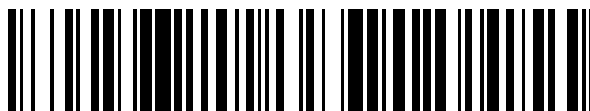


19



OFICINA ESPAÑOLA DE  
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 663 691**

51 Int. Cl.:

**H04N 19/105** (2014.01)

**H04N 19/172** (2014.01)

**H04N 19/147** (2014.01)

**H04N 19/19** (2014.01)

**H04N 19/573** (2014.01)

**H04N 19/577** (2014.01)

12

## TRADUCCIÓN DE PATENTE EUROPEA

T3

86 Fecha de presentación y número de la solicitud internacional: **06.10.2009 PCT/US2009/059698**

87 Fecha y número de publicación internacional: **15.04.2010 WO10042518**

96 Fecha de presentación y número de la solicitud europea: **06.10.2009 E 09736529 (0)**

97 Fecha y número de publicación de la concesión europea: **24.01.2018 EP 2345255**

54 Título: **Selección eficiente de modo de predicción**

30 Prioridad:

**06.10.2008 US 103100 P**

**15.04.2009 US 424472**

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

**16.04.2018**

73 Titular/es:

**QUALCOMM INCORPORATED (100.0%)**

**5775 Morehouse Drive**

**San Diego, CA 92121, US**

72 Inventor/es:

**PANCHAL, RAHUL y**

**KARCZEWICZ, MARTA**

74 Agente/Representante:

**FORTEA LAGUNA, Juan José**

ES 2 663 691 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

Selección eficiente de modo de predicción

5 **[1]** Esta solicitud reivindica la prioridad de la solicitud provisional estadounidense n. ° 61/103.100 presentada el 6 de octubre de 2008.

10 **[2]** Un documento titulado "H.264/MPEG-4 AVC Reference Software Enhancements", Documento N. ° VCEG-N19 del Grupo de expertos en codificación de vídeo de la UIT-T, del 17 al 21 de enero de 2005, describe la predicción ponderada por defecto y los modos de predicción ponderada implícita para unidades de datos de vídeo. El documento WO 2007/092215 describe técnicas que usan un modo de predicción ponderada. Ambos documentos no divulgan una selección entre un modo de predicción ponderada implícita y un modo de predicción ponderada explícita en base a una compensación, antes de comparar con el modo de predicción ponderada por defecto.

### CAMPO TÉCNICO

20 **[3]** Esta divulgación se refiere al vídeo digital y, más particularmente, a técnicas para codificar datos de vídeos digitales.

### ANTECEDENTES

25 **[4]** Se han desarrollado varias técnicas de codificación y decodificación de vídeo para codificar y decodificar datos de vídeo digital. El Grupo de Expertos en Imágenes en Movimiento (MPEG, por sus siglas en inglés), por ejemplo, ha desarrollado varias técnicas que incluyen MPEG-1, MPEG-2 y MPEG-4. Otros ejemplos incluyen la norma de la Unión Internacional de Telecomunicaciones (UIT) -T H.263 y la norma UIT-T H.264 y su equivalente, ISO/CEI MPEG-4, Parte 10, es decir, la Codificación de Vídeo Avanzada (AVC, por sus siglas en inglés). Estas normas de vídeo soportan una transmisión y un almacenamiento eficientes de datos de vídeo mediante la codificación de datos de una manera comprimida para reducir la cantidad de datos.

30 **[5]** La compresión de vídeo puede implicar una predicción espacial y/o temporal para reducir la redundancia inherente a las secuencias de vídeo. La intracodificación usa la predicción espacial para reducir la redundancia espacial de bloques de vídeo dentro de la misma trama de vídeo. La intercodificación usa la predicción temporal para reducir la redundancia temporal entre bloques de vídeo en tramas de vídeo sucesivos. Para la intercodificación, un codificador de vídeo realiza una estimación de movimiento para generar vectores de movimiento que indiquen el desplazamiento de los bloques de vídeo con relación a los correspondientes bloques de vídeo de predicción en una o más tramas de referencia.

35 **[6]** Un dispositivo de origen puede emplear una de las técnicas de codificación de vídeo anteriores para codificar los datos de vídeo digital. El dispositivo de origen archiva los datos de vídeo codificados y/o transmite los datos de vídeo codificados a un dispositivo de destino a través de un canal de transmisión. El canal de transmisión puede hacer uso de medios de comunicación alámbricos y/o inalámbricos. El dispositivo de destino recibe los datos de vídeo codificados y decodifica los datos de vídeo recibidos para recuperar los datos de vídeo digital originales para la reproducción. Muchos dispositivos incluyen tanto un codificador como un decodificador, que pueden combinarse en el denominado códec.

### SUMARIO

50 **[7]** En general, esta divulgación describe técnicas para seleccionar eficientemente un modo de predicción mediante el que predecir, a partir de al menos dos unidades codificadas de referencia, una tercera unidad codificada de datos de vídeo, que puede denominarse unidad codificada predictiva. El término "unidad predictiva" como se usa en el presente documento se refiere a una unidad que se ha "predicho". Las unidades codificadas de referencia se refieren a unidades codificadas que se han codificado previamente y que proporcionan una base a partir de la que predecir la unidad codificada predictiva. Especialmente, estas unidades codificadas de referencia pueden residir antes o después de la unidad codificada predictiva temporalmente. Típicamente, un codificador de vídeo implementa las técnicas para mejorar la eficiencia de codificación al reducir un número de pases de codificación predictivos requeridos para seleccionar un modo de predicción. Como las técnicas pueden reducir un número de pases de codificación, que generalmente pueden denominarse operaciones computacionales, necesarias para seleccionar un modo de predicción apropiado, las técnicas también pueden promover un consumo de energía eficiente. En otras palabras, el codificador de vídeo puede preseleccionar de manera inteligente un modo de predicción a partir de una pluralidad de modos de predicción sin tener que implementar todos y cada uno de los modos de predicción y entonces realizar un análisis comparativo de los resultados de cada predicción para seleccionar el modo de predicción más apropiado de la pluralidad.

65 **[8]** Más particularmente, el codificador de vídeo puede implementar las técnicas de esta divulgación para

seleccionar eficientemente un modo de predicción bidireccional compensada por movimiento por el que predecir unas unidades de datos de vídeo de un tipo particular de unidad codificada predictiva denominada trama B. Como ejemplo, dentro de la norma UIT-T H.264 y de su equivalente ISO/IEC MPEG-4, Parte 10, es decir, la norma de la Codificación de Vídeo Avanzada (AVC), se definen tres modos de predicción bidireccional compensada por movimiento para predecir una unidad de datos de vídeo de una trama B. Cada uno de estos modos de predicción bidireccional predice la unidad de datos de vídeo de la trama B a partir de al menos dos unidades de datos de vídeo de referencia, donde cada una de las unidades de datos de vídeo de referencia se incluye en unidades codificadas de referencia separadas o diferentes, tales como una trama I o una trama P, que se produce antes o después de la trama B temporalmente en los datos de vídeo.

**[9]** El primer modo se denomina modo de predicción ponderada por defecto, que aplica ponderaciones por defecto o conocidas a cada uno de los bloques de referencia para, por ejemplo, promediar los valores de píxel de la unidad de vídeo de referencia y de ese modo predecir la unidad de datos de vídeo predictivos de la trama B. El segundo modo se denomina modo de predicción ponderada explícita, donde las ponderaciones se definen explícitamente por el codificador de vídeo y se usan para adaptar la predicción de la unidad de datos de vídeo predictivos. El tercer modo se denomina modo de predicción ponderada implícita, donde las ponderaciones se determinan implícitamente mediante el codificador de vídeo a través del análisis de las unidades codificadas de referencia o de las unidades de datos de vídeo de referencia de las unidades codificadas de referencia, respectivamente, para predecir la unidad codificada predictiva.

**[10]** El codificador de vídeo puede, de acuerdo con las técnicas de esta divulgación, realizar eficientemente solamente dos de estos tres modos de predicción para generar solamente dos versiones de la tercera unidad codificada, en lugar de realizar los tres de estos modos de predicción para generar tres versiones. Como resultado, es posible que el codificador de vídeo solamente necesite analizar dos versiones en lugar de tres para determinar cuál de las dos versiones representa más apropiadamente una porción correspondiente de los datos de vídeo. A este respecto, el codificador de vídeo puede reducir el número de operaciones computacionales, por ejemplo, pases de codificación de predicción, al tiempo que promueve el consumo de energía eficiente.

**[11]** La presente invención se define en las reivindicaciones adjuntas.

**[12]** Los detalles de uno o más modos de realización de la invención se exponen en los dibujos adjuntos y en la siguiente descripción. Otras características, objetivos y ventajas de la invención resultarán evidentes a partir de la descripción, de los dibujos y de las reivindicaciones.

## BREVE DESCRIPCIÓN DE LOS DIBUJOS

### [13]

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 las técnicas de esta divulgación.

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

La FIG. 3 es un diagrama de bloques que ilustra un ejemplo de la unidad de compensación por movimiento mostrada en la FIG. 2 en más detalle.

La FIG. 4 es un diagrama de flujo que ilustra el ejemplo de funcionamiento de un codificador de vídeo al realizar las técnicas descritas en esta divulgación.

La FIG. 5 es un diagrama de flujo que ilustra un ejemplo de funcionamiento de un codificador de vídeo al realizar las técnicas descritas en esta divulgación en más detalle.

**[14]** producir temporalmente antes o después de la unidad codificada predictiva, calcular un valor de compensación para la primera versión de la unidad de datos de vídeo predictivos, seleccionar, en base al valor de compensación calculado, un modo de predicción ponderada implícita o un modo de predicción ponderada explícita, realizar el modo seleccionado para predecir una segunda versión de la unidad de datos de vídeo predictivos a partir de las primera y segunda unidades de datos de vídeo de referencia y codificar la unidad de datos de vídeo predictivos como la primera versión o la segunda versión.

**[15]** Los detalles de uno o más modos de realización de la invención se exponen en los dibujos adjuntos y en la siguiente descripción. Otras características, objetivos y ventajas de la invención resultarán evidentes a partir de la descripción y de los dibujos y a partir de las reivindicaciones.

## BREVE DESCRIPCIÓN DE LOS DIBUJOS

**[16]**

5 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 las técnicas de esta divulgación.

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

10 La FIG. 3 es un diagrama de bloques que ilustra un ejemplo de una unidad de compensación por movimiento de la FIG. 2 en más detalle.

La FIG. 4 es un diagrama de flujo que ilustra el ejemplo de funcionamiento de un codificador de vídeo al realizar las técnicas descritas en esta divulgación.

15 La FIG. 5 es un diagrama de flujo que ilustra un ejemplo de operación de funcionamiento de un codificador de vídeo al realizar las técnicas descritas en esta divulgación en más detalle.

20 La FIG. 6 es un diagrama que ilustra un ejemplo de secuencia ordenada temporalmente de unidades codificadas.

**DESCRIPCIÓN DETALLADA**

25 **[17]** En general, esta divulgación describe técnicas para seleccionar eficientemente un modo de predicción mediante el que predecir, a partir de al menos dos unidades codificadas de referencia, una tercera unidad codificada de datos de vídeo, que puede denominarse unidad codificada predictiva. Las unidades codificadas de referencia se refieren a unidades codificadas que se han codificado previamente y que proporcionan una base a partir de la que predecir la tercera unidad codificada. Especialmente, estas unidades codificadas de referencia pueden residir antes o después de la tercera unidad codificada temporalmente. Típicamente, un  
30 codificador de vídeo implementa las técnicas para mejorar la eficiencia de codificación al reducir un número de pases de codificación predictivos requeridos para seleccionar un modo de predicción. Como las técnicas pueden reducir un número de pases de codificación, que en general pueden denominarse operaciones computacionales, necesarias para seleccionar un modo de predicción óptimo o más apropiado, las técnicas también pueden promover un consumo de energía eficiente. En otras palabras, el codificador de vídeo puede  
35 preseleccionar de manera inteligente un modo de predicción a partir de una pluralidad de modos de predicción sin tener que implementar todos y cada uno de los modos de predicción y entonces realizar un análisis comparativo de los resultados de cada predicción para seleccionar el modo de predicción más apropiado.

40 **[18]** Más particularmente, el codificador de vídeo puede implementar las técnicas de esta divulgación para seleccionar eficientemente un modo de predicción bidireccional compensada por movimiento por el que predecir unas unidades de datos de vídeo de un tipo particular de unidad codificada predictiva denominada trama B. Como ejemplo, dentro de la norma UIT-T H.264 y de su equivalente ISO/IEC MPEG-4, Parte 10, es decir, la norma de la Codificación de Vídeo Avanzada (AVC), se definen tres modos de predicción bidireccional compensada por movimiento para predecir una unidad de datos de vídeo de una trama B. Cada uno de estos  
45 modos de predicción bidireccional predice la unidad de datos de vídeo de la trama B a partir de al menos dos unidades de datos de vídeo de referencia, donde cada una de las unidades de datos de vídeo de referencia se incluye dentro de unidades codificadas de referencia separadas o diferentes, tal como una trama I o una trama P, que se producen antes o después de la trama B temporalmente.

50 **[19]** El primer modo se denomina modo de predicción ponderada por defecto, que aplica ponderaciones por defecto o conocidas a cada uno de los bloques de referencia para, por ejemplo, promediar los valores de píxel de la unidad de vídeo de referencia y de ese modo predecir la unidad de datos de vídeo predictivos de la trama B. El segundo modo se denomina modo de predicción ponderada explícita, donde las ponderaciones se definen explícitamente por el codificador de vídeo. El tercer modo se denomina modo de predicción ponderada implícita, donde las ponderaciones se determinan implícitamente mediante el codificador de vídeo a través del  
55 análisis de las unidades codificadas de referencia o de las unidades de datos de vídeo de referencia de las unidades codificadas de referencia respectivamente para predecir la unidad codificada predictiva. Con respecto al modo de predicción ponderada explícita, el codificador de vídeo codifica explícitamente las ponderaciones, de ahí el nombre de predicción ponderada "explícita". Con respecto al modo de predicción ponderada implícita,  
60 el codificador de vídeo no codifica explícitamente las ponderaciones, sino que el decodificador utiliza el mismo algoritmo de determinación de ponderación que el usado por el codificador de vídeo para determinar las ponderaciones. En otras palabras, el codificador de vídeo codifica implícitamente las ponderaciones y requiere que el decodificador realice operaciones adicionales para determinar las mismas ponderaciones, en lugar de codificar explícitamente las ponderaciones y de ese modo aliviar al decodificador de tener que realizar estas  
65 operaciones adicionales.

[20] El codificador de vídeo puede, de acuerdo con las técnicas de esta divulgación, realizar eficientemente solamente dos de estos tres modos de predicción para generar solamente dos versiones de la tercera unidad codificada, en lugar de realizar los tres de estos modos de predicción para generar tres versiones. Como resultado, es posible que el codificador de vídeo solamente necesite analizar dos versiones en lugar de tres para determinar cuál de las dos versiones representa más apropiadamente una porción correspondiente de los datos de vídeo. A este respecto, el codificador de vídeo puede reducir el número de operaciones computacionales, por ejemplo, pases de codificación de predicción, al tiempo que promueve el consumo de energía eficiente.

