(19) OFICINA ESPAÑOLA DE PATENTES Y MARCAS

ESPAÑA

(11) Número de publicación: **2 710 782**

(51) Int. CI.:

H04N 19/105 (2014.01)
H04N 19/52 (2014.01)
H04N 19/139 (2014.01)
H04N 19/176 (2014.01)
H04N 19/513 (2014.01)
H04N 19/43 (2014.01)

(12) TRADUCCIÓN DE PATENTE EUROPEA T3

(96) Fecha de presentación y número de la solicitud europea: **08.03.2012 E 16175112 (8)**

(97) Fecha y número de publicación de la concesión europea: **26.09.2018 EP 3099070**

(54) Título: **Codificación y descodificación de video**

(30) Prioridad:

09.03.2011 GB 201104033

(45) Fecha de publicación y mención en BOPI de la traducción de la patente:
26.04.2019

(73) Titular/es:

CANON KABUSHIKI KAISHA (100.0%)
30-2, Shimomaruko 3-chome Ohta-ku
Tokyo, 146-8501, JP

(72) Inventor/es:

LAROCHE, GUILLAUME y
ONNO, PATRICE

(74) Agente/Representante:

DURAN-CORRETJER, S.L.P

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

Codificación y descodificación de video

Sector técnico de la invención

La invención se refiere a un procedimiento y un dispositivo para la codificación de una secuencia de imágenes digitales y a un procedimiento y un dispositivo para la descodificación de una secuencia de bits correspondiente.

La invención pertenece al sector del proceso de señales digitales y, en concreto, al sector de la compresión de video utilizando compensación del movimiento para reducir las redundancias espaciales y temporales en las secuencias de video.

Descripción de la técnica anterior

Muchos formatos de compresión de video, por ejemplo, H.263, H.264, MPEG-1, MPEG-2, MPEG-4, SVC, utilizan la transformada discreta del coseno (DCT, Discrete Cosine Transform) basada en bloques y la compensación del movimiento para eliminar redundancias espaciales y temporales. Estos se pueden denominar formatos predictivos de video. Cada fotograma o imagen de la señal de video está dividido en segmentos que están codificados y pueden ser descodificados de manera independiente. Un segmento es habitualmente una porción rectangular del fotograma o, de manera más general, una porción de un fotograma o un fotograma entero. Además, cada segmento está dividido en macrobloques (MB) y cada macrobloque está además dividido en bloques, habitualmente bloques de 8x8 píxeles. En la más reciente Codificación de video de alta eficiencia (HEVC, High Efficiency Video Coding) actualmente en proceso de estandarización, cada segmento está dividido en Unidades máximas de codificación (LCU, Largest Coding Units) no superpuestas, generalmente bloques de 64 píxeles x 64 píxeles de tamaño. Cada LCU se puede, a su vez, dividir iterativamente en Unidades de codificación (CU - Coding Units) menores de tamaño variable utilizando descomposición en árbol cuaternario. Cada CU se puede partir asimismo en un máximo de 2 Unidades de partición rectangulares simétricas. Los fotogramas codificados son de dos tipos: fotogramas predichos temporales (predichos a partir de un fotograma de referencia llamados fotogramas P, o predichos a partir de dos fotogramas de referencia llamados fotogramas B) y fotogramas predichos no temporales (llamados Intra fotogramas o fotogramas I).

La predicción temporal consiste en encontrar en un fotograma de referencia, bien un fotograma anterior o un fotograma futuro de la secuencia de video, una porción de imagen o una zona de referencia que sea la más próxima al bloque a codificar. Esta etapa se conoce como estimación del movimiento. A continuación, la diferencia entre el bloque a codificar y la porción de referencia es codificada (compensación del movimiento), junto con un elemento de información del movimiento relativo al vector de movimiento, que indica la zona de referencia que se debe utilizar para la compensación del movimiento.

Con el fin de reducir más el coste de la codificación de la información del movimiento, se ha propuesto codificar un vector de movimiento mediante la diferencia con respecto a un predictor del vector de movimiento, calculado habitualmente a partir de los vectores de movimiento de los bloques que rodean al bloque a codificar.

En H.264, los vectores de movimiento son codificados con respecto a un predictor medio calculado a partir de vectores de movimiento situados en un entorno causal con el bloque a codificar, por ejemplo, de los bloques situados encima y a la izquierda del bloque a codificar. Solo se codifica la diferencia, también llamada vector de movimiento residual, entre el predictor medio y el vector de movimiento del bloque actual.

La codificación utilizando vectores de movimiento residuales ahorra tasa de bits, pero requiere que el descodificador efectúe el mismo cálculo del predictor del vector de movimiento para descodificar el valor del vector de movimiento de un bloque a descodificar.

Recientemente, se han propuesto otras mejoras, tales como la utilización de una serie de posibles predictores del vector de movimiento. Este procedimiento, denominado competición de vectores de movimiento, comprende la determinación entre varios predictores del vector de movimiento o candidatos de qué predictor del vector de movimiento minimiza el coste de la codificación, habitualmente un coste de distorsión de la velocidad, de la información residual del movimiento. La información residual del movimiento se compone del vector de movimiento residual, es decir, la diferencia entre el vector de movimiento real del bloque a codificar y el predictor del vector de movimiento seleccionado, y un elemento de información que indica el predictor del vector de movimiento seleccionado, tal como por ejemplo un valor codificado del índice del predictor del vector de movimiento seleccionado.

En la Codificación de video de alta eficiencia (HEVC, High Efficiency Video Coding) actualmente en proceso de estandarización, se ha propuesto la selección de un primer predictor espacial del vector de movimiento de un grupo de bloques predeterminado por encima del bloque actual a codificar, y un segundo predictor espacial del vector de movimiento de otro grupo de bloques predeterminado a la izquierda del bloque actual, y un predictor temporal del

vector de movimiento de un bloque contiguo en un fotograma de referencia. Un bloque contiguo es un bloque situado en la misma posición que el bloque a codificar, pero en otra imagen. Este esquema de selección del predictor del vector de movimiento se denomina Predicción avanzada de vectores de movimiento (AMVP, Advanced Motion Vector Prediction).

Como resultado de la selección de predictor del vector de movimiento, se obtiene un conjunto de candidatos de predictor del vector de movimiento que mezclan predictores espaciales y predictores temporales, comprendiendo el conjunto en candidatos 0, 1, 2 o 3 dependiendo de la disponibilidad de los diferentes predictores espaciales del vector de movimiento y temporales. Con el fin de reducir más la sobrecarga de señalización del predictor del vector de movimiento en la secuencia de bits, el conjunto de predictores de vectores de movimiento se reduce mediante la aplicación de un proceso de supresión que elimina vectores de movimiento duplicados, es decir, vectores de movimiento que tienen el mismo valor. Por ejemplo, si el primer y el segundo predictores espaciales del vector de movimiento mencionados anteriormente son iguales, solo uno de ellos, así como el predictor temporal del vector de movimiento, se deben mantener como candidatos para la predicción del vector de movimiento. En este caso, solo es necesario un bit para indicar el índice del predictor del vector de movimiento al descodificador.

Los candidatos de predictor del vector de movimiento pueden ser asimismo ordenados para situar el predictor del vector de movimiento más probable en la primera posición, dado que la mínima sobrecarga se produce si el primer candidato se elige como el mejor predictor.

Una reducción suficiente del conjunto de predictores conduce a una ganancia en la sobrecarga de la señalización, dado que la indicación del predictor del vector de movimiento seleccionado se puede codificar utilizando menos bits. En el límite, el conjunto de candidatos se puede reducir a 1, por ejemplo, si todos los predictores del vector de movimiento son iguales y, por lo tanto, no es necesario introducir ninguna información relativa al predictor del vector de movimiento seleccionado en la secuencia de bits.

Tal como se ha descrito anteriormente, en la propuesta de HEVC actual, los vectores de movimiento son descodificados mediante codificación predictiva, utilizando una serie de predictores del vector de movimiento que podrían ser predictores espaciales o temporales. Por lo tanto, para cada fotograma que se utiliza como fotograma de referencia para la obtención del predictor del vector de movimiento contiguo, es necesario almacenar junto al codificador y al descodificador sus vectores de movimiento correspondientes. Por defecto, el tamaño de la memoria de vectores de movimiento es importante, considerando en primer lugar la granularidad del mapeo del movimiento (en el diseño de la HEVC actual el tamaño mínimo Inter-bloques es 4x4) y, en segundo lugar, que existen hasta dos vectores por bloque de movimiento para un SEGMENTO_B. Se estima que para imágenes con resolución de 4Kx2K, y utilizando una granularidad de un conjunto de vectores de movimiento por bloque de 4x4 píxeles, son necesarios 26 bits por fotograma.

La publicación de Yeping Su y Andrew Segall, “On motion vector competition”, (“Sobre la competición de los vectores de movimiento”) JCTVC-C257, Guangzhou, CN, 7-15 de octubre de 2010, y la publicación de Yeping Su y Andrew Segall “CE9: Reduced resolution storage of motion vector data” (“CE9: Almacenamiento de resolución reducida de los datos de los vectores de movimiento”), JCTVC-D072, Daegu, KR, 20-28 de enero de 2011, propone la reducción del tamaño de la memoria de vectores de movimiento necesaria en el lado del descodificador para el almacenamiento de vectores contiguos en el tiempo e índices de referencia. La solución propuesta comprende una simple compactación por bloques de los vectores de movimiento, tal como se muestra en la figura 1. En la implementación actual de la HEVC, se mantiene un vector de movimiento de entre 16 en cada cuadrado de 64x64 píxeles, y se considera el vector de movimiento del bloque de 4x4 superior izquierdo. Esto reduce las necesidades de memoria en un factor 16. Este factor de reducción se señaliza en el conjunto de parámetros de secuencia. El proceso de compresión de la memoria de vectores de movimiento se realiza tras un proceso de filtrado de bucle adaptativo, y antes de que la imagen descodificada se coloque en la memoria temporal de imágenes descodificadas (DPB, Descoded Picture Buffer).

Esta solución permite una reducción significativa de la memoria de vectores de movimiento, con una penalización en la eficiencia de la codificación. Sin embargo, esta solución no considera la diversidad de los predictores en el conjunto para la compactación por bloques en la compresión de la memoria de MV (Vectores de movimiento, Motion Vector). Por otra parte, se probaron varios criterios, pero sin considerar la diversidad. Por otra parte, esta solución no es ni adaptativa ni flexible. Además, cuando es necesaria más compresión, es necesaria una mayor relación de reducción en la eliminación de vectores de movimiento, lo que puede conducir a una pérdida significativa de eficiencia en la codificación.

El documento WO 2004/012459A describe otra propuesta para limitar el número de vectores de movimiento almacenados. La información de limitación se codifica como información de cabecera (posiblemente a un nivel global, tal como SPS, a nivel de imagen, tal como PPS o a nivel de segmento, tal como cabecera de segmento). Esta información de limitación indica el número de imágenes o el número de macrobloques para los que están almacenados vectores de movimiento. Si un bloque codificado hace referencia a un vector de movimiento que está almacenado, entonces se utiliza este vector de movimiento; de lo contrario, se utiliza un vector de movimiento nulo (ambos componentes son iguales a 0). No obstante, esta solución no considera la compactación por bloques.

Aparte de las propuestas para reducir las necesidades de almacenamiento de vectores de movimiento, otras propuestas de la técnica anterior han perseguido mejoras en los predictores para AMVP.

Por ejemplo, Jung, G. Clare, (Orange Labs), "Temporal MV predictor modification for MV-Comp, Skip, Direct and Merge" ("Modificación de predictores temporales de MV para esquemas de MV-comp, Salto, Dirección y Fusión"), JCTVC-D164, Daegu, KR, 20-28 de enero de 2011, propone utilizar un predictor temporal centrado; véase también el documento WO 2011/001077 A. En esta propuesta, el bloque contiguo se encuentra mediante el mapeo del centro del bloque a un fotograma de referencia contiguo, en lugar de al origen (parte superior izquierda) del bloque del bloque (tal como la versión previa de HEVC). Los vectores de movimiento central y superior izquierdo tienen el mismo valor si el bloque contiguo tiene el mismo tamaño o un tamaño mayor que el bloque actual. La figura 2 muestra un bloque contiguo subdividido en 64 bloques de 4x4, que tienen cada uno su propio vector de movimiento. Este bloque contiguo se utiliza para un bloque actual de 32x32 a codificar, y el predictor temporal seleccionado es el vector de movimiento del bloque de 4x4 sombreado en el centro del bloque contiguo.

El objetivo de seleccionar el centro de la partición en lugar de la parte superior izquierda es encontrar una mejor representación del movimiento del bloque contiguo. Teóricamente, el centro del bloque es una representación mejor del bloque, debido a que es el baricentro del bloque. Pero el predictor temporal a partir del bloque contiguo no es el único predictor del conjunto de predictores AMVP.

Jian-Liang Lin, Yu-Pao Tsai, Yu-Wen Huang, Shawmin Lei, (Media Tek Inc.), en "Improved Advanced Motion Vector Prediction" ("Predicción avanzada mejorada de vectores de movimiento"), JCTVC-D125, Daegu, KR, 20-28 de enero de 2011, proponen añadir varios bloques temporales tal como se representa en la figura 3. Se proponen 4 predictores, el superior izquierdo contiguo temporal (T) y otras tres esquinas (inferior izquierdo (I), inferior derecho (H), superior derecho (E)). De esta manera, se añaden predictores al conjunto de predictores y compiten entre sí según el criterio de distorsión de velocidad.

Ninguna de las dos propuestas anteriores para mejorar el conjunto de predictores de vectores de movimiento implica la compresión de la memoria de vectores de movimiento.

Es deseable solucionar uno o varios de los inconvenientes de la técnica anterior.

Es deseable asimismo reducir las necesidades de la memoria de vectores de movimiento. Preferentemente, esta reducción se debería conseguir sin ninguna, o sin ninguna significativa, penalización adicional en la eficiencia de la codificación. Incluso más Preferentemente, esta reducción se debería conseguir en combinación con mejoras en la eficiencia de la codificación.

Es deseable asimismo conseguir mejoras en la eficiencia de la codificación. Preferentemente, estas mejoras se deberían conseguir sin ninguna penalización, o sin ninguna significativa, en las necesidades de la memoria de vectores de movimiento. Incluso, más Preferentemente, estas mejoras se deberían conseguir en combinación con reducciones en las necesidades de la memoria de vectores de movimiento.

Es deseable asimismo encontrar buenos vectores de movimiento representativos en un fotograma de referencia para utilizar como predictores, no obstante, estos vectores de movimiento representativos se utilizan, es decir, incluso si se utilizan con propósitos diferentes a reducir las necesidades de la memoria de vectores de movimiento y/o a la mejora de la eficiencia de la codificación.

CARACTERÍSTICAS DE LA INVENCIÓN

La invención solo se define en las reivindicaciones, correspondientes a la duodécima realización en la descripción. Las diversas realizaciones que se describen a continuación engloban, por lo tanto, ejemplos que no realizan directamente la presente invención.

Según un primer aspecto de la presente invención, se proporciona un procedimiento para determinar un predictor temporal del vector de movimiento que se puede incluir, junto con uno o varios predictores espaciales del vector de movimiento, en un conjunto de predictores de vectores de movimiento para un bloque a codificar de un fotograma actual, comprendiendo el procedimiento seleccionar como dicho predictor temporal un vector de movimiento de entre los vectores de movimiento en un bloque de referencia de un fotograma de referencia diferente de dicho fotograma actual, siendo el bloque de referencia un bloque del fotograma de referencia contiguo a dicho bloque a codificar o un bloque del fotograma de referencia próximo al bloque contiguo, y estando basada la selección en un criterio de diversidad para conseguir la diversidad entre los predictores de dicho conjunto.

Dicho procedimiento puede ser utilizado para encontrar el mejor vector de movimiento para la compactación de vectores por bloques durante el proceso de compresión de la memoria. La selección del mejor predictor puede basarse en una distancia espacial y/o espacio-temporal entre las posiciones del predictor que podría ser utilizada para la predicción del vector de movimiento actual o para el predictor de los modos de Fusión y Salto.

En una realización, una posición del vector de movimiento se selecciona de entre un conjunto de posiciones que, en términos de distancia espacial, es la posición más alejada en comparación con las posiciones de los otros predictores. Además, las posiciones consideradas deben ser una buena representación de los bloques vecinos o del bloque contiguo. De esta manera es posible obtener diversidad en el conjunto de predictores.

Según un segundo aspecto de la presente invención, se proporciona un procedimiento para determinar un predictor temporal del vector de movimiento que se puede incluir, junto con uno o varios predictores espaciales del vector de movimiento, en un conjunto de predictores de vectores de movimiento para un bloque a codificar de un fotograma actual, comprendiendo el procedimiento seleccionar como dicho predictor temporal el vector de movimiento en una posición inferior derecha de un bloque del fotograma de referencia contiguo a dicho bloque a codificar.

