de vídeo 22 del dispositivo de origen 12 puede configurarse para codificar uno o más bloques de vídeo de acuerdo con un modo de fusión bipredictivo. Con el modo de fusión bipredictivo, un bloque de vídeo hereda su información de movimiento de dos bloques vecinos diferentes, en el que los dos bloques vecinos diferentes se codificaron cada uno en un modo unipredictivo. Se pueden generar elementos sintácticos en el codificador de vídeo 22 con el fin de identificar los dos bloques de vídeo vecinos diferentes codificados en los modos unipredictivos. De esta forma, un decodificador de vídeo puede reconstruir un bloque de vídeo bipredictivo basándose en la información de movimiento de los dos bloques de vídeo vecinos diferentes identificados por los elementos sintácticos.

30 [0026] Más específicamente, el codificador de vídeo 22 puede seleccionar un modo de fusión bipredictivo para codificar un bloque de vídeo actual, e identificar dos bloques de vídeo vecinos diferentes codificados en los modos unipredictivos. El codificador de vídeo 22 puede usar la información de movimiento de los dos bloques de vídeo vecinos diferentes para codificar el bloque de vídeo actual de acuerdo con el modo de fusión bipredictivo, y generar uno o más elementos sintácticos para identificar los dos bloques de vídeo vecinos diferentes para un decodificador de vídeo.

40 [0027] La fuente de vídeo 20 puede comprender un dispositivo de captura de vídeo, tal como una cámara de vídeo, un archivo de vídeo que contenga vídeo capturado previamente, un vídeo procedente de un proveedor de contenido de vídeo o de otra fuente de vídeo. Como otra alternativa, la fuente de vídeo 20 puede generar datos basados en gráficos por ordenador como el vídeo de origen, 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 videocámara, el dispositivo de origen 12 y el dispositivo de destino 16 pueden formar los llamados teléfonos con cámara o videoteléfonos. En cada caso, el vídeo capturado, precapturado o generado por ordenador puede codificarse mediante el codificador de vídeo 22.

45 [0028] Una vez que los datos de vídeo se codifican por el codificador de vídeo 22, la información de vídeo codificada puede modularse entonces por el módem 23 de acuerdo con una norma de comunicación, por ejemplo, tal como el acceso múltiple por división de código (CDMA), el multiplexado por división de frecuencia ortogonal (OFDM) o cualquier otra norma o técnica de comunicación. Los datos codificados y modulados se pueden transmitir entonces al dispositivo de destino 16 a través del transmisor 24. El módem 23 puede incluir diversos mezcladores, filtros, amplificadores u otros componentes diseñados para la modulación de señales. El transmisor 24 puede incluir circuitos diseñados para transmitir datos, incluidos amplificadores, filtros y una o más antenas. El receptor 26 del dispositivo de destino 16 recibe información a través del canal 15, y el módem 27 demodula la información.

55 [0029] El proceso de decodificación de vídeo realizado por el decodificador de vídeo 28 puede incluir técnicas recíprocas a las técnicas de codificación realizadas por el codificador de vídeo 22. En particular, el decodificador de vídeo 28 puede recibir uno o más elementos sintácticos para un bloque de vídeo actual, en el que el bloque de vídeo actual se codifica de acuerdo con un modo de fusión bipredictivo y, basándose en uno o más elementos sintácticos, identificar dos bloques de vídeo vecinos diferentes codificados en los modos unipredictivos. El decodificador de vídeo puede usar la información de movimiento de los dos bloques de vídeo vecinos diferentes para decodificar un bloque de vídeo actual de acuerdo con el modo de fusión bipredictivo.

5 [0030] El canal de comunicación 15 puede comprender cualquier medio de comunicación inalámbrica o cableada, tal como un espectro de radiofrecuencia (RF) o una o más líneas de transmisión física, o cualquier combinación de medios inalámbricos y cableados. El canal de comunicación 15 puede formar parte de una red basada en paquetes, tal como una red de área local, una red de área extensa o una red global tal como Internet. El canal de comunicación 15 representa en general cualquier medio de comunicación adecuado o un conjunto de diferentes medios de comunicación, para transmitir datos de vídeo desde el dispositivo de origen 12 hasta el dispositivo de destino 16. De nuevo, la FIG. 1 es meramente a modo de ejemplo y las técnicas de esta divulgación pueden aplicarse a configuraciones de codificación de vídeo (por ejemplo, codificación de vídeo o decodificación de vídeo) que no incluyan necesariamente ninguna comunicación de datos entre los dispositivos de codificación y decodificación. En otros ejemplos, los datos podría recuperarse de una memoria local, transmitirse a través de una red o similar.

15 [0031] En algunos casos, el codificador de vídeo 22 y el decodificador de vídeo 28 pueden funcionar sustancialmente de acuerdo con una norma de compresión de vídeo tal como la norma HEVC emergente. Sin embargo, las técnicas de esta divulgación también se pueden aplicar en el contexto de una variedad de otras normas de codificación de vídeo, incluyendo algunas normas antiguas, o normas nuevas o emergentes. Aunque no se muestra en la FIG. 1, en algunos casos, tanto el codificador de vídeo 22 como el decodificador de vídeo 28 pueden estar integrados cada uno en un codificador y decodificador de audio, y pueden incluir unidades MUX-DEMUX apropiadas, u otro tipo de hardware y software, para manejar la codificación tanto de audio como de vídeo en un flujo de datos común o en flujos de datos independientes. Si procede, las unidades MUX-DEMUX pueden ajustarse al protocolo de multiplexador H.223 de la ITU o a otros protocolos tales como el protocolo de datagramas de usuario (UDP).

25 [0032] El codificador de vídeo 22 y el decodificador de vídeo 28 pueden 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, software, hardware, firmware o combinaciones de los mismos. Cada uno del codificador de vídeo 22 y del decodificador de vídeo 28 pueden incluirse en uno o más codificadores o decodificadores, donde cualquiera de los mismos puede estar integrado como parte de un codificador/decodificador combinado (CODEC) en un respectivo dispositivo móvil, dispositivo de abonado, dispositivo de radiodifusión, servidor o similar. En esta divulgación, el término codificador se refiere a un codificador, un decodificador, o CODEC, y los términos codificador, decodificador y CODEC se refieren todos a máquinas específicas diseñadas para la codificación (codificación y/o decodificación) de datos de vídeo compatibles con esta divulgación.

35 [0033] En algunos casos, los dispositivos 12, 16 pueden funcionar de una manera sustancialmente simétrica. Por ejemplo, cada uno de los dispositivos 12, 16 puede incluir componentes de codificación y decodificación de vídeo. Por lo tanto, el sistema 10 puede prestar soporte a 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.

40 [0034] Durante el proceso de codificación, el codificador de vídeo 22 puede ejecutar varias técnicas u operaciones de codificación. En general, el codificador de vídeo 22 funciona en bloques de datos de vídeo compatibles con la norma HEVC. Coherente con la norma HEVC, los bloques de vídeo se denominan unidades de codificación (CU) y existen muchas CU dentro de tramas de vídeo individuales (u otras unidades de vídeo definidas independientemente, tales como segmentos). Las tramas, los segmentos, las porciones de tramas, los grupos de imágenes u otras estructuras de datos pueden definirse como unidades de información de vídeo que incluyan una pluralidad de CU. Las CU pueden tener distintos tamaños compatibles con la norma HEVC, y el flujo de bits puede definir unidades de codificación (LCU) más grandes como el mayor tamaño de CU. El modo de fusión bipredictivo se puede usar para codificar LCU, CU o posiblemente otros tipos de bloques de vídeo. Con la norma HEVC, las LCU se pueden dividir en CU cada vez más pequeñas de acuerdo con un sistema de partición en árbol cuaternario, y las diferentes CU que se definen en el sistema se pueden particionar además en las llamadas unidades de predicción (PU). Las LCU, las CU y las PU son todas bloques de vídeo en el sentido de esta divulgación.

50 [0035] El codificador de vídeo 22 puede realizar una codificación predictiva en la que un bloque de vídeo que se codifique (por ejemplo, una PU de una CU dentro de una LCU) se compare con uno o más candidatos predictivos para identificar un bloque predictivo. Este proceso de codificación predictiva puede ser intra (en cuyo caso los datos predictivos se generan basándose en los intradatos vecinos dentro de la misma trama o segmento de vídeo) o inter (en cuyo caso los datos predictivos se generan basándose en los datos de vídeo en tramas o segmentos anteriores o posteriores). Se puede prestar soporte a muchos modos de codificación diferentes, y el codificador de vídeo 22 puede seleccionar un modo de codificación de vídeo deseable. De acuerdo con esta divulgación, al menos algunos bloques de vídeo pueden codificarse usando el modo de fusión bipredictivo descrito en el presente documento.

60 [0036] Después de generar el bloque predictivo, las diferencias entre el bloque de vídeo actual que se esté codificando y el bloque predictivo se codifican como un bloque residual, y la sintaxis de predicción (tal como un vector

de movimiento en el caso de la intercodificación, o un modo predictivo en el caso de la intracodificación) se usa para identificar el bloque predictivo. Además, con el modo de fusión bipredictivo descrito en el presente documento, la sintaxis de predicción (por ejemplo, los elementos sintácticos) puede identificar dos bloques de vídeo vecinos diferentes a un decodificador de vídeo. Por consiguiente, el decodificador puede identificar dos bloques de vídeo vecinos diferentes codificados en los modos unipredictivos, basándose en los elementos sintácticos, y usar la información de movimiento de los dos bloques de vídeo vecinos diferentes para decodificar un bloque de vídeo actual de acuerdo con el modo de fusión bipredictivo.

[0037] El valor residual puede transformarse y cuantizarse. Las técnicas de transformada pueden comprender un proceso DCT o un proceso conceptualmente similar, transformadas de enteros, transformadas de ondículas u otros tipos de transformadas. En un proceso DCT, como ejemplo, el proceso de transformada convierte un conjunto de valores de píxeles (por ejemplo, valores de píxeles residuales) en coeficientes de transformada, que pueden representar la energía de los valores de píxel en el dominio de frecuencia. La nueva norma HEVC admite transformaciones de acuerdo con las unidades de transformada (TU), que pueden ser diferentes para diferentes CU. El tamaño de las TU se basa típicamente en el tamaño de las PU dentro de una CU definida dada para una LCU particionada, aunque puede que no sea siempre así. Las TU son típicamente del mismo tamaño o de un tamaño más pequeño que las PU.

[0038] La cuantización puede aplicarse a los coeficientes de transformada, y en general implica un proceso que limita el número de bits asociados con cualquier coeficiente de transformada dado. Más específicamente, la cuantización puede aplicarse de acuerdo con un parámetro de cuantización (QP) definido al nivel de LCU. Por consiguiente, el mismo nivel de cuantización puede aplicarse a todos los coeficientes de transformada en las TU asociadas con diferentes PU de CU dentro de una LCU. Sin embargo, en lugar de señalar el propio QP, un cambio (es decir, un delta) en el QP puede señalarse con la LCU para indicar el cambio en el QP en relación con el de una LCU anterior.

