

19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 773 691**

51 Int. Cl.:

H04N 19/523 (2014.01)
H04N 19/109 (2014.01)
H04N 19/136 (2014.01)
H04N 19/176 (2014.01)
H04N 19/61 (2014.01)
H04N 19/146 (2014.01)
H04N 19/19 (2014.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

96 Fecha de presentación y número de la solicitud europea: **14.09.2012** **E 18202093 (3)**

97 Fecha y número de publicación de la concesión europea: **05.02.2020** **EP 3454561**

54 Título: **Procedimiento y dispositivo de codificación de una unidad de predicción (PU) según su tamaño y dispositivo de decodificación correspondiente**

30 Prioridad:

14.09.2011 US 201161534499 P

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

14.07.2020

73 Titular/es:

SAMSUNG ELECTRONICS CO., LTD. (100.0%)
129, Samsung-ro, Yeongtong-gu
Suwon-si, Gyeonggi-do 16677, KR

72 Inventor/es:

ALSHINA, ELENA;
CHEN, JIANLE y
HAN, WOO-JIN

74 Agente/Representante:

GONZÁLEZ PECES, Gustavo Adolfo

ES 2 773 691 T3

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

DESCRIPCIÓN

Procedimiento y dispositivo de codificación de una unidad de predicción (PU) según su tamaño y dispositivo de decodificación correspondiente

Campo de la invención

- 5 La presente invención se refiere a codificación y decodificación de vídeo, y más particularmente, a un procedimiento y a un aparato para la codificación y decodificación de un vídeo para reducir el acceso de memoria para la lectura de datos de un cuadro de referencia durante la compensación de movimiento.

Descripción de la técnica relacionada

- 10 Con el fin de codificar un vídeo, una imagen se divide en unidades de datos predeterminado tal como macrobloques. Cada uno de los macrobloques está codificado por predicción mediante el uso de inter predicción o intra predicción.

- 15 La inter predicción, que es un procedimiento para comprimir una imagen eliminando la redundancia temporal entre imágenes, utiliza la estimación de movimiento y la compensación de movimiento. La estimación de movimiento predice los bloques de una imagen actual utilizando al menos una imagen de referencia. La estimación de movimiento y la compensación de movimiento son procesos de búsqueda de un bloque de referencia que es más similar a un bloque actual en un rango de búsqueda predeterminado al usar una función de evaluación predeterminada y leer los datos del bloque de referencia que es más similar al bloque actual. Para realizar con más precisión la estimación de movimiento y la compensación de movimiento, se generan subpíxeles entre píxeles enteros realizando una interpolación en el rango de búsqueda predeterminado de la imagen de referencia, y la estimación de movimiento y la compensación de movimiento se realizan en función de la imagen de referencia que tiene precisión de subpíxel. Para realizar la estimación de movimiento y la compensación de movimiento en la precisión de subpíxel se necesita un proceso de interpolación de los datos de la imagen de referencia.

- 20 En una técnica convencional, la interpolación se lleva a cabo mediante el uso de un filtro de respuesta de impulso finito (FIR) que tiene un número de derivación predeterminado. Un filtro FIR de N derivaciones interpola un subpíxel utilizando N píxeles enteros vecinos. Para interpolar un subpíxel en un proceso de interpolación mediante el uso de un filtro FIR de N derivaciones, los N píxeles de referencia deben leerse desde una memoria. Si tanto un componente horizontal como un componente vertical de un vector de movimiento tienen una precisión de subpíxel, ya que un proceso de interpolación bidimensional (2D) debe realizarse en direcciones horizontal y vertical, se aumenta el número de veces que se realiza el acceso a la memoria. El número de veces que el acceso a la memoria realizado para el movimiento bidireccional es dos veces mayor que el número de veces que el acceso a la memoria se realiza para la estimación de movimiento unidireccional porque se utilizan dos imágenes de referencia.

- 25 El documento US 2007/217513 A1 describe un procedimiento de codificación de imágenes para codificar una imagen cuyos macrobloques están divididos en partes de varios tamaños. El procedimiento de codificación de imágenes se proporciona para reducir adecuadamente la cantidad de codificación de un vector de movimiento, sin disminuir la eficiencia de codificación. Con el fin de reducir una cantidad de procesamiento y un aumento de una cantidad de codificación, que resulta de la generación de un gran número de vectores de movimiento debido a la predicción bidireccional seleccionada para una parte de tamaño pequeño, una unidad de estimación de movimiento selecciona el procedimiento de predicción óptimo de candidatos del procedimiento de predicción decididos por una unidad de determinación de candidatos del procedimiento de predicción. De este modo, cuando se realiza la compensación de movimiento para una parte de gran tamaño, se puede permitir la predicción hacia adelante, hacia atrás y direccional y bidireccional. Por otro lado, cuando se realiza la compensación de movimiento para una parte de tamaño pequeño, se prohíbe la predicción bidireccional, pero se permite la predicción hacia adelante o hacia atrás.

- 45 El documento US 2003/112864 A1 desvela la reducción del ancho de banda de la memoria realizando una interpolación "a demanda" tal que solo se interpolan de antemano subpíxeles con una resolución de media unidad para el cuadro completo y se almacenan en la memoria. Sin embargo, los cuartos de píxeles se siguen calculando en medios píxeles y píxeles completos.

- 50 El documento de RUSANOVKSY D ET AL: "Adaptive interpolation with directional filters", 33. VCEG MEETING; 82. MPEG MEETING; 20-10-2007 - 20-10-2007; SHENZHEN;(VÍDEO CODING EXPERTS GROUP OF ITU-T SG.16), n.º VCEG- AG21, del 20 de octubre de 2007 desvela la reducción de la complejidad de la interpolación obteniendo las ubicaciones de cuarto de píxel utilizando un filtro de interpolación de una dimensión basado en píxeles enteros.

Sumario de la invención

- La presente invención proporciona un procedimiento y un aparato para la codificación de un vídeo y un aparato para la decodificación de un vídeo que pueden reducir el acceso de memoria para la lectura de datos de referencia durante la estimación de movimiento y la compensación de movimiento.
- 55 De acuerdo con un aspecto de la presente invención, se proporciona un procedimiento de codificación de un vídeo,

un aparato para codificar un vídeo y un aparato para decodificar un vídeo como se indica en las reivindicaciones adjuntas.

5 El acceso a la memoria para leer datos de un marco de referencia puede limitarse a un ancho de banda predeterminado determinando si se usa un modo de estimación de movimiento unidireccional y un modo de estimación bidireccional en base a un tamaño de una unidad de predicción y realizando una estimación de movimiento y una compensación de movimiento utilizando el modo o modos de estimación de movimiento determinados.

Breve descripción de los dibujos

10 Las anteriores y otras características y ventajas de la presente invención serán más evidentes a partir de la descripción detallada de las realizaciones de ejemplo con referencia a los dibujos adjuntos, en los que:

- La figura 1 es un diagrama de bloques de un aparato para codificar un vídeo, de acuerdo con una realización de la presente invención;
- La figura 2 es un diagrama de bloques de un aparato para decodificar un vídeo, de acuerdo con una realización de la presente invención;
- 15 La figura 3 es un diagrama para describir un concepto de unidades de codificación de acuerdo con una realización de la presente invención;
- La figura 4 es un diagrama de bloques de un codificador de imágenes basado en unidades de codificación de acuerdo con una realización de la presente invención;
- 20 La figura 5 es un diagrama de bloques de un decodificador de imágenes basado en unidades de codificación de acuerdo con una realización de la presente invención;
- La figura 6 es un diagrama que ilustra unidades de codificación más profundas según las profundidades, y particiones de acuerdo con una realización de la presente invención;
- La figura 7 es un diagrama para describir una relación entre una unidad de codificación y unidades de transformación, de acuerdo con una realización de la presente invención;
- 25 La figura 8 es un diagrama para describir información de codificación de unidades de codificación correspondientes a una profundidad codificada, de acuerdo con una realización de la presente invención;
- La figura 9 es un diagrama de unidades de codificación más profundas según las profundidades, de acuerdo con una realización de la presente invención;
- 30 Las figuras 10 a 12 son diagramas para describir una relación entre unidades de codificación, unidades de predicción y unidades de transformación, de acuerdo con una realización de la presente invención;
- La figura 13 es un diagrama para describir una relación entre una unidad de codificación, una unidad de predicción o una partición y una unidad de transformación, de acuerdo con la información del modo de codificación de la Tabla 1;
- 35 La figura 14 es un diagrama que ilustra un aparato de estimación de movimiento de acuerdo con una realización de la presente invención;
- La figura 15 es un diagrama que ilustra las posiciones de un subpíxel y un píxel entero seleccionables durante la compensación de movimiento con una precisión de 1/4 de píxel, de acuerdo con una realización de la presente invención;
- 40 La figura 16 es un diagrama para explicar un proceso de interpolación de un marco de referencia, de acuerdo con una realización de la presente invención;
- Las figuras 17A a 17C son diagramas que ilustran datos de un marco de referencia necesario para un proceso de compensación de movimiento, de acuerdo con una realización de la presente invención;
- 45 Las figuras 18A a 18G son tablas que ilustran un número de píxeles de un cuadro de referencia necesario para la compensación de movimiento por píxel de una unidad de predicción actual de acuerdo con el tamaño de una unidad de predicción, un modo de predicción y si los componentes horizontales y verticales de un vector de movimiento de la unidad de predicción tienen valores enteros;
- 50 Las figuras 19 y 20 son tablas que ilustran en orden descendente comenzando desde un valor alto un número de píxeles de un cuadro de referencia necesario para la compensación de movimiento por píxel de una unidad de predicción actual de acuerdo con el tamaño de una unidad de predicción de un componente de luminancia, un tamaño de unidad de predicción de un componente de croma, un modo de predicción y si las componentes horizontales y verticales de un vector de movimiento de la unidad de predicción tienen valores enteros;
- 55 Las figuras 21 a 23 son diagramas de referencia para explicar un proceso de interpolación para reducir un proceso de lectura de datos de un marco de referencia, de acuerdo con una realización de la presente invención;
- Las figuras 24 y 25 son tablas que ilustran en orden descendente comenzando desde un valor alto un número de píxeles de un cuadro de referencia necesario para la compensación de movimiento por píxel de una unidad de predicción actual de acuerdo con el tamaño de una unidad de predicción de un componente de luminancia, un tamaño de unidad de predicción de un componente de croma, un modo de estimación de movimiento y si las componentes horizontales y verticales de un vector de movimiento de la unidad de predicción tienen valores enteros;
- 60 La figura 26 es un diagrama de flujo que ilustra un procedimiento de codificación de vídeo de acuerdo con una realización de la presente invención;
- La figura 27 es un diagrama de bloques que ilustra un aparato de compensación de movimiento de acuerdo con una realización de la presente invención;

La figura 28 es un diagrama de flujo que ilustra un procedimiento de decodificación de vídeo de acuerdo con una realización de la presente invención;

La figura 29A ilustra una estructura física de un disco que almacena un programa, de acuerdo con una realización de la presente invención;

5 La figura 29B ilustra una unidad de disco que graba y lee un programa utilizando un disco, de acuerdo con una realización de la presente invención;

La figura 30 ilustra una estructura completa de un sistema de suministro de contenido que proporciona un servicio de distribución de contenido, de acuerdo con una realización de la presente invención;

10 Las figuras 31 y 32 ilustran estructuras externas e internas de un teléfono móvil a las que se aplican un procedimiento de codificación de vídeo y un procedimiento de decodificación de vídeo, de acuerdo con una realización de la presente invención;

La figura 33 ilustra un sistema de transmisión digital que emplea un sistema de comunicación, vídeo de acuerdo con una realización de la presente invención; y

15 La figura 34 ilustra una estructura de red de un sistema informático en la nube que usa un aparato de codificación de vídeo y un aparato de decodificación de vídeo, de acuerdo con una realización de la presente invención.

Descripción detallada de la invención

La invención se describirá ahora más completamente con referencia a los dibujos adjuntos, en los que se muestran realizaciones de ejemplo de la invención.

20 La figura 1 es un diagrama de bloques de un aparato 100 de codificación de vídeo de acuerdo con una realización de la presente invención.

El aparato 100 de codificación de vídeo incluye un divisor 110 de unidad de codificación máxima, un determinador 120 de unidad de codificación, y una unidad 130 de salida.

25 El divisor 110 de unidad de codificación máxima puede dividir una imagen actual basada en una unidad máxima de codificación para la imagen actual de una imagen. Si la imagen actual es más grande que la unidad de codificación máxima, los datos de imagen de la imagen actual pueden dividirse en al menos una unidad de codificación máxima. La unidad de codificación máxima, de acuerdo con una realización de la presente invención, puede ser una unidad de datos que tenga un tamaño de 32x32, 64x64, 128x128, 256x256, etc., en la que una forma de la unidad de datos es un cuadrado que tiene una anchura y longitud en cuadrados de 2. Los datos de imagen pueden enviarse al
30 determinador 120 de la unidad de codificación según la al menos una unidad de codificación máxima.

Una unidad de codificación, de acuerdo con una realización de la presente invención, puede estar caracterizada por un tamaño máximo y una profundidad. La profundidad indica un número de veces que la unidad de codificación se divide espacialmente de la unidad de codificación máxima, y a medida que la profundidad se profundiza, las
35 unidades de codificación más profundas según las profundidades pueden dividirse de la unidad de codificación máxima a una unidad de codificación mínima. Una profundidad de la unidad de codificación máxima es una profundidad superior y una profundidad de la unidad de codificación mínima es una profundidad inferior. Dado que el tamaño de una unidad de codificación correspondiente a cada profundidad disminuye a medida que la profundidad de la unidad de codificación máxima se profundiza, una unidad de codificación correspondiente a una profundidad superior puede incluir una pluralidad de unidades de codificación correspondientes a profundidades más bajas.

40 Como se ha descrito anteriormente, los datos de imagen de la imagen actual se dividen en las unidades de codificación máxima de acuerdo con un tamaño máximo de la unidad de codificación, y cada una de las unidades de codificación máxima puede incluir unidades de codificación más profundas que se dividen según las profundidades. Como la unidad de codificación máxima, de acuerdo con una realización de la presente invención, se divide de acuerdo con las profundidades, los datos de imagen de un dominio espacial incluidos en la unidad de codificación
45 máxima se pueden clasificar jerárquicamente según las profundidades.

Una profundidad máxima y un tamaño máximo de una unidad de codificación, que limitan el número total de veces que una altura y una anchura de la unidad de codificación máxima se dividen jerárquicamente pueden estar predeterminados.

50 El determinador 120 de la unidad de codificación codifica al menos una región dividida obtenida dividiendo una región de la unidad máxima de codificación según las profundidades, y determina una profundidad para emitir datos de imagen finalmente codificados según la al menos una región dividida. En otras palabras, el determinador 120 de la unidad de codificación determina una profundidad codificada codificando los datos de imagen en las unidades de codificación más profundas de acuerdo con las profundidades según la unidad de codificación máxima de la imagen actual, y seleccionando una profundidad que tiene el menor error de codificación. Por lo tanto, los datos de la
55 imagen codificada de la unidad de codificación correspondiente a la profundidad codificada determinada se emiten finalmente. Además, las unidades de codificación correspondientes a la profundidad codificada pueden considerarse unidades de codificación codificadas. La profundidad codificada determinada y los datos de imagen codificados según la profundidad codificada determinada se envían a la unidad 130 de salida.

Los datos de imagen en la unidad de codificación máxima se codifican basándose en las unidades de codificación más profundas correspondientes a al menos una profundidad igual o menor que la profundidad máxima, y los resultados de codificación de los datos de imagen se comparan en función de cada una de las unidades de codificación más profundas. Se puede seleccionar una profundidad que tenga el menor error de codificación después de comparar los errores de codificación de las unidades de codificación más profundas. Se puede seleccionar al menos una profundidad codificada para cada unidad de codificación máxima.

El tamaño de la unidad de codificación máxima se divide, ya que una unidad de codificación se divide jerárquicamente de acuerdo con las profundidades, y a medida que aumenta el número de unidades de codificación. Además, incluso si las unidades de codificación corresponden a la misma profundidad en una unidad de codificación máxima, se determina la división de cada una de las unidades de codificación correspondientes a la misma profundidad a una profundidad más baja midiendo un error de codificación de los datos de imagen de cada unidad de codificación, por separado. En consecuencia, incluso cuando los datos de imagen se incluyen en una unidad de codificación máxima, los datos de imagen se dividen en regiones según las profundidades y los errores de codificación pueden diferir según las regiones en la unidad de codificación máxima, y por lo tanto las profundidades codificadas pueden diferir según regiones en los datos de la imagen. Así, se pueden determinar una o más profundidades codificadas en una unidad de codificación máxima, y los datos de imagen de la unidad de codificación máxima se pueden dividir de acuerdo con unidades de codificación de al menos una profundidad codificada.

Por consiguiente, el determinador 120 de la unidad de codificación puede determinar las unidades de codificación que tienen una estructura de árbol incluida en la unidad de codificación máxima. Las 'unidades de codificación que tienen una estructura de árbol', de acuerdo con una realización de la presente invención, incluyen unidades de codificación correspondientes a una profundidad determinada como la profundidad codificada, entre todas las unidades de codificación más profundas incluidas en la unidad de codificación máxima. Una unidad de codificación de una profundidad codificada se puede determinar jerárquicamente según las profundidades en la misma región de la unidad de codificación máxima, y se puede determinar independientemente en diferentes regiones. De manera similar, una profundidad codificada en una región actual puede determinarse independientemente a partir de una profundidad codificada en otra región.

Una profundidad máxima, de acuerdo con una realización de la presente invención, es un índice relacionado con el número de veces que la división se realiza a partir de una unidad máxima de codificación para una unidad de codificación mínima. Una primera profundidad máxima, de acuerdo con una realización de la presente invención, puede indicar el número total de veces que se realiza la división desde la unidad de codificación máxima a la unidad de codificación mínima. Una segunda profundidad máxima, de acuerdo con una realización de la presente invención, puede indicar el número total de niveles de profundidad desde la unidad de codificación máxima hasta la unidad de codificación mínima. Por ejemplo, cuando una profundidad de la unidad de codificación máxima es 0, una profundidad de una unidad de codificación, en la que la unidad de codificación máxima se divide una vez, se puede establecer en 1, y una profundidad de una unidad de codificación, en la que la unidad de codificación máxima se divide dos veces, se puede establecer en 2. Aquí, si la unidad de codificación mínima es una unidad de codificación en la que la unidad de codificación máxima se divide cuatro veces, existen 5 niveles de profundidad de profundidades 0, 1, 2, 3 y 4, por lo que la primera profundidad máxima se puede establecer en 4 y la segunda profundidad máxima se puede establecer en 5.

La codificación de la predicción y la transformación se pueden realizar de acuerdo con la unidad de codificación máxima. La codificación de predicción y la transformación también se realizan en función de las unidades de codificación más profundas según una profundidad igual o inferior a la profundidad máxima, según la unidad de codificación máxima. La transformación se puede realizar de acuerdo con un procedimiento de transformación ortogonal o transformación de enteros.

Puesto que el número de unidades de codificación más profunda aumenta siempre que la unidad de codificación máxima se divide de acuerdo con profundidades, la codificación incluyendo la codificación de predicción y la transformación se realiza en todas las unidades de codificación más profunda generadas cuando la profundidad se profundiza. Por conveniencia de la descripción, la codificación de predicción y la transformación se describirán ahora basándose en una unidad de codificación de una profundidad actual, en una unidad de codificación máxima.

El aparato 100 de codificación de imágenes puede seleccionar diversamente un tamaño o forma de una unidad de datos para codificar los datos de imagen. Para codificar los datos de la imagen, se realizan operaciones, como la codificación de predicción, la transformación y la codificación de entropía, y en este momento, se puede usar la misma unidad de datos para todas las operaciones o se pueden usar diferentes unidades de datos para cada operación.

Por ejemplo, el aparato 100 de codificación de vídeo puede seleccionar no solo una unidad de codificación para codificar los datos de imagen, sino también una unidad de datos diferente de la unidad de codificación para realizar la codificación de predicción en los datos de imagen en la unidad de codificación.

Para realizar codificación de predicción en la unidad máxima de codificación, la codificación de predicción puede realizarse basándose en una unidad de codificación correspondiente a una profundidad codificada, es decir, sobre la

base de una unidad de codificación que ya no se divide en unidades de codificación correspondientes a una menor profundidad. A continuación, la unidad de codificación ya no se divide y se convierte en una unidad base para codificación de predicción se denominará ahora "unidad de predicción". Una partición obtenida dividiendo la unidad de predicción puede incluir una unidad de predicción o una unidad de datos obtenida dividiendo al menos una entre una altura y una anchura de la unidad de predicción.

Por ejemplo, cuando una unidad de codificación de $2N \times 2N$ (donde N es un entero positivo) ya no se divide y se convierte en una unidad de predicción de $2N \times 2N$, y el tamaño de una partición puede ser $2N \times 2N$, $2N \times N$, $N \times 2N$, o $N \times N$. Los ejemplos de un tipo de partición incluyen particiones simétricas que se obtienen dividiendo simétricamente una altura o anchura de la unidad de predicción, particiones obtenidas dividiendo asimétricamente la altura o anchura de la unidad de predicción, como $1:n$ o $n:1$, particiones que se obtienen dividiendo geoméricamente la unidad de predicción y particiones que tienen formas arbitrarias.

Un modo de predicción de la unidad de predicción puede ser al menos uno de un modo intra, un modo inter, y un modo de salto. Por ejemplo, el modo intra o el modo inter se puede realizar en la partición de $2N \times 2N$, $2N \times N$, $N \times 2N$, o $N \times N$. Además, el modo de salto solo se puede realizar en la partición de $2N \times 2N$. La codificación se realiza independientemente en una unidad de predicción en una unidad de codificación, seleccionando así un modo de predicción que tiene un error de codificación mínimo.

El aparato 100 de codificación de vídeo también puede realizar la transformación en los datos de imagen en una unidad de codificación basado no solo en la unidad de codificación para codificar los datos de imagen, sino también sobre la base de una unidad de datos que es diferente de la unidad de codificación.

Para realizar la transformación en la unidad de codificación, la transformación puede realizarse a base de una unidad de datos que tiene un tamaño menor o igual que la unidad de codificación. Por ejemplo, la unidad de datos para la transformación puede incluir una unidad de datos para un modo intra y una unidad de datos para un modo inter.

Una unidad de datos utilizada como base de la transformación se denominará ahora "unidad de transformación". De manera similar a la unidad de codificación, la unidad de transformación en la unidad de codificación se puede dividir recursivamente en regiones de tamaño más pequeño, de modo que la unidad de transformación se puede determinar independientemente en unidades de regiones. Así, los datos residuales en la unidad de codificación se pueden dividir de acuerdo con la unidad de transformación que tiene la estructura de árbol de acuerdo con las profundidades de transformación.

Se realiza una profundidad de transformación que indica el número de tiempos de división para alcanzar la unidad de transformación dividiendo la altura y la anchura de la unidad de codificación también puede establecerse en la unidad de transformación. Por ejemplo, en una unidad de codificación actual de $2N \times 2N$, una profundidad de transformación puede ser 0 cuando el tamaño de una unidad de transformación es $2N \times 2N$, puede ser 1 cuando el tamaño de una unidad de transformación es $N \times N$, y puede ser 2 cuando el tamaño de una unidad de transformación es $N/2 \times N/2$. Es decir, la unidad de transformación que tiene la estructura de árbol también puede establecerse de acuerdo con las profundidades de transformación.

La información de codificación según las unidades de codificación correspondientes a una profundidad codificada requiere no solo información sobre la profundidad codificada, sino también información relacionada con la codificación de predicción y la transformación. Por consiguiente, el determinador 120 de la unidad de codificación no solo determina una profundidad codificada que tiene un error de codificación mínimo, sino que también determina un tipo de partición en una unidad de predicción, un modo de predicción según las unidades de predicción y el tamaño de una unidad de transformación para la transformación.

Las unidades de codificación que tienen una estructura de árbol en una unidad de codificación máxima y un procedimiento para determinar una partición, de acuerdo con una realización de la presente invención, se describirán en detalle más adelante con referencia a las figuras 3 a 12.

El determinador 120 de la unidad de codificación puede medir un error de codificación de unidades de codificación más profundas de acuerdo con las profundidades mediante el uso de optimización de distorsión de frecuencia (RD) basada en multiplicadores lagrangianos.

La unidad 130 de salida emite los datos de imagen de la unidad de codificación máxima, que está codificada basada en la al menos una profundidad codificada determinada por el determinador 120 de la unidad de codificación, y la información sobre el modo de codificación de acuerdo con la profundidad codificada, en corrientes de bits.

Los datos de imagen codificados pueden ser obtenidos mediante la codificación de los datos residuales de una imagen.

La información sobre el modo de codificación de acuerdo con la profundidad codificada puede incluir información acerca de la profundidad codificada, sobre el tipo de partición en la unidad de predicción, el modo de predicción, y el tamaño de la unidad de transformación.

- La información acerca de la profundidad codificada puede definirse mediante el uso de información de división de acuerdo con profundidades, que indica si la codificación se realiza en unidades de una profundidad inferior en lugar de una profundidad actual de codificación. Si la profundidad actual de la unidad de codificación actual es la profundidad codificada, los datos de imagen en la unidad de codificación actual se codifican y emiten, y así la información de división puede definirse para no dividir la unidad de codificación actual a una profundidad menor. Alternativamente, si la profundidad actual de la unidad de codificación actual no es la profundidad codificada, la codificación se realiza en la unidad de codificación de la profundidad más baja, y así la información de división puede definirse para dividir la unidad de codificación actual para obtener las unidades de codificación de la profundidad más baja.
- 5
- 10 Si la profundidad actual no es la profundidad codificada, la codificación se realiza en la unidad de codificación que se divide en la unidad de codificación de la profundidad más baja. Como al menos una unidad de codificación de la profundidad inferior existe en una unidad de codificación de la profundidad actual, la codificación se realiza repetidamente en cada unidad de codificación de la profundidad inferior y, por lo tanto, la codificación puede realizarse recursivamente para las unidades de codificación que tienen la misma profundidad.
- 15 Dado que las unidades de codificación que tienen una estructura de árbol se determinan para una unidad de codificación máxima, y se determina información sobre al menos un modo de codificación para una unidad de codificación de una profundidad codificada, la información sobre al menos un modo de codificación se puede determinar para una unidad de codificación máxima. Además, una profundidad codificada de los datos de imagen de la unidad de codificación máxima puede ser diferente según las ubicaciones, ya que los datos de imagen se dividen jerárquicamente según las profundidades, y así puede establecerse información sobre la profundidad codificada y el modo de codificación para los datos de imagen.
- 20
- Por consiguiente, la unidad 130 de salida puede asignar información de codificación sobre una profundidad codificada correspondiente y un modo de codificación a al menos una de la unidad de codificación, la unidad de predicción y una unidad mínima incluida en la unidad de codificación máxima.
- 25 La unidad mínima, de acuerdo con una realización de la presente invención, es una unidad de datos rectangular obtenida dividiendo la unidad de codificación mínima que constituye la profundidad más baja por 4. Alternativamente, la unidad mínima puede ser una unidad de datos rectangular máxima que puede incluirse en todas las unidades de codificación, unidades de predicción, unidades de partición y unidades de transformación incluidas en la unidad de codificación máxima.
- 30 Por ejemplo, la información de codificación emitida a través de la unidad 130 de salida puede clasificarse en información de codificación de acuerdo con unidades de codificación, e información de codificación de acuerdo con unidades de predicción. La información de codificación según las unidades de codificación puede incluir información sobre el modo de predicción y sobre el tamaño de las particiones. La información de codificación de acuerdo con las unidades de predicción puede incluir información sobre una dirección estimada de un modo intermedio, sobre un índice de imagen de referencia del modo inter, sobre un vector de movimiento, sobre un componente de croma de un modo intra, y sobre un procedimiento de interpolación del modo intra. Además, la información sobre un tamaño máximo de la unidad de codificación definida de acuerdo con imágenes, cortes o GOP, y la información sobre una profundidad máxima pueden insertarse en un encabezado de una corriente de bits.
- 35
- 40 En el aparato 100 de codificación de vídeo, la unidad de codificación más profunda puede ser una unidad de codificación obtenida dividiendo una altura o anchura de una unidad de codificación de una profundidad superior, que es la capa superior, por dos. En otras palabras, cuando el tamaño de la unidad de codificación de la profundidad actual es $2N \times 2N$, el tamaño de la unidad de codificación de la profundidad inferior es $N \times N$. Además, la unidad de codificación de la profundidad actual que tiene el tamaño de $2N \times 2N$ puede incluir un número máximo de 4 unidades de codificación de la profundidad más baja.
- 45 Por consiguiente, el aparato 100 de codificación de vídeo puede formar las unidades de codificación que tienen la estructura de árbol mediante la determinación de unidades de codificación que tienen una forma óptima y un tamaño óptimo para cada unidad de codificación máxima, basándose en el tamaño de la unidad de codificación máxima y la profundidad máxima determinadas teniendo en cuenta las características de la imagen actual. Además, dado que la codificación se puede realizar en cada unidad de codificación máxima mediante el uso de cualquiera de los diversos modos de predicción y transformaciones, se puede determinar un modo de codificación óptimo considerando las características de la unidad de codificación de diversos tamaños de imagen.
- 50
- Así, si una imagen que tiene alta resolución o una gran cantidad de datos está codificada en un macrobloque convencional, un número de macrobloques por imagen aumenta excesivamente. En consecuencia, se incrementan varias piezas de información comprimida generadas para cada macrobloque, y por lo tanto es difícil transmitir la información comprimida y la eficacia de compresión de datos disminuye. Sin embargo, usando el aparato 100 de codificación de vídeo, la eficacia de compresión de la imagen puede aumentar, ya que se ajusta una unidad de codificación al considerar las características de una imagen al tiempo que se aumenta el tamaño máximo de una unidad de codificación al considerar el tamaño de la imagen.
- 55

La figura 2 es un diagrama de bloques de un aparato 200 de decodificación de vídeo, de acuerdo con una realización de la presente invención.