Una selección de este tipo funciona bien con tamaños de bloque relativamente pequeños y tiene la ventaja de no tener carga de procesamiento, porque la selección está predeterminada.

Según un tercer aspecto de la presente invención, se proporciona un procedimiento para determinar un predictor temporal del vector de movimiento que se puede incluir, junto con uno o varios predictores espaciales del vector de movimiento, en un conjunto de predictores de vectores de movimiento para un bloque a codificar de un fotograma actual, comprendiendo el procedimiento seleccionar como dicho predictor temporal el vector de movimiento en una posición inferior derecha de un bloque del fotograma de referencia contiguo a dicho bloque a codificar o en una posición predeterminada, en la proximidad de dicha posición inferior derecha, en un bloque del fotograma de referencia próximo al bloque contiguo.

En este aspecto de la invención, la selección se expande para incluir posiciones en la proximidad de dicha posición inferior derecha, en un bloque del fotograma de referencia próximo al bloque contiguo.

Una selección de este tipo funciona bien asimismo con tamaños de bloque relativamente pequeños y tiene la ventaja de no tener carga de procesamiento, porque la selección está predeterminada.

Según un cuarto aspecto de la presente invención, se proporciona un procedimiento para determinar un predictor temporal del vector de movimiento que se puede incluir, junto con uno o varios predictores espaciales del vector de movimiento, en un conjunto de predictores de vectores de movimiento para un bloque a codificar de un fotograma actual, comprendiendo el procedimiento seleccionar como dicho predictor temporal un vector de movimiento de entre los vectores de movimiento en un bloque de referencia de un fotograma de referencia diferente de dicho fotograma actual, siendo el bloque de referencia un bloque del fotograma de referencia contiguo a dicho bloque a codificar o un bloque del fotograma de referencia próximo al bloque contiguo, y dependiendo la selección de un tamaño de bloque del bloque a codificar.

Un procedimiento de este tipo permite que el vector de movimiento sea seleccionado de manera diferente para diferentes tamaños de bloque. En experimentos, se consiguieron buenos resultados con este procedimiento.

Por ejemplo, en el caso de un primer tamaño de bloque, el vector de movimiento en una primera posición predeterminada en el bloque de referencia puede ser seleccionado como dicho predictor temporal y, en el caso de un segundo tamaño de bloque diferente del primer tamaño de bloque, el vector de movimiento en una segunda posición predeterminada en el bloque de referencia, diferente de la primera posición, puede ser seleccionado como dicho predictor temporal. Cuando el primer tamaño de bloque es más pequeño que el segundo tamaño de bloque, es efectivo hacer que la primera posición sea una posición inferior derecha en el bloque contiguo y hacer que la segunda posición sea una posición central en el bloque contiguo.

De manera alternativa, en el caso de un primer tamaño de bloque, un vector de movimiento en el bloque de referencia puede ser seleccionado como dicho predictor temporal utilizando un primer procedimiento de selección y, en el caso de un segundo tamaño de bloque diferente del primer tamaño de bloque, un vector de movimiento en el bloque de referencia puede ser seleccionado utilizando un segundo procedimiento de selección diferente de dicho primer procedimiento.

Algunas realizaciones implican determinar posiciones relativas de por lo menos uno de dichos predictores espaciales y por lo menos un vector de movimiento de dicho bloque de referencia, y seleccionar dicho predictor temporal en función de dichas posiciones relativas. Esta es una forma efectiva de garantizar la diversidad controlada entre los predictores.

Otras realizaciones implican determinar la disponibilidad, por lo menos, de uno de dichos predictores espaciales, y seleccionar dicho predictor temporal en función de dicha disponibilidad. De nuevo, esta es una manera efectiva de garantizar la diversidad controlada. Por ejemplo, en algunas implementaciones se pueden excluir uno o varios predictores espaciales. Asimismo, para algunas posiciones en el fotograma, por ejemplo, la parte superior del fotograma o el borde del lado izquierdo del fotograma, se sabe de antemano que uno o varios predictores espaciales no están disponibles.

Otra realización utiliza una tabla de búsqueda para seleccionar dicho predictor temporal, dicha tabla de búsqueda almacenando para una o varias diferentes combinaciones seleccionables de predictores espaciales información que identifica el vector de movimiento del bloque de referencia a seleccionar para la combinación en cuestión. Esto tiene una complejidad baja y puede funcionar de manera rápida.

Los procedimientos de los aspectos primero a cuarto pueden ser utilizados para lograr una reducción en las necesidades de memoria para almacenar vectores de movimiento. Asimismo, se pueden utilizar para mejorar la eficiencia de la codificación.

En el caso de reducción de la memoria, el predictor temporal se selecciona antes de que se seleccione, por lo menos, un predictor espacial. Por consiguiente, en este caso, los procedimientos pueden comprender además estimar la posición del predictor espacial en cuestión, dependiendo de una posición media entre las posiciones potenciales del predictor espacial en cuestión. La media puede ser ponderada, si se desea, según las frecuencias relativas en las que se seleccionarán las posiciones potenciales. Esto, a su vez, depende del proceso de selección aplicado, por ejemplo, en AMVP se realiza una búsqueda de derecha a izquierda para el predictor superior y de abajo a arriba para el predictor izquierdo. Esto afecta a las frecuencias relativas.

Según un quinto aspecto de la presente invención, se proporciona un procedimiento para almacenar vectores de movimiento para ser utilizados como predictores temporales del vector de movimiento, que comprende almacenar, entre los vectores de movimiento del bloque a codificar de un fotograma actual, el vector de movimiento que tiene la misma posición que el predictor temporal determinado por el procedimiento de uno cualquiera de los aspectos primero a cuarto de la presente invención, de tal modo que el vector de movimiento almacenado esté disponible como un predictor temporal para la codificación de otro fotograma haciendo referencia a dicho fotograma actual, y no almacenar otros de dichos vectores de movimiento del bloque a codificar del fotograma actual.

Esto mejora la eficiencia de la codificación en comparación con los procedimientos anteriores de compactación por bloques.

Según un sexto aspecto de la presente invención, se proporciona un procedimiento para generar un conjunto de predictores de vectores de movimiento para un bloque a codificar de un fotograma actual, permitiéndose al conjunto incluir, por lo menos, un predictor espacial del vector de movimiento y, por lo menos, un predictor temporal del vector de movimiento, que comprende utilizar el procedimiento de uno cualquiera de los primero a cuarto aspectos de la presente invención para determinar dicho predictor temporal del vector de movimiento o uno de dichos predictores temporales del vector de movimiento del conjunto.

Esto puede proporcionar mejoras en la eficiencia de la codificación. En este aspecto de la invención es posible almacenar todos los vectores de movimiento, es decir, no es necesario reducir el número de vectores de movimiento almacenados, aunque esto, por supuesto, es preferible.

En este caso, por lo menos, uno de dichos predictores espaciales se selecciona antes de que se determine el predictor temporal, y la determinación del predictor temporal se realiza preferentemente dependiendo del predictor o predictores espaciales seleccionados. A diferencia de la aplicación de reducción de la memoria, por lo tanto, los resultados reales de las selecciones de predictores espaciales se pueden tener en cuenta, lo que da buenos resultados en términos de mejora de la diversidad.

Según un séptimo aspecto de la presente invención, se proporciona un procedimiento para codificar una secuencia de imágenes digitales en una secuencia de bits, estando codificada, por lo menos una porción de una imagen, por compensación del movimiento con respecto a una porción de la imagen de referencia, comprendiendo el procedimiento de determinación un predictor temporal del vector de movimiento según cualquiera de los aspectos primero a cuarto de la presente invención.

Según un octavo aspecto de la presente invención, se proporciona un procedimiento para codificar una secuencia de imágenes digitales en una secuencia de bits, estando codificada, por lo menos una porción de una imagen, por compensación del movimiento con respecto a una porción de la imagen de referencia, comprendiendo el procedimiento de almacenamiento vectores de movimiento según el quinto aspecto de la presente invención.

Según un noveno aspecto de la presente invención se proporciona un procedimiento para codificar una secuencia de imágenes digitales en una secuencia de bits, estando codificada, por lo menos una porción de una imagen, por compensación del movimiento con respecto a una porción de la imagen de referencia, comprendiendo el procedimiento de generación un conjunto de predictores de vectores de movimiento según el sexto aspecto de la presente invención.

En los aspectos décimo a decimoquinto de la presente invención, los procedimientos de los aspectos primero a sexto también se pueden utilizar en un procedimiento de descodificación de una secuencia de bits que comprende una secuencia codificada de imágenes digitales, estando codificada, por lo menos una porción de una imagen, por

compensación del movimiento con respecto a una imagen de referencia.

Según un decimosexto aspecto de la presente invención, se proporciona un dispositivo para determinar un predictor temporal del vector de movimiento que se puede incluir, junto con uno o varios predictores espaciales del vector de movimiento, en un conjunto de predictores de vectores de movimiento para un bloque a codificar de un fotograma actual, comprendiendo el dispositivo medios para seleccionar como dicho predictor temporal un vector de movimiento de entre los vectores de movimiento en un bloque de referencia de un fotograma de referencia diferente de dicho fotograma actual, siendo el bloque de referencia un bloque del fotograma de referencia contiguo a dicho bloque a codificar o un bloque del fotograma de referencia próximo al bloque contiguo, y la selección se basa en un criterio de diversidad para conseguir la diversidad entre los predictores de dicho conjunto.

Según un decimoséptimo aspecto de la presente invención, se proporciona un dispositivo para determinar un predictor temporal del vector de movimiento que se puede incluir, junto con uno o varios predictores espaciales del vector de movimiento, en un conjunto de predictores de vectores de movimiento para un bloque a codificar de un fotograma actual, comprendiendo el dispositivo medios para seleccionar como dicho predictor temporal el vector de movimiento en la posición inferior derecha de un bloque del fotograma de referencia contiguo a dicho bloque a codificar.

Según un decimoctavo aspecto de la presente invención, se proporciona un dispositivo para determinar un predictor temporal del vector de movimiento que se puede incluir, junto con uno o varios predictores espaciales del vector de movimiento, en un conjunto de predictores de vectores de movimiento para un bloque a codificar de un fotograma actual, comprendiendo el dispositivo medios para seleccionar como dicho predictor temporal el vector de movimiento en una posición inferior derecha de un bloque del fotograma de referencia contiguo a dicho bloque a codificar o en una posición predeterminada, en la proximidad de dicha posición inferior derecha, en un bloque del fotograma de referencia próximo al bloque contiguo.

Según un decimonoveno aspecto de la presente invención, se proporciona un dispositivo para determinar un predictor temporal del vector de movimiento que se puede incluir, junto con uno o varios predictores espaciales del vector de movimiento, en un conjunto de predictores de vectores de movimiento para un bloque a codificar de un fotograma actual, comprendiendo el dispositivo medios para seleccionar como dicho predictor temporal un vector de movimiento de entre los vectores de movimiento en un bloque de referencia de un fotograma de referencia diferente de dicho fotograma actual, siendo el bloque de referencia un bloque del fotograma de referencia contiguo a dicho bloque a codificar o un bloque del fotograma de referencia próximo al bloque contiguo, y dependiendo la selección del tamaño de bloque del bloque a codificar.

Según un vigésimo aspecto de la presente invención, se proporciona un dispositivo para almacenar vectores de movimiento para ser utilizados como predictores temporales de vectores de movimiento, que comprende medios de control para provocar el almacenamiento, de entre los vectores de movimiento del bloque a codificar de un fotograma actual, del vector de movimiento que tiene la misma posición que el predictor temporal determinado por el dispositivo de cualquiera de los aspectos decimosexto a decimonoveno de la presente invención, de modo que el vector de movimiento almacenado está disponible como un predictor temporal para la codificación de otro fotograma haciendo referencia a dicho fotograma actual, y para evitar el almacenamiento de otros vectores de movimiento del bloque a codificar del fotograma actual.

Según un vigesimoprimer aspecto de la presente invención, se proporciona un dispositivo para generar un conjunto de predictores de vectores de movimiento para un bloque a codificar de un fotograma actual, permitiéndose al conjunto incluir, por lo menos, un predictor espacial del vector de movimiento y, por lo menos, un predictor temporal del vector de movimiento, comprendiendo el dispositivo de uno cualquiera de los aspectos decimosexto a decimonoveno para determinar dicho predictor temporal del vector de movimiento o uno de dichos predictores temporales de vectores de movimiento del conjunto.

Según un vigesimosegundo aspecto de la presente invención, se proporciona un dispositivo para codificar una secuencia de imágenes digitales en una secuencia de bits, estando codificada, por lo menos una porción de una imagen, por compensación del movimiento con respecto a una porción de la imagen de referencia, comprendiendo el dispositivo de uno cualquiera de los aspectos decimosexto a decimonoveno para determinar un predictor temporal del vector de movimiento.

Según un vigesimotercer aspecto de la presente invención, se proporciona un dispositivo para codificar una secuencia de imágenes digitales en una secuencia de bits, estando codificada, por lo menos una porción de una imagen por compensación del movimiento con respecto a una porción de la imagen de referencia, comprendiendo el dispositivo de almacenamiento vectores de movimiento según el vigésimo aspecto de la presente invención.

Según un vigesimocuarto aspecto de la presente invención, se proporciona un dispositivo para codificar una secuencia de imágenes digitales en una secuencia de bits, estando codificada, por lo menos una porción de una imagen por compensación del movimiento con respecto a una porción de la imagen de referencia, comprendiendo el dispositivo de generación un conjunto de predictores de vectores de movimiento según el primer aspecto de la

presente invención.

Según un vigesimoquinto aspecto de la presente invención, se proporciona un dispositivo para descodificar una secuencia de bits que comprende una secuencia codificada de imágenes digitales, estando codificada, por lo menos una porción de una imagen por compensación del movimiento con respecto a una porción de la imagen de referencia, comprendiendo el dispositivo según uno cualquiera de los aspectos decimosexto a decimonoveno de determinación de un predictor temporal del vector de movimiento.

Según un vigesimosexto aspecto de la presente invención, se proporciona un dispositivo para descodificar una secuencia de bits que comprende una secuencia codificada de imágenes digitales, estando codificada, por lo menos una porción de una imagen por compensación del movimiento con respecto a una porción de la imagen de referencia, comprendiendo el dispositivo de almacenamiento vectores de movimiento según el vigésimo aspecto de la presente invención.

Según un vigesimoséptimo aspecto de la presente invención, se proporciona un dispositivo para descodificar una secuencia de bits que comprende una secuencia codificada de imágenes digitales, estando por lo menos una parte de una imagen codificada por compensación del movimiento con respecto a una imagen de referencia, comprendiendo el dispositivo de generación un conjunto de predictores de vectores de movimiento según el vigesimoprimer aspecto de la presente invención.

La presente invención se extiende asimismo a programas que, cuando son ejecutados en un ordenador o procesador, hacen que el ordenador o procesador lleve a cabo cualquiera de los procedimientos descritos anteriormente o que, cuando se cargan en un dispositivo programable, hacen que el dispositivo sea cualquiera de los dispositivos descritos anteriormente. Se puede proporcionar el propio programa, o éste contenido en un medio portador. El medio portador puede ser un medio de almacenamiento o de grabación, o puede ser un medio de transmisión tal como una señal. Un programa que realiza la presente invención puede ser transitorio o no transitorio.

Tal como se ha descrito anteriormente, en las realizaciones de la presente invención se tienen en cuenta las posiciones de los predictores espaciales para producir diversidad en el conjunto de predictores, en lugar de utilizar la mejor representación del conjunto de predictores.

Para la compresión de la memoria de MV, se obtienen resultados significativos en comparación con otras propuestas. Además, no hay impacto en la complejidad informática en comparación con la técnica anterior.

La invención se puede asimismo utilizar para el proceso de obtención de vectores de movimiento con el fin de crear diversidad en el conjunto de predictores.

La invención se puede utilizar asimismo para aumentar la eficiencia de codificación de la compresión de la memoria del vector de movimiento. De manera más precisa, en este caso es posible reducir la pérdida (en términos de eficiencia de codificación) generada por la compactación por bloques de la compresión de la memoria del vector de movimiento, o aumentar la relación de compresión (de la compresión de la memoria del vector de movimiento) con una eficiencia de codificación similar a la compactación por bloques clásica.

Por extensión, las realizaciones de la invención pueden mejorar la eficiencia de la codificación del esquema de AMVP y del modo de Fusión.