[0039] Después de la transformada y la cuantización, la codificación por entropía se puede realizar en los bloques de vídeo residuales cuantizados y transformados. Los elementos sintácticos también pueden incluirse en el flujo de bits codificado por entropía. En general, la codificación por entropía comprende uno o más procesos que comprimen colectivamente una secuencia de coeficientes de transformada cuantizados y/u otra información sintáctica. Las técnicas de escaneo se pueden realizar en los coeficientes de transformada cuantizados para definir uno o más vectores unidimensionales de coeficientes a partir de bloques de vídeo bidimensionales. Los coeficientes escaneados se codifican por entropía entonces junto con cualquier información sintáctica, por ejemplo, a través de la codificación de longitud variable adaptativa al contenido (CAVLC), de la codificación aritmética binaria adaptativa al contexto (CABAC) o de cualquier otro proceso de codificación por entropía.

[0040] Como parte del proceso de codificación, los bloques de vídeo codificados pueden decodificarse con el fin de generar los datos de vídeo que se usan para la posterior codificación basada en la predicción de los bloques de vídeo posteriores. A menudo, esto se denomina bucle de decodificación del proceso de codificación, y en general imita la decodificación que se realiza por un dispositivo decodificador. En el bucle de decodificación de un codificador o decodificador, se pueden usar técnicas de filtrado para mejorar la calidad del vídeo y, por ejemplo, suavizar los límites de píxeles y posiblemente eliminar los fallos del vídeo decodificado. Este filtrado se puede realizar en bucle o posterior al bucle. Con el filtrado en bucle, el filtrado de los datos de vídeo reconstruidos se produce en el bucle de codificación, lo que significa que los datos filtrados se almacenan por un codificador o un decodificador para su uso posterior en la predicción de los datos de imagen posteriores. En contraste, con el filtrado posterior al bucle, el filtrado de los datos de vídeo reconstruidos se produce fuera del bucle de codificación, lo que significa que las versiones sin filtrar de los datos se almacenan por un codificador o un decodificador para su uso posterior en la predicción de los datos de imagen posteriores. El filtrado de bucle sigue a menudo un proceso de filtrado de desbloqueo independiente, que aplica típicamente el filtrado a los píxeles que están en o cerca de los límites de bloques de vídeo adyacentes con el fin de eliminar los fallos de bloqueo que se manifiesten en los límites del bloque de vídeo.

[0041] En relación con las normas de codificación anteriores, la nueva norma HEVC introduce nuevos términos y tamaños de bloque para los bloques de vídeo. En particular, la HEVC se refiere a unidades de codificación (CU), que se pueden particionar de acuerdo con un sistema de partición en árbol cuaternario. Una "LCU" se refiere a la unidad de codificación de mayor tamaño (por ejemplo, la "unidad de codificación más grande") soportada en una situación dada. El tamaño de la LCU puede señalarse como parte del flujo de bits, por ejemplo, como sintaxis de nivel de secuencia. La LCU se puede particionar en CU más pequeñas. Además, las CU se pueden particionar en unidades de predicción (PU) con fines de predicción. Las PU pueden tener forma cuadrada o rectangular. Las transformadas no se fijan en la nueva norma HEVC, pero se definen de acuerdo con los tamaños de unidad de transformada (TU), que pueden ser del mismo tamaño que una CU determinada, o posiblemente más pequeños. Los datos residuales para

una CU dada pueden comunicarse en las TU. Los elementos sintácticos se pueden definir en el nivel de LCU, el nivel de CU, el nivel de PU y el nivel de TU.

5 **[0042]** Para ilustrar bloques de vídeo de acuerdo con el estándar HEVC, la FIG. 2 muestra conceptualmente una LCU de profundidad 64 por 64, que se particiona entonces en CU más pequeñas de acuerdo con un sistema de partición en árbol cuaternario. Los elementos llamados "indicadores divididos" pueden incluirse como una sintaxis de nivel de CU para indicar si alguna CU determinada está subdividida en cuatro CU más. En la FIG. 2, CU₀ puede comprender la LCU, CU₁ a CU₄ puede comprender subCU de la LCU. Los elementos sintácticos de modo de fusión bipredictivos, como se describen en esta divulgación, pueden definirse a nivel de CU (o posiblemente al nivel de LCU si la LCU no está dividida en CU más pequeñas). El modo de fusión bipredictivo también podría recibir soporte para las PU de CU, en algunos ejemplos.

15 **[0043]** La FIG. 3 ilustra además una posible relación entre CU, PU y TU, que puede ser compatible con la norma HEVC emergente o con otras normas. Sin embargo, también son posibles otras relaciones, y la FIG. 3 se muestra meramente como un posible ejemplo. En este caso, cualquier CU dada de una LCU puede particionarse en PU, tales como las ilustradas en la FIG. 3. El tipo de PU para una CU dada se puede señalar como sintaxis de nivel de CU. Como se muestra en la FIG. 3, las PU de tipo simétrico y las PU de tipo asimétrico se pueden definir para una CU determinada. Además, se pueden definir dos estructuras de TU diferentes para cada una de las cuatro PU de tipo simétrico y las PU de tipo asimétrico. Por tanto, un elemento sintáctico de un bit (indicador de tamaño de TU) se puede usar para señalar el tamaño de TU, que también puede depender del tipo de PU (simétrico o asimétrico). Los patrones de bloques codificados (CBP) se pueden definir para una LCU para indicar si alguna CU dada incluye coeficientes de transformada distintos de cero (por ejemplo, si hay TU presentes).

25 **[0044]** En otros ejemplos, las TU pueden definirse de manera diferente a la mostrada en la FIG. 3. En algunos ejemplos, las muestras residuales correspondientes a una CU pueden subdividirse en unidades más pequeñas mediante una estructura en árbol cuaternario conocida como «árbol cuaternario residual» (RQT). Los nodos hoja del RQT pueden denominarse unidades de transformada (TU). En algunos casos, las TU pueden definirse para una CU de acuerdo con una estructura de árbol cuaternario, pero las TU pueden no depender necesariamente de las PU definidas para cualquier CU determinada. Las PU usadas para la predicción pueden definirse por separado de las TU de cualquier UC. Son posibles varios tipos diferentes de sistemas de partición para CU, TU y PU.

35 **[0045]** La FIG. 4 es un diagrama de bloques que ilustra un codificador de vídeo 50 compatible con esta divulgación. El codificador de vídeo 50 puede corresponder al codificador de vídeo 22 del dispositivo 20 o a un codificador de vídeo de un dispositivo diferente. Como se muestra en la FIG. 4, el codificador de vídeo 50 incluye una unidad de codificación de predicción 32, una unidad de partición en árbol cuaternario 31, sumadores 48 y 51 y una memoria 34. El codificador de vídeo 50 también incluye una unidad de transformada 38 y una unidad de cuantización 40, así como una unidad de cuantización inversa 42 y una unidad de transformada inversa 44. El codificador de vídeo 50 también incluye una unidad de codificación por entropía 46 y una unidad de filtro 47, que puede incluir filtros de desbloqueo y filtros de bucle posterior y/o en bucle. Los datos de vídeo codificados y la información sintáctica que define la manera de la codificación pueden comunicarse a la unidad de codificación por entropía 46, que realiza la codificación por entropía en el flujo de bits.

45 **[0046]** Como se muestra en la FIG. 4, la unidad de codificación de predicción 32 puede prestar soporte a una pluralidad de diferentes modos de codificación 35 en la codificación de bloques de vídeo. Los modos 35 pueden incluir modos de intercodificación que definan datos predictivos de diferentes tramas de vídeo (o segmentos). Los modos de intercodificación pueden ser bipredictivos, lo que significa que se usan dos listas diferentes (por ejemplo, Lista 0 y Lista 1) de datos predictivos (y típicamente dos vectores de movimiento diferentes) para identificar los datos predictivos. Los modos de intercodificación pueden ser de forma alternativa unipredictivos, lo que significa que se usa una lista (por ejemplo, Lista 0) de datos predictivos (y típicamente un vector de movimiento) para identificar los datos predictivos. Las interpolaciones, las compensaciones u otras técnicas se pueden realizar junto con la generación de datos predictivos. También se puede prestar soporte a los llamados modos SKIP y DIRECT, que heredan la información de movimiento asociada con un bloque codificado de otra trama (o segmento). Los bloques del modo SKIP no incluyen información residual, mientras que los bloques del modo DIRECT incluyen información residual.

55 **[0047]** Además, los modos 35 pueden incluir modos de intercodificación, que definen datos predictivos basados en datos dentro de la misma trama de vídeo (o segmento) que la que se esté codificando. Los modos de intracodificación pueden incluir modos direccionales que definan datos predictivos basados en datos en una dirección particular dentro de la misma trama, así como modos DC y/o planares que definan datos predictivos basados en el promedio o en el promedio ponderado de los datos vecinos. La unidad de codificación de predicción 32 puede seleccionar el modo para un bloque dado basándose en algunos criterios, tal como el que se basa en un análisis de distorsión de la velocidad o en algunas características del bloque, tal como el tamaño del bloque, la textura u otras características.

60

5 [0048] De acuerdo con esta divulgación, la unidad de codificación de predicción 32 presta soporte a un modo de fusión bipredictivo 35X. Con el modo de fusión bipredictivo 35X, un bloque de vídeo que se codifique hereda la información de movimiento de dos bloques vecinos diferentes, en la que los dos bloques vecinos diferentes se codificaron cada uno en un modo unipredictivo. Por tanto, el bloque de vídeo se codifica con dos vectores de movimiento diferentes que proceden de dos bloques de vídeo vecinos diferentes. En este caso, la unidad de codificación de predicción 32 emite una indicación de que el modo de fusión bipredictivo se usó para un bloque dado, y emite elementos sintácticos que identifican los dos vecinos unipredictivos diferentes que definen colectivamente la información de movimiento para el bloque bipredictivo actual. Los bloques predictivos asociados con el modo de fusión bipredictivo se pueden combinar en un bloque bipredictivo (posiblemente usando factores de ponderación), y el bloque bipredictivo se puede restar del bloque que se esté codificando (a través del sumador 48) para definir los datos residuales asociados con el bloque codificado en el modo de fusión bipredictivo.

15 [0049] La información de movimiento puede comprender dos vectores de movimiento unipredictivos diferentes asociados con los dos bloques de vídeo vecinos diferentes. Estos dos vectores de movimiento unipredictivos diferentes se pueden usar como los dos vectores de movimiento bipredictivos del bloque de vídeo actual. La información de movimiento puede comprender además dos valores de índice de referencia asociados con los dos vectores de movimiento unipredictivos diferentes, en la que los valores de índice de referencia identifican una o más listas de datos predictivos asociados con los dos vectores de movimiento unipredictivos diferentes. De nuevo, los datos residuales pueden generarse como la diferencia entre el bloque que se esté codificando y los datos predictivos definidos por los dos vectores de movimiento unipredictivos diferentes que definan colectivamente el bloque de fusión bipredictivo usado en la predicción.

25 [0050] En el caso de HEVC, el bloque de vídeo actual que se esté codificando puede comprender una llamada CU definida en relación con una LCU de acuerdo con un sistema de partición en árbol cuaternario. En este caso, la unidad de partición en árbol cuaternario 31 puede generar datos sintácticos de LCU que definan el sistema de partición en árbol cuaternario, y la unidad de codificación de predicción 32 puede generar información de modo para la CU que defina el modo de fusión bipredictivo, en el que uno o más elementos sintácticos (que identifique los dos vecinos unipredictivos) se incluyen en la información de modo para la CU.