5 El aparato 200 de decodificación de vídeo incluye un receptor 210, un extractor 220 de datos de imagen y de información de codificación, y un decodificador 230 de datos de imagen. Las definiciones de diversos términos, tales como una unidad de codificación, una profundidad, una unidad de predicción, una unidad de transformación e información sobre varios modos de codificación, para varias operaciones del aparato 200 de decodificación de vídeo son idénticas a las descritas con referencia a la figura 1 y al aparato 100 de codificación de vídeo.

10 El receptor 210 recibe y analiza una corriente de bits de un vídeo codificado. El extractor 220 de datos de imagen y de información de codificación extrae los datos de imagen codificados para cada unidad de codificación de la corriente de bits analizada, en el que las unidades de codificación tienen una estructura de árbol de acuerdo con cada unidad de codificación máxima, y emiten los datos de imagen extraídos al decodificador 230 de datos de imagen. El extractor 220 de datos de imagen y de información de codificación pueden extraer información sobre un tamaño máximo de una unidad de codificación de una imagen actual, desde un encabezado sobre la imagen actual.

15 Además, el extractor 220 de datos de imagen y de información de codificación extrae información sobre una profundidad codificada y un modo de codificación para las unidades de codificación que tienen una estructura de árbol de acuerdo con cada unidad de codificación máxima, a partir de la corriente de bits analizada. La información extraída sobre la profundidad codificada y el modo de codificación se envía al decodificador 230 de datos de imágenes. En otras palabras, los datos de imagen en una corriente de bits se dividen en la unidad de codificación máxima, de modo que el decodificador 230 de datos de imagen decodifica los datos de imagen para cada unidad de codificación máxima.

20 La información sobre la profundidad codificada y el modo de codificación según la unidad de codificación máxima puede establecerse para información sobre al menos una unidad de codificación correspondiente a la profundidad codificada, y la información sobre un modo de codificación puede incluir información sobre un tipo de partición de una unidad de codificación correspondiente que corresponde a la profundidad codificada, sobre un modo de predicción, y un tamaño de una unidad de transformación. Además, la información de división según profundidades se puede extraer como la información sobre la profundidad codificada.

25 La información acerca de la profundidad codificada y el modo de codificación de acuerdo a cada unidad de codificación máxima extraída por el extractor 220 de datos de imagen y de información de codificación es información sobre una profundidad codificada y un modo de codificación determinado para generar un error de codificación mínimo cuando un codificador, tal como el aparato 100 de codificación de vídeo, realiza repetidamente codificación para cada unidad de codificación más profunda de acuerdo con las profundidades de acuerdo con cada unidad de codificación máxima. En consecuencia, el aparato 200 de decodificación de vídeo puede restaurar una imagen decodificando los datos de imágenes de acuerdo con una profundidad codificada y un modo de codificación que genera el mínimo error de codificación.

30 Dado que la información de codificación sobre la profundidad codificada y el modo de codificación puede asignarse a una unidad de datos predeterminada de entre una unidad de codificación, una unidad de predicción y una unidad mínima correspondiente, el extractor 220 de datos de imagen y de información de codificación puede extraer la información sobre la profundidad codificada y el modo de codificación de acuerdo con las unidades de datos predeterminadas. Las unidades de datos predeterminadas a las que se les asigna la misma información sobre la profundidad codificada y el modo de codificación pueden inferirse para ser las unidades de datos incluidas en la misma unidad de codificación máxima.

35 El decodificador 230 de datos de imagen restaura la imagen actual decodificando los datos de imagen en cada unidad de codificación máxima basándose en la información sobre la profundidad codificada y el modo de codificación según las unidades de codificación máximas. En otras palabras, el decodificador 230 de datos de imagen puede decodificar los datos de imagen codificados basándose en la información extraída sobre el tipo de partición, el modo de predicción y la unidad de transformación para cada unidad de codificación de entre las unidades de codificación que tienen la estructura de árbol incluida en cada unidad de codificación máxima. Un proceso de decodificación puede incluir predicción que incluye intra predicción y compensación de movimiento, y transformación inversa. La transformación inversa se puede realizar de acuerdo con un procedimiento de transformación ortogonal inversa o transformación entera inversa.

40 El decodificador 230 de datos de imagen puede realizar intra predicción o compensación de movimiento de acuerdo con una partición y un modo de predicción de cada unidad de codificación, basándose en la información sobre el tipo de partición y el modo de predicción de la unidad de predicción de la unidad de codificación de acuerdo con profundidades codificadas.

45 Además, el decodificador 230 de datos de imagen puede realizar la transformación inversa de acuerdo con cada unidad de transformación en la unidad de codificación, basándose en la información sobre el tamaño de la unidad de transformación de la unidad de codificación de acuerdo con profundidades codificadas, así como para realizar la transformación inversa según unidades de codificación máxima.

- 5 El decodificador 230 de datos de imagen puede determinar al menos una profundidad codificada de una unidad de codificación máxima actual utilizando información de división de acuerdo con las profundidades. Si la información de división indica que los datos de la imagen ya no se dividen en la profundidad actual, la profundidad actual es una profundidad codificada. Por consiguiente, el decodificador 230 de datos de imagen puede decodificar datos codificados de al menos una unidad de codificación correspondiente a cada profundidad codificada en la unidad de codificación máxima actual utilizando la información sobre el tipo de partición de la unidad de predicción, el modo de predicción y el tamaño de la unidad de transformación para cada unidad de codificación correspondiente a la profundidad codificada, y emite los datos de imagen de la unidad de codificación máxima actual.
- 10 En otras palabras, las unidades de datos que contienen la información de codificación que incluye la misma información de división puede ser recopiladas por la observación del conjunto de información de codificación asignado para la unidad de datos predeterminada de entre la unidad de codificación, la unidad de predicción, y la unidad mínima, y las unidades de datos recopiladas pueden considerarse como una unidad de datos que se decodifica por el decodificador 230 de datos de imagen en el mismo modo de codificación.
- 15 El aparato 200 de decodificación de imágenes puede obtener información sobre al menos una unidad de codificación que genera el mínimo error de codificación cuando la codificación se realiza recursivamente para cada unidad de codificación máxima, y puede usar la información para decodificar la imagen actual. En otras palabras, las unidades de codificación que tienen la estructura de árbol determinada para ser las unidades de codificación óptimas en cada unidad de codificación máxima pueden decodificarse. Además, el tamaño máximo de la unidad de codificación se determina teniendo en cuenta la resolución y una cantidad de datos de imagen.
- 20 Por consiguiente, incluso si los datos de imagen tienen una alta resolución y una gran cantidad de datos, los datos de imagen se pueden decodificar y restaurar de manera eficiente utilizando un tamaño de una unidad de codificación y un modo de codificación, que se determinan de forma adaptativa según las características de los datos de imagen, mediante el uso de información sobre un modo de codificación óptimo recibido desde un codificador.
- 25 Un procedimiento de determinación de las unidades de codificación que tienen una estructura de árbol, una unidad de predicción, y una unidad de transformación, de acuerdo con una realización de la presente invención, se describirá ahora con referencia a las figuras 3 a 13.
- La figura 3 es un diagrama para describir un concepto de unidades de codificación, de acuerdo con una realización de la presente invención.
- 30 Un tamaño de una unidad de codificación puede expresarse en anchura x altura, y puede ser de 64x64, 32x32, 16x16 y 8x8. Una unidad de codificación de 64x64 puede dividirse en particiones de 64x64, 64x32, 32x64, o 32x32, y una unidad de codificación de 32x32 puede dividirse en particiones de 32x32, 32x16, 16x32, o 16x16, una unidad de codificación de 16x16 puede dividirse en particiones de 16x16, 16x8, 8x16 u 8x8 y una unidad de codificación de 8x8 se puede dividir en particiones de 8x8, 8x4, 4x8 o 4x4.
- 35 En datos 310 de vídeo, una resolución es 1920x1080, un tamaño máximo de una unidad de codificación es 64, y una profundidad máxima es 2. En datos 320 de vídeo, una resolución es 1920x1080, un tamaño máximo de una unidad de codificación es 64, y una profundidad máxima es 3. En datos 330 de vídeo, una resolución es 352x288, un tamaño máximo de una unidad de codificación es 16, y una profundidad máxima es 1. La profundidad máxima mostrada en la figura 3 indica un número total de divisiones desde una unidad de codificación máxima a una unidad de decodificación mínima.
- 40 Si una resolución es alta o la cantidad de datos es grande, un tamaño máximo de una unidad de codificación puede ser grande para no solo aumentar la eficacia de la codificación, sino también para reflejar con precisión las características de una imagen. En consecuencia, el tamaño máximo de la unidad de codificación de los datos de vídeo 310 y 320 que tiene la resolución más alta que los datos de vídeo 330 puede ser 64.
- 45 Como la profundidad máxima de los datos 310 de vídeo es 2, las unidades 315 de codificación de los datos 310 de vídeo pueden incluir una unidad de codificación máxima que tiene un tamaño de eje largo de 64 y unidades de codificación que tienen tamaños de eje largo de 32 y 16, ya que las profundidades se profundizan en dos capas dividiendo la unidad de codificación máxima dos veces. Mientras tanto, como la profundidad máxima de los datos 330 de vídeo es 1, las unidades 335 de codificación de los datos de vídeo 330 pueden incluir una unidad de codificación máxima que tiene un tamaño de eje largo de 16 y unidades de codificación que tienen un tamaño de eje largo de 8, ya que las profundidades se profundizan en una capa dividiendo la unidad de codificación máxima una vez.
- 50 Como la profundidad máxima de los datos 320 de vídeo es 3, las unidades 325 de codificación de los datos de vídeo 320 pueden incluir una unidad de codificación máxima que tiene un tamaño de eje largo de 64 y unidades de codificación que tienen tamaños de eje largo de 32, 16 y 8 porque las profundidades se profundizan en 3 capas dividiendo la unidad de codificación máxima tres veces. A medida que se profundiza una profundidad, la información detallada se puede expresar con precisión.
- 55 La figura 4 es un diagrama de bloques de un codificador 400 de imágenes 400 basado en unidades de codificación,

de acuerdo con una realización de la presente invención.

5 El codificador 400 de imágenes realiza operaciones del determinador 120 de unidad de codificación del aparato 100 de codificación de vídeo para codificar datos de imágenes. En otras palabras, un intra predictor 410 realiza intra predicción en unidades de codificación en un modo intra, entre un marco 405 actual, y un estimador 420 de movimiento y un compensador 425 de movimiento realiza una inter estimación y una compensación de movimiento en unidades de codificación en un modo inter entre el cuadro actual 405 utilizando el marco 405 actual, y un marco 495 de referencia.

10 La salida de datos desde el intra predictor 410, estimador 420 de movimiento, y el compensador 420 de movimiento se emiten como un coeficiente de transformación cuantificado a través de un transformador 430 y un cuantificador 440. El coeficiente de transformación cuantificado se restablece como datos en un dominio espacial a través de un cuantificador 460 inverso y un transformador 470 inverso, y los datos restaurados en el dominio espacial salen como la trama 495 de referencia después de ser procesados a través de una unidad 480 de desbloqueo y una unidad 490 de filtrado de bucle. El coeficiente de transformación cuantificado puede emitirse como un flujo de bits 455 a través de un codificador 450 de entropía.

15 Para que el codificador 400 de imágenes se aplique en el aparato 100 de codificación de vídeo, todos los elementos del codificador 400 de imágenes, es decir, el intra predictor 410, 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 operaciones basadas en cada unidad de codificación entre las unidades de codificación que tienen una estructura de árbol mientras consideran la profundidad máxima de cada unidad de codificación máxima.

20 Específicamente, el intra predictor 410, el estimador 420 de movimiento, y el compensador 425 de movimiento determina particiones y un modo de predicción de cada unidad de codificación entre las unidades de codificación que tienen una estructura de árbol considerando el tamaño máximo y la profundidad máxima de una unidad de codificación máxima actual, y el transformador 430 determina el tamaño de la unidad de transformación en cada unidad de codificación entre las unidades de codificación que tienen una estructura de árbol.

La figura 5 es un diagrama de bloques de un decodificador 500 de imágenes basado en unidades de codificación, de acuerdo con una realización de la presente invención.

30 Un analizador 510 analiza los datos de imagen codificados a decodificar e información sobre la codificación requerida para decodificar desde un flujo de bits 505. Los datos de imagen codificados se emiten como datos cuantificados inversos a través de un decodificador 520 de entropía y un cuantificador 530 inverso, y los datos cuantificados inversos se restauran a datos de imagen en un dominio espacial a través de un transformador 540 inverso.

35 Un intra predictor 550 realiza intra predicción en las unidades de codificación en un modo intra con respecto a los datos de imagen en el dominio espacial, y un compensador 560 de movimiento realiza la compensación de movimiento en las unidades de codificación en un modo inter mediante el uso de un marco 585 de referencia.

40 Los datos de imagen en el dominio espacial, que pasaron a través del intra predictor 550 y el compensador 560 de movimiento, pueden emitirse como un cuadro 595 restaurado después de ser procesados posteriormente a través de una unidad 570 de desbloqueo y una unidad 580 de filtrado de bucle. Además, los datos de imagen, que se procesan posteriormente a través de la unidad 570 de desbloqueo y la unidad 580 de filtrado de bucle, pueden emitirse como marco 585 de referencia.

Para decodificar los datos de imagen en el decodificador 230 de datos de imagen del aparato 200 de decodificación de vídeo, el descodificador 500 de imágenes puede realizar las operaciones que se realizan después de realizarse las operaciones del analizador 510.

45 Para que el decodificador 500 de imágenes que debe aplicarse en el aparato 200 de decodificación de vídeo, todos los elementos del decodificador 500 de imágenes, es decir, el analizador 510, el decodificador 520 de entropía, el cuantificador 530 inverso, el transformador 540 inverso, el intra predictor 550, el compensador 560 de movimiento, la unidad 570 de desbloqueo y la unidad 580 de filtrado de bucle realizan operaciones basadas en unidades de codificación que tienen una estructura de árbol para cada unidad de codificación máxima.

50 Específicamente, el intra predictor 550 y el compensador 560 de movimiento realizan operaciones basadas en particiones y un modo de predicción para cada una de las unidades de codificación que tienen una estructura de árbol, y el transformador 540 inverso realiza operaciones sobre la base de un tamaño de una unidad de transformación para cada unidad de codificación.

La figura 6 es un diagrama que ilustra unidades de codificación más profundas según profundidades y particiones, de acuerdo con una realización de la presente invención.

55 El aparato 100 de codificación de vídeo y el aparato 200 de decodificación de vídeo usan unidades de codificación

- 5 jerárquicas para considerar las características de una imagen. Una altura máxima, una anchura máxima y una profundidad máxima de unidades de codificación pueden determinarse de manera adaptativa según las características de la imagen, o pueden ser configuradas de manera diferente por un usuario. Los tamaños de las unidades de codificación más profundas según las profundidades se pueden determinar de acuerdo con el tamaño máximo predeterminado de la unidad de codificación.
- 10 En una estructura 600 jerárquica de unidades de codificación, de acuerdo con una realización de la presente invención, la altura máxima y la anchura máxima de las unidades de codificación son 64 y la profundidad máxima es 4. Dado que una profundidad se profundiza a lo largo de un eje vertical de la estructura 600 jerárquica, se dividen una altura y una anchura de la unidad de codificación más profunda. Además, una unidad de predicción y particiones, que son las bases para la codificación de predicción de cada unidad de codificación más profunda, se muestran a lo largo de un eje horizontal de la estructura 600 jerárquica.
- 15 En otras palabras, una unidad 610 de codificación es una unidad máxima de codificación en la estructura 600 jerárquica, en el que una profundidad es 0 y un tamaño, es decir, una altura por anchura es de 64x64. La profundidad se profundiza a lo largo del eje vertical, y existe una unidad 620 de codificación que tiene un tamaño de 32x32 y una profundidad de 1, una unidad 630 de codificación que tiene un tamaño de 16x16 y una profundidad de 2, una unidad 640 de codificación que tiene un tamaño de 8x8 y una profundidad de 3, y una unidad 650 de codificación que tiene un tamaño de 4x4 y una profundidad de 4. La unidad 650 de codificación que tiene el tamaño de 4x4 y la profundidad de 4 es una unidad de codificación mínima.
- 20 La unidad de predicción y las particiones de una unidad de codificación están dispuestas a lo largo del eje horizontal de acuerdo con cada profundidad. En otras palabras, si la unidad 610 de codificación que tiene el tamaño de 64x64 y la profundidad de 0 es una unidad de predicción, la unidad de predicción se puede dividir en particiones incluidas en la unidad 610 de codificación, es decir, una partición 610 que tiene un tamaño de 64x64, particiones 612 que tienen el tamaño de 64x32, particiones 614 que tienen el tamaño de 32x64, o particiones 616 que tienen el tamaño de 32x32.
- 25 Del mismo modo, una unidad de predicción de la unidad 620 de codificación que tiene el tamaño de 32x32 y la profundidad de 1 podrá dividir en particiones incluidas en la unidad 620 de codificación, es decir, una partición 620 que tiene un tamaño de 32x32, particiones 622 que tienen un tamaño de 32x16, particiones 624 que tienen un tamaño de 16x32, y particiones 626 que tienen un tamaño de 16x16.
- 30 Del mismo modo, una unidad de predicción de la unidad 630 de codificación que tiene el tamaño de 16x16 y la profundidad de 2 se puede dividir en particiones incluidas en la unidad 630 de codificación, es decir, una partición que tiene un tamaño de 16x16 incluida en la unidad de codificación 630, teniendo las particiones 632 un tamaño de 16x8, teniendo las particiones 634 un tamaño de 8x16, y teniendo las particiones 636 un tamaño de 8x8.
- 35 De manera similar, una unidad de predicción de la unidad 640 de codificación que tiene el tamaño de 8x8 y la profundidad de 3 puede dividirse en particiones incluidas en la unidad 640 de codificación, es decir, una partición que tiene un tamaño de 8x8 incluida en la unidad 640 de codificación, particiones 642 que tienen un tamaño de 8x4, particiones 644 tienen un tamaño de 4x8 y particiones 646 que tienen un tamaño de 4x4.
- La unidad 650 de codificación que tiene el tamaño de 4x4 y la profundidad de 4 es la unidad de codificación mínima y una unidad de codificación de la profundidad más baja. Una unidad de predicción de la unidad 650 de codificación solo está asignada a una partición que tiene un tamaño de 4x4.
- 40 Para determinar la al menos una profundidad codificada de las unidades de codificación que constituyen la unidad 610 de codificación máxima, el determinador 120 de la unidad de codificación del aparato 100 de codificación de vídeo realiza codificación para las unidades de codificación correspondientes a cada profundidad incluida en la unidad 610 de codificación máxima.
- 45 Un número de unidades de codificación más profundas de acuerdo con las profundidades, incluyendo los datos en el mismo rango y el mismo tamaño aumenta a medida que la profundidad se profundiza. Por ejemplo, se requieren cuatro unidades de codificación correspondientes a una profundidad de 2 para cubrir los datos que se incluyen en una unidad de codificación correspondiente a una profundidad de 1. En consecuencia, para comparar los resultados de codificación de los mismos datos de acuerdo con las profundidades, la unidad de codificación correspondiente a la profundidad de 1 y cuatro unidades de codificación correspondientes a la profundidad de 2 están codificadas.
- 50 Para realizar la codificación para una profundidad actual de entre las profundidades, al menos un error de codificación puede ser seleccionado para la profundidad actual de ejecución de la codificación para cada unidad de predicción en las unidades de codificación correspondientes a la profundidad actual, a lo largo del eje horizontal de la estructura 600 jerárquica. Alternativamente, se puede buscar el error de codificación mínimo comparando los errores de codificación mínimos según las profundidades y realizando una codificación para cada profundidad a medida que la profundidad se profundiza a lo largo del eje vertical de la estructura 600 jerárquica. Se puede seleccionar una profundidad y una partición que tengan el mínimo error de codificación en la unidad 610 de codificación como la profundidad codificada y un tipo de partición de la unidad 610 de codificación.
- 55

La figura 7 es un diagrama para describir una relación entre una unidad 710 de codificación y unidades 720 de transformación, de acuerdo con una realización de la presente invención.

5 El aparato 100 o 200 de codificación de vídeo codifica o decodifica una imagen de acuerdo con unidades de codificación que tienen tamaños menores que o iguales a una unidad de codificación máxima para cada unidad de codificación máxima. Los tamaños de las unidades de transformación para la transformación durante la codificación pueden seleccionarse basándose en unidades de datos que no son más grandes que una unidad de codificación correspondiente.

10 Por ejemplo, en el aparato 100 o 200 de codificación de vídeo, si un tamaño de la unidad 710 de codificación es 64x64, la transformación puede realizarse mediante el uso de las unidades 720 de transformación que tienen un tamaño de 32x32.

Además, los datos de la unidad 710 de codificación que tienen el tamaño de 64x64 pueden codificarse mediante la realización de la transformación en cada una de las unidades de transformación que tienen el tamaño de 32x32, 16x16, 8x8, y 4x4, que son más pequeñas que 64x64, y luego puede seleccionarse una unidad de transformación que tiene el menor error de codificación.

15 La figura 8 es un diagrama para describir información de codificación de unidades de codificación correspondientes a una profundidad codificada, de acuerdo con una realización de la presente invención.

20 La unidad 130 de salida del aparato 100 de codificación de vídeo puede codificar y transmitir información 800 acerca de un tipo de partición, la información 810 acerca de un modo de predicción, y la información 820 sobre un tamaño de una unidad de transformación para cada unidad de codificación correspondiente a una profundidad codificada, como información sobre un modo de codificación.

25 La información 800 indica información acerca de una forma de una partición obtenida por la división de una unidad de predicción de una unidad de codificación actual, en el que la partición es una unidad de datos para la predicción que codifica la unidad de codificación actual. Por ejemplo, una unidad de codificación actual CU_0 que tiene un tamaño de 2Nx2N puede dividirse en cualquiera de una partición 802 que tiene un tamaño de 2Nx2N, una partición 804 que tiene un tamaño de 2NxN, una partición 806 que tiene un tamaño de Nx2N, y una partición 808 que tiene un tamaño de NxN. Aquí, la información 800 sobre un tipo de partición se establece para indicar que una partición 804 tiene un tamaño de 2NxN, la partición 806 que tiene un tamaño de Nx2N y la partición 808 que tiene un tamaño de NxN. La información 810 indica un modo de predicción de cada partición. Por ejemplo, la información 810 puede indicar un modo de codificación de predicción realizada en una partición indicada por la información 800, es decir, un modo intra 812, un modo inter 814 o un modo de salto 816.

30 La información 820 indica una unidad de transformación en función de cuándo se realiza la transformación en una unidad de codificación actual. Por ejemplo, la unidad de transformación puede ser una primera unidad 822 de transformación, una segunda unidad 824 de transformación, una primera unidad 826 de transformación o una segunda unidad 828 de transformación.

35 El extractor 220 de datos de imagen y de información de codificación del aparato 200 de decodificación de vídeo puede extraer y usar la información 800, 810 y 820 para la decodificación, de acuerdo con cada unidad de codificación más profunda.

La figura 9 es un diagrama de unidades de codificación más profundas de acuerdo con las profundidades, de acuerdo con una realización de la presente invención.

40 La información de división se puede usar para indicar un cambio de profundidad. La información de división indica si una unidad de codificación de una profundidad actual se divide en unidades de codificación de una profundidad inferior.

45 Una unidad 910 de predicción para la codificación de predicción de una unidad 900 de codificación que tiene una profundidad de 0 y un tamaño de 2N_0x2N_0 puede incluir particiones de un tipo de partición 912 que tiene un tamaño de 2N_0x2N_0, un tipo de partición 914 que tiene un tamaño de 2N_0xN_0, un tipo de partición 916 que tiene un tamaño de N_0x2N_0, y un tipo de partición 918 que tiene un tamaño de N_0xN_0. La figura 9 solo ilustra los tipos de partición 912 a 918 que se obtienen al dividir simétricamente la unidad 910 de predicción, pero el tipo de partición no está limitado a los mismos, y las particiones de la unidad 910 de predicción pueden incluir particiones asimétricas, particiones que tienen una forma predeterminada y particiones que tienen una forma geométrica.

50 La codificación de predicción se realiza repetidamente en una partición que tiene un tamaño de 2N_0x2N_0, dos particiones que tienen un tamaño de 2N_0xN_0, dos particiones que tienen un tamaño de N_0x2N_0 y cuatro particiones que tienen un tamaño de N_0xN_0, según cada tipo de partición. La codificación de predicción en un modo intra y un modo inter se puede realizar en las particiones que tienen los tamaños de 2N_0x2N_0, N_0x2N_0, 2N_0xN_0, y N_0xN_0. La codificación de predicción en un modo de salto se realiza solo en la partición que tiene el tamaño de 2N_0x2N_0.

55

Si el error de codificación es más pequeño en uno de los tipos de partición 912 a 916 que tienen los tamaños de $2N_0 \times 2N_0$, $2N_0 \times N_0$ y $N_0 \times 2N_0$, la unidad 910 de predicción no se puede dividir en una profundidad menor.

5 Si el error de codificación es el más pequeño en el tipo de partición 918, que tiene un tamaño de $N_0 \times N_0$, una profundidad cambia de 0 a 1 para dividir el tipo de partición 918 en la operación 920, y la codificación se realiza repetidamente en las unidades 930 de codificación que tienen una profundidad de 2 y un tamaño de $N_0 \times N_0$ para buscar un error de codificación mínimo.

10 Una unidad 940 de predicción para la codificación de predicción de la unidad 930 de codificación que tiene una profundidad de 1 y un tamaño de $2N_1 \times 2N_1 (=N_0 \times N_0)$ puede incluir particiones de un tipo de partición 942 que tiene un tamaño de $2N_1 \times 2N_1$, un tipo de partición 944 que tiene un tamaño de $2N_1 \times N_1$, un tipo de partición 946 que tiene un tamaño de $N_1 \times 2N_1$, y un tipo de partición 948 que tiene un tamaño de $N_1 \times N_1$.

Si un error de codificación es el más pequeño en el tipo de partición 948, que tiene un tamaño de $N_1 \times N_1$, una profundidad cambia de 1 a 2 para dividir el tipo de partición 948 en la operación 950, y la codificación se realiza repetidamente en las unidades 960 de codificación que tienen una profundidad de 2 y un tamaño de $N_2 \times N_2$ para buscar un error de codificación mínimo.

15 Cuando una profundidad máxima es d , la operación dividida de acuerdo con cada profundidad se puede realizar hasta cuando una profundidad se convierte en $d-1$, y la información de división se puede codificar como hasta cuando una profundidad es de 0 a $d-2$. En otras palabras, codificación se realiza hasta cuando la profundidad es $d-1$ después de que una unidad de codificación correspondiente a una profundidad de $d-2$ se divide en la operación 970, una unidad 990 de predicción para la codificación de predicción de una unidad 980 de codificación que tiene una
20 profundidad de $d-1$ y un tamaño de $2N_{(d-1)} \times 2N_{(d-1)}$ puede incluir particiones de un tipo de partición 992 que tiene un tamaño de $2N_{(d-1)} \times 2N_{(d-1)}$, un tipo de partición 994 que tiene un tamaño de $2N_{(d-1)} \times N_{(d-1)}$, un tipo de partición 996 que tiene un tamaño de $N_{(d-1)} \times 2N_{(d-1)}$, y un tipo de partición 998 que tiene un tamaño de $N_{(d-1)} \times N_{(d-1)}$.

25 La codificación de predicción se puede realizar repetidamente en una partición que tiene un tamaño de $2N_{(d-1)} \times 2N_{(d-1)}$, dos particiones que tienen un tamaño de $2N_{(d-1)} \times N_{(d-1)}$, dos particiones que tienen un tamaño de $N_{(d-1)} \times 2N_{(d-1)}$, cuatro particiones que tienen un tamaño de $N_{(d-1)} \times N_{(d-1)}$ entre los tipos de partición 992 a 998 para buscar un tipo de partición que tenga un error de codificación mínimo.

30 Incluso cuando el tipo de partición 998 que tiene el tamaño de $N_{(d-1)} \times N_{(d-1)}$ tiene el error de codificación mínimo, como una profundidad máxima es d , una unidad de codificación $CU_{(d-1)}$ que tiene una profundidad de $d-1$ ya no se divide a una profundidad menor, y se determina que una profundidad codificada para las unidades de codificación que constituyen una unidad 900 de codificación máxima actual es $d-1$ y una partición el tipo de la unidad 900 de codificación máxima actual se puede determinar que sea $N_{(d-1)} \times N_{(d-1)}$. Además, dado que la profundidad máxima es d y una unidad 980 de codificación mínima que tiene una profundidad más baja de $d-1$ ya no se divide a una profundidad más baja, la información de división para la unidad 980 de codificación mínima no está
35 configurada.

Una unidad 999 de datos puede ser una 'unidad mínima' para la unidad de codificación máxima actual. Una unidad mínima, de acuerdo con una realización de la presente invención, puede ser una unidad de datos rectangular obtenida dividiendo una unidad 980 de codificación mínima por 4. Al realizar la codificación repetidamente, el
40 aparato 100 de codificación de vídeo puede seleccionar una profundidad que tenga el menor error de codificación comparando errores de codificación según las profundidades de la unidad 900 de codificación para determinar una profundidad codificada, y establecer un tipo de partición correspondiente y un modo de predicción como un modo de codificación de la profundidad codificada.