BREVE DESCRIPCIÓN DE LOS DIBUJOS

A continuación, se hará ahora referencia, a modo de ejemplo, a los dibujos adjuntos, en los que:

- la figura 1, descrita anteriormente en esta memoria, es un diagrama esquemático para su utilización en la explicación de una propuesta de la técnica anterior para reducir las necesidades de la memoria de vectores de movimiento;

- la figura 2, descrita también anteriormente en esta memoria, es un diagrama esquemático para su utilización en la explicación de una propuesta de la técnica anterior para mejorar el conjunto de predictores del vector de movimiento;

- la figura 3, descrita también anteriormente en esta memoria, es un diagrama esquemático para su utilización en la explicación de otra propuesta de la técnica anterior para mejorar el conjunto de predictores del vector de movimiento;

- la figura 4 muestra partes del dispositivo adecuadas para implementar un codificador o un descodificador según una realización de la presente invención;

- la figura 5 muestra un diagrama de bloques de partes de un codificador según una realización de la presente invención;

- la figura 6 muestra una secuencia de imágenes procesada mediante el codificador de la figura 5;

- la figura 7 muestra un diagrama de bloques de partes de un descodificador según una realización de la invención;

- la figura 8 es un diagrama esquemático para su utilización en la explicación de un procedimiento de determinación de un conjunto de predictores de vectores de movimiento que pueden ser utilizados por el codificador de la figura 5 y el descodificador de la figura 7;

- la figura 9 es un diagrama de flujo de las etapas llevadas a cabo por el codificador de la figura 5 cuando se utiliza el procedimiento de la figura 8;

- la figura 10 es un diagrama de flujo de las etapas llevadas a cabo por el descodificador de la figura 7 cuando se utiliza el procedimiento de la figura 8;

- la figura 11 es un diagrama de flujo de un procedimiento de compactación por bloques del vector de movimiento utilizable por el codificador de la figura 5 y el descodificador de la figura 7;

- la figura 12 es un diagrama esquemático para utilizar en la explicación de un principio que subyace al procedimiento de la figura 11; y

- las figuras 13A a 13C muestran matrices de distancias respectivas adecuadas para su utilización en el procedimiento de la figura 11.

DESCRIPCIÓN DETALLADA DE LAS REALIZACIONES

La **figura 4** muestra un diagrama del dispositivo -1000- adaptado para implementar un codificador según una realización de la presente invención o para implementar un descodificador según una realización de la presente invención. El dispositivo -1000- es, por ejemplo, un microordenador, una estación de trabajo o un dispositivo portátil ligero.

El dispositivo -1000- comprende un bus de comunicaciones -1113- al que se conectan preferentemente:

- una unidad central de proceso -1111-, tal como un microprocesador, indicada como CPU;

- una memoria de solo lectura (ROM, Read Only Memory) -1107- que almacena uno o varios programas informáticos para implementar la invención;

- una memoria de acceso aleatorio (RAM, Random Access Memory) -1112- que almacena código ejecutable del procedimiento de la invención y proporciona registros adaptados para grabar las variables y los parámetros necesarios para implementar el procedimiento de codificación de una secuencia de imágenes digitales y/o el procedimiento de descodificación de una secuencia de bits; y

- una interfaz de comunicación -1102- conectada a una red de comunicación -1103- sobre la que se transmiten los datos digitales a procesar.

Una memoria de vectores de movimiento (MVM, Motion Vector Memory) -1112a- forma parte de la RAM -1112- y se utiliza para almacenar los vectores de movimiento de fotogramas de referencia.

Opcionalmente, el dispositivo -1000- puede tener también los siguientes componentes:

- un medio de almacenamiento de datos -1104- tal como un disco duro, capaz de contener los programas que implementan la invención y los datos utilizados o producidos durante la implementación de la invención;

- una unidad de disco -1105- para un disco -1106-, estando la unidad de disco adaptada para leer datos del disco -1106- o para escribir datos en dicho disco;

- una pantalla -1109- para mostrar y/o servir como interfaz gráfica con el usuario, por medio de un teclado -1110- o de cualquier otro medio de señalización.

El dispositivo -1000- puede ser conectado a diversos periféricos, tales como por ejemplo una cámara digital -1100- o un micrófono -1108-, siendo cada uno conectado a una tarjeta de entrada/salida (no mostrada) con el fin de proporcionar datos multimedia al dispositivo -1000-.

El bus de comunicación permite la comunicación e interoperatividad entre los diferentes elementos incluidos en el dispositivo -1000- o conectados al mismo. El mapeo del bus no es limitativo y, en particular, la unidad central de

proceso es capaz de comunicar instrucciones a cualquier elemento del dispositivo -1000-, directamente o por medio de otro elemento del dispositivo -1000-.

El disco -1106- puede ser reemplazado por cualquier medio de información tal como, por ejemplo, un disco compacto (CD-ROM), regrabable o no, un disco ZIP o una tarjeta de memoria y, en términos generales, por medios de almacenamiento de información que pueden ser leídos por un microordenador o por un microprocesador, integrados o no en el dispositivo, opcionalmente extraíble y adaptado para almacenar uno o varios programas cuya ejecución permita la implementación del procedimiento de codificación de una secuencia de imágenes digitales y/o del procedimiento de descodificación de una secuencia de bits según la invención.

El código ejecutable puede ser almacenado bien en una memoria de solo lectura -1107-, en el disco duro -1104- o en un medio digital extraíble tal como por ejemplo un disco -1106- tal como se ha descrito anteriormente. Según una variante, el código ejecutable de los programas puede ser recibido por medio de la red de comunicación -1103-, a través de la interfaz -1102-, con el fin de ser almacenado en uno de los medios de almacenamiento del dispositivo -1000- antes de ser ejecutado, tal como el disco duro -1104-.

La unidad central de proceso -1111- está adaptada para controlar y dirigir la ejecución de las instrucciones o porciones de código del software del programa o programas según la invención, instrucciones que están almacenadas en uno de los medios de almacenamiento mencionados anteriormente. Cuando se conecta el equipo, el programa o programas que están almacenados en una memoria no volátil, por ejemplo, en el disco duro -1104- o en la memoria de solo lectura -1107-, son transferidos a la memoria de acceso aleatorio -1112-, que contiene entonces el código ejecutable del programa o programas, así como los registros para almacenar las variables y parámetros necesarios para la implementación de la invención.

En esta realización, el dispositivo es un dispositivo programable que utiliza software para implementar la invención. Sin embargo, de manera alternativa, la presente invención puede ser implementada en hardware (por ejemplo, en forma de un circuito integrado de aplicación específica, o ASIC, Application Specific Integrated Circuit).

La **figura 5** muestra un diagrama de bloques de un codificador -30- según una realización de la invención. El codificador está representado por módulos conectados, estando cada módulo adaptado para implementar, por ejemplo, en forma de instrucciones de programación para ser ejecutadas por la CPU -1111- del dispositivo -1000-, una etapa correspondiente de un procedimiento que implementa una realización de la invención.

Una secuencia original de imágenes digitales i_0 a i_n -301- es recibida como entrada por el codificador -30-. Cada imagen digital está representada por un conjunto de muestras, conocidas como píxeles.

Una secuencia de bits -310- es emitida por el codificador -30-.

La secuencia de bits -310- comprende una serie de unidades de codificación o segmentos, comprendiendo cada segmento una cabecera de segmento para la codificación de los valores de los parámetros de codificación utilizados para codificar el segmento y el cuerpo del segmento, que comprende datos de video codificados. En HEVC, estos segmentos están divididos en Unidades máximas de codificación (LCU, Largest Coding Units) no superpuestas, generalmente bloques de tamaño 64 píxeles x 64 píxeles. Cada LCU puede, a su vez, ser dividida iterativamente en Unidades de codificación (CU, Coding Units) menores de tamaño variable, utilizando una descomposición de árbol cuaternario. Cada CU puede ser además dividida en un máximo de 2 Unidades de partición (PU, Partition Units) rectangulares simétricas.

La **figura 6** muestra la secuencia -301- de imágenes digitales i, segmentos -103-, LCU -104-, CU -105-, PU -106- y TU -107-. Una TU (Unidad transformada, Transform Unit) se define independientemente de la PU para transformación y cuantificación en la CU.

Se debe observar que, en la siguiente descripción se utiliza el término “bloque” en lugar de la terminología específica CU y PU utilizada en HEVCA. CU o PU es un bloque de píxeles.

Volviendo a la figura 5, las imágenes digitales de entrada i son divididas en bloques mediante el módulo -302-. Estos bloques son porciones de imagen y pueden ser de tamaños variables (por ejemplo, 4x4, 8x8, 16x16, 32x32, 64x64).

Mediante el módulo -306-, se selecciona un modo de codificación para cada bloque de entrada. El módulo -306- se describe más adelante.

Existen dos familias de modos de codificación, codificación de predicción espacial o Intra-codificación, y codificación de predicción temporal o Inter-codificación. Los modos de codificación posibles están probados.

El módulo -303- implementa la Intra-predicción, en la que el bloque determinado a codificar es predicho por medio de un predictor “Intra”, un bloque de píxeles construido a partir de la información ya codificada, por ejemplo, calculada a partir de los píxeles del entorno de dicho bloque a codificar. Una indicación del predictor Intra

seleccionado y la diferencia entre el bloque determinado y su predictor se codifica si se selecciona la Intra-codificación mediante el módulo -306-.

La predicción temporal se implementa mediante los módulos -304- y -305-. En primer lugar, se selecciona una imagen de referencia entre un conjunto de imágenes de referencia -316-, y se selecciona una porción de la imagen de referencia, llamada también zona de referencia, que es la zona más próxima al bloque determinado que se debe codificar, mediante el módulo -304- de estimación del movimiento. Generalmente, el módulo de estimación del movimiento -304- utiliza un algoritmo de ajuste de bloques (BMA, Block Matching Algorithm).

Con respecto a la "Inter" codificación, son posibles dos tipos de predicción. La Mono-predicción (tipo P) consiste en predecir el bloque haciendo referencia a una zona de referencia de una imagen de referencia. La Bi-predicción (tipo B) consiste en predecir el bloque haciendo referencia a dos zonas de referencia de una o dos imágenes de referencia. En el módulo -304- se realiza una estimación del movimiento entre el bloque actual y las imágenes de referencia -316- con el fin de identificar, en una o varias de estas imágenes de referencia, un bloque de píxeles (tipo P), o varios (tipo B), para su utilización como predictores de este bloque actual. En el caso en que se utilicen varios predictores de bloque (tipo B), estos son fusionados para generar un solo bloque de predicción. Las imágenes de referencia utilizadas son imágenes en la secuencia de video que han sido ya codificadas y después reconstruidas (mediante descodificación).

La diferencia entre la zona de referencia seleccionada y el bloque determinado, llamado también bloque residual, se calcula mediante el módulo de compensación del movimiento -305-. La zona de referencia seleccionada se indica mediante un vector de movimiento.

La información relativa al vector de movimiento y al bloque residual se codifica si se selecciona Inter-predicción mediante el módulo -306-. Para reducir más la tasa de bits, el vector de movimiento es codificado por diferencia con respecto a un predictor del vector de movimiento. Se obtiene un conjunto de predictores del vector de movimiento, llamados también predictores de información del movimiento, a partir del campo de vectores de movimiento -318- mediante una predicción del vector de movimiento y del módulo de codificación -317-. La operación del módulo -317- se describirá a continuación con detalle, con respecto a las figuras 8 y 9.

El módulo de selección del modo de codificación -306- utiliza un criterio de coste de codificación, tal como un criterio de distorsión de velocidad, para determinar cuál es el mejor modo entre los modos de predicción Intra e Inter. Se aplica una transformada -307- al bloque residual, los datos transformados obtenidos se cuantifican a continuación mediante el módulo -308- y se codifican por entropía mediante el módulo -309-. La transformada se aplica a la Unidad de transformada (TU) mencionada anteriormente que está incluida en un bloque. Una TU puede ser dividida además en varias TU más pequeñas utilizando una llamada descomposición de Árbol cuaternario residual (RQT, Residual Árbol cuaternario), tal como se muestra en la figura 6. En HEVC, generalmente se utilizan 2 o 3 niveles de descomposiciones, y los tamaños autorizados de transformadas son de 32x32, 16x16, 8x8 y 4x4. La base de la transformada se deriva a partir de una transformada del coseno discreta, DCT (Discrete Cosine Transform).

Finalmente, el bloque residual codificado del bloque actual a codificar es introducido en la secuencia de bits -310-, junto con la información relativa al predictor utilizado. Para los bloques codificados en modo 'SALTO', solo se codifica una referencia al predictor en la secuencia de bits, sin ningún bloque residual.

Para calcular los predictores "Intra" o realizar una estimación del movimiento para los predictores "Inter", el codificador realiza una descodificación de los bloques ya codificados por medio de un llamado bucle de "descodificación" -311- - -315-. Este bucle de descodificación hace posible la reconstrucción de los bloques e imágenes a partir de residuos transformados cuantificados.

El residuo transformado cuantificado es descuantificado en el módulo -311- mediante la aplicación de la cuantificación inversa a la proporcionada por el módulo -308-, y es reconstruido en el módulo -312- mediante la aplicación de la transformada inversa a la del módulo -307-.

Si el residuo procede de una codificación "Intra", entonces en el módulo -313- el predictor "Intra" utilizado se añade a este residuo con el fin de obtener un bloque reconstruido correspondiente al bloque original modificado por las pérdidas resultantes de una transformación con pérdidas, en este caso, operaciones de cuantificación.

Si el residuo, por otro lado, procede de una codificación "Inter", los bloques señalados por los vectores de movimiento actuales (estos bloques pertenecen a las imágenes de referencia -316- indicadas por los índices de imagen actuales) son fusionados y, añadidos a continuación a este residuo descodificado en el módulo -314-. De esta manera, se obtiene el bloque original, modificado por las pérdidas resultantes de las operaciones de cuantificación.

Un filtro de bucle -315- final se aplica a la señal reconstruida con el fin de reducir los efectos creados por la fuerte cuantificación de los residuos obtenidos, y para mejorar la calidad de la señal. El filtro de bucle comprende dos etapas, un filtro de "desbloqueo" y un filtrado lineal. El filtrado de desbloqueo suaviza los límites entre los bloques

con el fin de atenuar visualmente estas altas frecuencias creadas por la codificación. El filtrado lineal mejora además la señal utilizando coeficientes de filtro determinados adaptativamente en el codificador. El filtrado mediante el módulo -315- se aplica, de este modo, a una imagen cuando todos los bloques de píxeles de esta imagen han sido descodificados.

Las imágenes filtradas, también llamadas imágenes reconstruidas, se almacenan a continuación como imágenes de referencia -316- con el fin de permitir la realización de predicciones "Inter" sucesivas durante la compresión de las imágenes siguientes de la secuencia de video actual.

En el contexto de HEVC, es posible utilizar varias imágenes de referencia -316- para la estimación y la compensación del movimiento de la imagen actual. En otras palabras, la estimación del movimiento se realiza sobre N imágenes. De este modo, los mejores predictores "Inter" del bloque actual, para la compensación del movimiento, se seleccionan en algunas de las múltiples imágenes de referencia. En consecuencia, dos bloques adyacentes pueden tener dos bloques de predictores que proceden de dos imágenes de referencia distintas. Esta es, en concreto, la razón por la que, en la secuencia de bits comprimida, se indica el índice de la imagen de referencia (además del vector de movimiento) utilizado para el bloque del predictor.

La utilización de múltiples imágenes de referencia es tanto una herramienta para evitar los errores como una herramienta para mejorar la eficacia de la compresión. El grupo VCEG recomienda limitar el número de imágenes de referencia a cuatro.

La **figura 7** muestra un diagrama de bloques de un descodificador -40- según una realización de la invención. El descodificador se representa mediante módulos conectados, estando cada módulo adaptado para implementar, por ejemplo, en forma de instrucciones de programación para ser ejecutadas por la CPU -1111- del dispositivo -1000-, una etapa correspondiente de un procedimiento que implementa una realización de la invención.

El descodificador -40- recibe una secuencia de bits -401- que comprende unidades de codificación, estando cada unidad compuesta por una cabecera que contiene información sobre los parámetros de codificación y un cuerpo que contiene los datos de video codificados. Tal como se ha explicado anteriormente con respecto a la figura 5, los datos de video codificados se codifican por entropía, y los índices de los predictores del vector de movimiento se codifican, para un bloque determinado, en un número de bits predeterminado. Los datos de video codificados recibidos son descodificados por entropía mediante un módulo -402-, descuantificados mediante un módulo -403- y a continuación se aplica una transformada inversa mediante un módulo -404-.

En concreto, cuando los datos del video codificado recibidos corresponden a un bloque residual de un bloque actual a codificar, el descodificador descodifica también la información de la predicción del movimiento a partir de la secuencia de bits, con el fin de encontrar la zona de referencia utilizada por el codificador.

Un módulo -410- aplica la descodificación del vector de movimiento para cada bloque actual codificado mediante la predicción de movimiento. De manera similar al módulo -317- del descodificador de la figura 5, el módulo de descodificación -410- del vector de movimiento utiliza información (el campo de los vectores de movimiento -411-, que es similar al campo de los vectores de movimiento -318- de la figura 5) relativa a los vectores de movimiento del fotograma actual y de los fotogramas de referencia para generar un conjunto de predictores de vectores de movimiento. La operación del módulo -410- se describirá con más detalle a continuación, haciendo referencia a la figura 10. Si la secuencia de bits se recibe sin pérdidas, el descodificador genera exactamente el mismo conjunto de predictores de vectores de movimiento que el codificador. Una vez que se ha obtenido el índice del predictor del vector de movimiento para el bloque actual, si no se han producido pérdidas, el valor real del vector de movimiento asociado con el bloque actual puede ser descodificado y proporcionado a un módulo -406- que aplica compensación del movimiento inverso. La zona de referencia indicada por el vector de movimiento descodificado es extraída de una imagen de referencia entre las imágenes de referencia almacenadas -408-, y proporcionada también al módulo -406- para que éste pueda aplicar la compensación del movimiento inversa.