[21] 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 las técnicas de esta divulgación. Como se muestra en la FIG. 1, el sistema 10 incluye un dispositivo de origen 12 que transmite un 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 cualquiera de una amplia gama de dispositivos. En algunos casos, el dispositivo de origen 12 y el dispositivo de destino 16 comprenden dispositivos de comunicación inalámbrica, tales como microteléfonos inalámbricos, los denominados radioteléfonos celulares o vía satélite o cualquier dispositivo inalámbrico que pueda comunicar información de vídeo a través de un canal de comunicación 15, en cuyo caso el canal de comunicación 15 es inalámbrico. Sin embargo, las técnicas de la presente divulgación, que se refieren a la selección eficiente de un algoritmo o modo de predicción que prediga las unidades codificadas, no se limitan necesariamente a aplicaciones o configuraciones inalámbricas.

[22] En el ejemplo de la FIG. 1, el dispositivo de origen 12 puede incluir una fuente de vídeo 20, un codificador de vídeo 22, un modulador/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 estar configurado para aplicar una o más de las técnicas de selección eficientes de esta divulgación como parte de un proceso de codificación de vídeo.

[23] El sistema 10 ilustrado de la FIG. 1, es simplemente a modo de ejemplo. Las técnicas de selección eficientes de esta divulgación pueden realizarse mediante cualquier dispositivo de codificación que soporte la predicción compensada por movimiento. El dispositivo de origen 12 y el dispositivo de destino 16 son simplemente ejemplos de dichos dispositivos de codificación en los que el dispositivo de origen 12 genera datos de vídeo codificados para su transmisión al dispositivo de destino 16. Los dispositivos 12, 16 pueden funcionar de manera sustancialmente simétrica, de manera que cada uno de los dispositivos 12, 16 incluya componentes de codificación y de decodificación de vídeo. Por lo tanto, el sistema 10 puede soportar una transmisión de vídeo unidireccional o bidireccional entre los dispositivos de vídeo 12, 16, por ejemplo, para la transmisión de vídeo, la reproducción de vídeo, la difusión de vídeo o la videotelefonía.

[24] La fuente de vídeo 20 del dispositivo de origen 12 incluye un dispositivo de captura de vídeo, tal como una cámara de vídeo, un archivo de vídeo que contenga vídeo capturado previamente o vídeo procedente de un proveedor de contenido de vídeo. Como alternativa adicional, la fuente de vídeo 20 puede generar datos, basados en gráficos de 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 denominados teléfonos con cámara o videoteléfonos. En cada caso, el vídeo capturado, precapturado o generado por ordenador puede codificarse mediante el codificador de vídeo 22. La información de vídeo codificada puede modularse entonces mediante un módem 23 de acuerdo con una norma de comunicación, por ejemplo, tal como el acceso múltiple por división de código (CDMA) u otra norma de comunicación, y transmitirse 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.

[25] El receptor 26 del dispositivo de destino 16 recibe información a través del canal 15, y el módem 27 demodula la información. De nuevo, el proceso de codificación de vídeo puede implementar una o más de las técnicas descritas en el presente documento para codificar eficientemente una unidad codificada durante la compensación por movimiento. La información comunicada a través del canal 15 puede incluir información definida por el codificador de vídeo 22, que también se usa por el decodificador de vídeo 28. El dispositivo de visualización 30 visualiza los datos de vídeo decodificados a un usuario y puede comprender cualquiera de una variedad de dispositivos de visualización, tales como un tubo de rayos catódicos, una pantalla de cristal líquido (LCD), una pantalla de plasma, una pantalla de diodos orgánicos emisores de luz (OLED) u otro tipo de dispositivo de visualización.

[26] En el ejemplo de la FIG. 1, el canal de comunicación 15 puede comprender cualquier medio de comunicación inalámbrico o alámbrico, tal como un espectro de radiofrecuencia (RF) o una o más líneas de transmisión física, o cualquier combinación de medios inalámbricos y alámbricos. El canal de comunicación 15 puede formar parte de una red basada en paquetes, tal como una red de área local (LAN), una red de área

extensa (WAN) o una red global, tal como Internet, que comprenda una interconexión de una o más redes. El canal de comunicación 15 representa en general cualquier medio de comunicación adecuado, o un conjunto de diferentes medios de comunicación, para transmitir datos de vídeo desde el dispositivo de origen 12 hasta el dispositivo de destino 16. El canal de comunicación 15 puede incluir routers, conmutadores, estaciones base o cualquier otro equipo que pueda ser útil para facilitar la comunicación desde el dispositivo de origen 12 al dispositivo de destino 16.

**[27]** El codificador de vídeo 22 y el decodificador de vídeo 28 pueden funcionar de acuerdo con una norma de compresión de vídeo, tal como la norma UIT-T H.264, descrita de forma alternativa como MPEG-4, Parte 10, Codificación de Vídeo Avanzada (AVC). Sin embargo, las técnicas de esta divulgación no están limitadas a ninguna norma de codificación particular. Aunque no se muestra en la FIG. 1, en algunos aspectos, el codificador de vídeo 22 y el decodificador de vídeo 28 pueden estar integrados cada uno de ellos con un codificador y un 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 separados. Si procede, las unidades MUX-DEMUX pueden ajustarse al protocolo de multiplexación UIT H.223 o a otros protocolos, tales como el protocolo de datagramas de usuario (UDP).

**[28]** La norma UIT-T H.264/MPEG-4 (AVC) se formuló por el Grupo de Expertos en Codificación de Vídeo de UIT-T (VCEG) junto al Grupo de Expertos en Imágenes en Movimiento de ISO/IEC (MPEG), como el producto de una asociación colectiva conocida como Equipo Mixto de Vídeo (JVT, por sus siglas en inglés). En algunos aspectos, las técnicas descritas en esta divulgación pueden aplicarse a dispositivos que se ajusten en general a la norma H.264. La norma H.264 se describe en la Recomendación UIT-T H.264, Codificación de Vídeo Avanzada para los servicios audiovisuales genéricos, por el Grupo de Estudio de la UIT-T, y con fecha de marzo de 2005, y que puede denominarse en el presente documento norma H.264 o memoria descriptiva H.264, o norma o memoria descriptiva H.264/AVC. El Equipo Mixto de Vídeo (JVT) continúa trabajando en ampliaciones a H.264/MPEG-4 AVC.

**[29]** Recientemente, el trabajo para avanzar en la norma H.264/MPEG-4 AVC ha comenzado en diversos foros de la UIT-T, tal como el foro Área de Tecnologías Clave (KTA, por sus siglas en inglés). El foro KTA busca, parcialmente, tecnologías de codificación que exhiban una mayor eficiencia de codificación que la que se exhibe por la norma H.264/AVC. Las técnicas descritas en esta divulgación pueden proporcionar una selección más eficiente de modos de predicción que la proporcionada en la norma H.264/AVC. Recientemente, el foro KTA recibió un documento que detalla estas técnicas en diversos aspectos como una presentación numerada VCEG-AJ25, titulada "Experimental Results on Simplified JMKA 2.0 Software JMKA 2.0", de Rahul Panchal y Marta Karczewicz (que se presentó en la 36ª reunión en San Diego, California, produciéndose del 8 al 10 de octubre de 2008).

**[30]** 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 cualquier combinación 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 (CÓDEC) en un respectivo dispositivo móvil, dispositivo de abonado, dispositivo de radiodifusión, servidor o similar.

**[31]** Una secuencia de vídeo incluye típicamente una serie de tramas de vídeo. Un codificador de vídeo 22 actúa en bloques de vídeo de tramas de vídeo individuales con el fin de codificar los datos de vídeo. Los bloques de vídeo pueden presentar tamaños fijos o variables y pueden diferir en tamaño de acuerdo con una norma de codificación especificada. Cada trama de vídeo incluye una serie de segmentos. Cada segmento puede incluir una serie de macrobloques, que pueden disponerse en subbloques. Como ejemplo, la norma UIT-T H.264 soporta la intrapredicción en diversos tamaños de bloque, tales como 16 por 16, 8 por 8 o 4 por 4 para componentes de luminancia, y 8x8 para componentes de croma, así como interpredicción en diversos tamaños de bloque, tales como 16 por 16, 16 por 8, 8 por 16, 8 por 8, 8 por 4, 4 por 8 y 4 por 4 para componentes de luminancia y tamaños escalados correspondientes para componentes de croma. Los bloques de vídeo pueden comprender bloques de datos de píxel, o bloques de coeficientes de transformada, por ejemplo, tras un proceso de transformada tal como de transformada de coseno discreta) o un proceso de transformada conceptualmente similar.

**[32]** Los bloques de vídeo más pequeños pueden proporcionar una mejor resolución y pueden usarse en ubicaciones de una trama de vídeo que incluyan altos niveles de detalle. En general, los macrobloques y los diversos subbloques pueden considerarse como bloques de vídeo. Además, un segmento puede considerarse una serie de bloques de vídeo, tales como macrobloques y/o subbloques. Cada segmento puede ser una unidad independientemente decodificable de una trama de vídeo. De forma alternativa, las propias tramas pueden ser unidades decodificables, o pueden definirse otras partes de una trama como unidades decodificables. El término "unidad codificada" se refiere a cualquier unidad de una trama de vídeo decodificable de manera independiente tal como una trama completa, un segmento de una trama, un grupo de imágenes

(GOP, por sus siglas en inglés) u otra unidad decodificable de manera independiente definida de acuerdo con las técnicas de codificación usadas.

**[33]** Tras la codificación predictiva interbasada (que incluye la interpolación y las técnicas de esta divulgación para seleccionar eficientemente un algoritmo o modo de predicción por el que predecir una unidad codificada) y tras cualquier transformada (como la transformada entera de 4x4 u 8x8 usada en H. 264/AVC o una transformada de coseno discreta o DCT, por sus siglas en inglés), puede realizarse la cuantificación. La cuantificación se refiere en general a un proceso en el que los coeficientes se cuantifican para reducir posiblemente la cantidad de datos usados para representar los coeficientes. El proceso de cuantificación puede reducir la profundidad de bits asociada con algunos o con la totalidad de los coeficientes. Por ejemplo, un valor de 16 bits puede redondearse a la baja hasta un valor de 15 bits durante la cuantificación. Tras la cuantificación, puede realizarse la codificación por entropía, por ejemplo de acuerdo con la codificación de longitud variable adaptativa al contenido (CAVLC), con la codificación aritmética binaria adaptativa al contexto (CABAC) o con otra metodología de codificación por entropía.

**[34]** De acuerdo con las técnicas de esta divulgación, el codificador de vídeo 22 puede seleccionar al menos dos unidades codificadas de referencia diferentes de los datos de vídeo. Estas dos unidades codificadas pueden comprender una primera trama y una segunda trama diferente de una primera trama. De forma alternativa, el codificador de vídeo 22 puede seleccionar una primera unidad de datos de vídeo de referencia, por ejemplo, un macrobloque o cualquier otra unidad de datos de vídeo de tamaño, de la primera trama de referencia y una segunda unidad de datos de vídeo de referencia de la segunda trama de referencia. Con fines de ilustración, las técnicas de la divulgación se describen a continuación con respecto a los bloques y a las tramas de vídeo. Sin embargo, las técnicas pueden aplicarse en general a unidades codificadas completas, o a partes de las mismas, tales como unidades de datos de vídeo de unidades codificadas.

**[35]** En algunos aspectos, el codificador de vídeo 22 puede almacenar un primer conjunto de tramas de referencia, que puede denominarse colectivamente lista cero (0), y un segundo conjunto de tramas de referencia, que puede denominarse colectivamente lista uno (1), donde cada uno de los conjuntos incluye tramas de referencia diferentes de las tramas de referencia del otro conjunto. El codificador de vídeo 22 puede seleccionar una o más tramas de referencia a partir de la lista 0 y una o más tramas de referencia a partir de la lista 1. A este respecto, el codificador de vídeo 22 puede recibir al menos dos tramas de referencia diferentes de los datos de vídeo.

**[36]** Después de seleccionar las tramas de referencia, el codificador de vídeo 22 puede realizar la predicción ponderada por defecto para predecir una primera versión de un bloque de vídeo predictivo de una trama predictiva de los datos de vídeo a partir de al menos dos bloques de vídeo de referencia, cada uno seleccionado a partir de uno diferente de al menos dos tramas de referencia diferentes. En particular, las dos tramas de referencia están temporalmente alejadas de la trama predictiva. Asimismo, los al menos dos bloques de vídeo de referencia están temporalmente alejados del bloque de vídeo predictivo. En otras palabras, las tramas de referencia (y los bloques de vídeo de referencia seleccionados a partir de estas tramas de referencia) del primer conjunto de tramas de referencia pueden comprender tramas de referencia que se produzcan en el tiempo secuencial o temporalmente antes o después o tanto antes como después de la trama predictiva. Asimismo, las tramas de referencia (y los bloques de vídeo de referencia seleccionados a partir de estas tramas de referencia) del segundo conjunto de tramas de referencia pueden comprender tramas que se produzcan en el tiempo secuencial o temporalmente antes o después o tanto antes como después de la trama predictiva.

**[37]** El codificador de vídeo 22 puede, por lo tanto, seleccionar una primera trama de referencia que se produzca temporalmente antes o después de la trama predictiva a partir de la lista 0 y una segunda trama de referencia que se produzca temporalmente antes o después de la trama predictiva a partir de la lista 1. Las tramas predictivas predichas a partir de estas dos unidades seleccionadas a menudo se denominan tramas bidireccionales o de imágenes (tramas B o imágenes B, para abreviar) ya que la trama predicha se predice a partir de ambas direcciones temporales, por ejemplo, ambas unidades de referencia que se producen antes y después de la trama predictiva. Aunque se las denomine "tramas B" por esta razón, las tramas B también pueden, en diversos aspectos, predecirse a partir de dos tramas de referencia que se produzcan antes de la trama B o, de forma alternativa, a partir de dos tramas de referencia que se produzcan después de la trama.

**[38]** Típicamente, una trama B se predice de esta manera bloque por bloque y el codificador de vídeo 22 puede seleccionar un primer bloque de vídeo de referencia desde la primera trama de referencia y un segundo bloque de vídeo desde la segunda trama de referencia. Para seleccionar estos bloques, el codificador de vídeo 22 puede identificar los primer y segundo bloques como aquellos que mejor coinciden con el bloque de vídeo predictivo o exhiben valores de píxel similares a los del bloque de vídeo predictivo. El codificador de vídeo 22 puede realizar entonces la predicción ponderada por defecto para predecir una primera versión de un bloque de vídeo predictivo del predictivo a partir de los primer y segundo bloques de vídeo de referencia de la primera y segunda tramas de referencia, respectivamente.

[39] Para realizar la predicción ponderada por defecto, el codificador de vídeo 22 puede multiplicar el primer bloque de vídeo de referencia por una primera ponderación para determinar un primer bloque de vídeo ponderado y el segundo bloque de vídeo de referencia por una segunda ponderación para determinar un segundo bloque de vídeo ponderado. El codificador de vídeo 22 puede añadir a continuación el primer bloque de vídeo ponderado al segundo bloque de vídeo ponderado para generar un bloque de vídeo total ponderado. El codificador de vídeo 22 puede predecir la primera versión del bloque de vídeo predictivo de acuerdo con el algoritmo de predicción ponderada por defecto dividiendo el bloque de vídeo total ponderado por el número de bloques de vídeo de referencia seleccionados para predecir la primera versión del bloque de vídeo predictivo, que es dos en este caso, por ejemplo, los primer y segundo bloques de vídeo de referencia. Típicamente, las primera y segunda ponderaciones son iguales entre sí, por ejemplo, la primera ponderación es igual a 0,5 o al 50 % y la segunda ponderación es igual a 0,5 o al 50 %. Para más de dos ponderaciones, las ponderaciones pueden ser solamente aproximadamente iguales entre sí, por ejemplo, una primera ponderación equivale al 33,33 %, una segunda ponderación equivale al 33,33 % y una tercera ponderación equivale al 33,34 %. Como las ponderaciones no varían habitualmente, esta forma de predicción ponderada puede denominarse predicción recta, por defecto o ponderada por igual.

