

19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 668 472**

51 Int. Cl.:

H04N 19/96 (2014.01)

H04N 19/70 (2014.01)

H04N 19/625 (2014.01)

H04N 19/61 (2014.01)

H04N 19/176 (2014.01)

H04N 19/122 (2014.01)

H04N 19/119 (2014.01)

G06T 9/00 (2006.01)

H04N 19/107 (2014.01)

H04N 19/169 (2014.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

96 Fecha de presentación y número de la solicitud europea: **13.08.2010** **E 13167744 (5)**

97 Fecha y número de publicación de la concesión europea: **02.05.2018** **EP 2629526**

54 Título: **Aparato de decodificación de una imagen usando unidades de transformación grandes**

30 Prioridad:

13.08.2009 KR 20090074895

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

18.05.2018

73 Titular/es:

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

72 Inventor/es:

LEE, TAMMY;
HAN, WOO-JIN;
CHEN, JIANLE y
JUNG, HAE-KYUNG

74 Agente/Representante:

CARPINTERO LÓPEZ, Mario

ES 2 668 472 T3

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

DESCRIPCIÓN

Aparato de decodificación de una imagen usando unidades de transformación grandes

[Campo técnico]

Las realizaciones ejemplares se refieren a un aparato para decodificar una imagen.

5 **[Antecedentes de la técnica]**

Para realizar compresión de imagen, la mayoría de los procedimientos y aparatos de codificación y decodificación de imágenes codifican una imagen transformando una imagen de un dominio de píxel en coeficientes de un dominio de frecuencia. Una transformada de coseno discreta (DCT), que es una de las técnicas de transformada de frecuencia, es una técnica bien conocida que se usa ampliamente en compresión de imagen o sonido. Un procedimiento de codificación de imagen que usa la DCT implica realizar la DCT en una imagen de un dominio de píxel, generar coeficientes de coseno discretos, cuantificar los coeficientes de coseno discretos generados, y realizar codificación por entropía en los coeficientes de coseno discretos generados.

10 En la norma de codificación de vídeo H.264/AVC (véase Wiegand y col: "Overview of the H.264/ AVC video coding standard", IEEE Transactions on Circuits and Systems for Video Technology, vol. 13, n.º 7, 1 de julio 2003, páginas 560-576) una imagen se divide en macrobloques de tamaño 16x16. Para predicción de movimiento compensado se soportan particiones de macrobloques de tamaño 16x16, 16x8, 8x16 y 8x8 y las particiones de 8x8 pueden dividirse adicionalmente en particiones de 8x4, 4x8 o 4x4 si se indican por un elemento de sintaxis. Se aplica la transformada de números enteros de 4x4 de tamaño fijo independientemente del tamaño de las particiones usadas para predicción.

20 **[Divulgación]**

A continuación, la expresión "realización ejemplar" debería entenderse como que significa "ejemplo".

[Solución técnica]

Las realizaciones ejemplares proporcionan un aparato para decodificar una imagen.

[Efectos ventajosos]

25 De acuerdo con la una o más realizaciones ejemplares, es posible establecer la unidad de transformación para que sea mayor que la unidad de predicción, y para realizar la DCT, de modo que una imagen pueda comprimirse y codificarse de manera eficaz.

[Descripción de los dibujos]

30 Las anteriores y otras características de las realizaciones ejemplares se harán más evidentes describiendo en detalle realizaciones ejemplares de las mismas con referencia a los dibujos adjuntos en los que:

La Figura 1 es un diagrama de bloques de un aparato de codificación de imagen de acuerdo con una realización ejemplar;

La Figura 2 es un diagrama de un aparato de decodificación de imagen de acuerdo con otra realización ejemplar;

La Figura 3 es un diagrama de una unidad de codificación jerárquica de acuerdo con otra realización ejemplar;

35 La Figura 4 es un diagrama de bloques de un codificador de imagen basado en una unidad de codificación de acuerdo con otra realización ejemplar;

La Figura 5 es un diagrama de bloques de un decodificador de imagen basado en una unidad de codificación de acuerdo con otra realización ejemplar;

40 La Figura 6 ilustra una unidad de codificación máxima, sub-unidades de codificación y unidades de predicción de acuerdo con otra realización ejemplar;

La Figura 7 es un diagrama de una unidad de codificación y una unidad de transformación de acuerdo con otra realización ejemplar;

Las Figuras 8A y 8B ilustran formas de división de una unidad de codificación máxima, una unidad de predicción y una unidad de transformación de acuerdo con otra realización ejemplar;

45 La Figura 9 es un diagrama de bloques de un aparato de codificación de imagen de acuerdo con otra realización ejemplar;

La Figura 10 es un diagrama del transformador;

Las Figuras 11A a 11C ilustran tipos de una unidad de transformación de acuerdo con otra realización ejemplar;

La Figura 12 ilustra diferentes unidades de transformación de acuerdo con otra realización ejemplar;

5 La Figura 13 es un diagrama de bloques de un aparato de decodificación de imagen de acuerdo con otra realización ejemplar; y

La Figura 14 es un diagrama de flujo de un procedimiento de codificación de imagen, de acuerdo con una realización ejemplar.

La Figura 15 es un diagrama de flujo de un procedimiento de decodificación de imagen, de acuerdo con otra realización ejemplar.

10 **[Mejor modo]**

De acuerdo con una realización ejemplar, se proporciona un aparato de decodificación de imagen como se expone en la reivindicación adjunta 1.

[Modo para la invención]

15 En lo sucesivo, las realizaciones ejemplares se describirán en detalle con referencia a los dibujos adjuntos. En las realizaciones ejemplares, “unidad” puede o no hacer referencia a una unidad de tamaño, dependiendo de su contexto e “imagen” puede indicar una imagen fija para un vídeo o una imagen en movimiento, es decir, el mismo vídeo.

La Figura 1 es un diagrama de bloques de un aparato 100 para codificar una imagen, de acuerdo con una realización ejemplar.

20 Haciendo referencia a la Figura 1, el aparato 100 incluye una unidad 110 de división de unidad de codificación máxima, una unidad 120 de determinación de profundidad de codificación, un codificador 130 de datos de imagen y un codificador 140 de información de codificación.

25 La unidad 110 de división de unidad de codificación máxima puede dividir una instantánea o corte actual basándose en una unidad de codificación máxima que es una unidad de codificación del tamaño más grande. Es decir, la unidad 110 de división de unidad de codificación máxima puede dividir la instantánea o corte actual para obtener al menos una unidad de codificación máxima.

30 De acuerdo con una realización ejemplar, una unidad de codificación puede representarse usando una unidad de codificación máxima y una profundidad. Como se ha descrito anteriormente, la unidad de codificación máxima indica una unidad de codificación que tiene el tamaño más grande entre las unidades de codificación de la instantánea actual, y la profundidad indica el tamaño de una subunidad de codificación obtenida reduciendo jerárquicamente la unidad de codificación. A medida que una profundidad se hace profunda, una unidad de codificación puede reducirse en tamaño desde una unidad de codificación máxima a una unidad de codificación mínima, en el que una profundidad de la unidad de codificación máxima se define como una profundidad mínima y una profundidad de la unidad de codificación mínima se define como una profundidad máxima. Puesto que el tamaño de una unidad de codificación se reduce desde una unidad de codificación máxima a medida que una profundidad se hace profunda, una subunidad de codificación de una profundidad de orden k puede incluir una pluralidad de subunidades de codificación de una profundidad de orden $(k+n)$ (k y n son números enteros iguales a 0 o mayores que 1).