En caso de que se haya aplicado Intra-predicción, un módulo -405- aplica una Intra-predicción inversa.

Como resultado de la descodificación según el modo Inter o Intra, se obtiene un bloque descodificado. Se aplica un filtro de desbloqueo mediante un módulo -407-, de manera similar al filtro de desbloqueo -315- aplicado en el codificador. Finalmente, el descodificador -40- proporciona una señal de video descodificada -409-.

La **figura 8** es un diagrama esquemático para su utilización en la explicación de la generación del conjunto de predictores de vectores de movimiento o de candidatos de vector de movimiento en la implementación de HEVC actual.

En el diseño de HEVC actual, los vectores de movimiento son codificados mediante codificación predictiva, utilizando una serie de vectores de movimiento. Este procedimiento se denomina Predicción avanzada del vector de movimiento (AMVP, Advanced Motion Vector Prediction) y ha sido adaptado para considerar el nuevo contexto de HEVC con una estructura de bloque grande. Este esquema se aplica a los modos de Salto, Inter y Fusión.

El procedimiento permite la selección del mejor predictor para un conjunto determinado, en el que el conjunto está compuesto de vectores de movimiento espaciales y vectores de movimiento temporales. El número óptimo de predictores espaciales y temporales se está evaluando todavía en el proceso de estandarización de HEVC. No obstante, en la fecha de presentación de la presente solicitud, la implementación actual incluye 2 predictores espaciales y un predictor temporal contiguo para los modos Salto e Inter, y 4 predictores espaciales y un predictor temporal para el modo Fusión. La presente invención no está limitada a ser utilizada con la implementación actual de AMVP. La implementación de AMVP puede cambiar con respecto a la actual que se describe a continuación, pero se considera que las realizaciones de la presente invención que se van a describir a continuación proporcionarán los mismos efectos y resultados ventajosos con otras implementaciones que pueden ser adoptadas.

Además, en la publicación JCTVC-D072 mencionada en la introducción, se propuso utilizar más predictores temporales en lugar de utilizar solo uno en la versión actual. La invención se puede aplicar asimismo con esta modificación.

En el conjunto de predictores representado en la figura 8, ambos vectores de movimiento espaciales se eligen entre los bloques de arriba, y los de la izquierda, incluyendo los bloques de las esquinas superiores y el bloque de la esquina izquierda.

El predictor izquierdo se selecciona entre los bloques I, H, G, F. El predictor del vector de movimiento se considera disponible si el vector existe, y si el índice del fotograma de referencia es el mismo que el índice del fotograma de referencia del bloque actual (lo que quiere decir que el vector de movimiento utilizado como predictor apunta al mismo fotograma de referencia que el vector de movimiento del bloque actual). La selección se realiza por medio de una búsqueda de abajo (I) arriba (F). El primer predictor que cumple los criterios de disponibilidad anteriores se selecciona como predictor izquierdo (solo se añade un predictor izquierdo al conjunto de predictores). Si ningún predictor cumple los criterios, el predictor izquierdo se considera no disponible.

Un inter-bloque puede ser mono-predictivo (tipo P) o bi-predictivo (tipo B). En un fotograma-P, los inter-bloques son solo de tipo P. En un fotograma-B, los inter-bloques son de tipo P o B. En un inter-bloque de tipo P, se utiliza una lista L0 de fotogramas de referencia. Su vector de movimiento hace referencia a un fotograma de referencia de esta lista. Un índice de referencia se asocia, por lo tanto, con el vector de movimiento. En un inter-bloque de tipo P, se utilizan dos listas L0 y L1 de los fotogramas de referencia. Uno de sus dos vectores de movimiento hace referencia a un fotograma de referencia de la lista L0, y el otro de sus dos vectores de movimiento hace referencia a un fotograma de referencia de la lista L1. Un índice de referencia se asocia por lo tanto con cada uno de ambos vectores de movimiento.

La no existencia de un vector de movimiento significa que el bloque relacionado estaba Intra-codificado o que no existe ningún vector de movimiento en la lista con el que esté asociado el vector de movimiento codificado. Por ejemplo, para un bloque en un fotograma B, si un bloque vecino tiene solo un vector de movimiento en la lista ‘L1’ y el vector de movimiento actual está en ‘L0’, el vector de movimiento vecino se considera no existente para la predicción del vector de movimiento actual.

El predictor superior se selecciona entre los bloques E, D, C, B, A, de nuevo como resultado de una búsqueda, en este caso de derecha a izquierda. El primer vector de movimiento, de derecha a izquierda, que cumple los criterios de disponibilidad definidos anteriormente (si el predictor existe y tiene el mismo fotograma de referencia que el vector de movimiento actual) se selecciona como predictor superior. Si ningún predictor cumple los criterios, el predictor superior se considera no disponible.

El predictor temporal del vector de movimiento procede del fotograma de referencia más cercano cuando los fotogramas no están ordenados de manera diferente para la codificación y para la visualización (están codificados sucesivamente sin reordenación). Esta configuración corresponde a una configuración de bajo retardo (no hay retardo entre el proceso de descodificación y el proceso de visualización). En el caso de los fotogramas B, se consideran 2 vectores de movimiento para el bloque contiguo. Uno está en la primera lista “L0” de imágenes de referencia y otro en la segunda lista “L1” de imágenes de referencia. Si ambos vectores de movimiento existen, se selecciona el movimiento que tiene la distancia temporal más corta. Si ambos predictores tienen la misma distancia temporal, se selecciona el movimiento forma de “L0”. El vector de movimiento contiguo seleccionado, a continuación, se escala, si es necesario, según la distancia temporal entre la imagen de referencia y la imagen que contiene el fotograma que se va a codificar. Si no existe ningún predictor contiguo, el predictor se considera no disponible.

Para la codificación de fotogramas B jerárquicos, que implica la reordenación de los fotogramas y, por lo tanto, más retardo de descodificación, se pueden considerar 2 vectores de movimiento contiguos. Ambos proceden del fotograma de referencia futuro. Se selecciona el vector de movimiento que atraviesa el fotograma actual. Si ambos predictores atraviesan el fotograma actual, se selecciona el bloque que contiene el vector de movimiento que tiene la distancia temporal más corta. Si ambos predictores tienen la misma distancia temporal, entonces se selecciona el vector de movimiento de la primera lista “L0”. A continuación, el vector de movimiento contiguo seleccionado como el predictor temporal del vector de movimiento se escala, si es necesario, según la distancia temporal entre la imagen

de referencia y la imagen que contiene el bloque a descodificar. Si no existe ningún predictor contiguo, el predictor se considera no disponible.

Para los casos tanto de bajo retardo como jerárquicos, cuando el bloque contiguo está dividido en una serie de particiones (potencialmente, el bloque contiguo contiene una serie de vectores de movimiento), el vector de movimiento seleccionado procede de la partición central, tal como se ha mencionado en la introducción a la presente memoria, véase Jung, G. Clare, (Orange Labs), "Temporal MV predictor modification for MV-Comp, Skip, Direct and Merge schemes" ("Modificación de predictores temporales de MV para esquemas de MV-Comp, salto, dirección y fusión"), JCTVC-S164, Daegu, KR, 20-28 de enero de 2011, que propone utilizar un predictor temporal centrado, y el documento WO 2011/001077 A.

Como resultado de este procedimiento de generación de predictores de vectores de movimiento, el conjunto de predictores generado puede contener 0, 1, 2 o 3 predictores. Si no se ha podido incluir ningún predictor en el conjunto, el vector de movimiento no se predice. Las componentes tanto vertical como horizontal se codifican sin predicción. (Esto corresponde a una predicción mediante un predictor igual al valor cero). En la implementación de HEVC actual, el índice del predictor es igual a 0.

El modo Fusión es una Inter-codificación concreta, similar al modo Salto habitual bien conocido por los expertos en la materia. La principal diferencia en comparación con el modo Salto habitual es que el modo Fusión propaga el valor del índice del fotograma de referencia, la dirección (Bi-direccional o uni-direccional) y la lista (con la dirección uni-direccional) de los predictores de vectores de movimiento al bloque predicho. El modo Fusión utiliza un predictor del vector de movimiento y su índice del fotograma de referencia, a menos que el predictor sea un predictor temporal en el que el fotograma de referencia considerado es siempre el fotograma de referencia inmediatamente anterior, llamado también Ref0 (y siempre predicción bi para fotogramas B). De esta manera, los predictores de bloques (los bloques copiados) proceden de los fotogramas de referencia señalados por los predictores del vector de movimiento.

La ordenación de candidatos en el conjunto es importante para reducir la sobrecarga de señalizar el mejor predictor del movimiento en el conjunto de predictores. La ordenación del conjunto se adapta en función del modo de predicción actual para situar el predictor más probable del movimiento en la primera posición, dado que se produce una mínima sobrecarga si el primer candidato se elige como mejor predictor. En la implementación actual de HEVC, el predictor temporal es la primera posición.

La sobrecarga de señalizar el índice del mejor predictor se puede reducir más minimizando el número de candidatos en el conjunto. Los vectores de movimiento duplicados simplemente se eliminan del conjunto.

Para el caso concreto del modo Fusión, el proceso de supresión tiene en cuenta los valores del vector de movimiento y su fotograma de referencia. Según esto, para determinar si dos predictores son predictores duplicados, las dos componentes del vector de movimiento y su índice de referencia se comparan para ambos predictores y, solo si estos tres valores son iguales, se elimina un predictor del conjunto. Para un fotograma B, este criterio de igualdad se extiende a la dirección y las listas. Así que dos predictores se consideran predictores duplicados si ambos utilizan la misma dirección, las mismas listas (L0, L1 o L0 y L1), los índices del fotograma de referencia y tienen los mismos valores de los vectores de movimiento (MV_L0 y MV_L1 para predicción bi).

En AMVP, la señalización del índice depende del resultado del proceso de supresión de predictores del vector de movimiento descrito anteriormente. De hecho, el número de bits asignados a la señalización depende del número de vectores de movimiento que quedan tras la supresión. Por ejemplo, si al final del proceso de supresión solo queda un vector de movimiento, no es necesario ninguna sobrecarga para señalizar el índice del predictores del vector de movimiento, dado que el índice puede ser obtenido fácilmente por el descodificador. La Tabla 1 que sigue muestra la palabra de código para cada codificación de índice según el número de predictores tras el proceso de supresión.

Tabla 1

	Palabra de código según el número N de predictores en el conjunto				
Índice	N=1	N=2	N=3	N=4	N=5
0	(inferido)	0	0	0	0
1		1	10	10	10
2			11	110	110
3				111	1110
4					1111

La **figura 9** es un diagrama de flujo para su utilización en la explicación del esquema de operación de la AMVP en el lado del codificador. Las operaciones de la figura 9 se llevan a cabo mediante el módulo -317- de la figura 5, excepto

donde se indique otra cosa, y se puede considerar que este módulo -317- comprende los módulos -603-, -605-, -607-, -610- y -615- de la figura 9. El campo de los vectores de movimiento -601- en la figura 9 corresponde al campo de los vectores de movimiento -318- de la figura 5. El módulo de codificador entrópico -612- en la figura 9 corresponde al módulo de codificador entrópico -309- en la figura 5. Todas las operaciones en la figura 9 pueden ser implementadas en software y ejecutadas por la unidad central de proceso -1111- del dispositivo -1000-.

Un módulo de generación de predictores del vector de movimiento -603- recibe un índice del fotograma de referencia -613- del vector de movimiento actual a codificar, y recibe también el campo de los vectores de movimiento -601-. El módulo -603- genera un conjunto de predictores de vectores de movimiento -604- tal como se ha descrito anteriormente haciendo referencia a la figura 8, teniendo en cuenta el índice del fotograma de referencia -613-. A continuación, el proceso de supresión es aplicado por un módulo -605-, tal como se ha descrito también anteriormente haciendo referencia a la figura 8. El módulo -605- produce un conjunto reducido de predictores del vector de movimiento -606-. Se entrega asimismo el número de predictores del vector de movimiento -616- en el conjunto reducido -606-. Un módulo -607- recibe el vector de movimiento a codificar -602- y aplica una selección de distorsión de velocidad (RD, Rate Distortion) del mejor predictor del conjunto reducido de predictores de vectores de movimiento -606-. Si se selecciona un mejor predictor, el módulo-607- produce un índice de predictor del vector de movimiento -608- y el predictor del vector de movimiento seleccionado -609-. A continuación, un módulo -610- forma la diferencia entre el vector de movimiento a codificar -602- y el predictor del vector de movimiento seleccionado -609-. Esta diferencia es un vector de movimiento residual -611-. Este vector de movimiento residual es a continuación codificado entrópicamente en un módulo -612-. Un módulo -614- convierte el índice del predictor del vector de movimiento -608- en una palabra de código -615- según el número de predictores -616- en el conjunto reducido de predictores del vector de movimiento -606-, tal como se ha descrito anteriormente haciendo referencia a la Tabla 1. Tal como se ha descrito anteriormente, si este conjunto contiene solo un predictor, no se transmite ningún índice al lado del descodificador y no se genera ninguna palabra de código. Si el conjunto contiene uno o varios predictores, la palabra de código se genera en el módulo -614- y, a continuación, se codifica por entropía en el módulo -612-.

La **figura 10** es un diagrama de flujo para su utilización explicando el funcionamiento del esquema de la AMVP en el lado del descodificador. Las operaciones de la figura 10 son realizadas mediante el módulo -410- en la figura 7, excepto donde se indique otra cosa, y se puede considerar que este módulo -410- comprende los módulos -702-, -704-, -711- y -715- en la figura 10. Un campo de vectores de movimiento -701- en la figura 10 corresponde al campo de vectores de movimiento -411- en la figura 7. Un módulo descodificador entrópico -706- en la figura 10 corresponde al módulo descodificador entrópico -402- en la figura 7. Todas las operaciones en la figura 10 pueden ser implementadas en software y ejecutadas por la unidad central de proceso -1111- del dispositivo -1000-.

Un módulo-702- recibe el campo de vectores de movimiento -701- del fotograma actual y de los fotogramas descodificados anteriormente. El módulo -702- recibe también un índice del fotograma de referencia -713- del vector de movimiento actual a descodificar. El módulo -702- genera un conjunto de predictores de movimiento -703- en base al campo de vectores de movimiento -701- y al índice del fotograma de referencia -713-. Este proceso es el mismo que el descrito en relación con el módulo -603- en el lado del codificador. A continuación, un módulo -704- aplica un proceso de supresión. Este proceso es el mismo que el descrito en relación con el módulo -605- en el lado del codificador. El módulo -704- produce un conjunto reducido de predictores de vectores de movimiento -708-. Se entrega asimismo el número de predictores del vector de movimiento -716- en el conjunto reducido -708-.

El módulo de descodificador entrópico -706- extrae un vector de movimiento residual -707- de la secuencia de bits -705- y lo descodifica. El número de predictores -716- en el conjunto reducido -708- es utilizado a continuación por el módulo -706- para extraer (si es necesario) la palabra de código del predictor del vector de movimiento -714-. Esta palabra de código (si existe) es transformada por el módulo -715- en un valor de índice de predictor -709- según el número de los predictores -716- en el conjunto reducido, utilizando la Tabla 1 anterior para la conversión. El predictor del vector de movimiento -710- se extrae a continuación del conjunto reducido -708- según el valor del índice del predictor -709-. Un módulo -711- añade el predictor del vector de movimiento al residuo del movimiento -707- con el fin de producir el vector de movimiento descodificado -712-.

A partir de lo anterior, es evidente que para cada fotograma que se utiliza como fotograma de referencia para la obtención del predictor del vector de movimiento contiguo, es necesario almacenar junto al codificador y el descodificador sus vectores de movimiento correspondientes. Esto conduce a que el tamaño de la memoria de vectores de movimiento sea importante, considerando en primer lugar la granularidad del mapeo del movimiento (en el diseño de HEVC actual, el tamaño mínimo de bloque en el modo Inter es 4x4) y, en segundo lugar, que existen dos vectores por bloque de movimiento para el SEGMENTO_B. Se estima que para imágenes de 4Kx2K de resolución, y utilizando una granularidad de un conjunto de vectores de movimiento por bloque de 4x4, son necesarios 26 Mbits por fotograma. Esta memoria de vectores de movimiento tiene que ser una memoria rápida, y forma parte habitualmente de la RAM, por ejemplo, la RAM -1112- en la figura 4. Esto es costoso, sobre todo para dispositivos portátiles.