45 Como tales, los errores de codificación mínimos según las profundidades se comparan en todas las profundidades de 1 a d , y una profundidad que tiene el menor error de codificación se puede determinar como una profundidad codificada. La profundidad codificada, el tipo de partición de la unidad de predicción y el modo de predicción pueden codificarse y transmitirse como información sobre un modo de codificación. Además, como una unidad de codificación se divide desde una profundidad de 0 a una profundidad codificada, solo la información de división de la profundidad codificada se establece en 0, y la información de división de profundidades excluyendo la profundidad codificada se establece en 1.

50 El extractor 220 de datos de imagen e información de codificación del aparato 200 de decodificación de vídeo puede extraer y utilizar la información acerca de la profundidad codificada y la unidad de predicción de la unidad 900 de codificación para decodificar la partición 912. El aparato 200 de decodificación de vídeo puede determinar una profundidad, en la que la información de división es 0, como una profundidad codificada mediante el uso de información de división de acuerdo con las profundidades, y el uso de información sobre un modo de codificación de
55 la profundidad correspondiente para la decodificación.

Las figuras 10 a 12 son diagramas para describir una relación entre unidades 1060 de codificación, unidades 1060 de predicción y unidades 1070 de transformación, de acuerdo con una realización de la presente invención.

Las unidades 1010 de codificación son unidades de codificación que tienen una estructura de árbol, correspondiente a profundidades codificadas determinadas por el aparato 100 de codificación de vídeo, en una unidad de codificación máxima. Las unidades 1060 de predicción son particiones de unidades de predicción de cada una de las unidades 1010 de codificación, y las unidades 1070 de transformación son unidades de transformación de cada una de las unidades 1010 de codificación.

Cuando una profundidad de una unidad de codificación máxima es 0 en las unidades 1010 de codificación, las profundidades de las unidades 1012 y 1054 de codificación son 1, las profundidades de las unidades 1014, 1016, 1018, 1028, 1050 y 1052 de codificación son 2, las profundidades de las unidades 1020, 1022, 1024, 1026, 1030, 1032 y 1048 de codificación son 3, y las profundidades de las unidades 1040, 1042, 1044 y 1046 de codificación son 4.

En las unidades 1060 de predicción, algunas unidades 1014, 1016, 1022, 1032, 1048, 1050, 1052 y 1054 de codificación se obtienen dividiendo las unidades de codificación. En otras palabras, los tipos de partición en las unidades 1014, 1022, 1050 y 1054 de codificación tienen un tamaño de $2N \times N$, los tipos de partición en las unidades 1016, 1048 y 1052 de codificación tienen un tamaño de $N \times 2N$, y un tipo de partición de la unidad 1032 de codificación tiene un tamaño de $N \times N$. Las unidades de predicción y las particiones de las unidades 1010 de codificación son más pequeñas que o iguales a cada unidad de codificación.

La transformación o transformación inversa se realiza en datos de imagen de la unidad 1052 de codificación en las unidades 1070 de transformación en una unidad de datos que es más pequeña que la unidad 1052 de codificación. Además, las unidades 1014, 1016, 1022, 1032, 1048, 1050 y 1052 de codificación en las unidades 1070 de transformación son diferentes de las unidades 1060 de predicción en términos de tamaños y formas. En otras palabras, los aparatos 100 y 200 de codificación y decodificación de vídeo pueden realizar intra predicción, estimación de movimiento, compensación de movimiento, transformación y transformación inversa individualmente en una unidad de datos en la misma unidad de codificación.

En consecuencia, la codificación se realiza recursivamente en cada una de las unidades de codificación que tienen una estructura jerárquica en cada región de una unidad de codificación máxima para determinar una unidad de codificación óptima, y de este modo se pueden obtener unidades de codificación que tienen una estructura de árbol recursiva. La información de codificación puede incluir información dividida sobre una unidad de codificación, información sobre un tipo de partición, información sobre un modo de predicción e información sobre el tamaño de una unidad de transformación. La Tabla 1 muestra la información de codificación que puede establecerse mediante los aparatos 100 y 200 de codificación y decodificación de vídeo.

Tabla 1

Información dividida 0 (Codificación en la unidad de codificación con tamaño de $2N \times 2N$ y profundidad actual de d)					Información dividida 1
Modo de predicción	Tipo de partición		Tamaño de la Unidad de Transformación		Repetidamente codifica unidades de codificación con profundidad inferior de $d + 1$
Intra Inter	Tipo de partición simétrica	Tipo de partición asimétrica	Información dividida 0 de la unidad de transformación	Información dividida 1 de la unidad de transformación	
	Omisión (solo $2N \times 2N$)	$2N \times 2N$	$2N \times nU$	$2N \times 2N$	
$2N \times N$		$2N \times nD$	$N/2 \times N/2$ (Tipo asimétrico)		
$N \times 2N$		$nL \times 2N$			
$N \times N$	$nR \times 2N$				

La unidad 130 de salida del aparato 100 de codificación de vídeo puede emitir la información de codificación acerca de las unidades de codificación que tienen una estructura de árbol, y el extractor 220 de datos de imagen y de información de codificación del aparato 200 de decodificación de vídeo puede extraer la información de codificación sobre las unidades de codificación que tienen una estructura de árbol de una corriente de bits recibida.

La información de división indica si una unidad de codificación actual se divide en unidades de codificación de una profundidad más baja. Si la información dividida de una profundidad actual d es 0, una profundidad, en la que una unidad de codificación actual ya no se divide en una profundidad inferior, es una profundidad codificada, y por lo

tanto información sobre un tipo de partición, modo de predicción y un tamaño de La unidad de transformación se puede definir para la profundidad codificada. Si la unidad de codificación actual se divide de acuerdo con la información dividida, la codificación se realiza de forma independiente en cuatro unidades de codificación divididas de una profundidad menor.

- 5 Un modo de predicción puede ser uno de un modo intra, un modo inter y un modo de salto. El modo intra y el modo inter se pueden definir en todos los tipos de partición, y el modo de salto se define solo en un tipo de partición que tenga un tamaño de $2N \times 2N$.

La información sobre el tipo de partición puede indicar tipos de particiones simétricas que tienen tamaños de $2N \times 2N$, $2N \times N$, $N \times 2N$, y $N \times N$, que se obtienen mediante la división simétricamente de una altura o una anchura de una unidad de predicción, y tipos de particiones asimétricas que tienen tamaños de $2N \times nU$, $2N \times nD$, $nL \times 2N$ y $nR \times 2N$, que se obtienen al dividir asimétricamente la altura o la anchura de la unidad de predicción. Los tipos de particiones asimétricas que tienen los tamaños de $2N \times nU$ y $2N \times nD$ pueden obtenerse respectivamente dividiendo la altura de la unidad de predicción en 1:3 y 3:1, y los tipos de partición asimétricos que tienen los tamaños de $nL \times 2N$ y $nR \times 2N$ pueden obtenerse respectivamente dividiendo la anchura de la unidad de predicción en 1:3 y 3:1.

15 El tamaño de la unidad de transformación puede establecerse para ser dos tipos en el modo intra y dos tipos en el modo inter. En otras palabras, si la información dividida de la unidad de transformación es 0, el tamaño de la unidad de transformación puede ser $2N \times 2N$, que es el tamaño de la unidad de codificación actual. Si la información dividida de la unidad de transformación es 1, las unidades de transformación se pueden obtener dividiendo la unidad de codificación actual. Además, si un tipo de partición de la unidad de codificación actual que tiene el tamaño de $2N \times 2N$ es un tipo de partición simétrica, el tamaño de una unidad de transformación puede ser $N \times N$, y si el tipo de partición de la unidad de codificación actual es un tipo de partición asimétrica, el tamaño de la unidad de transformación puede ser $N/2 \times N/2$.

La información de codificación sobre unidades de codificación que tienen una estructura de árbol puede incluir al menos una de una unidad de codificación correspondiente a una profundidad codificada, una unidad de predicción, y una unidad mínima. La unidad de codificación correspondiente a la profundidad codificada puede incluir al menos una de una unidad de predicción y una unidad mínima que contiene la misma información de codificación.

En consecuencia, se determina si las unidades de datos adyacentes se incluyen en la misma unidad de codificación que corresponde a la profundidad codificada mediante la comparación de la información de codificación de las unidades de datos adyacentes. Además, una unidad de codificación correspondiente que corresponde a una profundidad codificada se determina usando información de codificación de una unidad de datos, y así se puede determinar una distribución de profundidades codificadas en una unidad de codificación máxima.

Por consiguiente, si una unidad de codificación actual se predice basándose en la información de codificación de las unidades de datos adyacentes, la información de codificación de las unidades de datos en unidades de codificación más profundas adyacentes a la unidad de codificación actual se puede referir y usar directamente.

35 Alternativamente, se predice una unidad de codificación actual en función de la información de codificación de las unidades de datos adyacentes, las unidades de datos adyacentes a la unidad de codificación actual se buscan utilizando la información de codificación de las unidades de datos, y las unidades de codificación adyacentes buscadas pueden referirse para predecir la unidad de codificación actual.

La figura 13 es un diagrama para describir una relación entre una unidad de codificación, una unidad de predicción o una partición y una unidad de transformación, de acuerdo con la información del modo de codificación de la Tabla 1.

Una unidad 1300 de codificación máxima incluye unidades 1302, 1304, 1306, 1312, 1314, 1316 y 1318 de codificación de profundidades codificadas. Aquí, dado que la unidad 1318 de codificación es una unidad de codificación de una profundidad codificada, la información de división puede establecerse en 0. La información sobre un tipo de partición de la unidad 1318 de codificación que tiene un tamaño de $2N \times 2N$ puede configurarse para ser una de un tipo de partición 1322 que tiene un tamaño de $2N \times 2N$, un tipo de partición 1324 que tiene un tamaño de $2N \times N$, un tipo de partición 1326 que tiene un tamaño de $N \times 2N$, un tipo de partición 1328 que tiene un tamaño de $N \times N$, un tipo de partición 1332 que tiene un tamaño de $2N \times nU$, un tipo de partición 1334 que tiene un tamaño de $2N \times nD$, un tipo de partición 1336 que tiene un tamaño de $nL \times 2N$, y un tipo de partición 1338 que tiene un tamaño de $nR \times 2N$.

50 Cuando el tipo de partición está configurado para ser simétrico, es decir, el tipo de partición 1322, 1324, 1326 o 1328, se establece una unidad 1342 de transformación que tiene un tamaño de $2N \times 2N$ si la información de división (indicador de tamaño de TU) de una unidad de transformación es 0, y se establece una unidad 1344 de transformación que tiene un tamaño de $N \times N$ si un indicador de tamaño de TU es 1.

55 Cuando el tipo de partición está configurado para ser asimétrico, es decir, el tipo de partición 1332, 1334, 1336 o 1338, una unidad 1352 de transformación que tiene un tamaño de $2N \times 2N$ se establece si un indicador de tamaño de TU es 0, y se establece una unidad 1354 de transformación que tiene un tamaño de $N/2 \times N/2$ si un indicador de tamaño de TU es 1.

Un procedimiento de compensación de movimiento realizado por el compensador 425 de movimiento del codificador 400 de imagen de la figura 4 y el intra predictor 550 del decodificador 500 de imágenes de la figura 5 se explicará en detalle.

5 La figura 14 es un diagrama que ilustra un aparato 1400 de estimación de movimiento, de acuerdo con una realización de la presente invención. El aparato 1400 de estimación de movimiento de la figura 14 corresponde al estimador 420 de movimiento del codificador 400 de imágenes de la figura 4.

Una unidad 1410 de interpolación genera un marco de referencia interpolado mediante la interpolación de un marco de referencia y la generación de un valor de píxel de una posición de subpíxel. Una unidad 1420 de ejecución de estimación de movimiento realiza una estimación de movimiento con una precisión de subpíxel utilizando el marco de referencia interpolado, y determina y genera un vector de movimiento que indica una región correspondiente del marco de referencia que es más similar a una unidad de predicción. El compensador 425 de movimiento de la figura 4 obtiene un valor de predicción de píxeles de la unidad de predicción al obtener un subpíxel interpolado o un píxel entero del marco de referencia indicado por el vector de movimiento. Aunque la unidad 1410 de interpolación está incluida en el aparato 1400 de estimación de movimiento en la figura 14, la presente realización no se limita a esto y 15 la unidad 1410 de interpolación puede estar separada del aparato 1400 de estimación de movimiento.

La figura 15 es un diagrama que ilustra las posiciones de un subpíxel y un píxel entero seleccionables en un proceso de compensación de movimiento con una precisión de 1/4 de píxel, de acuerdo con una realización de la presente invención.

Con referencia a la figura 15, A 1501, B 1502, C 1503 y D 1504 etiquetados con letras mayúsculas indican píxeles enteros, y 1505 a 1519 etiquetados con letras minúsculas indican subpíxeles entre píxeles enteros. Cuando un vector de movimiento indica uno de los píxeles enteros A 1501, B 1502, C 1503 y D 1504, es decir, cuando las componentes horizontal y vertical del vector de movimiento tienen valores enteros, un valor de predicción de una unidad de predicción puede ser obtenido leyendo directamente los datos de un cuadro de referencia almacenado en una memoria durante un proceso de compensación de movimiento. En consecuencia, cuando los componentes 25 horizontales y verticales de un vector de movimiento tienen valores enteros, no se necesita un proceso de interpolación. Por ejemplo, cuando una unidad de predicción actual es LxM (L y M son enteros) y un componente horizontal MVx y un componente vertical MVy de un vector de movimiento MV obtenido después de la estimación de movimiento realizada en la unidad de predicción actual LxM tienen valores enteros, el compensador 425 de movimiento puede obtener un valor de predicción de la unidad de predicción actual LxM al obtener píxeles enteros de LxM correspondientes a una región correspondiente de una posición obtenida al cambiar la unidad de predicción actual LxM del marco de referencia almacenado en la memoria por un vector de movimiento (MVx, MVy). 30

Cuando el vector de movimiento indica un subpíxel cualquiera entre 1505 y 1519 en lugar de un píxel entero, se necesita un proceso de interpolación para obtener el subpíxel correspondiente. El proceso de interpolación del subpíxel se puede realizar aplicando un filtro de interpolación que tiene un número de referencia de filtro 35 predeterminado a píxeles enteros del marco de referencia. En un 1505, b 1506 y c 1507 que son subpíxeles dispuestos entre los píxeles enteros A 1501 y B 1502 que son horizontalmente adyacentes entre sí, solo un componente horizontal tiene un valor fraccional, y un componente vertical tiene un valor entero. Por consiguiente, para interpolar los subpíxeles a 1505, b 1506 y c 1507, se realiza un proceso de interpolación horizontal unidimensional (1 D). Del mismo modo, en d 1508, h 1512 e l 1516, que son subpíxeles dispuestos entre los píxeles enteros A 1501 y C 1503 que están adyacentes verticalmente entre sí, solo un componente vertical tiene un valor fraccional y un componente horizontal tiene un valor entero. Por consiguiente, para interpolar los subpíxeles d 1508, h 1512 y l 1516, se realiza un proceso de interpolación vertical 1D. 40

En subpíxeles e 1509, f 1.510, g 1511, i 1513, J 1514, k 1515, m 1517, n 1518, y o 1519 aparte de los subpíxeles un 1505, B 1506, C 1507, D 1508, h 1512 y l 1516, dado que tanto los componentes horizontales como los verticales 45 tienen valores fraccionarios, se realiza un proceso de interpolación bidimensional (2D) horizontal y vertical.

Como tal, un proceso de interpolación 1D usado en el presente documento se refiere a un proceso de realización de la interpolación solamente en una dirección de entre una dirección horizontal y una dirección vertical durante un proceso de compensación de movimiento, ya que uno cualquiera de un componente horizontal y un componente vertical de un vector de movimiento tiene un valor fraccionario en lugar de un valor entero. Un proceso de 50 interpolación 2D utilizado aquí se refiere a un proceso de interpolación en direcciones horizontal y vertical, ya que tanto un componente horizontal como un componente vertical de un vector de movimiento tienen valores fraccionarios en lugar de valores enteros.

La figura 16 es un diagrama para explicar un proceso de interpolación de un marco de referencia, de acuerdo con una realización de la presente invención. En la figura 16, las letras mayúsculas de la A a la Z indican píxeles enteros, y las letras en minúscula de la a a la z indican 1/2 o 1/4 de subpíxeles. 55

La unidad 1410 de interpolación aplica un filtro de interpolación que tiene un número de coeficientes del filtro predeterminado a píxeles enteros para generar un subpíxel necesario para la estimación de movimiento y la compensación de movimiento. Por ejemplo, con referencia a la figura 16, la unidad 1410 de interpolación puede

generar subpíxeles al aplicar un filtro de interpolación de N-(N es un número entero) a las señales de píxeles enteros de la A a la Z en una dirección horizontal o vertical. Se puede usar un filtro de interpolación de 8 derivaciones para interpolar un componente de luminancia, y se puede usar un filtro de interpolación de 4 derivaciones para interpolar un componente de croma. Un número de derivación de filtro N no se limita al mismo, y puede cambiarse.

5 Un proceso de subpíxeles de interpolación se explicará en más detalle con referencia a la figura 16.

Para interpolar a, b, c, d, h y n que son subpíxeles ubicados en la misma fila o columna que los píxeles enteros, se pueden generar los subpíxeles a, b, c, d, h, y n mediante la lectura de N píxeles enteros según un número de derivaciones de filtro en cualquier dirección entre direcciones horizontales y verticales, y aplicando un filtro de interpolación a los píxeles enteros N de lectura como se describió anteriormente. Por ejemplo, el 1/2 subpíxel b 1610 puede generarse aplicando un filtro de interpolación de 8 derivaciones que tiene un coeficiente de derivación de {p1, p2, p3, p4, p5, p6, p7, p8} a G, H, I, J, K, L, M y N, que son 8 píxeles enteros horizontalmente adyacentes al subpíxel b 1610 utilizando la Ecuación $b = (G*p1+H*p2+I*p3+J*p4+K*p5+L*p6+M*p7+N*p8)$. Del mismo modo, el 1/2 subpíxel h 1620 se puede generar aplicando un filtro de interpolación de 8 derivaciones que tiene un coeficiente de derivación de {p1, p2, p3, p4, p5, p6, p7, p8} a A, C, E, J, R, W, Y, AA que son 8 píxeles enteros verticalmente adyacentes al subpíxel h 1620 utilizando la Ecuación $h=(A*p1 +C*p2+E*p3+J*p4+R*p5+W*p6+Y*p7+AA*p8)$. Un número de derivación y un coeficiente de derivación se pueden cambiar como se describe anteriormente.

1/4 subpíxeles a, c, d, y n se encuentra en la misma fila o columna como píxeles enteros pueden ser generados mediante el uso de un valor medio entre el medio subpíxel b 1610 y el medio subpíxel h 1620 que se generan mediante el proceso de interpolación anterior y un píxel entero. Por ejemplo, el 1/4 subpíxel a 1640 se puede generar utilizando la Ecuación $a=(J+b)/2$. Como tal, para interpolar el 1/2 subpíxel b o no los 1/4 subpíxeles a, c, d, y n ubicados en la misma fila o columna como píxeles enteros, es necesario un proceso de lectura de N (N es un número de derivación del filtro) píxeles enteros vecinos.

A continuación, se explicará un proceso de interpolación e, f, g, i, j, k, p, q, y r que son subpíxeles que no se encuentran en la misma fila o columna como píxeles enteros.

25 El subpíxel 1/2 j 1630 puede generarse aplicando un filtro de interpolación en una dirección arbitraria entre una dirección horizontal y una dirección vertical. Por ejemplo, el 1/2 subpíxel j 1630 puede generarse aplicando un filtro de interpolación de 8 derivaciones a los 1/2 subpíxeles vecinos nn, mm, ll, h, hh, ii, jj y kk que son horizontalmente adyacentes, o pueden se genera aplicando un filtro de interpolación de 8 derivaciones a los 1/2 subpíxeles adyacentes aa, bb, cc, b, dd, ee, ff y gg que son verticalmente adyacentes. Aquí, los 1/2 subpíxeles nn, mm, ll, h, hh, ii, jj y kk que son horizontalmente adyacentes se pueden generar de una manera similar a la utilizada para obtener el 1/2 subpíxel h 1620, y los 1/2 subpíxeles aa, bb, cc, b, dd, ee, ff y gg que son verticalmente adyacentes se pueden generar de una manera similar a la utilizada para obtener el 1/2 subpíxel b 1610.

1/4 subpíxeles e, f, g, i, k, p, q, y r que no están ubicados en la misma fila o columna que los píxeles enteros pueden generarse utilizando un filtro de valor promedio con un píxel o 1/2 subpíxel entero adyacente vecino. Por ejemplo, el 1/4 subpíxel f 1650 se puede obtener usando un valor promedio entre el 1/2 subpíxel b 1610 y el 1/2 subpíxel j 1630, y el 1/4 subpíxel r 1660 se puede generar usando un filtro promedio diagonal, como la Ecuación $r=(hh+dd)/2$.

Las figuras 17A a 17C son diagramas que ilustran datos de un marco de referencia necesario para un proceso de compensación de movimiento, de acuerdo con una realización de la presente invención.

40 Con referencia a la figura 17A, cuando un estimador 420 de movimiento determina un vector de movimiento MV_1 de una unidad 1711 de predicción actual con un tamaño de $L \times M$ (L y M son enteros), el compensador 425 de movimiento genera un valor de predicción de la unidad 1711 de predicción actual leyendo una región de correspondencia indicada por el vector de movimiento MV_1 en un marco 1720 de referencia desde una memoria. Se supone que un componente horizontal del vector de movimiento MV_1 es MV_{x1} , un componente vertical del vector de movimiento MV_1 es MV_{y1} , el componente horizontal MV_{x1} tiene un valor entero y el componente vertical MV_{y1} tiene un valor fraccional. En este caso, para poder leer la región 1721 de correspondencia, se debe generar un subpíxel entre píxeles enteros que sean adyacentes horizontalmente. Si se supone que un número de derivación de un filtro de interpolación aplicado en una dirección vertical es N_L , ya que un número de píxeles enteros verticales necesarios para la interpolación vertical por píxel es N_L y una columna de la región 1721 de correspondencia incluye L píxeles, para leer la región 1721 de correspondencia, puede ser necesario leer $(L+N_L-1)*M$ píxeles enteros que incluyen los píxeles enteros verticales necesarios para la interpolación.

55 Del mismo modo, con referencia a la figura 17B, cuando el estimador 420 de movimiento determina un vector de movimiento MV_2 de una unidad 1711 de predicción actual, el compensador 425 de movimiento genera un valor de predicción de la unidad 1711 de predicción actual leyendo la región 1721 de correspondencia indicada por el vector de movimiento MV_2 en el marco 1720 de referencia desde la memoria. Se supone que un componente horizontal del vector de movimiento MV_2 es MV_{x2} , un componente vertical del vector de movimiento MV_2 es MV_{y2} , el componente horizontal MV_{x2} tiene un valor fraccional y el componente vertical MV_{y2} tiene un valor entero. En este caso, para leer la región 1721 de correspondencia, puede ser necesario generar un subpíxel entre píxeles enteros adyacentes horizontalmente. Si se supone que un número de derivación de un filtro de interpolación aplicado en una dirección

horizontal es N_M , un número de píxeles enteros horizontales necesarios para la interpolación horizontal por píxel es N_M y una fila de la región 1721 de correspondencia incluye M píxeles, para leer la región 1721 de correspondencia, es posible que se deban leer los píxeles enteros $L*(M+N_M-1)$ que incluyen los píxeles enteros horizontales necesarios para la interpolación.

- 5 Con referencia a la figura 17C, cuando el estimador 420 de movimiento determina un vector de movimiento MV_3 de una unidad 1711 de predicción actual, el compensador 425 de movimiento genera un valor de predicción de la unidad 1711 de predicción actual leyendo la región 1721 de correspondencia indicada por el vector de movimiento MV_3 en el marco 1720 de referencia desde la memoria. Se supone que un componente horizontal del vector de movimiento MV_3 es MVx_3 , un componente vertical del vector de movimiento MV_3 es MVy_3 , y tanto el componente horizontal MVx_3 como el componente vertical MVy_3 tienen valores fraccionarios. En este caso, para leer la región 10 1721 de correspondencia, puede ser necesario generar un subpíxel entre píxeles enteros realizando una interpolación en una dirección horizontal y una dirección vertical. Como se describió anteriormente, cuando se supone que un número de derivación de un filtro de interpolación aplicado en una dirección vertical es N_L y un número de derivación de un filtro de interpolación aplicado en una dirección horizontal es N_M , para leer la región 15 1721 de correspondencia, puede ser necesario leer $(L+N_L-1)*(M+N_M-1)$ incluyendo los píxeles enteros horizontales y verticales necesarios para la interpolación.

Como tal, un número de píxeles enteros de un marco de referencia que son leídos para interpolar subpíxeles se determina de acuerdo con un componente horizontal y un componente vertical de un vector de movimiento de una unidad de predicción actual.

- 20 Desde un punto de vista de acceso de memoria, se puede obtener un número de píxeles enteros de un cuadro de referencia necesario por píxel de una unidad de predicción actual durante la compensación de movimiento utilizando la Ecuación Número = (número de píxeles de un cuadro de referencia necesario para compensación de movimiento de una unidad de predicción actual)/(número total de píxeles incluidos en la unidad de predicción actual). Como se describió anteriormente, se supone que un número de derivación de un filtro de interpolación aplicado en una 25 dirección vertical es N_L , un número de toma de un filtro de interpolación aplicado en una dirección horizontal es N_M , un componente horizontal de un vector de movimiento de una unidad de predicción actual es MVx , y un componente vertical del vector de movimiento de la unidad de predicción actual es MVy . Cuando solo el componente horizontal MVx del vector de movimiento tiene un valor fraccionario en lugar de un valor entero, ya que de un número de píxeles de un cuadro de referencia necesario para la compensación de movimiento de la unidad de predicción actual que tiene un tamaño de LxM es $(L+N_L-1)*M$, un número Num de píxeles enteros del cuadro de referencia necesario por píxel de la unidad de predicción actual durante la compensación de movimiento es $Num=(L+N_L-1)*M/(L*M)$.
- 30

Del mismo modo, cuando solo el componente MVy vertical del vector de movimiento tiene un valor fraccionario en lugar de un valor entero, ya que un número de píxeles de la imagen de referencia necesaria para la compensación de movimiento de la unidad de predicción de corriente que tiene el tamaño de LxM es $L*(M+N_M-1)$, un número de 35 píxeles enteros del cuadro de referencia necesario por píxel de la unidad de predicción actual durante la compensación de movimiento es $Num=L*(M+N_M-1)/(L*M)$.

Cuando el componente horizontal MVx y el componente vertical MVy del vector de movimiento tienen valores fraccionarios, en lugar de valores enteros, ya que un número de píxeles del cuadro de referencia necesario para la compensación de movimiento de la unidad de predicción actual que tiene el tamaño de LxM es $(L+N_L-1)*(M+N_M-1)$, un número de píxeles enteros del cuadro de referencia necesario por píxel de la unidad de predicción actual durante 40 la compensación de movimiento es $Num=(L+N_L-1)*(M+N_M-1)/(L*M)$.

Como tal, un número Num de píxeles enteros de un cuadro de referencia por píxel de una unidad de predicción actual durante la compensación de movimiento corresponde a la del caso de la predicción unidireccional. En el caso de predicción bidireccional, dado que se utilizan dos marcos de referencia, una cantidad de píxeles enteros 45 necesarios es $2*Num$.

Como se describirá más adelante, un número de píxeles de un cuadro de referencia necesario para la compensación de movimiento por un píxel de una unidad de predicción actual se clasifica de acuerdo con el tamaño de la unidad de predicción actual y un modo de estimación de movimiento unidireccional y un modo de estimación de movimiento bidireccional aplicado a la unidad de predicción actual, y el acceso a un dispositivo de memoria predeterminado que almacena el marco de referencia se limita aplicando tanto el modo de estimación de movimiento unidireccional como el modo de estimación de movimiento bidireccional a la unidad de predicción actual cuando el número de píxeles del marco de referencia necesario para la compensación de movimiento por un píxel de la unidad de predicción actual es igual o menor que un valor de umbral predeterminado y aplica solo el modo de estimación de movimiento unidireccional a la unidad de predicción actual cuando el número de píxeles del marco de referencia necesario para la compensación de movimiento por un píxel de la unidad de predicción actual es mayor que el valor umbral predeterminado. Es decir, durante la compensación de movimiento, cuando un número de píxeles de un cuadro de referencia necesario por un píxel es demasiado grande, el acceso a la memoria para leer los datos del cuadro de referencia en un proceso de compensación de movimiento se limita al limitar el proceso de compensación de movimiento.