35 De acuerdo con un aumento del tamaño de una instantánea a codificarse, codificar una imagen en una unidad de codificación mayor puede dar como resultado una relación de compresión de imagen superior. Sin embargo, si se fija una unidad de codificación mayor, una imagen no puede codificarse eficazmente teniendo en cuenta las características de la imagen que cambian de manera continua.

Por ejemplo, cuando se codifica un área suave tal como el mar o el cielo, cuanto más grande es una unidad de codificación, puede aumentar la relación de compresión. Sin embargo, cuando se codifica un área compleja tal como personas o edificios, cuanto menor es una unidad de codificación, más puede aumentar una relación de compresión.

45 Por consiguiente, de acuerdo con una realización ejemplar, una unidad de codificación de imagen máxima diferente y una profundidad máxima diferente se establecen para cada instantánea o corte. Puesto que una profundidad máxima indica el número máximo de veces por las que puede reducirse una unidad de codificación, el tamaño de cada unidad de codificación mínima incluida en una unidad de codificación de imagen máxima puede establecerse de manera variable de acuerdo con una profundidad máxima.

50 La unidad 120 de determinación de profundidad de codificación determina una profundidad máxima. La profundidad máxima puede determinarse basándose en el cálculo del coste Tasa-Distorsión (R-D). La profundidad máxima puede determinarse de manera diferente para cada instantánea o corte o para cada unidad de codificación máxima.

La profundidad máxima determinada se proporciona al codificador 140 de información de codificación, y los datos de imagen de acuerdo con las unidades de codificación máxima se proporcionan al codificador 130 de datos de imagen.

La profundidad máxima indica una unidad de codificación que tiene el tamaño más pequeño, que puede incluirse en una unidad de codificación máxima, es decir, una unidad de codificación mínima. En otras palabras, una unidad de codificación máxima puede dividirse en subunidades de codificación que tienen diferentes tamaños de acuerdo con diferentes profundidades. Esto se describe en detalle más adelante con referencia a las Figuras 8A y 8B. Además, las subunidades de codificación que tienen diferentes tamaños, que se incluyen en la unidad de codificación máxima, pueden predecirse o transformarse basándose en unidades de procesamiento que tienen diferentes tamaños. En otras palabras, el aparato 100 puede realizar una pluralidad de operaciones de procesamiento para codificación de imagen basándose en unidades de procesamiento que tienen diversos tamaños y diversas formas. Para codificar datos de imagen, se realizan operaciones de procesamiento tales como predicción, transformación y codificación por entropía, en las que las unidades de procesamiento que tienen el mismo tamaño pueden usarse para cada operación o las unidades de procesamiento que tienen diferentes tamaños pueden usarse para cada operación.

Por ejemplo, el aparato 100 puede seleccionar una unidad de procesamiento que es diferente de una unidad de codificación para predecir la unidad de codificación.

Cuando el tamaño de una unidad de codificación es $2N \times 2N$ (donde N es un número entero positivo), las unidades de procesamiento para predicción pueden ser $2N \times 2N$, $2N \times N$, $N \times 2N$ y $N \times N$. En otras palabras, puede realizarse predicción de movimiento basándose en una unidad de procesamiento que tiene una forma mediante la cual al menos una de la altura y anchura de una unidad de codificación se divide igualmente en dos. En lo sucesivo, una unidad de procesamiento, que es la base de la predicción, se define como una 'unidad de predicción'.

Un modo de predicción puede ser al menos uno de un intra modo, un inter modo y un modo de salto y un modo de predicción específico puede realizarse para únicamente una unidad de predicción que tiene un tamaño o forma específicos. Por ejemplo, el intra modo puede realizarse para únicamente unidades de predicción que tienen los tamaños de $2N \times 2N$ y $N \times N$ de los cuales la forma es un cuadrado. Además, el modo de salto puede realizarse para únicamente una unidad de predicción que tiene el tamaño de $2N \times 2N$. Si existe una pluralidad de unidades de predicción en una unidad de codificación, el modo de predicción con los mínimos errores de codificación puede seleccionarse después de realizar la predicción para cada unidad de predicción.

Como alternativa, el aparato 100 puede realizar transformación de frecuencia en datos de imagen basándose en una unidad de procesamiento que tiene un tamaño diferente de una unidad de codificación. Para la transformación de frecuencia en la unidad de codificación, la transformación de frecuencia puede realizarse basándose en una unidad de procesamiento que tiene un tamaño igual a o menor que el de la unidad de codificación. En lo sucesivo, una unidad de procesamiento, que es la base de la transformación de frecuencia, se define como una 'unidad de transformación'. La transformación de frecuencia puede ser una Transformada de Coseno Discreta (DCT) o una Transformada de Karhunen Loeve (KLT).

La unidad 120 de determinación de profundidad de codificación puede determinar subunidades de codificación incluidas en una unidad de codificación máxima usando optimización de RD en un multiplicador Langrangiano. En otras palabras, la unidad 120 de determinación de profundidad de codificación puede determinar las formas de una pluralidad de subunidades de codificación divididas desde la unidad de codificación máxima, en las que la pluralidad de subunidades de codificación tienen diferentes tamaños de acuerdo con sus profundidades. El codificador 130 de datos de imagen emite una secuencia de bits codificando la unidad de codificación máxima basándose en las formas de división, es decir, las formas que dividen la unidad de codificación máxima, según se determina por la unidad 120 de determinación de profundidad de codificación.

El codificador 140 de información de codificación codifica información acerca de un modo de codificación de la unidad de codificación máxima determinada por la unidad 120 de determinación de profundidad de codificación. En otras palabras, el codificador 140 de información de codificación emite una secuencia de bits codificando información acerca de una forma de división de la unidad de codificación máxima, información acerca de la profundidad máxima e información acerca de un modo de codificación de una subunidad de codificación para cada profundidad. La información acerca del modo de codificación de la subunidad de codificación puede incluir información acerca de una unidad de predicción de la subunidad de codificación, información acerca de un modo de predicción para cada unidad de predicción e información acerca de una unidad de transformación de la subunidad de codificación.

Puesto que existen subunidades de codificación que tienen diferentes tamaños para cada unidad de codificación máxima y la información acerca de un modo de codificación debe determinarse para cada subunidad de codificación, la información acerca de al menos un modo de codificación puede determinarse para una unidad de codificación máxima.

El aparato 100 puede generar subunidades de codificación dividiendo igualmente tanto la altura como anchura de una unidad de codificación máxima en dos de acuerdo con un aumento de la profundidad. Es decir, cuando el tamaño de una unidad de codificación de una profundidad de orden k es $2N \times 2N$, el tamaño de una unidad de

codificación de una profundidad de orden (k+1) es NxN.

Por consiguiente, el aparato 100 de acuerdo con una realización ejemplar puede determinar una forma de división óptima para cada unidad de codificación máxima basándose en tamaños de unidades de codificación máxima y una profundidad máxima en consideración de las características de imagen. Ajustando de manera variable el tamaño de una unidad de codificación máxima en consideración de las características de imagen y codificando una imagen a través de la división de una unidad de codificación máxima en subunidades de codificación de diferentes profundidades, pueden codificarse más eficazmente imágenes que tienen diversas resoluciones.

