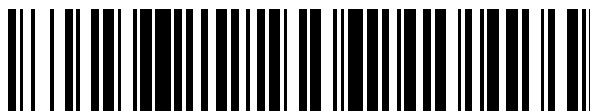


19



OFICINA ESPAÑOLA DE  
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 751 523**

51 Int. Cl.:

**H04N 19/105** (2014.01)  
**H04N 19/52** (2014.01)  
**H04N 19/139** (2014.01)  
**H04N 19/176** (2014.01)  
**H04N 19/70** (2014.01)  
**H04N 19/147** (2014.01)  
**H04N 19/513** (2014.01)  
**H04N 19/61** (2014.01)  
**H04N 19/96** (2014.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

96 Fecha de presentación y número de la solicitud europea: **13.08.2010** **E 15168128 (5)**

97 Fecha y número de publicación de la concesión europea: **02.10.2019** **EP 2928186**

54 Título: **Aparato y procedimiento de decodificación de una imagen**

30 Prioridad:

**13.08.2009 KR 20090074896**

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

**31.03.2020**

73 Titular/es:

**SAMSUNG ELECTRONICS CO., LTD. (100.0%)**  
**129, Samsung-ro, Yeongtong-gu**  
**Suwon-si, Gyeonggi-do 443-742, KR**

72 Inventor/es:

**LEE, TAMMY;**  
**HAN, WOO-JIN y**  
**LEE, KYO-HYUK**

74 Agente/Representante:

**CARPINTERO LÓPEZ, Mario**

**ES 2 751 523 T3**

Aviso: En el plazo de nueve meses a contar desde la fecha de publicación en el Boletín Europeo de Patentes, de la mención de concesión de la patente europea, cualquier persona podrá oponerse ante la Oficina Europea de Patentes a la patente concedida. La oposición deberá formularse por escrito y estar motivada; sólo se considerará como formulada una vez que se haya realizado el pago de la tasa de oposición (art. 99.1 del Convenio sobre Concesión de Patentes Europeas).

## DESCRIPCIÓN

Aparato y procedimiento de decodificación de una imagen

### **Campo técnico**

5 Los procedimientos consistentes con realizaciones ejemplares se refieren a un procedimiento para decodificar una imagen.

### **Técnica anterior**

10 Un codificador-decodificador, tal como Grupo de expertos de Imagen en Movimiento (MPEG) -4 H.264/MPEG-4 Codificación de Vídeo Avanzada (AVC), utiliza vectores de movimiento de bloques previamente codificados adyacentes a un bloque actual para predecir un vector de movimiento del bloque actual. Es decir, se usa una mediana de vectores de movimiento de bloques codificados previamente adyacentes a los lados izquierdo, superior y superior derecho de un bloque actual como un predictor del vector de movimiento del bloque actual.

15 La publicación de revista LAROCHE G ET AL: "RD Optimized Coding for Motion Vector Predictor Selection", IEEE TRANSACTIONS ON CIRCUITS AND SYSTEMS FOR VIDEO TECHNOLOGY, IEEE SERVICE CENTER, PISCATAWAY, NJ, US, vol. 18, n.º 9, 1 de Septiembre de 2008 (2008-09-01), páginas 1247-1257, ISSN: 1051-8215, DOI: 10.1109/TCSVT.2008.928882 describe un procedimiento de codificación de vectores de movimiento en el contexto del estándar H.264/AVC que se basa en un esquema de competencia de vectores de movimiento que involucra un conjunto de predictores candidatos de vectores de movimiento derivados de bloques adyacentes.

### **Divulgación**

#### **Solución técnica**

20 Las realizaciones ejemplares proporcionan un procedimiento para decodificar una imagen.

#### **Efectos ventajosos**

De acuerdo con la presente solicitud, un vector de movimiento se descodifica de manera eficiente en base a un predictor del vector de movimiento.

#### **Descripción de los dibujos**

25 Los aspectos anteriores y otros se harán más evidentes describiendo en detalle ejemplos de realización con referencia a los dibujos adjuntos en los que:

La figura 1 es un diagrama de bloques de un aparato para codificar una imagen;

La figura 2 es un diagrama de bloques de un aparato para decodificar una imagen;

La figura 3 ilustra unidades de codificación jerárquica;

30 La figura 4 es un diagrama de bloques de un codificador de imagen basado en una unidad de codificación;

La figura 5 es un diagrama de bloques de un decodificador de imagen basado en una unidad de codificación;

La figura 6 ilustra una unidad de codificación máxima, una unidad de subcodificación y una unidad de predicción;

La figura 7 ilustra una unidad de codificación y una unidad de transformación;

35 Las figuras 8A y 8B ilustran formas de división de una unidad de codificación, una unidad de predicción y una unidad de transformación;

La figura 9 es un diagrama de bloques de un aparato para codificar un vector de movimiento;

Las figuras 10A y 10B ilustran candidatos predictores del vector de movimiento de un modo explícito;

Las figuras 11A a 11C ilustran candidatos predictores del vector de movimiento de un modo implícito;

La figura 12 ilustra un procedimiento para generar un predictor del vector de movimiento en un modo implícito;

40 La figura 13 es un diagrama de bloques de un aparato para decodificar un vector de movimiento;

La figura 14 es un diagrama de flujo de un procedimiento para codificar un vector de movimiento; y

La figura 15 es un diagrama de flujo que ilustra un procedimiento para decodificar un vector de movimiento, de acuerdo con una realización ejemplar.

**Mejor modo**

Según un aspecto de una realización ejemplar, se proporciona un procedimiento de decodificación de acuerdo con las reivindicaciones 1-3 adjuntas.

La divulgación que permite la invención se encuentra en la realización de la figura 15.

- 5 Las restantes "realizaciones" o "realizaciones ejemplares" han de entenderse como ejemplos que no describen partes de la presente invención.

**Modo de la invención**

- 10 La siguiente descripción se hace con referencia a los dibujos adjuntos, en los que números de referencia iguales se refieren a elementos similares de principio a fin. Las expresiones como "al menos uno de", cuando preceden a una lista de elementos, modifican toda la lista de elementos y no modifican los elementos individuales de la lista. En la presente especificación, una "imagen" puede denotar una imagen fija para un vídeo o una imagen en movimiento, es decir, el vídeo en sí.

- 15 La figura 1 es un diagrama de bloques de un aparato 100 para codificar una imagen, de acuerdo con una realización ejemplar. Con referencia a la figura 1, el aparato 100 incluye un divisor 110 de unidad de codificación máxima, un determinante 120 de profundidad de codificación, un codificador 130 de datos de imagen y un codificador 140 de información de codificación.

- 20 El divisor 110 de unidad de codificación máxima puede dividir una imagen actual o porción basado en una unidad de codificación máxima que es una unidad de codificación de un tamaño más grande. Es decir, el divisor 110 de unidad de codificación máxima puede dividir la imagen o segmento actual para obtener al menos una unidad de codificación máxima.

- 25 De acuerdo con una realización ejemplar, una unidad de codificación se puede representar utilizando una unidad de codificación máxima y una profundidad. Como se describió anteriormente, la unidad de codificación máxima indica una unidad de codificación que tiene el tamaño más grande entre las unidades de codificación de la imagen actual, y la profundidad indica el tamaño de una unidad de subcodificación obtenida disminuyendo jerárquicamente la unidad de codificación. A medida que aumenta la profundidad, la unidad de codificación puede disminuir de una unidad de codificación máxima a una unidad de codificación mínima, en el que una profundidad de la unidad de codificación máxima se define como una profundidad mínima y una profundidad de la unidad de codificación mínima se define como una profundidad máxima. Dado que el tamaño de la unidad de codificación disminuye desde la unidad de codificación máxima a medida que aumenta la profundidad, una unidad de subcodificación de una profundidad k puede incluir una pluralidad de unidades de subcodificación de una profundidad (k + n)ésima (donde k y n son enteros igual o mayor que 1).

- 35 De acuerdo con un aumento del tamaño de una imagen a codificar, la codificación de una imagen en una unidad de codificación mayor puede causar una relación de compresión de imagen superior. Sin embargo, si se fija una unidad de codificación mayor, una imagen puede no codificarse de manera eficiente al reflejar las características de la imagen que cambian continuamente.

Por ejemplo, cuando se codifica un área lisa como el mar o en el cielo, a mayor sea una unidad de codificación, más puede aumentar una relación de compresión. Sin embargo, cuando se codifica un área compleja como personas o edificios, cuanto más pequeña es una unidad de codificación, más puede aumentar la relación de compresión.

- 40 Por consiguiente, según una realización ejemplar, se establece una unidad de codificación de imagen máxima diferente y una profundidad máxima diferente para cada imagen o porción. Dado que una profundidad máxima indica el número máximo de veces por el cual una unidad de codificación puede disminuir, el tamaño de cada unidad de codificación mínima incluida en una unidad de codificación de imagen máxima puede establecerse de manera variable de acuerdo con una profundidad máxima.

- 45 El determinante 120 de profundidad de codificación determina una profundidad máxima. Por ejemplo, la profundidad máxima se puede determinar en función del cálculo del coste de la tasa de distorsión (R-D). Además, la profundidad máxima se puede determinar de manera diferente para cada imagen o porción o para cada unidad de codificación máxima. La profundidad máxima determinada se proporciona al codificador 140 de información de codificación, y los datos de imagen según las unidades de codificación máximas se proporcionan al codificador 130 de datos de imagen.

- 50 La profundidad máxima indica una unidad de codificación que tiene el tamaño más pequeño que puede ser incluido en una unidad de codificación máximo, es decir, una unidad mínima de codificación. En otras palabras, una unidad de codificación máxima se puede dividir en unidades de subcodificación que tienen diferentes tamaños según diferentes profundidades. Esto se describirá en detalle con referencia a las figuras 8A y 8B. Además, las unidades de subcodificación que tienen diferentes tamaños, que se incluyen en la unidad de codificación máxima, pueden predecirse o transformarse basándose en unidades de procesamiento que tienen diferentes tamaños. En otras
- 55

5 palabras, el aparato 100 puede realizar una pluralidad de operaciones de procesamiento para la codificación de imágenes basadas en unidades de procesamiento que tienen varios tamaños y diversas formas. Para codificar datos de imagen, se realizan operaciones de procesamiento como predicción, transformación y codificación de entropía, en las que se pueden usar unidades de procesamiento que tengan el mismo tamaño para cada operación o unidades de procesamiento que tengan diferentes tamaños para cada operación.

10 Por ejemplo, el aparato 100 puede seleccionar una unidad de procesamiento que es diferente de una unidad de codificación para predecir la unidad de codificación. Cuando el tamaño de una unidad de codificación es  $2N \times 2N$  (donde  $N$  es un número entero positivo), las unidades de procesamiento para la predicción pueden ser  $2N \times 2N$ ,  $2N \times N$ ,  $N \times 2N$  y  $N \times N$ . En otras palabras, la predicción de movimiento se puede realizar basándose en una unidad de procesamiento que tiene una forma en la que al menos uno de la altura y el ancho de una unidad de codificación se divide por igual por dos. En lo sucesivo, una unidad de procesamiento, que es la base de la predicción, se denomina unidad de predicción.

15 A modo de predicción puede ser al menos uno de un modo intra, un modo inter, y un modo de salto, y un modo de predicción específica puede llevarse a cabo por sólo una unidad de predicción que tiene un tamaño o forma específica. Por ejemplo, el modo intra puede realizarse solo para unidades de predicción que tienen tamaños de  $2N \times 2N$  y  $N \times N$ , cuya forma es un cuadrado. Además, el modo de omisión puede realizarse solo para una unidad de predicción que tiene un tamaño de  $2N \times 2N$ . Si existe una pluralidad de unidades de predicción en una unidad de codificación, el modo de predicción con los menores errores de codificación puede seleccionarse después de realizar la predicción para cada unidad de predicción.

20 Alternativamente, el aparato 100 puede llevar a cabo una transformación de frecuencia en datos de imagen en base a una unidad de procesamiento que tiene un tamaño diferente de una unidad de codificación. Para la transformación de frecuencia en la unidad de codificación, la transformación de frecuencia puede realizarse basándose en una unidad de procesamiento que tiene un tamaño igual o menor que el de la unidad de codificación. En lo sucesivo, una unidad de procesamiento, que es la base de la transformación de frecuencia, se denomina unidad de transformación. La transformación de frecuencia puede ser una Transformación discreta de coseno (DCT) o una Transformación de Karhunen-Loeve (KLT).