30 [0051] El modo de fusión bipredictivo descrito puede aumentar el número de candidatos bipredictivos que pueden usarse en el contexto de la codificación del modo de fusión. Por ejemplo, si ninguno de los vecinos se codifica en el modo bipredictivo, el modo de fusión bipredictivo descrito puede permitir que se aproveche la bipredicción combinando la información de movimiento de dos vecinos para predecir el bloque de vídeo actual. Además, incluso si uno o más vecinos se codifican en el modo bipredictivo, la combinación de dos vecinos unidireccionales de acuerdo con el modo de fusión bipredictivo descrito puede proporcionar ganancias de codificación en algunas situaciones.

40 [0052] En general, durante el proceso de codificación, el codificador de vídeo 50 recibe datos de vídeo de entrada. La unidad de codificación de predicción 32 realiza técnicas de codificación predictiva en bloques de vídeo (por ejemplo, CU y PU). La unidad de partición en árbol cuaternario 31 puede romper una LCU en CU y PU más pequeñas de acuerdo con la partición de HEVC explicada anteriormente con referencia a las FIGS. 2 y 3. Para la intercodificación, la unidad de codificación de predicción 32 compara CU o PU con diversos candidatos predictivos en una o más tramas o segmentos de referencia de vídeo (por ejemplo, una o más "listas" de datos de referencia) para definir un bloque predictivo. Para la intracodificación, la unidad de codificación de predicción 32 genera un bloque predictivo basado en datos vecinos dentro de la misma trama o segmento de vídeo. La unidad de codificación de predicción 32 genera el bloque de predicción y el sumador 48 resta el bloque de predicción de la CU o la PU que se estén codificando para generar un bloque residual. De nuevo, al menos algunos bloques de vídeo pueden codificarse usando el modo de fusión bipredictivo descrito en el presente documento.

50 [0053] La FIG. 5 ilustra un ejemplo de la unidad de codificación de predicción 32 del codificador de vídeo 50 con mayor detalle. La unidad de codificación de predicción 32 puede incluir una unidad de selección de modo 75 que seleccione el modo deseado de los modos 35, que incluyen el modo de fusión bipredictivo 35X como una posibilidad. Para la intercodificación, la unidad de codificación de predicción 32 puede comprender la unidad de estimación de movimiento (ME) 76 y la unidad de compensación de movimiento (MC) 77 que identifican uno o más vectores de movimiento que apuntan a datos predictivos, y genera el bloque de predicción basado en el vector de movimiento. Típicamente, la estimación de movimiento se considera el proceso de generar uno o más vectores de movimiento, que estiman el movimiento. Por ejemplo, el vector de movimiento puede indicar el desplazamiento de un bloque predictivo dentro de una trama predictiva en relación con el bloque actual que se esté codificando dentro de la trama actual. En el caso del modo de fusión bipredictivo 35X, dos vectores de movimiento unidireccionales de dos vecinos se combinan para crear una predicción bidireccional.

5 **[0054]** La compensación por movimiento se considera típicamente el proceso de extraer o generar el bloque (o bloques) predictivo basándose en el vector de movimiento determinado por la estimación de movimiento. En algunos casos, la compensación de movimiento para la intercodificación puede incluir interpolaciones a la resolución de subpíxeles, lo que permite que el proceso de estimación de movimiento estime el movimiento de los bloques de vídeo a dicha resolución de subpíxeles. También se pueden usar combinaciones ponderadas de dos bloques (en el caso de la bipredicción).

10 **[0055]** Para la intracodificación, la unidad de codificación de predicción 32 puede comprender la unidad de intrapredicción 78. En este caso, los datos predictivos se pueden generar basándose en los datos dentro del bloque de vídeo actual (por ejemplo, adyacente al bloque de vídeo que se esté codificando). De nuevo, los modos de intracodificación pueden incluir modos direccionales que definan datos predictivos basados en datos en una dirección particular dentro de la misma trama, así como modos de DC y/o planares que definan datos predictivos basados en el promedio o en el promedio ponderado de los datos vecinos.

15 **[0056]** La unidad de distorsión de frecuencia (R-D) 79 puede comparar los resultados de codificación de bloques de vídeo (por ejemplo, CU o PU) en diferentes modos. Además, la unidad R-D 79 puede permitir otros tipos de ajustes de parámetros, tales como ajustes a interpolaciones, compensaciones, parámetros de cuantización u otros factores que puedan afectar la velocidad de codificación. La unidad de selección de modo 75 puede analizar los resultados de codificación en términos de velocidad de codificación (es decir, bits de codificación requeridos para el bloque) y distorsión (por ejemplo, que represente la calidad del vídeo del bloque codificado en relación con el bloque original) para realizar selecciones de modo para bloques de vídeo. De esta manera, la unidad R-D 79 proporciona un análisis de los resultados de diferentes modos para permitir que la unidad de selección de modo 75 seleccione el modo deseado para diferentes bloques de vídeo. Compatible con esta divulgación, el modo de fusión bipredictivo 35X puede seleccionarse cuando la unidad R-D 79 lo identifique como el modo deseado para un bloque de vídeo dado, por ejemplo, debido a las ganancias de codificación o a la eficiencia de codificación.

30 **[0057]** Con referencia de nuevo a la FIG. 4, después de que la unidad de codificación de predicción 32 emita el bloque de predicción, y después de que el sumador 48 reste el bloque de predicción del bloque de vídeo que se esté codificando con el fin de generar un bloque residual de valores de píxeles residuales, la unidad de transformada 38 aplica una transformada al bloque residual. La transformada puede comprender una transformada de coseno discreta (DCT) o una transformada conceptualmente similar a la definida por la norma H.264 de la ITU o la norma HEVC. Las llamadas estructuras "mariposa" pueden definirse para realizar las transformadas, o también podría usarse la multiplicación basada en matrices. En algunos ejemplos, de acuerdo con la norma HEVC, el tamaño de la transformada puede variar para diferentes CU, por ejemplo, dependiendo del nivel de partición que se produzca con respecto a una LCU determinada. Las unidades de transformada (TU) se pueden definir para establecer el tamaño de transformada aplicado por la unidad de transformada 38. También podrían usarse transformadas de ondículas, transformadas de números enteros, transformadas de subbandas u otros tipos de transformadas. En cualquier caso, la unidad de transformada 52 aplica la transformada al bloque residual, produciendo un bloque de coeficientes de transformada residuales. La transformada puede convertir en general la información residual de un dominio de píxel en un dominio de frecuencia.

45 **[0058]** La unidad de cuantización 40 cuantiza entonces los coeficientes de transformada residuales para reducir más la tasa de bits. La unidad de cuantización 40, por ejemplo, puede limitar el número de bits usados para codificar cada uno de los coeficientes. En particular, la unidad de cuantización 40 puede aplicar el QP delta definido para la LCU a fin de definir el nivel de cuantización que se vaya a aplicar (por ejemplo, combinando el QP delta con el QP de la LCU anterior o algún otro QP conocido). Después de realizar la cuantización en muestras residuales, la unidad de codificación por entropía 46 puede escanear y codificar por entropía los datos.

50 **[0059]** CAVLC es un tipo de técnica de codificación por entropía soportada por la norma H.264 de la ITU y la nueva norma HEVC, que puede aplicarse de forma vectorializada por la unidad de codificación por entropía 46. CAVLC usa tablas de codificación de longitud variable (VLC) de una manera que comprime efectivamente las "ejecuciones" serializadas de coeficientes y/o elementos sintácticos. CABAC es otro tipo de técnica de codificación por entropía soportada por la norma H.264 de la ITU o la norma HEVC, que puede aplicarse de forma vectorializada por la unidad de codificación por entropía 46. CABAC puede implicar varias etapas, incluida la binarización, la selección del modelo de contexto y la codificación aritmética binaria. En este caso, la unidad de codificación por entropía 46 codifica los coeficientes y los elementos sintácticos de acuerdo con CABAC. También existen muchos otros tipos de técnicas de codificación por entropía, y en el futuro probablemente aparecerán nuevas técnicas de codificación por entropía. Esta divulgación no se limita a ninguna técnica específica de codificación por entropía.

60 **[0060]** Tras la codificación por entropía realizada por la unidad de codificación por entropía 46, el vídeo codificado puede transmitirse a otro dispositivo o archivarse para su transmisión o recuperación posterior. El vídeo codificado

5 puede comprender los vectores codificados por entropía y diversas informaciones de sintaxis (incluida la información sintáctica que define dos vecinos en el caso del modo de fusión bipredictivo). El decodificador puede usar dicha información para configurar correctamente el proceso de decodificación. La unidad de cuantización inversa 42 y la unidad de transformada inversa 44 aplican la cuantización inversa y la transformada inversa, respectivamente, para
 10 reconstruir el bloque residual en el dominio del píxel. El sumador 51 suma el bloque residual reconstruido al bloque de predicción producido por la unidad de codificación de predicción 32 para producir un bloque de vídeo reconstruido para su almacenamiento en la memoria 34. Sin embargo, antes de dicho almacenamiento, la unidad de filtro 47 puede aplicar un filtrado al bloque de vídeo para mejorar la calidad de vídeo. El filtrado aplicado por la unidad de filtro 47 puede reducir los fallos y suavizar los límites de los píxeles. Además, el filtrado puede mejorar la compresión generando bloques de vídeo predictivos que comprendan coincidencias aproximadas con los bloques de vídeo que se estén codificando.

15 **[0061]** De acuerdo con esta divulgación, se presta soporte al modo de fusión bipredictivo 35X que hereda la información de movimiento de dos bloques vecinos diferentes, en el que los dos bloques vecinos diferentes se codificaron cada uno en un modo unipredictivo. El modo de fusión bipredictivo 35X descrito puede aumentar el número de candidatos bipredictivos que pueden usarse en el contexto de la codificación del modo de fusión. Por consiguiente, la unidad R-D 79 (FIG. 5) puede identificar el modo de fusión bipredictivo 35X como el modo de codificación más deseable debido a las ganancias de codificación logradas por este modo en relación con otros modos. En dichos casos, la unidad de selección de modo 75 puede seleccionar el modo de fusión bipredictivo 35X para codificar bloques de vídeo.
 20

25 **[0062]** La FIG. 6 es un diagrama de bloques que ilustra un ejemplo de un decodificador de vídeo 60, que puede decodificar una secuencia de vídeo que se codifique de la manera descrita en el presente documento. Las técnicas de esta divulgación pueden realizarse por el decodificador de vídeo 60 en algunos ejemplos. En particular, el decodificador de vídeo 60 recibe uno o más elementos sintácticos para un bloque de vídeo actual, en el que el bloque de vídeo actual se codifica de acuerdo con un modo de fusión bipredictivo y, basándose en uno o más elementos sintácticos, identifica dos bloques de vídeo vecinos diferentes codificados en los modos unipredictivos. El decodificador de vídeo 60 usa entonces la información de movimiento de los dos bloques de vídeo vecinos diferentes para decodificar un bloque de vídeo actual de acuerdo con el modo de fusión bipredictivo.
 30

35 **[0063]** Una secuencia de vídeo recibida en el decodificador de vídeo 60 puede comprender un conjunto codificado de tramas de imágenes, un conjunto de segmentos de trama, un grupo codificado comúnmente de imágenes (GOP) o una gran variedad de unidades de información de vídeo que incluyan LCU codificadas (u otros bloques de vídeo) e información sintáctica para definir cómo decodificar dichas LCU. El proceso de decodificación de las LCU puede incluir la decodificación de una indicación del modo de codificación, que puede ser el modo de fusión bipredictivo descrito en el presente documento.