Primera realización

A continuación, se describirá una primera realización de la presente invención.

En la primera realización, tal como en las propuestas JCTVC-C257 y JCTVC-D072 mencionadas en la introducción y mostradas en la figura 1, la posición de un bloque se utiliza para la compactación por bloques de un conjunto de NxN vectores de movimiento. El vector de movimiento del bloque correspondiente a esta posición de un bloque es almacenado como vector de movimiento representativo para todo el conjunto de NxN. Estas propuestas de la técnica anterior utilizan exclusivamente la posición del bloque superior izquierdo, tal como se muestra en la figura 1. No obstante, esta manera de compactación no tiene en cuenta que el predictor del vector de movimiento contiguo se utiliza en un esquema competitivo con otros predictores espaciales. Como resultado, aunque estas propuestas de la técnica anterior ofrecen una importante reducción en las necesidades de la memoria de vectores de movimiento (por ejemplo, un factor de reducción de 16 en el ejemplo de la figura 1), implican una penalización en términos de eficiencia de la codificación. Por ejemplo, se estima que la utilización de una reducción de 16 veces en las necesidades de memoria tiene como resultado una pérdida de 0,5% en la eficiencia de la codificación en TMuC v0.9. Esto ocurre porque a menudo el vector de movimiento en la posición del bloque superior izquierdo del bloque contiguo es similar a los predictores espaciales de los bloques vecinos del bloque actual a codificar. Esto significa que el predictor del vector de movimiento contiguo no aporta suficiente diversidad al conjunto de predictores en conjunto.

Utilizando una posición de bloque diferente dentro del bloque contiguo, o incluso utilizando una posición de bloque en otro bloque vecino del bloque contiguo, se puede obtener un mayor grado de diversidad. El efecto de esto es que, a pesar de seguir consiguiendo la misma reducción en las necesidades de memoria del vector de movimiento que las propuestas de la técnica anterior mencionadas anteriormente, la presente realización no incurre en ninguna penalización, o en ninguna importante, en la eficiencia de la codificación en comparación con un sistema en el que se almacenan todos los vectores de movimiento y no se utiliza ninguna compactación por bloques.

Las realizaciones de la presente invención reconocen que la diversidad del conjunto de predictores que incluye una mezcla de predictores espaciales y temporales depende de las posiciones relativas de los bloques de los que proceden los predictores. Cuanto mayor sea la distancia de un predictor a otro, mayor es la probabilidad de que los predictores sean diferentes. De este modo, puede ser preferible seleccionar como posición del bloque en el bloque contiguo una posición de bloque más alejada de las posiciones de bloque de los predictores espaciales. Por supuesto, la distancia no debe ser tan grande que el predictor temporal ya no sea representativo de los vectores de movimiento en el bloque contiguo. Después de todo, el bloque contiguo se selecciona como fuente del predictor temporal precisamente porque es contiguo al bloque actual. Según esto, lo que se necesita es elegir una posición de bloque para el predictor temporal que permita diversidad controlada entre ese predictor y los predictores espaciales.

Será evidente que en AMVP los predictores espaciales se seleccionan entre ciertas posiciones (las posiciones A a I mostradas en la figura 8) según los criterios de disponibilidad. Esto significa que, dependiendo de la disponibilidad, se pueden seleccionar 0, 1 o 2 predictores espaciales. La posición del bloque a elegir para la compactación por bloques del bloque contiguo no puede tener en cuenta la selección o selecciones de predictores espaciales reales, dado que se lleva a cabo antes que la selección o selecciones de predictores espaciales. Para posibilitar tener en cuenta la selección o selecciones del predictores espaciales cuando se elige la posición de bloque del bloque contiguo sería necesario almacenar los vectores de movimiento en todas las posiciones de bloque en el bloque contiguo, lo que es incompatible con la reducción de las necesidades de la memoria de vectores de movimiento.

A este respecto, no obstante, cuando el objetivo no es reducir las necesidades de memoria sino mejorar la eficiencia en la codificación, es posible tener en cuenta la selección o selecciones de predictores espaciales cuando se elige la posición de bloque del bloque contiguo, tal como se describirá en las realizaciones siguientes.

En la implementación actual del modelo de prueba de HEVC, para un bloque de Salto o Inter actual, se utilizan dos predictores espaciales, tal como se ha descrito anteriormente haciendo referencia a la figura 8, mientras que se utilizan cuatro predictores espaciales para el modo de Fusión. No obstante, el modo de Fusión se selecciona con menor frecuencia que los modos Salto e Inter. Según esto, en la presente realización se supone que se tienen 2 predictores espaciales, por ejemplo, el predictor izquierdo y los predictores superiores, tal como se muestra en la figura 8. La presente realización se puede extender fácilmente cuando se utilizan más o menos predictores en el conjunto de predictores espaciales. Además, la descripción de esta realización se refiere a la compresión de la memoria de vectores de movimiento sobre la base de una compactación por bloques de 4x4. Esto se puede extender fácilmente a todos los demás tamaños utilizados para la compactación por bloques.

La **figura 11** muestra las etapas de un procedimiento de determinación de la posición de un bloque 4x4 cuyo vector de movimiento se seleccionará como el único vector de movimiento en el proceso de compresión de la memoria de vectores de movimiento de los primeros ejemplos. Tal como se ha explicado anteriormente, esto se puede considerar un proceso de compactación por bloques, dado que tiene como resultado que se selecciona un vector de movimiento como representación del bloque como conjunto. Los vectores de movimiento no seleccionados del bloque (por ejemplo 15 vectores de movimiento en el caso presente), que tienen posiciones de bloque distintas de la

posición de bloque determinada, no necesitan ser almacenados, lo que tiene como resultado una significativa compresión de la memoria de vectores de movimiento (por ejemplo, en un factor de 16 en el presente caso).

A este respecto, además de los propios vectores de movimiento (que tienen cada uno componentes vertical y horizontal y un índice de fotograma de referencia), la selección de la posición del bloque se podría aplicar a todos los demás datos necesarios para el proceso de derivación del predictor contiguo (la implementación actual de la compresión de la memoria de MV no tiene en cuenta los otros datos). Estos datos son: el modo (Inter/Intra), el índice del fotograma de referencia del vector de movimiento, la lista 'L0' 'L1' de imágenes de referencia y la dirección Bi o uni-direccional.

En una primera etapa S1001, se recibe una lista de posiciones potenciales de los predictores espaciales. Estas posiciones son posiciones relativas con respecto al bloque actual. Con el propósito de compactación por bloques, se puede considerar que el tamaño (en píxeles) del bloque actual es 16x16 cuando se aplica la compactación del bloque de 4x4 de vectores de movimiento tal como se representa en la figura 1. De este modo, las posiciones potenciales A, B, C, D, E para el predictor superior y las posiciones potenciales F, G, H, I para el predictor izquierdo de la figura 8 se reciben mediante el algoritmo. Por sencillez, se puede considerar que los bloques vecinos tienen un tamaño de CU mínimo igual a 4x4 (en píxeles) tal como se representa en la figura 12. En ese caso, todos los bloques de predictores espaciales tienen el mismo tamaño, que es el tamaño de los bloques que se compactarán en la compresión de la memoria MV. Tal como se representa en la figura 12, las posiciones potenciales del predictor superior son TP1-TP6, y las posiciones potenciales del predictor izquierdo son LP1-LP5. A continuación, en la etapa S1002, el algoritmo determina la posición media para cada predictor espacial. En el presente ejemplo simplificado de la figura 12, la posición media para el predictor izquierdo es la posición del bloque LP3, y la posición media para el predictor superior es la frontera entre los bloques TP3 y TP4. Dado que la posición media para el predictor superior es la frontera entre los bloques TP3 y TP4, las posiciones de ambos bloques TP3 y TP4 se consideran en el proceso siguiente, mientras que para el predictor izquierdo solo se considera la posición del bloque LP3 en el proceso siguiente. A continuación, se genera una matriz de distancias para cada predictor en la etapa S1003. Esta matriz tiene una celda para cada posición posible de bloque del bloque que se compacta y contiene en cada celda la distancia mínima entre la posición media del predictor correspondiente (el predictor superior o el predictor izquierdo, según sea el caso) y la posición de bloque posible de la celda correspondiente. Las posiciones de bloque posibles se numeran de 1 a 16 en la figura 12, según una representación en zigzag adoptada para la memoria de vectores de movimiento. La construcción de esta matriz se explica a continuación haciendo referencia a las figuras 13A a 13C. A continuación, en la etapa S1004 las matrices de distancia de ambos predictores se suman celda a celda con el fin de obtener una única matriz de distancias global. En la etapa S1005, se identifica el valor máximo en esta matriz de distancias global. La posición correspondiente a este valor máximo es la posición utilizada para la compactación por bloques de la compresión de la memoria de MV (1006).

Para generar la matriz de distancias mínima entre la posición media y cada posición de bloque posible numerada de 1 a 16 en la figura 12, se puede considerar el número de bloques a ser atravesados para llegar a la posición de bloque cuando se empieza desde la posición media para el predictor correspondiente. Por ejemplo, con esta medida, tal como se representa en la figura 13A, que muestra la matriz de distancias para el predictor izquierdo, la distancia mínima entre la posición media correspondiente LP3 y la posición de bloque numerada 7 en la figura 12 es 4. Esto es consecuencia de que, para alcanzar esta posición, se necesita atravesar las posiciones de bloque numeradas 9, 3, 4 y 7, que son 4 bloques. En esta medida, no se ha considerado el caso diagonal para calcular la distancia, pero dicho caso se podría utilizar si se desea. Además, se podrían utilizar otras medidas de distancia en el presente algoritmo.

Las figuras 13A y 13B muestran respectivamente las matrices de distancias calculadas con esta medida para los predictores izquierdo y superior. Además, la figura 13C muestra la matriz de distancias global formada sumando celda a celda estas dos matrices de las figuras 13A y 13B.

Se puede observar que el valor máximo es 10 para la posición de bloque numerada 16 (véase la figura 12), que es la posición de bloque inferior derecha. Esta es la posición de bloque identificada mediante el algoritmo en el caso de que se utilicen los predictores superior e izquierdo.

En la primera realización, se ejecuta el mismo algoritmo también mediante el descodificador, y el descodificador selecciona la misma posición de bloque que el codificador y evita almacenar los vectores de movimiento de las posiciones de bloque no seleccionadas. De este modo, se consigue una reducción similar en las necesidades de la memoria de vectores de movimiento también en el descodificador.

Segunda realización

En la primera realización, no se tiene en cuenta la probabilidad de selección de los predictores espaciales en las diferentes posiciones TP1-TP6 y LP1-LP5. No obstante, en el proceso de AMVP, la búsqueda del predictor superior se inicia a partir de la posición más a la derecha y se desplaza hasta la posición más a la izquierda, tal como ya se ha descrito haciendo referencia a la figura 8. De este modo, se espera que el predictor en la posición TP6 sea seleccionado con mayor frecuencia que el predictor en la posición TP5, y así sucesivamente, de derecha a izquierda

en la figura 12. Ocurre lo mismo para el predictor izquierdo, para el que la búsqueda se inicia en la posición inferior y se desplaza hacia arriba hasta la posición superior. Teniendo en cuenta las frecuencias de selección relativas, puede ser preferible ajustar la posición media para el predictor superior hacia la derecha y la posición media para el predictor inferior hacia abajo. La posición media para cada predictor se podría calcular como una media ponderada de las posiciones potenciales, reflejando los pesos las frecuencias de selección relativas.

Tercera realización

En las realizaciones primera y segunda descritas anteriormente, el codificador utiliza dos predictores espaciales (los predictores superior e izquierdo). En la tercera realización, el codificador utiliza solo un predictor espacial en el conjunto de predictores espaciales. Por ejemplo, si solo se utiliza el predictor izquierdo, se utiliza la matriz de distancias del predictor izquierdo proporcionada en la figura 13A para encontrar la mejor posición. En ese caso, la posición seleccionada es la posición de bloque número 6, el bloque superior derecho. Si solo se utiliza el predictor superior, dos posiciones de bloque tienen al valor máximo (posición de bloque número 11 y posición de bloque número 16). En ese caso, si se utiliza AMVP, tal como se ha descrito haciendo referencia a la figura 8, teóricamente los predictores del lado derecho se seleccionan con mayor frecuencia que los predictores del lado izquierdo (tal como se ha explicado, los predictores se seleccionan de E a A en el proceso de derivación de HEVC). De este modo, el algoritmo puede considerar la utilización del bloque numerado 11, dado que sería el bloque más lejano.

Cuarta realización

En las realizaciones primera, segunda y tercera descritas anteriormente, se supone que todos los bloques a codificar pueden tener dos predictores espaciales (los predictores superior e izquierdo). Según esto, no se realiza ninguna distinción entre bloques para la compactación por bloques.

En la cuarta realización, al menos un bloque se trata de manera diferente de otro bloque para la compactación por bloques.

Por ejemplo, el codificador y el descodificador pueden conocer alguna zona de la imagen en la que existe solamente un predictor espacial, o ninguno. En concreto, el codificador y el descodificador conocen la frontera del fotograma (la frontera del fotograma tiene la misma posición para todos los fotogramas). De este modo, el algoritmo puede tener en cuenta este conocimiento específico. Por ejemplo, para el bloque superior izquierdo del fotograma, no existen predictores disponibles. En ese caso, el algoritmo puede encontrar que la mejor posición de bloque para la compactación por bloques es el centro del bloque. Para bloques a lo largo de la frontera superior, podría estar disponible solamente el predictor izquierdo, y la posición de bloque seleccionada por el algoritmo es la posición de bloque superior derecha numerada 6 (tal como en la tercera realización). Para bloques a lo largo de la frontera izquierda, solo están disponibles algunos predictores superiores. El algoritmo puede encontrar que la posición es el número de bloque 11, es decir, la posición inferior izquierda. De hecho, la posición TP1 en la figura 12 nunca está disponible para la frontera izquierda, de modo que la media de las posiciones es TP4 (no la frontera entre TP3 y TP4) y, en consecuencia, el algoritmo encuentra el número de bloque 11 como la posición más lejana.

Esta realización se puede extender para la frontera del segmento si todos los fotogramas en la secuencia de bits tienen exactamente la misma fragmentación (es decir, si cada segmento tiene exactamente la misma posición para todos los fotogramas).

Quinta realización

En las realizaciones anteriores, el descodificador lleva a cabo el mismo proceso que el codificador para determinar la posición de bloque para la compactación por bloques.

En la quinta realización, solo el codificador utiliza el algoritmo, y la posición del bloque utilizado para la compactación por bloques se transmite al descodificador en la secuencia de bits (Secuencia/imagen/cabecera de segmento).

Sexta realización

En las realizaciones anteriores, el codificador (y, excepto para el quinto ejemplo, el descodificador) ejecuta el algoritmo de la figura 11 para determinar la posición del bloque.

En la sexta realización, el codificador y el descodificador hacen referencia a una tabla que contiene la posición del bloque según la utilización o no de los predictores espaciales. Un ejemplo de tal tabla se proporciona en la Tabla 2 siguiente. Esta tabla se determina por adelantado utilizando el mismo algoritmo que el de la figura 11. Esta solución es menos compleja que aplicar el algoritmo todo el tiempo. La siguiente tabla muestra esta tabla concreta para el ejemplo descrito anteriormente. Las posiciones hacen referencia a la figura 12.

Predictores espaciales disponibles	Posiciones de bloque para la compactación CON N=4
X	4 (centro)
Izquierdo	6 (Superior derecha)
Superior	11 (inferior izquierda)
Izquierdo y superior	16 (inferior derecho)

Séptima realización

Esta realización extiende aún más la idea de la sexta realización. La posición de bloque inferior derecha se utiliza siempre para la compactación por bloques cualquiera que sea el bloque espacial utilizado para AMVP o modo de fusión. Esto ahorra carga de proceso y no hay necesidad de señalizar al descodificador qué posición de predictor ha sido determinada.

Octava realización

Las realizaciones anteriores suponen que la compactación por bloques se aplica de manera uniforme a todos los bloques. No obstante, en HEVC, se permiten bloques de diferentes tamaños. Las posiciones de bloque para un tamaño de bloque pueden no ser óptimas para otro tamaño de bloque. Por ejemplo, se ha encontrado que las posiciones de bloque mostradas en la Tabla 2 funcionan bien para un bloque pequeño, por ejemplo, un bloque de 4x4, pero en el caso de un bloque grande puede ser mejor utilizar otras posiciones de bloque.

En la octava realización, es posible tener en cuenta el tamaño del bloque cuando se determina la posición del bloque. Esto se puede conseguir teniendo una matriz de distancias similar a la de la figura 13C para cada tamaño de bloque diferente.