La Figura 2 es un diagrama de bloques de un aparato 200 para decodificar una imagen de acuerdo con una realización ejemplar.

Haciendo referencia a la Figura 2, el aparato 200 incluye una unidad 210 de obtención de datos de imagen, una unidad 220 de extracción de información de codificación y un decodificador 230 de datos de imagen.

La unidad 210 de obtención de datos de imagen obtiene datos de imagen de acuerdo con unidades de codificación máxima analizando una secuencia de bits recibida por el aparato 200 y emite los datos de imagen al decodificador 230 de datos de imagen. La unidad 210 de obtención de datos de imagen puede extraer información acerca de una unidad de codificación máxima de una instantánea o corte actual desde un encabezamiento de la instantánea o corte actual. En otras palabras, la unidad 210 de obtención de datos de imagen divide la secuencia de bits en la unidad de codificación máxima de modo que el decodificador 230 de datos de imagen puede decodificar los datos de imagen de acuerdo con las unidades de codificación máxima.

La unidad 220 de extracción de información de codificación extrae información acerca de una unidad de codificación máxima, una profundidad máxima, una forma de división de la unidad de codificación máxima, un modo de codificación de subunidades de codificación desde el encabezamiento de la instantánea actual analizando la secuencia de bits recibida por el aparato 200. La información acerca de una forma de división y la información acerca de un modo de codificación se proporcionan al decodificador 230 de datos de imagen.

La información acerca de una forma de división de la unidad de codificación máxima puede incluir información acerca de subunidades de codificación que tienen diferentes tamaños de acuerdo con las profundidades incluidas en la unidad de codificación máxima, y la información acerca de un modo de codificación puede incluir información acerca de una unidad de predicción de acuerdo con la subunidad de codificación, información acerca de un modo de predicción e información acerca de unas unidades de transformación.

El decodificador 230 de datos de imagen restaura la instantánea actual decodificando datos de imagen de cada unidad de codificación máxima basándose en la información extraída por la unidad 220 de extracción de información de codificación. El decodificador 230 de datos de imagen puede decodificar subunidades de codificación incluidas en una unidad de codificación máxima basándose en la información acerca de una forma de división de la unidad de codificación máxima. Un procedimiento de decodificación puede incluir un procedimiento de predicción que incluye intra predicción y compensación de movimiento y un procedimiento de transformación inversa.

El decodificador 230 de datos de imagen puede realizar intra predicción o inter predicción basándose en información acerca de una unidad de predicción e información acerca de un modo de predicción para predecir una unidad de predicción. El decodificador 230 de datos de imagen puede realizar también transformación inversa para cada subunidad de codificación basándose en información acerca de una unidad de transformación de una subunidad de codificación.

La Figura 3 ilustra unidades de codificación jerárquica de acuerdo con una realización ejemplar.

Haciendo referencia a la Figura 3, las unidades de codificación jerárquica de acuerdo con una realización ejemplar pueden incluir unidades de codificación cuyas anchuras x alturas son 64x64, 32x32, 16x16, 8x8 y 4x4. Aparte de estas unidades de codificación que tienen formas cuadradas perfectas, pueden existir también unidades de codificación cuyas anchuras x alturas son 64x32, 32x64, 32x16, 16x32, 16x8, 8x16, 8x4 y 4x8.

Haciendo referencia a la Figura 3, para datos 310 de imagen cuya resolución es 1920x1080, el tamaño de una unidad de codificación máxima se establece a 64x64 y una profundidad máxima se establece a 2.

Para datos 320 de imagen cuya resolución es 1920x1080, el tamaño de una unidad de codificación máxima se establece a 64x64, y una profundidad máxima se establece a 4. Para datos 330 de imagen cuya resolución es 352x288, el tamaño de una unidad de codificación máxima se establece a 16x16 y una profundidad máxima se establece a 1.

Cuando la resolución es alta o la cantidad de datos es grande, se proporciona, aunque no es necesario, que un tamaño máximo de una unidad de codificación sea relativamente grande para aumentar una relación de compresión y refleje exactamente las características de imagen. Por consiguiente, para los datos 310 y 320 de imagen que tienen resolución superior a los datos 330 de imagen, puede seleccionarse 64x64 como el tamaño de una unidad de codificación máxima.

Una profundidad máxima indica el número total de capas en las unidades de codificación jerárquica. Puesto que la profundidad máxima de los datos 310 de imagen es 2, una unidad 315 de codificación de los datos 310 de imagen puede incluir una unidad de codificación máxima cuyo tamaño de eje más largo es 64 y subunidades de codificación cuyos tamaños de eje más largos son 32 y 16, de acuerdo con un aumento de una profundidad.

- 5 Por otra parte, puesto que la profundidad máxima de los datos 330 de imagen es 1, una unidad 335 de codificación de los datos 330 de imagen puede incluir una unidad de codificación máxima cuyo tamaño de eje más largo es 16 y unidades de codificación cuyos tamaños de eje más largo son 8, de acuerdo con un aumento de una profundidad.

- 10 Sin embargo, puesto que la profundidad máxima de los datos 320 de imagen es 4, una unidad 325 de codificación de los datos 320 de imagen puede incluir una unidad de codificación máxima cuyo tamaño de eje más largo es 64 y subunidades de codificación cuyos tamaños de eje más largos son 32, 16, 8 y 4 de acuerdo con un aumento de una profundidad. Puesto que se codifica una imagen basándose en una subunidad de codificación más pequeña a medida que una profundidad se hace profunda, la realización ejemplar es adecuada para codificar una imagen que incluye más detalles minuciosos en las escenas.

- 15 La Figura 4 es un diagrama de bloques de un codificador 400 de imagen basado en una unidad de codificación, de acuerdo con una realización ejemplar.

Un intra predictor 410 realiza intra predicción en unidades de predicción del intra modo en un fotograma 405 actual, y una unidad 420 de estimación de movimiento y una unidad 425 de compensación de movimiento realizan inter predicción y compensación de movimiento en unidades de predicción del inter modo usando el fotograma 405 actual y un fotograma 495 de referencia.

- 20 Se generan valores residuales basándose en las unidades de predicción emitidas desde el intra predictor 410, la unidad 420 de estimación de movimiento y la unidad 425 de compensación de movimiento, y los valores residuales generados se emiten como coeficientes de transformada cuantificados pasando a través de un transformador 430 y una unidad 440 de cuantificación.

- 25 Los coeficientes de transformada cuantificados se restauran a valores residuales pasando a través de una unidad 460 de cuantificación inversa y un transformador 470 inverso de frecuencia, y los valores residuales restaurados se procesan posteriormente pasando a través de una unidad 480 de desbloqueo y una unidad 490 de filtración de bucle y se emiten como el fotograma 495 de referencia. Los coeficientes de transformada cuantificados pueden emitirse como una secuencia de bits 455 pasando a través de un codificador 450 por entropía.

- 30 Para realizar codificación basándose en un procedimiento de codificación de acuerdo con una realización ejemplar, los componentes del codificador 400 de imagen, es decir, el intra predictor 410, la unidad 420 de estimación de movimiento, la unidad 425 de compensación de movimiento, el transformador 430, la unidad 440 de cuantificación, el codificador 450 por entropía, la unidad 460 de cuantificación inversa, el transformador 470 inverso de frecuencia, la unidad 480 de desbloqueo y la unidad 490 de filtración en bucle, realizan procedimientos de codificación de imagen basándose en una unidad de codificación máxima, una subunidad de codificación de acuerdo con las profundidades, una unidad de predicción y una unidad de transformación.