40 **[0064]** El decodificador de vídeo 60 incluye una unidad de decodificación por entropía 52, que realiza la función de decodificación recíproca de la codificación realizada por la unidad de codificación por entropía 46 de la FIG. 2. En particular, la unidad de decodificación por entropía 52 puede realizar la decodificación CAVLC o CABAC, o cualquier otro tipo de decodificación de entropía usada por el codificador de vídeo 50. El decodificador de vídeo 60 también incluye una unidad de decodificación de predicción 54, una unidad de cuantización inversa 56, una unidad de transformada inversa 58, una memoria 62 y un sumador 64. En particular, como el codificador de vídeo 50, el decodificador de vídeo 60 incluye una unidad de decodificación de predicción 54 y una unidad de filtro 57. La unidad de decodificación de predicción 54 del decodificador de vídeo 60 puede incluir la unidad de compensación de movimiento 86, que decodifica bloques intercodificados y posiblemente incluye uno o más filtros de interpolación para la interpolación de subpíxeles en el proceso de compensación de movimiento. La unidad de decodificación de predicción 54 también puede incluir una unidad de intrapredicción para decodificar intramodos. La unidad de decodificación de predicción 54 puede prestar soporte a una pluralidad de modos 35, incluido el modo de fusión bipredictivo 55X. La unidad de filtro 57 puede filtrar la salida del sumador 64, y puede recibir información de filtro decodificado por entropía a fin de definir los coeficientes de filtro aplicados en el filtrado de bucle.
 45
 50

55 **[0065]** Tras recibir datos de vídeo codificados, la unidad de decodificación de entropía 52 realiza una decodificación recíproca a la codificación realizada por la unidad de codificación por entropía 46 (del codificador 50 en la FIG. 4). En el decodificador, la unidad de decodificación por entropía 52 analiza el flujo de bits para determinar las LCU y la partición correspondiente asociada con las LCU. En algunos ejemplos, una LCU o las CU de la LCU pueden definir los modos de codificación que se usaron, y estos modos de codificación pueden incluir el modo de fusión bipredictivo. Por consiguiente, la unidad de decodificación por entropía 52 puede enviar la información sintáctica a la unidad de predicción que identifique el modo de fusión bipredictivo. En este caso, la información sintáctica puede incluir uno o más elementos sintácticos que identifiquen dos bloques de vídeo vecinos diferentes codificados en los modos unipredictivos. En este caso, la unidad MC 86 de la unidad de decodificación de predicción 54 puede usar la
 60

información de movimiento de los dos bloques de vídeo vecinos diferentes para decodificar un bloque de vídeo actual de acuerdo con el modo de fusión bipredictivo. Es decir, la unidad MC 86 puede buscar datos predictivos identificados por la información de movimiento de los dos bloques de vídeo vecinos diferentes, y usar alguna combinación de estos datos predictivos para decodificar el bloque de vídeo actual en el modo de fusión bipredictivo.

5 **[0066]** La FIG. 7 es una ilustración conceptual que muestra un ejemplo de cinco vecinos diferentes que se pueden tener en cuenta con los fines del modo de fusión bipredictivo. En este ejemplo, el vecino superior (T), el vecino superior derecho (TR), el vecino izquierdo (L), el vecino inferior izquierdo (BL) y el vecino temporal coubicado (Temp) desde otra trama de vídeo se pueden tener en cuenta con los fines del modo de fusión bipredictivo. Por supuesto, otros vecinos (espaciales o temporales) también podrían usarse para cualquier herencia de modo de fusión de la información de movimiento.

10 **[0067]** De nuevo, con el modo de fusión, el bloque de vídeo actual puede heredar toda la información de movimiento de un bloque candidato vecino. Eso significa que el bloque actual tendrá el mismo vector de movimiento, la misma trama de referencia y el mismo modo de predicción (unipredicción o bipredicción) que el bloque vecino seleccionado. El bloque vecino seleccionado puede señalarse como parte de un flujo de bits codificado, pero la información de movimiento no necesita señalarse porque el decodificador puede obtener la información de movimiento del bloque vecino seleccionado.

20 **[0068]** Compatible con esta divulgación, se presta soporte a un modo de fusión bipredictivo, que hereda la información de movimiento de dos bloques vecinos diferentes, en el que cada uno de los dos bloques vecinos diferentes se codificó en un modo unipredictivo. El modo de fusión bipredictivo descrito puede aumentar el número de candidatos bipredictivos que pueden usarse en el contexto del modo de fusión. En lugar de señalar a un vecino, el modo de fusión bipredictivo puede señalar a dos vecinos diferentes. El modo de fusión bipredictivo puede ser una extensión de un modo de fusión convencional incrementando simplemente los bloques vecinos candidatos para incluir combinaciones de los mismos, o puede ser un modo completamente independiente en relación con un modo de fusión convencional.

30 **[0069]** Suponiendo los bloques vecinos candidatos espaciales y temporales mostrados en la FIG. 1, el modo de fusión bipredictivo descrito puede funcionar en al menos dos escenarios. En un primer escenario, todos los bloques candidatos vecinos están codificados en los modos unipredictivos. En este caso, se puede seleccionar cualquiera de los dos bloques candidatos, y la información de movimiento de ambos candidatos seleccionados se puede combinar para lograr una bipredicción. Por ejemplo, supongamos que los bloques vecinos mostrados en la FIG. 7 se codifican de acuerdo con la siguiente información:

35 L: unipred, L0, reldx = 0
 T: unipred, L1, reldx = 0
 TR: unipred, L0, reldx = 1
 BL: unipred, L0, reldx = 0
 40 Temp: unipred, L1, reldx = 1

En este caso, hay 10 combinaciones de dos de cada cinco candidatos. L0 se refiere a una primera lista de datos predictivos, y L1 se refiere a una segunda lista de datos predictivos. El reldx puede comprender un índice a una imagen particular en la respectiva lista. El codificador de vídeo puede seleccionar la mejor combinación (por ejemplo, en términos de velocidad de codificación y distorsión) y puede enviar información sintáctica que identifique los dos bloques vecinos seleccionados. El decodificador puede decodificar la información sintáctica y obtener la información de movimiento de los bloques vecinos seleccionados.

50 **[0070]** En un segundo escenario, al menos uno de los bloques vecinos candidatos puede codificarse en un modo bipredictivo. En este caso, cualquiera de los dos bloques vecinos unipredictivos se puede tener en cuenta (en combinación) para obtener un candidato de modo de fusión bipredictivo. Sin embargo, cualquier bloque bipredictivo también se puede usar solo por consideración como un candidato de fusión bipredictivo. Por ejemplo, supongamos que los bloques vecinos mostrados en la FIG. 7 se codifican de acuerdo con la siguiente información:

55 L: bipred, L0, reldx = 0, L1, reldx = 0
 T: unipred, L1, reldx = 0
 TR: unipred, L0, reldx = 1
 BL: unipred, L0, reldx = 0
 60 Temp: bipred, L0, reldx = 0, L1, reldx = 1

De nuevo, L0 puede comprender un valor que se refiera a una primera lista de datos predictivos, L1 puede comprender un valor que se refiera a una segunda lista de datos predictivos, y refldx pueden ser valores que definan índices para una imagen particular en la respectiva lista. En este segundo ejemplo, dos de los cinco candidatos ya están en el modo bipredictivo, por lo que estos pueden tenerse en cuenta solos con los fines de los modos de fusión bipredictivos. Además, se pueden tener en cuenta diferentes combinaciones de los tres candidatos unipredictivos restantes. Entonces, en este caso, habrá 5 posibilidades de modo bipredictivo:

1. L
2. Temp
3. T + TR
4. T + BL
5. TR + BL

[0071] En este segundo ejemplo, el codificador puede seleccionar el mejor vecino (o combinación de vecinos) de estas cinco posibilidades (por ejemplo, en términos de velocidad de codificación y distorsión), y puede enviar información sintáctica que identifique qué vecino individual o combinación de vecinos se usó en el modo de fusión. El decodificador puede decodificar la información sintáctica y obtener la información de movimiento del o de los bloques vecinos seleccionados.

[0072] En el ejemplo de la FIG. 7, se muestran cinco candidatos. Sin embargo, también pueden tenerse en cuenta candidatos adicionales en las mismas zonas que los candidatos de la FIG. 7 o en otras zonas. En algunos casos, puede haber varios candidatos superiores (T), varios candidatos superiores izquierdos (TL), varios candidatos izquierdos (L), varios candidatos inferiores izquierdos (BL) y varios candidatos temporales (T). En algunos casos, el tamaño del bloque actual puede diferir del de los candidatos, en cuyo caso el borde superior o el borde izquierdo del bloque actual pueden ser adyacentes a varios candidatos. En otros casos, los candidatos que se encuentran a distancias incluso más alejadas del bloque de vídeo actual pueden tenerse en cuenta con los fines del modo de fusión bipredictivo descrito en esta divulgación. Muchos escenarios diferentes que usen muchos candidatos diferentes son posiblemente compatibles con esta divulgación. Por tanto, la FIG. 7 es meramente un ejemplo que ilustra cinco candidatos vecinos en relación con el bloque de vídeo actual.

[0073] La FIG. 8 es un diagrama de flujo que ilustra una técnica de decodificación compatible con esta divulgación. La FIG. 8 se describirá desde la perspectiva del decodificador de vídeo 60 de la FIG. 6, aunque otros dispositivos pueden realizar técnicas similares. Como se muestra en la FIG. 8, el decodificador de vídeo 60 recibe una LCU que incluye elementos sintácticos para la LCU y para las CU dentro de la LCU (801). En particular, la unidad de decodificación por entropía 52 puede recibir un flujo de bits que incluye la LCU y analizar el flujo de bits para identificar los elementos sintácticos, que pueden enviarse a la unidad de decodificación de predicción 54. Por consiguiente, la unidad de decodificación de predicción 54 tiene modos de las CU basadas en los elementos sintácticos. En otros ejemplos, los modos pueden definirse en el nivel de PU en lugar de en el nivel de CU.

[0074] Al determinar los modos, la unidad de decodificación de predicción 54 identifica cualquier CU codificada en el modo de fusión bipredictivo (803). Si no se codifica ninguna CU en el modo de fusión bipredictivo ("no" 803), entonces esa CU se decodifica de acuerdo con su modo (804). Por ejemplo, se puede prestar soporte a muchos intramodos diferentes y a muchos intermodos diferentes. Si alguna CU se codifica en el modo de fusión bipredictivo ("sí" 803), entonces esta CU es bipredictiva. Sin embargo, sus vectores de movimiento para la bipredicción proceden de dos vecinos unidireccionales como se analiza en el presente documento. En este caso, la unidad MC 86 de la unidad de decodificación de predicción 54 identifica dos vecinos unipredictivos diferentes a la CU basada en el elemento sintáctico de la CU (805), y usa la información de movimiento de los vecinos unipredictivos para decodificar la CU bipredictiva.

[0075] La FIG. 9 es un diagrama de flujo que ilustra una técnica de codificación compatible con esta divulgación. La FIG. 9 se describirá desde la perspectiva del decodificador de vídeo 50 de la FIG. 4 aunque otros dispositivos pueden realizar técnicas similares. Como se muestra en la FIG. 9, la unidad de codificación de predicción 32 selecciona un modo de fusión bipredictivo para una CU. Por ejemplo, la unidad de codificación de predicción 32 (véase la FIG. 5) puede incluir una unidad de selección de modo 75 que selecciona el modo de fusión bipredictivo 35X para una CU basada en una pluralidad de modos posibles 35. La unidad R-D 75 puede identificar la velocidad de codificación y el nivel de calidad o distorsión asociados con diferentes modos al analizar los resultados de la codificación por la unidad ME 76 y la unidad MC 77 para diversos intermodos y al analizar los resultados de la codificación por la unidad de intrapredicción 78 para los intramodos. De esta forma, la unidad de selección de modo 75 puede identificar el mejor modo para cualquier situación dada.