Novena realización

Es posible simplificar la octava realización simplemente cambiando de una posición de bloque predeterminada para un primer tamaño de bloque a otra posición de bloque predeterminada para un segundo tamaño de bloque. Por ejemplo, cuando el tamaño del bloque es pequeño, por ejemplo, menor o igual que 16x16 (píxeles), se podría seleccionar siempre la posición de bloque inferior derecha, mientras que cuando el tamaño del bloque es grande, por ejemplo, un 32x32 (píxeles), o mayor, se podría seleccionar siempre la posición de bloque central, véase la posición de bloque sombreada en la figura 2.

Los resultados de la simulación para esta realización son prometedores. Se consiguió un ahorro en tasa de bits (mejora en la eficiencia de la codificación), en comparación con utilizar la posición de bloque central en todas las ocasiones, tal como en la propuesta anterior JCT-VC D164 mencionada en la introducción.

Si se desea, se puede disponer una tabla del tipo mostrado en la Tabla 2 para cada tamaño de bloque, que contenga las posiciones de bloque adecuadas para cada permutación de los predictores disponibles.

Se considera que para bloques mayores la posición de bloque central funciona mejor que la posición inferior derecha, porque la posición inferior derecha en el caso de un bloque grande no puede ser tan representativa del bloque como conjunto como el bloque central. Este no es el caso con bloques pequeños, para los que la posición inferior derecha sigue siendo representativa del bloque como conjunto debido a las menores distancias implicadas, pero sigue introduciendo una diversidad controlada con respecto a los predictores espaciales.

Décima realización

En la décima realización, la posición o posiciones de los bloques utilizadas en JCT-VC D164 se utilizan para bloques grandes, y la posición de bloque inferior derecha se utiliza para bloques pequeños.

Undécima realización

En las realizaciones anteriores, la posición de bloque seleccionada para la compactación por bloques ha sido una de las posiciones de bloque en el bloque contiguo. No obstante, no es necesario limitar la posición de bloque seleccionada a una en el bloque contiguo. Por ejemplo, en lugar de la posición inferior derecha del bloque contiguo sería posible utilizar una posición de bloque en la proximidad de la posición inferior derecha de otro bloque en el mismo fotograma de referencia vecino del bloque contiguo. Como ejemplos, la posición del bloque podría ser la posición inferior izquierda en el bloque inmediatamente a la derecha del bloque contiguo, o la posición superior derecha en el bloque inmediatamente por debajo del bloque contiguo, o la posición superior izquierda en el bloque diagonalmente por debajo y a la derecha del bloque contiguo.

Duodécima realización

En el documento JCT-VC D125, descrito en la introducción, se utilizan 4 predictores temporales, uno (T) del bloque contiguo y otros tres (E, H e I) exteriores al bloque contiguo en las esquinas superior izquierda, inferior izquierda e inferior derecha del bloque contiguo. En lugar de utilizar T como uno de los predictores temporales sería posible utilizar la posición de bloque inferior derecha en el bloque contiguo. Esto probablemente no proporciona un buen compromiso en términos de diversidad, pero en esto el centro sería lo mejor. Pero en términos de codificación del índice del predictor esto tiene una ventaja, porque la inferior derecha puede con más frecuencia ser igual al predictor H (a menudo se elimina un predictor). De hecho, en la contribución del documento JCT-VC D125 se utilizan demasiados predictores, de modo que D125 aumenta el coeficiente dedicado al índice del predictor.

Decimotercera realización

En las realizaciones anteriores la posición del bloque se ha seleccionado con el propósito de reducir las necesidades de la memoria de vectores de movimiento.

En la decimotercera realización, la posición del bloque se selecciona con el propósito de mejorar la eficiencia en la codificación. Esto es posible dado que, si se elige el predictor temporal del fotograma de referencia, por ejemplo, de una posición de bloque adecuada en el bloque contiguo para conseguir una diversidad controlada con los predictores espaciales, entonces el predictor temporal tiende a ser un buen competidor para el predictor espacial, cuando se realiza la selección de distorsión de coeficiente del predictor. Si el predictor temporal es un predictor mejor que los predictores espaciales, como es a menudo el caso en la práctica, entonces, mejorar este predictor (en el sentido de hacerlo más próximo al vector de movimiento requerido para el bloque a codificar, y por ello minimizar el residuo) es bueno para la eficiencia de la codificación.

En principio, se pueden utilizar todas las anteriores primera a duodécima realizaciones, para seleccionar la posición del bloque con el propósito de mejorar la eficiencia en la codificación. En la decimotercera realización, aún pueden ser almacenados todos los vectores de movimiento, es decir, no es necesario emplear reducción de la memoria de vectores de movimiento. No obstante, tal reducción es deseable en muchos casos.

Decimocuarta realización

Si se almacenan los vectores de movimiento en todas las posiciones de bloque, es posible tener en cuenta el resultado real de la selección de los predictores espaciales cuando se selecciona la mejor posición de bloque. Por ejemplo, en la figura 12, si solo está disponible el predictor superior y se selecciona la posición TP6, se podría seleccionar la posición de bloque numerada 11. Si, por otro lado, solo está disponible el predictor superior y se selecciona la posición TP1, entonces se podría seleccionar la posición de bloque numerada 16. Otras posiciones de bloque serían mejores si estuviese disponible una combinación de los predictores superior e izquierdo, de nuevo en función de las posiciones de los predictores superior e izquierdo seleccionados. En general para cada posible resultado de la selección de predictor espacial se puede determinar una mejor posición de bloque. Las mejores posiciones de bloque se podrían almacenar en una tabla de consulta que tiene una entrada para cada resultado diferente.

Las realizaciones descritas anteriormente se basan en particiones de bloques de imágenes de entrada, pero, de manera más general, se puede considerar cualquier tipo de porciones de imagen para codificar o descodificar, en concreto porciones rectangulares, o de manera más general porciones geométricas.

REIVINDICACIONES

1. Procedimiento de generación de predictores de vectores de movimiento para un bloque a descodificar que forma parte de un fotograma de video actual que debe ser descodificado, en el que el número de predictores temporales de vectores de movimiento generados no es mayor de uno, y el número de predictores espaciales de vectores de movimiento generados es al menos uno, comprendiendo el procedimiento generar dicho un predictor temporal del vector de movimiento seleccionando, si está disponible, como el predictor temporal del vector de movimiento el vector de movimiento asociado con una posición superior izquierda, en un bloque del fotograma de referencia vecino y diagonalmente por debajo y a la derecha de un bloque del fotograma de referencia contiguo a dicho bloque a descodificar, teniendo dicho bloque del fotograma de referencia la misma posición y el mismo tamaño que el bloque a descodificar.

2. Procedimiento, según la reivindicación 1, en el que el por lo menos uno de dichos predictores espaciales se selecciona antes de la generación del predictor temporal de vector de movimiento, y la generación del predictor temporal se lleva a cabo dependiendo del predictor o predictores espaciales seleccionados.

3. Procedimiento de descodificación de una secuencia de bits que comprende una secuencia codificada de imágenes digitales, siendo codificada al menos una porción de una imagen mediante compensación del movimiento con respecto a una imagen de referencia, que comprende el procedimiento de generación de predictores de vectores de movimiento según una de las reivindicaciones 1 o 2.

4. Procedimiento según la reivindicación 3, que comprende descodificar la secuencia de bits utilizando los predictores de vectores de movimiento.

5. Dispositivo de generación de predictores de vectores de movimiento para un bloque a descodificar que forma parte de un fotograma de video actual que debe ser descodificado, en el que el número de predictores temporales de vectores de movimiento generados no es mayor de uno, y el número de predictores espaciales de vectores de movimiento generados es al menos uno, comprendiendo dicho dispositivo medios dispuestos para generar el predictor temporal del vector de movimiento seleccionando, si está disponible, como el predictor temporal del vector de movimiento, el vector de movimiento en una posición superior izquierda, en un bloque del fotograma de referencia vecino y diagonalmente por debajo y a la derecha de un bloque del fotograma de referencia contiguo a dicho bloque a descodificar, teniendo dicho bloque del fotograma de referencia la misma posición y el mismo tamaño que el bloque a descodificar.

6. Dispositivo, según la reivindicación 5, que comprende medios para seleccionar al menos dicho predictor espacial antes de que sea generado el predictor temporal, pudiendo funcionar los medios de generación del predictor temporal de vector de movimiento para generar el predictor temporal de vector de movimiento en función del predictor o predictores espaciales seleccionados.

7. Dispositivo para la descodificación de una secuencia de bits que comprende una secuencia codificada de imágenes digitales, siendo codificada por lo menos una porción de una imagen mediante compensación del movimiento con respecto a una imagen de referencia, comprendiendo el dispositivo de generación un conjunto de predictores de vectores de movimiento según cualquiera de las reivindicaciones 5 o 6.

8. Dispositivo según la reivindicación 7, que comprende medios para la descodificación de la secuencia de bits utilizando predictores de vectores de movimiento.

9. Procedimiento de generación de predictores de vectores de movimiento para un bloque a codificar que forma parte de un fotograma de video actual que debe ser codificado, en el que el número de predictores temporales de vectores de movimiento generados no es mayor de uno, y el número de predictores espaciales de vectores de movimiento generados es al menos uno, comprendiendo el procedimiento la generación de dicho un predictor temporal del vector de movimiento, si está disponible, seleccionando como el predictor temporal del vector de movimiento el vector de movimiento asociado con una posición superior izquierda, en un bloque del fotograma de referencia vecino y diagonalmente por debajo y a la derecha de un bloque del fotograma de referencia contiguo a dicho bloque a codificar, teniendo dicho bloque del fotograma de referencia la misma posición y el mismo tamaño que el bloque a codificar.

10. Dispositivo de generación de predictores de vectores de movimiento para un bloque a codificar que forma parte de un fotograma de video actual que debe ser codificado, en el que el número de predictores temporales de vectores de movimiento generados no es mayor de uno, y el número de predictores espaciales de vectores de movimiento generados es al menos uno, comprendiendo dicho dispositivo medios dispuestos para generar el predictor temporal del vector de movimiento seleccionando, si está disponible, como el predictor temporal del vector de movimiento, el vector de movimiento en una posición superior izquierda, en un bloque del fotograma de referencia vecino y diagonalmente por abajo y a la derecha de un bloque del fotograma de referencia contiguo a dicho bloque a codificar, teniendo dicho bloque del fotograma de referencia la misma posición y el mismo tamaño que el bloque a codificar.

11. Dispositivo para codificación de una secuencia de imágenes digitales en una secuencia de bits, estando codificada al menos una porción de una imagen por compensación del movimiento con respecto a una imagen de referencia, comprendiendo el dispositivo de generación predictores de vectores de movimiento según la reivindicación 10.

12. Programa que, cuando es ejecutado en un ordenador o un procesador, hace que el ordenador o el procesador lleven a cabo un procedimiento según una cualquiera de las reivindicaciones 1 a 4 y la reivindicación 9.

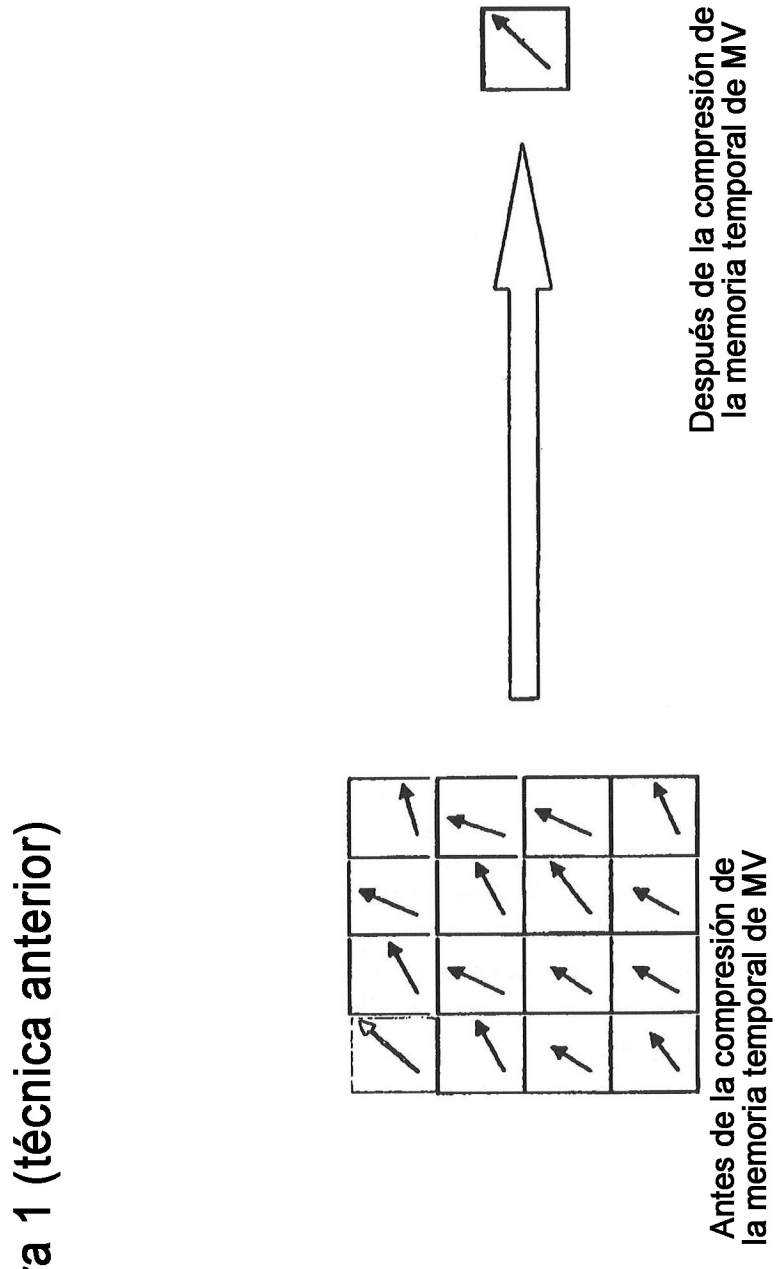


Figura 1 (técnica anterior)

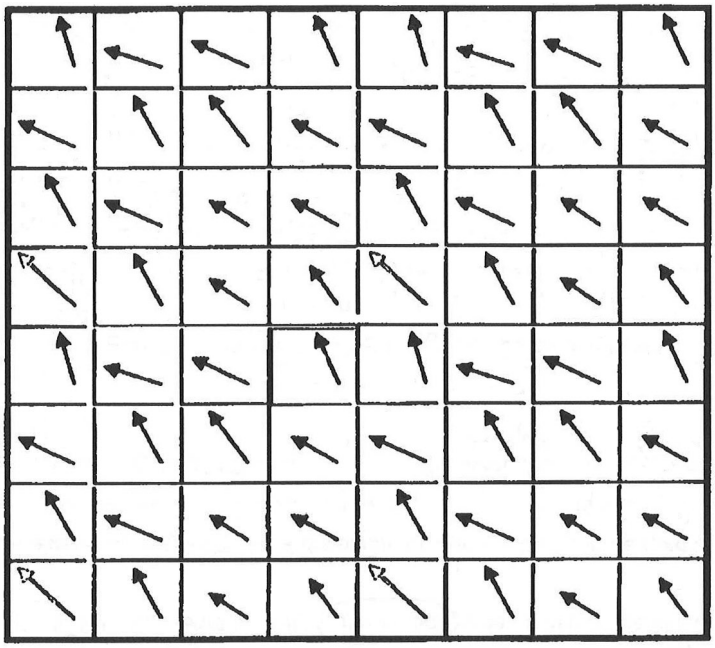


Figura 2 (técnica anterior)

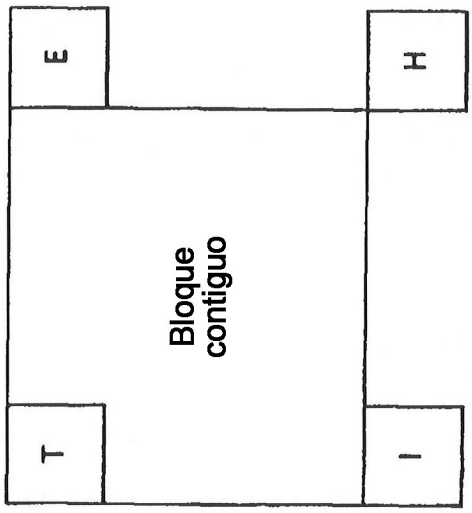


Figura 3 (técnica anterior)