[40] Después de calcular, determinar o predecir de otra forma la primera versión del bloque de vídeo predictivo usando el algoritmo de predicción ponderada por defecto, el codificador de vídeo 22 puede calcular un valor de compensación para el bloque de vídeo predictivo. Este valor de compensación puede comprender una compensación de CC o cualquier otro valor de compensación. En la compensación anterior, "CC" se refiere a una compensación de una corriente continua dada en el sentido eléctrico, pero desde entonces se ha adaptado para su uso por un gran número de contextos fuera del contexto eléctrico.

[41] En el contexto de la codificación de vídeo, puede calcularse una compensación de CC mediante un codificador de vídeo 22 que promedie en primer lugar un componente de luminancia de valores de píxel para uno de los bloques de vídeo de referencia. El codificador de vídeo 22 puede promediar a continuación un componente de luminancia de valores de píxel para el bloque de vídeo predictivo. Cada uno de estos valores calculados puede comprender un valor de CC. El codificador de vídeo 22 puede entonces calcular la compensación de CC sustrayendo los valores de CC entre sí. A este respecto, los valores de CC pueden comprender una luminancia promedio del bloque de vídeo de referencia y del bloque de vídeo predictivo, respectivamente. Este promedio es similar a la tensión de CC en el contexto eléctrico. La diferencia entre estos dos valores de CC promedio puede comprender una compensación a partir del valor de CC estándar, de forma muy similar a cómo una compensación de CC en el contexto eléctrico representa una compensación a partir del valor de CC estándar o promedio, de ahí el nombre "compensación de CC".

[42] En cualquier caso, el codificador de vídeo 22 puede realizar entonces una predicción ponderada implícita o una predicción ponderada explícita para predecir una segunda versión del bloque de vídeo predictivo de la trama predictiva. Para determinar cuál de las predicciones ponderadas implícitas o explícitas llevar a cabo, el codificador de vídeo 22 puede comparar la compensación de CC calculada con un valor de umbral y en base a esta comparación realizar una predicción ponderada implícita o explícita. Por ejemplo, el codificador de vídeo 22 puede comparar la compensación de CC calculada con un valor de umbral, que típicamente es cero (0). De forma alternativa, el codificador de vídeo 22 puede determinar en primer lugar un valor absoluto de la compensación de CC calculada y comparar este valor absoluto de la compensación de CC con el valor umbral, que de nuevo típicamente puede establecerse en cero (0). En base a la comparación, el codificador de vídeo 22 puede realizar entonces una predicción ponderada implícita o explícita para predecir la segunda versión del bloque de vídeo predictivo de la trama predictiva.

[43] La predicción ponderada "implícita" o "explícita" se refiere a una forma de predicción en la que las dos o más ponderaciones usadas para predecir la trama predictiva, por ejemplo, se codifican o no dentro de la unidad codificada. Para la predicción ponderada implícita, las ponderaciones no se codifican y, por lo tanto, pueden estar implícitas en el bloque de vídeo predictivo. El decodificador de vídeo 26 puede derivar entonces, al determinar que el bloque de vídeo predictivo se predice usando la predicción ponderada implícita, las ponderaciones a través del análisis de los primer y segundo bloques de vídeo de referencia o, en algunos casos, del análisis de las primera y segunda tramas de referencia. En la predicción ponderada explícita, el codificador de vídeo 22 puede codificar explícitamente las ponderaciones usadas en la predicción del bloque de vídeo predictivo dentro del bloque de vídeo predictivo, o en algunos casos, dentro de la trama predictiva. La predicción ponderada explícita puede proporcionar al codificador de vídeo 22 más latitud para seleccionar las ponderaciones y, como resultado, puede permitir que el codificador de vídeo 22 adapte la predicción de la trama predictiva, por ejemplo, una trama B, para adaptarse a un contexto o forma particular de datos de vídeo.

[44] Como ejemplo de realizar una predicción ponderada implícita o explícita en base a las ponderaciones, el codificador de vídeo 22 puede realizar una predicción ponderada explícita para predecir la segunda versión del bloque de vídeo predictivo cuando el valor absoluto del valor de compensación de CC calculado exceda el valor umbral. Sin embargo, cuando el valor absoluto del valor de compensación de CC calculado no exceda el valor umbral, es decir, sea más bajo o igual que el valor umbral, el codificador de vídeo 22 puede realizar una predicción ponderada implícita para predecir la segunda versión del bloque de vídeo predictivo. De esta



manera, el codificador de vídeo 22 puede realizar una predicción ponderada implícita o explícita, pero no ambas, para predecir la segunda versión del bloque de vídeo predictivo en base al valor de compensación calculado.

5 **[45]** Al predecir la segunda versión del bloque de vídeo predictivo, el codificador de vídeo 22 codifica el bloque de vídeo predictivo como la primera versión o la segunda versión. El codificador de vídeo 22 puede, en algunos casos, realizar un análisis de costes, tal como un análisis de costes de distorsión-velocidad (R-D, por sus siglas en inglés), de las primera y segunda versiones del bloque de vídeo predictivo para seleccionar cuál de las primera y segunda versiones codifica más apropiadamente una parte correspondiente de los datos de vídeo. De nuevo, el codificador de vídeo 22 codifica la primera versión del bloque de vídeo realizando la predicción ponderada por defecto y codifica la segunda versión del bloque de vídeo predictivo realizando la predicción ponderada explícita o implícita, pero no ambas.

15 **[46]** El codificador de vídeo 22 puede seleccionar una de las primera y segunda versiones que estén asociadas con el coste más bajo. En cualquier caso, el codificador de vídeo 22 codifica el bloque de vídeo predictivo como la primera versión o la segunda versión dentro del flujo de bits codificado. El transmisor 24 del dispositivo de origen 12 puede transmitir el flujo de bits codificado al receptor 26 del dispositivo de destino 16. En el dispositivo de destino 16, el decodificador de vídeo 28 recibe el flujo de bits codificado y decodifica el flujo de bits codificado para reconstruir la secuencia de vídeo.

20 **[47]** De esta manera, el codificador de vídeo 22 puede implementar las técnicas descritas en esta divulgación para seleccionar más eficientemente entre una pluralidad de algoritmos o modos de predicción sin generar realmente una versión del bloque de vídeo predictivo realizando cada uno de los algoritmos de predicción. Como se describió anteriormente, el codificador de vídeo 22 puede realizar en primer lugar una predicción ponderada por defecto y entonces realizar una predicción ponderada implícita o explícita, pero no ambas. Por tanto, en lugar de realizar las tres formas o tipos de predicción, el codificador de vídeo 22 realiza solamente dos de las tres predicciones ponderadas por defecto y explícitas o implícitas, pero no ambas. A este respecto, el codificador de vídeo 22 puede seleccionar de manera más eficiente entre tres o más modos de predicción sin tener que realizar todos y cada uno de los tres o más modos de predicción.

25 **[48]** El codificador de vídeo 22 puede, de nuevo como se describió anteriormente, seleccionar entre realizar la predicción ponderada implícita o explícita para predecir una segunda versión del bloque de vídeo predictivo de la trama B en base a una compensación calculada para la primera versión del bloque de vídeo predictivo de una trama B. Aunque se describe en esta divulgación con respecto a una compensación, cualquier característica comúnmente accesible o determinable por el codificador de vídeo 22 puede usarse por el codificador de vídeo 22 como base para la selección entre realizar una predicción ponderada implícita o explícita.

30 **[49]** La FIG. 2 es un diagrama de bloques que ilustra un ejemplo de un codificador de vídeo 50 que puede realizar las técnicas de compensación coherentes 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. El codificador de vídeo 50 puede realizar la intra e intercodificación de bloques dentro de tramas de vídeo, aunque los componentes de intracodificación no se muestran en la FIG. 2 para facilitar la ilustración. La intracodificación se basa en la predicción espacial para reducir o eliminar la redundancia espacial en el vídeo en una trama de vídeo dada. La intercodificación se basa en la predicción temporal para reducir o eliminar la redundancia temporal en el vídeo dentro de tramas adyacentes de una secuencia de vídeo. El intramodo (modo I) puede referirse al modo de compresión con base espacial, y los intermodos, tales como de predicción (modo P) o bidireccional (modo B), pueden referirse a los modos de compresión con base temporal. Las técnicas de esta divulgación se aplican durante la intercodificación y, por lo tanto, las unidades de intracodificación tales como la unidad de predicción espacial no se ilustran en la FIG. 2 con fines de simplicidad de y para facilitar la ilustración.

35 **[50]** Como se muestra en la FIG. 2, el codificador de vídeo 50 recibe un bloque de vídeo actual dentro de una trama de vídeo que vaya a codificarse. En el ejemplo de la FIG. 2, el codificador de vídeo 50 incluye una unidad de estimación de modo 32, una unidad de compensación por movimiento 35, un almacén de tramas de referencia 34, un sumador 48, una unidad de transformada 38, una unidad de cuantificación 40 y una unidad de codificación por entropía 46. Para la reconstrucción de bloques de vídeo, el codificador de vídeo 50 incluye también una unidad de cuantificación inversa 42, una unidad de transformada inversa 44 y un sumador 51. El decodificador de vídeo 50 puede incluir también un filtro de desbloqueo (no mostrado) para filtrar los límites de bloque para eliminar distorsiones de efecto pixelado del vídeo reconstruido. Si se desea, el filtro de desbloqueo filtraría típicamente la salida del sumador 51.

40 **[51]** Durante el proceso de codificación, el codificador de vídeo 50 recibe un bloque de vídeo para codificarse, y la unidad de estimación de movimiento 32 y la unidad de compensación por movimiento 35 realizan una codificación interpredictiva. La unidad de estimación de movimiento 32 y la unidad de compensación por movimiento 35 pueden estar altamente integradas, pero se ilustran por separado con fines

conceptuales. Se considera típicamente que la estimación de movimiento es el proceso de generar vectores de movimiento, que estiman un movimiento para bloques de vídeo. Un vector de movimiento, por ejemplo, puede indicar el desplazamiento de un bloque predictivo dentro de una trama predictiva (u otra unidad codificada) con relación al bloque actual que esté codificándose en la trama actual (u otra unidad codificada). La compensación por movimiento se considera típicamente el proceso de extraer o generar el bloque predictivo en base al vector de movimiento determinado por la estimación de movimiento. De nuevo, la unidad de estimación de movimiento 32 y la unidad de compensación por movimiento 35 pueden integrarse funcionalmente. Con fines demostrativos, las técnicas descritas en esta divulgación se describen como realizadas por la unidad de compensación por movimiento 35.

**[52]** La unidad de estimación de movimiento 32 selecciona el vector de movimiento apropiado para que se codifique el bloque de vídeo comparando el bloque de vídeo con los bloques de vídeo de una o más tramas de referencia (por ejemplo, una trama previa y/o futura en términos de tiempo o temporalmente). La unidad de estimación de movimiento 32 puede, como ejemplo, seleccionar un vector de movimiento para una trama B de varias formas. De una forma, la unidad de estimación de movimiento 32 puede seleccionar una trama previa o futura de un primer conjunto de tramas (denominado lista 0) y determinar un vector de movimiento usando solamente esta trama previa o futura a partir de la lista 0. De forma alternativa, la unidad de estimación de movimiento 32 puede seleccionar una trama previa o futura de un segundo conjunto de tramas (denominado lista 1) y determinar un vector de movimiento usando solamente esta trama previa o futura a partir de la lista 1. De otra forma más, la unidad de estimación de movimiento 32 puede seleccionar una primera trama a partir de la lista 0 y una segunda trama a partir de la lista 1 y seleccionar uno o más vectores de movimiento de la primera trama de la lista 0 y de la segunda trama de la lista 1. Esta forma de predicción puede, como se mencionó anteriormente, denominarse estimación de movimiento bipredictivo. Las técnicas de esta divulgación pueden implementarse para seleccionar eficientemente un modo de bipredicción compensada por movimiento.

**[53]** De acuerdo con la norma H.264/AVC, pueden usarse tres algoritmos o modos bipredictivos compensados por movimiento para predecir una trama B o partes de la misma, tales como bloques de vídeo, macrobloques o cualquier otra parte discreta y/o contigua de una trama B. Un primer algoritmo o modo bipredictivo compensado por movimiento, que comúnmente se denomina predicción ponderada por defecto, puede implicar la aplicación de ponderaciones aproximadamente iguales a cada bloque de vídeo identificado de la primera trama de la lista 0 y de la segunda trama de la lista 1. Los bloques ponderados de las primera y segunda tramas se añaden y se dividen entonces por el número total de tramas usadas para predecir la trama B, por ejemplo, dos en este caso. A menudo, esta división se logra sumando 1 a la adición de los bloques ponderados de las primera y segunda tramas y moviendo entonces el resultado a la derecha en un bit.

**[54]** El segundo algoritmo o modo bipredictivo compensado por movimiento, que comúnmente se denomina predicción ponderada explícita, puede implicar la determinación de ponderaciones para cada bloque de vídeo identificado de las primera y segunda tramas y entonces realizar una multiplicación similar de las ponderaciones para generar los bloques ponderados y la adición a lo que se ha descrito anteriormente con respecto al algoritmo de predicción ponderada por defecto. Con respecto a la predicción ponderada explícita, sin embargo, pueden añadirse uno o más valores o compensaciones en redondeo antes de mover la suma de los primer y segundo bloques ponderados a la derecha por uno u otro número para garantizar una división apropiada por el número total de tramas usadas para predecir la tercera unidad codificada o trama B.

**[55]** El tercer algoritmo o modo bipredictivo compensado por movimiento, que comúnmente se denomina predicción ponderada implícita, puede implicar la determinación de ponderaciones para cada bloque identificado de acuerdo con un algoritmo establecido, que puede, por ejemplo, analizar cada uno de los primer y segundo bloques para determinar las primera y segunda ponderaciones. Una vez determinado, este algoritmo puede aplicar o, de otro modo, multiplicar los primer y segundo bloques por las primera y segunda ponderaciones determinadas respectivamente para generar los primer y segundo bloques ponderados. Después de determinar los bloques ponderados, el algoritmo puede añadir a continuación los bloques ponderados junto con un valor o una compensación en redondeo antes de mover la suma por un valor que garantice una división por el número total de tramas usadas para predecir la trama B.

**[56]** La unidad de compensación por movimiento 35 puede implementar cada uno de estos algoritmos bipredictivos compensados por movimiento para generar datos predictivos representativos de un bloque bipredictivo de una trama B. Por ejemplo, la unidad de compensación por movimiento 32 puede implementar la predicción ponderada por defecto de acuerdo con la siguiente ecuación (1):

$$\text{pred}(i,j) = (\text{pred}0(i,j) + \text{pred}1(i,j)+1) \gg 1 \quad (1).$$

Pred (i, j) se refiere a datos predictivos representativos de un bloque de vídeo en la i-ésima fila y en la j-ésima columna de la trama predictiva, por ejemplo, la trama B. Pred0 (i, j) se refiere a datos representativos de un bloque de vídeo en la i-ésima fila y en j-ésima columna de una primera trama de referencia, donde la primera trama de referencia se selecciona a partir de la lista 0. Pred1 (i, j) se refiere a los datos representativos de un bloque de vídeo en la i-ésima fila y en la j-ésima columna de una segunda trama de referencia, donde la

primera trama de referencia se selecciona a partir de la lista 1. Como las ponderaciones son iguales o aproximadamente iguales, la adición de uno en la ecuación (1) y en el movimiento (>>) hacia la derecha en un bit divide efectivamente la suma de pred0 (i, j) + pred1 (i, j) entre dos, por ejemplo, la cantidad total de tramas o bloques de vídeo usados para predecir el bloque de vídeo predictivo de la trama B.