30 El determinante 120 de profundidad de codificación puede determinar las unidades de subcodificación incluidas en una unidad máxima de codificación mediante la optimización R-D basado en un multiplicador de Lagrange. En otras palabras, el determinante 120 de profundidad de codificación puede determinar qué forma tiene una pluralidad de unidades de subcodificación divididas de la unidad de codificación máxima, en el que la pluralidad de unidades de subcodificación tienen diferentes tamaños según sus profundidades. El codificador 130 de datos de imagen emite una corriente de bits codificando la unidad de codificación máxima en base a las formas de división determinadas por el determinante 120 de profundidad de codificación.

35 El codificador 140 de información de codificación codifica la información acerca de un modo de codificación de la unidad máxima de codificación determinado por el determinante 120 de profundidad de codificación. En otras palabras, el codificador 140 de información de codificación emite una corriente de bits al codificar información sobre una forma de división de la unidad de codificación máxima, información sobre la profundidad máxima e información sobre un modo de codificación de una unidad de subcodificación para cada profundidad. La información sobre el modo de codificación de la unidad de subcodificación puede incluir al menos una información sobre una unidad de predicción de la unidad de subcodificación, información sobre un modo de predicción para cada unidad de predicción e información sobre una unidad de transformación de la unidad de subcodificación.

Dado que las unidades de subcodificación que tienen existen diferentes tamaños para cada unidad de codificación máximo e información de un modo de codificación se determinan para cada unidad de subcodificación, información sobre al menos un modo de codificación puede ser determinada por una unidad máxima de codificación.

45 El aparato 100 puede generar unidades de subcodificación dividiendo igualmente tanto la altura y la anchura de una unidad máxima de codificación por dos de acuerdo con un aumento de la profundidad. Es decir, cuando el tamaño de una unidad de codificación de una  $k$ -ésima profundidad es  $2N \times 2N$ , el tamaño de una unidad de codificación de una profundidad  $(k + 1)$  puede ser  $N \times N$ .

50 Por consiguiente, el aparato 100 de acuerdo con una realización ejemplar puede determinar una forma de división óptima para cada unidad de codificación máxima basada en tamaños de unidades de codificación máximas y una profundidad máxima en consideración de las características de la imagen. Al ajustar de forma variable el tamaño de una unidad de codificación máxima en consideración de las características de la imagen y codificar una imagen a través de la división de una unidad de codificación máxima en unidades de subcodificación de diferentes profundidades, las imágenes que tienen varias resoluciones pueden codificarse de manera más eficiente.

55 La figura 2 es un diagrama de bloques de un aparato 200 para decodificar una imagen de acuerdo con una realización ejemplar. Con referencia a la figura 2, el aparato 200 incluye una unidad 210 de adquisición de datos de imagen, un extractor 220 de información de codificación y un decodificador 230 de datos de imagen.

La unidad 210 de adquisición de datos de imagen adquiere datos de imagen de acuerdo a las unidades máximas de

- codificación mediante el análisis de una corriente de bits recibida por el aparato 200 y emite los datos de imagen al decodificador 230 de datos de imagen. La unidad 210 de adquisición de datos de imagen puede extraer información sobre una unidad de codificación máxima de una imagen o porción actual desde un encabezado de la imagen o porción actual. En otras palabras, la unidad 210 de adquisición de datos de imagen divide la corriente de bits en la
- 5 unidad de codificación máxima para que el decodificador 230 de datos de imagen pueda decodificar los datos de imagen de acuerdo con las unidades de codificación máximas.
- El extractor 220 de información de codificación extrae información acerca de una unidad máxima de codificación, una profundidad máxima, una forma de división de la unidad máxima de codificación, y un modo de codificación de unidades de subcodificación mediante el análisis de la corriente de bits recibido por el aparato 200. Por ejemplo, el
- 10 extractor 220 de información de codificación puede extraer la información descrita anteriormente del encabezado de la imagen actual. La información sobre la forma de división y la información sobre el modo de codificación se proporcionan al decodificador 230 de datos de imagen.
- La información sobre la forma división de la unidad máxima de codificación puede incluir información sobre unidades de subcodificación que tienen diferentes tamaños de acuerdo con profundidades incluidos en la unidad máxima de
- 15 codificación, y la información sobre el modo de codificación puede incluir al menos una de información acerca de una unidad de predicción según la unidad de subcodificación, información sobre un modo de predicción e información sobre una unidad de transformación.
- El decodificador 230 de datos de imagen restaura la imagen actual mediante la decodificación de datos de imagen de cada unidad máxima de codificación en base a la información extraída por el extractor 220 de información de
- 20 codificación. El decodificador 230 de datos de imagen puede decodificar unidades de subcodificación incluidas en una unidad de codificación máxima basándose en la información sobre la forma de división de la unidad de codificación máxima. Un procedimiento de decodificación puede incluir al menos uno de un procedimiento de predicción que incluye predicción intra y compensación de movimiento y un procedimiento de transformación inversa.
- Por otra parte, el decodificador 230 de datos de imagen puede realizar la predicción intra o predicción inter basada en la información sobre la unidad de predicción y la información sobre el modo de predicción a fin de predecir una
- 25 unidad de predicción. El decodificador 230 de datos de imagen también puede realizar una transformación inversa para cada unidad de subcodificación basándose en la información sobre la unidad de transformación de una unidad de subcodificación.
- La figura 3 ilustra unidades de codificación jerárquica de acuerdo con una realización ejemplar. Con referencia a la figura 3, las unidades de codificación jerárquicas ejemplares incluyen unidades de codificación cuyos tamaños son
- 30  $64 \times 64$ ,  $32 \times 32$ ,  $16 \times 16$ ,  $8 \times 8$  y  $4 \times 4$ . Además, también pueden existir unidades de codificación cuyos tamaños son  $64 \times 32$ ,  $32 \times 64$ ,  $32 \times 16$ ,  $16 \times 32$ ,  $16 \times 8$ ,  $8 \times 16$ ,  $8 \times 4$  y  $4 \times 8$ .
- En la realización ilustrada en la figura 3, para los primeros datos 310 de imagen cuya resolución es  $1920 \times 1080$ , el tamaño de una unidad de codificación máxima se establece en  $64 \times 64$ , y la profundidad máxima se establece en 2. Para los datos de la segunda imagen 320 cuya resolución es  $1920 \times 1080$ , el tamaño de una unidad de codificación
- 35 máxima se establece en  $64 \times 64$ , y la profundidad máxima se establece en 3. Para los datos de la tercera imagen 330 cuya resolución es  $352 \times 288$ , el tamaño de una unidad de codificación máxima se establece en  $16 \times 16$ , y la profundidad máxima se establece en 1.
- Cuando la resolución es alta o la cantidad de datos es grande, un tamaño máximo de una unidad de codificación puede ser relativamente grande para aumentar una relación de compresión y reflejar exactamente características de la imagen. En consecuencia, para los datos 310 y 320 de imagen primero y segundo que tienen una resolución más alta que el tercer dato 330 de imagen, se puede seleccionar  $64 \times 64$  como el tamaño de la unidad de codificación
- 40 máxima.
- Una profundidad máxima indica el número total de capas en las unidades de codificación jerárquica. Como la profundidad máxima de los primeros datos 310 de imagen es 2, una unidad 315 de codificación de los datos 310 de imagen puede incluir una unidad de codificación máxima cuyo tamaño de eje más largo es 64 y unidades de subcodificación cuyos tamaños de eje más largos son 32 y 16, de acuerdo con un aumento de una profundidad.
- Por otro lado, puesto que la profundidad máxima de los terceros datos 330 de imagen es 1, una unidad 335 de
- 50 codificación de los datos 330 de imagen puede incluir una unidad de codificación de máxima cuyo eje más largo tamaño es 16 y las unidades de codificación cuyos tamaños eje más largo es 8, según un aumento de una profundidad.
- Sin embargo, puesto que la profundidad máxima de los segundos datos 320 de imagen es 3, una unidad 325 de
- 55 codificación de los datos 320 de imagen puede incluir una unidad de codificación de máxima cuyo eje más largo tamaño es 64 y las unidades de subcodificación cuyos tamaños eje más largo son 32, 16 y 8 según un aumento de una profundidad. Dado que una imagen se codifica en base a una unidad de subcodificación más pequeña a medida que aumenta la profundidad, las realizaciones ejemplares son adecuadas para codificar una imagen que incluye más escenas diminutas.

- La figura 4 es un diagrama de bloques de un codificador 400 de imagen basado en una unidad de codificación, de acuerdo con una realización ejemplar. Con referencia a la figura 4, un predictor 410 intra realiza predicción intra en unidades de predicción del modo intra en un cuadro 405 actual, y un estimador 420 de movimiento y un compensador 425 de movimiento realizan predicción inter y compensación de movimiento en unidades de predicción del modo inter usando el cuadro 405 actual y un cuadro 495 de referencia.
- Los valores residuales se generan basándose en la salida de unidades de predicción desde el predictor 410 intra, el estimador 420 de movimiento y el compensador 425 de movimiento. Los valores residuales generados se emiten como coeficientes de transformación cuantificados pasando a través de un transformador 430 y un cuantificador 440.
- Los coeficientes de transformación cuantificados se restauran a los valores residuales mediante el paso a través de un cuantificador 460 inverso y un transformador 470 inverso. Los valores residuales restaurados se procesan posteriormente pasando a través de una unidad 480 de desbloqueo y una unidad 490 de filtrado de bucle y se emiten como el cuadro 495 de referencia. Los coeficientes de transformación cuantificados pueden salir como un flujo 455 de bits pasando a través de un codificador 450 de entropía.
- Para realizar la codificación basada en un procedimiento de codificación de acuerdo con una realización de ejemplo, los componentes del codificador 400 de imagen, es decir, el predictor 410 intra, el estimador 420 de movimiento, el compensador 425 de movimiento, el transformador 430, el cuantificador 440, el codificador 450 de entropía, el cuantificador 460 inverso, el transformador 470 inverso, la unidad 480 de desbloqueo y la unidad 490 de filtrado de bucle, realizan procedimientos de codificación de imagen basados en una unidad de codificación máxima, una unidad de subcodificación según las profundidades, una unidad de predicción, y una unidad de transformación.
- La figura 5 es un diagrama de bloques de un decodificador 500 de imagen basado en una unidad de codificación, de acuerdo con una realización ejemplar. Con referencia a la figura 5, un flujo 505 de bits pasa a través de un analizador 510 de modo que se analizan los datos de imagen codificados para decodificar y se analiza la información de codificación utilizada para decodificar. Los datos de imagen codificados se emiten como datos cuantificados inversamente pasando a través de un decodificador 520 de entropía y un cuantificador 530 inverso y restaurados a valores residuales pasando a través de un transformador 540 inverso. Los valores residuales se restauran de acuerdo con las unidades de codificación mediante la adición a un resultado de predicción intra de un predictor 550 intra o un resultado de compensación de movimiento de un compensador 560 de movimiento. Las unidades de codificación restauradas se usan para predecir las siguientes unidades de codificación o una imagen siguiente pasando a través de una unidad 570 de desbloqueo y una unidad 580 de filtrado de bucle.
- Para realizar la decodificación basado en un procedimiento de decodificación de acuerdo con un ejemplo de realización, los componentes del decodificador 500 de imagen, es decir, el analizador 510, el decodificador 520 de entropía, el cuantificador 530 inverso, el transformador 540 inverso, el predictor 550 intra, el compensador 560 de movimiento, la unidad 570 de desbloqueo y la unidad 580 de filtrado de bucle, realizan procedimientos de decodificación de imágenes basados en una unidad de codificación máxima, una unidad de subcodificación según las profundidades, una unidad de predicción y una unidad de transformación.
- En particular, el predictor 550 intra y el compensador 560 de movimiento determina una unidad de predicción y un modo de predicción en una unidad de subcodificación considerando una unidad de codificación máxima y una profundidad, y el transformador 540 inverso realiza la transformación inversa considerando el tamaño de una unidad de transformación.
- La figura 6 ilustra una unidad de codificación máxima, una unidad de subcodificación y una unidad de predicción, de acuerdo con una realización ejemplar.
- Como se ha descrito anteriormente, el aparato 100 de codificación y el aparato 200 de decodificación de acuerdo con una o más realizaciones ejemplares utilizan unidades de codificación jerárquica para llevar a cabo la codificación y decodificación en consideración de las características de imagen. Una unidad de codificación máxima y una profundidad máxima se pueden configurar de forma adaptativa según las características de la imagen o se pueden configurar de acuerdo con los requisitos de un usuario.
- Con referencia a la figura 6, una estructura 600 de unidad de codificación jerárquica de acuerdo con una realización ejemplar ilustra una unidad 610 de codificación máxima cuya altura y anchura son 64 y la profundidad máxima es 4. Una profundidad aumenta a lo largo de un eje vertical de la estructura 600 de unidad de codificación jerárquica, y a medida que aumenta la profundidad, las alturas y anchuras de las unidades 620 a 650 de subcodificación disminuyen. Las unidades de predicción de la unidad 610 de codificación máxima y las unidades 620 a 650 de subcodificación se muestran a lo largo de un eje horizontal de la estructura 600 de unidad de codificación jerárquica.
- La unidad 610 de codificación máxima tiene una profundidad de 0 y un tamaño, es decir, altura y anchura, de  $64 \times 64$ . La profundidad aumenta a lo largo del eje vertical, de modo que existe una unidad 620 de subcodificación cuyo tamaño es  $32 \times 32$  y la profundidad es 1, una unidad 630 de subcodificación cuyo tamaño es  $16 \times 16$  y la profundidad es 2, una unidad 640 de subcodificación cuyo tamaño es  $8 \times 8$  y la profundidad es 3, y una unidad 650 de subcodificación cuyo tamaño es  $4 \times 4$  y la profundidad es 4. La unidad 650 de subcodificación cuyo tamaño es  $4 \times 4$  y la profundidad es 4 es una unidad mínima de codificación. La unidad 650 de codificación mínima puede

dividirse en unidades de predicción, cada una de las cuales es menor que la unidad de codificación mínima.