- 35 La Figura 5 es un diagrama de bloques de un decodificador 500 de imagen basado en una unidad de codificación, de acuerdo con una realización ejemplar.

- 40 Una secuencia de bits 505 pasa a través de una unidad 510 de análisis de modo que se analiza los datos de imagen codificados a decodificarse y la información de codificación necesaria para decodificación. Los datos de imagen codificados se emiten como datos cuantificados a la inversa pasando a través de un decodificador 520 por entropía y una unidad 530 de cuantificación inversa y se restauran a valores residuales pasando a través de un transformador 540 inverso de frecuencia. Los valores residuales se restauran de acuerdo con unidades de codificación añadiéndose a un resultado de intra predicción de un intra predictor 550 o un resultado de compensación de movimiento de una unidad 560 de compensación de movimiento. Las unidades de codificación restauradas se usan para predicción de las siguientes unidades de codificación o de una siguiente instantánea pasando a través de una unidad 570 de desbloqueo y una unidad 580 de filtración de bucle.

- 50 Para realizar la decodificación basándose en un procedimiento de decodificación de acuerdo con una realización ejemplar, los componentes del decodificador 500 de imagen, es decir, la unidad 510 de análisis, el decodificador 520 por entropía, la unidad 530 de cuantificación inversa, el transformador 540 inverso de frecuencia, el intra predictor 550, la unidad 560 de compensación de movimiento, la unidad 570 de desbloqueo y la unidad de filtrado en bucle 580, realizan procedimientos de decodificación de imagen basándose en una unidad de codificación máxima, una subunidad de codificación de acuerdo con las profundidades, una unidad de predicción y una unidad de transformación.

- 55 En particular, el intra predictor 550 y la unidad 560 de compensación de movimiento determinan una unidad de predicción y un modo de predicción en una subunidad de codificación considerando una unidad de codificación máxima y una profundidad, y el transformador 540 inverso de frecuencia realiza transformación inversa considerando el tamaño de una unidad de transformación.

La Figura 6 ilustra una unidad de codificación máxima, una subunidad de codificación y una unidad de predicción, de acuerdo con una realización ejemplar.

El aparato 100 y el aparato 200 de acuerdo con una realización ejemplar usan unidades de codificación jerárquica para realizar codificación y decodificación en consideración de las características de imagen. Una unidad de codificación máxima y una profundidad máxima pueden establecerse de manera adaptativa de acuerdo con las características de imagen o establecerse de manera variable de acuerdo con requisitos de un usuario.

Una estructura 600 de unidad de codificación jerárquica de acuerdo con una realización ejemplar ilustra una unidad 610 de codificación máxima cuya altura y anchura son 64 y la profundidad máxima es 4. Una profundidad aumenta a lo largo de un eje vertical de la estructura 600 de unidad de codificación jerárquica, y a medida que una profundidad se hace profunda, se reducen las alturas y anchuras de las subunidades 620 a 650 de codificación. Las unidades de predicción de la unidad 610 de codificación máxima y las subunidades 620 a 650 de codificación se muestran a lo largo de un eje horizontal de la estructura 600 de unidad de codificación jerárquica.

La unidad 610 de codificación máxima tiene una profundidad de 0 y el tamaño de una unidad de codificación, es decir, altura y anchura, de 64x64. Una profundidad aumenta a lo largo del eje vertical, y existe una subunidad 620 de codificación cuyo tamaño es 32x32 y profundidad es 1, una subunidad 630 de codificación cuyo tamaño es 16x16 y profundidad es 2, una subunidad 640 de codificación cuyo tamaño es 8x8 y profundidad es 3, y una subunidad 650 de codificación cuyo tamaño es 4x4 y profundidad es 4. La subunidad 650 de codificación cuyo tamaño es 4x4 y profundidad es 4 es una unidad de codificación mínima, y la unidad de codificación mínima puede dividirse en unidades de predicción, cada una de las cuales es menor que la unidad de codificación mínima.

Haciendo referencia a la Figura 6, se muestran ejemplos de una unidad de predicción a lo largo del eje horizontal de acuerdo con cada profundidad. Es decir, una unidad de predicción de la unidad 610 de codificación máxima cuya profundidad es 0 puede ser una unidad de predicción cuyo tamaño es igual a la unidad 610 de codificación, es decir, 64x64, o una unidad 612 de predicción cuyo tamaño es 64x32, una unidad 614 de predicción cuyo tamaño es 32x64, o una unidad 616 de predicción cuyo tamaño es 32x32, que tiene un tamaño menor que la unidad 610 de codificación cuyo tamaño es 64x64.

Una unidad de predicción de la unidad 620 de codificación cuya profundidad es 1 y el tamaño es 32x32 puede ser una unidad de predicción cuyo tamaño es igual a la unidad 620 de codificación, es decir, 32x32, o una unidad 622 de predicción cuyo tamaño es 32x16, una unidad 624 de predicción cuyo tamaño es 16x32, o una unidad 626 de predicción cuyo tamaño es 16x16, que tiene un tamaño menor que la unidad 620 de codificación cuyo tamaño es 32x32.

Una unidad de predicción de la unidad 630 de codificación cuya profundidad es 2 y el tamaño es 16x16 puede ser una unidad de predicción cuyo tamaño es igual a la unidad 630 de codificación, es decir, 16x16, o una unidad 632 de predicción cuyo tamaño es 16x8, una unidad 634 de predicción cuyo tamaño es 8x16, o una unidad 636 de predicción cuyo tamaño es 8x8, que tiene un tamaño menor que la unidad 630 de codificación cuyo tamaño es 16x16.

Una unidad de predicción de la unidad 640 de codificación cuya profundidad es 3 y el tamaño es 8x8 puede ser una unidad de predicción cuyo tamaño es igual a la unidad 640 de codificación, es decir, 8x8, o una unidad 642 de predicción cuyo tamaño es 8x4, una unidad 644 de predicción cuyo tamaño es 4x8, o una unidad 646 de predicción cuyo tamaño es 4x4, que tiene un tamaño menor que la unidad 640 de codificación cuyo tamaño es 8x8.

Finalmente, la unidad 650 de codificación cuya profundidad es 4 y el tamaño es 4x4 es una unidad de codificación mínima y una unidad de codificación de una profundidad máxima, y una unidad de predicción de la unidad 650 de codificación puede ser una unidad 650 de predicción cuyo tamaño es 4x4, una unidad 652 de predicción que tiene un tamaño de 4x2, una unidad 654 de predicción que tiene un tamaño de 2x4, o una unidad 656 de predicción que tiene un tamaño de 2x2.

La Figura 7 ilustra una unidad de codificación y una unidad de transformación, de acuerdo con una realización ejemplar.

El aparato 100 y el aparato 200, de acuerdo con una realización ejemplar, realizan codificación con una misma unidad de codificación máxima o con subunidades de codificación, que son iguales a o menores que la unidad de codificación máxima, divididas a partir de la unidad de codificación máxima.

En el procedimiento de codificación, el tamaño de una unidad de transformación para transformación de frecuencia se selecciona para que no sea mayor que el de una unidad de codificación correspondiente. Por ejemplo, cuando una unidad 710 de codificación tiene el tamaño de 64x64, puede realizarse transformación de frecuencia usando una unidad 720 de transformación que tiene el tamaño de 32x32.