[57] La unidad de compensación por movimiento 35 puede implementar el segundo algoritmo bipredictivo compensado por movimiento o la predicción ponderada explícita de acuerdo con la siguiente ecuación (2):

$$\text{pred}(i,j)=(\text{pred0}(i,j)*w_0+\text{pred1}(i,j)*w_1+2^r)\gg(r+1)+((o_1+o_2+1)\gg1) \quad (2).$$

[58] De nuevo, cada uno de pred (i, j), pred0 (i, j) y pred1 (i, j) se refieren a los mismos bloques de vídeo de referencia descritos anteriormente con respecto a la ecuación (1). Como las ponderaciones están determinadas y pueden no ser aproximadamente iguales, la ecuación (2) incluye las ponderaciones w0 y w1, que se aplican de manera multiplicativa a los respectivos pred0 (i, j) y pred1 (i, j). La variable "r" puede representar un número entero seleccionado para garantizar que las ponderaciones w0 y w1 den como resultado un número entero. Las variables o1 y o2 representan cada una una compensación en redondeo, donde la ecuación (2) proporciona un promedio de las compensaciones en redondeo o1 y o2 a través de la adición de las compensaciones en redondeo y una, seguida de un movimiento hacia la derecha de uno (1). La suma ponderada de los bloques también puede moverse antes de la adición del promedio de las compensaciones en redondeo para garantizar una división apropiada por el número total de bloques de vídeo de referencia usados para predecir el bloque de vídeo predictivo de la trama B.

[59] La unidad de compensación por movimiento 35 puede implementar el tercer algoritmo bipredictivo compensado por movimiento o la predicción ponderada implícita de acuerdo con la siguiente ecuación (3):

$$\text{pred}(i,j) = (\text{pred0}(i,j)*w_0+\text{pred1}(i,j)*w_1+32)\gg6 \quad (3).$$

[60] De nuevo, cada uno de pred (i, j), pred0 (i, j) y pred1 (i, j) se refieren a los mismos bloques de vídeo de referencia descritos anteriormente con respecto a las ecuaciones (1) y (2). Como las ponderaciones están determinadas y pueden no ser aproximadamente iguales, la ecuación (3) incluye las ponderaciones w0 y w1, que se aplican de manera multiplicativa a los respectivos pred0 (i, j) y pred1 (i, j). El valor "32" puede comprender una compensación en redondeo estática usada para garantizar que las ponderaciones w0 y w1 son números enteros y el movimiento hacia la derecha (>>) por seis (6) puede representar una división, dada la compensación en redondeo de 32, que efectivamente divide la suma de los bloques ponderados totales por el número total de tramas usadas para predecir el bloque de vídeo predictivo de la trama B.

[61] De acuerdo con las técnicas de esta divulgación, la unidad de estimación de movimiento 32 puede seleccionar al menos dos tramas para determinar vectores de movimiento para una trama B. Particularmente, como se describió anteriormente, el almacén de tramas de referencia 34 puede comprender una memoria para almacenar un primer conjunto de tramas de referencia de los datos de vídeo denominado lista 0 y un segundo conjunto de tramas de referencia denominado lista 1. Las tramas de referencia de la lista 1 y de la lista 0 pueden comprender cada una tramas I o tramas P. La unidad de estimación de movimiento 32 puede acceder al almacén de tramas de referencia 34 y seleccionar una o más tramas de referencia a partir de la lista 0 y una o más tramas de referencia a partir de la lista 1. De esta manera, la unidad de estimación de movimiento 32 puede, en un aspecto, seleccionar al menos dos unidades codificadas, por ejemplo, tramas, de los datos de vídeo.

[62] La unidad de estimación de movimiento 32 puede entonces determinar vectores de movimiento para un primer bloque de una trama B predictiva. La unidad de estimación de movimiento 32 puede identificar un correspondiente primer bloque en una primera de las al menos dos tramas de referencia que corresponda al bloque predictivo de la trama B y un segundo bloque correspondiente en una segunda de las al menos dos tramas de referencia que correspondan al bloque predictivo de la trama B. Típicamente, la primera trama y la segunda trama están respectivamente temporalmente alejadas de la tercera trama B. A menudo, la primera trama de referencia reside en la secuencia de vídeo previamente a o antes de la trama B, mientras que la segunda trama de referencia reside en la secuencia de vídeo después de la trama B. Sin embargo, en algunos casos, tanto la primera como la segunda trama de referencia pueden residir o producirse en diferentes momentos antes o después de la trama B. En cualquier caso, la unidad de estimación de movimiento 32 puede calcular entonces los vectores de movimiento en base al primer bloque de referencia de la primera trama de referencia y al segundo bloque de referencia a partir de la segunda trama de referencia, ambos de los cuales están determinados en relación con el tercer bloque predictivo de la trama B predictiva.

[63] Una vez que la unidad de estimación de movimiento 32 ha seleccionado los vectores de movimiento para el bloque de vídeo que vaya a codificarse, la unidad de compensación por movimiento 35 genera el bloque de vídeo predictivo asociado con estos vectores de movimiento. La unidad de compensación por movimiento 35 puede generar una primera versión del bloque de vídeo predictivo de la trama B de acuerdo con el algoritmo predictivo ponderado por defecto, como se representa por la ecuación (1) anterior. Para hacerlo, la

unidad de compensación por movimiento 35 puede recuperar los diversos bloques de referencia especificados por la ecuación (1) del almacén de tramas de referencia 34 e implementar la ecuación (1) para generar datos predictivos representativos del bloque de vídeo predictivo de la trama B. A continuación, la unidad de compensación por movimiento 35 puede calcular el valor de compensación de CC anterior u otro valor de compensación para la primera versión del bloque de vídeo predictivo de la trama B.

**[64]** En un ejemplo, la unidad de compensación por movimiento 35 puede calcular un primer conjunto de métricas (por ejemplo, valores medios) asociado respectivamente con cada ubicación de píxel entero y subentero para los bloques de vídeo de referencia de las primera y segunda tramas de referencia que se usen para predecir el bloque de vídeo predictivo. La unidad de compensación por movimiento 35 también puede calcular un segundo conjunto de métricas (por ejemplo, valores medios o una diferencia entre la suma de valores de luminancia y/o valores de crominancia) asociados respectivamente con cada ubicación de píxel entero y subentero para el bloque de vídeo predictivo. La unidad de compensación por movimiento 35 calcula entonces los valores de compensación en base a los primer y segundo conjunto de métricas. La unidad de compensación por movimiento 35 puede, por ejemplo, calcular una primera compensación de CC como la diferencia entre los valores medios calculados para el primer bloque de vídeo de referencia y el bloque de vídeo predictivo. La unidad de compensación por movimiento 35 puede calcular a continuación una segunda compensación de CC como la diferencia entre los valores medios calculados para el segundo bloque de vídeo de referencia y el bloque de vídeo predictivo. La unidad de compensación por movimiento 35 puede calcular entonces un promedio de estas primera y segunda compensaciones de CC para generar una compensación de CC promedio. Los valores de compensación pueden comprender valores absolutos o valores firmados que puedan reflejar una polarización hacia arriba o hacia abajo de los valores de píxel de un bloque de vídeo correspondiente, que de nuevo pueden ser muy útiles para indicar cambios de escena o destellos encontrados en la codificación de vídeo.

**[65]** En otras palabras, el primer conjunto de métricas puede comprender un conjunto de valores medios que correspondan a la media de los valores de píxel en cada ubicación de píxeles enteros y subenteros de los bloques de vídeo de una unidad codificada dada. El segundo conjunto de métricas puede comprender un conjunto de valores medios que correspondan a la media de los valores de píxel en cada ubicación de píxeles enteros y subenteros de los bloques predictivos usados para predecir los bloques actuales que estén codificándose en esa unidad codificada. La pluralidad de valores de compensación puede comprender diferencias entre el primer conjunto de valores medios y el segundo conjunto de valores medios. Cada ubicación de macrobloque puede definirse por un único píxel, por ejemplo, el píxel respectivo en la esquina superior izquierda del macrobloque. Sin embargo, cada macrobloque puede definir dieciséis valores de píxeles que contribuyan a un valor medio particular en el primer conjunto de valores medios. Por supuesto, estos valores de compensación también pueden calcularse para bloques de vídeo de otro tamaño.

**[66]** En general, la compensación para cualquier posición dada puede calcularse como la diferencia entre la media de todos los píxeles en la trama actual (u otra unidad codificada) que tengan una precisión de vector de movimiento correspondiente a esa posición de píxel o subpíxel y la media de los valores interpolados de datos predictivos correspondientes a esa posición de píxel o posición subpíxel. Por tanto, cada compensación respectiva puede verse como la diferencia promedio de píxeles de la unidad codificada con respecto a los datos predictivos para cada posición entera, interpolada o extrapolada respectiva de los datos.

**[67]** La unidad de compensación por movimiento 35 puede calcular los valores de compensación con respecto a bloques de luminancia, bloques de croma o ambos. Pueden definirse diferentes compensaciones para cada ubicación de píxeles enteros y subenteros asociados con cada tipo de bloque de vídeo (por ejemplo, bloques de luminancia y croma). Además, podrían asignarse diferentes compensaciones a cada bloque en cada tamaño particular, partición o subpartición de cada bloque.

**[68]** Después de calcular los valores de compensación, la unidad de compensación por movimiento 35 puede comparar cada uno de los valores de compensación individuales o el valor de compensación promedio con un valor umbral. El valor umbral puede establecerse de manera programática por un usuario del codificador de vídeo 50 o establecerse de manera estática por un diseñador de hardware del codificador de vídeo 50. En algunos casos, la unidad de compensación por movimiento 35 puede establecer automáticamente, por ejemplo, sin requerir ninguna supervisión o entrada del usuario, el valor umbral en base a algún análisis de los diversos bloques de una o más de las primera, segunda y tercera unidades codificadas. A este respecto, la unidad de compensación por movimiento 35 puede adaptar o adaptar automáticamente el valor umbral al realizar la comparación. Típicamente, sin embargo, el diseñador o el usuario establece de manera programática el valor umbral en cero (0). En base a esta comparación, la unidad de estimación de movimiento 35 puede realizar una predicción ponderada explícita o implícita de acuerdo con, por ejemplo, cualquiera de las ecuaciones (2) o (3), respectivamente, para predecir una segunda versión del tercer bloque de la trama B.

**[69]** Por ejemplo, la unidad de compensación por movimiento 35 puede realizar una predicción ponderada explícita de acuerdo con la ecuación (2) para predecir una segunda versión del bloque predictivo cuando el

valor de compensación calculado o, en algunos casos, un valor absoluto del valor de compensación calculado exceda el valor umbral. La unidad de compensación por movimiento 35 puede usar las ponderaciones por defecto, por ejemplo, las ponderaciones usadas durante la predicción ponderada por defecto, para que las ponderaciones  $w_0$  y  $w_1$  predigan el tercer o el bloque predictivo de la trama B. Además, cuando se realice una predicción ponderada explícita, la unidad de compensación por movimiento 35 puede, en algunos aspectos, usar las compensaciones calculadas. [

**[70]** Cuando los valores de compensación calculados o, en algunos casos, el valor absoluto de los valores de compensación calculados, no excedan, es decir, sean menores o iguales que el valor umbral, la unidad de compensación por movimiento 35 puede realizar una predicción ponderada implícita de acuerdo con la ecuación (3) para predecir la segunda versión del bloque de vídeo predictivo de la trama B. Como se describió anteriormente con respecto a la predicción ponderada por defecto, la unidad de compensación por movimiento 35 puede recuperar los diversos bloques de referencia especificados por la ecuación (2) o (3) a partir del almacén de tramas de referencia 34. De forma alternativa, la unidad de compensación por movimiento 35 puede almacenar en una memoria caché o mantener de otra forma o almacenar los diversos bloques de vídeo de referencia de manera local después de recuperar en primer lugar estos bloques para calcular la primera versión del bloque de vídeo predictivo de acuerdo con el algoritmo de predicción ponderada por defecto, por ejemplo, como se representa por la ecuación (1) anteriormente.

**[71]** Independientemente de si la segunda versión del bloque de vídeo predictivo de la trama B se predice de acuerdo con la predicción ponderada explícita o implícita, la unidad de compensación por movimiento 35 compara la primera versión del bloque de vídeo predictivo predicho al realizar la predicción ponderada por defecto con la segunda versión del bloque de vídeo predictivo predicho al realizar una predicción ponderada explícita o implícita entre sí y selecciona la versión que representa de manera más apropiada la parte correspondiente del vídeo. La unidad de compensación por movimiento 35 puede, para realizar esta comparación, realizar un análisis de distorsión-velocidad (R-D) en las primera y segunda versiones. Como se describió anteriormente, el análisis R-D puede implicar el cálculo de un coste para cada una de las versiones y la selección de una de las primera y segunda versiones para las que se calculó un coste más bajo. Este análisis R-D estima en general la velocidad y la distorsión o, mejor dicho, equilibra la cantidad de datos usados en la codificación de las primera y segunda versiones de los bloques de vídeo predictivos frente a algún nivel cuantificado de calidad, por ejemplo, un nivel o cantidad de distorsión.

**[72]** Como ejemplo, la unidad de compensación por movimiento 35 puede basar el análisis R-D en un coste calculado para cada una de las primera y segunda versiones de acuerdo con la siguiente ecuación (4) que proporciona una función de coste lagrangiana:

$$F_c = d + (\lambda^0 * R) \quad (4).$$

$F_c$  comprende una variable representativa de la función de coste. La letra "d" comprende una variable representativa del primer o del segundo valor de distorsión calculado comparando las primera y la segunda versiones, respectivamente, con la parte correspondiente del vídeo. La unidad de compensación por movimiento 35 calcula o mide típicamente las primera y segunda distorsiones como un error cuadrático medio para maximizar una relación pico de señal-ruido (PSNR).  $\lambda^0$  o " $\lambda^0$ " comprende una variable representativa de un multiplicador lagrangiano, que es un valor que representa una relación entre el coste del bit y la calidad para un nivel de calidad particular. La letra "R" comprende una variable representativa de una tasa de bits a la que se codifica el vídeo.

**[73]** La unidad de compensación por movimiento 35 puede, por lo tanto, calcular un primer y un segundo valor de distorsión para las respectivas primera y segunda versiones del bloque de vídeo predictivo y almacenar datos que definan valores tanto para el multiplicador lagrangiano como para la tasa de bits. Para cada una de las primera y segunda versiones, la unidad de compensación por movimiento 35 puede determinar un coste de acuerdo con la ecuación (4) y comparar estos primer y segundo costes entre sí. La unidad de compensación por movimiento 35 selecciona entonces una de las primera y segunda versiones en base a la comparación de los primer y segundo costes. En algunos aspectos, la unidad de compensación por movimiento 35 selecciona una de las primera y segunda versiones para las que determinó el coste más bajo.