50

55

Las figuras 18A a 18G son tablas que ilustran un número de píxeles de un cuadro de referencia necesario para la compensación de movimiento por un píxel de una unidad de predicción actual de acuerdo con el tamaño de una unidad de predicción, un modo de predicción y si los componentes horizontales y verticales de un vector de movimiento de la unidad de predicción tienen valores enteros. En las figuras 18A a 18G, I indica un caso donde tanto un componente horizontal como un componente vertical de un vector de movimiento tienen valores enteros, ID indica un caso donde cualquiera de un componente horizontal y un componente vertical de un vector de movimiento tienen un valor fraccionario en lugar de un valor entero, y 2D indica un caso en el que tanto un componente horizontal como un componente vertical de un vector de movimiento tienen valores fraccionarios en lugar de valores enteros. Además, en el caso de predicción bidireccional que incluye la predicción L0 que se refiere a una primera lista de píxeles de referencia y la predicción L1 que se refiere a una segunda lista de imágenes de referencia, I-I indica un caso en el que tanto un componente horizontal como un componente vertical de cada uno de los dos vectores de movimiento utilizados para la predicción L0 y la predicción L1 tienen valores enteros, ID-ID indica un caso en el que dos vectores de movimiento utilizados para la predicción L0 y la predicción L1 tienen un componente horizontal o vertical que tiene un valor fraccionario en lugar de un valor entero, y 2D -2D indica un caso donde un componente horizontal y un valor vertical de dos vectores de movimiento utilizados para la predicción L0 y la predicción L1 tienen valores fraccionarios en lugar de valores enteros. Además, I-1D indica un caso en el que tanto un componente horizontal como un componente vertical de un vector de movimiento de entre dos vectores de movimiento utilizados para la predicción L0 y la predicción L1 tienen valores enteros, y cualquiera de un componente horizontal y un componente vertical del vector de movimiento restante tiene un valor fraccionario en lugar de un valor entero. Además, I-2D indica un caso en el que tanto un componente horizontal como un componente vertical de un vector de movimiento de entre dos vectores de movimiento utilizados para la predicción L0 y la predicción L1 tienen valores enteros, y tanto un componente horizontal como un componente vertical del vector de movimiento restante tiene valores fraccionarios en lugar de valores enteros. Además, 1D-2D indica un caso en el que cualquiera de un componente horizontal y un componente vertical de un vector de movimiento de entre dos vectores de movimiento utilizados para la predicción L0 y la predicción L1 tiene un valor fraccionario en lugar de un valor entero, y ambos de un componente horizontal y un componente vertical del vector de movimiento restante tienen valores fraccionarios en lugar de valores enteros.

En las figuras 18A a 18G, un filtro de interpolación que tiene el mismo número de derivación N se usa como un filtro de interpolación en una dirección horizontal y una dirección vertical, y el número de derivación N tiene un valor que va de 8 a 2.

Con referencia a la figura 18A, cuando se usa un filtro de interpolación de 8 derivaciones, cualquiera de un componente horizontal y un componente vertical de un vector de movimiento de una unidad de predicción que tiene un tamaño de 4x8 tiene un valor fraccionario en lugar de un valor entero, y la estimación de movimiento unidireccional se realiza en la unidad de predicción que tiene el tamaño de 4x8, ya que un número total de píxeles enteros necesarios para la compensación de movimiento de la unidad de predicción que tiene el tamaño de 4x8 es $(4+8-1)*8$, se puede obtener 2,75 dividiendo $(4+8-1)*8$ por 32 píxeles incluidos en la unidad de predicción que tiene el tamaño de $4x8((4+8-1)*8/(4x8) = 2,75)$. A través de este cálculo, como se muestra en la figura 18A, cuando un tamaño de una unidad de predicción (bloque) es 4x8, solo uno de un componente horizontal y un componente vertical de un vector de movimiento tiene un valor fraccionario en lugar de un valor entero (en el caso 1 D), y se realiza un modo predicción unidireccional, se puede obtener 2,75, que es un número de píxeles enteros de un marco de referencia necesario por un píxel. Del mismo modo, las tablas de las figuras 18A a 18G se pueden obtener utilizando uno cualquiera de $Num = (L+N_L-1) * M/(L*M)$, $Num = L*(M+N_M-1)/(L*M)$, y $Num = (L+N_L-1)*(M+N_M-1)/(L*M)$ que calcula un número de píxeles enteros de un cuadro de referencia necesario por un píxel de una unidad de predicción actual durante la compensación de movimiento de acuerdo con si un componente horizontal y un componente vertical de un vector de movimiento son valores enteros o valores fraccionarios.

Analizando las tablas de las figuras 18A a 18G, se encuentra que una cantidad de píxeles enteros de un marco de referencia necesario por un píxel de una unidad de predicción actual durante la compensación de movimiento disminuye a medida que aumenta el tamaño de una unidad de predicción. Por ejemplo, con referencia a la figura 18A, en un modo de predicción bidireccional (predicción doble), se encuentra que cuando un componente horizontal y un componente vertical de un vector de movimiento tienen valores fraccionarios en lugar de valores enteros (en el caso 2D-2D) y un tamaño de predicción la unidad aumenta de 2x2 a 64x64, un número de píxeles enteros de un marco de referencia necesario por píxel disminuye de 40,50 a 2,46. Como tal, cuando la estimación de movimiento y la compensación de movimiento se realizan en una unidad de predicción que tiene un tamaño demasiado pequeño, ya que se incrementa un número de píxeles enteros de un marco de referencia necesario por un píxel, es ineficiente desde el punto de vista de acceso a la memoria. Por consiguiente, el acceso a la memoria para leer los datos del cuadro de referencia se puede limitar a un ancho de banda predeterminado al limitar el modo de estimación de movimiento utilizado cuando un tamaño de una unidad de predicción tiene un tamaño igual o menor que un valor umbral predeterminado utilizando tanto un modo de estimación de movimiento unidireccional como un modo de estimación de movimiento bidireccional cuando el tamaño de la unidad de predicción es mayor que el tamaño predeterminado y utilizando solo un modo de estimación de movimiento unidireccional cuando el tamaño de la unidad de predicción es igual o menor que el tamaño predeterminado.

Las figuras 19 a 20 son tablas que ilustran en orden descendente comenzando desde un valor alto un número de píxeles de un cuadro de referencia necesario para la compensación de movimiento por un píxel de una unidad de

predicción actual de acuerdo con el tamaño de una unidad de predicción de un componente de luminancia, un tamaño de unidad de predicción de un componente de croma, un modo de predicción y si un componente horizontal y un componente vertical de un vector de movimiento de la unidad de predicción tienen valores enteros. En la figura 19, se supone que un filtro de interpolación de 8 derivaciones se utiliza como filtro de interpolación para el componente de luminancia, un filtro de interpolación de 4 derivaciones se usa como componente de interpolación para el componente de croma, un tamaño de una unidad de predicción mínima de la luminancia el componente es 4x8, y el tamaño de una unidad de predicción mínima del componente de croma es 2x4. Cuando se compara con la figura 19, en la figura 20, un tamaño de una unidad de predicción mínima del componente de luminancia es 4x4 y un tamaño de una unidad de predicción mínima del componente de croma es 2x2.

Con referencia a la figura 19, dado que una posición superior en la tabla es más ineficiente desde el punto de vista del acceso a la memoria, el estimador 420 de movimiento y el compensador 425 de movimiento pueden limitar el acceso a la memoria en un proceso de compensación de movimiento limitando o cambiando un modo de estimación de movimiento en el caso de una posición superior. Por ejemplo, el estimador 420 de movimiento y el compensador 425 de movimiento pueden realizar solo un modo de estimación de movimiento unidireccional que incluye predicción L0 y predicción L1 cuando un tamaño de una unidad de predicción del componente de luminancia es 4x8 u 8x4, que es una unidad de predicción mínima, y puede realizar un modo de estimación de movimiento bidireccional que no sea un modo de estimación de movimiento unidireccional solo cuando el tamaño de una unidad de predicción del componente de luminancia sea mayor que 4x8 u 8x4, que es un modo de estimación de movimiento mínimo. Del mismo modo, el estimador 420 de movimiento y el compensador 425 de movimiento pueden realizar solo un modo de estimación de movimiento unidireccional que incluye predicción L0 y predicción L1 cuando un tamaño de una unidad de predicción del componente de croma es 2x4 o 4x2, que es una unidad de predicción mínima, y puede realizar un modo de estimación de movimiento bidireccional que no sea un modo de estimación de movimiento unidireccional solo cuando el tamaño de una unidad de predicción del componente de croma es mayor que 2x4 o 4x2, que es una unidad de predicción mínima.

Del mismo modo, con referencia a la figura 19, el estimador 420 de movimiento y el compensador 425 de movimiento pueden limitar el acceso a la memoria en un proceso de compensación de movimiento limitando o cambiando un modo de estimación de movimiento en el caso de una posición superior.

Como se ha descrito anteriormente, el estimador 420 de movimiento y el compensador 425 de movimiento limitan un modo de estimación de movimiento, cuando un número de píxeles de un cuadro de referencia necesario para la compensación de movimiento por un píxel de una unidad de predicción actual es igual o mayor que un valor umbral predeterminado desde un punto de vista de acceso a la memoria. Es decir, el estimador 420 de movimiento y el compensador 425 de movimiento determinan que tanto un modo de estimación de movimiento unidireccional como un modo de estimación de movimiento bidireccional se usan cuando un tamaño de una unidad de predicción actual es mayor que un tamaño predeterminado, y determina que solo un movimiento unidireccional el modo de estimación se utiliza cuando un tamaño de una unidad de predicción actual es igual o menor que un tamaño predeterminado. El estimador 420 de movimiento y el compensador 425 de movimiento pueden determinar que un modo de estimación de movimiento no se usa cuando el tamaño de una unidad de predicción actual es demasiado pequeño. Por ejemplo, el estimador 420 de movimiento y el compensador 425 de movimiento pueden establecer un tamaño de una unidad de predicción mínima en 4x8 u 8x4 para una unidad de predicción del componente de luminancia, y pueden no realizar una estimación de movimiento cuando una unidad de predicción actual del componente de luminancia es menor que 4x8 u 8x4. Además, el estimador 420 de movimiento y el compensador 425 de movimiento pueden establecer un tamaño de una unidad de predicción mínima en 2x4 o 4x2 para una unidad de predicción del componente de croma, y pueden no realizar la estimación de movimiento cuando una unidad de predicción actual del componente de croma es menor que 2x4 o 4x2. El estimador 420 de movimiento y el compensador 425 de movimiento pueden clasificar un modo de estimación de movimiento que se realizará incluso cuando un tamaño es igual o mayor que un tamaño de una unidad de predicción mínima, y puede determinar que solo se utiliza un modo de estimación de movimiento unidireccional cuando el tamaño de una unidad de predicción actual más ineficiente es un tamaño de una unidad de predicción mínima y determina que tanto el modo de estimación de movimiento unidireccional como el modo de estimación de movimiento bidireccional se usan solo cuando el tamaño de una unidad de predicción actual no es el tamaño de una predicción mínima unidad, es decir, un tamaño de una unidad de predicción actual es mayor que un tamaño de una unidad de predicción mínima.

Como tal, el estimador 420 de movimiento y el compensador 425 de movimiento determinan si un modo de estimación de movimiento unidireccional y un modo de estimación de movimiento bidireccional aplicado a la estimación de movimiento y a la compensación de movimiento de una unidad de predicción actual se utilizan basadas en un tamaño de la predicción actual y realizan la estimación de movimiento y la compensación de movimiento para la unidad de predicción actual utilizando al menos un modo de estimación de movimiento entre un modo de estimación de movimiento unidireccional y un modo de estimación de movimiento bidireccional según si se usa el modo de estimación de movimiento unidireccional y el modo de estimación de movimiento bidireccional. Un controlador no mostrado en la figura 4 determina un modo de estimación de movimiento óptimo para ser aplicado a una unidad de predicción actual basándose en un coste de codificación, por ejemplo, un coste de RD, de la unidad de predicción actual obtenida a través de la estimación de movimiento y la compensación de movimiento. Un codificador 460 de entropía codifica determinada información del modo de estimación de movimiento en función del tamaño de la unidad de predicción actual.

Por ejemplo, el codificador 460 de entropía determina una sintaxis `inter_pred_idc` que indica un modo de estimación de movimiento final para la unidad de predicción actual, basado en un tamaño de la unidad de predicción actual, basado en la Tabla 2, y codifica la sintaxis `inter_pred_idc`.

Tabla 2

inter_pred_idc	Modo de estimación de movimiento correspondiente	
	Cuando un tamaño de una unidad de predicción es mayor que un tamaño de una unidad de predicción mínima	Cuando un tamaño de una unidad de predicción es igual a un tamaño de la unidad de predicción mínima
0	Predicción L0	Predicción L0
1	Predicción L1	Predicción L1
2	Predicción bidireccional	No permitido

5

Con referencia a la Tabla 2, el codificador 460 de entropía codifica por entropía `inter_pred_idc`, que indica un modo de estimación de movimiento final para una unidad de predicción actual basada en un tamaño de la unidad de predicción actual. En detalle, cuando la unidad de predicción actual tiene un tamaño mayor que 4x8 u 8x4, que es una unidad de predicción mínima como unidad de predicción de un componente de luminancia, se pueden usar tanto la predicción unidireccional (predicción L0 y predicción L1) como la predicción bidireccional. Cuando la predicción L1 se determina como un modo de estimación de movimiento final de una unidad de predicción actual, el codificador 460 de entropía codifica por entropía 1 como `inter_pred_idc`. El decodificador 520 de entropía en el lado del decodificador puede determinar un modo de estimación de movimiento usado para la unidad de predicción actual de acuerdo con la Tabla 2, basado en `inter_pred_idc` e información de tamaño de la unidad de predicción actual obtenida de una corriente de bits. Alternativamente, cuando una unidad de predicción actual es una unidad de predicción mínima de 4x8 u 8x4 que es una unidad de predicción mínima como unidad de predicción de un componente de luminancia, solo se puede usar la predicción unidireccional (predicción L0 y predicción L1). Cuando la predicción L1 se determina como un modo de estimación de movimiento final de una unidad de predicción actual, el codificador 460 de entropía codifica por entropía 1 como `inter_pred_idc`.

10

15

20

25

La compensación de movimiento se realiza mediante la limitación de si un modo de estimación de movimiento unidireccional y un modo de estimación de movimiento bidireccional se utilizan sobre la base de un tamaño de una unidad de predicción a partir de un punto de acceso a la memoria de la vista. Alternativamente, se puede reducir el número de derivaciones de un filtro de interpolación utilizado para la estimación de movimiento. Por ejemplo, se puede usar un filtro de interpolación de 4 derivaciones en lugar de un filtro de interpolación de 8 derivaciones usado para interpolar un componente de luminancia, y un filtro de interpolación de 2 derivaciones en lugar de un filtro de interpolación de 4 derivaciones usado para interpolar un componente de croma.

Alternativamente, un proceso de lectura de los datos de marcos de referencia se puede reducir mediante la sustitución de un proceso de interpolación 2D con un proceso de interpolación 1D.

30

Las figuras 21 a 23 son diagramas de referencia para explicar un proceso de interpolación para reducir un proceso de lectura de los datos del marco de referencia, de acuerdo con una realización de la presente invención.

35

Con referencia a las figuras 21 a 23, como los subpíxeles e, f, g, i, j, k, m, n, y o que no están ubicados en la misma fila o columna que los píxeles enteros pueden no obtenerse usando solo un filtro de interpolación 1D y son obtenidos a través de un proceso de interpolación 2D, se incrementa el acceso a la memoria. En consecuencia, de acuerdo con una realización de la presente invención, los valores de los subpíxeles e, f, g, i, j, k, m, n, y o pueden obtenerse utilizando un subpíxel vecino o un píxel entero obtenido a través de un filtro de interpolación 1D sin realizar un proceso de interpolación 2D en los subpíxeles e, f, g, i, j, k, m, n y o.

Con referencia a las figuras 21 a 23, el estimador 420 de movimiento y el compensador 425 de movimiento pueden usar subpíxeles adyacentes o píxeles enteros marcados con flechas para generar valores de los subpíxeles e, f, g, i, j, k, m, n, y o.

40

45

Las figuras 24 y 25 son tablas que ilustran en orden descendente comenzando desde un valor alto un número de píxeles de un cuadro de referencia necesario para la compensación de movimiento por un píxel de una unidad de predicción actual de acuerdo con el tamaño de una unidad de predicción de un componente de luminancia, un tamaño de una unidad de predicción de un componente de croma, un modo de estimación de movimiento y si los componentes horizontales y verticales de un vector de movimiento de la unidad de predicción tienen valores enteros, cuando la interpolación se realiza utilizando un subpíxel vecino o un píxel entero que se puede obtener utilizando un filtro de interpolación 1D en lugar de un proceso de interpolación 2D o reduciendo el número de derivaciones de un filtro de interpolación de acuerdo con una realización de la presente invención. La figura 24 ilustra un caso donde un

tamaño de una unidad de predicción mínima del componente de luminancia es 4x8 y un tamaño de una unidad de predicción mínima del componente de croma es 2x4. La figura 25 ilustra un caso donde un tamaño de una unidad de predicción mínima del componente de luminancia es 4x4 y un tamaño de una unidad de predicción mínima del componente de croma es 2x2. Además, en las figuras 24 y 25, A indica un caso donde se reduce el número de derivaciones de un filtro de interpolación, y B indica un caso donde se realiza la interpolación utilizando un subpíxel o un píxel entero que puede obtenerse utilizando un filtro de interpolación 1D en lugar de un proceso de interpolación 2D.

Con referencia a las figuras 24 y 25, se encuentra que cuando la interpolación se realiza utilizando un subpíxel vecino o un píxel entero que puede obtenerse utilizando un filtro de interpolación 1D en lugar de un proceso de interpolación 2D o reduciendo el número de pulsación de un filtro de interpolación, un número de derivación de los píxeles de un cuadro de referencia necesarios para la compensación de movimiento por un píxel de una unidad de predicción actual pueden ser menores que los de un caso original.

La figura 26 es un diagrama de flujo que ilustra un procedimiento de codificación de vídeo, de acuerdo con una realización de la presente invención.

Con referencia a la figura 26, en la operación 2610, el estimador 420 de movimiento y el compensador 425 de movimiento determinan si se utiliza un modo de estimación de movimiento unidireccional y un modo de estimación de movimiento bidireccional aplicado a la estimación de movimiento y a la compensación de movimiento de una unidad de predicción actual en función del tamaño de la unidad de predicción actual a codificar. Como se ha descrito anteriormente, el estimador 420 de movimiento y el compensador 425 de movimiento limitan un modo de estimación de movimiento, cuando un número de píxeles de un cuadro de referencia necesario para la compensación de movimiento por un píxel de la unidad de predicción actual es igual o mayor que un valor umbral predeterminado desde un punto de vista de acceso a la memoria. Es decir, el estimador 420 de movimiento y el compensador 425 de movimiento pueden determinar que tanto un modo de estimación de movimiento unidireccional como un modo de estimación de movimiento bidireccional se usan cuando un tamaño de la unidad de predicción actual es mayor que un tamaño predeterminado, y pueden determinar que solo un movimiento unidireccional el modo de estimación se utiliza cuando un tamaño de la unidad de predicción actual es igual o menor que un tamaño predeterminado.

En la operación 2620, el estimador 420 de movimiento y el compensador 425 de movimiento realizan la estimación de movimiento y la compensación de movimiento en la unidad de predicción actual aplicando el modo de estimación de al menos un movimiento de entre el modo de estimación de movimiento unidireccional y el modo de estimación de movimiento bidireccional según si se utilizan el modo de estimación de movimiento unidireccional y el modo de estimación de movimiento bidireccional.

En la operación 2630, un controlador determina un modo de estimación de movimiento de la unidad de predicción actual, basado en un coste de codificación de la unidad de predicción actual obtenida a través de la estimación de movimiento y de la compensación de movimiento. En la operación 2640, un codificador 460 de entropía codifica por entropía determinada información del modo de estimación de movimiento en función del tamaño de la unidad de predicción actual. Como se describió anteriormente, el codificador 460 de entropía determina una sintaxis `inter_pred_idc` que indica un modo de estimación de movimiento final para la unidad de predicción actual basada en la Tabla 2, y la entropía codifica la sintaxis `inter_pred_idc`.

La figura 27 es un diagrama de bloques que ilustra un aparato 2700 de compensación de movimiento, de acuerdo con una realización de la presente invención. El aparato 2700 de compensación de movimiento de la figura 27 corresponde al compensador 560 de movimiento del decodificador 500 de imágenes de la figura 5.

Con referencia a las figuras 5 y 27, el decodificador 520 de entropía obtiene información del modo de estimación de movimiento que indica un modo de estimación de movimiento aplicado a una unidad de predicción actual y la información de tamaño de la unidad de predicción actual para ser decodificada a partir de una corriente de bits.

Como se describió con referencia a las figuras 1 a 13, información que indica el tamaño de una unidad de codificación máxima obtenida al dividir un cuadro que constituye un vídeo de una corriente de bits, una profundidad que indica varias veces que la unidad de codificación máxima se divide espacialmente y una estructura de unidades de transformación que tiene una estructura jerárquica y se puede obtener una unidad de predicción utilizada para la codificación de predicción de unidades de codificación que se clasifican jerárquicamente de acuerdo con las profundidades, y para cada al menos una unidad de codificación máxima obtenida dividiendo un cuadro en una unidad de codificación de un tamaño máximo basado en la información obtenida, de entre las unidades de codificación que se clasifican jerárquicamente según las profundidades de acuerdo con la profundidad que indica el número de veces que la unidad de codificación máxima se divide espacialmente, se pueden determinar las unidades de codificación que tienen una estructura de árbol que incluye unidades de codificación de una profundidad codificada, se puede determinar una unidad de predicción para la codificación de predicción para cada unidad de codificación de la profundidad codificada, y se pueden determinar las unidades de transformación que tienen una estructura de árbol.

Además, como información de modo de estimación de movimiento, se puede usar la sintaxis `inter_pred_idc` que indica un modo de estimación de movimiento se aplica a la unidad de predicción actual, basado en un tamaño de la unidad de predicción actual.

5 Una unidad 2720 de ejecución de compensación de movimiento determina mediante el uso de la Tabla 2 un modo de estimación de movimiento que se aplica a la unidad de predicción actual de entre un modo de estimación de movimiento unidireccional y un modo de estimación de movimiento bidireccional basado en un tamaño de la unidad de predicción actual y la información del modo de estimación movimiento obtenida, es decir, la sintaxis `inter_pred_idc`.

10 Con referencia a la Tabla 2, cuando la unidad de predicción actual tiene un tamaño mayor que 4x8 u 8x4, que es una unidad de predicción mínima como unidad de predicción de un componente de luminancia y un valor de la sintaxis `inter_pred_idc` extraída de la corriente de bits es 1, la unidad 2720 de ejecución de compensación de movimiento determina un modo de estimación de movimiento de la unidad de predicción actual como predicción L1 unidireccional. Alternativamente, cuando la unidad de predicción actual es una unidad de predicción mínima de 4x8 u 8x4 que es una unidad de predicción mínima como unidad de predicción de un componente de luminancia, la
15 unidad 2720 de ejecución de compensación de movimiento determina que el modo de estimación de movimiento final de la unidad de predicción actual es predicción L1 cuando un valor de la sintaxis `inter_pred_idc` extraído de la corriente de bits es 1, y determina que no se aplica un modo de estimación de movimiento a la unidad de predicción actual cuando un valor de la sintaxis `inter_pred_idc` es 2.

20 La figura 28 es un diagrama de flujo que ilustra un procedimiento de decodificación de vídeo de acuerdo con una realización de la presente invención.

Con referencia a la figura 28, en la operación 2810, el decodificador 520 de entropía obtiene información del modo de estimación de movimiento que indica un modo de estimación de movimiento aplicado a una unidad de predicción actual y la información de tamaño de la unidad de predicción actual para ser decodificada a partir de una corriente de bits.

25 En la operación 2820, el compensador 560 de movimiento determina un modo de estimación de movimiento que se aplica a la unidad de predicción actual de entre un modo de estimación de movimiento unidireccional y un modo de estimación de movimiento bidireccional basado en un tamaño de la unidad de predicción actual y la información del modo de estimación de movimiento obtenida. Como se describió anteriormente, el compensador 560 de movimiento puede determinar un modo de estimación de movimiento que se usará para la unidad de predicción actual basándose en el tamaño de la unidad de predicción y a la sintaxis `inter_pred_idc` determinada de acuerdo con la
30 Tabla 2.

En la operación 2830, el compensador 560 de movimiento realiza la compensación de movimiento que obtiene un valor de predicción para la unidad de predicción actual mediante la lectura de datos de un cuadro de referencia de acuerdo con el modo de estimación de movimiento determinado.

35 Como un programa para ejecutar cada uno de los procedimientos de codificación y decodificación de vídeo descritos con referencia a las figuras 1 a 28 se almacena en un medio de grabación legible por ordenador, un sistema informático independiente puede realizar fácilmente las operaciones de acuerdo con cada realización almacenada en el medio de grabación legible por ordenador.

40 Una realización en la que un medio de grabación legible por ordenador que almacena un programa es un disco 260 se explicará en detalle.

La figura 29A ilustra una estructura física de un disco 260 que almacena un programa, de acuerdo con una realización de la presente invención. El disco 260, que es un medio de almacenamiento, puede ser un disco duro, un disco de memoria de solo lectura (CD-ROM), un disco Blu-ray o un disco versátil digital (DVD). El disco 260 incluye una pluralidad de pistas concéntricas T_f , cada una dividida en un número específico de sectores S_e en una dirección
45 circunferencial del disco 260. En una región específica del disco 260, un programa que ejecuta un procedimiento de predicción de vídeo de vista múltiple, un procedimiento de predicción de restauración de vídeo de vista múltiple, un procedimiento de codificación de vídeo de vista múltiple y un procedimiento de decodificación de vídeo de vista múltiple como se describe anteriormente puede ser asignado y almacenado.

50 Ahora se describirá un sistema informático realizado utilizando un medio de almacenamiento que almacena un programa para ejecutar un procedimiento de codificación de vídeo y un procedimiento de decodificación de vídeo como se describió anteriormente con referencia a la figura 29B.

La figura 29B ilustra una unidad de disco 263 que graba y lee un programa utilizando un disco 260. Un sistema 265 informático puede almacenar un programa que ejecuta al menos uno de un procedimiento de codificación de vídeo y un procedimiento de decodificación de vídeo de acuerdo con una realización de la presente invención, en un disco
55 260 a través de la unidad 263 de disco. Para ejecutar el programa almacenado en el disco 260 en el sistema 265 informático, el programa puede leerse desde el disco 260 y transmitirse al sistema 265 informático usando la unidad de disco 263.

El programa que ejecuta al menos uno de un procedimiento de codificación de vídeo y un procedimiento de decodificación de vídeo de acuerdo con una realización de la presente invención puede almacenarse no solo en el disco 260 que se ilustra en la figura 29A o 29B, sino también en una tarjeta de memoria, un casete ROM o una unidad de estado sólido (SSD).

5 Un sistema al que el procedimiento de codificación de vídeo y un procedimiento de descodificación de vídeo se ha descrito anteriormente se aplican se describirá a continuación.

La figura 30 ilustra una estructura completa de un sistema 1100 de suministro de contenido 1100 que proporciona un servicio de distribución de contenido. Una región de servicio de un sistema de comunicación se divide en células de tamaño predeterminado, y las estaciones 1170, 1180, 1190 y 1200 base inalámbricas están instaladas en estas
10 células, respectivamente.

El sistema 1100 de suministro de contenido incluye una pluralidad de dispositivos independientes. Por ejemplo, la pluralidad de dispositivos independientes, tal como un ordenador 1210, un asistente 1220 digital personal (PDA), una cámara 1230 de vídeo y un teléfono 1250 móvil, están conectados a Internet 1110 a través de un proveedor 1120 de servicios de Internet, una red 1140 de comunicación y las estaciones 1170, 1180, 1190 y 1200 base
15 inalámbricas.

Sin embargo, el sistema 1100 de suministro de contenido no se limita a lo ilustrado en la figura 30, y dispositivos pueden estar conectados selectivamente a los mismos. La pluralidad de dispositivos independientes puede conectarse directamente a la red 1140 de comunicación, en lugar de a través de las estaciones 1170, 1180, 1190 y 1200 base inalámbricas.

20 La cámara 1230 de vídeo es un dispositivo de imagen, por ejemplo, una cámara de vídeo digital, que es capaz de capturar imágenes de vídeo. El teléfono 1250 móvil puede emplear al menos un procedimiento de comunicación entre varios protocolos, por ejemplo, comunicaciones digitales personales (PDC), acceso múltiple por división de código (CDMA), acceso múltiple por división de código de banda ancha (W-CDMA), sistema global para comunicaciones móviles (GSM), y sistema de teléfono de mano personal (PHS).

25 La cámara 1230 de vídeo puede estar conectada a un servidor 1130 de transmisión a través de la estación 1190 base inalámbrica y la red 1140 de comunicación. El servidor 1130 de transmisión permite que el contenido recibido de un usuario a través de la cámara 1230 de vídeo se transmita mediante transmisión en tiempo real. El contenido recibido de la cámara 1230 de vídeo puede codificarse utilizando la cámara 1230 de vídeo o el servidor 1130 de transmisión. Los datos de vídeo capturados por la cámara 1230 de vídeo pueden transmitirse al servidor 1130 de
30 transmisión a través del ordenador 1210.

Los datos de vídeo captados por una cámara 1230 también pueden ser transmitidos al servidor 1130 de transmisión a través del ordenador 1210. La cámara 1260 es un dispositivo de imágenes capaz de capturar imágenes fijas e imágenes de vídeo, similar a una cámara digital. Los datos de vídeo capturados por la cámara 1260 pueden codificarse utilizando la cámara 1260 o el ordenador 1210. El software que codifica y decodifica el vídeo se puede
35 almacenar en un medio de grabación legible por ordenador, por ejemplo, un CD-ROM, un disquete, un disco duro, un SSD o una tarjeta de memoria, que puede ser accesible por el ordenador 1210.

Si los datos de vídeo son capturados por una cámara integrada en el teléfono 1250 móvil, los datos de vídeo se pueden recibir desde el teléfono 1250 móvil.