[0076] Una vez que la unidad de codificación de predicción 32 selecciona el modo de fusión bipredictivo 35X para una CU, la unidad de predicción identifica dos bloques vecinos diferentes codificados en los modos unipredictivos (902). Este proceso de identificación de los dos bloques vecinos diferentes codificados en los modos unipredictivos puede realizarse por la unidad ME 76 de manera similar a la descrita anteriormente. Por ejemplo, la unidad ME 76 y la unidad MC 77 pueden generar resultados de codificación de diferentes combinaciones de información de movimiento de diferentes vecinos unidireccionales, y estos resultados pueden analizarse por la unidad R-D 75 para determinar la velocidad de codificación y la calidad o la distorsión asociada con dichas combinaciones diferentes. Finalmente, la unidad R-D 75 puede determinar qué combinación de modo de fusión bipredictivo de vecinos unipredictivos produce los mejores resultados de codificación.

[0077] Por consiguiente, la unidad de codificación de predicción 32 usa la información de movimiento de la mejor combinación de dos vecinos unipredictivos para codificar la CU como bipredictiva (903). Por supuesto, cualquier vecino bipredictivo también puede tenerse en cuenta, y posiblemente usarse para la codificación del modo de fusión si los resultados de distorsión de velocidad son mejores que usar dos vecinos unipredictivos. La unidad de predicción 32 (por ejemplo, la unidad ME 76 o la unidad de MC 77) genera uno o más elementos sintácticos para que la CU identifique los dos vecinos unipredictivos diferentes usados para la codificación del modo de fusión bipredictivo de la CU. Los elementos sintácticos, por ejemplo, pueden comprender valores de índice que identifiquen a dos de los vecinos de la CU, tal como el vecino izquierdo (L), el vecino inferior izquierdo (BL), el vecino superior (T), el vecino superior derecho (TR), o el vecino temporal (T) cubricado como se ilustra conceptualmente en la FIG. 7. Sin embargo, también se pueden usar muchos otros sistemas de señalización para los elementos sintácticos.

[0078] Las técnicas de esta divulgación se pueden realizar en una gran diversidad de dispositivos o aparatos, incluidos un teléfono 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 la realización por diferentes unidades de hardware.

[0079] Por consiguiente, las técnicas descritas en el presente documento pueden implementarse en hardware, software, firmware o cualquier combinación de los mismos. Todas las características descritas como módulos o componentes pueden implementarse juntas en un dispositivo lógico integrado o por separado, como dispositivos lógicos discretos pero interoperables. Si se implementan en software, las técnicas pueden realizarse al menos parcialmente 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 de almacenamiento de datos legible por ordenador puede formar parte de un producto de programa informático, que puede incluir materiales de embalaje.

[0080] El medio legible por ordenador puede comprender un medio tangible de almacenamiento legible por ordenador, tal una como memoria de acceso aleatorio (RAM) tal como memoria de acceso aleatorio dinámica síncrona (SDRAM), una memoria de sólo lectura (ROM), una memoria de acceso aleatorio no volátil (NVRAM), una memoria de sólo lectura programable y borrable eléctricamente (EEPROM), una memoria FLASH, medios de almacenamiento de datos magnéticos u ópticos, y similares. Las técnicas pueden realizarse adicionalmente, o de forma alternativa, al menos parcialmente 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 pueda accederse, y que pueda leerse y/o ejecutarse por un ordenador.

[0081] Las instrucciones pueden ejecutarse por uno o más procesadores, tales como uno o más procesadores de señales digitales (DSP), microprocesadores de uso general, circuitos integrados específicos de la aplicación (ASIC), matrices lógicas programables por campo (FPGA) u otros circuitos lógicos, integrados o discretos, equivalentes. El término «procesador», como se usa en el presente documento, puede referirse a cualquiera de las estructuras anteriores o a cualquiera de 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 puede proporcionarse dentro de software o hardware dedicados configurados para codificar y decodificar, o incorporarse en un codificador-decodificador de vídeo combinado (CODEC). También, las técnicas se podrían implementar completamente en uno o más circuitos o elementos lógicos.

REIVINDICACIONES

1. Un procedimiento para codificación de datos de vídeo, el procedimiento que comprende:
 - 5 seleccionar un modo de fusión bipredictivo para codificar un bloque de vídeo actual;
 - identificar un primer conjunto de dos bloques de vídeo vecinos diferentes codificados en los modos unipredictivos de un conjunto de bloques vecinos candidatos, el conjunto de bloques vecinos candidatos que comprende un vecino superior, un vecino superior derecho, un vecino izquierdo, un vecino inferior izquierdo y un vecino temporal coubicado desde otra trama de vídeo;
 - 10 usar dos vectores de movimiento unipredictivos diferentes, uno asociado con cada uno de los dos bloques de vídeo vecinos diferentes de dicho primer conjunto, como un primer conjunto candidato de vectores de movimiento para codificar el bloque de vídeo actual de acuerdo con el modo de fusión bipredictivo;
 - 15 identificar un segundo conjunto de dos bloques de vídeo vecinos diferentes codificados en los modos unipredictivos del conjunto de bloques vecinos candidatos, y usar dos vectores de movimiento unipredictivos diferentes, uno asociado con cada uno de los dos bloques de vídeo vecinos diferentes de dicho segundo conjunto, como un segundo conjunto candidato de vectores de movimiento para codificar el bloque de vídeo actual de acuerdo con el modo de fusión bipredictivo;
 - 20 identificar al menos uno de los bloques vecinos del conjunto de bloques vecinos candidatos que se codifique en el modo bipredictivo, y usar los dos vectores de movimiento bipredictivos de dicho al menos uno de los bloques vecinos como un tercer conjunto candidato de vectores de movimiento para codificar el bloque de vídeo actual de acuerdo con el modo de fusión bipredictivo;
 - 25 seleccionar uno de los conjuntos candidatos de vectores de movimiento para codificar el bloque de vídeo actual de acuerdo con el modo de fusión bipredictivo; y
 - 30 generar uno o más elementos sintácticos para identificar el conjunto candidato seleccionado de vectores de movimiento.
2. El procedimiento según la reivindicación 1, en el que la información de movimiento comprende además al menos dos valores asociados con el primer conjunto de dos vectores de movimiento unipredictivos diferentes, en el que los valores identifican una o más listas de datos predictivos asociados con el conjunto de dos vectores de movimiento unipredictivos diferentes.
3. El procedimiento según la reivindicación 1, en el que el bloque de vídeo actual comprende una unidad de codificación (CU) definida de acuerdo con una norma de codificación de vídeo de alta eficiencia (HEVC), el procedimiento que comprende además:
 - 40 definir la CU en relación con una unidad de codificación más grande (LCU) de acuerdo con un sistema de partición en árbol cuaternario;
 - 45 generar datos sintácticos de LCU que definan el sistema de partición en árbol cuaternario; y
 - generar información de modo para la CU que defina el modo de fusión bipredictivo, en el que uno o más elementos sintácticos se incluyen en la información de modo para la CU.
4. El procedimiento según la reivindicación 1, en el que el bloque de vídeo actual comprende una unidad de predicción (PU) de una unidad de codificación (CU) que se define de acuerdo con una norma de codificación de vídeo de alta eficiencia (HEVC).
5. El procedimiento según la reivindicación 1, en el que la selección de un conjunto de vectores de movimiento candidatos se basa en una determinación de velocidad y distorsión.
6. Un procedimiento para decodificación de datos de vídeo, el procedimiento que comprende:
 - 60 recibir uno o más elementos sintácticos para un bloque de vídeo actual, en el que el bloque de vídeo actual se codifica de acuerdo con un modo de fusión bipredictivo;

basándose en uno o más elementos sintácticos, identificar:

- 5 uno de una pluralidad de conjuntos de dos bloques de vídeo vecinos diferentes codificados en los modos unipredictivos, o un bloque vecino del conjunto de bloques vecinos candidatos que se codifique en el modo bipredictivo, en el que los bloques de vídeo vecinos son de un conjunto de bloques vecinos candidatos, el conjunto de bloques vecinos candidatos que comprende un vecino superior, un vecino superior derecho, un vecino izquierdo, un vecino inferior izquierdo y un vecino temporal coubicado desde otra trama de vídeo; donde el elemento sintáctico identifica uno o una pluralidad de conjuntos de dos bloques de vídeo vecinos diferentes codificados en el modo unipredictivo, usando dos vectores de movimiento unipredictivos diferentes, uno asociado con cada uno de los dos bloques de vídeo vecinos diferentes para decodificar un bloque de vídeo actual de acuerdo con el modo de fusión bipredictivo; y,
- 10 donde el elemento sintáctico identifica un bloque de vídeo vecino codificado en el modo bipredictivo usando los dos vectores de movimiento bipredictivo asociados con el bloque de vídeo vecino para decodificar un bloque de vídeo actual de acuerdo con el modo de fusión bipredictivo.
- 15
7. El procedimiento según la reivindicación 6, en el que la información de movimiento comprende además al menos dos valores asociados con los dos vectores de movimiento unipredictivos diferentes, en el que los valores identifican una o más listas de datos predictivos asociados con los dos vectores de movimiento unipredictivos diferentes.
- 20
8. El procedimiento según la reivindicación 6, en el que el bloque de vídeo actual comprende una unidad de codificación (CU) definida de acuerdo con una norma de codificación de vídeo de alta eficiencia (HEVC), en el que la CU se define con relación a una unidad de codificación más grande (LCU) de acuerdo con un sistema de partición en árbol cuaternario, el procedimiento que comprende además:
- 25
- recibir datos sintácticos de LCU que definan el sistema de partición en árbol cuaternario; y
- recibir información de modo para la CU que defina el modo de fusión bipredictivo, en el que el uno o más elementos sintácticos se incluyen en la información de modo para la CU.
- 30
9. El procedimiento según la reivindicación 6, en el que el bloque de vídeo actual comprende una unidad de predicción (PU) de una unidad de codificación (CU) que se define de acuerdo con una norma de codificación de vídeo de alta eficiencia (HEVC).
- 35
10. Un dispositivo para decodificar datos de vídeo, el dispositivo que comprende:
- 40
- medios para recibir uno o más elementos sintácticos para un bloque de vídeo actual, en el que el bloque de vídeo actual se codifica de acuerdo con un modo de fusión bipredictivo;
- medios para identificar uno de una pluralidad de conjuntos de dos bloques de vídeo vecinos diferentes codificados en los modos unipredictivos o de un bloque vecino del conjunto de bloques vecinos candidatos que se codifique en el modo bipredictivo, basados en uno o más elementos sintácticos, en los que los bloques de vídeo vecinos son de un conjunto de bloques vecinos candidatos, el conjunto de bloques vecinos candidatos que comprende un vecino superior, un vecino superior derecho, un vecino izquierdo, un vecino inferior izquierdo y un vecino temporal coubicado desde otra trama de vídeo;
- 45
- medios para usar dos vectores de movimiento unipredictivos diferentes, uno asociado con cada uno de los conjuntos de dos bloques de vídeo vecinos diferentes para decodificar un bloque de vídeo actual de acuerdo con el modo de fusión bipredictivo donde el elemento sintáctico identifica uno o una pluralidad de conjuntos de dos bloques de vídeo vecinos diferentes codificados en el modo unipredictivo o para usar los dos vectores de movimiento bipredictivos asociados con el bloque de vídeo vecino para decodificar un bloque de vídeo actual de acuerdo con el modo de fusión bipredictivo donde el elemento sintáctico identifica un bloque de vídeo vecino codificado en el modo bipredictivo.
- 50
- 55
11. Un dispositivo para codificar datos de vídeo, el dispositivo que comprende:
- medios para seleccionar un modo de fusión bipredictivo para codificar un bloque de vídeo actual;
- 60
- medios para identificar un primer conjunto de dos bloques de vídeo vecinos diferentes codificados en los modos unipredictivos de un conjunto de bloques vecinos candidatos, el conjunto de bloques vecinos

candidatos que comprende un vecino superior, un vecino superior derecho, un vecino izquierdo, un vecino inferior izquierdo y un vecino temporal coubicado desde otra trama de vídeo;