**[74]** Si la unidad de compensación por movimiento 35 selecciona la segunda versión predicha por medio del algoritmo de predicción ponderada explícita, la unidad de compensación por movimiento 35 puede aplicar los valores de compensación de CC a bloques de vídeo predictivos originales para generar bloques de vídeo predictivos compensados y codificar bloques de vídeo de la unidad codificada en los bloques de vídeo predictivos de compensación. Al añadir la compensación a los valores de píxel de los bloques predictivos de una manera basada en la ubicación de acuerdo con las ubicaciones de píxeles de los bloques predictivos (ubicación de enteros o una de una pluralidad de posibles ubicaciones de píxeles subenteros), los bloques predictivos pueden ser más similares a los bloques codificados, lo que puede mejorar la eficiencia de la codificación. Además, puesto que los valores de compensación se definen de manera diferente para diferentes ubicaciones de píxeles, estos valores de compensación pueden proporcionar una capacidad para lograr la

segmentación entre la codificación de datos asociados con diferentes tipos de interpolación. Si, sin embargo, la unidad de compensación por movimiento 35 selecciona la primera versión o la segunda versión predicha por medio de la predicción ponderada implícita, la unidad de compensación por movimiento 35 puede codificar los bloques sin añadir los valores de compensación calculados. De esta manera, la unidad de compensación por movimiento 35 puede generar datos predictivos representativos de un tercer bloque de vídeo predictivo de una trama B.

**[75]** El codificador de vídeo 50 forma entonces un bloque de vídeo residual sustrayendo los datos de predicción del bloque de vídeo original que esté codificándose. El sumador 48 representa el componente o los componentes que realizan esta operación de sustracción. La unidad de transformada 38 aplica una transformada, tal como una transformada discreta de coseno (DCT) o una transformada conceptualmente similar, al bloque residual, produciendo un bloque de vídeo que comprende coeficientes de bloque de transformada residual. La unidad de transformada 38 puede realizar, por ejemplo, otras transformadas, tales como las definidas por la norma H.264, que sean conceptualmente similares a la DCT. También podrían usarse transformadas de ondículas, transformadas de enteros, transformadas de subbandas u otros tipos de transformadas. En cualquier caso, la unidad de transformada 38 aplica la transformada al bloque residual, produciendo un bloque de coeficientes de transformada residuales. La transformada puede convertir la información residual de un dominio de píxel en un dominio de frecuencia.

**[76]** La unidad de cuantificación 40 cuantifica los coeficientes de transformada residuales para reducir más la tasa de bits. El proceso de cuantificación puede reducir la profundidad de bits asociada a algunos o con la totalidad de los coeficientes. Por ejemplo, un valor de 16 bits puede redondearse a la baja hasta un valor de 15 bits durante la cuantificación. Además, la unidad de cuantificación 40 puede cuantificar también las diferentes compensaciones en el caso en el que la segunda versión se prediga por medio de una predicción ponderada explícita para asignar un número deseado de bits al número entero respectivo y partes fraccionarias de las diferentes compensaciones. En este caso, la unidad de cuantificación puede asignar, para cada uno de los valores de compensación, un primer número de bits a una parte entera de un valor de compensación dado y asignar un segundo número de bits a una parte fraccionada del valor de compensación dado, en donde los primer y segundo números de bits se determinan en base a una magnitud de la parte entera. La unidad de codificación por entropía 46 puede codificar el primer número de bits de manera diferente que el segundo número de bits.

**[77]** Tras la cuantificación, la unidad de codificación por entropía 46 realiza la codificación por entropía de los coeficientes de transformada cuantificados. Por ejemplo, la unidad de codificación por entropía 46 puede realizar la codificación de longitud variable adaptativa al contenido (CAVLC), la codificación aritmética binaria adaptativa al contexto (CABAC) u otra técnica de codificación por entropía. 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 flujo de bits codificado puede incluir bloques residuales codificados por entropía, vectores de movimiento para dichos bloques y otra sintaxis que incluya los valores de compensación que identifiquen la pluralidad de compensación diferentes en diferentes ubicaciones de píxeles enteros y de píxeles subenteros dentro de la unidad codificada.

**[78]** La unidad de cuantificación inversa 42 y la unidad de transformada inversa 44 aplican la cuantificación inversa y la transformada inversa, respectivamente, para reconstruir el bloque residual en el dominio de píxeles, por ejemplo, para su uso posterior como bloque de referencia de la manera descrita anteriormente. El sumador 51 añade el bloque residual reconstruido al bloque de predicción compensado por movimiento producido por la unidad de compensación por movimiento 35 para producir un bloque de vídeo reconstruido para su almacenamiento en el almacén de tramas de referencia 34. El bloque de vídeo reconstruido puede usarse por la unidad de estimación de movimiento 32 y por la unidad de compensación por movimiento 35 como bloque de referencia para intercodificar un bloque en una trama de vídeo posterior.

**[79]** De esta manera, el codificador de vídeo 50 puede implementar las técnicas de esta divulgación para seleccionar eficientemente entre tres algoritmos bipredictivos compensados por movimiento. La selección es eficiente de manera que la unidad de compensación por movimiento 35 solamente realiza dos de los tres algoritmos bipredictivos compensados por movimiento en lugar de realizar los tres y selecciona entonces una de las tres versiones resultantes que representan de manera más apropiada los datos de vídeo. Las técnicas, por lo tanto, no solamente facilitan una selección más eficiente al eliminar el rendimiento de al menos uno de los algoritmos, sino que también eliminan un cálculo adicional de la función de costes del análisis R-D ya que solamente las primera y segunda versiones son el tema de este análisis. A este respecto, las técnicas pueden implementarse mediante el codificador de vídeo 50 para seleccionar de manera más eficiente entre los tres algoritmos bipredictivos compensados por movimiento.

**[80]** La FIG. 3 es un diagrama de bloques que ilustra un ejemplo de unidad de compensación por movimiento 35 de la FIG. 2 en más detalle. Como se muestra en el ejemplo de la FIG. 3, la unidad de compensación por movimiento 35 se acopla al almacén de tramas de referencia 34, que almacena los primer y segundo conjuntos de unidades codificadas o tramas de referencia descritos anteriormente como lista 052A y

lista 1 52B. La unidad de compensación por movimiento 35 puede recuperar en primer lugar al menos dos tramas de referencia del almacén de tramas de referencia 34. Típicamente, la unidad de compensación por movimiento 35 recupera al menos una trama de la lista 0 52A y al menos una trama de la lista 1 52A. Estas tramas de las listas 0 y 1 52A y 52B pueden referirse en esta divulgación como una primera trama de referencia y una segunda trama de referencia, respectivamente. A menudo, la unidad de compensación por movimiento 35 recupera las primera y segunda tramas de referencia indicadas por la unidad de estimación de movimiento 32.

**[81]** Como se muestra además en la FIG. 3, la unidad de compensación por movimiento 54 incluye un módulo de predicción por defecto 54, un módulo de cálculo de compensación 56, un comparador 58, un módulo de predicción explícita 60, un módulo de predicción implícita 62 y un módulo de análisis (R-D) de distorsión-velocidad 64 ("Módulo de análisis R-D 64"). El módulo de predicción por defecto 54 puede representar un módulo de hardware y/o software que implemente el algoritmo de predicción ponderada por defecto descrita anteriormente con respecto a la ecuación (1). El módulo de cálculo de compensación 56 puede representar un módulo de hardware y/o software que calcule valores de compensación, tales como valores de compensación de CC, de la manera descrita anteriormente.

**[82]** El comparador 58 puede representar un módulo de hardware y/o software que compara los valores de compensación con un valor umbral, que se muestra en la FIG. 3 como umbral 66. El módulo de predicción explícita 60 puede representar un módulo de hardware y/o software que implemente el algoritmo de predicción ponderada explícita descrita anteriormente con respecto a la ecuación (2). El módulo de predicción implícita 62 puede representar un módulo de hardware y/o software que implemente el algoritmo de predicción ponderada implícita descrita anteriormente con respecto a la ecuación (3). El módulo de análisis R-D 64 puede representar un módulo de hardware y/o software que implemente el análisis R-D descrito anteriormente. El módulo de análisis R-D 64 puede implementar una función de coste 68, tal como la representada por la ecuación (4), y emplear la función de coste 68 para calcular los costes 70A y 70B ("costes 70"). El análisis R-D puede basar el análisis en estos costes 70.

**[83]** Como se muestra en la FIG. 3 que comprende diversos módulos 54-64, estos módulos 54-64 se proporcionan con fines demostrativos. Uno o más de los módulos 54-64 pueden implementarse como un módulo integrado dentro de la unidad de compensación por movimiento 35. De forma alternativa, la unidad de compensación por movimiento 35 puede comprender en general un programa de software o informático que se ejecute en un procesador. Los diversos módulos, en este caso, pueden representar módulos de software o componentes del programa de software más grande. El programa de software puede comprender una pluralidad de instrucciones ejecutables por el procesador y que hagan que el procesador realice las técnicas descritas en esta divulgación.

**[84]** En cualquier caso, después de recuperar o recibir las primera y segunda tramas de referencia, la unidad de compensación por movimiento 35 puede invocar o por el contrario hacer que el módulo de predicción por defecto 54 genere datos predictivos representativos de una primera versión de un bloque de vídeo predictivo de una trama B (o más generalmente, de una trama predictiva), de acuerdo con la ecuación (1) anterior. El módulo de predicción por defecto 54 puede emitir esta primera versión para compensar el módulo de cálculo 56, que puede calcular los valores de compensación de la manera descrita anteriormente. Es decir, el módulo de cálculo de compensación 56 puede calcular los valores de compensación en base a una comparación entre la primera versión y los datos de vídeo originales o el bloque correspondiente dentro de las primera y segunda tramas de referencia. En algunos casos, el módulo de cálculo de compensación 56 puede determinar un valor absoluto de los valores de compensación y emitir estos valores de compensación en esta forma de valor absoluto. El módulo de cálculo de compensación 56 puede enviar estos valores de compensación al comparador 58, que puede comparar los valores de compensación con el valor umbral 66. El valor umbral 66 puede configurarse de manera programática, establecerse automáticamente o establecerse de manera estática. Típicamente, el valor umbral 66 se establece en un valor de cero (0).

**[85]** En base a la comparación, el comparador 58 puede emitir un código de activación tanto para el módulo de predicción explícita 60 como para el módulo de predicción implícita 62. El código de activación puede activar uno u otro, pero no ambos, del módulo de predicción explícita 60 y del módulo de predicción implícita 62. Como se describe en el ejemplo anterior, cuando uno o más de los valores de compensación excedan el valor umbral 66, el comparador 58 puede generar y transmitir un código de activación que active el módulo de predicción explícita 60 pero no el módulo de predicción implícita 62. Cuando uno o más de los valores de compensación no excedan el valor umbral 66, el comparador 58 puede generar y transmitir un código de activación que active el módulo de predicción implícita 62 pero no el módulo de predicción explícita 60.

**[86]** En base al código de activación, los módulos de predicción explícita o implícita 60 o 62 generan una segunda versión del bloque de vídeo predictivo de la trama predictiva. El activado de los módulos de predicción explícita o implícita 60 o 62 puede transmitir esta segunda versión al módulo de análisis R-D 64. Aunque se describa con respecto a un código de activación, la activación de los módulos de predicción explícita o implícita

60 o 62 puede producirse de varias formas. En algunos modos de realización, tanto los módulos de predicción explícita como implícita 60 y 62 pueden generar diferentes versiones del bloque de vídeo predictivo. En estos modos de realización, un multiplexor u otra lógica de selección puede seleccionar la segunda versión de estas versiones diferentes en base al código de activación y transmitir la versión seleccionada o la segunda versión al módulo de análisis R-D 64.

[87] Independientemente de cómo se obtenga o determine la segunda versión, el módulo de análisis R-D 64 también puede recibir, en algún momento, la primera versión del mismo bloque predictivo de vídeo. El módulo de análisis R-D 64 puede determinar los primer y segundo valores de distorsión 72A, 72B ("Dist. 72A" y "Dist. 72B") de la manera descrita anteriormente para las primera y segunda versiones. El módulo de análisis R-D 64 también puede almacenar o por el contrario mantener el valor  $\lambda$  74 ("lamda 74") y el valor de tasa de bits 76 ("tasa 76"). El módulo de análisis de R-D 64 puede proporcionar uno apropiado de valores de distorsión 72A, 72B (valores de distorsión 72), un valor  $\lambda$  74 y un valor de tasa de bits 76 como entradas en la función de coste 68, que emite un valor de coste 70A asociado con la primera versión. El módulo de análisis R-D 64 también puede proporcionar uno apropiado de valores de distorsión 72, de un valor  $\lambda$  74 y de un valor de tasa de bits 76 como entradas en la función de coste 68, que emite un valor de coste 70B asociado con la segunda versión.

[88] El módulo de análisis R-D 76 puede comparar entonces los costes 70 entre sí para determinar cuál de los costes 70 es más bajo que el otro. El módulo de análisis R-D 76 puede emitir entonces la primera versión o la segunda versión del bloque de vídeo predictivo para el que se calculó el más bajo de los costes 70. El módulo de análisis R-D 76 puede emitir este coste más bajo, una de las primera y segunda versiones como datos predictivos representativos del bloque de vídeo predictivo. En algunos casos, el módulo de análisis R-D 64 emite como datos predictivos la segunda versión predicha por el módulo de predicción explícita 60 y los valores de compensación calculados por el módulo de cálculo de compensación 56, como se describió anteriormente.

[89] Como resultado de tener que realizar solamente el análisis R-D en dos en lugar de en tres versiones de la tercera unidad codificada, la unidad de estimación de movimiento 35 puede codificar de manera más eficiente el bloque de vídeo predictivo. Esta eficiencia puede referirse a la eficiencia computacional, que también puede traducirse en un consumo de energía más eficiente.

[90] Aunque se describió anteriormente con respecto a un bloque de vídeo predictivo, las técnicas pueden aplicarse a cualquier otra parte de una trama, por ejemplo, un segmento, así como a toda la trama. Como se indicó anteriormente con respecto a las ecuaciones (1) a (3), las variables  $proj$ ,  $proj_0$  y  $proj_1$  se refieren a un bloque particular de las respectivas tramas de referencia. Las técnicas pueden aplicarse a un bloque particular seleccionado de la trama como representativo de todo el bloque de esa trama y, dependiendo de la versión seleccionada, la unidad de compensación por movimiento 35 puede aplicar el mismo algoritmo predictivo que el determinado con respecto al bloque de referencia seleccionado para codificar toda la trama predictiva. De forma alternativa, cada bloque individual de la trama predictiva puede someterse a las técnicas descritas en esta divulgación para seleccionar eficientemente un modo de predicción para cada bloque predictivo en una trama predictiva dada. En otro aspecto, las técnicas pueden aplicarse al nivel de segmento en una de las dos formas descritas anteriormente. Como resultado, la referencia en esta divulgación para realizar uno de los algoritmos de predicción descritos anteriormente a un bloque de vídeo predictivo no debería considerarse limitante. Por el contrario, la referencia a un bloque de vídeo predictivo puede abarcar en general cualquiera de los ejemplos anteriores así como ejemplos que no se describan explícitamente en esta divulgación, pero que se entiendan o extrapolen fácilmente a partir de esta divulgación.