40 Los datos de vídeo también pueden ser codificados por un sistema a gran escala de circuitos integrados (LSI) instalado en la cámara 1230 de vídeo, el teléfono 1250 móvil, o la cámara 1260.

De acuerdo con una realización de la presente invención, el sistema 1100 de suministro de contenido puede codificar datos de contenido grabados por un usuario usando la cámara 1230 de vídeo, la cámara 1260, el teléfono 1250 móvil, u otro dispositivo de imagen, por ejemplo, el contenido grabado durante un concierto, y transmitir los datos de contenido codificado al servidor 1130 de transmisión. El servidor 1130 de transmisión puede transmitir los
45 datos de contenido codificados a otros clientes que solicitan los datos de contenido.

Los clientes son dispositivos capaces de decodificar los datos codificados de contenido, por ejemplo, el ordenador 1210, el PDA 1220, la cámara 1230 de vídeo, o el teléfono 1250 móvil. Por lo tanto, el sistema 1100 de suministro de contenido permite a los clientes recibir y reproducir los datos de contenido codificados. Además, el sistema 1100 de suministro de contenido permite a los clientes recibir los datos de contenido codificados y decodificar y reproducir los
50 datos de contenido codificados en tiempo real, permitiendo así la difusión personal.

Las operaciones de codificación y decodificación de la pluralidad de dispositivos independientes incluidos en el sistema 1100 de suministro de contenido pueden ser similares a las de un aparato de codificación de vídeo y un aparato de decodificación de vídeo de acuerdo con una realización de la presente invención.

55 El teléfono 1250 móvil incluido en el sistema 1100 de suministro de contenido, de acuerdo con una realización de la presente invención, se describirá ahora con mayor detalle con referencia a las figuras 31 y 32.

La figura 31 ilustra una estructura externa de un teléfono 1250 móvil a la que se aplica un procedimiento de codificación de vídeo y un procedimiento de decodificación de vídeo, de acuerdo con una realización de la presente invención. El teléfono 1250 móvil puede ser un teléfono inteligente, cuyas funciones no están limitadas y una gran parte de las funciones que se pueden cambiar o ampliar.

5 El teléfono 1250 móvil incluye una antena 1251 interna a través de la cual una señal de radiofrecuencia (RF) puede intercambiarse con la estación 1200 base inalámbrica de la figura 26, e incluye una pantalla 1252 de visualización para mostrar imágenes capturadas por una cámara 1253 o imágenes que se reciben a través de la antena 1251 y se decodifican, por ejemplo, una pantalla de cristal líquido (LCD) o una pantalla de diodos orgánicos emisores de luz (OLED). El teléfono 1251 inteligente incluye un panel 1254 de operación que incluye un botón de control y un panel táctil. Si la pantalla 1252 de visualización es una pantalla táctil, el panel 1254 de operación incluye además un panel de detección táctil de la pantalla 1252 de visualización. El teléfono 1251 inteligente incluye un altavoz para emitir voz y sonido u otro tipo de unidad de salida de sonido, y un micrófono 1255 para ingresar voz y sonido u otro tipo de unidad de entrada de sonido. El teléfono 1251 inteligente incluye además la cámara 1253, como una cámara con dispositivo de carga acoplada (CCD), para capturar vídeo e imágenes fijas. El teléfono 1251 inteligente puede incluir además un medio 1257 de almacenamiento para almacenar datos codificados/decodificados, por ejemplo, vídeo o imágenes fijas capturadas por la cámara 1253, recibidas por correo electrónico u obtenidas de acuerdo con otro procedimiento; y una ranura 1256 a través de la cual el medio 1257 de almacenamiento se carga en el teléfono 1250 móvil. El medio 1257 de almacenamiento puede ser una memoria flash, por ejemplo, una tarjeta digital segura (SD) o una memoria de solo lectura programable y borrrable eléctricamente (EEPROM) incluida en una carcasa de plástico.

La figura 32 ilustra una estructura interna del teléfono 1250 móvil, de acuerdo con una realización de la presente invención. Para controlar de manera sistémica partes del teléfono 1250 móvil que incluye la pantalla 1252 de visualización y el panel 1254 de operación, un circuito 1270 de fuente de alimentación, un controlador 1264 de entrada de operación, un codificador 1272 de imágenes, una interfaz 1263 de cámara, un controlador 1262 LCD, un decodificador 1269 de imágenes, un multiplexor/demultiplexor 1268, una unidad 1267 de grabación/lectura, una unidad 1266 de modulación/demodulación y un procesador 1265 de sonido están conectados a un controlador 1271 central a través de un bus 1273 de sincronización.

Si un usuario acciona un botón de encendido y cambia de un estado de 'apagado' a un estado de 'encendido', el circuito 1270 de suministro de energía suministra energía a todas las partes del teléfono 1250 móvil desde una batería, configurando así el teléfono 1250 móvil en modo de operación.

El controlador 1271 central incluye una unidad central de procesamiento (CPU), una ROM, y una memoria de acceso aleatorio (RAM).

Mientras que el teléfono 1250 móvil transmite datos de comunicación al exterior, se genera una señal digital en el teléfono 1250 móvil bajo el control del controlador central. Por ejemplo, el procesador 1265 de sonido puede generar una señal de sonido digital, el codificador 1272 de imágenes puede generar una señal de imagen digital, y los datos de texto de un mensaje pueden generarse a través del panel 1254 de operación y el controlador 1264 de entrada de operación. Cuando se envía una señal digital a la unidad 1266 de modulación/demodulación bajo el control del controlador 1271 central, la unidad 1266 de modulación/demodulación realiza la modulación en una banda de frecuencia de la señal digital, y un circuito 1261 de comunicación realiza la conversión de digital a analógico (DAC) y conversión de frecuencia en la señal digital modulada en banda de frecuencia. Una señal de transmisión emitida desde el circuito 1261 de comunicación puede transmitirse a una estación base de comunicación de voz o la estación 1200 base inalámbrica a través de la antena 1251.

Por ejemplo, cuando el teléfono 1250 móvil está en un modo de conversación, una señal de sonido obtenida a través del micrófono 1255 se transforma en una señal de sonido digital mediante el procesador 1265 de sonido, bajo el control del controlador 1271 central. La señal de sonido digital puede transformarse en una señal de transformación a través de la unidad 1266 de modulación/demodulación y el circuito 1261 de comunicación, y puede transmitirse a través de la antena 1251.

Cuando un mensaje de texto, por ejemplo, un correo electrónico, se transmite en un modo de comunicación de datos, los datos de texto del mensaje de texto se ingresan a través del panel 1254 de operación y se transmiten al controlador 1261 central a través del controlador 1264 de entrada de operación. Bajo el control del controlador 1261 central, los datos de texto se transforman en una señal de transmisión a través de la unidad 1266 de modulación/demodulación y el circuito 1261 de comunicación y se transmiten a la estación 1200 base inalámbrica a través de la antena 1251.

Para transmitir datos de imagen en el modo de comunicación de datos, datos de imagen capturados por la cámara 1253 se proporcionan al codificador 1272 de imágenes a través de la interfaz 1263 de la cámara. Los datos de la imagen capturada pueden visualizarse directamente en la pantalla 1252 de visualización a través de la interfaz 1263 de la cámara y el controlador 1262 LCD.

Una estructura del codificador 1272 de imágenes puede corresponder a la del aparato 100 de codificación de vídeo

descrito anteriormente. El codificador 1272 de imágenes puede transformar los datos de imagen recibidos desde la cámara 1253 en datos de imagen codificados de manera compresiva según un procedimiento de codificación de vídeo empleado por el aparato 100 de codificación de vídeo o el codificador 400 de imágenes descrito anteriormente, y luego enviar los datos de imagen codificados al multiplexor/demultiplexor 1268. Durante una operación de grabación de la cámara 1253, una señal de sonido obtenida por el micrófono 1255 del teléfono 1250 móvil puede transformarse en datos de sonido digital a través del procesador 1265 de sonido, y los datos de sonido digital pueden entregarse al multiplexor/demultiplexor 1268.

El multiplexor/demultiplexor 1268 multiplexa los datos de imagen codificados recibidos desde el codificador 1272 de imágenes, junto con los datos de sonido recibidos desde el procesador 1265 de sonido. Un resultado de la multiplexación de los datos puede transformarse en una señal de transmisión a través de la unidad 1266 de modulación/demodulación y el circuito 1261 de comunicación, y luego puede transmitirse a través de la antena 1251.

Mientras que el teléfono 1250 móvil recibe datos de comunicación desde el exterior, la recuperación de frecuencia y el ADC se realizan en una señal recibida a través de la antena 1251 para transformar la señal en una señal digital. La unidad 1266 de modulación/demodulación modula una banda de frecuencia de la señal digital. La señal digital modulada en banda de frecuencia se transmite al decodificador 1269 de vídeo, al procesador 1265 de sonido o al controlador 1262 LCD, según el tipo de señal digital.

En el modo de conversación, el teléfono 1250 móvil amplifica una señal recibida a través de la antena 1251, y obtiene una señal de sonido digital mediante la realización de la conversión de frecuencia y ADC en la señal amplificada. Una señal de sonido digital recibida se transforma en una señal de sonido analógica a través de la unidad 1266 de modulación/demodulación y el procesador 1265 de sonido, y la señal de sonido analógica se emite a través del altavoz 1258, bajo el control del controlador 1271 central.

Cuando está en el modo de comunicación de datos, se reciben datos de un archivo de vídeo al que se accede a través de un sitio web de Internet, una señal recibida desde la estación 1200 base inalámbrica a través de la antena 1251 se envía como datos multiplexados a través de la unidad 1266 de modulación/demodulación, y los datos multiplexados se transmiten al multiplexor/demultiplexor 1268.

Para decodificar los datos multiplexados recibidos a través de la antena 1251, el multiplexor/demultiplexor 1268 demultiplexa los datos multiplexados en una corriente de datos de vídeo codificados y una corriente de datos de audio codificada. A través del bus 1273 de sincronización, la corriente de datos de vídeo codificado y la corriente de datos de audio codificados se proporcionan al decodificador 1269 de vídeo y al procesador 1265 de sonido, respectivamente.

Una estructura del decodificador 1269 de imágenes puede corresponder a la del aparato 200 de decodificación de vídeo descrito anteriormente. El decodificador 1269 de imágenes puede decodificar los datos de vídeo codificados para obtener datos de vídeo restaurados y proporcionar los datos de vídeo restaurados a la pantalla 1252 de visualización a través del controlador 1262 LCD, de acuerdo con un procedimiento de decodificación de vídeo empleado por el aparato 200 de decodificación de vídeo o el decodificador 500 de imágenes descrito anteriormente.

Por lo tanto, los datos del archivo de vídeo al que se accede a través del sitio web de Internet pueden mostrarse en la pantalla 1252 de visualización. Al mismo tiempo, el procesador 1265 de sonido puede transformar los datos de audio en una señal de sonido analógica, y proporcionar la señal de sonido analógica al altavoz 1258. Por lo tanto, los datos de audio contenidos en el archivo de vídeo al que se accede a través del sitio web de Internet también se pueden reproducir a través del altavoz 1258.

El teléfono 1250 móvil u otro tipo de terminal de comunicación puede ser un terminal transceptor que incluye un aparato de codificación de vídeo y un aparato de decodificación de vídeo de acuerdo con una realización de la presente invención, puede ser un terminal transceptor que incluya solo el aparato de codificación de vídeo, o puede ser un terminal transceptor que incluye solo el aparato de decodificación de vídeo.

Un sistema de comunicación según la presente invención no se limita al sistema de comunicación descrito anteriormente con referencia a la figura 30. Por ejemplo, la figura 33 ilustra un sistema de transmisión digital que emplea un sistema de comunicación, de acuerdo con una realización de la presente invención. El sistema de transmisión digital de la figura 32 puede recibir una transmisión digital transmitida a través de un satélite o una red terrestre usando un aparato de codificación de vídeo y un aparato de decodificación de vídeo de acuerdo con una realización de la presente invención.

Específicamente, una estación 1289 de transmisión transmite una corriente de datos de vídeo a un satélite de comunicación o un satélite 1290 de transmisión utilizando ondas electrónicas. El satélite de transmisión 1290 transmite una señal de transmisión, y la señal de transmisión se transmite a un receptor de transmisión vía satélite a través de una antena 1286 doméstica. En cada casa, un receptor de TV 1281, un decodificador 1287 u otro dispositivo pueden decodificar y reproducir una corriente de vídeo codificado.

Cuando se incluye un aparato de decodificación de vídeo de acuerdo con una realización de la presente invención en un aparato 1283 de reproducción, el aparato 1283 de reproducción puede analizar y decodificar una corriente de

vídeo codificada grabado en un medio 1282 de almacenamiento, tal como un disco o una tarjeta de memoria para restaurar la señal de vídeo original. Por lo tanto, la señal de vídeo restaurada se puede reproducir, por ejemplo, en un monitor 1284.

5 En la antena 1286 para una transmisión satelital/terrestre o el decodificador 1287 conectado a una antena 1285 de cable para recibir programas de televisión por cable (TV), puede instalarse un aparato de decodificación de vídeo de acuerdo con una realización de la presente invención. Los datos emitidos desde el decodificador 1287 también pueden reproducirse en un monitor de TV 1288.

Como otro ejemplo, un aparato de decodificación de vídeo de acuerdo con una realización de la presente invención puede instalarse en el receptor de TV 1281 en lugar del decodificador 1287.

10 Un automóvil 1292 que incluye una antena 1291 apropiada puede recibir una señal transmitida desde el satélite 1290 o la estación 1170 base inalámbrica de la figura 26. Un vídeo decodificado se puede reproducir en una pantalla de visualización de un sistema 1293 de navegación de automóvil integrado en el automóvil 1292.

15 Una señal de vídeo puede ser codificada por un aparato de codificación de vídeo de acuerdo con una realización de la presente invención y puede entonces ser almacenado en un medio de almacenamiento. Específicamente, una señal de imagen puede almacenarse en un disco DVD 1296 mediante una grabadora de DVD o puede almacenarse en un disco duro mediante una grabadora 1295 de disco duro. Como otro ejemplo, la señal de vídeo se puede almacenar en una tarjeta SD 1297. Si la grabadora 1295 de disco duro incluye un aparato de decodificación de vídeo de acuerdo con una realización de la presente invención, se puede reproducir una señal de vídeo grabada en el disco DVD 1296, la tarjeta SD 1297 u otro medio de almacenamiento en el monitor 1288 de TV.

20 El sistema 1293 de navegación de automóvil puede no incluir la cámara 1253, la interfaz 1263 de la cámara, y el codificador 1272 de imágenes de la figura 32. Por ejemplo, el ordenador 1210 y el receptor de TV 1281 pueden no estar incluidos en la cámara 1253, la interfaz 1263 de la cámara y el codificador 1272 de imágenes de la figura 32.

La figura 34 ilustra una estructura de red de un sistema informático en la nube que usa un aparato de codificación de vídeo y un aparato de decodificación de vídeo de acuerdo con una realización de la presente invención.

25 El sistema informático en la nube puede incluir un servidor 1400 informático en la nube, una base 1410 de datos de usuario (DB), una pluralidad de los recursos 1420 informáticos, y un terminal de usuario.

30 El sistema informático en la nube proporciona un servicio de externalización bajo demanda de la pluralidad de recursos 1420 informáticos a través de una red de comunicación de información, por ejemplo, Internet, en respuesta a una solicitud del terminal de usuario. Bajo un entorno informático en la nube, un proveedor de servicios brinda a los usuarios los servicios deseados al combinar los recursos informáticos en centros de datos ubicados en ubicaciones físicamente diferentes mediante el uso de la tecnología de virtualización. Un usuario del servicio no tiene que instalar recursos informáticos, por ejemplo, una aplicación, una instalación de almacenamiento, un sistema operativo (OS) o características de seguridad, en su propio terminal para usarlos, pero puede seleccionar y usar los servicios deseados de entre los servicios en un espacio virtual generado a través de la tecnología de virtualización, en el momento deseado.

35 Un terminal de usuario de un usuario de servicio específico está conectado al servidor 1410 informático en la nube a través de una red de comunicación de información que incluye Internet y una red de telecomunicaciones móviles. Los terminales de usuario pueden recibir servicios informáticos en la nube y, en particular, servicios de reproducción de imágenes en movimiento, desde el servidor 1410 informático en la nube. Los terminales de usuario pueden ser 40 varios tipos de dispositivos electrónicos que pueden conectarse a Internet, por ejemplo, una PC 1430 de escritorio, un televisor 1440 inteligente, un teléfono 1450 inteligente, un ordenador 1460 portátil, un reproductor 1470 multimedia portátil (PMP), un PC 1480 de tableta, y similares.

45 El servidor 1410 informático en la nube puede combinar la pluralidad de los recursos 1420 informáticos distribuidos en una red de la nube y proporcionar terminales de usuario con un resultado de la combinación. La pluralidad de recursos 1420 informáticos puede incluir diversos servicios de datos, y puede incluir datos cargados desde terminales de usuario. Como se describió anteriormente, el servidor 1410 informático en la nube puede proporcionar a los terminales de usuario los servicios deseados al combinar una base de datos de imágenes en movimiento distribuida en diferentes regiones de acuerdo con la tecnología de virtualización.

50 La información del usuario sobre los usuarios que se han suscrito a un servicio informático en la nube se almacena en la base 1410 de datos de usuario. La información del usuario puede incluir información de registro, direcciones, nombres e información de crédito personal de los usuarios. La información del usuario puede incluir además índices de imágenes en movimiento. Aquí, los índices pueden incluir una lista de imágenes en movimiento que ya se han reproducido, una lista de imágenes en movimiento que se están reproduciendo, un punto de pausa de una imagen en movimiento que se estaba reproduciendo, y similares.

55 La información sobre una imagen en movimiento almacenada en la base 1410 de datos del usuario puede compartirse entre los dispositivos del usuario. Por ejemplo, cuando se proporciona un servicio de imágenes en

movimiento al ordenador 1460 portátil en respuesta a una solicitud del ordenador 1460 portátil, se almacena un historial de reproducción del servicio de imágenes en movimiento en la base 1410 de datos del usuario. Cuando se recibe una solicitud para reproducir este servicio de imágenes en movimiento desde el teléfono 1450 inteligente, el servidor 1410 informático en la nube busca y reproduce este servicio de imágenes en movimiento, basado en la base 1410 de datos del usuario. Cuando el teléfono 1450 inteligente recibe un flujo de datos de imágenes en movimiento del servidor 1410 informático en la nube, un proceso de reproducción de vídeo mediante la decodificación del flujo de datos de imágenes en movimiento es similar a una operación del teléfono 1250 móvil descrito anteriormente con referencia a la figura 28.

El servidor 1410 informático en la nube puede referirse a un historial de reproducción de un servicio de imagen en movimiento deseado, almacenado en la DB 1410 del usuario. Por ejemplo, el servidor 1410 informático en la nube recibe una solicitud para reproducir una imagen en movimiento almacenada en la base 1410 de datos del usuario, desde un terminal de usuario. Si esta imagen en movimiento se estaba reproduciendo, entonces un procedimiento de transmisión de esta imagen en movimiento, realizado por el servidor 1410 informático en la nube puede variar de acuerdo con la solicitud del terminal del usuario, es decir, según si la imagen en movimiento se reproducirá, comenzando desde un inicio del mismo o un punto de pausa del mismo. Por ejemplo, si el terminal de usuario solicita reproducir la imagen en movimiento, comenzando desde el principio, el servidor 1410 informático en la nube transmite la imagen en movimiento desde el primer marco del mismo al terminal de usuario. Si el terminal de usuario solicita reproducir la imagen en movimiento, comenzando desde el punto de pausa del mismo, el servidor 1410 informático en la nube transmite la transmisión de la imagen en movimiento desde un cuadro correspondiente al punto de pausa, al terminal de usuario.

En este caso, el terminal de usuario puede incluir un aparato de decodificación de vídeo como se describe anteriormente con referencia a las figuras 1 a 28. Como otro ejemplo, el terminal de usuario puede incluir un aparato de codificación de vídeo como se describe anteriormente con referencia a las figuras 1 a 28. Alternativamente, el terminal del usuario puede incluir tanto el aparato de decodificación de vídeo como el aparato de codificación de vídeo como se describió anteriormente con referencia a las figuras 1 a 28.

Varias aplicaciones de un procedimiento de codificación de vídeo, un procedimiento de decodificación de vídeo, un aparato de codificación de vídeo, y un aparato de decodificación de vídeo, de acuerdo con una realización de la presente invención, se describen anteriormente con referencia a las figuras 1 a 28 se ha descrito anteriormente con referencia a las figuras 25A a 30. Sin embargo, los procedimientos para almacenar el procedimiento de codificación de vídeo y el procedimiento de decodificación de vídeo en un medio de almacenamiento o los procedimientos de incluir el aparato de codificación de vídeo y el aparato de decodificación de vídeo en un dispositivo de acuerdo con varias realizaciones de la presente invención, no están limitados a las realizaciones descritas anteriormente con referencia a las figuras 29A a 34.

Los expertos en la técnica entenderán los diagramas de bloques descritos en la presente divulgación como diagramas conceptuales de circuitos para la realización de los principios de la presente invención. De manera similar, a los expertos en la técnica les resultará evidente que diagramas de flujo arbitrarios, diagramas de flujo, diagramas de transición de estado, pseudocódigo y otros similares indican diversos procesos que pueden almacenarse sustancialmente en un medio de grabación legible por ordenador y que pueden ser realizados por un ordenador o un procesador, independientemente de si el ordenador o el procesador están explícitamente ilustrados o no. Por lo tanto, las realizaciones de la presente invención descritas anteriormente pueden realizarse como un programa de ordenador. El programa de ordenador puede almacenarse en un medio de grabación legible por ordenador y ejecutarse usando un ordenador digital general. Ejemplos del medio legible por ordenador son un medio de grabación magnético (una ROM, un disquete, un disco duro, etc.) y un medio de grabación óptico (un CD-ROM, un DVD, etc.).

Las funciones de los diversos elementos ilustrados en los dibujos pueden estar relacionados con el software apropiado, y puede proporcionar a través de no solo hardware capaz de ejecutar el software, sino también hardware exclusivo. Estas funciones también pueden proporcionarse mediante un único procesador exclusivo, mediante un único procesador compartido, o mediante una pluralidad de procesadores individuales, algunos de los cuales pueden compartirse. Además, el uso explícito del término 'procesador' o 'controlador' no se limita a usar exclusivamente hardware capaz de ejecutar software, y puede incluir implícitamente hardware como un procesador de señal digital (DSP) y una memoria de solo lectura (ROM), una memoria de acceso aleatorio (RAM) o un medio de almacenamiento no volátil para almacenar software.

En las reivindicaciones de la presente memoria descriptiva, un elemento sugerido como un elemento para realizar una operación específica incluye cualesquiera procedimientos arbitrarios de la realización de la operación específica. Ejemplos de este elemento pueden incluir una combinación de elementos de circuito capaces de realizar la operación específica, o software que tenga una forma arbitraria, por ejemplo, firmware o microcódigo, que se combina con un circuito apropiado para ejecutar software para realizar la operación específica.

En la presente descripción, la expresión 'una realización' de los principios de la presente invención y diversas modificaciones de esta expresión significa que características específicas, la estructura y características relacionadas con esta realización se incluyen en al menos una realización de los principios de la presente invención.

Por lo tanto, la expresión 'una realización' y otras modificaciones arbitrarias de la misma descritas en la presente divulgación no siempre indican la misma realización.

5 En la presente descripción, la expresión 'al menos uno de' de 'al menos uno de A y B' se utiliza para significar inclusive que solo se selecciona la primera opción (A), solo se selecciona la segunda opción (B), o ambas operaciones, la primera y la segunda (A y B) están seleccionadas. Además, la expresión 'al menos uno de A, B y C' se usa para indicar de manera inclusiva que solo se selecciona la primera opción (A), solo se selecciona la segunda opción (B), solo se selecciona la tercera opción (C), solo se seleccionan las opciones primera y segunda (A y B), solo se seleccionan las opciones segunda y tercera (B y C), solo se seleccionan la primera y la tercera (A y C) o se seleccionan las tres opciones (A , B y C). Cuando se enumeran más de tres elementos en relación con esta
10 expresión, el significado de los mismos sería evidente para los expertos en la técnica.

Se han descrito anteriormente realizaciones de ejemplo de la presente invención.

Si bien la presente invención se ha mostrado y descrito en particular con referencia a realizaciones de ejemplo de la misma, los expertos en la materia deberán entender que se pueden efectuar diversos cambios en la forma y los detalles de la misma sin desviarse del alcance de la invención tal y como está definida en las reivindicaciones adjuntas. Las realizaciones de ejemplo deberán considerarse solo en un sentido descriptivo y no a efectos de limitación. Por lo tanto, el alcance de la invención está definido no por la descripción detallada de la invención, sino por las reivindicaciones adjuntas y deberá interpretarse que todas las diferencias dentro del alcance de la misma están incluidas en la presente invención.
15

REIVINDICACIONES

1. Un aparato de decodificación de un vídeo, comprendiendo el aparato:

un dispositivo de obtención que está configurado para:

5 obtener, basándose en el tamaño de una unidad de predicción actual, información del modo de predicción de movimiento que indica un modo de predicción de movimiento aplicado a la unidad de predicción actual entre un modo de predicción de movimiento unidireccional y un modo de predicción de movimiento bidireccional, en el que el modo de predicción de movimiento unidireccional comprende uno de entre un modo de predicción L0 y un modo de predicción L1; y

10 determinar el modo de predicción de movimiento aplicado a la unidad de predicción actual entre el modo de predicción L0, el modo de predicción L1 y el modo de predicción de movimiento bidireccional basado en la información del modo de predicción de movimiento de la unidad de predicción actual; y

un compensador de movimiento que está configurado para realizar, basándose en el modo de predicción de movimiento determinado, una compensación de movimiento en la unidad de predicción actual mediante el uso de valores de predicción de la unidad de predicción actual,

15 **caracterizado porque** el compensador de movimiento está configurado para:

determinar una ubicación (x_1, y_1) con x_1 igual a $N+1/4$ e y_1 igual a $M+1/4$, en la que N y M son números enteros, de un primer subpíxel en un marco de referencia de la unidad de predicción actual indicado por un vector de movimiento de la unidad de predicción actual, en el que el marco de referencia de la unidad de predicción actual se obtiene de acuerdo con el modo de predicción de movimiento determinado, en el que N y M son números enteros,

20 generar un valor de píxel del primer subpíxel utilizando un subpíxel vecino ubicado en una ubicación del lado superior (x_2, y_2) , con x_2 igual a $N+1/4$ e y_2 igual a M, del primer subpíxel, y obtener los valores de predicción de la unidad de predicción actual, incluyendo el valor de píxel generado del primer subpíxel.

25 2. Un procedimiento de codificación de un vídeo, comprendiendo el procedimiento:

generar información del modo de predicción de movimiento que indica un modo de predicción de movimiento aplicado a una unidad de predicción actual entre un modo de predicción de movimiento unidireccional y un modo de predicción de movimiento bidireccional, en el que el modo de predicción de movimiento unidireccional comprende uno de entre un modo de predicción L0 y un modo de predicción L1; y

30 realizar, basándose en el modo de predicción de movimiento, una compensación de movimiento en la unidad de predicción actual usando valores de predicción de la unidad de predicción actual,

caracterizado porque

35 el modo de predicción de movimiento se determina, basándose en un tamaño de la unidad de predicción actual, entre el modo de predicción L0, el modo de predicción L1 y el modo de predicción de movimiento bidireccional, la realización de la compensación de movimiento en la unidad de predicción actual comprende:

determinar una ubicación (x_1, y_1) con x_1 igual a $N+1/4$ e y_1 igual a $M+1/4$, en la que N y M son números enteros, de un primer subpíxel en un marco de referencia de la unidad de predicción actual, en el que el marco de referencia de la unidad de predicción actual se obtiene de acuerdo con el modo de predicción de movimiento determinado, en el que N y M son números enteros;

40 generar un valor de píxel del primer subpíxel utilizando un subpíxel vecino ubicado en una ubicación del lado superior (x_2, y_2) , con x_2 igual a $N+1/4$ e y_2 igual a M, del primer subpíxel; y

obtener los valores de predicción de la unidad de predicción actual, incluyendo el valor de píxel generado del primer subpíxel.

3. Un aparato de codificación de un vídeo, comprendiendo el aparato:

45 un codificador que está configurado para:

generar información del modo de predicción de movimiento que indica un modo de predicción de movimiento aplicado a una unidad de predicción actual entre un modo de predicción de movimiento unidireccional y un modo de predicción de movimiento bidireccional, en el que el modo de predicción de movimiento unidireccional comprende uno de entre un modo de predicción L0 y un modo de predicción L1; y

50 un compensador de movimiento que está configurado para realizar, basándose en el modo de predicción de movimiento, una compensación de movimiento en la unidad de predicción actual mediante el uso de valores de predicción de la unidad de predicción actual,

caracterizado porque

el modo de predicción de movimiento se determina, basándose en un tamaño de la unidad de predicción actual, entre el modo de predicción L0, el modo de predicción L1 y el modo de predicción de movimiento bidireccional,

en el que el compensador de movimiento está configurado para:

- 5 determinar una ubicación (x_1, y_1) con x_1 igual a $N+1/4$ e y_1 igual a $M+1/4$, en la que N y M son números enteros, de un primer subpíxel en un marco de referencia de la unidad de predicción actual, en el que el marco de referencia de la unidad de predicción actual se obtiene de acuerdo con el modo de predicción de movimiento determinado, en el que N y M son números enteros;
- 10 generar un valor de píxel del primer subpíxel utilizando un subpíxel vecino ubicado en una ubicación del lado superior (x_2, y_2) , con x_2 igual a $N+1/4$ e y_2 igual a M, del primer subpíxel; y
- obtener los valores de predicción de la unidad de predicción actual, incluyendo el valor de píxel generado del primer subpíxel.