En la realización ilustrada en la figura 6, se muestran ejemplos de una unidad de predicción a lo largo del eje horizontal de acuerdo con cada profundidad. Es decir, una unidad de predicción de la unidad 610 de codificación máxima cuya profundidad es 0 puede ser una unidad de predicción cuyo tamaño es igual a la unidad 610 de codificación, es decir,  $64 \times 64$ , o una unidad 612 de predicción cuyo tamaño es  $64 \times 32$ , una unidad 614 de predicción cuyo tamaño es  $32 \times 64$ , o una unidad 616 de predicción cuyo tamaño es  $32 \times 32$ , que tiene un tamaño menor que la unidad 610 de codificación cuyo tamaño es  $64 \times 64$ .

Una unidad de predicción de la unidad 620 de codificación cuya profundidad es 1 y el tamaño es de  $32 \times 32$  puede ser una unidad de predicción cuyo tamaño es igual a la unidad 620 de codificación, es decir,  $32 \times 32$ , o una unidad 622 de predicción cuyo tamaño es  $32 \times 16$ , una unidad 624 de predicción cuyo tamaño es  $16 \times 32$ , o una unidad 626 de predicción cuyo tamaño es  $16 \times 16$ , que tiene un tamaño menor que la unidad 620 de codificación cuyo tamaño es  $32 \times 32$ .

Una unidad de predicción de la unidad 630 de codificación cuya profundidad es 2 y el tamaño es de  $16 \times 16$  puede ser una unidad de predicción cuyo tamaño es igual a la unidad 630 de codificación, es decir,  $16 \times 16$ , o una unidad 632 de predicción cuyo tamaño es  $16 \times 8$ , una unidad 634 de predicción cuyo tamaño es  $8 \times 16$ , o una unidad de predicción 636 cuyo tamaño es  $8 \times 8$ , que tiene un tamaño menor que la unidad 630 de codificación cuyo tamaño es  $16 \times 16$ .

Una unidad de predicción de la unidad 640 de codificación cuya profundidad es 3 y el tamaño es de  $8 \times 8$  puede ser una unidad de predicción cuyo tamaño es igual a la unidad 640 de codificación, es decir,  $8 \times 8$ , o una unidad 642 de predicción, cuyo tamaño es de  $8 \times 4$ , una unidad 644 de predicción cuyo tamaño es  $4 \times 8$ , o una unidad 646 de predicción cuyo tamaño es  $4 \times 4$ , que tiene un tamaño menor que la unidad 640 de codificación cuyo tamaño es  $8 \times 8$ .

La unidad 650 de codificación cuya profundidad es 4 y el tamaño es de  $4 \times 4$  es una unidad de codificación mínimo y una unidad de codificación de una profundidad máxima. Una unidad de predicción de la unidad 650 de codificación puede ser una unidad de predicción 650 cuyo tamaño es  $4 \times 4$ , una unidad 652 de predicción que tiene un tamaño de  $4 \times 2$ , una unidad 654 de predicción que tiene un tamaño de  $2 \times 4$  o una unidad 656 de predicción que tiene un tamaño de  $2 \times 2$ .

La figura 7 ilustra una unidad de codificación y una unidad de transformación, de acuerdo con una realización ejemplar. El aparato 100 de codificación y el aparato 200 de decodificación, de acuerdo con una o más realizaciones ejemplares, realizan la codificación con una unidad de codificación máxima en sí misma o con unidades de subcodificación, que son iguales o menores que la unidad de codificación máxima y divididas de la unidad de codificación máxima.

En el procedimiento de codificación, el tamaño de una unidad de transformación para la transformación de frecuencia se selecciona para ser mayor que la de una unidad de codificación correspondiente. Por ejemplo, cuando una unidad 710 de codificación actual tiene un tamaño de  $64 \times 64$ , la transformación de frecuencia puede realizarse usando una unidad 720 de transformación que tiene un tamaño de  $32 \times 32$ .

Las figuras 8A y 8B ilustran formas de división de una unidad de codificación, una unidad de predicción y una unidad de transformación, de acuerdo con una realización ejemplar. La figura 8A ilustra una unidad de codificación y una unidad de predicción, de acuerdo con una realización ejemplar.

Un lado izquierdo de la figura 8A muestra una forma de división seleccionada por un aparato 100 de codificación de acuerdo con una realización ejemplar para codificar una unidad 810 de codificación máxima. El aparato 100 divide la unidad 810 de codificación máxima en varias formas, realiza la codificación y selecciona una forma de división óptima comparando los resultados de codificación de varias formas de división entre sí en base al coste de R-D. Cuando es óptimo que la unidad 810 de codificación máxima esté codificada tal cual, la unidad 810 de codificación máxima puede codificarse sin dividir la unidad 810 de codificación máxima como se ilustra en las figuras 8A y 8B.

Con referencia al lado izquierdo de la figura 8A, la unidad 810 de codificación máxima cuya profundidad es 0 se codifica dividiendo la unidad de codificación máxima en unidades de subcodificación cuyas profundidades son iguales o mayores que 1. Es decir, la unidad 810 de codificación máxima se divide en 4 unidades de subcodificación cuyas profundidades son 1, y todas o algunas de las unidades de subcodificación cuyas profundidades son 1 se dividen en unidades de subcodificación cuyas profundidades son 2.

Una unidad de subcodificación situada en un lado superior derecho y una unidad de subcodificación situada en un lado inferior izquierdo entre las unidades de subcodificación cuyas profundidades son 1 están divididos en unidades de subcodificación cuyas profundidades son igual o mayor que 2. Algunas de las unidades de subcodificación cuyas profundidades son iguales o superiores a 2 pueden dividirse en unidades de subcodificación cuyas profundidades son iguales o superiores a 3.

El lado derecho de la figura 8A muestra una forma de división de una unidad de predicción para la unidad 810 de

codificación máxima. Con referencia al lado derecho de la figura 8A, una unidad 860 de predicción para la unidad 810 de codificación máxima puede dividirse de manera diferente de la unidad 810 de codificación máxima. En otras palabras, una unidad de predicción para cada una de las unidades de subcodificación puede ser más pequeña que una unidad de subcodificación correspondiente.

5 Por ejemplo, una unidad de predicción para una unidad 854 de subcodificación situada en un lado inferior derecho entre las unidades de subcodificación cuyas profundidades son de 1 puede ser menor que la unidad 854 de subcodificación. Además, las unidades de predicción para algunas unidades 814, 816, 850 y 852 de subcodificación de las unidades 814, 816, 818, 828, 850 y 852 de subcodificación cuyas profundidades son 2 pueden ser más pequeñas que las unidades 814, 816, 850 y 852 de subcodificación, respectivamente. Además, las unidades de predicción para las unidades 822, 832 y 848 de subcodificación cuyas profundidades son 3 pueden ser más pequeñas que las unidades 822, 832 y 848 de subcodificación, respectivamente. Las unidades de predicción pueden tener una forma en la que las respectivas unidades de subcodificación se dividan por igual en dos en una dirección de altura o anchura o pueden tener una forma en la que las unidades de subcodificación respectivas se dividan igualmente por cuatro en direcciones de altura y anchura.

15 La figura 8B ilustra una unidad de predicción y una unidad de transformación, de acuerdo con una realización ejemplar. Un lado izquierdo de la figura 8B muestra una forma de división de una unidad de predicción para la unidad 810 de codificación máxima mostrada en el lado derecho de la figura 8A, y un lado derecho de la figura 8B muestra una forma de división de una unidad de transformación de la unidad 810 de codificación máxima.

20 Con referencia al lado derecho de la figura 8B, una forma de división de una unidad 870 de transformación puede establecerse de manera diferente a la unidad 860 de predicción. Por ejemplo, aunque una unidad de predicción para la unidad 854 de codificación cuya profundidad es 1 se selecciona con una forma por la cual la altura de la unidad 854 de codificación se divide por igual entre dos, una unidad de transformación se puede seleccionar con el mismo tamaño que la unidad 854 de codificación. Del mismo modo, a pesar de que las unidades de predicción para las unidades 814 y 850 de codificación cuyas profundidades son 2 se seleccionan con una forma en la que la altura de cada una de las unidades 814 y 850 de codificación se divide por igual, se puede seleccionar una unidad de transformación con el mismo tamaño que El tamaño original de cada una de las unidades 814 y 850 de codificación.

25 Una unidad de transformación se puede seleccionar con un tamaño más pequeño que una unidad de predicción. Por ejemplo, cuando una unidad de predicción para la unidad 852 de codificación cuya profundidad es 2 se selecciona con una forma por la cual el ancho de la unidad 852 de codificación se divide por igual por dos, se puede seleccionar una unidad de transformación con una forma por la cual la unidad 852 de codificación es igualmente dividida por cuatro en direcciones de altura y anchura, que tiene un tamaño menor que la forma de la unidad de predicción.

30 La figura 9 es un diagrama de bloques de un aparato 900 para codificar una imagen, de acuerdo con una realización ejemplar. El aparato 900 para codificar un vector de movimiento puede incluirse en el aparato 100 descrito anteriormente con referencia a la figura 1 o el codificador 400 de imagen descrito anteriormente con referencia a la figura 4. Con referencia a la figura 9, el aparato 900 de codificación del vector de movimiento incluye un predictor 910, un primer codificador 920 y un segundo codificador 930.

35 Con el fin de descodificar un bloque codificado usando predicción inter, es decir, la predicción entre imágenes, se utiliza la información sobre un vector de movimiento que indica una diferencia de posición entre un bloque actual y un bloque similar en una imagen de referencia. Por lo tanto, la información sobre los vectores de movimiento se codifica y se inserta en una corriente de bits en un procedimiento de codificación de imágenes. Sin embargo, si la información sobre los vectores de movimiento se codifica e inserta como está, aumenta la sobrecarga para codificar la información sobre los vectores de movimiento, disminuyendo así la relación de compresión de los datos de imagen.

40 Por lo tanto, en un procedimiento de codificación de imagen, la información sobre un vector de movimiento se comprime mediante la predicción de un vector de movimiento de un bloque actual, que codifica solamente un vector diferencial entre un predictor del vector de movimiento generado como resultado de la predicción y un vector de movimiento original, e insertando el vector diferencial codificado en una corriente de bits. La figura 9 ilustra un aparato 900 para codificar un vector de movimiento que usa dicho predictor del vector de movimiento.

45 Con referencia a la figura 9, el predictor 910 determina si un vector de movimiento de un bloque actual está codificado por predicción en base a un modo explícito o un modo implícito.

50 Como se describió anteriormente, un codificador-decodificador, tal como MPEG-4 H.264/MPEG-4 AVC utiliza vectores de movimiento de bloques previamente codificados adyacente a un bloque actual para predecir un vector de movimiento del bloque actual. Es decir, se utiliza una mediana de vectores de movimiento de bloques codificados previamente adyacentes a los lados izquierdo, superior y superior derecho del bloque actual como un predictor del vector de movimiento del bloque actual. Dado que los vectores de movimiento de todos los bloques codificados usando predicción inter se predicen usando el mismo procedimiento, la información sobre un predictor del vector de movimiento no tiene que codificarse por separado. Sin embargo, el aparato 100 o el decodificador 400 de imagen de acuerdo con una o más realizaciones ejemplares usa un modo en el que la información sobre un predictor del vector



de movimiento no se codifica por separado y un modo en el que la información sobre un predictor del vector de movimiento se codifica para predecir más exactamente un vector de movimiento, que ahora se describirá en detalle.