[91] La FIG. 4 es un diagrama de flujo que ilustra el ejemplo de funcionamiento de un codificador de vídeo en la realización de las técnicas descritas en esta divulgación. Aunque se describan con respecto a un codificador de vídeo 50 de la FIG. 2 y más particularmente a un componente particular del codificador de vídeo 50, por ejemplo, la unidad de compensación por movimiento 35 de la FIG. 3, las técnicas pueden implementarse en hardware y/o software mediante cualquiera de los dispositivos mencionados anteriormente con respecto a la FIG. 1. Inicialmente, la unidad de compensación por movimiento 35 y, más particularmente, el módulo de predicción por defecto 54 realiza, por ejemplo, de acuerdo con la ecuación (1) anterior, una predicción ponderada por defecto para generar una primera versión de un bloque de vídeo predictivo de una trama B desde un primer y un segundo bloque de vídeo de referencia de las respectivas primera y segunda tramas de referencia (78). Como se describió anteriormente, la unidad de compensación por movimiento 35 puede recuperar las primera y segunda tramas de referencia en su totalidad o solamente los primer y segundo bloques de vídeo de referencia del almacén de tramas de referencia 34. El módulo de predicción por defecto 54 puede hacer pasar entonces la primera versión al módulo de cálculo de compensación 56, que puede calcular uno o más valores de compensación, por ejemplo, un valor de compensación de CC promedio en forma de valor absoluto (80). El módulo de cálculo de compensación 56 puede enviar entonces estos uno o más valores de compensación de CC al comparador 58.

[92] El comparador 58 puede comparar estos uno o más valores de compensación de CC con el valor



umbral 66 (82). En base a la comparación, el comparador 58 puede emitir una señal de activación para activar cualquiera del módulo de predicción explícita 60 o del módulo de predicción implícita 62, pero no ambos. En la operación de ejemplo mostrada en la FIG. 4, el comparador 58 puede activarse por medio del módulo de predicción explícita de la señal de activación 60 al determinar que los valores de compensación de CC exceden el valor umbral 66 ("SÍ" 82). Sin embargo, al determinar que la compensación de CC no excede el valor umbral 66 ("NO" 82), el comparador 58 puede activarse por medio del módulo de predicción implícita de señal de activación 62.

**[93]** En casos donde los valores de compensación DC exceden el valor umbral 66 y el comparador 58 emite una señal de activación para activar el módulo de predicción explícita 60, el módulo de predicción explícita 60 realiza la predicción ponderada explícita, por ejemplo, de acuerdo con la ecuación (2) anterior, para generar una segunda versión del bloque de vídeo de predicción de los primer y segundo bloques de vídeo de referencia, como se describió anteriormente (84). En los casos donde los valores de compensación DC no exceden el valor umbral 66 y el comparador 58 emite una señal de activación para activar el módulo de predicción implícita 62, el módulo de predicción implícita 62 realiza una predicción ponderada implícita, por ejemplo, de acuerdo con la ecuación (3) anterior, para generar la segunda versión del bloque de vídeo de predicción a partir de los primer y segundo bloques de vídeo de referencia, como se describió anteriormente (86). Independientemente de cuál de los módulos de predicción explícita o implícita 60, 62 genere la segunda versión, el activado respectivamente de los módulos 60, 62 hace pasar la segunda versión al módulo de análisis R-D 64.

**[94]** El módulo de análisis R-D 64 también recibe, como se describió anteriormente, la primera versión del bloque de vídeo predictivo y selecciona una más apropiada de las primera y segunda versiones (88). Es decir, el módulo de análisis R-D 64 puede realizar el análisis R-D descrito anteriormente, por ejemplo, de acuerdo con la ecuación (4) anterior, para generar costes 70A y 70B para cada una de las primera y segunda versiones, respectivamente y seleccionar una de las primera y segunda versiones asociadas con el más bajo de los costes 70A, 70B. Esta versión seleccionada "más apropiadamente" representa, para una tasa de bits dada, los datos de vídeo originales a los que correspondan los datos de vídeo de predicción. La unidad de compensación por movimiento 35 puede emitir entonces esta versión seleccionada, cuyo codificador de vídeo 22 puede proceder a codificar como el bloque de vídeo predictivo de la trama B predictiva.

**[95]** En un aspecto, la unidad de compensación por movimiento 35 puede, cuando el valor absoluto del valor de compensación de CC calculado exceda el valor umbral, realizar el análisis R-D de acuerdo con la ecuación (4) anterior para seleccionar entre una primera versión del bloque de vídeo predictivo generado por medio de la predicción ponderada por defecto y una segunda versión del mismo bloque de vídeo predictivo generado mediante la predicción ponderada explícita usando las ponderaciones por defecto. La unidad de compensación por movimiento 35 puede, en otras palabras, seleccionar entre codificar explícitamente la ponderación por defecto o no codificar explícitamente las ponderaciones por defecto. Esta selección se produce, como se sugirió anteriormente, porque, cuando se realiza una predicción ponderada explícita, la unidad de estimación de movimiento 35 también puede codificar los valores de compensación de CC calculados.

**[96]** Sin embargo, cuando el valor absoluto de la compensación de CC calculada no exceda el valor umbral, la unidad de compensación por movimiento 35 puede realizar un análisis R-D de acuerdo con la ecuación (4) anterior para seleccionar entre la primera versión del bloque de vídeo predictivo generado por medio de la predicción ponderada por defecto y una segunda versión del mismo bloque de vídeo predictivo generado por medio de la predicción ponderada implícita usando las ponderaciones determinadas implícitamente. A este respecto, la unidad de compensación por movimiento 35 puede determinar que las compensaciones de CC no son necesarias para codificar y, por lo tanto, seleccionar entre la primera y la segunda versión. Como resultado de comparar esta compensación de CC calculada con el umbral, la unidad de compensación por movimiento 35 puede, en otras palabras, seleccionar entre realizar la predicción ponderada implícita y explícita, realizar la seleccionada de la predicción ponderada implícita y explícita sin realizar la no seleccionada de la predicción ponderada implícita y explícita para generar la segunda versión y codificar el bloque de vídeo predictivo como la primera o la segunda versión.

**[97]** La FIG. 5 es un diagrama de flujo que ilustra un ejemplo de operación de un codificador de vídeo al realizar las técnicas descritas en esta divulgación en más detalle. De nuevo, aunque se describa con respecto al codificador de vídeo 50 particular de la FIG. 2 y más particularmente un componente particular del codificador de vídeo 50, por ejemplo, la unidad de compensación por movimiento 35 de la FIG. 3, las técnicas pueden implementarse en hardware y/o software mediante cualquiera de los dispositivos mencionados anteriormente con respecto a la FIG. 1.

**[98]** Inicialmente, la unidad de compensación por movimiento 35 recibe dos unidades de datos de vídeo de referencia, por ejemplo, bloques de vídeo de referencia, del almacén de tramas de referencia 34, como se describió anteriormente (100). De forma alternativa, la unidad de compensación por movimiento 35 puede recibir dos unidades codificadas de referencia, por ejemplo, tramas de referencia, que incluyan cada una de las

dos unidades de datos de vídeo de referencia. El módulo de predicción por defecto 54 puede realizar la predicción ponderada por defecto usando los bloques de vídeo de referencia de la manera descrita anteriormente, por ejemplo, con respecto a la ecuación (1) para generar una primera versión de la unidad de datos de vídeo predictivos, por ejemplo, un bloque de vídeo predictivo, a partir de los dos, por ejemplo, primer y segundo, bloques de vídeo de referencia (102, 104). Después de generar la primera versión, el módulo de predicción por defecto 54 puede emitir la primera versión para compensar el módulo de cálculo 56.

**[99]** El módulo de cálculo de compensación 56, de nuevo de la manera descrita anteriormente, puede calcular valores de compensación, tales como los valores de compensación de CC descritos anteriormente, para la primera versión del bloque de vídeo predictivo (106). El módulo de cálculo de compensación 56 puede emitir entonces los valores de compensación al comparador 58, que procede a comparar los valores de compensación con el valor umbral 66 (108). Si uno o más, un promedio, posiblemente todos o cualquier otro indicador derivado de los valores de compensación calculados exceden el valor umbral 66 ("SÍ" 110), el comparador 58 puede generar y emitir un código de activación que active el módulo de predicción explícita 60 en lugar del módulo de predicción implícita 62. Si uno o más, un promedio, una media, posiblemente todos o cualquier indicador derivado de los valores de compensación calculados no exceden el valor umbral 66 ("NO" 110), el comparador 58 puede generar y emitir un código de activación que active el módulo de predicción implícita 62 en lugar del módulo de predicción explícita 60.

**[100]** Si el módulo de predicción explícita 60 está activado, el módulo de predicción explícita 60 realiza la predicción ponderada explícita de la manera descrita anteriormente con respecto a, por ejemplo, la ecuación (2) usando los dos bloques de vídeo de referencia para predecir o generar una segunda versión de la unidad de datos de vídeo predictivos, por ejemplo, el bloque de vídeo predictivo, a partir de los dos bloques de vídeo de referencia (112, 116). Sin embargo, si el módulo de predicción implícita 62 está activado, el módulo de predicción implícita 52 realiza una predicción ponderada implícita de la manera descrita anteriormente con respecto a, por ejemplo, la ecuación (3) usando los bloques de vídeo de referencia para predecir o generar la segunda versión del bloque de vídeo predictivo a partir de los bloques de vídeo de referencia (114, 116). Independientemente de qué módulo genere la segunda versión del bloque de vídeo predictivo, el módulo de predicción explícita 60 o el módulo de predicción implícita 62 emiten la segunda versión al módulo de análisis R-D 64.

**[101]** El módulo de análisis R-D también puede recibir la primera versión del bloque de vídeo predictivo y puede realizar un análisis R-D de la manera descrita anteriormente con respecto, por ejemplo, a la ecuación (4) para determinar los primer y segundo costes 72 (118). El módulo de análisis R-D 64 puede seleccionar entonces uno más bajo de los costes 72 (120). El módulo de análisis R-D 64 puede, al seleccionar el más bajo de los costes 72, seleccionar cualquiera de las primera o segunda versiones asociadas con el más bajo de los costes 72 y codificar esta versión asociada con el más bajo de los costes 72 como la tercera unidad codificada (122).

**[102]** La FIG. 6 es un diagrama que ilustra un ejemplo de secuencia ordenada temporalmente de unidades codificadas 124A-124E. Es decir, la unidad codificada 124A se produce en el tiempo antes que la unidad codificada 124B, que se produce en el tiempo antes que la unidad codificada 124C, que se produce en el tiempo antes que la unidad codificada 124D, que a su vez se produce en el tiempo antes que la unidad codificada 124E. Aunque las unidades codificadas 124A-124E ("unidades codificadas 124") se produzcan en el tiempo una antes que la otra, una o más de las unidades codificadas 124 pueden codificarse antes que otra de las unidades codificadas 124. Por ejemplo, las unidades codificadas 124A y 124E representan cada una tramas I que se codifican independientemente, por lo tanto la "I" en la trama I, a partir de cada una de las unidades codificadas 124, mientras que las unidades codificadas 124B y 124D representan cada una tramas P que se codifican de forma predictiva, de ahí la "P" en la trama P, a partir de al menos otra de las unidades codificadas 124. Por tanto, aunque la trama I 124E se produzca después de la trama P 124D, la trama I 124E puede codificarse en el tiempo antes que la trama P 124D ya que la codificación de la trama I 124E puede no depender de otra trama, que puede o no codificarse todavía.

**[103]** En cualquier caso, la unidad codificada 124C puede representar una trama B que se codifique de acuerdo con uno de los tres algoritmos de predicción bidireccional compensada en movimiento descritos anteriormente, de ahí la "B" en la trama B. Como se muestra en la FIG. 3, una unidad de datos de vídeo 126 de la trama B 124C puede predecirse desde una parte o unidad de datos de vídeo de una trama anterior en el tiempo, por ejemplo, una unidad de datos de vídeo 128 de la trama P 124B y una parte o unidad de datos de vídeo de una trama posterior o futura en el tiempo, por ejemplo, una unidad de datos de vídeo 130 de la trama P 124D. Aunque se describe como predicha bidireccionalmente a partir de una primera trama anterior 124B en el tiempo y una segunda trama posterior 124D en el tiempo, la unidad de datos de vídeo 104 puede predecirse bidireccionalmente a partir de dos unidades codificadas posteriores 124 o dos unidades codificadas anteriores 124. Las unidades de datos de vídeo 126, 128 y 130 pueden representar cada una un bloque de vídeo correspondiente, tal como un macrobloque o cualquier otro bloque de vídeo de cualquier tamaño. Un bloque de vídeo correspondiente puede comprender un par de bloques de vídeo que sean similares entre sí porque cada bloque define datos de píxeles cuyos valores medios, como ejemplo, están dentro del conjunto o de las

tolerancias adaptativas. Además, aunque se muestra como predicha bidireccionalmente a partir de dos tramas 124B y 124D directamente adyacentes, la unidad de datos de vídeo 126 puede predecirse bidireccionalmente a partir de tramas no adyacentes a la trama B 124C, tal como la unidad codificada 124A y la unidad codificada 124E. A este respecto, las técnicas no deberían limitarse al ejemplo proporcionado en esta divulgación.

5 **[104]** La unidad de estimación de movimiento 32 puede ubicar estas partes o unidades de datos de vídeo 128 y 130 que corresponden a la unidad de datos de vídeo 126 y, con fines ilustrativos, determinar los vectores de movimiento 132A y 132B ("vectores de movimiento 132"). La unidad de compensación por movimiento 35 puede recibir estas unidades de datos de vídeo 128 y 130 accediendo al almacén de tramas de referencia 34 y recuperando las tramas P 124B y 124D ("tramas P 124") o directamente desde la unidad de estimación de movimiento 32. Especialmente, una de las tramas P 124 puede comprender una trama del conjunto de tramas denominado lista 0 52A, mientras que la otra de las tramas P 124 puede comprender una trama del conjunto de tramas denominado lista 1 52B.

15 **[105]** La unidad de compensación por movimiento 35 puede implementar entonces las técnicas descritas en esta divulgación para seleccionar eficientemente una de las dos versiones de la unidad de datos de vídeo 126 generada al realizar solamente dos de los posibles tres algoritmos de predicción bidireccional compensada en movimiento usando las partes 128 y 130. En este ejemplo, las técnicas se aplican para predecir una parte o unidad de datos de vídeo, tal como un macrobloque o un bloque de vídeo de cualquier otro tamaño, de una unidad codificada o de una trama B 124C. Ya se apliquen a una parte, tal como la unidad de datos de vídeo 126, o a una unidad codificada completa, tal como la unidad codificada 124C, las técnicas pueden seleccionar eficientemente entre dos en lugar de tres versiones.

25 **[106]** Las técnicas descritas en el presente documento pueden implementarse en hardware, software, firmware o en cualquier combinación de lo anterior. Cualquier característica descrita como módulos, unidades o componentes puede implementarse juntos en un dispositivo lógico integrado o por separado, como dispositivos lógicos discretos pero interoperables. En algunos casos, diversas características pueden implementarse como un dispositivo de circuito integrado, tal como un chip de circuito integrado o conjunto de chips. Si se implementan en software, las técnicas pueden realizarse, al menos parcialmente, mediante un medio legible por ordenador que comprenda instrucciones que, al ejecutarse, cause que un procesador realice uno o más de los procedimientos descritos anteriormente.

35 **[107]** Un medio legible por ordenador puede formar parte de un producto de programa informático, que pueda incluir materiales de embalaje. El medio legible por ordenador puede comprender un medio de almacenamiento de datos por ordenador tal como una memoria de acceso aleatorio (RAM), una memoria dinámica síncrona de acceso aleatorio (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 código en forma de instrucciones o estructuras de datos y al que pueda accederse, y que pueda leerse y/o ejecutarse por un ordenador.