FIG. 1

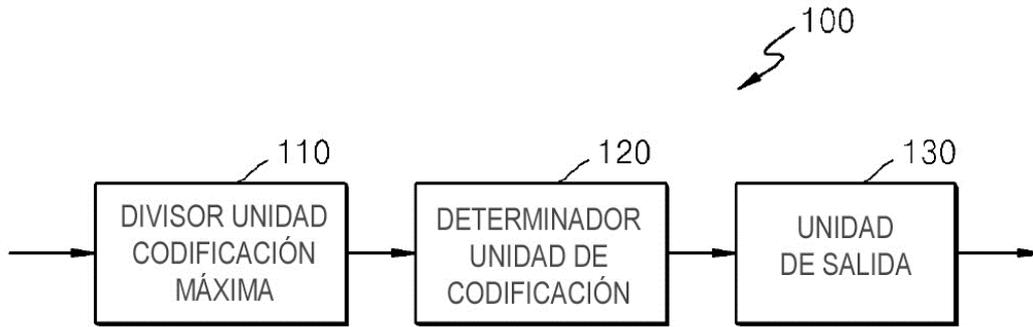


FIG. 2

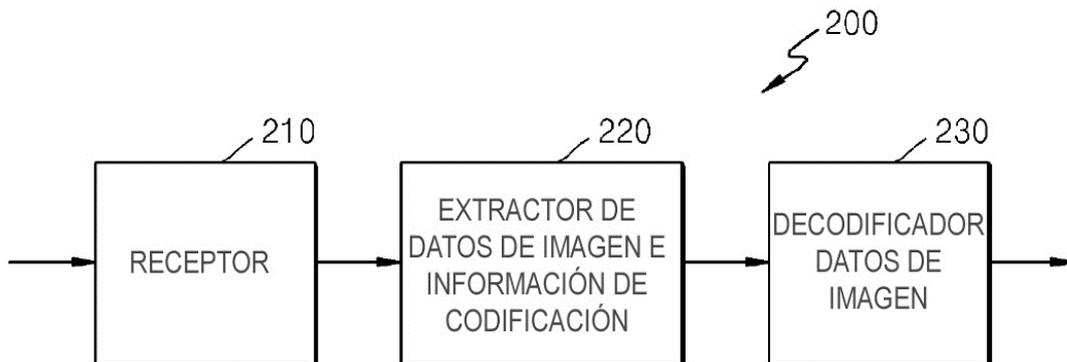


FIG. 3

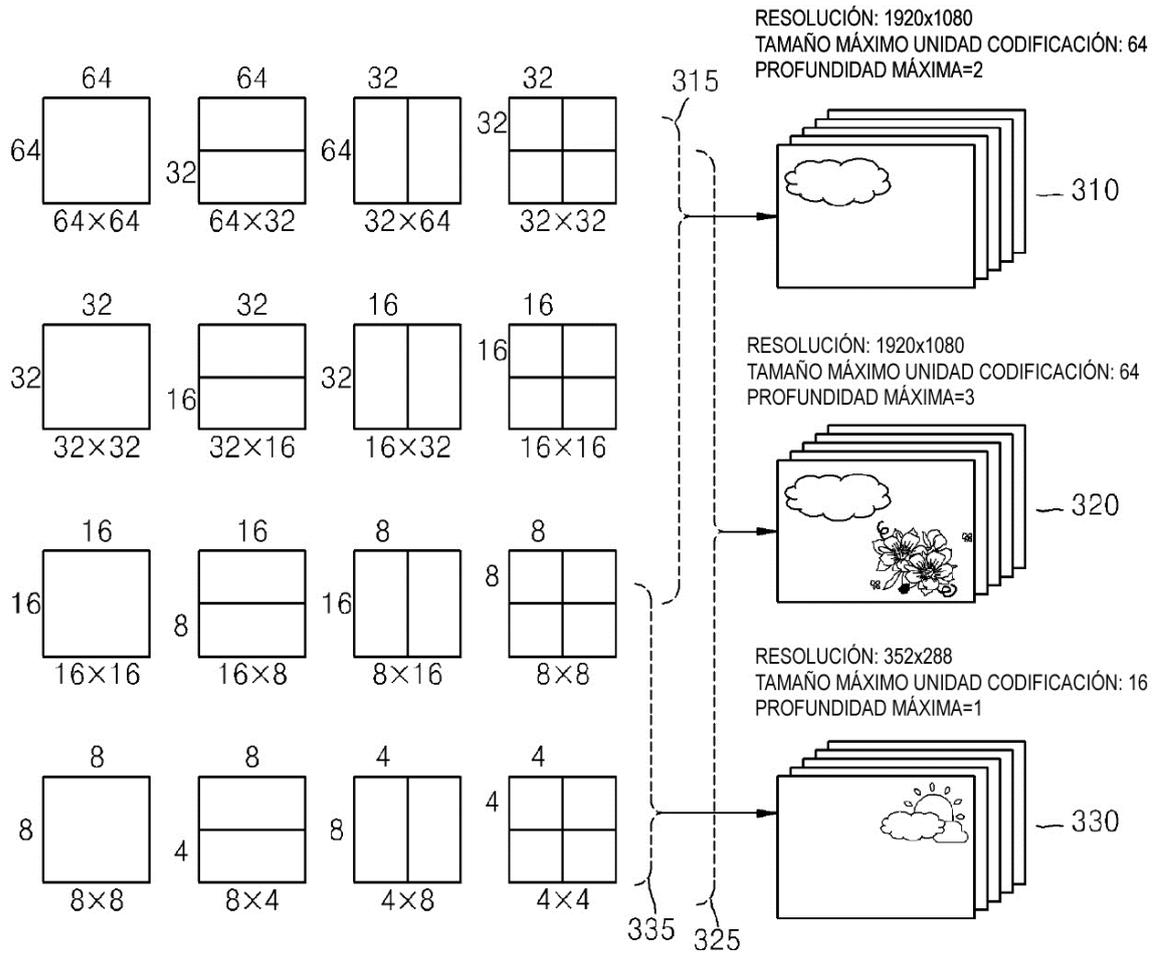


FIG. 4

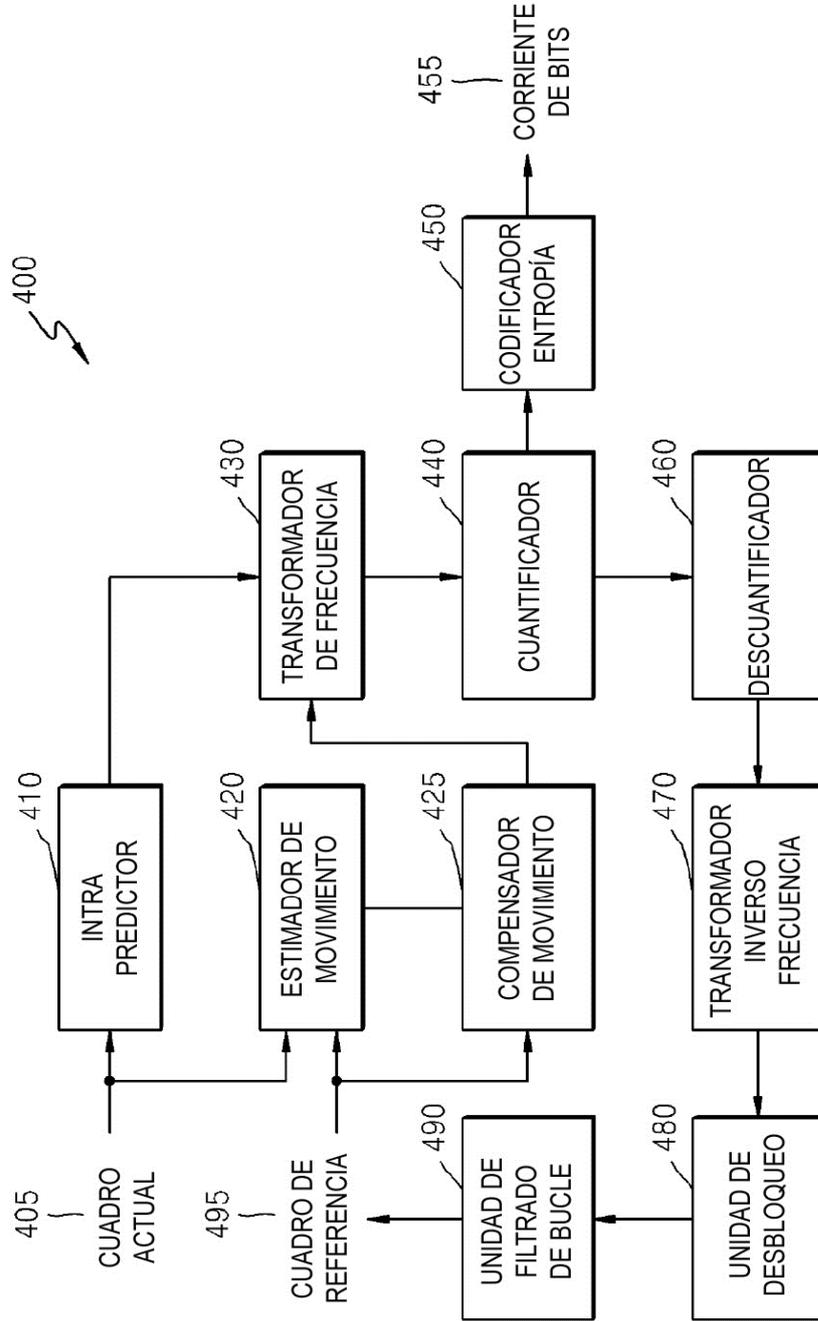


FIG. 5

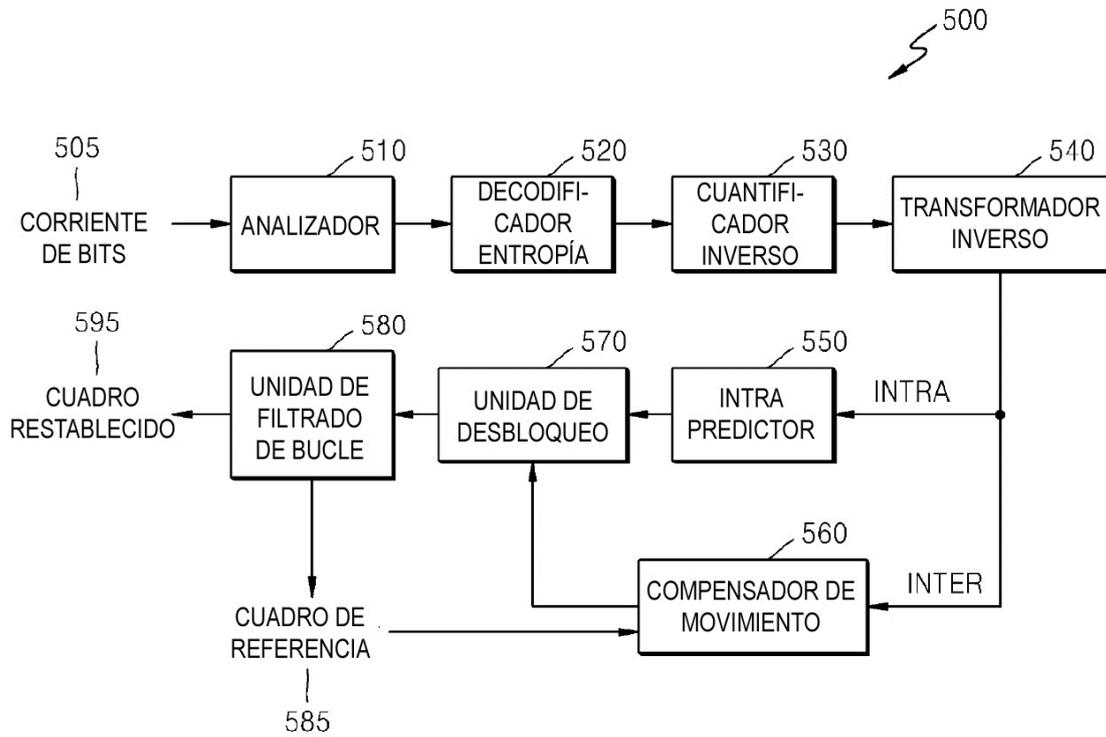


FIG. 6

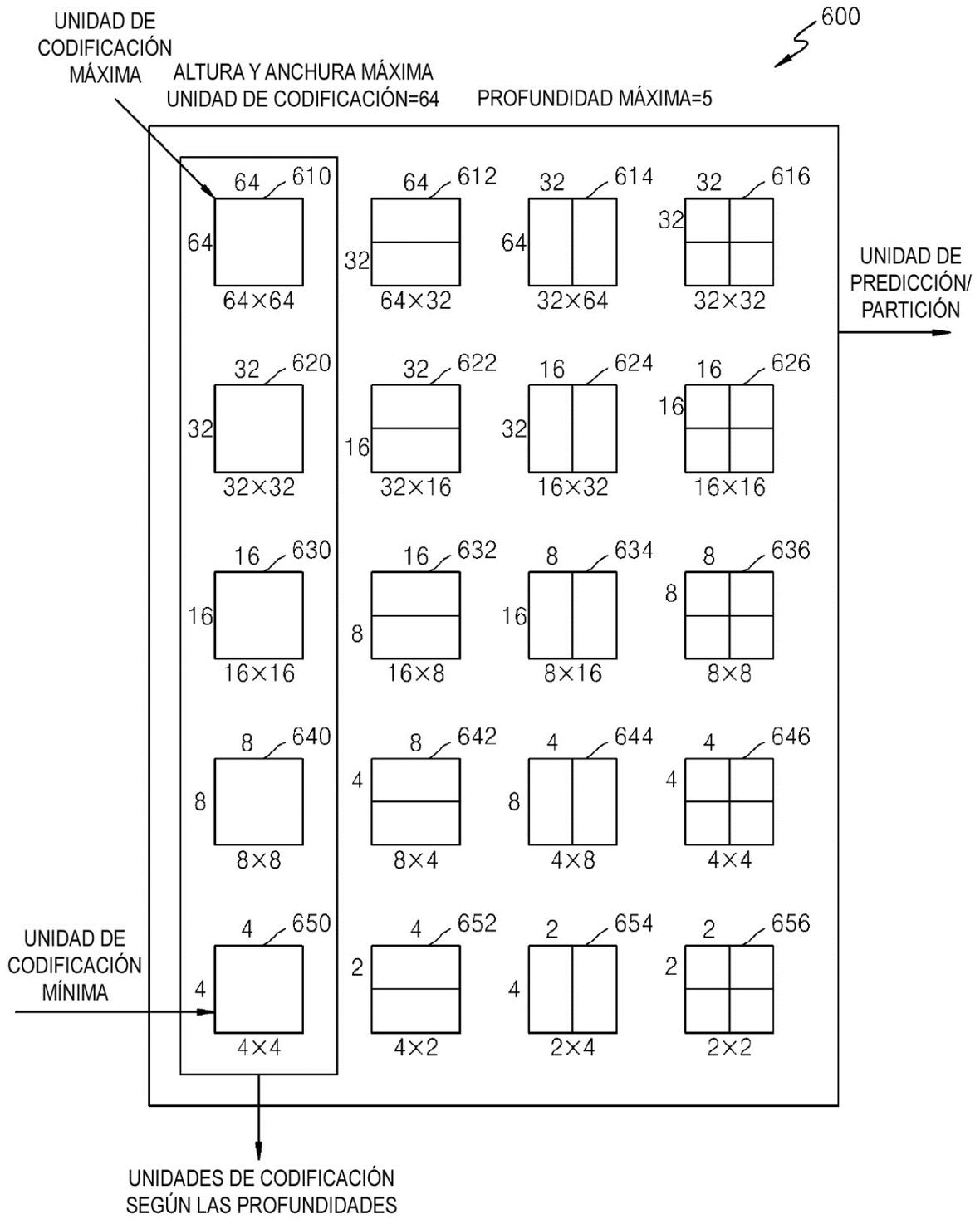


FIG. 7

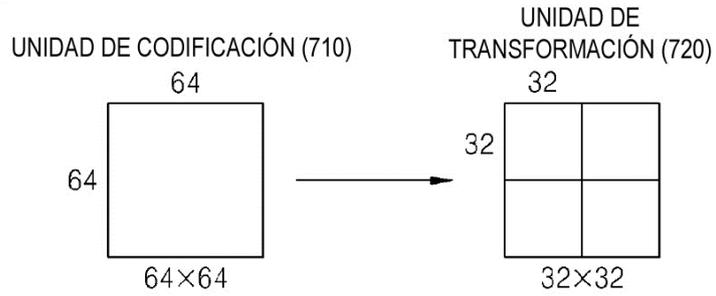
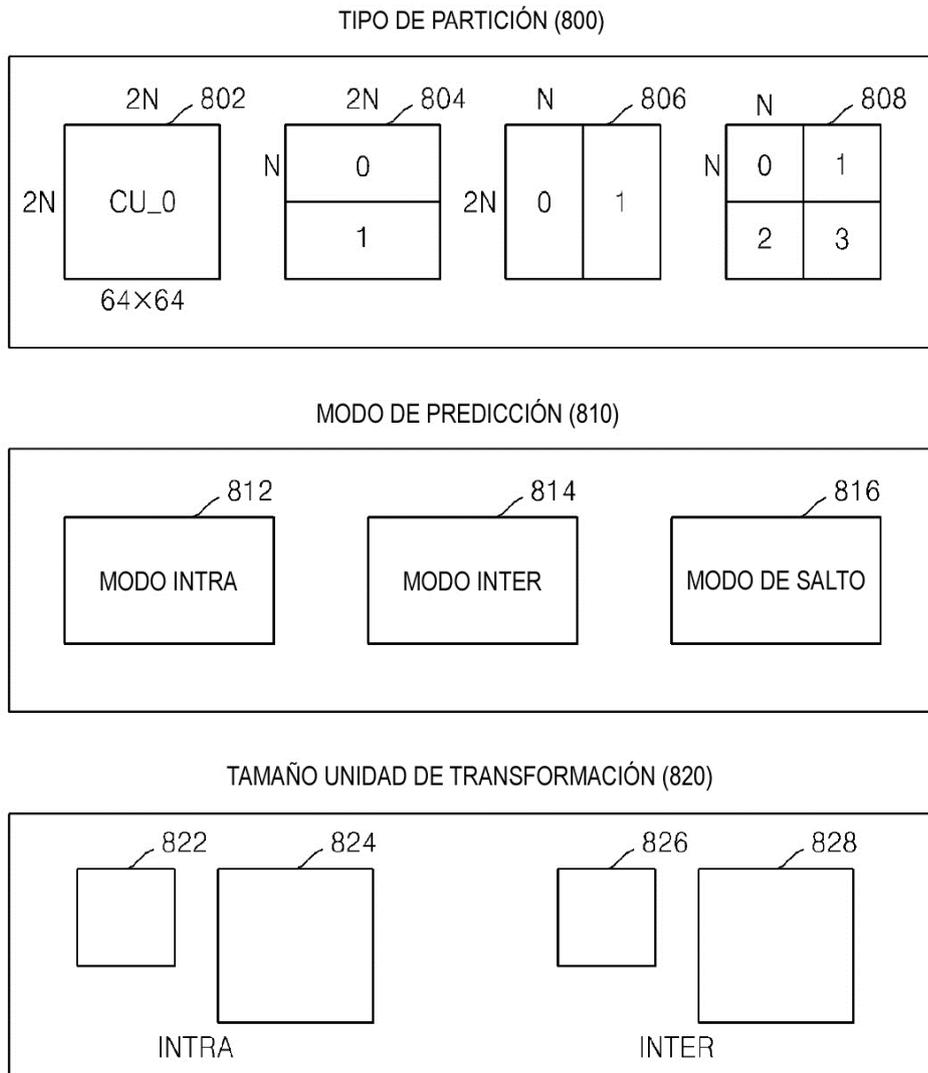


FIG. 8



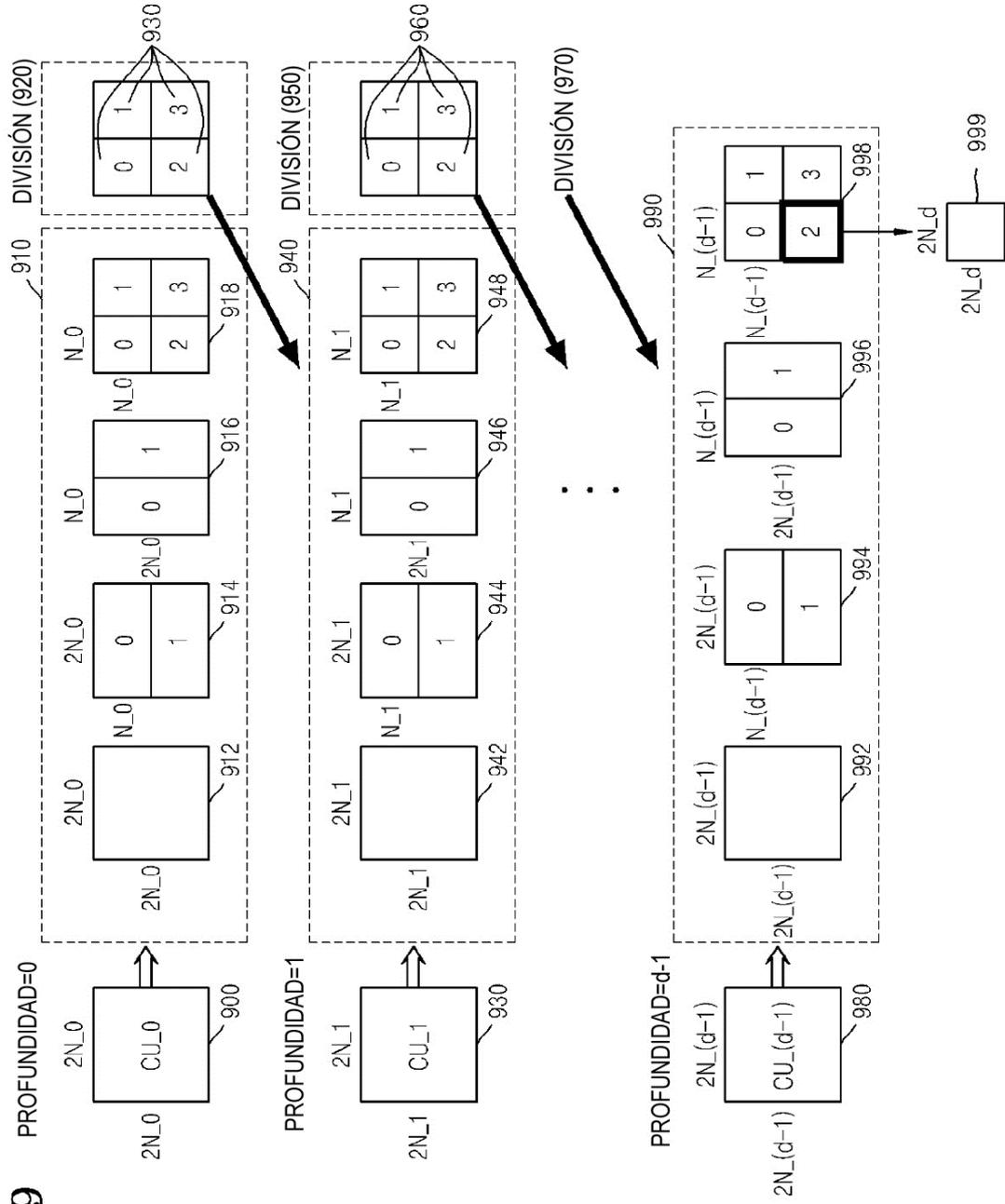
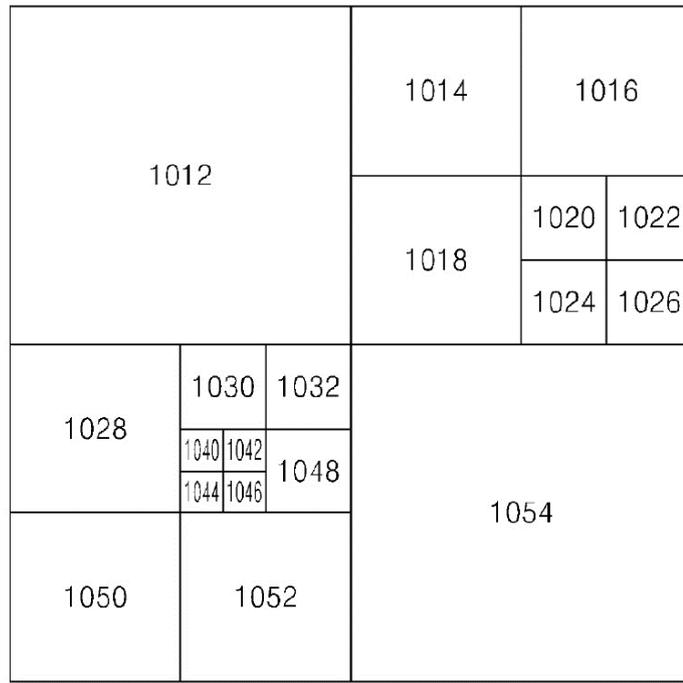
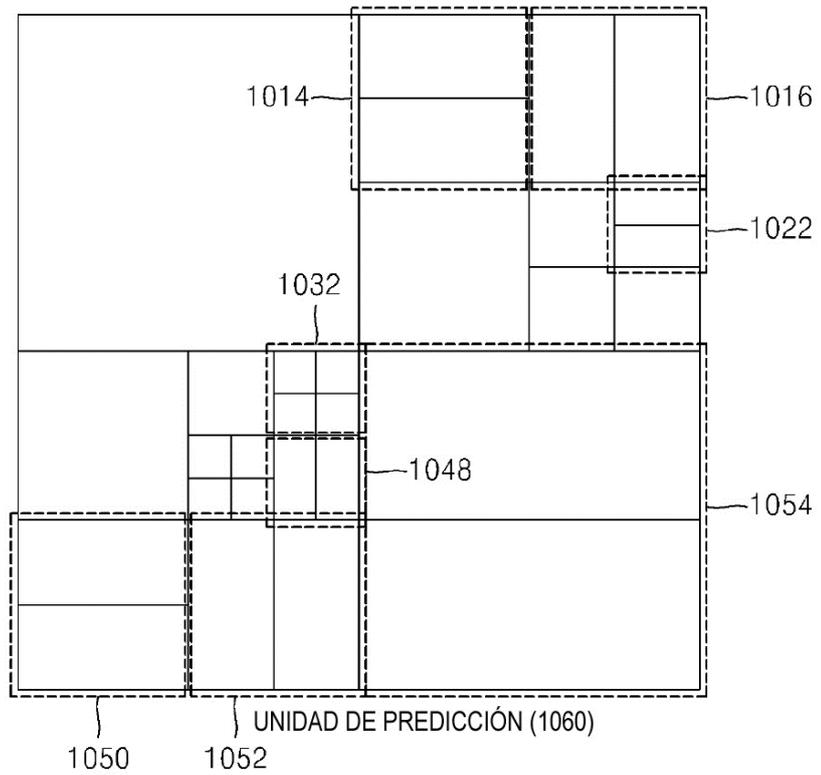


FIG. 10



UNIDAD DE CODIFICACIÓN (1010)

FIG. 11



UNIDAD DE PREDICCIÓN (1060)