(1) Modo explícito

5 Uno de los procedimientos de codificación de un predictor del vector de movimiento, que puede ser seleccionado por el predictor 910, puede implementar un modo de codificar explícitamente la información sobre un predictor del vector de movimiento de un bloque actual. Este modo explícito es un modo de calcular al menos un candidato predictor del vector de movimiento y codificar por separado la información que indica qué predictor del vector de movimiento se utiliza para predecir un vector de movimiento de un bloque actual. Los candidatos predictores del vector de movimiento según una o más realizaciones ejemplares se describirán ahora con referencia a las figuras 10A, 10B y 10C.

10 Las figuras 10A y 10B ilustran candidatos predictores del vector de movimiento de un modo explícito, de acuerdo con una o más realizaciones ejemplares. Con referencia a la figura 10A, un procedimiento de predicción de vector de movimiento de acuerdo con una realización ejemplar puede usar uno de los vectores de movimiento de bloques codificados previamente adyacentes a un bloque actual como un predictor del vector de movimiento del bloque actual. Un bloque a0 en el extremo izquierdo entre los bloques adyacentes a un lado superior del bloque actual, un bloque b0 en el extremo superior entre los bloques adyacentes a un lado izquierdo del mismo, un bloque c adyacente a un lado superior derecho del mismo, un bloque d adyacente a un lado superior izquierdo del mismo, y un bloque e adyacente a un lado inferior izquierdo del mismo puede usarse para predictores del vector de movimiento del bloque actual.

20 Con referencia a la figura 10B, los vectores de movimiento de todos los bloques adyacentes a un bloque actual pueden usarse como predictores del vector de movimiento del bloque actual. En otras palabras, los vectores de movimiento no solo de un bloque a0 en el extremo izquierdo entre los bloques adyacentes a un lado superior del bloque actual, sino que todos los bloques adyacentes al lado superior del mismo pueden usarse como predictores de vectores de movimiento del bloque actual. Además, los vectores de movimiento de no solo un bloque b0 en la parte superior entre los bloques adyacentes a un lado izquierdo del mismo, sino que todos los bloques adyacentes al lado izquierdo del mismo pueden usarse como predictores del vector de movimiento del bloque actual.

25 Como alternativa, se puede usar un valor medio de vectores de movimiento de bloques adyacentes como un predictor del vector de movimiento. Por ejemplo, la mediana (mv\_ a0, mv\_ b0, mv\_c) puede usarse como un predictor del vector de movimiento del bloque actual, en el que mv\_ a0 denota un vector de movimiento del bloque a0, mv\_ b0 denota un vector de movimiento del bloque b0 y mv\_c denota un vector de movimiento del bloque c.

30 Las figuras 11A a 11C ilustran candidatos predictores del vector de movimiento de un modo explícito, de acuerdo con otra realización ejemplar. La figura 11A ilustra un procedimiento para calcular un predictor del vector de movimiento de una imagen predictiva bidireccional (denominada imagen B), de acuerdo con una realización ejemplar. Cuando una imagen actual que incluye un bloque actual es una imagen B en la que se realiza una predicción bidireccional, un vector de movimiento generado en base a una distancia temporal puede ser un predictor del vector de movimiento.

35 Con referencia a la figura 11A, se puede generar un predictor del vector de movimiento de un bloque 1100 actual de una imagen 1110 actual usando un vector de movimiento de un bloque 1120 en una posición de ubicación conjunta de una imagen 1112 temporalmente anterior. Por ejemplo, si se genera un vector de movimiento mv\_colA del bloque 1120 en una posición de ubicación conjunta con el bloque 1100 actual para un bloque 1122 buscado de una imagen 1114 siguiente temporalmente de la imagen 1110 actual, los candidatos de predictores del vector de movimiento mv\_L0A y mv\_L1A del bloque 1100 actual se puede generar de acuerdo con las ecuaciones siguientes:

$$mv\_L1A=(t1/t2)*mv\_colA$$

$$mv\_L0A=mv\_L1A-mv\_colA$$

45 donde mv\_L0A denota un predictor del vector de movimiento del bloque 1100 actual para la imagen 1112 precedente temporalmente, y mv\_L1A denota un predictor del vector de movimiento del bloque 1100 actual para la imagen 1114 siguiente temporalmente.

50 La figura 11B ilustra un procedimiento para generar un predictor del vector de movimiento de una imagen B, de acuerdo con otra realización ejemplar. En comparación con el procedimiento ilustrado en la figura 11A, un bloque 1130 en una posición co-ubicada con el bloque 1100 actual existe en la imagen 1114 siguiente temporalmente en la figura 11B.

55 Con referencia a la figura 11B, se puede generar un predictor del vector de movimiento del bloque 1100 actual de la imagen 1110 actual usando un vector de movimiento de un bloque 1130 en una posición de ubicación conjunta de la imagen 1114 siguiente temporalmente. Por ejemplo, si se genera un vector de movimiento mv\_colB del bloque 1130 en una posición de ubicación conjunta con el bloque 1100 actual para un bloque 1132 buscado de la imagen 1112 precedente temporalmente de la imagen 1110 actual, los candidatos de predictores del vector de movimiento

mv\_L0B y mv\_L1B del bloque 1100 actual se puede generar de acuerdo con las ecuaciones siguientes:

$$mv\_L0B=(t3/t4)*mv\_colB$$

$$mv\_L1B=mv\_L0B-mv\_colB$$

5 donde mv\_L0B denota un predictor del vector de movimiento del bloque 1100 actual para la imagen 1112 precedente temporalmente, y mv\_L1B denota un predictor del vector de movimiento del bloque 1100 actual para la imagen 1114 siguiente temporalmente.

10 En la generación de un vector de movimiento del bloque 1100 actual de una imagen B, se puede usar al menos uno de los procedimientos ilustrados en las figuras 11A y 11B. En otras palabras, dado que se genera un predictor del vector de movimiento usando un vector de movimiento y una distancia temporal del bloque 1120 o 1130 en una posición de ubicación conjunta con el bloque 1100 actual, se pueden generar predictores del vector de movimiento usando los procedimientos ilustrados en las figuras 11A y 11B si existen vectores de movimiento de los bloques 1120 y 1130 en la posición de ubicación conjunta. Por lo tanto, el predictor 910 de acuerdo con una realización ejemplar puede generar un predictor del vector de movimiento del bloque 1100 actual usando solo un bloque que tiene un vector de movimiento entre los bloques 1120 y 1130 en la posición de ubicación conjunta.

15 Por ejemplo, cuando el bloque 1120 en una posición de ubicación conjunta de la imagen 1112 precedente temporalmente se codifica usando predicción intra en lugar de predicción inter, un vector de movimiento del bloque 1120 no existe, y por tanto un predictor del vector de movimiento del bloque 1100 actual no puede generarse usando el procedimiento de generar un predictor del vector de movimiento como se ilustra en la figura 11A.

20 La figura 11C ilustra un procedimiento para generar un predictor del vector de movimiento de una imagen B, de acuerdo con una realización ejemplar. Con referencia a la figura 11C, se puede generar un predictor del vector de movimiento del bloque 1100 actual de la imagen 1110 actual usando un vector de movimiento de un bloque 1140 en una posición de ubicación conjunta de la imagen 1112 temporalmente anterior. Por ejemplo, si se genera un vector de movimiento mv\_colC del bloque 1130 en una posición de ubicación conjunta con el bloque 1100 actual para un bloque 1142 buscado de otra imagen 1116 temporalmente anterior, se puede obtener un candidato predictor del vector de movimiento mv\_L0C del bloque 1100 actual generado de acuerdo con la siguiente ecuación:

$$mv\_L0C=(t6/t5)*mv\_colC.$$

Puesto que la imagen 1110 actual es una imagen P, el número de predictores del vector de movimiento del bloque 1100 actual es 1, a diferencia de las figuras 11A y 11B.

30 En resumen, un conjunto C de candidatos de predicción de vectores de movimiento según las figuras 10A, 10B y 11A a 11C se pueden generar de acuerdo con la siguiente ecuación:

$$C = \{\text{median}(mv\_a0, mv\_b0, mv\_c), mv\_a0, mv\_a1, \dots, mv\_aN, mv\_b0, mv\_b1, \dots, mv\_bN, mv\_c, mv\_d, mv\_e, mv\_temporal\}.$$

Alternativamente, el conjunto C puede generarse reduciendo el número de candidatos predictores del vector de movimiento de acuerdo con la siguiente ecuación:

$$C = \{\text{median}(mv\_a', mv\_b', mv\_c'), mv\_a', mv\_b', mv\_c', mv\_temporal\}.$$

35 Aquí, mv\_x denota un vector de movimiento de un bloque x, mediana () denota un valor mediano, y mv\_temporal denota candidatos predictores del vector de movimiento generados usando una distancia temporal descrita anteriormente en asociación con las figuras 11A a 11C.

40 Además, mv\_a' representa un vector de movimiento primera válida entre mv\_a0, mv\_a1 ..., mv\_aN. Por ejemplo, cuando un bloque a0 se codifica utilizando predicción intra, un vector de movimiento mv\_a0 del bloque a0 no es válido, y por lo tanto mv\_a' = mv\_a1, y si un vector de movimiento de un bloque a1 tampoco es válido, mv\_a' = mv\_a2.

Del mismo modo, mv\_b' denota un primer vector de movimiento válido entre mv\_b0, mv\_b1 ..., mv\_bN, y mv\_c' denota un primer vector de movimiento válido entre mv\_c, mv\_d, y mv\_e.

45 El modo explícito es un modo de codificación de la información que indica que el vector de movimiento se ha utilizado para un predictor del vector de movimiento de un bloque actual. Por ejemplo, cuando un vector de movimiento se codifica en el modo explícito, se puede asignar un número binario a cada uno de los elementos del conjunto C, es decir, los candidatos predictores del vector de movimiento, y si uno de los candidatos se utiliza como un predictor del vector de movimiento de un bloque actual, se puede emitir un número binario correspondiente.

50 Los expertos en la técnica entenderán fácilmente que se pueden usar otros candidatos predictores de vectores de movimiento además de los descritos anteriormente en asociación con el modo explícito.

## (2) Modo implícito

Otro de los procedimientos de codificación de un predictor del vector de movimiento, que puede ser seleccionado por el predictor 910, implementa un modo de codificación de la información que indica que un predictor del vector de movimiento de un bloque actual se genera sobre la base de bloques o píxeles incluidos en un área previamente codificada adyacente al bloque actual. A diferencia del modo explícito, este modo es un modo de codificación de información que indica la generación de un predictor del vector de movimiento en el modo implícito sin codificar información para especificar un predictor del vector de movimiento.

Como se describió anteriormente, un codificador-decodificador, tal como MPEG-4 H.264/MPEG-4 AVC utiliza vectores de movimiento de bloques previamente codificados adyacente a un bloque actual para predecir un vector de movimiento del bloque actual. Es decir, se utiliza una mediana de vectores de movimiento de bloques codificados previamente adyacentes a los lados izquierdo, superior y superior derecho del bloque actual como un predictor del vector de movimiento del bloque actual. En este caso, a diferencia del modo explícito, la información para seleccionar uno de los candidatos predictores del vector de movimiento puede no estar codificada.

En otras palabras, si sólo la información que indica que un predictor del vector de movimiento de un bloque actual ha sido codificada en el modo implícito se codifica en un procedimiento de codificación de imagen, un valor de la mediana de los vectores de movimiento de bloques previamente codificados adyacentes a izquierda, superior, y los lados superior derecho del bloque actual se pueden usar como un predictor del vector de movimiento del bloque actual en un procedimiento de decodificación de imágenes.

Además, un procedimiento de codificación de imágenes según una realización ejemplar proporciona un nuevo modo implícito además del procedimiento de utilización de un valor de la mediana de los vectores de movimiento de bloques previamente codificados adyacentes a izquierda, superior, y los lados superiores derecho de un bloque actual como un predictor del vector de movimiento del bloque actual. Esto se describirá ahora en detalle con referencia a la figura 12.