45 **[108]** El código 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 otra circuitería lógica integrada o discreta equivalente. Por consiguiente, el término "procesador", como se usa en el presente documento, puede referirse a cualquiera de las estructuras anteriores o a cualquier otra estructura adecuada para la implementación de las técnicas descritas en el presente documento. Además, en algunos aspectos, la funcionalidad descrita en el presente documento puede proporcionarse dentro del software o hardware dedicado. La divulgación también contempla cualquiera de una variedad de dispositivos de circuitos integrados que incluyan circuitos para implementar una o más de las técnicas descritas en esta divulgación. Dicha circuitería puede proporcionarse en un único chip de circuito integrado o en múltiples chips de circuitos integrados interoperables en un denominado conjunto de chips. Dichos dispositivos de circuitos integrados pueden usarse en una variedad de aplicaciones, algunas de las cuales pueden incluir el uso en dispositivos de comunicación inalámbrica, tales como teléfonos móviles.

55 **[109]** Se han descrito diversos modos de realización de las técnicas.

**REIVINDICACIONES**

1. Un procedimiento de codificación de datos de vídeo, comprendiendo el procedimiento:
  - 5 recibir las primera y segunda unidades codificadas de referencia de los datos de vídeo, en donde cada una de las primera y segunda unidades codificadas comprende cada una las respectivas primera y segunda unidades de datos de vídeo de referencia (100);
  - 10 realizar la predicción ponderada por defecto usando las ponderaciones por defecto para predecir una primera versión de una unidad de datos de vídeo predicha de una unidad codificada predicha de los datos de vídeo a partir de las primera y segunda unidades de datos de vídeo de referencia (102, 104), en donde cada una de las dos unidades codificadas de referencia se producen temporalmente antes o después de la unidad codificada predicha;
  - 15 calcular un valor de compensación para la primera versión de la unidad de datos de vídeo predicha (106) calculando la diferencia entre una luminancia promedio o una croma promedio de la unidad de datos de vídeo predicha y una respectiva luminancia promedio o una croma promedio de una de las unidades de datos de vídeo de referencia;
  - 20 comparar el valor de compensación calculado con un valor umbral;
  - 25 seleccionar, en base a la comparación del valor de compensación calculado con el valor umbral, un modo de predicción ponderada implícita usando ponderaciones determinadas a través del análisis de las unidades codificadas de referencia o del modo de predicción ponderada explícita usando ponderaciones explícitamente definidas y codificadas (108, 110);
  - 30 realizar el modo seleccionado para predecir una segunda versión de la unidad de datos de vídeo predicha a partir de las primera y segunda unidades de datos de vídeo de referencia (112, 114); y
  - 35 codificar la unidad de datos de vídeo predicha solamente como la primera versión o la segunda versión (122).
2. El procedimiento de la reivindicación 1, que comprende además determinar si la primera o la segunda versión de la unidad de datos de vídeo predicha codifica de manera más apropiada los datos de vídeo usando el análisis de distorsión-velocidad.
3. El procedimiento de la reivindicación 2, en donde determinar si la primera o la segunda versión de la unidad de datos de vídeo predicha codifica de manera más apropiada los datos de vídeo comprende:
  - 40 comparar cada una de las primera y segunda versiones de la unidad de datos de vídeo predicha con una parte correspondiente de los datos de vídeo para determinar los primer y segundo valores de distorsión, en donde los primer y segundo valores de distorsión indican respectivamente cada uno una cantidad de distorsión introducida por las primera y segunda versiones;
  - 45 calcular los primer y segundo costes para las primera y segunda versiones de la unidad de datos de vídeo predicha en base a los respectivos primer y segundo valores de distorsión;
  - 50 comparar los primer y segundo costes para determinar cuál de los costes es más bajo; y
  - 55 seleccionar la primera o la segunda versión en base a la comparación.
4. El procedimiento según la reivindicación 1, en donde la primera unidad codificada de referencia se produce en el tiempo antes de la unidad codificada predicha, en donde la segunda unidad codificada de referencia se produce en el tiempo después de la unidad codificada predicha, en donde la unidad codificada predicha comprende una trama predicha bidireccional (trama B) y en donde cada una de las primera y segunda unidades codificadas comprende una de una trama codificada independientemente (trama I), de una trama predicha (trama P) o de otra trama B.
5. Un medio de almacenamiento legible por ordenador que almacena instrucciones que, cuando se ejecutan mediante un procesador, provocan la ejecución por el procesador de las etapas del procedimiento de cualquiera de las reivindicaciones 1 a 4.
6. Un dispositivo para codificar datos de vídeo, comprendiendo el dispositivo:

medios para codificar datos de vídeo (50), en donde los medios para codificar datos de vídeo incluyen:

- 5 medios para almacenar (34) unas primera y segunda unidades codificadas de referencia de los datos de vídeo, en donde la primera unidad codificada de referencia incluye una primera unidad de datos de vídeo de referencia y la segunda unidad codificada de referencia incluye una segunda unidad de datos de vídeo de referencia;
- 10 medios para realizar la predicción ponderada por defecto (35) usando ponderaciones por defecto para predecir una primera versión de una unidad de datos de vídeo predicha de una unidad codificada predicha de los datos de vídeo a partir de las primera y segunda unidades de datos de vídeo de referencia, en donde cada una de las dos unidades codificadas de referencia se producen temporalmente antes o después de la unidad codificada predicha;
- 15 medios para calcular un valor de compensación (35) para la primera versión de la unidad de datos de vídeo predicha calculando la diferencia entre una luminancia promedio o un croma promedio de la unidad de datos de vídeo predicha y una luminancia media o un croma promedio respectivo de una de las unidades de datos de vídeo de referencia;
- 20 medios para comparar el valor de compensación calculado con un valor umbral;
- 25 medios para realizar, en base a la comparación del valor de compensación calculado con el valor umbral, cualquiera, pero no ambos, el modo de predicción ponderada implícita usando ponderaciones determinadas a través del análisis de las unidades codificadas de referencia o de la predicción ponderada explícita usando ponderaciones definidas explícitamente y codificadas (35) para predecir una segunda versión de la unidad de datos de vídeo predicha a partir de las primera y segunda unidades de datos de vídeo de referencia; y
- 30 medios para codificar la unidad de datos de vídeo predicha (35) solamente como la primera versión o la segunda versión.
7. El dispositivo según la reivindicación 6, que comprende además  
medios para determinar si la primera o la segunda versión de la unidad de datos de vídeo predicha  
35 codifica de manera más apropiada los datos de vídeo usando el análisis de distorsión-velocidad.
8. El dispositivo de la reivindicación 7, en donde el medio para determinar si la primera o la segunda versión de  
la unidad de datos de vídeo predicha codifica de manera más apropiada los datos de vídeo comprende:
- 40 medios para comparar cada una de las primera y segunda versiones de la unidad de datos de vídeo  
predicha con una parte correspondiente de los datos de vídeo para determinar los primer y segundo  
valores de distorsión, en donde los primer y segundo valores de distorsión indican respectivamente  
cada uno una cantidad de distorsión introducida por las primera y segunda versiones;
- 45 medios para calcular el primer y segundo coste para las primera y segunda versiones de la unidad  
de datos de vídeo predicha en base a los respectivos primer y segundo valores de distorsión;
- medios para comparar los primer y segundo costes para determinar cuál de los costes es más bajo;  
y
- 50 medios para seleccionar la primera o la segunda versión en base a la comparación.
9. El dispositivo según la reivindicación 6,  
en donde la primera unidad codificada de referencia se produce en el tiempo antes de la unidad  
codificada predicha,  
55 en donde la segunda unidad codificada de referencia se produce en el tiempo después de la unidad  
codificada predicha,  
en donde la unidad codificada predicha comprende una trama predicha bidireccional (trama B) y  
en donde cada una de las primera y segunda unidades codificadas comprende una de una trama  
codificada independientemente (trama I), una trama predicha (trama P) u otra trama B.
- 60 10. El dispositivo según la reivindicación 6, en donde los medios para codificar la unidad de datos de vídeo  
predicha codifican, cuando la segunda versión de la unidad de datos de vídeo predicha comprende la  
segunda versión predicha usando la predicción ponderada explícita, la unidad de datos de vídeo predicha  
como la segunda versión de la unidad de datos de vídeo predicha y el valor de compensación, además de la  
65 segunda versión.

11. El dispositivo según cualquiera de las reivindicaciones 8 a 10,  
en donde los medios para codificar datos de vídeo comprenden un codificador de vídeo o  
codificador/decodificador de vídeo, CODEC (50),  
5 en donde los medios para almacenar comprenden un almacén de tramas de referencia (34), en donde  
los medios para realizar la predicción ponderada por defecto comprenden una unidad de compensación  
de movimiento (35),  
en donde los medios para calcular un valor de compensación comprenden la unidad de compensación  
de movimiento (35),  
10 en donde los medios para realizar la predicción ponderada implícita o la predicción ponderada explícita,  
pero no ambos, comprenden la unidad de compensación por movimiento (35), y  
en donde los medios para codificar los datos de vídeo predichos comprenden la unidad de  
compensación por movimiento (35).