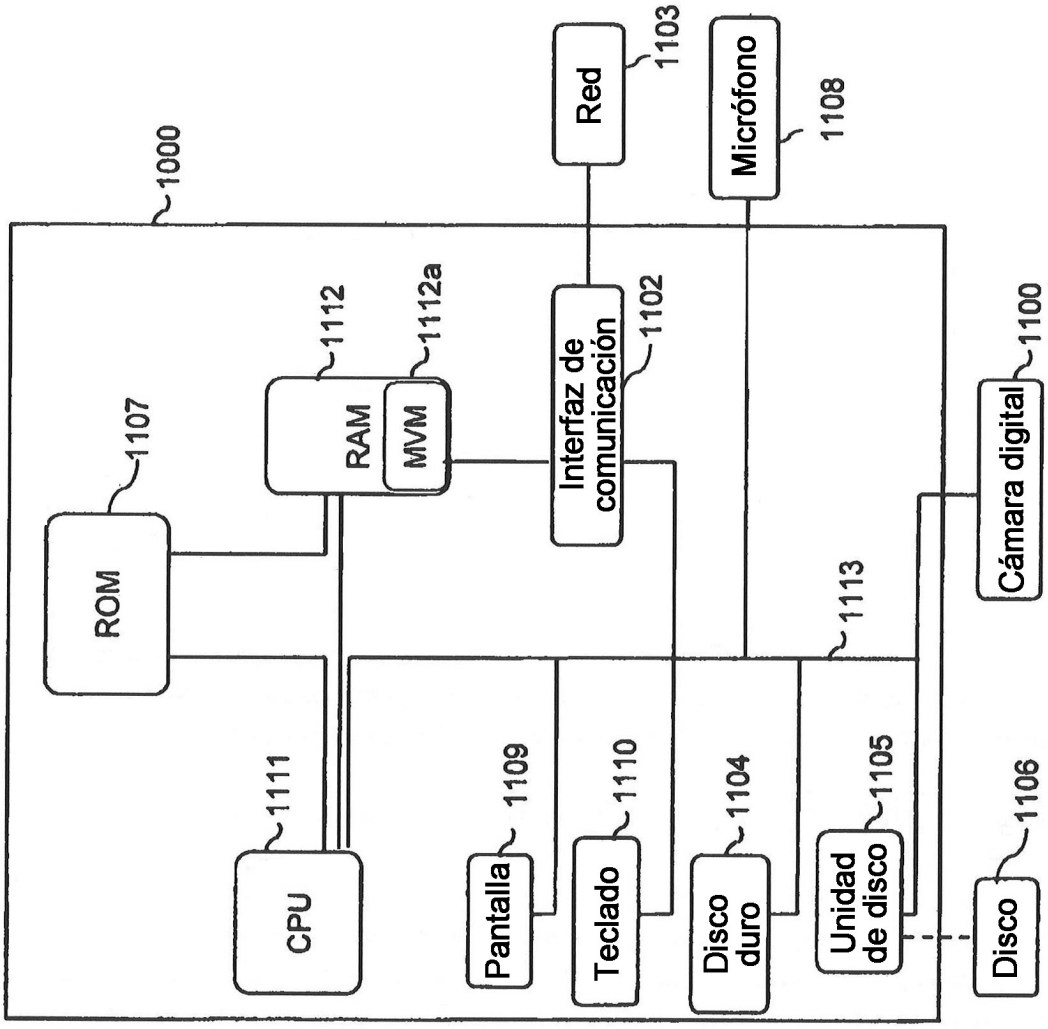

1000
Red
1103
Micrófono
1108
1112
1112a
RAM
MVM
Interfaz de comunicación
1102
1107
ROM
Cámara digital
1100
1113
1111
CPU
1109
Pantalla
1110
Teclado
1104
Disco duro
1105
Unidad de disco
1106
Disco

Figura 4

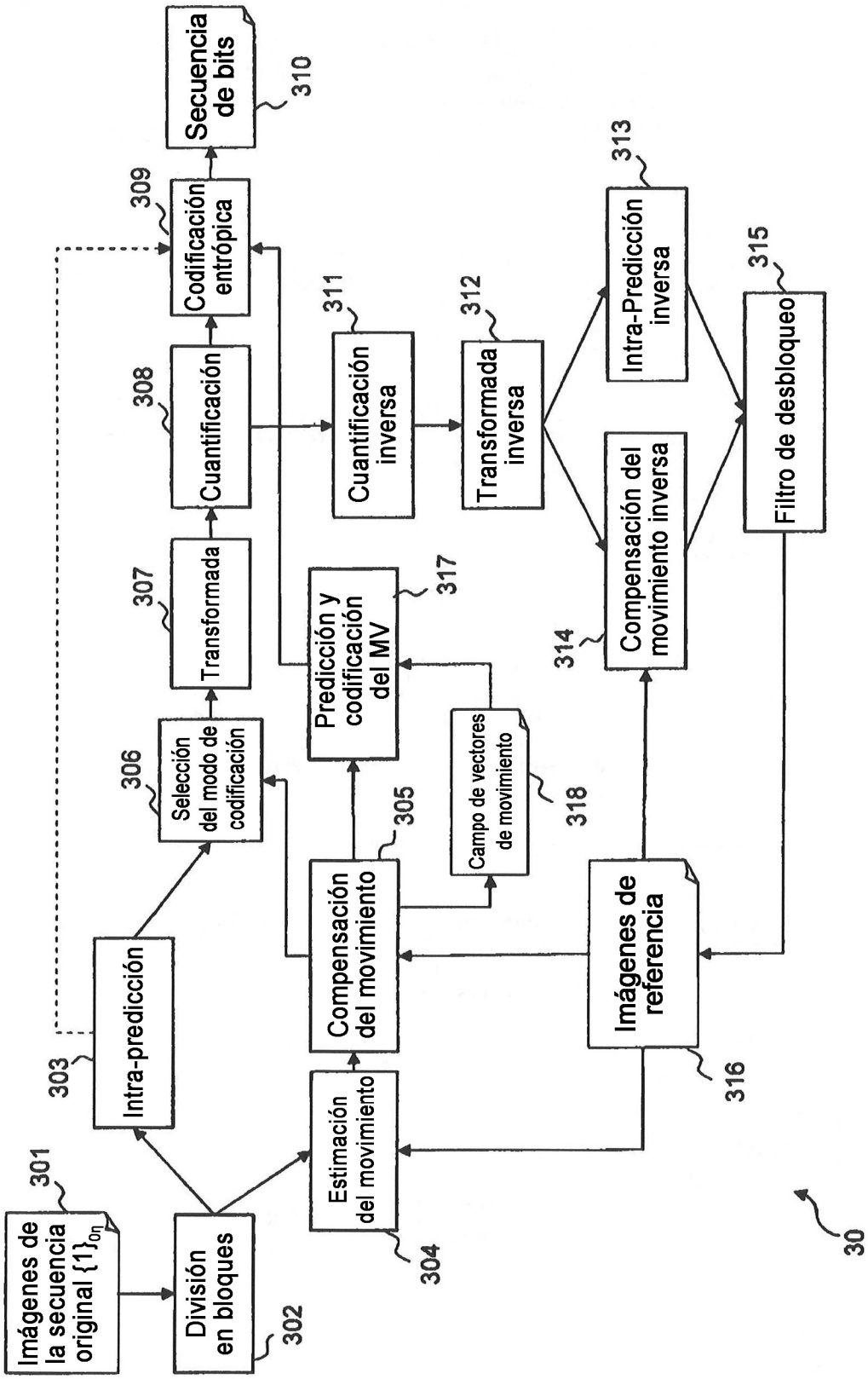

Imágenes de la secuencia original $\{I\}_{0n}$
301
División en bloques
302
Intra-predicción
303
Estimación del movimiento
304
Compensación del movimiento
305
Selección del modo de codificación
306
Transformada
307
Cuantificación
308
Codificación entrópica
309
Secuencia de bits
310
Predicción y codificación del MV
317
Campo de vectores de movimiento
318
Cuantificación inversa
311
Transformada inversa
312
Intra-Predicción inversa
313
Compensación del movimiento inversa
314
Filtro de desbloqueo
315
Imágenes de referencia
316
30

Figura 5

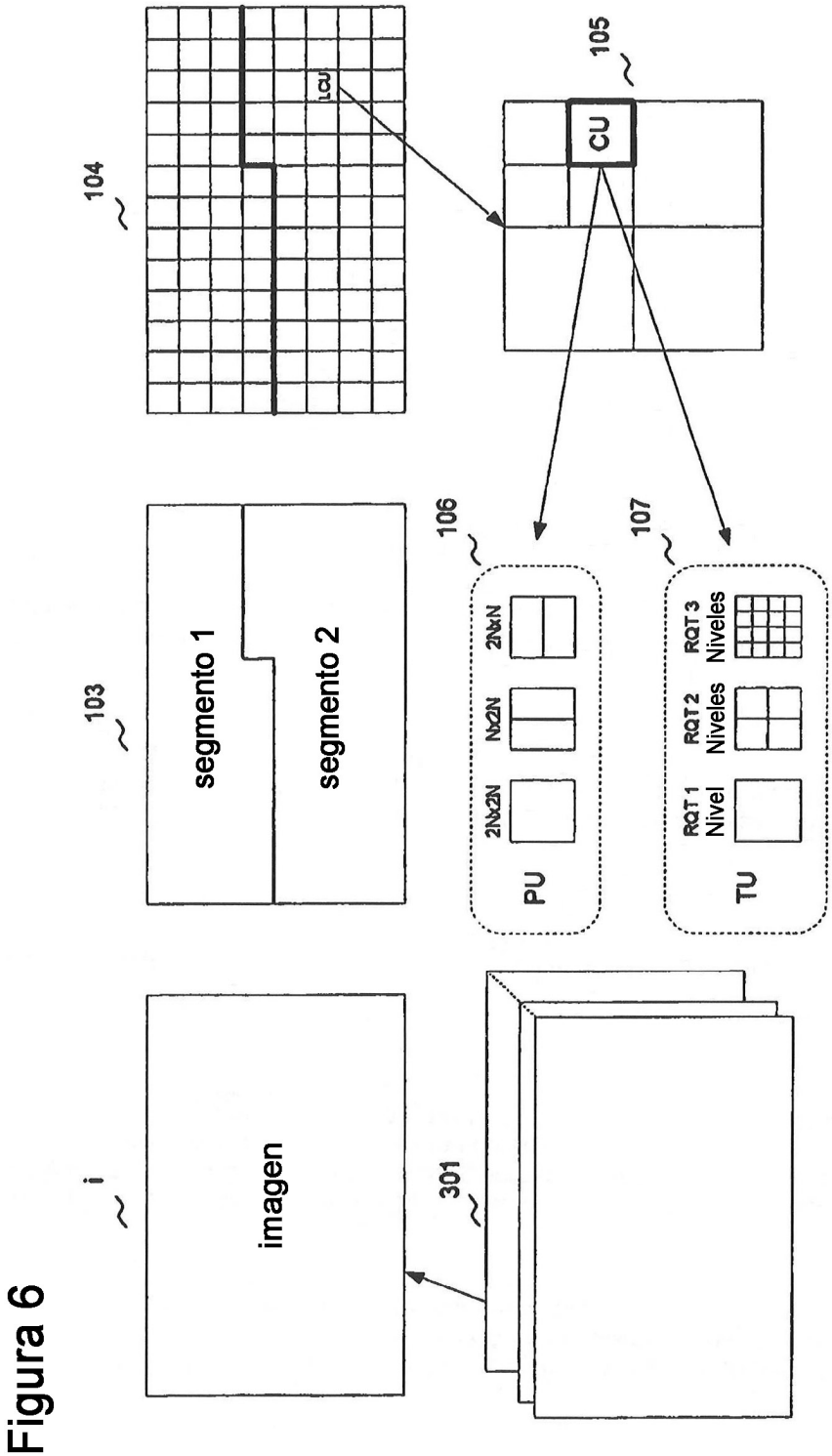

i
imagen
301
segmento 1
segmento 2
103
104
LCU
105
CU
106
PU
2Nx2N
NxN
2Nx2N
107
TU
RQT 1
Nivel
RQT 2
Niveles
RQT 3
Niveles

Figura 6

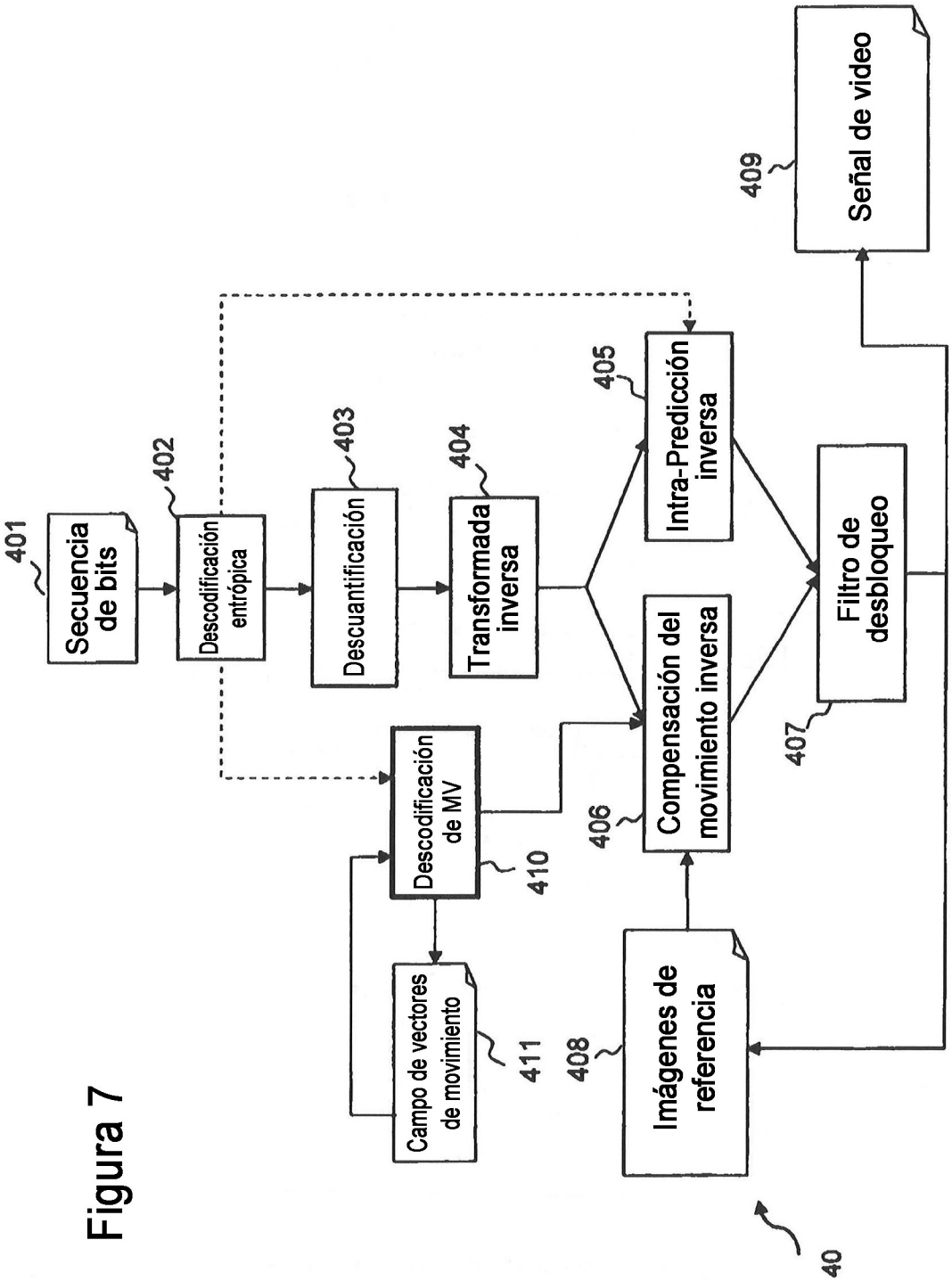

401
Secuencia de bits
402
Descodificación entrópica
403
Descuantificación
404
Transformada inversa
405
Intra-Predicción inversa
409
Señal de video
Filtro de desbloqueo
407
Compensación del movimiento inversa
406
Descodificación de MV
410
Campo de vectores de movimiento
411
Imágenes de referencia
408
40