FIG. 12

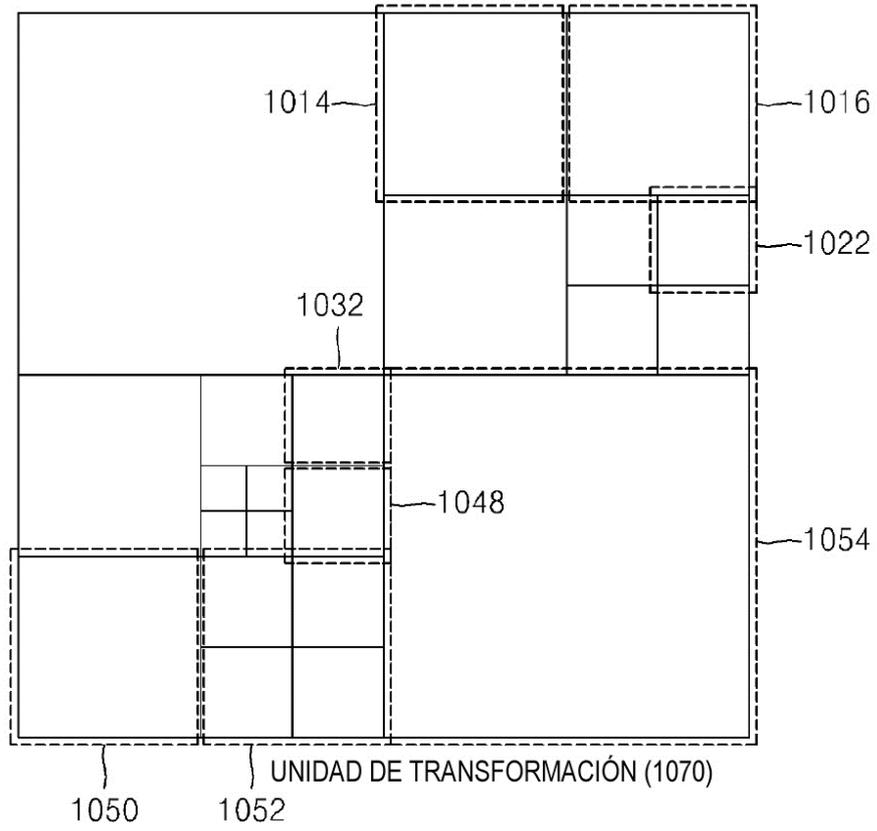


FIG. 13

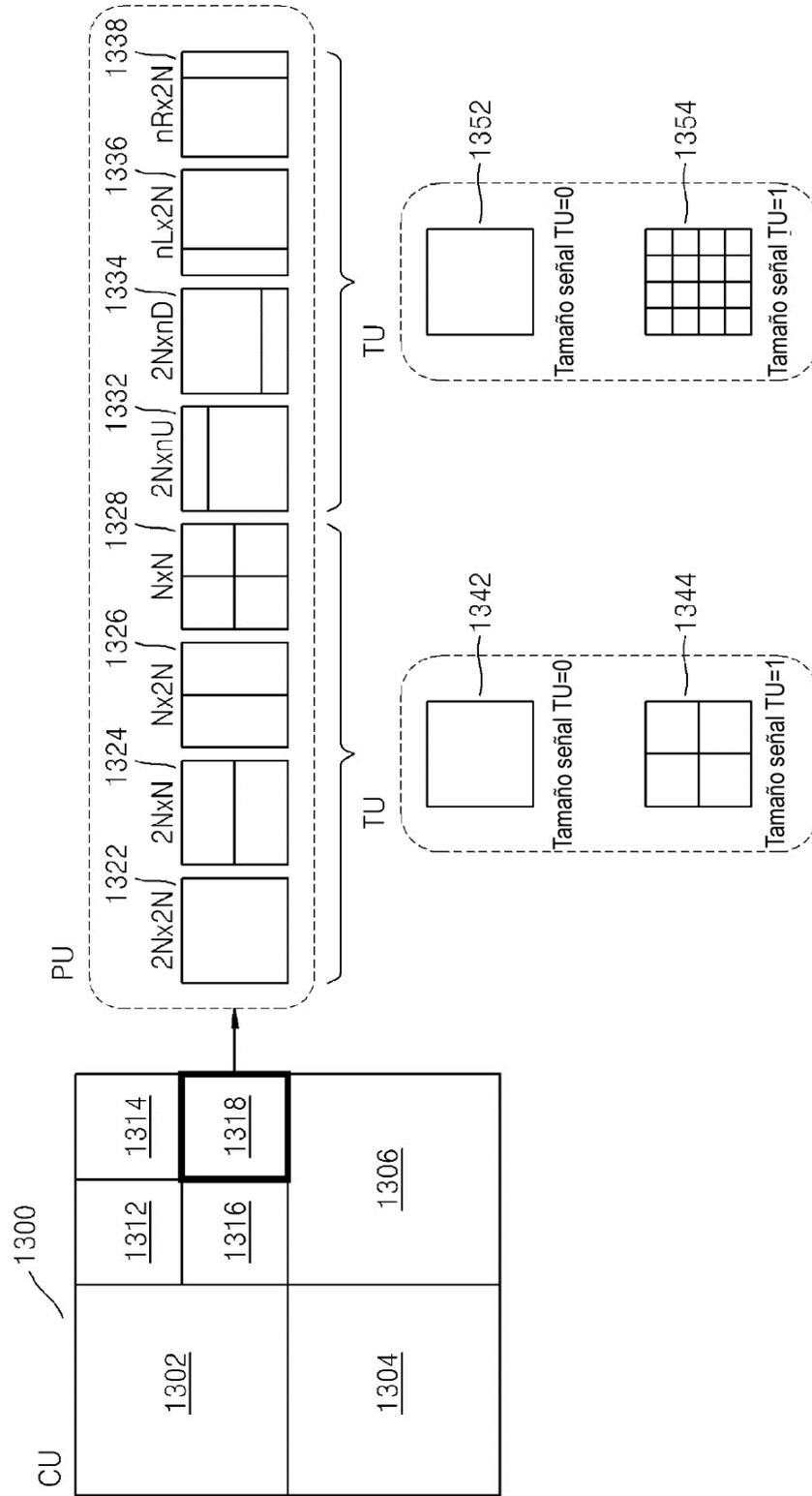


FIG. 14

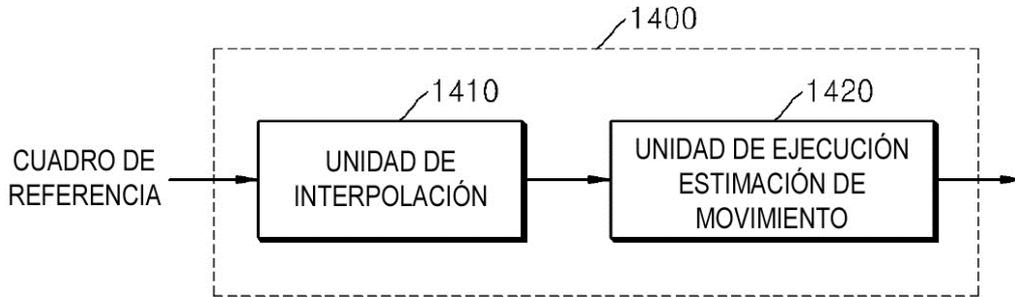


FIG. 15

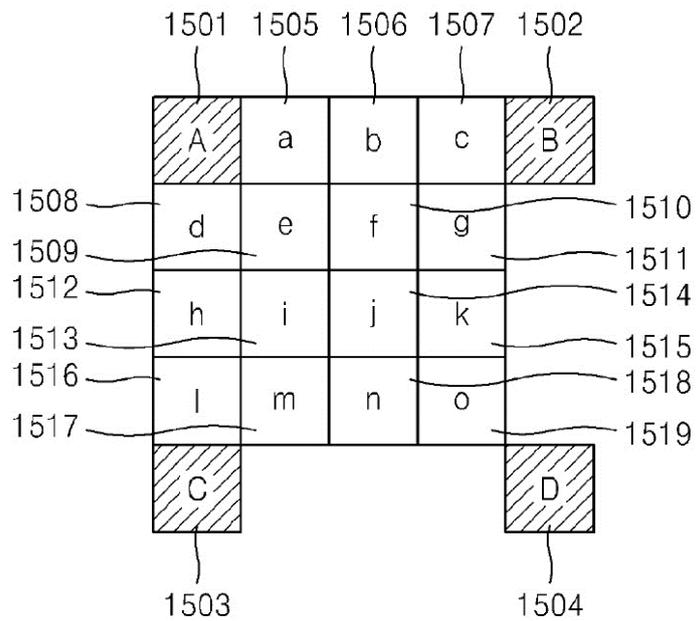


FIG. 16

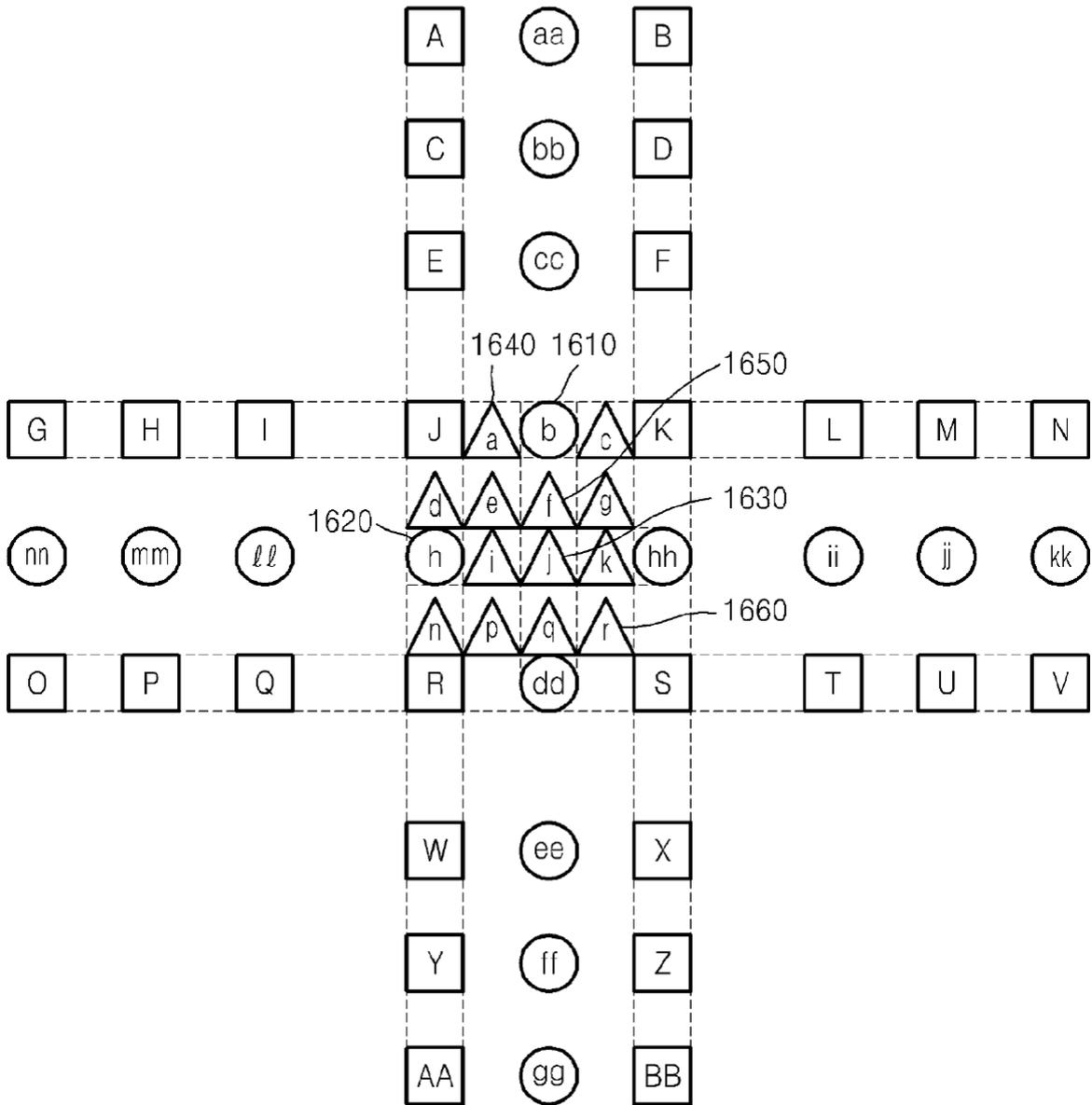


FIG. 17A

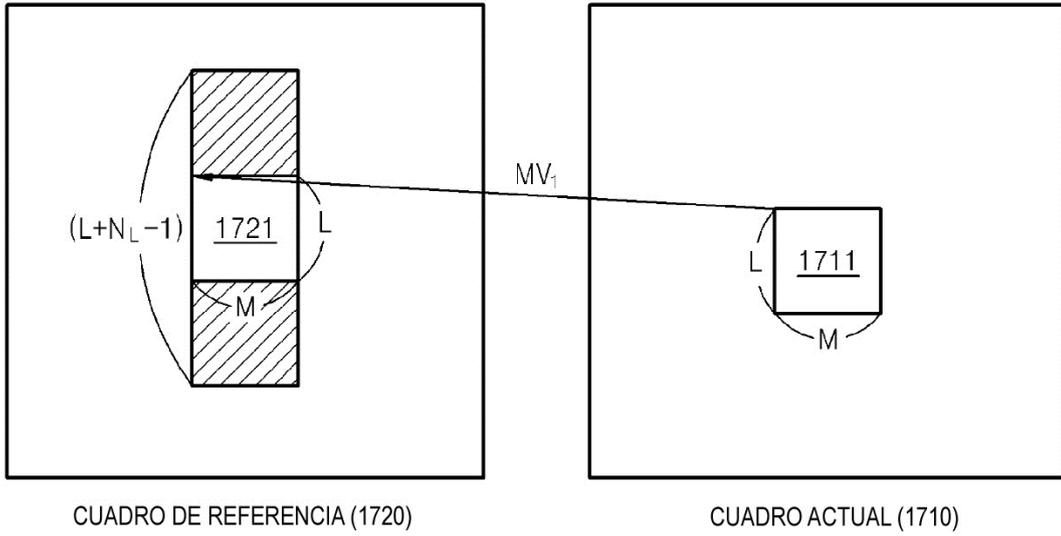


FIG. 17B

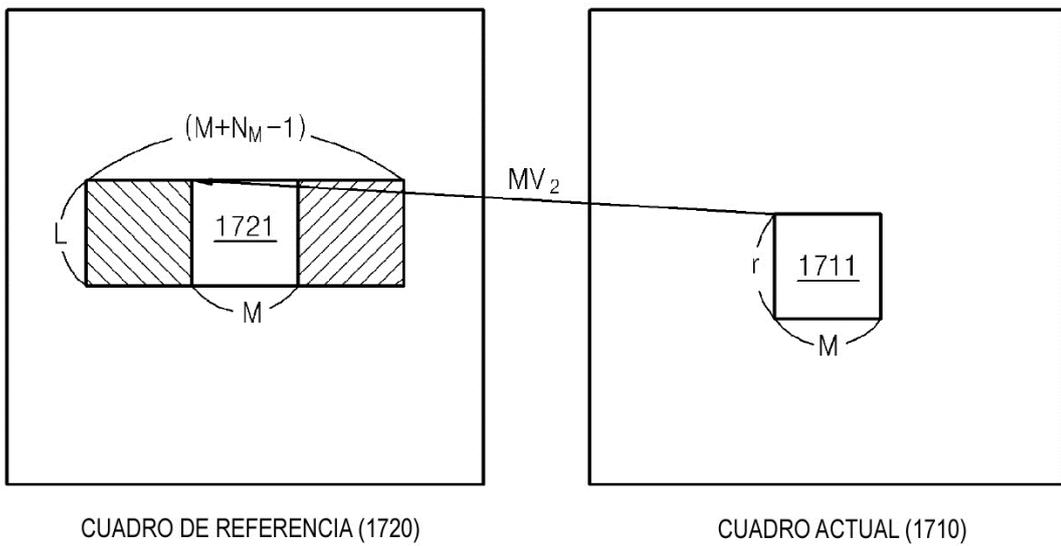


FIG. 17C

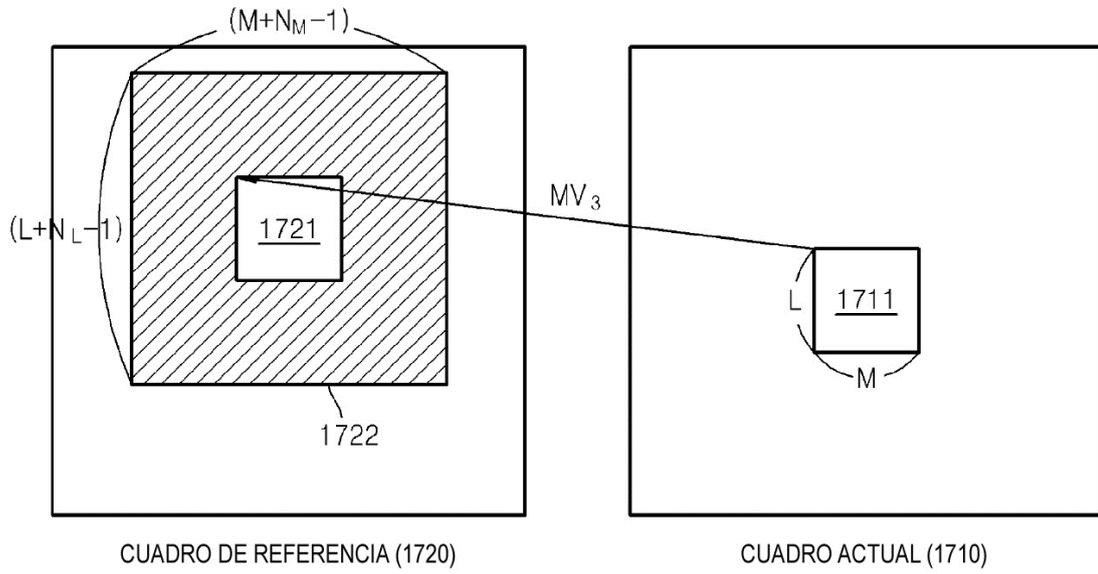


FIG. 18A

N=8

TAMAÑO BLOQUE		Unipredicción			Bipredicción					
		I	1D	2D	I-I	1D-1D	2D-2D	I-1D	I-2D	1D-2D
2	2	1	4,50	20,25	2	9,00	40,50	5,50	21,25	24,75
2	4	1	4,50	12,38	2	9,00	24,75	5,50	13,38	16,88
4	4	1	2,75	7,56	2	5,50	15,13	3,75	8,56	10,31
4	8	1	2,75	5,16	2	5,50	10,31	3,75	6,16	7,91
8	8	1	1,88	3,52	2	3,75	7,03	2,88	4,52	5,39
8	16	1	1,88	2,70	2	3,75	5,39	2,88	3,07	4,57
16	16	1	1,44	2,07	2	2,88	4,13	2,44	3,07	3,50
16	32	1	1,44	1,75	2	2,88	3,50	2,44	2,75	3,19
32	32	1	1,22	1,49	2	2,44	2,97	2,22	2,49	2,70
32	64	1	1,22	1,35	2	2,44	2,70	2,22	2,35	2,57
64	64	1	1,11	1,23	2	2,22	2,46	2,11	2,23	2,34

FIG. 18B

N=7

TAMAÑO BLOQUE		Unipredicción			Bipredicción					
		I	1D	2D	I-I	1D-1D	2D-2D	I-1D	I-2D	1D-2D
2	2	1	4,00	16,00	2	8,00	32,00	5,00	17,00	20,00
2	4	1	4,00	10,00	2	8,00	20,00	5,00	11,00	14,00
4	4	1	2,50	6,25	2	5,00	12,50	3,50	7,25	8,75
4	8	1	2,50	4,38	2	5,00	8,75	3,50	5,38	6,88
8	8	1	1,75	3,06	2	3,50	6,13	2,75	4,06	4,81
8	16	1	1,75	2,41	2	3,50	4,81	2,75	3,41	4,16
16	16	1	1,38	1,89	2	2,75	3,78	2,38	2,89	3,27
16	32	1	1,38	1,63	2	2,75	3,27	2,38	2,63	3,01
32	32	1	1,19	1,41	2	2,38	2,82	2,19	2,41	2,60
32	64	1	1,19	1,30	2	2,38	2,60	2,19	2,30	2,49
64	64	1	1,09	1,20	2	2,19	2,39	2,09	2,20	2,29

FIG. 18C

N=6

TAMAÑO BLOQUE		Unipredicción			Bipredicción					
		I	1D	2D	I-I	1D-1D	2D-2D	I-1D	I-2D	1D-2D
2	2	1	3,50	12,25	2	7,00	24,50	4,50	13,25	15,75
2	4	1	3,50	7,88	2	7,00	15,75	4,50	8,88	11,38
4	4	1	2,25	5,06	2	4,50	10,13	3,25	6,06	7,31
4	8	1	2,25	3,66	2	4,50	7,31	3,25	4,66	5,91
8	8	1	1,63	2,64	2	3,25	5,28	2,63	3,64	4,27
8	16	1	1,63	2,13	2	3,25	4,27	2,63	3,13	3,76
16	16	1	1,31	1,72	2	2,63	3,45	2,31	2,72	3,04
16	32	1	1,31	1,52	2	2,63	3,04	2,31	2,52	2,83
32	32	1	1,16	1,34	2	2,31	2,67	2,16	2,34	2,49
32	64	1	1,16	1,25	2	2,31	2,49	2,16	2,25	2,40

FIG. 18D

N=5

TAMAÑO BLOQUE		Unipredicción			Bipredicción					
		I	1D	2D	I-I	1D-1D	2D-2D	I-1D	I-2D	1D-2D
2	2	1	3,00	9,00	2	6,00	18,00	4,00	10,00	12,00
2	4	1	3,00	6,00	2	6,00	12,00	4,00	7,00	9,00
4	4	1	2,00	4,00	2	4,00	8,00	3,00	5,00	6,00
4	8	1	2,00	3,00	2	4,00	6,00	3,00	4,00	5,00
8	8	1	1,50	2,25	2	3,00	4,50	2,50	3,25	3,75
8	16	1	1,50	1,88	2	3,00	3,75	2,50	2,88	3,38
16	16	1	1,25	1,56	2	2,50	3,13	2,25	2,56	2,81
16	32	1	1,25	1,41	2	2,50	2,81	2,25	2,41	2,66
32	32	1	1,13	1,27	2	2,25	2,53	2,13	2,27	2,39
32	64	1	1,13	1,20	2	2,25	2,39	2,13	2,20	2,32
64	64	1	1,06	1,13	2	2,13	2,26	2,06	2,13	2,19

FIG. 18E

N=4

TAMAÑO BLOQUE		Unipredicción			Bipredicción					
		I	1D	2D	I-I	1D-1D	2D-2D	I-1D	I-2D	1D-2D
2	2	1	2,50	6,25	2	5,00	12,50	3,50	7,25	8,75
2	4	1	2,50	4,38	2	5,00	8,75	3,50	5,38	6,88
4	4	1	1,75	3,06	2	3,50	6,13	2,75	4,06	4,81
4	8	1	1,75	2,41	2	3,50	4,81	2,75	3,41	4,16
8	8	1	1,38	1,89	2	2,75	3,78	2,38	2,63	3,27
8	16	1	1,38	1,63	2	2,75	3,27	2,38	2,41	3,01
16	16	1	1,19	1,41	2	2,38	2,82	2,19	2,30	2,60
16	32	1	1,19	1,30	2	2,38	2,60	2,19	2,30	2,49
32	32	1	1,09	1,20	2	2,19	2,39	2,09	2,20	2,29
32	64	1	1,09	1,15	2	2,19	2,29	2,09	2,15	2,24
64	64	1	1,05	1,10	2	2,09	2,19	2,05	2,10	2,14

FIG. 18F

N=3

TAMAÑO BLOQUE		Unipredicción			Bipredicción					
		I	1D	2D	I-I	1D-1D	2D-2D	I-1D	I-2D	1D-2D
2	2	1	2,00	4,00	2	4,00	8,00	3,00	5,00	6,00
2	4	1	2,00	3,00	2	4,00	6,00	3,00	4,00	5,00
4	4	1	1,50	2,25	2	3,00	4,50	2,50	3,25	3,75
4	8	1	1,50	1,88	2	3,00	3,75	2,50	2,88	3,38
8	8	1	1,25	1,56	2	2,50	3,13	2,25	2,56	2,81
8	16	1	1,25	1,41	2	2,50	2,81	2,25	2,41	2,66
16	16	1	1,13	1,27	2	2,25	2,53	2,13	2,27	2,39
16	32	1	1,13	1,20	2	2,25	2,39	2,13	2,20	2,32
32	32	1	1,06	1,13	2	2,13	2,26	2,06	2,13	2,19

FIG. 18G

N=2

TAMAÑO BLOQUE		Unipredicción			Bipredicción					
		I	1D	2D	I-I	1D-1D	2D-2D	I-1D	I-2D	1D-2D
2	2	1	1,50	2,25	2	3,00	4,50	2,50	3,25	3,75
2	4	1	1,50	1,88	2	3,00	3,75	2,50	2,88	3,38
4	4	1	1,25	1,56	2	2,50	3,13	2,25	2,56	2,81
4	8	1	1,25	1,41	2	2,50	2,81	2,25	2,41	2,66
8	8	1	1,13	1,27	2	2,25	2,53	2,13	2,27	2,39
8	16	1	1,13	1,20	2	2,25	2,39	2,13	2,20	2,32
16	16	1	1,06	1,13	2	2,13	2,26	2,06	2,13	2,19
16	32	1	1,06	1,10	2	2,13	2,19	2,06	2,10	2,16
32	32	1	1,03	1,06	2	2,06	2,13	2,03	2,06	2,09
32	64	1	1,03	1,05	2	2,06	2,09	2,03	2,05	2,08
64	64	1	1,02	1,03	2	2,03	2,06	2,02	2,03	2,05

FIG. 19

N=8 (para Luma), N=4 (para croma)

#	Comp. color	Tamaño	Uni/Bi	Posición MV	Píxeles para leer desde la memoria por 1 píxel MC
1	Luma	4x8	bi	2D-2D	10,31
2	Croma	2x4	bi	2D-2D	8,75
3	Luma	4x8	bi	1D-2D	7,91
4	Luma	8x8	bi	2D-2D	7,03
5	Croma	2x4	bi	1D-2D	6,88
6	Luma	4x8	bi	1-2D	6,16
7	Croma	4x4	bi	2D-2D	6,13
8	Luma	4x8	bi	1D-1D	5,50
9	Luma	8x16 8x8	bi	2D-2D 1D-2D	5,39
10	Croma	2x4	bi	1-2D	5,38
11	Luma	8x8	uni	2D-2D	5,16
12

FIG. 20

N=8 (para Luma), N=4 (para croma)

#	Comp. color	Tamaño	Uni/Bi	Posición MV	Píxeles para leer desde la memoria por 1 píxel MC
1	Luma	4x4	bi	2D-2D	15,31
2	Croma	2x2	bi	2D-2D	12,50
3	Luma	4x4 4x8	bi	1D-2D 2D-2D	10,31
4	Croma	2x2 2x4	bi	1D-2D 2D-2D	8,75
5	Luma	4x4	bi	1D-2D	8,56
6	Luma	4x8	bi	1D-2D	7,91
7	Luma	4x4	uni	2D	7,56
8	Croma	2x2	bi	1-2D	7,25
9	Luma	8x8	bi	2D-2D	7,03
10	Croma	2x4	bi	1D-2D	6,88
11	Croma	2x2	uni	2D	6,25
12	Luma	4x8	bi	1-2D	6,16
13	Croma	4x4	bi	2D-2D	6,13
14	Luma	4x8	bi	1D-1D	5,50
15	Luma	8x16 8x8	bi	2D-2D 1D-2D	3,39
16	Croma	2x4	bi	1-2D	5,38
17	Luma	8x8	uni	2D	5,16
18

FIG. 21

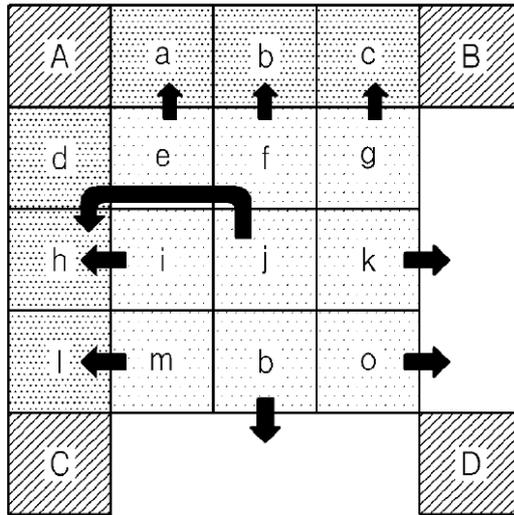


FIG. 22

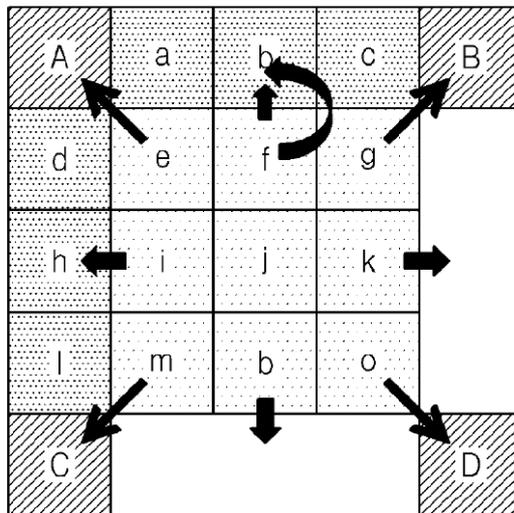


FIG. 23

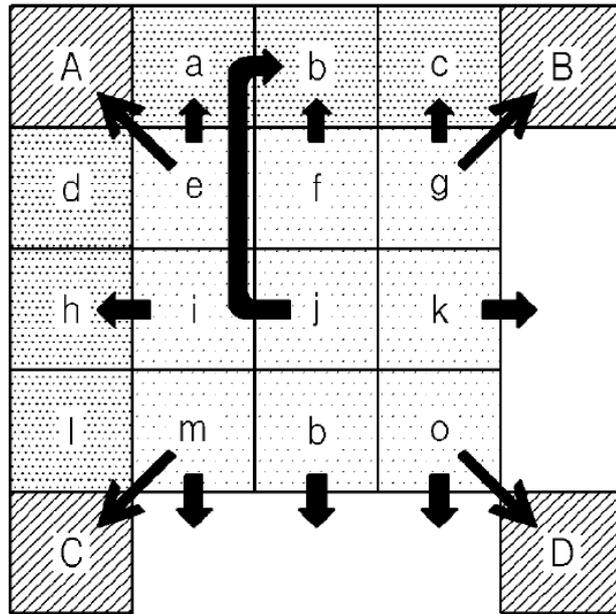


FIG. 24

#	Comp. color	Tamaño	Uni/Bi	Posición MV	Original	A	B
1	Luma	4x8	bi	2D-2D	10,31	4,81	5,50
2	Croma	2x4	bi	2D-2D	8,75	3,75	5,00
3	Luma	4x8	bi	1D-2D	7,91	4,16	5,50
4	Luma	8x8	bi	2D-2D	7,03	3,78	3,75
5	Croma	2x4	bi	1D-2D	6,88	3,38	5,00
6	Luma	4x8	bi	1-2D	6,16	3,41	3,75
7	Croma	4x4	bi	2D-2D	6,13	3,13	3,50
8	Luma	4x8	bi	1D-1D	5,50	3,50	5,50
9	Luma	8x16 8x8	bi	2D-2D 1D-2D	5,39	3,27	3,75
10	Croma	2x4	bi	1-2D	5,38	2,88	3,50
11	Luma	8x8	uni	2D-2D	5,16	2,41	2,75
12		