La figura 12 ilustra un procedimiento para generar un predictor del vector de movimiento en un modo implícito, de acuerdo con una realización ejemplar. Con referencia a la figura 12, los píxeles 1222 incluidos en un área 1220 previamente codificada adyacente a un bloque 1200 actual de una imagen 1210 actual se usan para generar un predictor del vector de movimiento del bloque 1200 actual. Los píxeles 1224 correspondientes se determinan buscando una imagen 1212 de referencia usando los píxeles 1222 adyacentes. Los píxeles 1224 correspondientes se pueden determinar calculando una Suma de Diferencias Absolutas (SAD). Cuando se determinan los píxeles 1224 correspondientes, se genera un vector de movimiento  $mv\_plantilla$  de los píxeles 1222 adyacentes, y el vector de movimiento  $mv\_plantilla$  puede usarse como un predictor del vector de movimiento del bloque 1200 actual.

Si un modo de usar una mediana de vectores de movimiento de bloques adyacentes como un predictor del vector de movimiento se define como "modo implícito\_1", y si un modo de generar un predictor del vector de movimiento usando píxeles adyacentes a un bloque actual se define como " modo implícito\_2 ", se puede generar un predictor del vector de movimiento usando uno de los dos modos implícitos modo implícito\_1 y modo implícito\_2 codificando información sobre uno de los dos modos implícitos en un procedimiento de codificación de imagen y haciendo referencia a la información sobre un modo en una decodificación de imagen procedimiento.

## (3) Selección de modo

Puede haber varios criterios para que el predictor 910 seleccione uno de los modos explícito e implícito descritos anteriormente.

Dado que uno de una pluralidad de candidatos de predictor del vector de movimiento se selecciona en el modo explícito, un predictor del vector de movimiento más similar a un vector de movimiento de un bloque actual puede ser seleccionado. Sin embargo, dado que la información que indica uno de una pluralidad de candidatos predictores del vector de movimiento está codificada, puede producirse una sobrecarga mayor que en los modos implícitos. Por lo tanto, para una unidad de codificación que tiene un gran tamaño, un vector de movimiento puede codificarse en el modo explícito porque la probabilidad de aumentar un error cuando un vector de movimiento se predice erróneamente es mayor para una unidad de codificación que tiene un gran tamaño que una unidad de codificación tener un tamaño pequeño y el número de tiempos de ocurrencia de sobrecarga disminuye para cada imagen.

Por ejemplo, cuando una imagen dividida en partes iguales en la codificación de  $m$  unidades que tienen un tamaño de  $64 \times 64$  se codifica en el modo explícito, el número de veces de la ocurrencia de arriba es  $m$ . Sin embargo, cuando una imagen, que tiene el mismo tamaño, dividida equitativamente en unidades de codificación de  $4m$  con un tamaño de  $32 \times 32$ , se codifica en el modo explícito, el número de veces de sobrecarga es de  $4m$ .

Por consiguiente, el predictor 910 de acuerdo con una realización ejemplar puede seleccionar uno del modo explícito y el modo implícito basado en el tamaño de una unidad de codificación cuando se codifica un vector de movimiento de un bloque actual.

Dado que el tamaño de una unidad de codificación en el procedimiento de codificación de imagen y el procedimiento

de decodificación de imagen de acuerdo con realizaciones ejemplares descritas anteriormente con referencia a las figuras 1 a 8 se representa usando una profundidad, el predictor 910 determina en base a la profundidad de un bloque actual si un vector de movimiento del bloque actual está codificado en el modo explícito o en el modo implícito. Por ejemplo, cuando las unidades de codificación cuyas profundidades son 0 y 1 se predicen entre sí, los vectores de movimiento de las unidades de codificación se codifican en el modo explícito, y cuando las unidades de codificación cuyas profundidades son iguales o mayores a 2 se predicen entre sí, los vectores de movimiento de las unidades de codificación están codificadas en el modo implícito.

Según otra realización ejemplar, el predictor 910 puede seleccionar el modo explícito o el modo implícito para cada imagen o unidad de porción. Dado que las características de la imagen son diferentes para cada imagen o unidad de porción, el modo explícito o el modo implícito se pueden seleccionar para cada imagen o unidad de porción considerando estas características de imagen. Los vectores de movimiento de las unidades de codificación incluidas en una imagen o segmento actual se pueden codificar mediante predicción seleccionando un modo óptimo entre el modo explícito y el modo implícito en consideración del coste de R-D.

Por ejemplo, si los vectores de movimiento de las unidades de codificación incluidos en una imagen o porción pueden predecirse exactamente sin usar el modo explícito, los vectores de movimiento de todas las unidades de codificación incluidas en la imagen o porción pueden codificarse con predicción en el modo implícito.

De acuerdo con otra realización ejemplar, el predictor 910 puede seleccionar el modo explícito o el modo implícito en función de si un bloque actual se ha codificado en el modo de omisión. El modo de omisión es un modo de codificación en el que la información del indicador que indica que se ha codificado un bloque actual en el modo de omisión se codifica sin codificar un valor de píxel.

Además, el modo de omisión es un modo en el que el valor de píxel de un bloque actual no está codificado ya que un bloque de predicción generado al realizar la compensación de movimiento usando un predictor del vector de movimiento como un vector de movimiento del bloque actual es similar al bloque actual. Por lo tanto, como un predictor del vector de movimiento se genera de manera más similar a un vector de movimiento de un bloque actual, la probabilidad de codificar el bloque actual en el modo de omisión es mayor. En consecuencia, un bloque codificado en el modo de omisión puede codificarse en el modo explícito.

Con referencia de nuevo a la figura 9, cuando el predictor 910 selecciona uno del modo explícito y el modo implícito y determina un predictor del vector de movimiento de acuerdo con el modo seleccionado, el primer codificador 920 y el segundo codificador 930 codifican información sobre un modo de codificación y un vector de movimiento.

Específicamente, el primer codificador 920 codifica información sobre un predictor del vector de movimiento de un bloque actual. Con más detalle, cuando el predictor 910 determina que un vector de movimiento del bloque actual está codificado en modo explícito, el primer codificador 920 codifica información que indica que se ha generado un predictor del vector de movimiento en modo explícito e información que indica qué predictor del vector de movimiento candidato se ha utilizado como el predictor del vector de movimiento del bloque actual.

Por el contrario, cuando el predictor 910 selecciona que el vector de movimiento del bloque actual se codifica en el modo implícito, el primer codificador 920 codifica la información que indica que el predictor del vector de movimiento del bloque actual se ha generado en el modo implícito. En otras palabras, el primer codificador 920 codifica información que indica que el predictor del vector de movimiento del bloque actual se ha generado usando bloques o píxeles adyacentes al bloque actual. Si se usan dos o más modos implícitos, el primer codificador 920 puede codificar adicionalmente información que indica qué modo implícito se ha usado para generar el predictor del vector de movimiento del bloque actual.

El segundo codificador 930 codifica un vector de movimiento de un bloque actual basado en un predictor del vector de movimiento determinado por el predictor 910. Alternativamente, el segundo codificador 930 genera un vector de diferencia restando el predictor del vector de movimiento generado por el predictor 910 del vector de movimiento del bloque actual generado como resultado de la compensación de movimiento y codifica información sobre el vector de diferencia.

La figura 13 es un diagrama de bloques de un aparato 1300 para codificar una imagen, de acuerdo con una realización ejemplar. El aparato 1300 para decodificar el vector de movimiento puede incluirse en el aparato 200 de decodificación de imagen descrito anteriormente con referencia a la figura 2 o el decodificador 500 de imagen descrito anteriormente con referencia a la figura 5. Con referencia a la figura 13, un aparato 1300 de decodificación del vector de movimiento incluye un primer decodificador 1310, un segundo decodificador 1320, un predictor 1330 y un restaurador 1340 de vector de movimiento.

El primer decodificador 1310 decodifica la información sobre un predictor del vector de movimiento de un bloque actual, que se incluye en una corriente de bits. En detalle, el primer decodificador 1310 decodifica información que indica si el predictor del vector de movimiento del bloque actual se ha codificado en el modo explícito o en el modo implícito. Cuando el predictor del vector de movimiento del bloque actual se ha codificado en el modo explícito, el primer decodificador 1310 descodifica adicionalmente información que indica un predictor del vector de movimiento utilizado como predictor del vector de movimiento del bloque actual entre una pluralidad de predictores del vector de

movimiento. Cuando el predictor del vector de movimiento del bloque actual se ha codificado en el modo implícito, el primer decodificador 1310 puede decodificar adicionalmente información que indica cuál de una pluralidad de modos implícitos se ha utilizado para codificar el predictor del vector de movimiento del bloque actual.

5 El segundo decodificador 1320 decodifica un vector de diferencia entre un vector de movimiento y el vector de movimiento de predicción del bloque actual incluido en la corriente de bits.

El predictor 1330 genera un predictor del vector de movimiento del bloque actual basado en la información sobre el predictor del vector de movimiento del bloque actual, que ha sido decodificado por el primer decodificador 1310.

10 Cuando se decodifica la información sobre el predictor del vector de movimiento del bloque actual, que ha sido codificado en el modo explícito, el predictor 1330 genera un predictor del vector de movimiento entre los candidatos de predictores del vector de movimiento descritos anteriormente con referencia a las figuras 10A, 10B y 11A a 11C y utiliza el predictor del vector de movimiento generado como el predictor del vector de movimiento del bloque actual.

15 Cuando se decodifica la información sobre el predictor del vector de movimiento del bloque actual, que ha sido codificado en el modo implícito, el predictor 1330 genera el predictor del vector de movimiento del bloque actual usando bloques o píxeles incluidos en un área previamente codificada adyacente al bloque actual. Con más detalle, el predictor 1330 genera un valor medio de vectores de movimiento de bloques adyacentes al bloque actual como el predictor del vector de movimiento del bloque actual o genera el predictor del vector de movimiento del bloque actual buscando una imagen de referencia usando píxeles adyacentes al bloque actual.

20 El vector de movimiento restaurador 1340 restaura un vector de movimiento del bloque actual sumando el predictor del vector de movimiento generado por el predictor 1330 y el vector de diferencia decodificado por el segundo decodificador 1320. El vector de movimiento restaurado se usa para la compensación de movimiento del bloque actual.

25 La figura 14 es un diagrama de flujo que ilustra un procedimiento para decodificar un vector de movimiento, de acuerdo con una realización ejemplar. Con referencia a la figura 14, un aparato 900 de codificación del vector de movimiento según una realización ejemplar de selecciona uno de un modo explícito y un modo implícito como un modo de codificación de información sobre un predictor del vector de movimiento en la operación 1410.

30 El modo explícito es un modo de codificación de información que indica un candidato predictor del vector de movimiento entre al menos un candidato predictor del vector de movimiento como información sobre un predictor del vector de movimiento, y el modo implícito es un modo de codificación de información que indica que un predictor del vector de movimiento se ha generado en base a bloques o píxeles incluidos en un área previamente codificada adyacente a un bloque actual como información sobre el predictor del vector de movimiento. Se han dado descripciones detalladas de los mismos anteriormente con referencia a las figuras 10A, 10B, 11A a 11C y 12.

Se puede seleccionar un modo basado en el tamaño de un bloque actual, es decir, una profundidad del bloque actual, o se puede seleccionar en una unidad de una imagen o porción actual en el que se incluye el bloque actual. Alternativamente, se puede seleccionar un modo según si el bloque actual se ha codificado en un modo de omisión.

35 En la operación 1420, el aparato 900 de codificación del vector de movimiento determina un predictor del vector de movimiento de acuerdo con el modo seleccionado en la operación 1410. En detalle, el aparato 900 de codificación del vector de movimiento determina un predictor del vector de movimiento del bloque actual en base al modo explícito o modo implícito seleccionado en la operación 1410. Con más detalle, el aparato 900 de codificación del vector de movimiento determina un candidato predictor del vector de movimiento entre al menos un candidato predictor del vector de movimiento como el predictor del vector de movimiento del bloque actual en el modo explícito o determina el predictor del vector de movimiento del bloque actual basado en bloques o píxeles adyacentes al bloque actual en el modo implícito.

45 En la operación 1430, el aparato 900 de codificación del vector de movimiento codifica información sobre el predictor del vector de movimiento determinado en la operación 1420. En el caso del modo explícito, el aparato 900 de codificación del vector de movimiento codifica información que indica un candidato predictor del vector de movimiento entre al menos un candidato predictor del vector de movimiento como el predictor del vector de movimiento del bloque actual y la información que indica esa información sobre el predictor del vector de movimiento del bloque actual ha sido codificado en el modo explícito. En el caso del modo implícito, el aparato 900 de codificación del vector de movimiento codifica información que indica que el predictor del vector de movimiento del bloque actual se ha generado en base a bloques o píxeles incluidos en un área previamente codificada adyacente al bloque actual. En el caso de una pluralidad de modos implícitos, el aparato 900 de codificación del vector de movimiento puede codificar adicionalmente información que indica uno de la pluralidad de modos implícitos.