5 medios para usar dos vectores de movimiento unipredictivos diferentes, uno asociado con cada uno de los dos bloques de vídeo vecinos diferentes de dicho primer conjunto, como un primer conjunto candidato de vectores de movimiento para codificar el bloque de vídeo actual de acuerdo con el modo de fusión bipredictivo;

10 medios para identificar un segundo conjunto de dos bloques de vídeo vecinos diferentes codificados en los modos unipredictivos del conjunto de bloques vecinos candidatos, y usar dos vectores de movimiento unipredictivos diferentes, uno asociado con cada uno de los dos bloques de vídeo vecinos diferentes de dicho segundo conjunto, como un segundo conjunto candidato de vectores de movimiento para codificar el bloque de vídeo actual de acuerdo con el modo de fusión bipredictivo;

15 medios para identificar al menos uno de los bloques vecinos del conjunto de bloques vecinos candidatos que se codifique en el modo bipredictivo, y usar los dos vectores de movimiento bipredictivos de dicho al menos uno de los bloques vecinos como un tercer conjunto candidato de vectores de movimiento para codificar el bloque de vídeo actual de acuerdo con el modo de fusión bipredictivo;

20 medios para seleccionar uno de los conjuntos candidatos de vectores de movimiento para codificar el bloque de vídeo actual de acuerdo con el modo de fusión bipredictivo; y,

25 medios para generar uno o más elementos sintácticos para identificar el conjunto candidato seleccionado de vectores de movimiento.

- 12.** Un medio legible por ordenador que comprende instrucciones que, cuando se ejecutan, hacen que un procesador codifique datos de vídeo, en el que las instrucciones hacen que el procesador realice un procedimiento de acuerdo con una cualquiera de las reivindicaciones 1 a 9.