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

La Figura 8A ilustra una unidad de codificación y una unidad de predicción, de acuerdo con una realización ejemplar.

El lado izquierdo de la Figura 8A muestra una forma de división seleccionada por el aparato 100, de acuerdo con una realización ejemplar, para codificar una unidad 810 de codificación máxima. El aparato 100 divide la unidad 810 de codificación máxima en diversas formas, realiza codificación y selecciona una forma de división óptima comparando resultados de codificación de diversas formas de división entre sí basándose en el coste de R-D. Cuando es óptimo codificar la unidad 810 de codificación máxima como está, la unidad 810 de codificación máxima puede codificarse sin dividir la unidad 810 de codificación máxima como se ilustra en las Figuras 8A y 8B.

Haciendo referencia al lado izquierdo de la Figura 8A, la unidad 810 de codificación máxima cuya profundidad es 0 se codifica dividiéndola en subunidades de codificación cuyas profundidades son igual a o mayores que 1. Es decir, la unidad 810 de codificación máxima se divide en 4 subunidades de codificación cuyas profundidades son 1, y todas o algunas de las subunidades de codificación cuyas profundidades son 1 se dividen en subunidades de codificación cuyas profundidades son 2.

Una subunidad de codificación localizada en un lado superior derecha y una subunidad de codificación localizada en un lado inferior izquierda entre las subunidades de codificación cuyas profundidades son 1 se dividen en subunidades de codificación cuyas profundidades son igual a o mayores que 2. Algunas de las subunidades de codificación cuyas profundidades son igual a o mayores que 2 pueden dividirse en subunidades de codificación cuyas profundidades son iguales a o mayores que 3.

El lado derecho de la Figura 8A muestra una forma de división de una unidad de predicción para la unidad 810 de codificación máxima.

Haciendo referencia al lado derecho de la Figura 8A, una unidad 860 de predicción para la unidad 810 de codificación máxima puede dividirse de manera diferente de la unidad 810 de codificación máxima. En otras palabras, una unidad de predicción para cada una de las subunidades de codificación puede ser menor que una correspondiente subunidad de codificación.

Por ejemplo, una unidad de predicción para una subunidad 854 de codificación localizada en un lado derecho inferior entre las subunidades de codificación cuyas profundidades son 1 puede ser menor que la subunidad 854 de codificación. Además, las unidades de predicción para algunas (814, 816, 850 y 852) de las subunidades 814, 816, 818, 828, 850 y 852 de codificación cuyas profundidades son 2 pueden ser menores que las subunidades 814, 816, 850, y 852 de codificación, respectivamente. Además, las unidades de predicción para las subunidades 822, 832 y 848 de codificación cuyas profundidades son 3 pueden ser menores que las subunidades 822, 832 y 848 de codificación, respectivamente. Las unidades de predicción pueden tener una forma mediante la cual las respectivas subunidades de codificación se dividen igualmente en dos en una dirección de altura o anchura o tienen una forma mediante la cual las respectivas subunidades de codificación se dividen igualmente en cuatro en direcciones de altura y anchura.

La Figura 8B ilustra una unidad de predicción y una unidad de transformación, de acuerdo con una realización ejemplar.

Un lado izquierdo de la Figura 8B muestra una forma de división de una unidad de predicción para la unidad 810 de codificación máxima mostrada en el lado derecho de la Figura 8A, y un lado derecho de la Figura 8B muestra una forma de división de una unidad de transformación de la unidad 810 de codificación máxima.

Haciendo referencia al lado derecho de la Figura 8B, una forma de división de una unidad 870 de transformación puede establecerse de manera diferente de la unidad 860 de predicción.

Por ejemplo, incluso aunque una unidad de predicción para la unidad 854 de codificación cuya profundidad es 1 se seleccione con una forma mediante la cual la altura de la unidad 854 de codificación se divide igualmente en dos, una unidad de transformación puede seleccionarse con el mismo tamaño que la unidad 854 de codificación. Análogamente, incluso aunque las unidades de predicción para las unidades 814 y 850 de codificación cuyas profundidades son 2 se seleccionen con una forma mediante la cual la altura de cada una de las unidades 814 y 850 de codificación se divide igualmente en dos, una unidad de transformación puede seleccionarse con el mismo tamaño que el tamaño original de cada una de las unidades 814 y 850 de codificación.

Una unidad de transformación puede seleccionarse con un tamaño menor que una unidad de predicción. Por ejemplo, cuando una unidad de predicción para la unidad 852 de codificación cuya profundidad es 2 se selecciona con una forma mediante la cual la anchura de la unidad 852 de codificación se divide igualmente en dos, una unidad de transformación puede seleccionarse con una forma mediante la cual la unidad 852 de codificación se divide igualmente en cuatro en direcciones de altura y anchura, que tienen un tamaño menor que la forma de la unidad de predicción.

La Figura 9 es un diagrama de bloques de un aparato 900 de codificación de imagen de acuerdo con otra realización ejemplar.

Haciendo referencia a la Figura 9, el aparato 900 de codificación de imagen de acuerdo con la presente realización ejemplar incluye un transformador 910, una unidad 920 de cuantificación y un codificador 930 por entropía.

El transformador 910 recibe una unidad de procesamiento de imagen de un dominio de píxel, y transforma la unidad de procesamiento de imagen en un dominio de frecuencia. El transformador 910 recibe una pluralidad de unidades de predicción que incluyen valores residuales generados debido a intra-predicción o inter-predicción, y transforma las unidades de predicción en un dominio de frecuencia. Como resultado de la transformada al dominio de la frecuencia, se generan coeficientes de componentes de frecuencia. De acuerdo con la presente realización ejemplar, la transformada al dominio de la frecuencia puede tener lugar mediante una transformada de coseno discreta (DCT) o Transformada de Karhunen Loeve (KLT), y como resultado de la DCT o KLT, se generan coeficientes del dominio de la frecuencia. En lo sucesivo, la transformada al dominio de la frecuencia puede ser la DCT, sin embargo, es evidente para un experto en la materia que la transformada al dominio de la frecuencia puede ser cualquier transformada que implica la transformación de una imagen de un dominio de píxel en un dominio de frecuencia.

También, de acuerdo con la presente realización ejemplar, el transformador 910 establece una unidad de transformación agrupando una pluralidad de unidades de predicción, y realiza la transformación de acuerdo con la unidad de transformación. Este procedimiento se describirá en detalle con referencia a las Figuras 10, 11A, 11B y 12.

La Figura 10 es un diagrama del transformador 910.

Haciendo referencia a la Figura 10, el transformador 910 incluye una unidad 1010 de selección y una unidad 1020 de realización de transformada.

La unidad 1010 de selección establece una unidad de transformación seleccionando una pluralidad de unidades de predicción adyacentes.

Un aparato de codificación de imagen de acuerdo con la técnica relacionada realiza intra-predicción o inter-predicción de acuerdo con un bloque que tiene un tamaño predeterminado, es decir, de acuerdo con una unidad de predicción, y realiza la DCT basándose en un tamaño que es menor que o igual al de la unidad de predicción. En otras palabras, el aparato de codificación de imagen de acuerdo con la técnica relacionada realiza la DCT usando unidades de transformación que son menores que o iguales a la unidad de predicción.