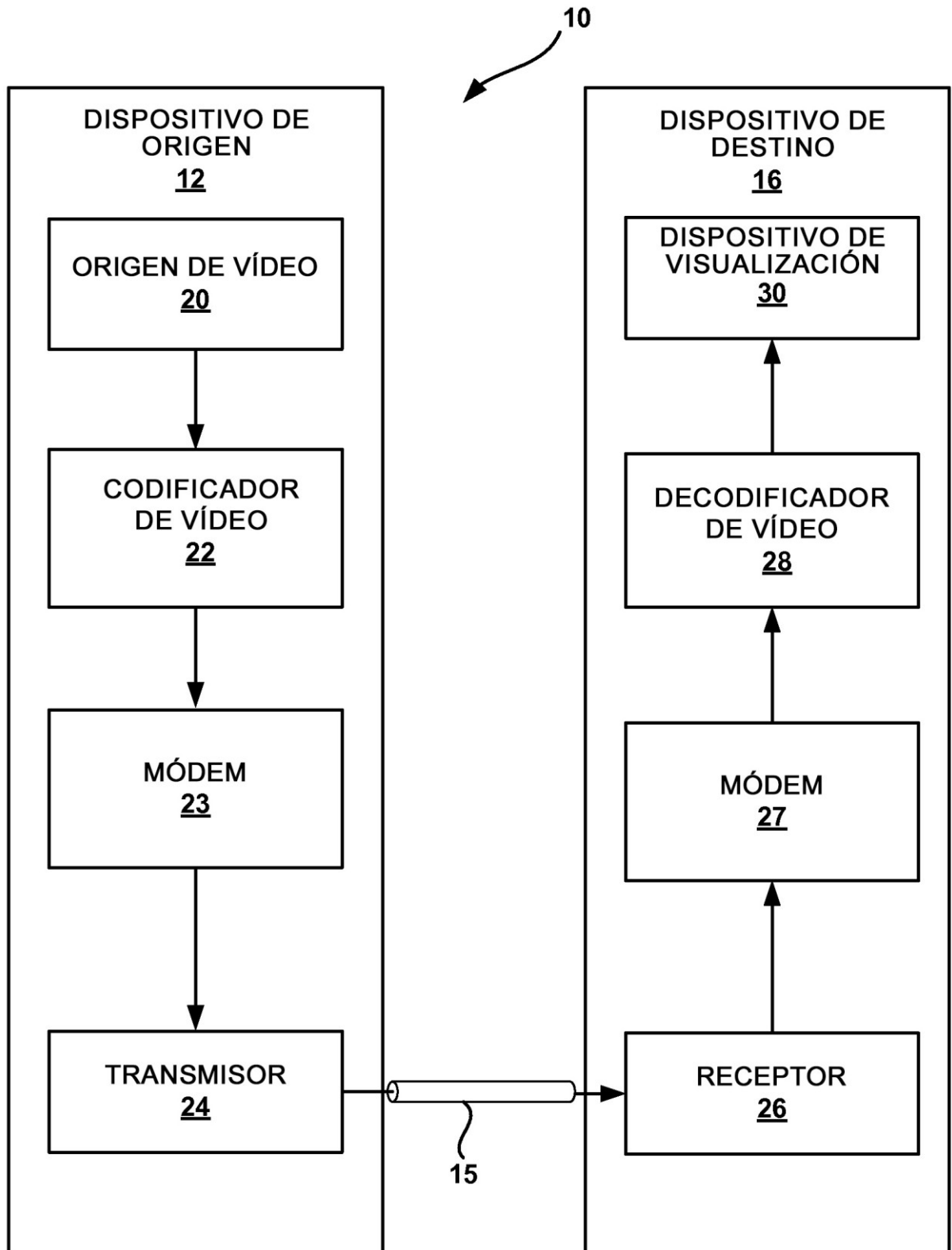


FIG. 1

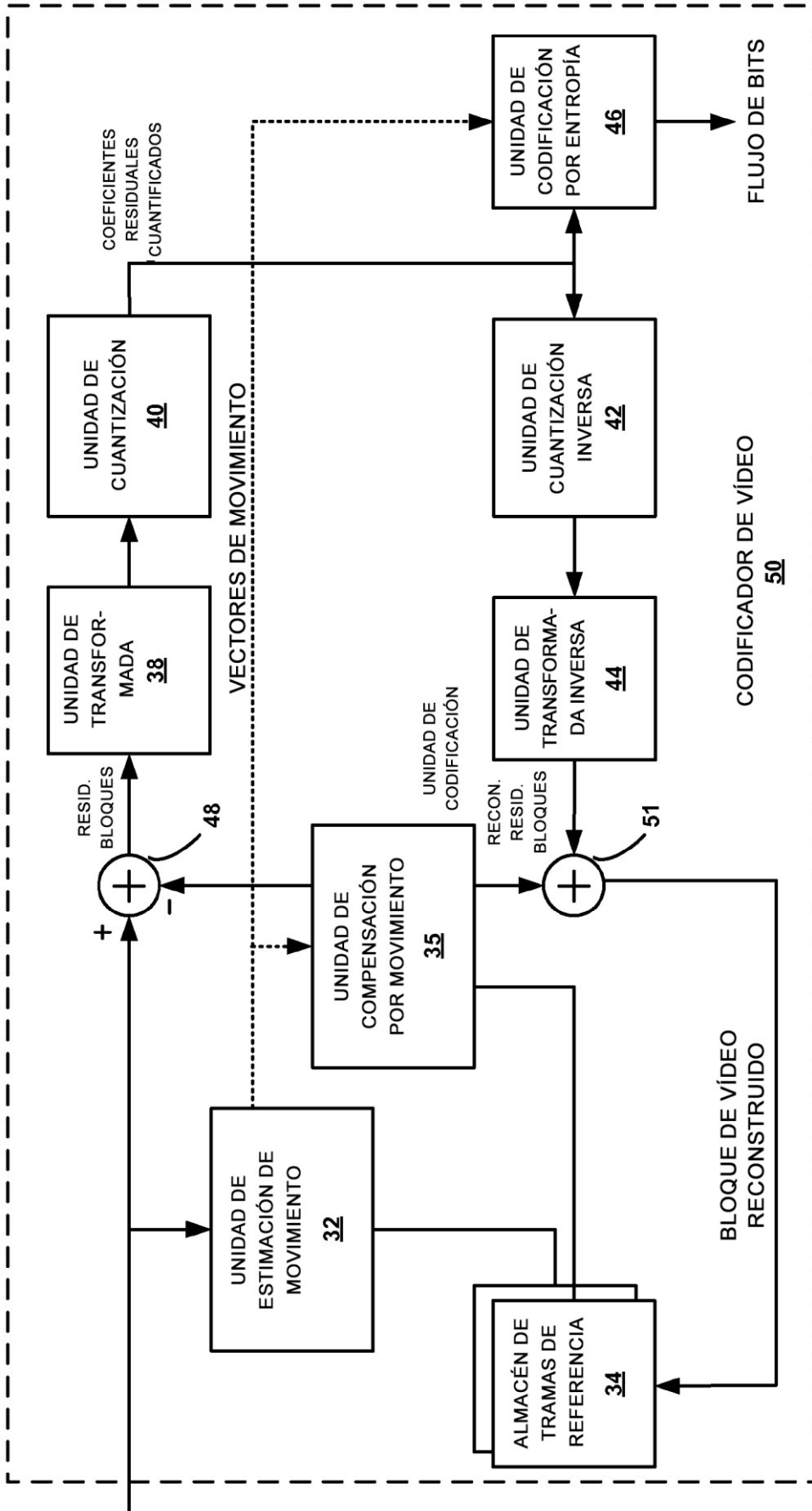


FIG. 2



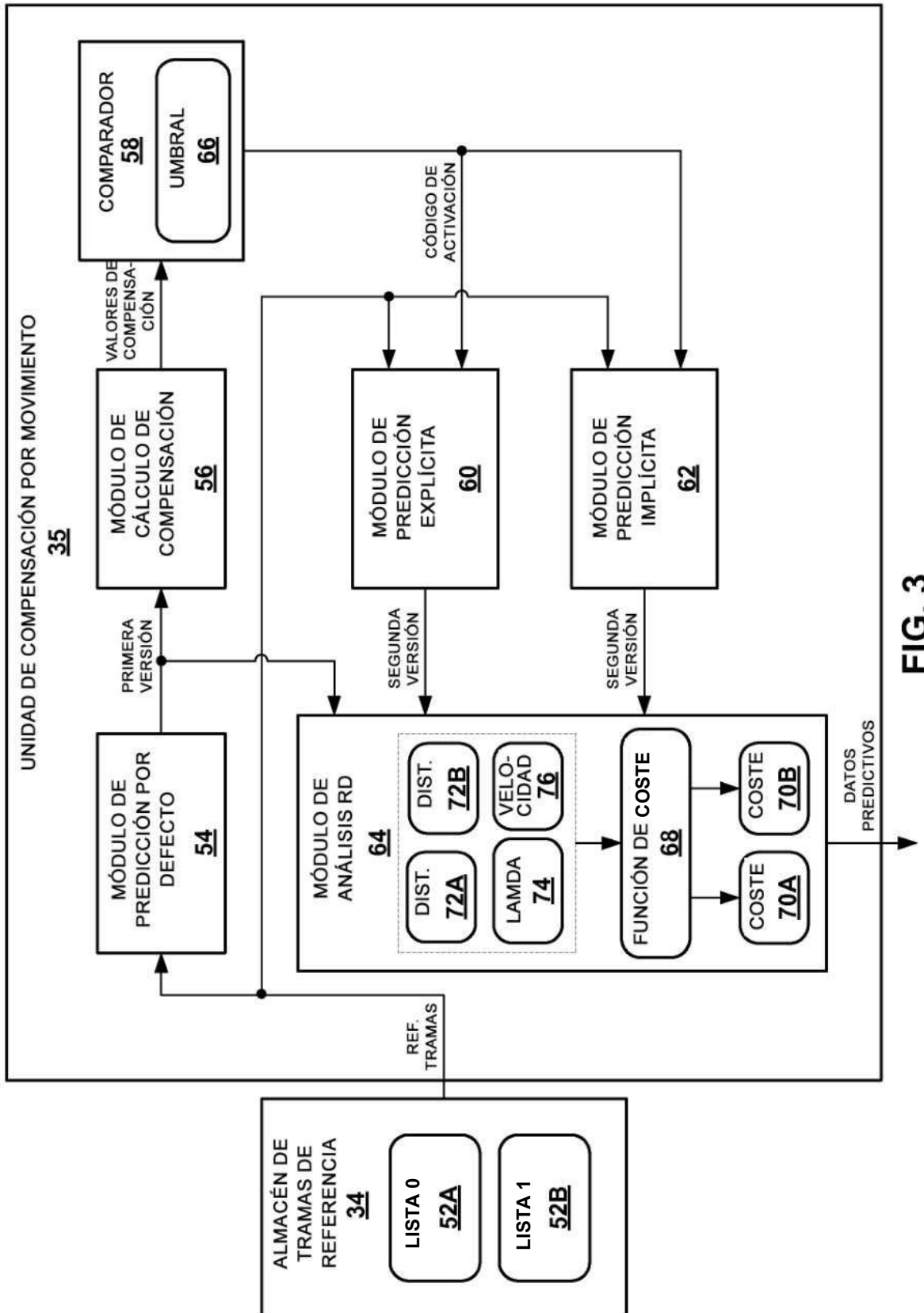


FIG. 3

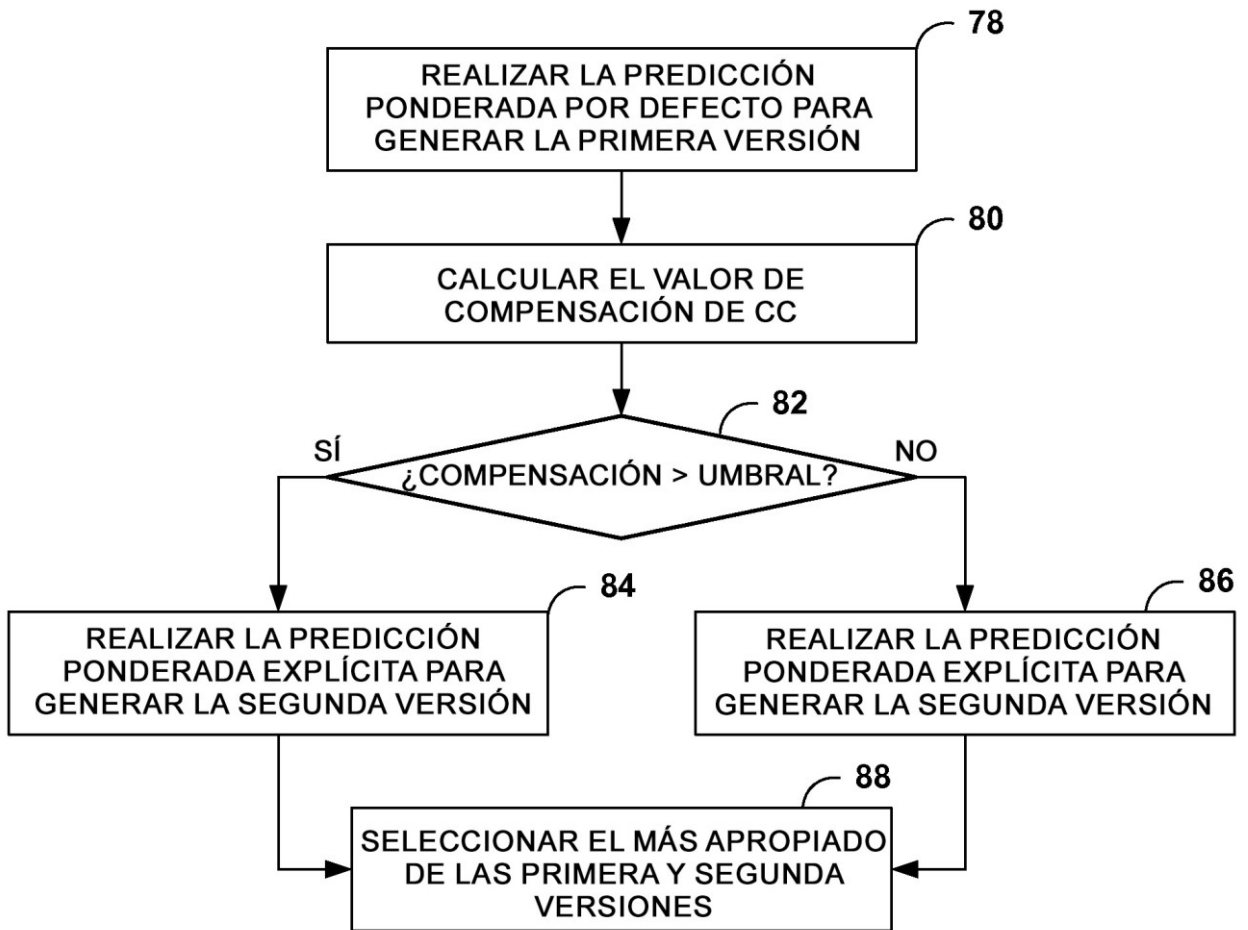


FIG. 4

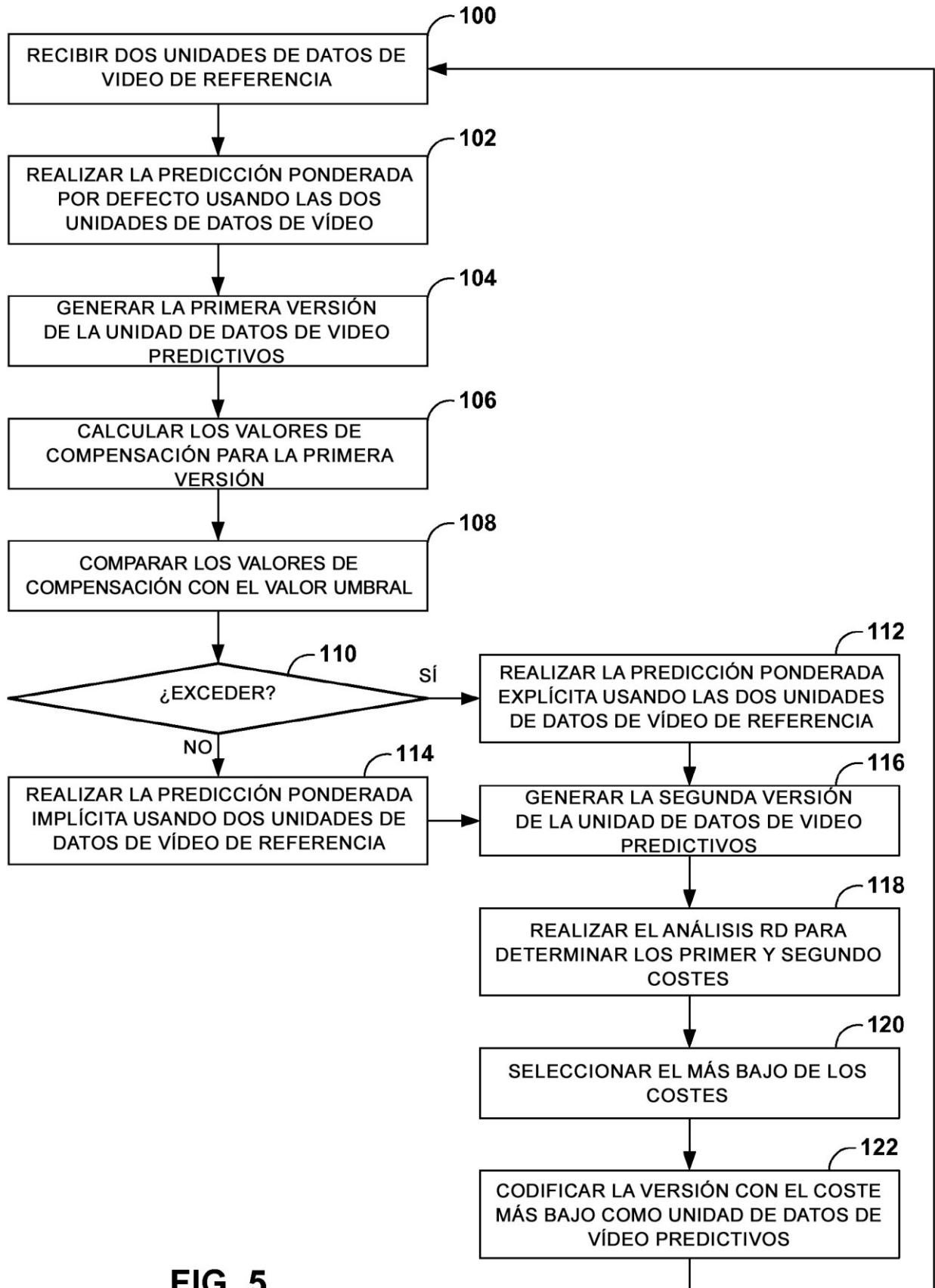


FIG. 5

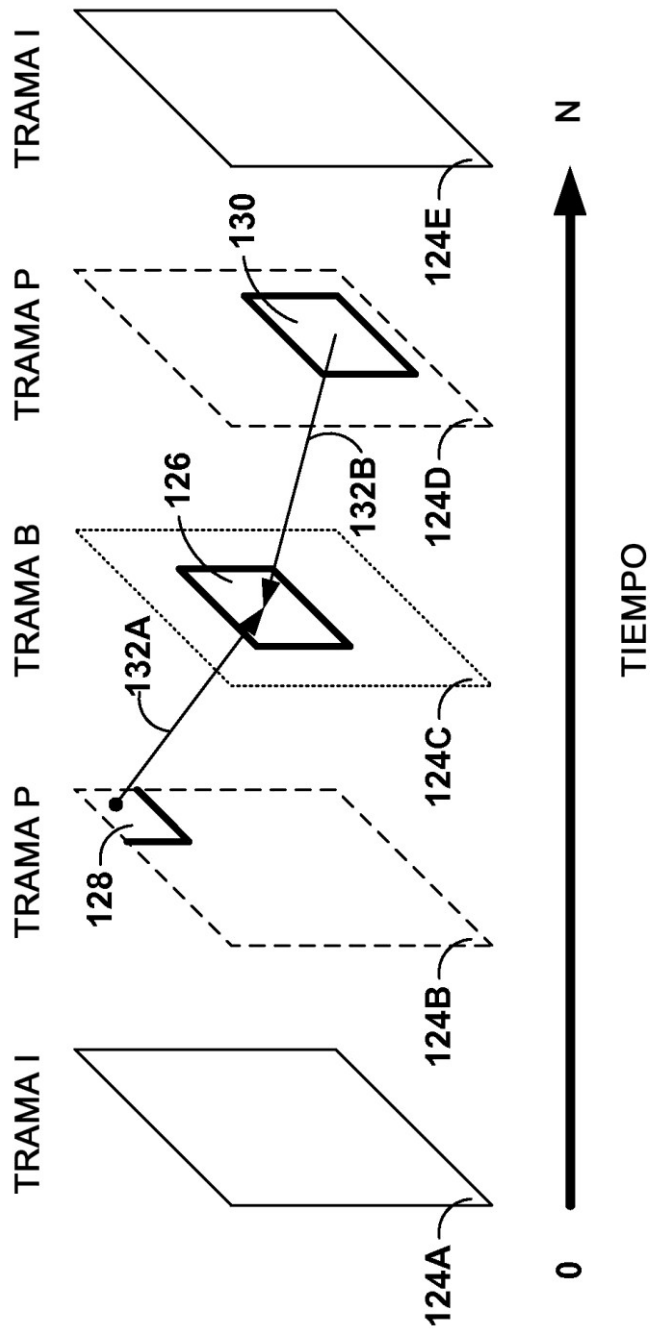


FIG. 6