55 En la operación 1440, el aparato 900 de codificación del vector de movimiento codifica un vector de diferencia generado al restar el predictor del vector de movimiento determinado en la operación 1420 de un vector de movimiento del bloque actual.

La figura 15 es un diagrama de flujo que ilustra un procedimiento para decodificar un vector de movimiento, de

5 acuerdo con una realización ejemplar. Con referencia a la figura 15, un aparato 1300 de decodificación del vector de movimiento de acuerdo con una realización ejemplar decodifica información sobre un predictor del vector de movimiento de un bloque actual, que está incluido en una corriente de bits, en la operación 1510. En detalle, el aparato 1300 de decodificación del vector de movimiento decodifica información sobre un modo usado para codificar el predictor del vector de movimiento del bloque actual entre un modo explícito y un modo implícito.

10 En el caso del modo explícito, el aparato 1300 de decodificación del vector de movimiento decodifica información que indica que el predictor del vector de movimiento del bloque actual se ha codificado en el modo explícito e información sobre un candidato predictor del vector de movimiento entre al menos un vector de movimiento candidato predictor. En el caso del modo implícito, el aparato 1300 de decodificación del vector de movimiento decodifica información que indica que el predictor del vector de movimiento del bloque actual se ha generado en base a bloques o píxeles incluidos en un área previamente decodificada adyacente al bloque actual. En el caso de una pluralidad de modos implícitos, el aparato 1300 de decodificación del vector de movimiento puede decodificar adicionalmente información que indica uno de la pluralidad de modos implícitos.

15 En la operación 1520, el aparato 1300 de decodificación del vector de movimiento decodifica información sobre un vector de diferencia. El vector de diferencia es un vector de diferencia entre el predictor del vector de movimiento del bloque actual y un vector de movimiento del bloque actual.

20 En la operación 1530, el aparato 1300 de decodificación del vector de movimiento genera el predictor del vector de movimiento del bloque actual en base a la información sobre el predictor del vector de movimiento, que se ha decodificado en la operación 1510. En detalle, el aparato 1300 de decodificación del vector de movimiento genera el predictor del vector de movimiento del bloque actual de acuerdo con el modo explícito o el modo implícito. Con más detalle, el aparato 1300 de decodificación del vector de movimiento genera el predictor del vector de movimiento del bloque actual seleccionando un candidato predictor del vector de movimiento entre al menos un candidato predictor del vector de movimiento o usando bloques o píxeles incluidos en un área previamente decodificada adyacente al bloque actual.

25 En la operación 1540, el aparato 1300 de decodificación del vector de movimiento restaura el vector de movimiento del bloque actual sumando el vector de diferencia decodificado en la operación 1520 y el predictor del vector de movimiento generado en la operación 1530.

30 Además, un sistema de acuerdo con un ejemplo de realización se puede implementar utilizando un código legible por ordenador en un medio de grabación legible por ordenador. Por ejemplo, al menos uno de un aparato 100 para codificar una imagen, un aparato 200 para decodificar una imagen, un codificador 400 de imagen, un decodificador 500 de imagen, un aparato 900 de codificación del vector de movimiento y un aparato 1300 de decodificación del vector de movimiento, de acuerdo con las realizaciones ejemplares pueden incluir un bus acoplado a unidades de cada uno de los dispositivos mostrados en las figuras 1, 2, 4, 5, 9 y 13 y al menos un procesador conectado al bus. Además, una memoria acoplada a al menos un procesador para realizar comandos como se describió anteriormente puede incluirse y conectarse al bus para almacenar los comandos y los mensajes recibidos o mensajes generados.

35 Un medio de grabación legible por ordenador no transitorio es cualquier dispositivo de almacenamiento de datos que puede almacenar datos que luego pueden ser leídos por un sistema informático. Los ejemplos de medios de grabación legibles por ordenador no transitorios incluyen una memoria de solo lectura (ROM), una memoria de acceso aleatorio (RAM), discos compactos (CD-ROM), cintas magnéticas, disquetes y dispositivos de almacenamiento de datos ópticos. El medio de grabación legible por ordenador no transitorio también se puede distribuir a través de sistemas informáticos acoplados a la red, de modo que el código legible por ordenador se almacena y ejecuta de manera distribuida.

**REIVINDICACIONES**

1. Un procedimiento de decodificación de una imagen, comprendiendo el procedimiento:

obtener una unidad de codificación actual cuadrada dividida jerárquicamente de una unidad de codificación máxima de acuerdo con una profundidad actual;

5 obtener una o más unidades de predicción de la unidad de codificación actual dividiendo por igual un ancho o una altura de la unidad de codificación actual por dos, o dividiendo por igual tanto la altura como el ancho de la unidad de codificación actual por dos, o determinando la unidad de codificación actual como la unidad de predicción;

10 obtener una información de modo de predicción de una unidad de predicción actual en la unidad de codificación actual de la corriente de bits;

cuando un modo de predicción de la unidad de predicción actual es el modo de inter-predicción y un predictor del vector de movimiento de la unidad de predicción actual se determina de acuerdo con un modo explícito en lugar de un modo implícito, determinar una pluralidad de candidatos de predictores del vector de movimiento entre los vectores de movimiento de unidades de predicción vecinas previamente decodificadas adyacentes a la unidad de predicción actual, en el que el modo explícito es un modo que determina la pluralidad de candidatos de predictores del vector de movimiento y determina el predictor del vector de movimiento de la unidad de predicción actual en base a una información del predictor de vector de movimiento que indica el predictor del vector de movimiento de la unidad de predicción actual entre la pluralidad de candidatos predictores del vector de movimiento, y el modo implícito es un modo que determina el predictor del vector de movimiento de la unidad de predicción actual sin usar la información del predictor del vector de movimiento; y

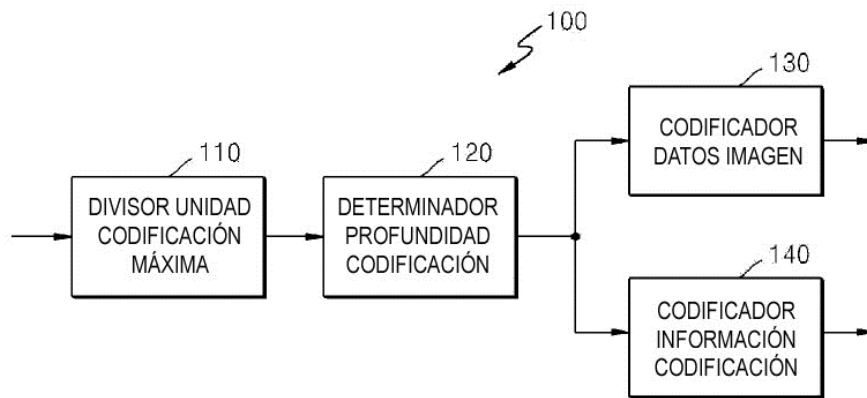
20 determinar un predictor del vector de movimiento de la unidad de predicción actual entre los candidatos de predictores del vector de movimiento en base a la información del predictor del vector de movimiento obtenida de la corriente de bits,

25 en el que la unidad de predicción vecina previamente decodificada comprende una primera unidad de predicción que se encuentra en un lado inferior izquierdo de la unidad de predicción actual, es decir, directamente debajo de una unidad de predicción más baja entre las unidades de predicción izquierdas adyacentes a un lado izquierdo de la unidad de predicción actual.

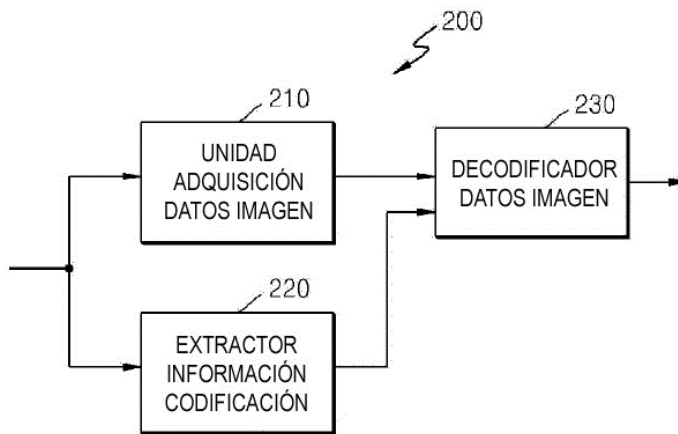
30 2. El procedimiento de la reivindicación 1, en el que los candidatos predictores del vector de movimiento comprenden además un vector de movimiento de una unidad de predicción ubicada en una posición en una imagen de referencia ubicada junto con una posición de la unidad de predicción actual en una imagen actual.

3. El procedimiento de la reivindicación 2, en el que el vector de movimiento de la unidad de predicción de ubicación conjunta se escala en base a una distancia temporal entre la imagen de referencia y la imagen actual.

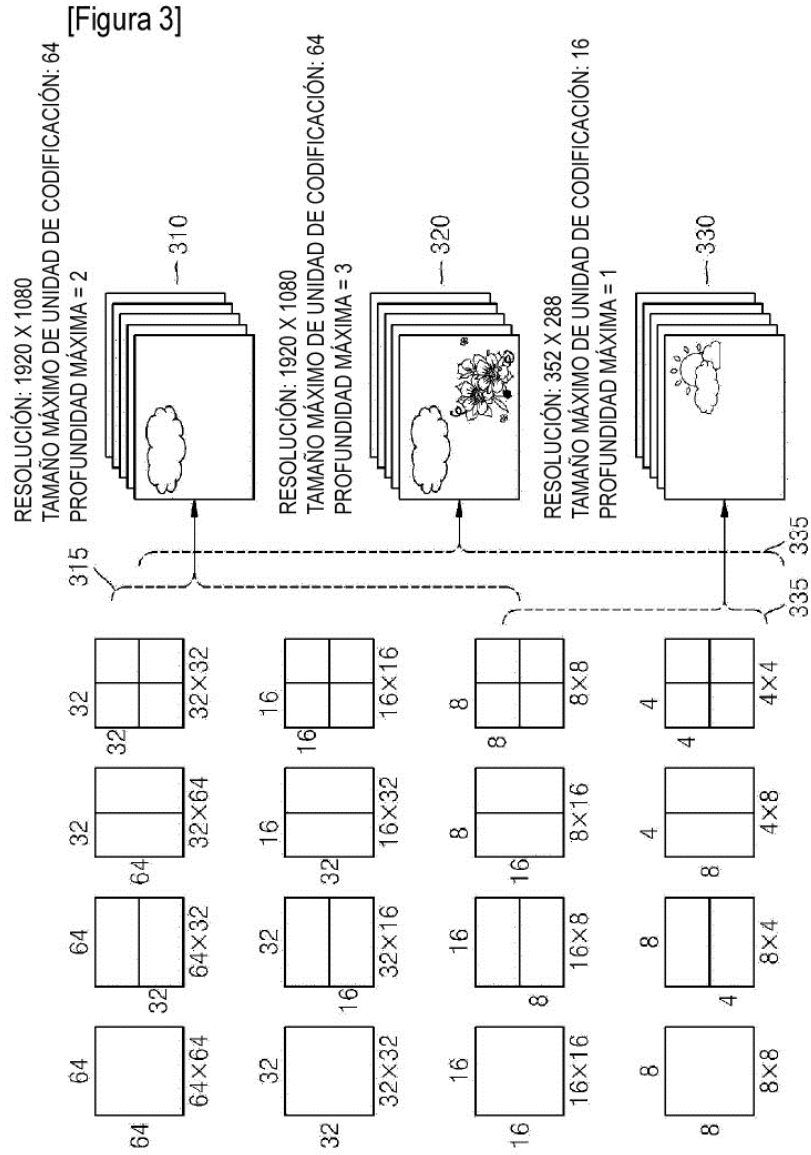
[Figura 1]