FIG. 25

#	Comp. color	Tamaño	Uni/Bi	Posición MV	Original	A	B
1	Luma	4x4	bi	2D-2D	15,31	6,13	5,50
2	Croma	2x2	bi	2D-2D	12,50	4,50	5,00
3	Luma	4x4 4x8	bi	1D-2D 2D-2D	10,31	4,81	5,50
4	Croma	2x2 2x4	bi	1D-2D 2D-2D	8,75	3,75	5,00
5	Luma	4x4	bi	1-2D	8,56	4,06	3,75
6	Luma	4x8	bi	1D-2D	7,01	4,16	5,50
7	Luma	4x4	uni	2D	7,56	3,06	2,75
8	Croma	2x2	bi	1-2D	7,25	3,25	3,50
9	Luma	8x8	bi	2D-2D	7,03	3,78	3,75
10	Croma	2x4	bi	1D-2D	6,88	3,38	5,00
11	Croma	2x2	uni	2D	6,25	1,50	2,50
12	Luma	4x8	bi	1-2D	6,16	3,41	3,75
13	Croma	4x4	bi	2D-2D	6,13	3,13	3,50
14	Luma	4x8	bi	1D-1D	5,50	3,50	5,50

FIG. 26

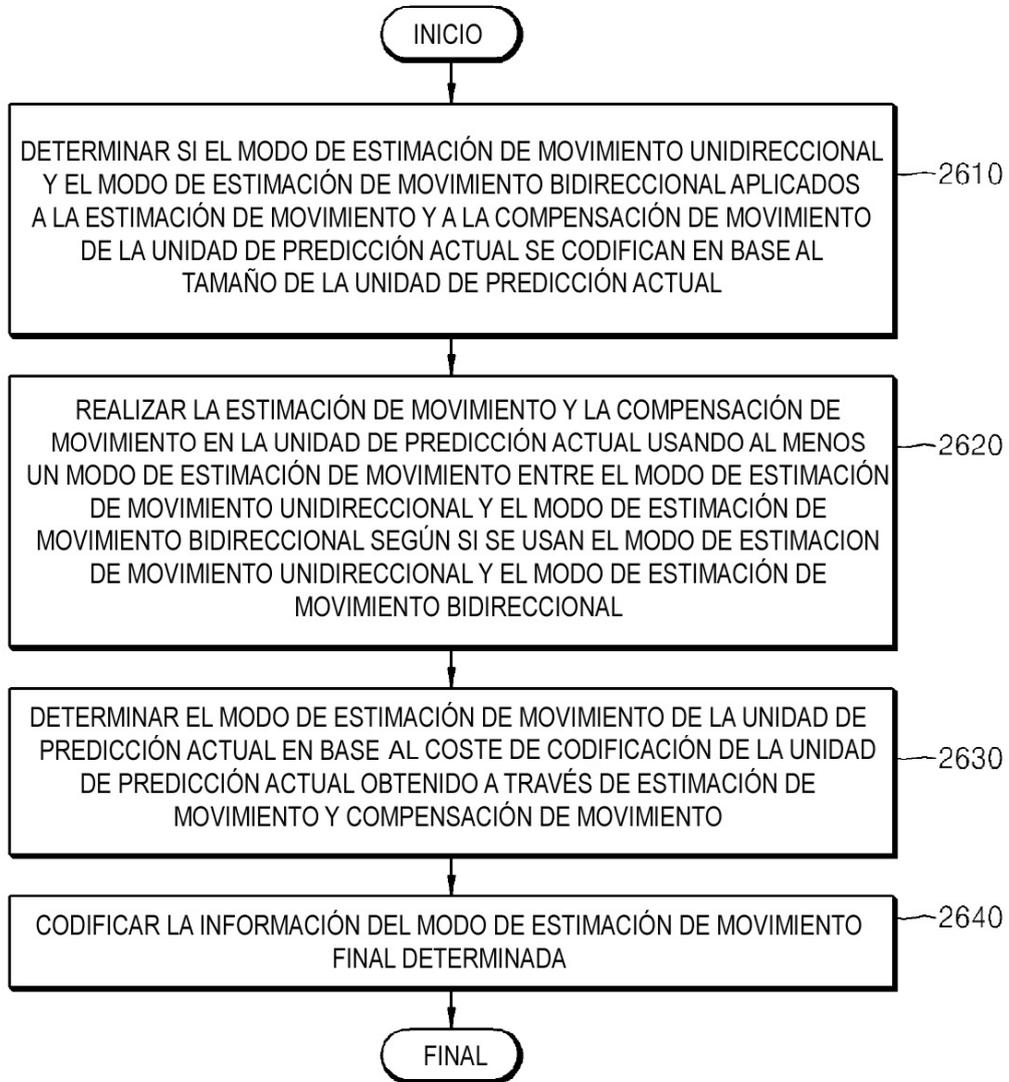


FIG. 27

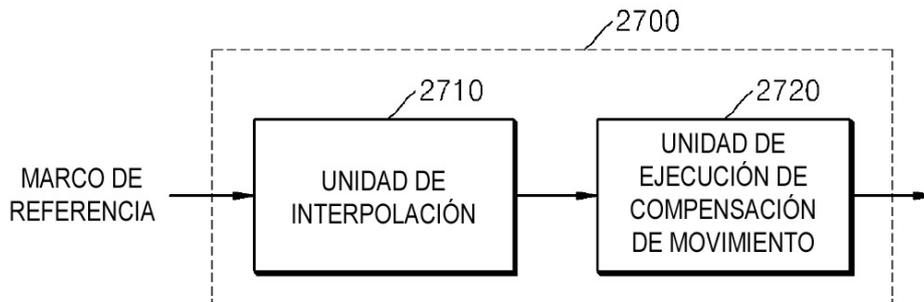


FIG. 28

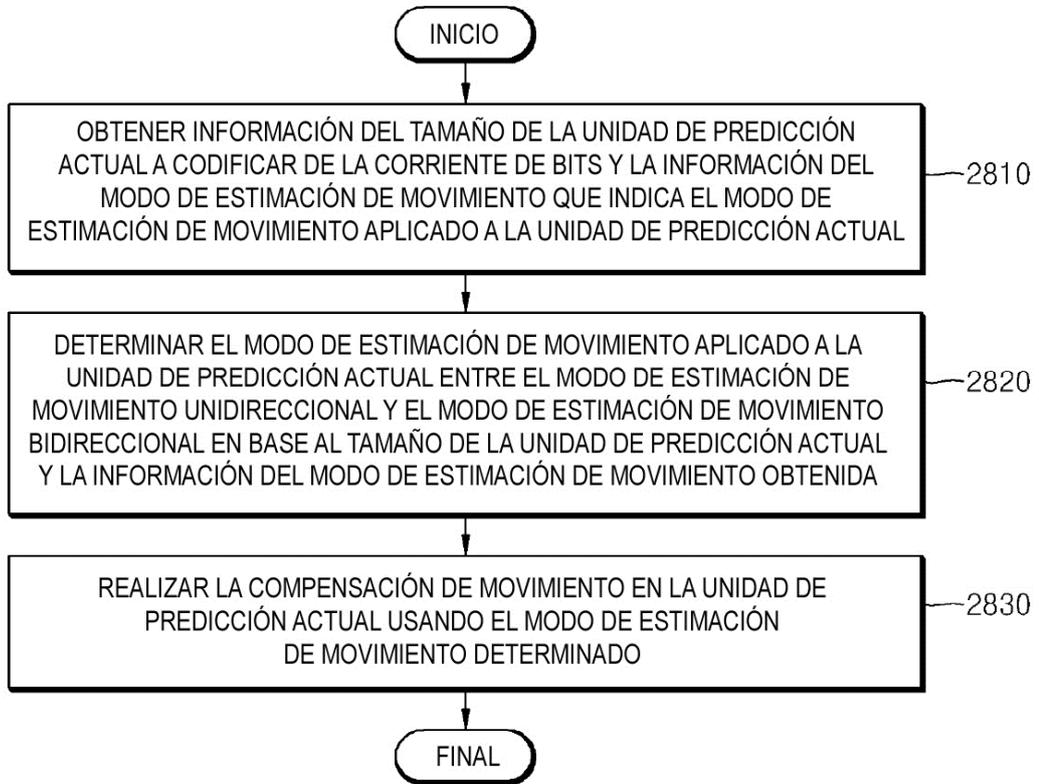


FIG. 29A

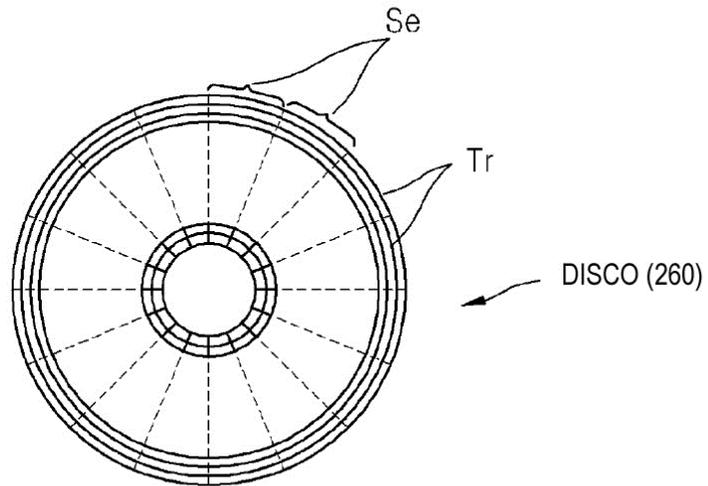
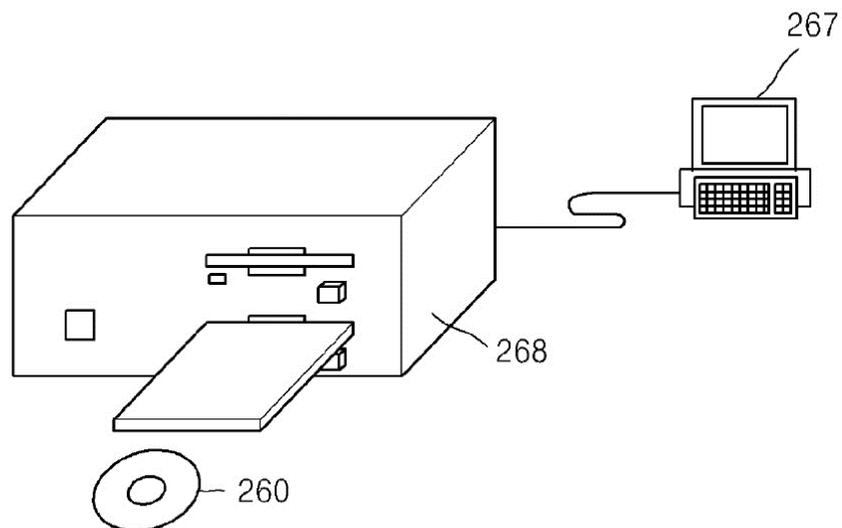


FIG. 29B



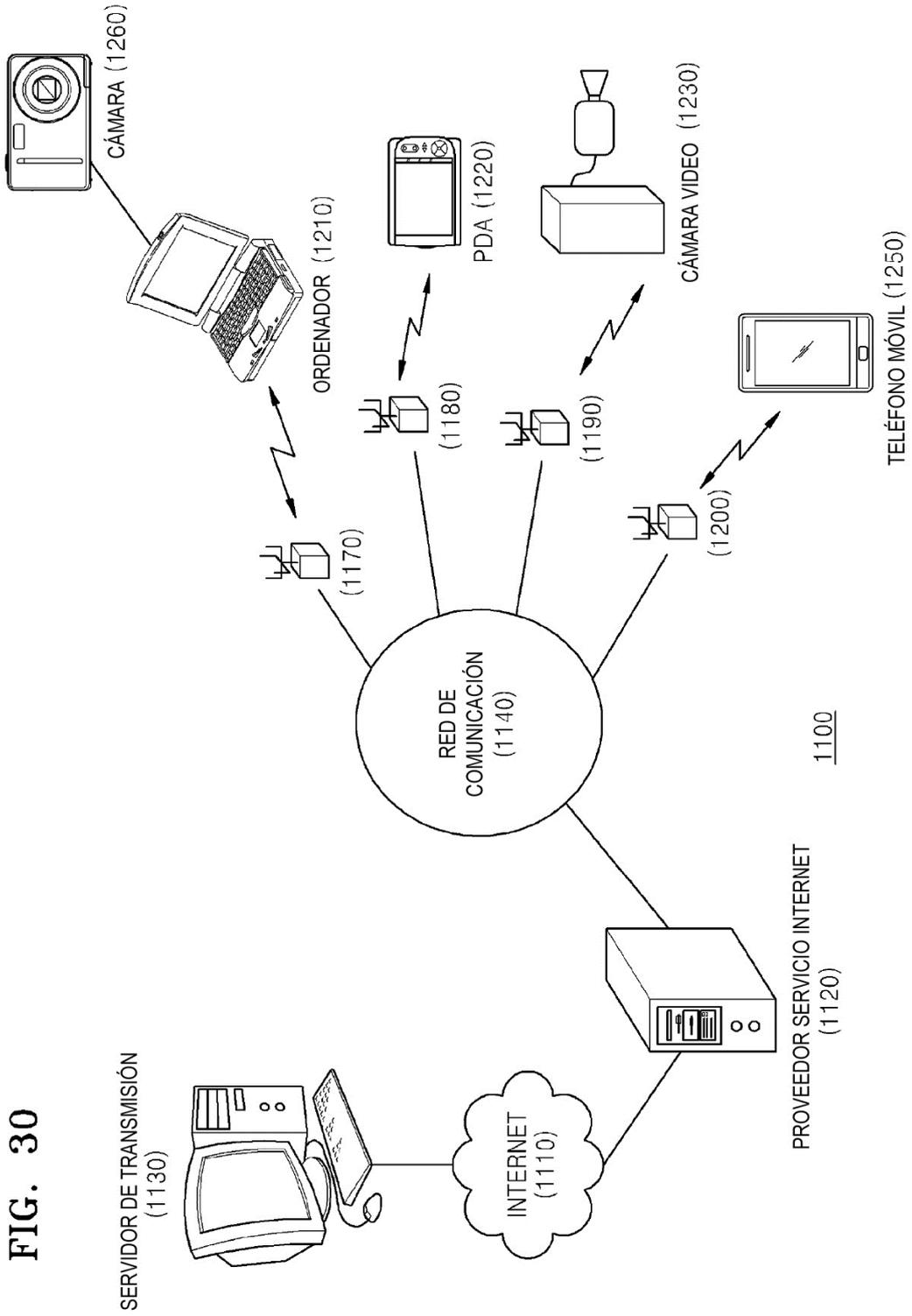
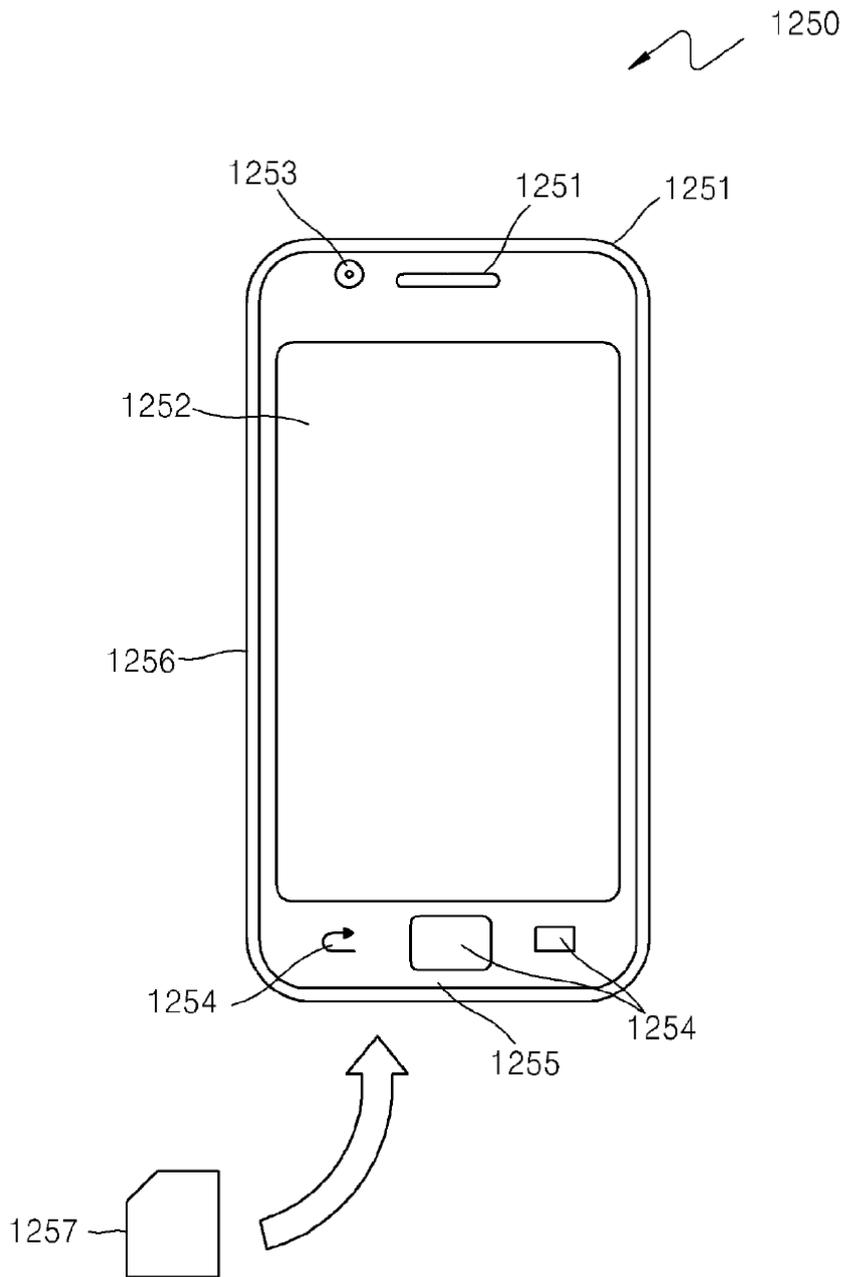


FIG. 30

FIG. 31



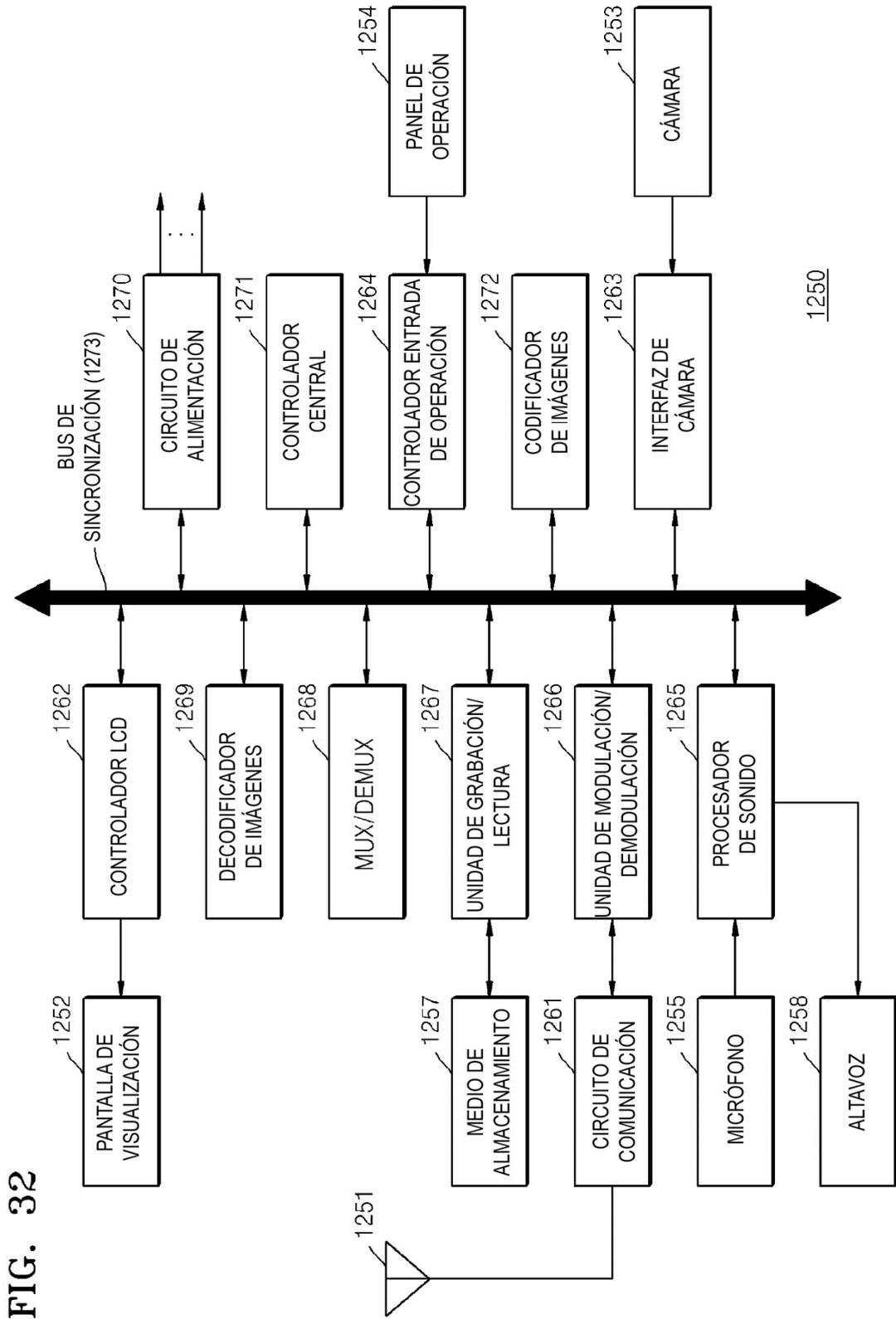
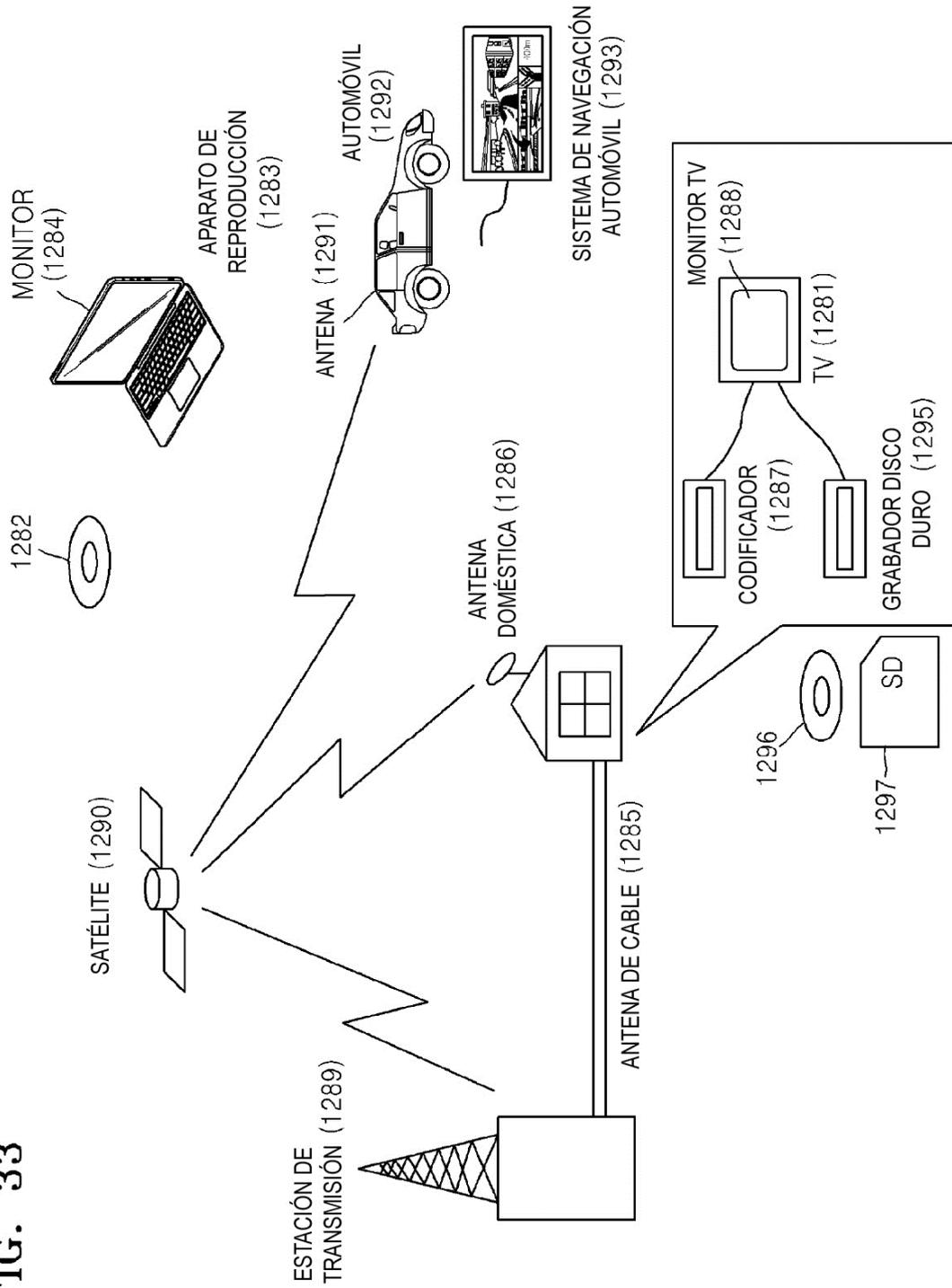


FIG. 32

FIG. 33



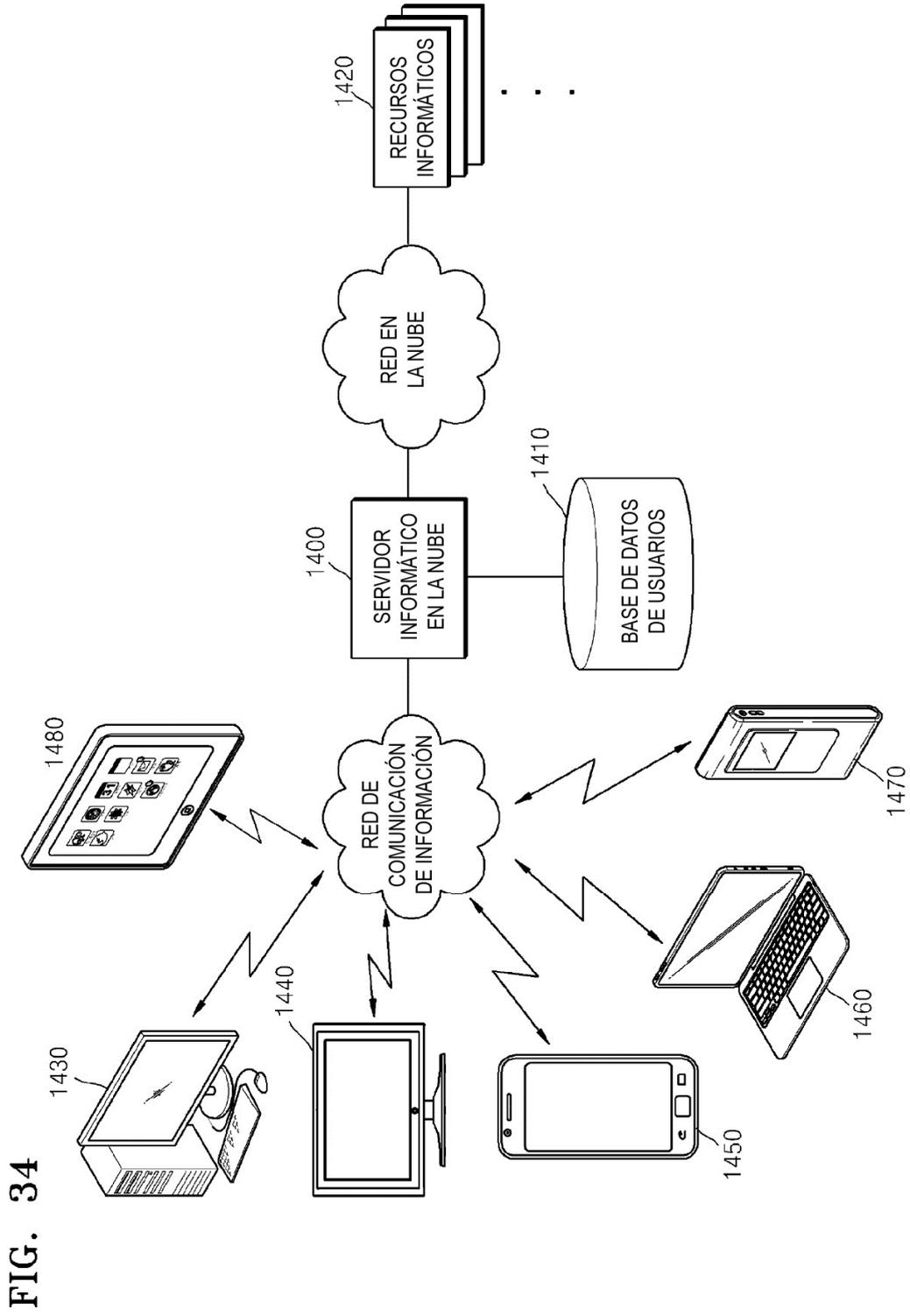


FIG. 34