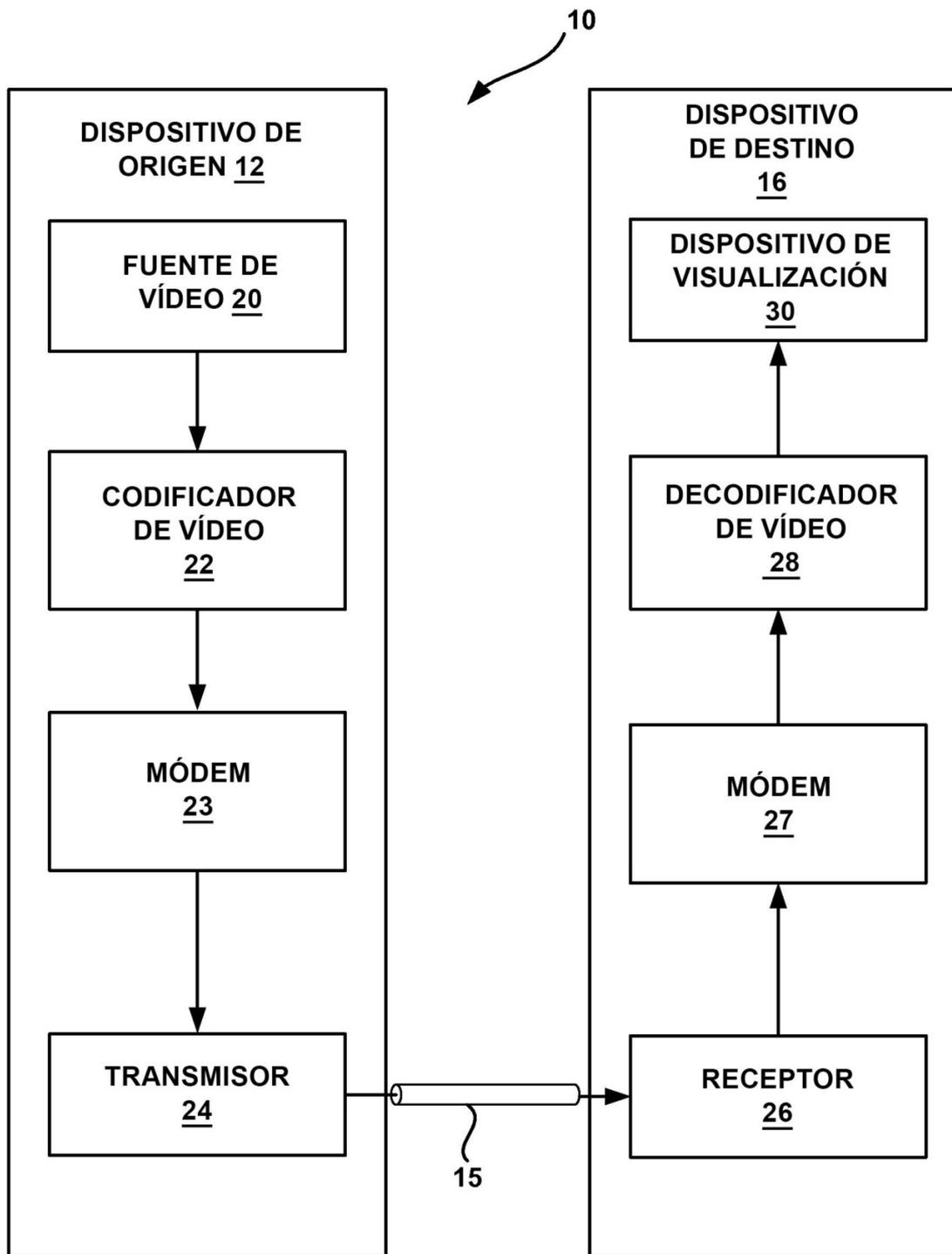


FIG. 1

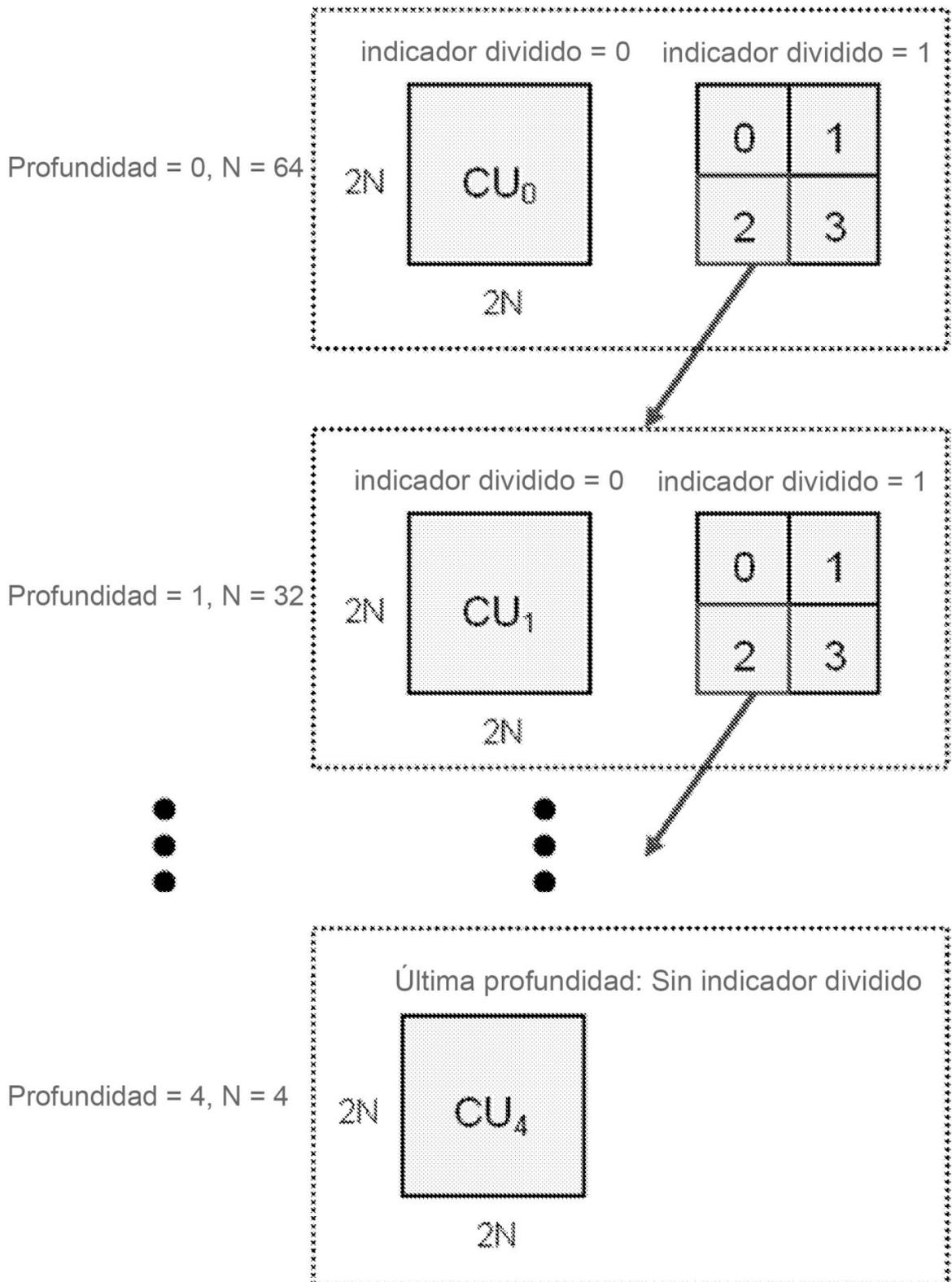


FIG. 2

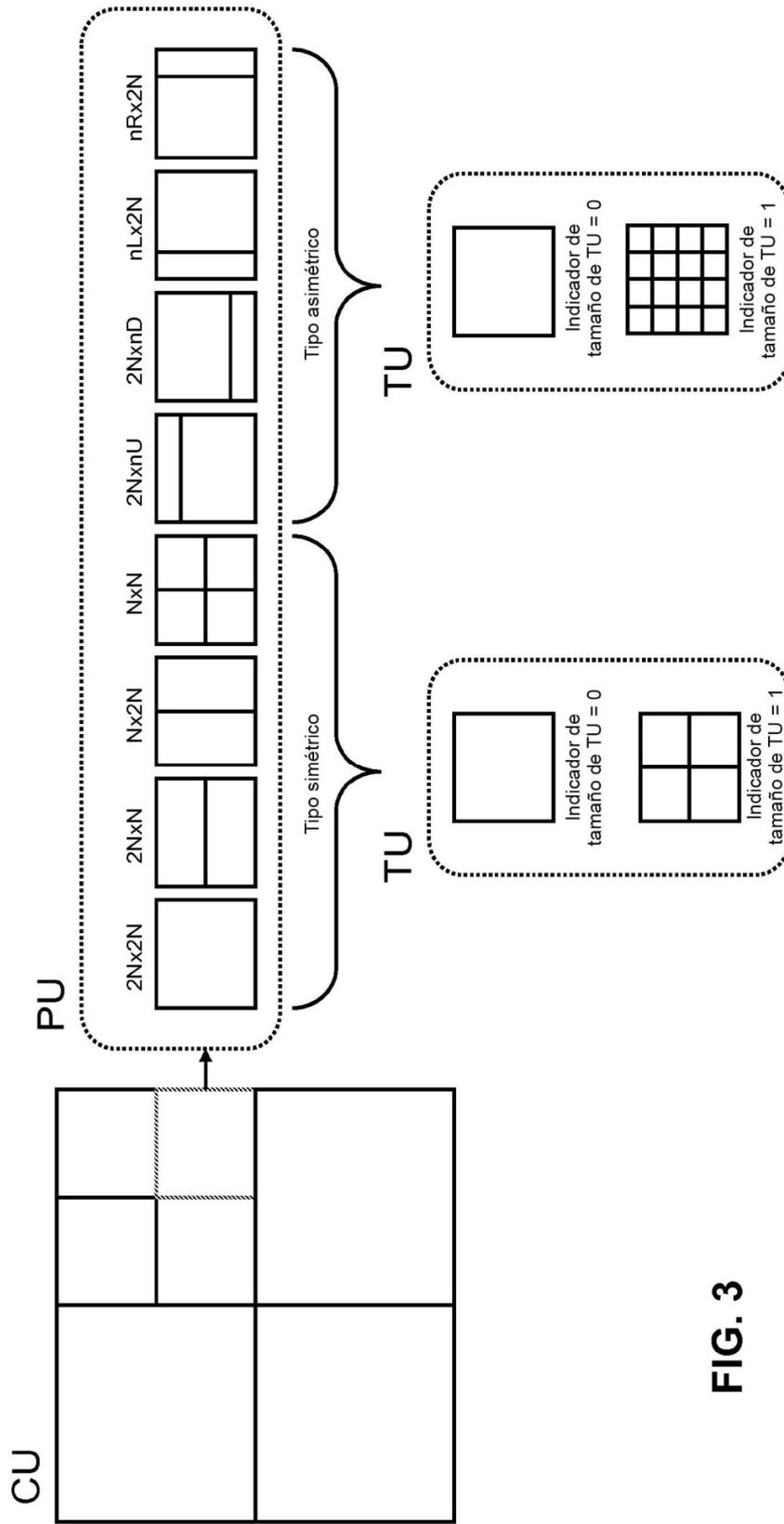


FIG. 3

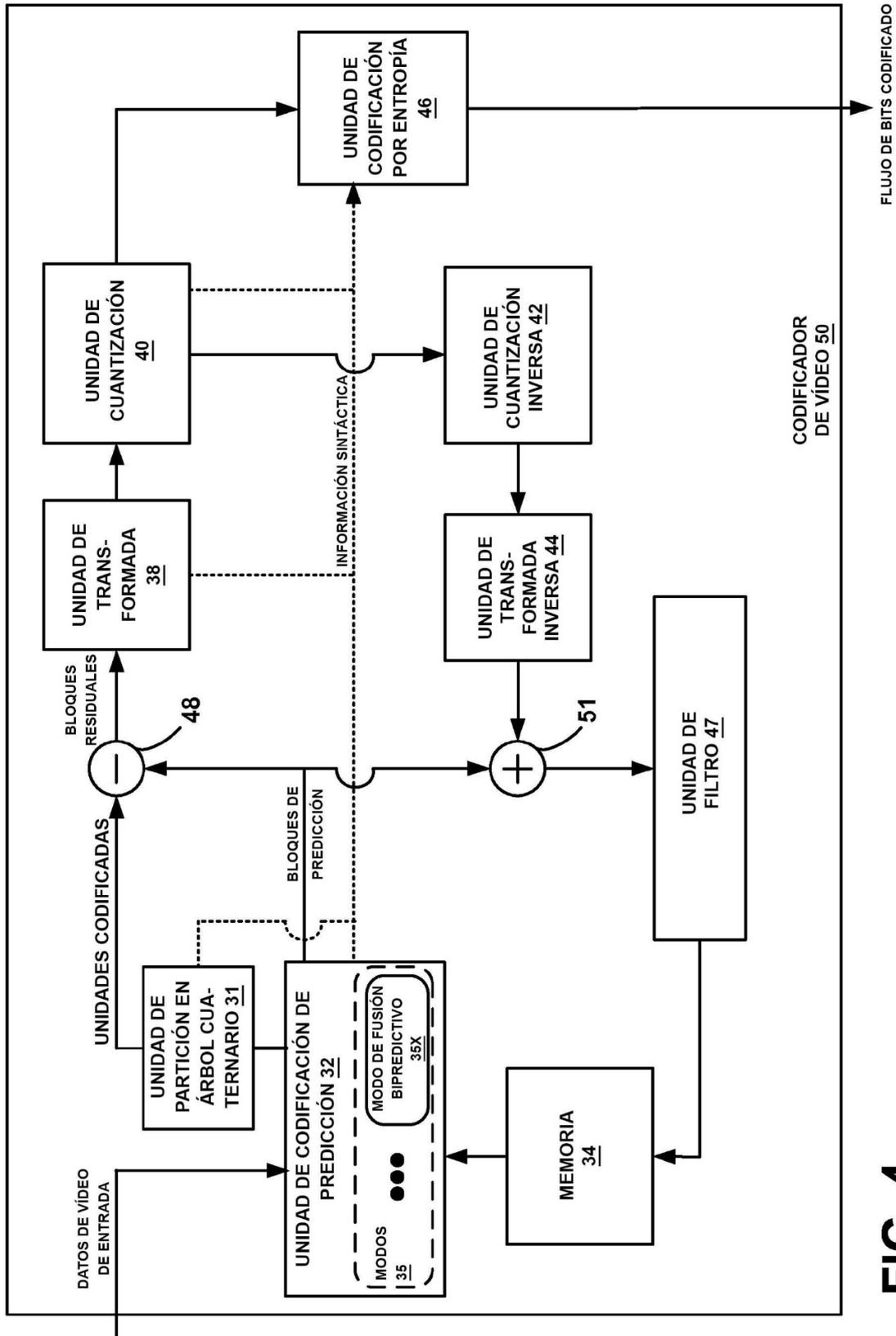


FIG. 4

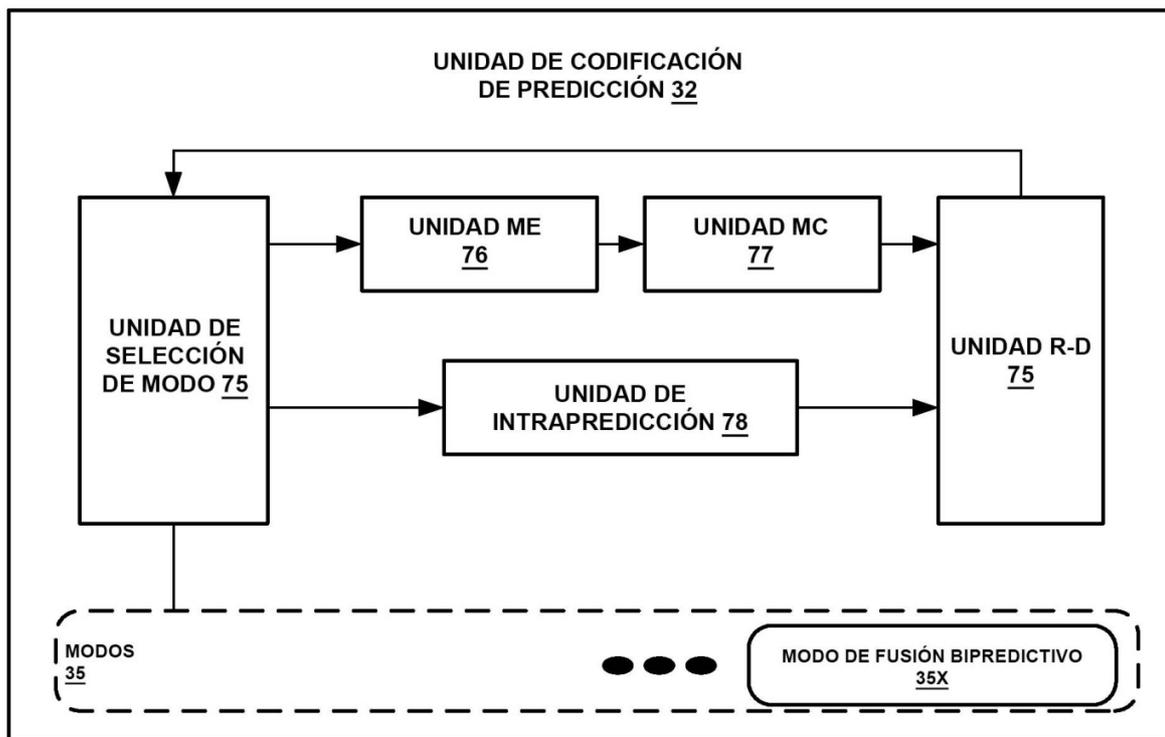