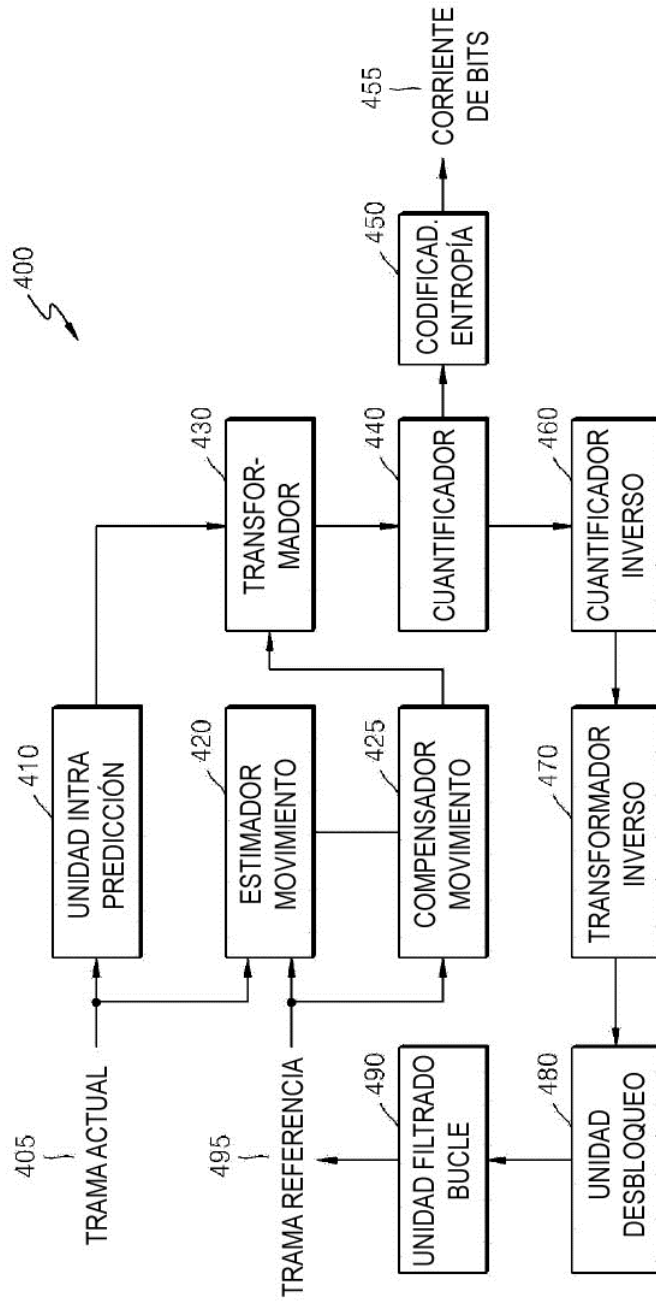
[Figura 2]



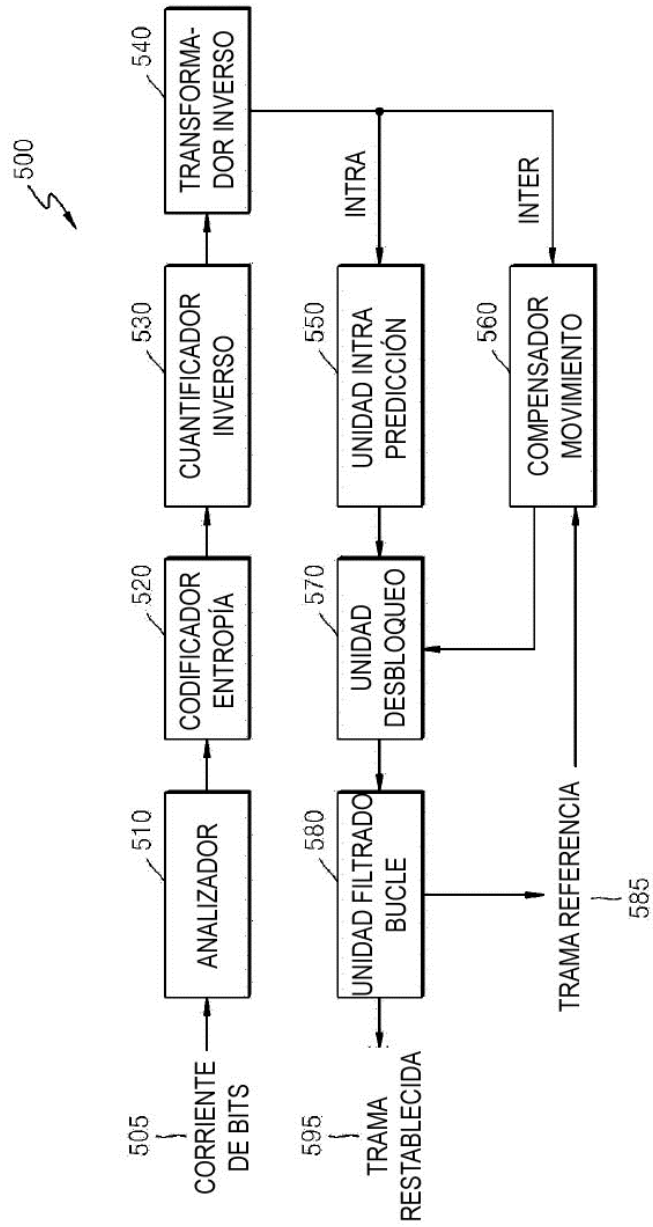




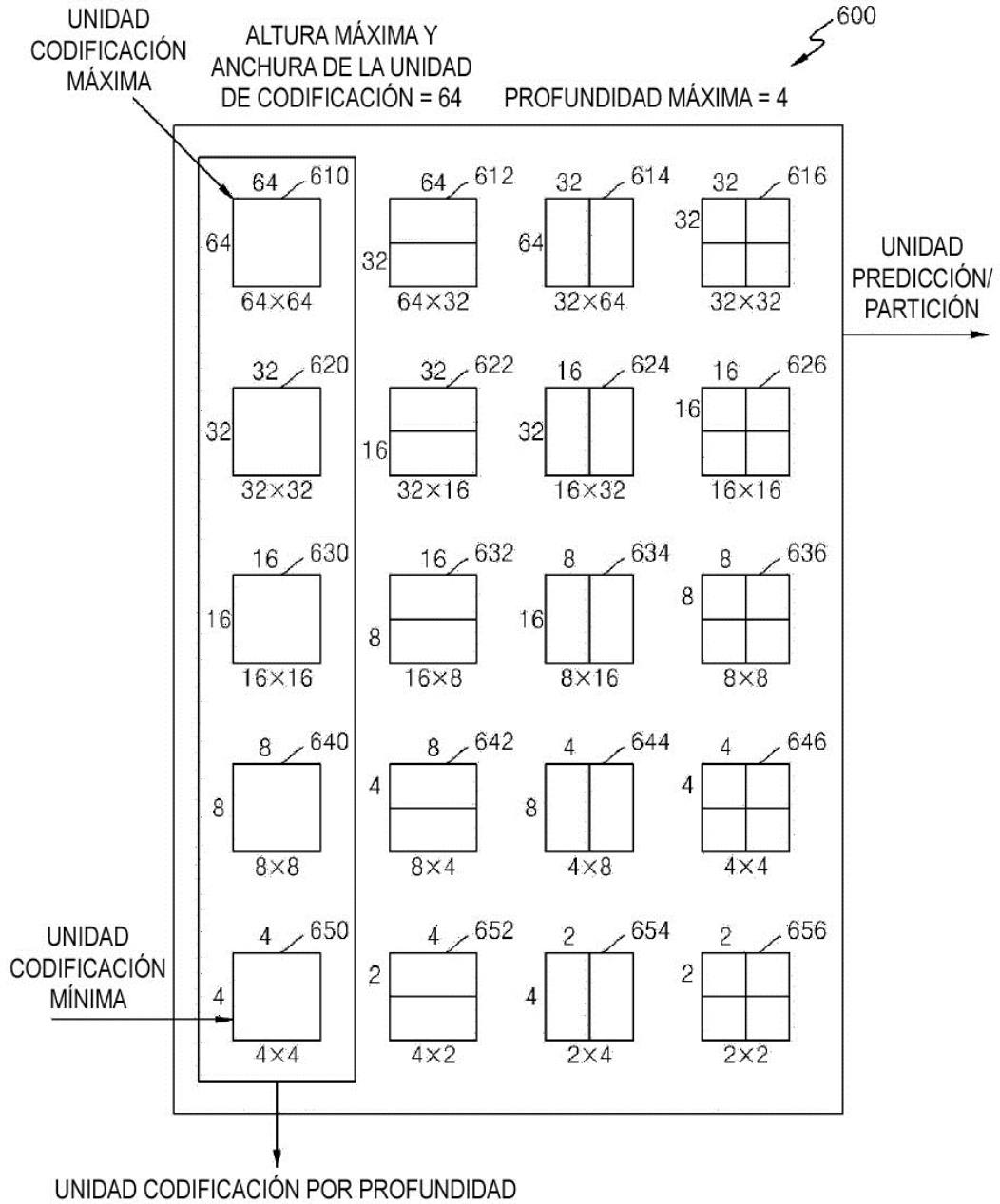
[Figura 4]



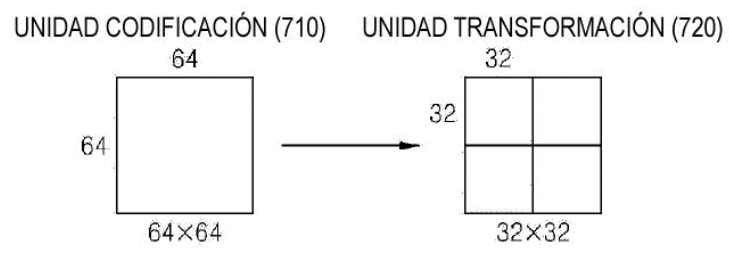
[Figura 5]



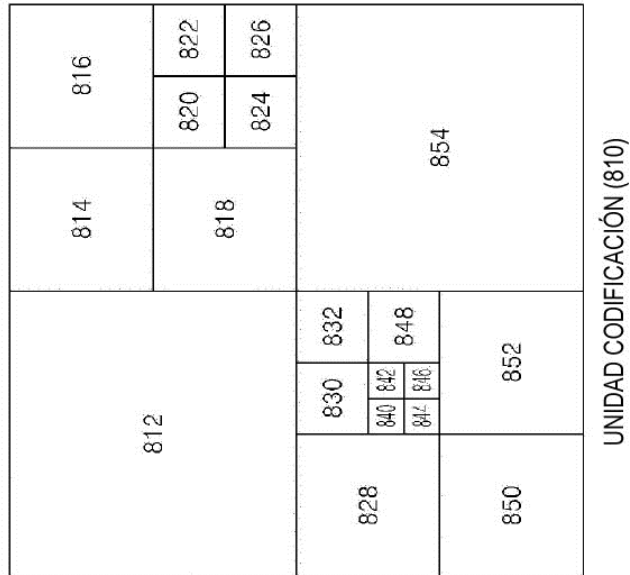
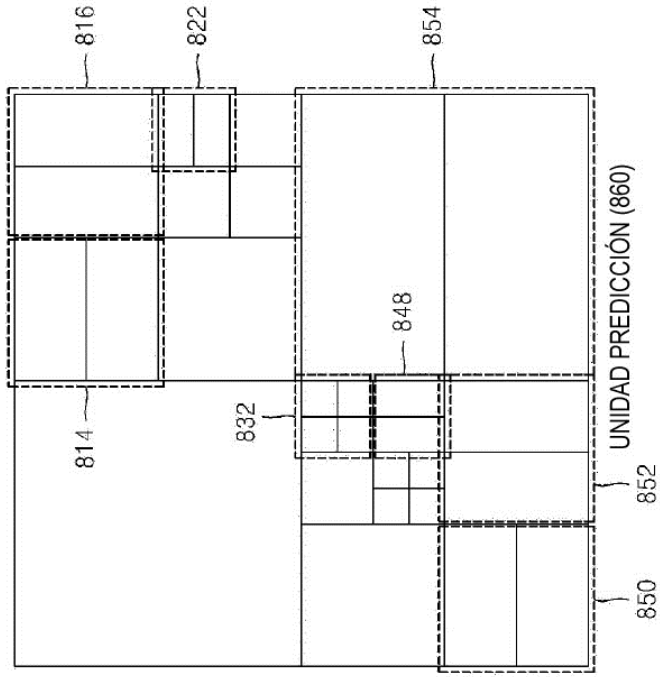
[Figura 6]