Sin embargo, debido a una pluralidad de piezas de información de encabezamiento añadidas a las unidades de transformación, se aumentan taras añadidas a medida que se reducen las unidades de transformación, de manera que se deteriora una tasa de compresión de una operación de codificación de imagen. Para resolver este problema, el aparato 900 de codificación de imagen de acuerdo con la presente realización ejemplar agrupa una pluralidad de unidades de predicción adyacentes en una unidad de transformación, y realiza la transformación de acuerdo con la unidad de transformación que se genera por el agrupamiento. Existe una alta posibilidad de que las unidades de predicción adyacentes puedan incluir valores residuales similares, de modo que, si las unidades de predicción adyacentes se agrupan en una unidad de transformación y a continuación la transformación se realiza en la misma, una tasa de compresión de una operación de codificación puede aumentarse de manera elevada.

Para este aumento, la unidad 1010 de selección selecciona las unidades de predicción adyacentes para que se agrupen en una unidad de transformación. Este procedimiento se describirá en detalle con referencia a las Figuras 11A a 11C y 12.

Las Figuras 11A a 11C ilustran tipos de una unidad de transformación de acuerdo con otra realización ejemplar.

Haciendo referencia a las Figuras 11A a 11C, una unidad 1120 de predicción con respecto a una unidad 1110 de codificación puede tener una forma de división obtenida dividiendo por la mitad una anchura de la unidad 1110 de codificación. La unidad 1110 de codificación puede ser una unidad de codificación máxima, o puede ser una sub-unidad de codificación que tiene un tamaño menor que la unidad de codificación máxima.

Como se ilustra en la Figura 11A, un tamaño de la unidad 1130 de transformación puede ser menor que la unidad 1120 de predicción, o como se ilustra en la Figura 11B, un tamaño de la unidad 1140 de transformación puede ser igual a la unidad 1120 de predicción. También, como se ilustra en la Figura 11C, un tamaño de la unidad 1150 de transformación pueden ser mayor que la unidad 1120 de predicción. Es decir, las unidades 1130 a 1150 de transformación pueden establecerse mientras que no tengan conexión con la unidad 1120 de predicción.

También, la Figura 11C ilustra un ejemplo en el que la unidad 1120 de predicción se establece agrupando una pluralidad de las unidades 1120 de predicción incluidas en la unidad 1110 de codificación. Sin embargo, una unidad de transformación puede establecerse para que sea mayor que una unidad de codificación de una manera en la que se establece una pluralidad de unidades de predicción, que no están incluidas en una unidad de codificación pero en una pluralidad de unidades de codificación, como una unidad de transformación. En otras palabras, como se describe con referencia a las Figuras 11A a 11C, una unidad de transformación puede establecerse para que sea igual a o menor que un tamaño de una unidad de codificación, o para que sea mayor que el tamaño de la unidad de

codificación. Es decir, la unidad de transformación puede establecerse mientras que no se tenga conexión con la unidad de predicción y la unidad de codificación.

Aunque las Figuras 11A a 11C ilustran ejemplos en los que la unidad de transformación tiene una forma cuadrada. Sin embargo, de acuerdo con un procedimiento de agrupación de unidades de predicción adyacentes, la unidad de transformación puede tener una forma rectangular. Por ejemplo, en un caso donde la unidad de predicción no se establece para que tenga formas rectangulares como se ilustra en las Figuras 11A a 11C pero se establece para que tenga cuatro formas cuadradas obtenidas seccionando en cuatro la unidad 1110 de codificación, se agrupan las unidades de predicción superior e inferior, o las unidades de predicción izquierda y derecha de modo que la unidad de transformación puede tener una forma rectangular cuyo lado horizontal o lado vertical es largo.

Haciendo referencia de vuelta a la Figura 10, no hay límite en un criterio mediante el cual la unidad 1010 de selección seleccione las unidades de predicción adyacentes. Sin embargo, de acuerdo con la realización ejemplar, la unidad 1010 de selección puede seleccionar la unidad de transformación de acuerdo con una profundidad. Como se ha descrito anteriormente, la profundidad indica un nivel de tamaño de reducción que se realiza gradualmente de una unidad de codificación máxima de un corte actual o una instantánea actual a una sub-unidad de codificación. Como se ha descrito con referencia a las Figuras 3 y 6, a medida que se aumenta la profundidad, se reduce un tamaño de una sub-unidad de codificación, y por lo tanto una unidad de predicción incluida en la sub-unidad de codificación también se reduce. En este caso, si la transformación se realiza de acuerdo con una unidad de transformación que es menor que o igual a la unidad de predicción, se deteriora una tasa de compresión de una operación de codificación de imagen puesto que se añade información de encabezamiento a cada unidad de transformación.

Por lo tanto, con respecto a una sub-unidad de codificación a una profundidad de un valor predeterminado, es preferible, pero no necesario, que las unidades de predicción incluidas en la sub-unidad de codificación se agrupen y se establezcan como una unidad de transformación, y a continuación la transformación se realiza en la mismas. Para esto, la unidad 1010 de selección establece la unidad de transformación basándose en la profundidad de la sub-unidad de codificación. Por ejemplo, en el caso donde una profundidad de la unidad 1110 de codificación en la Figura 11C sea mayor que k , la unidad 1010 de selección agrupa las unidades 1120 de predicción y las establece como una unidad 1150 de transformación.

También, de acuerdo con otra realización ejemplar, la unidad 1010 de selección puede agrupar una pluralidad de unidades de predicción adyacentes en las que se realiza predicción de acuerdo con el mismo modo de predicción, y puede establecerlas como una unidad de transformación. La unidad 1010 de selección agrupa las unidades de predicción adyacentes en las que se realiza predicción de acuerdo con intra-predicción o inter-predicción, y a continuación las establece como una unidad de transformación. Puesto que hay una alta posibilidad de que las unidades de predicción adyacentes en las que se realiza predicción de acuerdo con el mismo modo de predicción incluyan valores residuales similares, es posible agrupar las unidades de predicción adyacentes en la unidad de transformación y a continuación realizar la transformación en las unidades de predicción adyacentes.

Cuando la unidad 1010 de selección establece la unidad de transformación, la unidad 1020 de realización de transformada transforma las unidades de predicción adyacentes en un dominio de frecuencia, de acuerdo con la unidad de transformación. La unidad 1020 de realización de transformada realiza la DCT en las unidades de predicción adyacentes de acuerdo con la unidad de transformación, y genera coeficientes de coseno discretos.

Haciendo referencia de vuelta a la Figura 9, la unidad 920 de cuantificación cuantifica coeficientes de componente de frecuencia generados por el transformador 910, por ejemplo, los coeficientes de coseno discretos. La unidad 920 de cuantificación puede cuantificar los coeficientes de coseno discretos que se introducen de acuerdo con una etapa de cuantificación predeterminada.

El codificador 930 por entropía realiza codificación por entropía en los coeficientes de componente de frecuencia que se cuantifican por la unidad 920 de cuantificación. El codificador 930 por entropía puede realizar la codificación por entropía en los coeficientes de coseno discretos usando codificación aritmética variable adaptativa según contexto (CABAC) o codificación de longitud variable adaptativa según contexto (CAVLC).