Figura 7

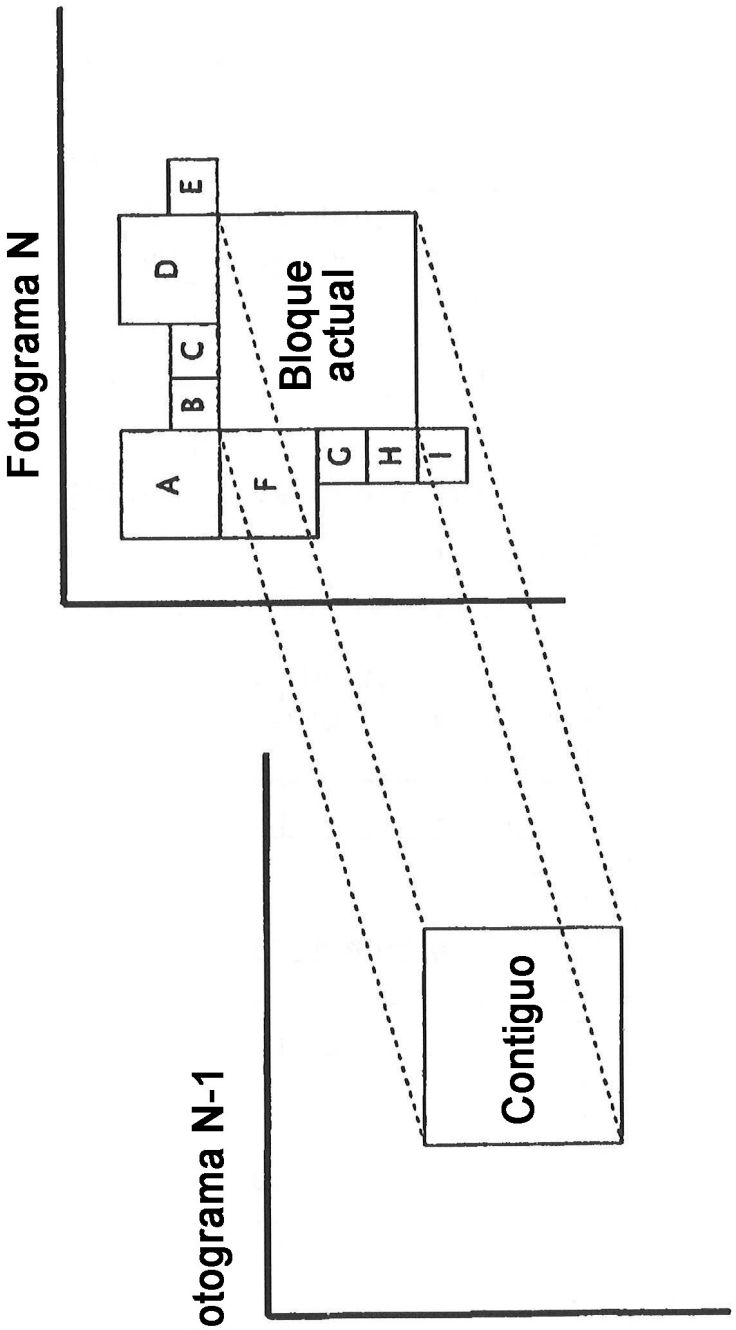

Fotograma N
A
B
C
D
E
F
G
H
I
Bloque actual
Fotograma N-1
Contiguo

Figura 8

Figura 9

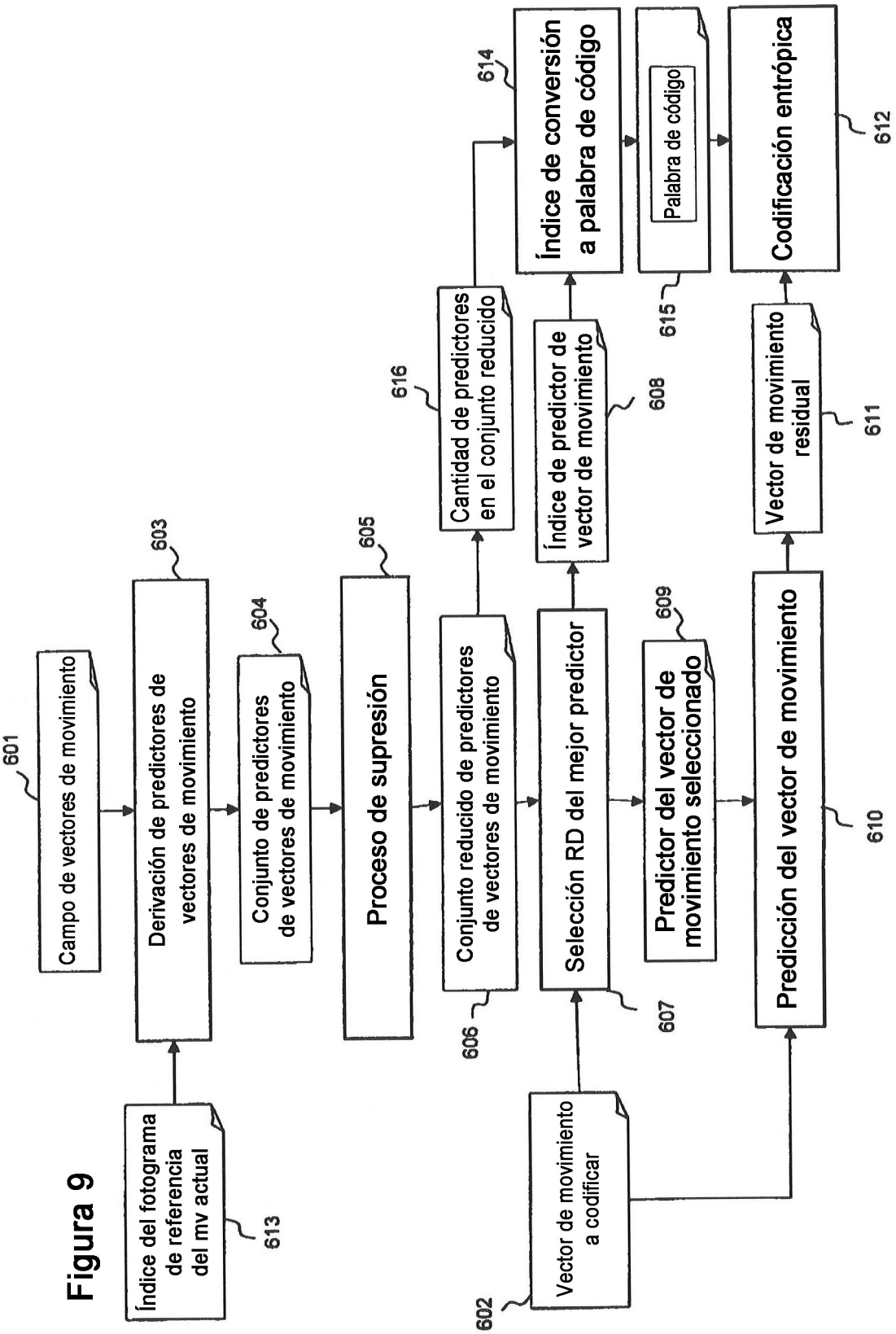
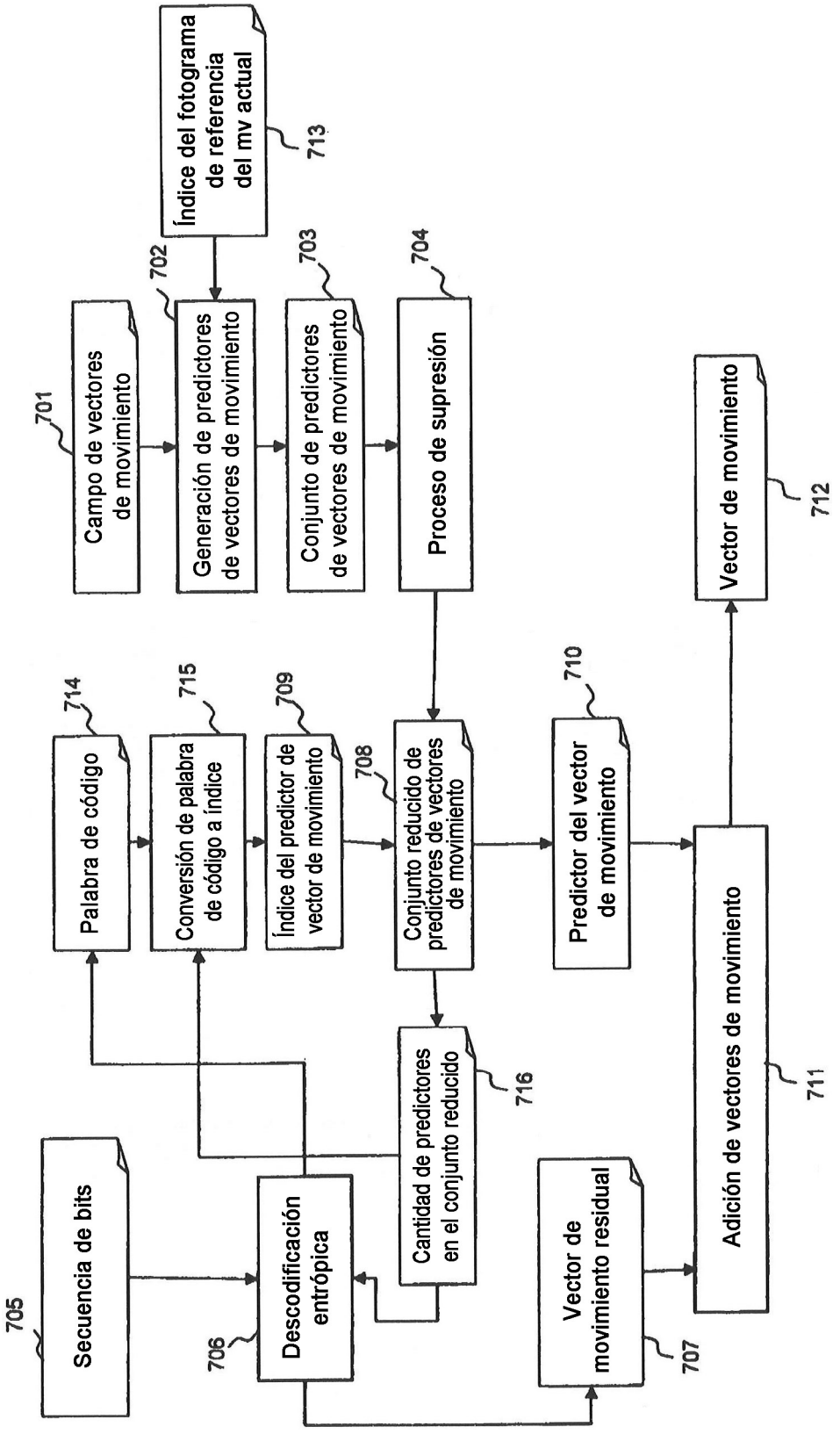

Índice del fotograma de referencia del mv actual
613
Campo de vectores de movimiento
601
Derivación de predictores de vectores de movimiento
603
Conjunto de predictores de vectores de movimiento
604
Proceso de supresión
605
Conjunto reducido de predictores de vectores de movimiento
606
Cantidad de predictores en el conjunto reducido
616
Índice de conversión a palabra de código
614
Palabra de código
615
Codificación entrópica
612
Índice de predictor de vector de movimiento
608
Selección RD del mejor predictor
607
Predictor del vector de movimiento seleccionado
609
Vector de movimiento residual
611
Predicción del vector de movimiento
610
Vector de movimiento a codificar
602

Figura 10


Índice del fotograma de referencia del mv actual
713
702
703
704
Campo de vectores de movimiento
Generación de predictores de vectores de movimiento
Conjunto de predictores de vectores de movimiento
Proceso de supresión
701
Vector de movimiento
712
714
715
709
710
Palabra de código
Conversión de palabra de código a índice
Índice del predictor de vector de movimiento
708
Conjunto reducido de predictores de vectores de movimiento
Predictor del vector de movimiento
Cantidad de predictores en el conjunto reducido
716
Secuencia de bits
705
706
Descodificación entrópica
Vector de movimiento residual
707
Adición de vectores de movimiento
711

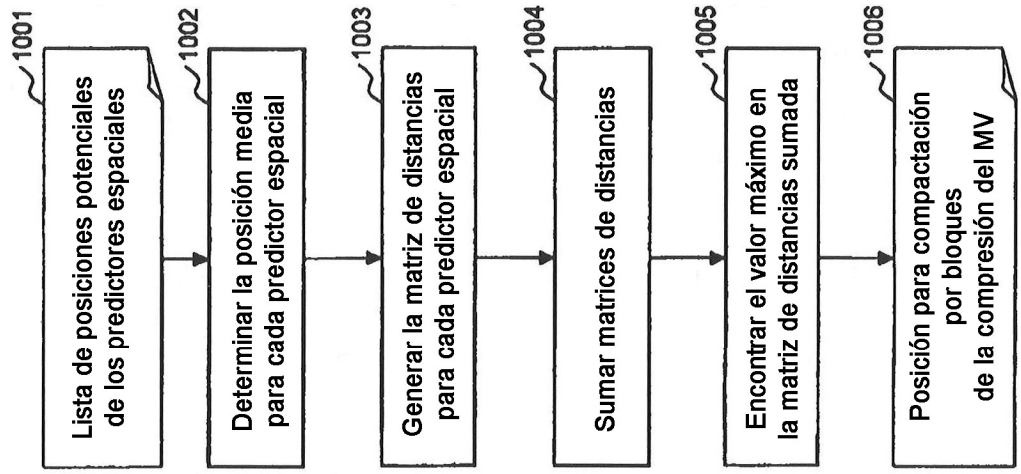

1001
Lista de posiciones potenciales de los predictores espaciales
1002
Determinar la posición media para cada predictor espacial
1003
Generar la matriz de distancias para cada predictor espacial
1004
Sumar matrices de distancias
1005
Encontrar el valor máximo en la matriz de distancias sumada
1006
Posición para compactación por bloques de la compresión del MV

Figura 11

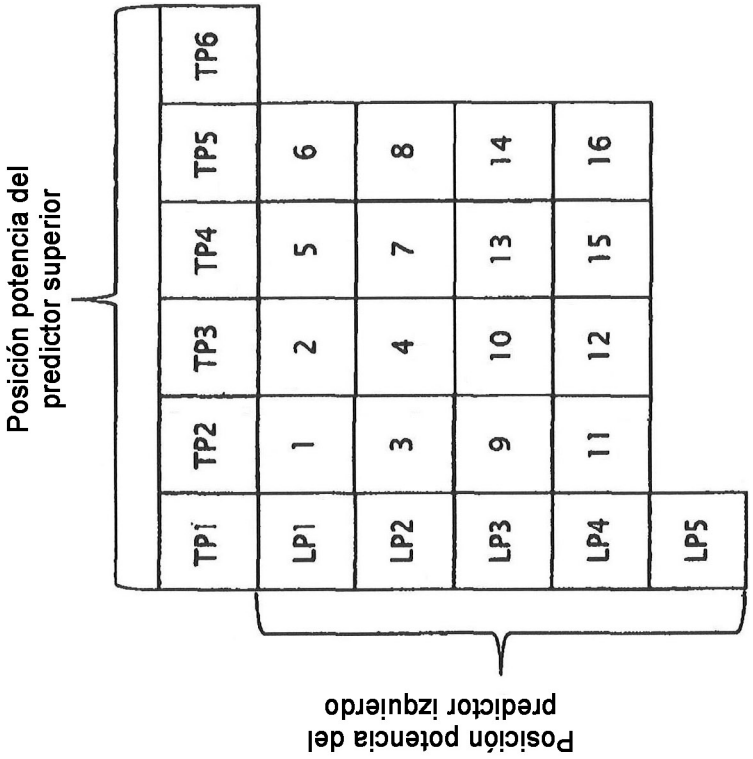
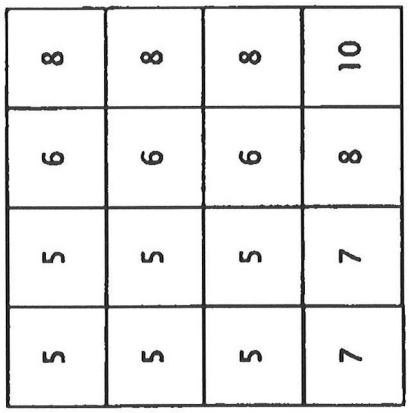
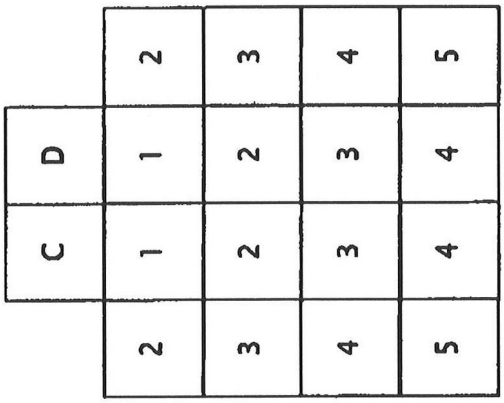

Posición potencia del predictor superior
TP1
TP2
TP3
TP4
TP5
TP6
LP1
LP2
LP3
LP4
LP5
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
Posición potencia del predictor izquierdo

Figura 12



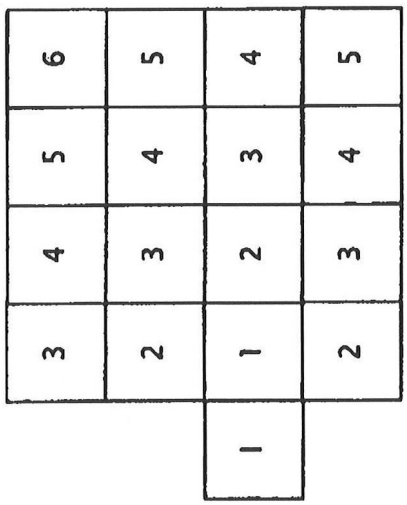
5	5	5	7
5	5	5	7
6	6	6	8
8	8	8	10

Figura 13C



	C	D		
2	1	1	2	
3	2	2	3	
4	3	3	4	
5	4	4	5	

Figura 13B



	1			
3	2	2	3	4
4	3	3	4	5
5	4	4	5	6

Figura 13A