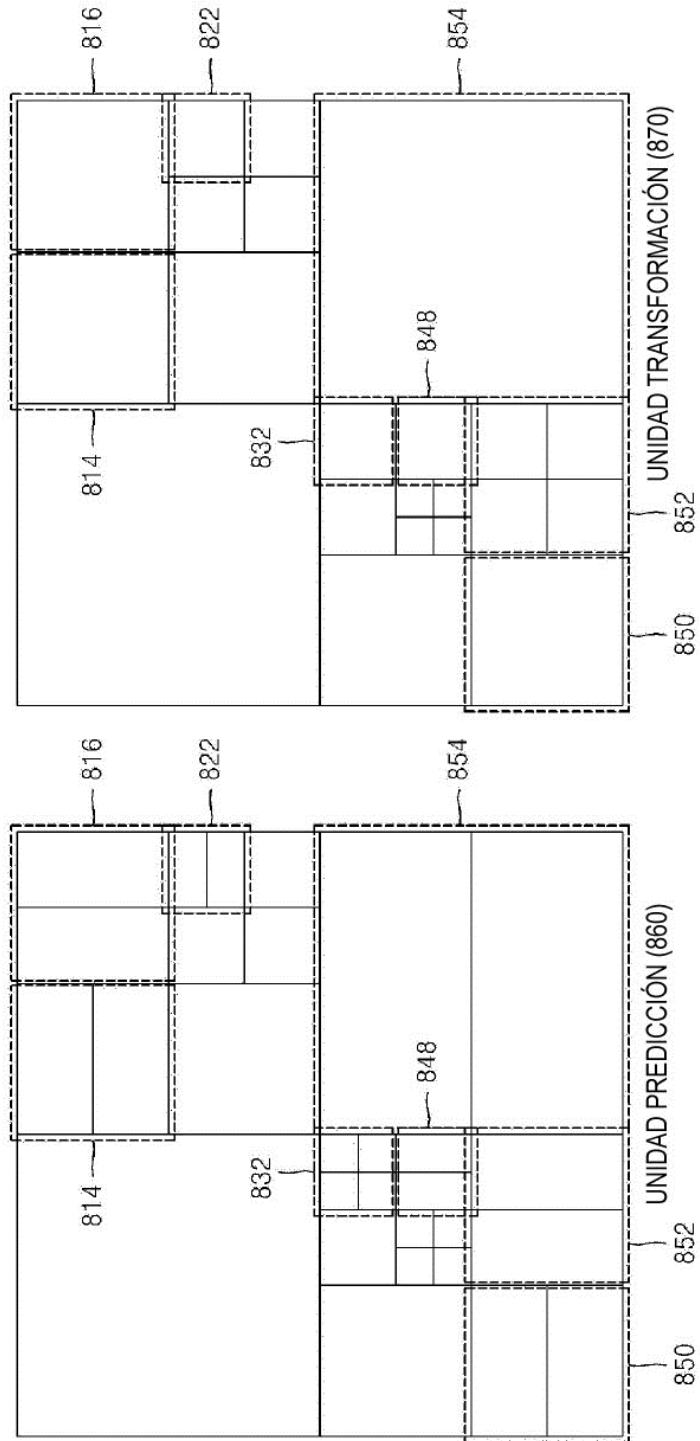
[Figura 7]



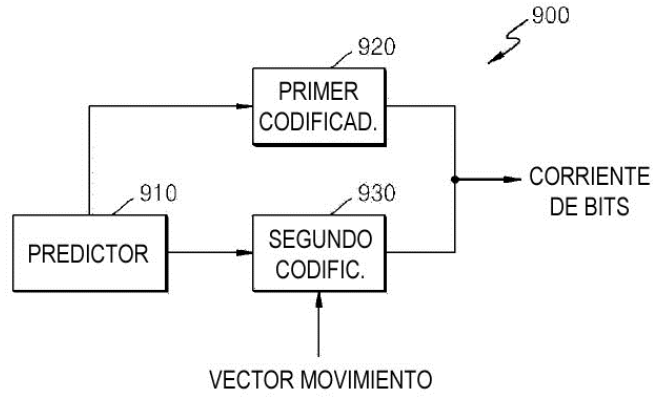
[Figura 8a]



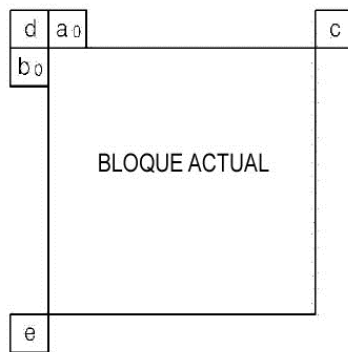
[Figura 8b]



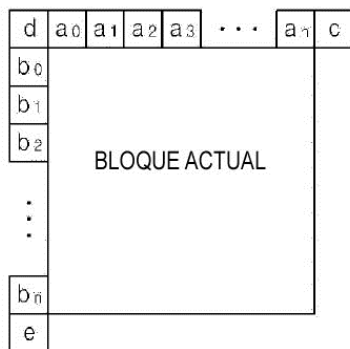
[Figura 9]



[Figura 10a]

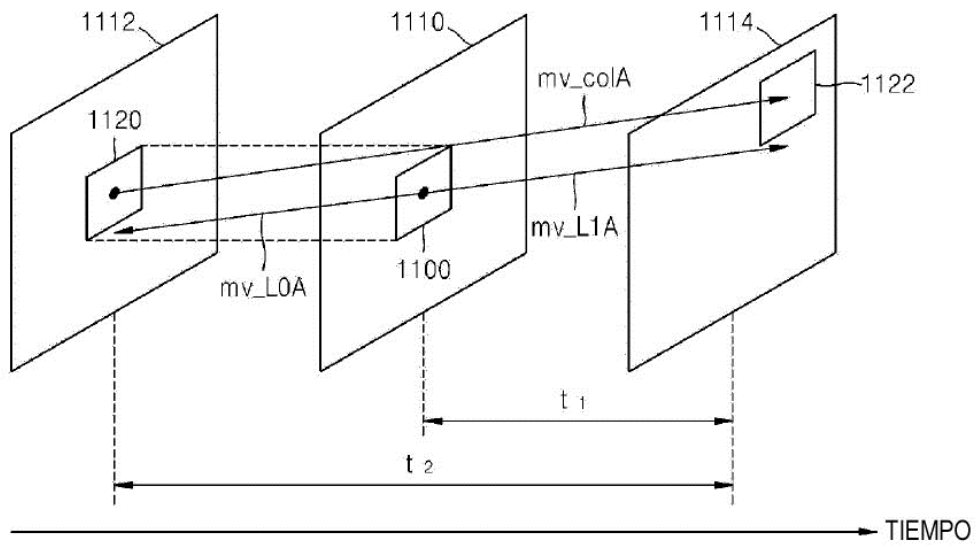


[Figura 10b]

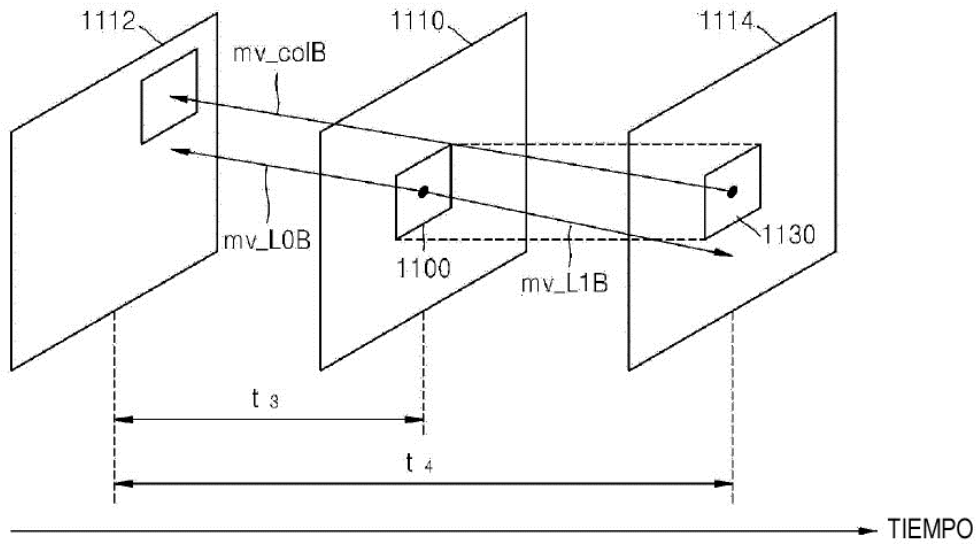




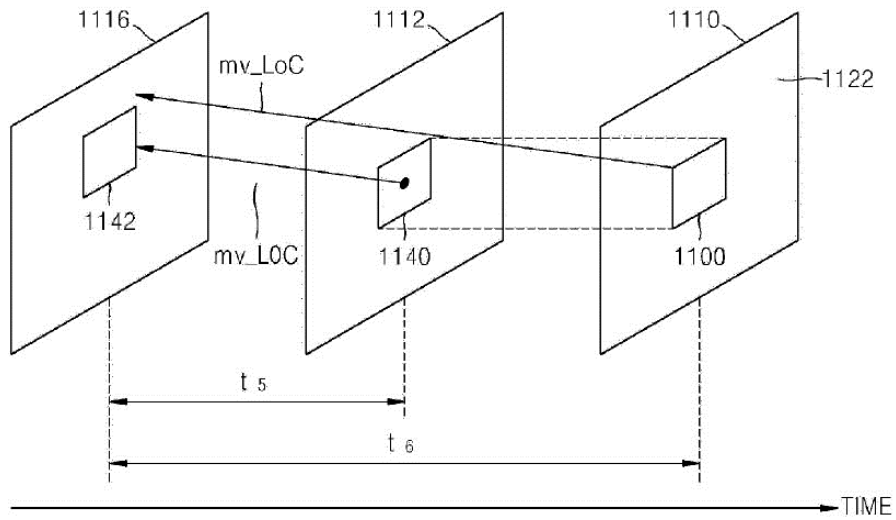
[Figura 11a]



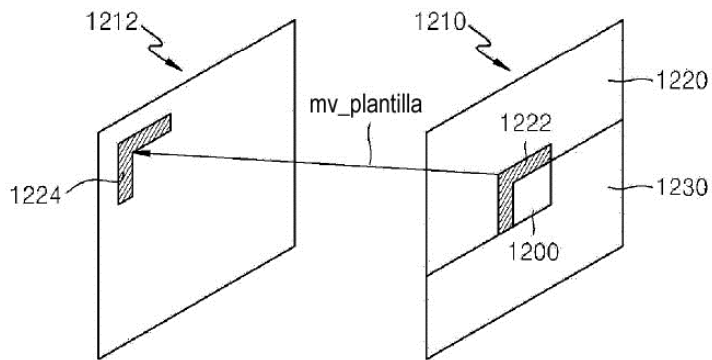
[Figura 11b]



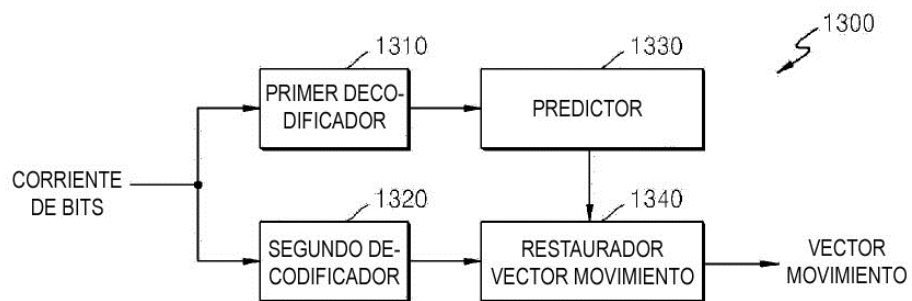
[Figura 11c]



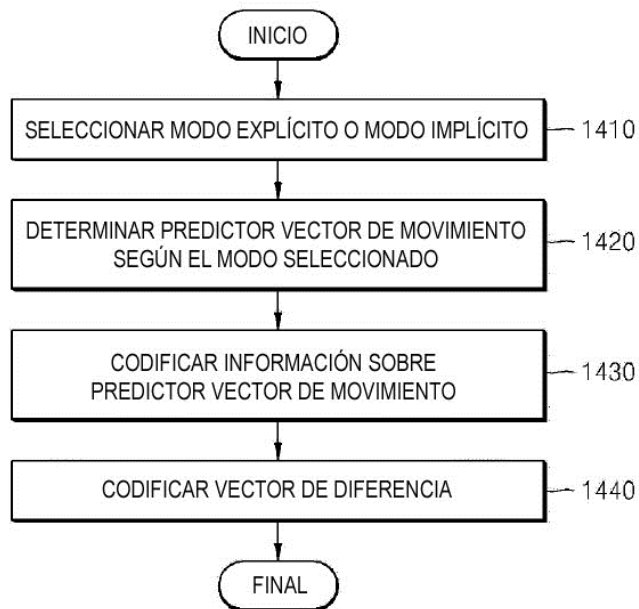
[Figura 12]



[Figura 13]



[Figura 14]



[Figura 15]