El aparato 900 de codificación de imagen puede determinar una unidad de transformación óptima realizando de manera repetitiva la DCT, la cuantificación y la codificación por entropía en diferentes unidades de transformación. Un procedimiento para seleccionar las unidades de predicción adyacentes puede repetirse para determinar la unidad de transformación óptima. La unidad de transformación óptima puede determinarse en consideración de un cálculo de coste de RD, y esto se describirá en detalle con referencia a la Figura 12.

La Figura 12 ilustra diferentes unidades de transformación de acuerdo con otra realización ejemplar.

Haciendo referencia a la Figura 12, el aparato 900 de codificación de imagen realiza de manera repetitiva una operación de codificación en las diferentes unidades de transformación.

Como se ilustra en la Figura 12, una unidad 1210 de codificación puede predecirse y codificarse basándose en una unidad 1220 de predicción que tiene un tamaño menor que la unidad 1210 de codificación. Se realiza una

transformación en valores residuales que se generan por un resultado de la predicción, y en este punto, como se ilustra en la Figura 12, la DCT puede realizarse en los valores residuales basándose en las diferentes unidades de transformación.

5 Una primera unidad 1230 de transformación ilustrada tiene el mismo tamaño que la unidad 1210 de codificación, y tiene un tamaño obtenido agrupando todas las unidades de predicción incluidas en la unidad 1210 de codificación.

Una segunda unidad 1240 de transformación ilustrada tiene tamaños obtenidos dividiendo por la mitad una anchura de la unidad 1210 de codificación, y los tamaños se obtienen agrupando cada dos unidades de predicción adyacentes entre sí en una dirección vertical, respectivamente.

10 Una tercera unidad 1250 de transformación ilustrada tiene tamaños obtenidos dividiendo por la mitad una altura de la unidad 1210 de codificación, y los tamaños se obtienen agrupando cada dos unidades de predicción adyacentes entre sí en una dirección horizontal, respectivamente.

Una cuarta unidad 1260 de transformación ilustrada se usa cuando la transformación se realiza basándose en la cuarta unidad 1260 de transformación ilustrada que tiene el mismo tamaño que la unidad 1220 de predicción.

15 La Figura 13 es un diagrama de bloques de un aparato 1300 de decodificación de imagen de acuerdo con otra realización ejemplar.

Haciendo referencia a la Figura 13, el aparato 1300 de decodificación de imagen de acuerdo con la presente realización ejemplar incluye un decodificador 1310 por entropía, una unidad 1320 de cuantificación inversa y un transformador 1330 inverso.

20 El decodificador 1310 por entropía realiza decodificación por entropía en coeficientes de componente de frecuencia con respecto a una unidad de transformación predeterminada. Como se ha descrito con referencia a las Figuras 11A a 11C y 12, la unidad de transformación predeterminada puede ser una unidad de transformación generada agrupando una pluralidad de unidades de predicción adyacentes.

25 Como se ha descrito con referencia al aparato 900 de codificación de imagen, la unidad de transformación puede generarse agrupando las unidades de predicción adyacentes de acuerdo con una profundidad, o puede generarse agrupando una pluralidad de unidades de predicción adyacentes en las que se realiza predicción de acuerdo con el mismo modo de predicción, es decir, de acuerdo con un intra-modo de predicción o un inter-modo de predicción.

30 La pluralidad de unidades de predicción pueden no incluirse en una unidad de codificación sino incluirse en una pluralidad de unidades de codificación. En otras palabras, como se ha descrito con referencia a las Figuras 11A a 11C, la unidad de transformación que se decodifica por entropía por el decodificador 1310 por entropía puede establecerse para que sea igual a o menor que un tamaño de una unidad de codificación, o para que sea mayor que el tamaño de la unidad de codificación.

35 También, como se ha descrito con referencia a la Figura 12, la unidad de transformación puede ser una unidad de transformación óptima seleccionada repitiendo un procedimiento para agrupar una pluralidad de unidades de predicción adyacentes, y realizando de manera repetitiva una transformación, cuantificación y decodificación por entropía en diferentes unidades de transformación.

La unidad 1320 de cuantificación inversa cuantifica a la inversa los coeficientes de componente de frecuencia que están decodificados por entropía por el decodificador 1310 por entropía.

40 La unidad 1320 de cuantificación inversa cuantifica a la inversa los coeficientes de componente de frecuencia decodificados por entropía de acuerdo con una etapa de cuantificación que se usa en la codificación de la unidad de transformación.

45 El transformador 1330 inverso transforma a la inversa los coeficientes de componente de frecuencia cuantificados a la inversa en un dominio de píxel. El transformador inverso puede realizar una DCT inversa en coeficientes de coseno discretos cuantificados a la inversa (es decir, los coeficientes de componente de frecuencia cuantificados a la inversa), y a continuación puede reconstruir una unidad de transformación del dominio de píxel. La unidad de transformación reconstruida puede incluir unidades de predicción adyacentes.

La Figura 14 es un diagrama de flujo de un procedimiento de codificación de imagen, de acuerdo con una realización ejemplar.

50 Haciendo referencia a la Figura 14, en la operación 1410, un aparato de codificación de imagen establece una unidad de transformación seleccionando una pluralidad de unidades de predicción adyacentes. El aparato de codificación de imagen puede seleccionar una pluralidad de unidades de predicción adyacentes de acuerdo con una profundidad, o puede seleccionar una pluralidad de unidades de predicción adyacentes en las que se realiza predicción de acuerdo con el mismo modo de predicción.

En la operación 1420, el aparato de codificación de imagen transforma las unidades de predicción adyacentes en un dominio de frecuencia de acuerdo con la unidad de transformación establecida en la operación 1420. El aparato de codificación de imagen agrupa las unidades de predicción adyacentes, realiza una DCT en las unidades de predicción adyacentes, y genera de esta manera coeficientes de coseno discretos.

- 5 En la operación 1430, el aparato de codificación de imagen cuantifica coeficientes de componente de frecuencia, generados en la operación 1420, de acuerdo con una etapa de cuantificación.

En la operación 1440, el aparato de codificación de imagen realiza codificación por entropía en los coeficientes de componente de frecuencia cuantificados en la operación 1430. El aparato de codificación de imagen realiza la codificación por entropía en los coeficientes de coseno discretos usando CABAC o CAVLC.

- 10 Un procedimiento de codificación de imagen de acuerdo con otra realización ejemplar puede incluir adicionalmente una operación de establecimiento de una unidad de transformación óptima realizando de manera repetitiva las operaciones 1410 a 1440 en diferentes unidades de transformación. Es decir, realizando de manera repetitiva la transformación, la cuantificación y la codificación por entropía en diferentes unidades de transformación como se ilustra en la Figura 12, es posible establecer la unidad de transformación óptima.

- 15 La Figura 15 es un diagrama de flujo de un procedimiento de decodificación de imagen, de acuerdo con otra realización ejemplar.

Haciendo referencia a la Figura 15, en la operación 1510, un aparato de decodificación de imagen realiza decodificación por entropía en coeficientes de componente de frecuencia con respecto a una unidad de transformación predeterminada. Los coeficientes de componente de frecuencia pueden ser coeficientes de coseno discretos.

- 20 En la operación 1520, el aparato de decodificación de imagen cuantifica a la inversa los coeficientes de componente de frecuencia que están decodificados por entropía en la operación 1510. El aparato de decodificación de imagen cuantifica a la inversa los coeficientes de coseno discretos usando una etapa de cuantificación usada en una operación de codificación.