FIG. 5

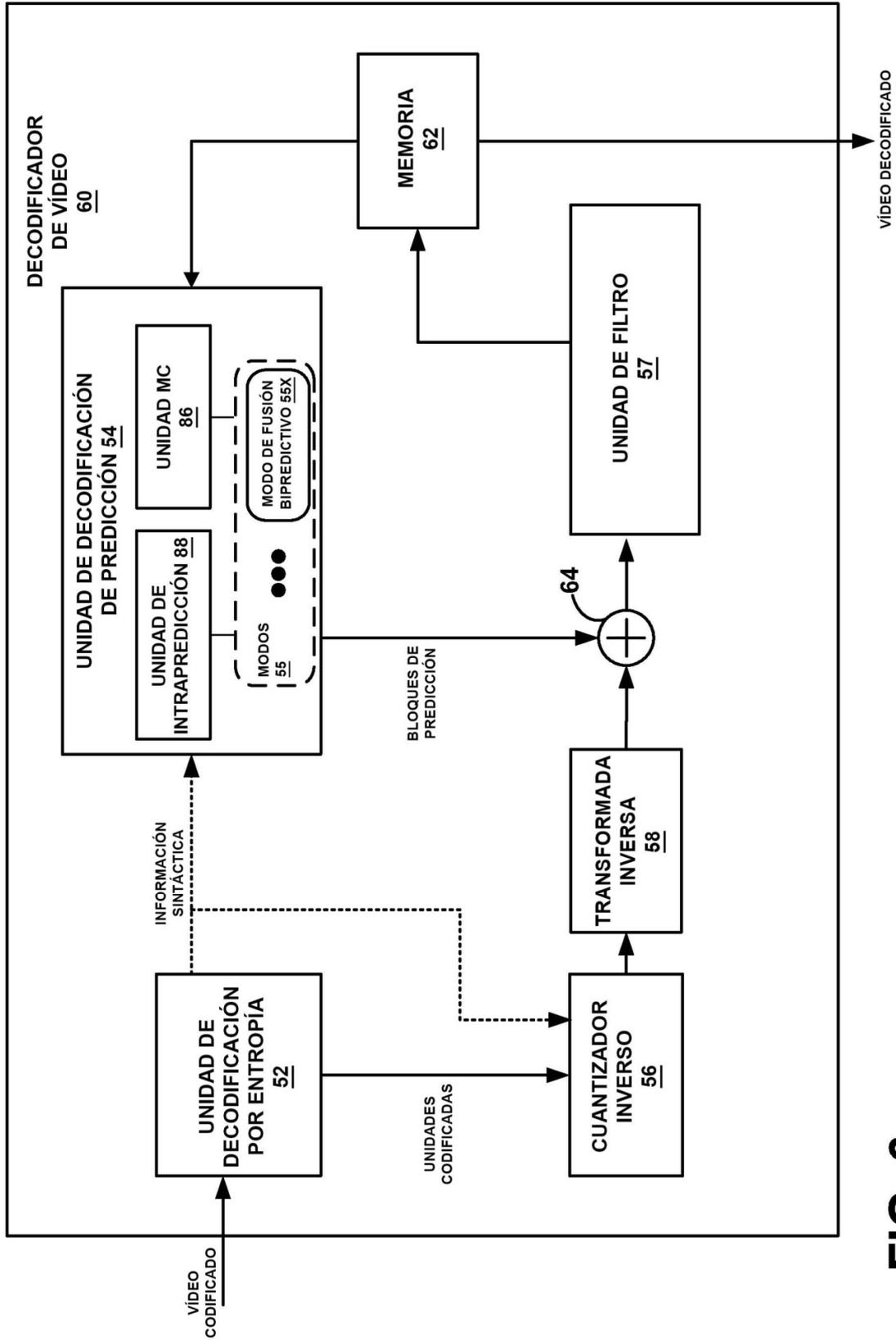


FIG. 6

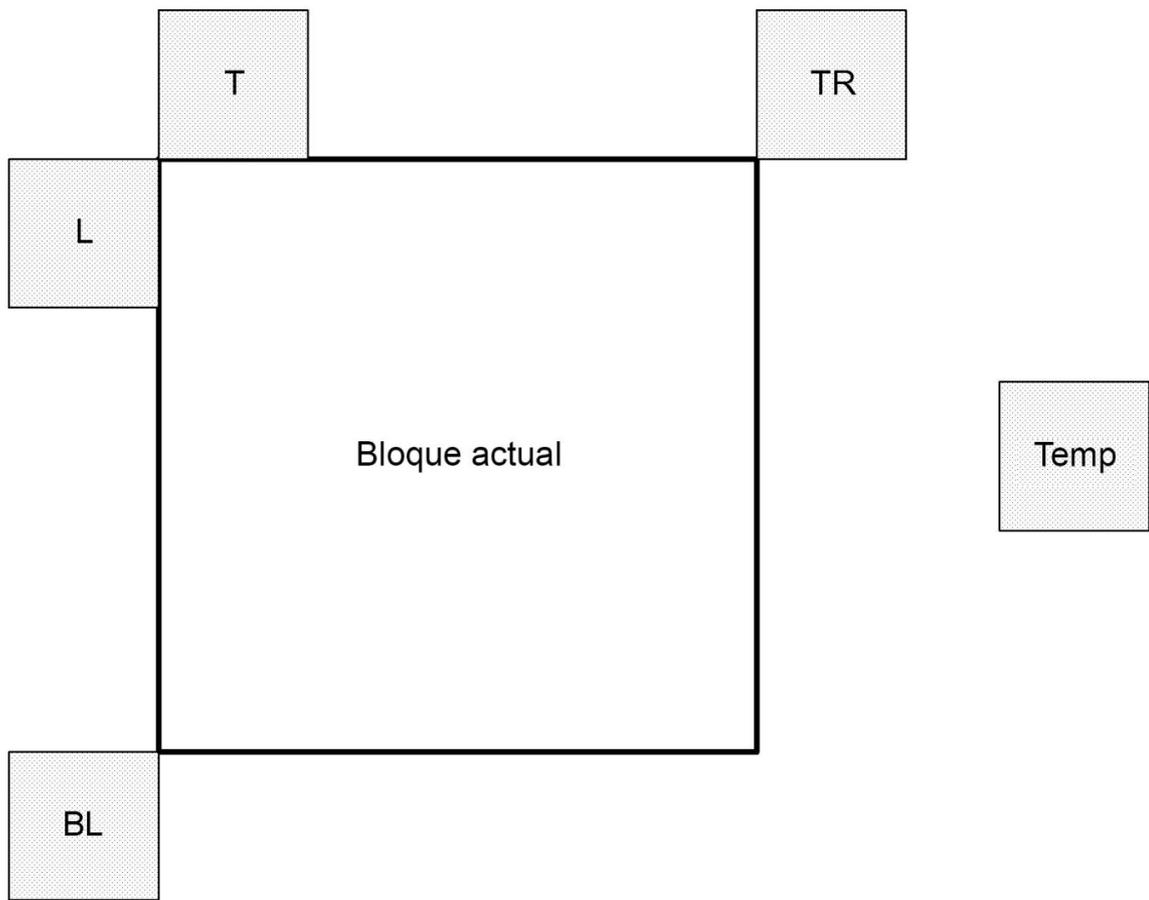


FIG. 7

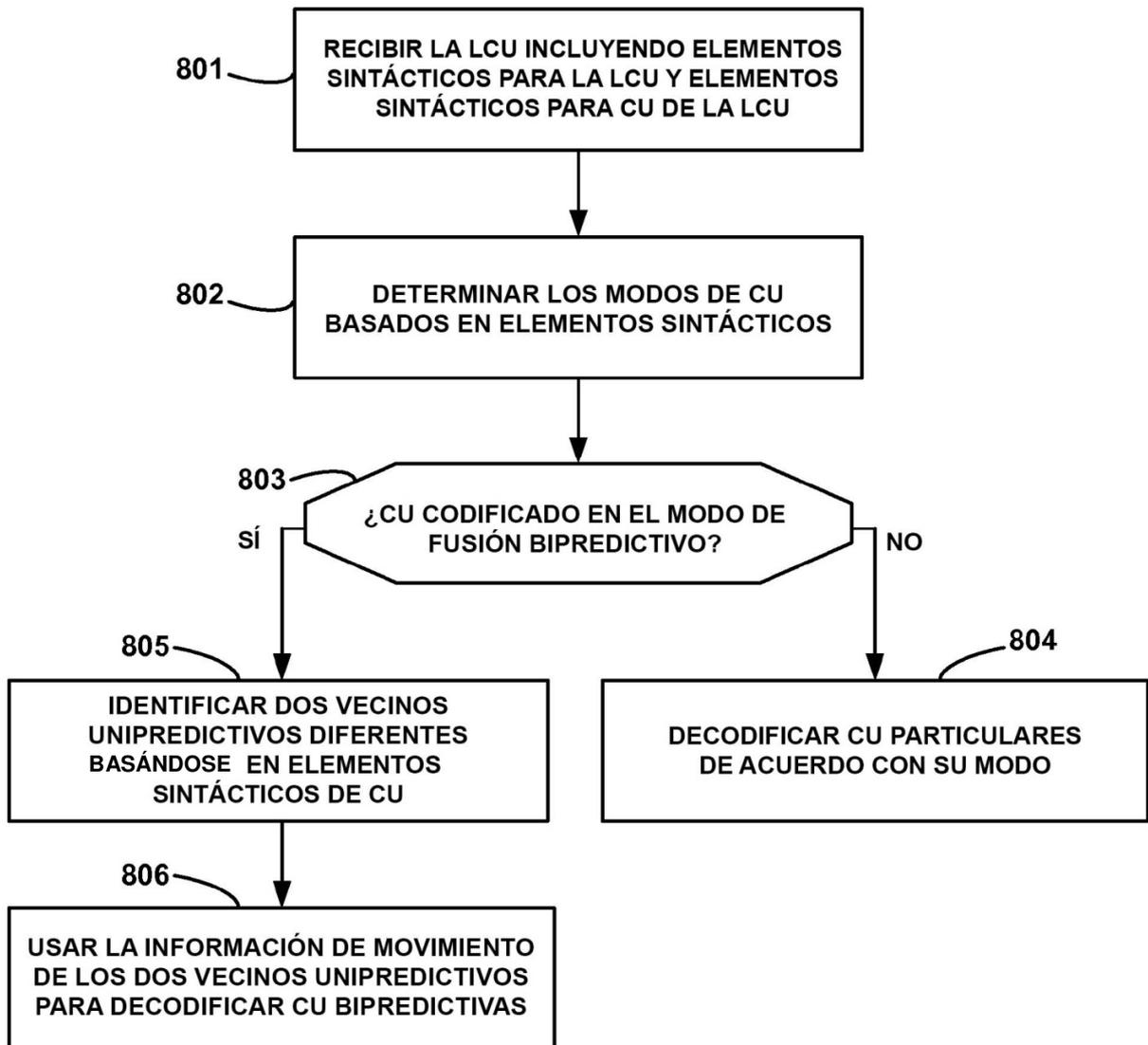


FIG. 8

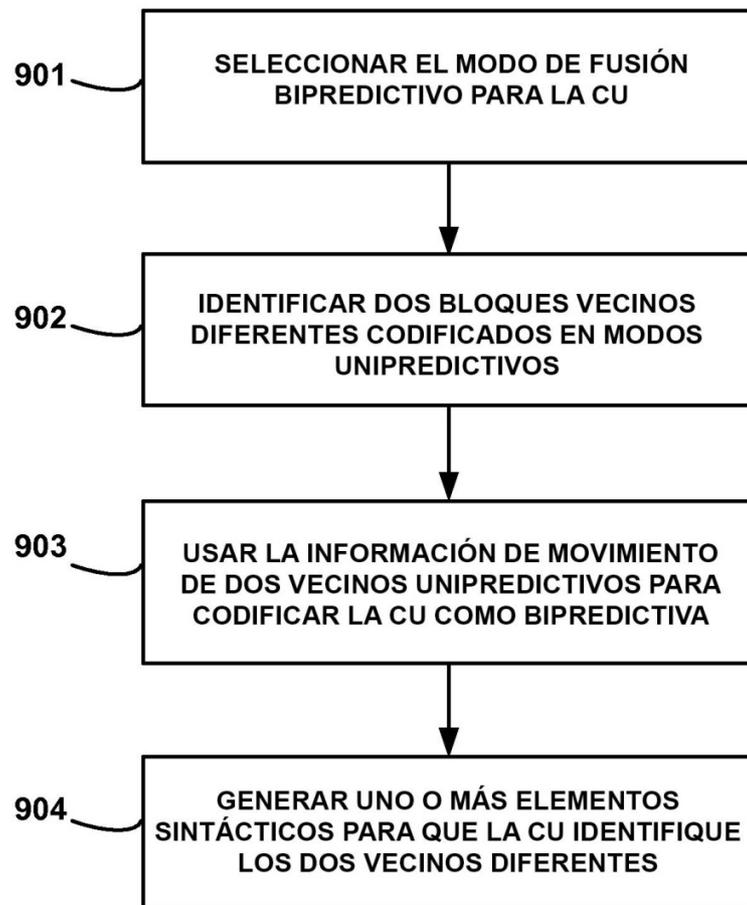


FIG. 9