- 25 En la operación 1530, el aparato de decodificación de imagen transforma a la inversa los coeficientes de componente de frecuencia, que se han cuantificado a la inversa en la operación 1520, en un dominio de píxel y a continuación reconstruye la unidad de transformación. La unidad de transformación reconstruida se establece agrupando una pluralidad de unidades de predicción adyacentes. Como se ha descrito anteriormente, la unidad de transformación puede establecerse agrupando las unidades de predicción adyacentes de acuerdo con una profundidad, o puede establecerse agrupando una pluralidad de unidades de predicción adyacentes en las que se realiza predicción de acuerdo con el mismo modo de predicción.

De acuerdo con la una o más realizaciones ejemplares, es posible establecer la unidad de transformación para que sea mayor que la unidad de predicción, y para realizar la DCT, de modo que una imagen puede comprimirse y codificarse de manera eficaz.

- 35 Las realizaciones ejemplares pueden realizarse también como códigos legibles por ordenador en un medio de grabación legible por ordenador. El medio de grabación legible por ordenador es cualquier dispositivo de almacenamiento de datos que puede almacenar datos, que pueden leerse posteriormente por un sistema informático. Ejemplos del medio de grabación legible por ordenador incluyen memoria de solo lectura (ROM), memoria de acceso aleatorio (RAM), CD-ROM, cintas magnéticas, discos flexibles y dispositivos de almacenamiento de datos ópticos. El medio de grabación legible por ordenador puede también distribuirse a través de sistemas informáticos acoplados en red de modo que el código legible por ordenador se almacena y se ejecuta de una manera distribuida.

- 45 Por ejemplo, cada uno del aparato de codificación de imagen, el aparato de decodificación de imagen, el codificador de imagen y el decodificador de imagen de acuerdo con la una o más realizaciones pueden incluir un bus acoplado a cada unidad en un aparato como se ilustra en las Figuras 1-2, 4-5, 9-10 y 14, y al menos un procesador acoplado al bus. También, cada uno del aparato de codificación de imagen, el aparato de decodificación de imagen, el codificador de imagen y el decodificador de imagen de acuerdo con la una o más realizaciones pueden incluir una memoria acoplada a al menos un procesador que está acoplado al bus para almacenar comandos, mensajes recibidos o mensajes generados y para ejecutar los comandos.

- 50 Aunque esta invención se ha mostrado y descrito particularmente con referencia a realizaciones ejemplares de la misma, se entenderá por los expertos en la materia que pueden realizarse en la misma diversos cambios en forma y detalles sin alejarse del alcance de la invención como se define por las reivindicaciones adjuntas.

REIVINDICACIONES

1. Un aparato de decodificación de imagen que comprende:

5 un procesador operable para obtener, de una secuencia de bits recibida, información acerca de un tamaño máximo de una unidad de codificación, información acerca de una unidad de predicción en la unidad de codificación e información acerca de una unidad de transformación para transformación de la unidad de codificación; y

10 un decodificador operable para determinar una pluralidad de unidades de codificación máxima cuadradas en una imagen usando la información acerca de un tamaño máximo de una unidad de codificación, y para determinar unidades de codificación cuadradas de una estructura de árbol que comprende unidades de codificación que ya no se dividen más entre las unidades de codificación divididas de una unidad de codificación máxima usando información acerca de unidades de codificación analizadas a partir de una secuencia de bits, para determinar al menos una unidad de predicción usando la información acerca de una unidad de predicción; y para determinar al menos una unidad de transformación usando la información acerca de una unidad de transformación,

en el que:

15 la unidad de codificación máxima entre la pluralidad de las unidades de codificación máxima se divide jerárquicamente en una o más unidades de codificación que incluyen al menos una de una unidad de codificación de una profundidad actual y una unidad de codificación de una profundidad mayor que la profundidad actual,

la unidad de codificación de la profundidad actual se divide en cuatro unidades de codificación cuadradas de la profundidad mayor que la profundidad actual, independientemente de las unidades de codificación vecinas de la profundidad actual, y

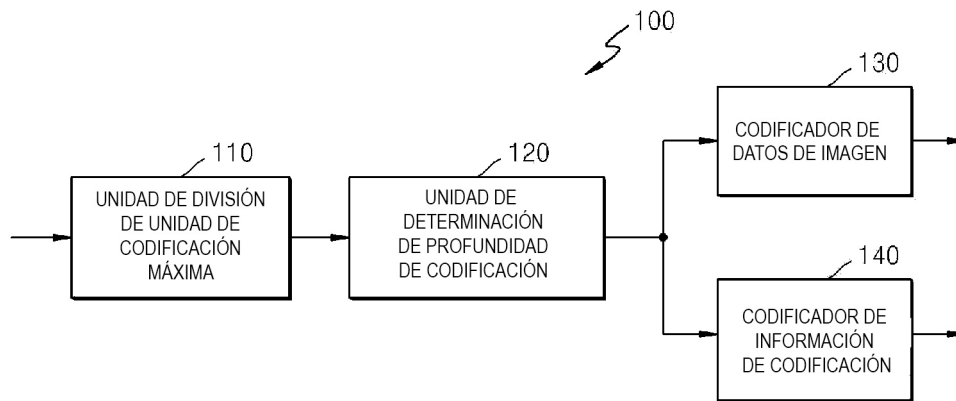
20 la unidad de codificación de la profundidad mayor que la profundidad actual se predice usando la al menos una unidad de predicción y se transforma a la inversa usando la unidad de transformación,

25 en el que la al menos una unidad de predicción es una entre los bloques que incluyen: un bloque igual en tamaño a la unidad de codificación de la profundidad mayor que la profundidad actual; y un bloque entre una pluralidad de bloques generados dividiendo igualmente al menos una de la altura y anchura de la unidad de codificación de la profundidad mayor que la profundidad actual,

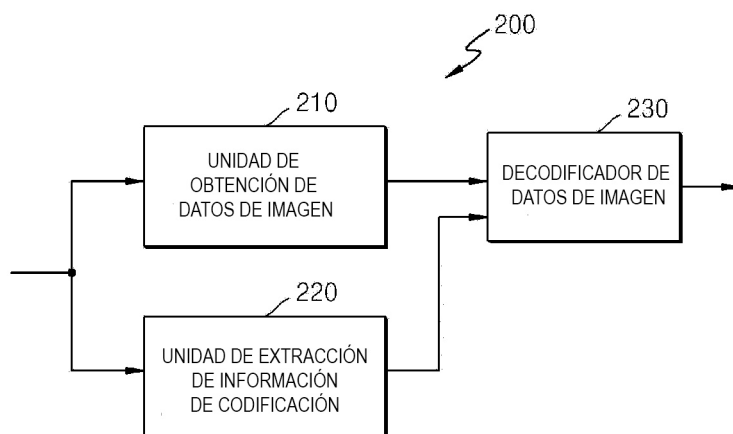
en el que la al menos una unidad de transformación es una entre los bloques que incluyen: un bloque igual en tamaño a la unidad de codificación de la profundidad mayor que la profundidad actual; y un bloque entre una pluralidad de bloques generados dividiendo la altura y anchura de la unidad de codificación de la profundidad mayor que la profundidad actual, y

30 **caracterizado porque** el aparato de decodificación está configurado para soportar una unidad de transformación, entre la al menos una unidad de transformación, que tiene un tamaño de $2N \times 2N$ que incluye cuatro unidades de predicción que tienen un tamaño de $N \times N$.

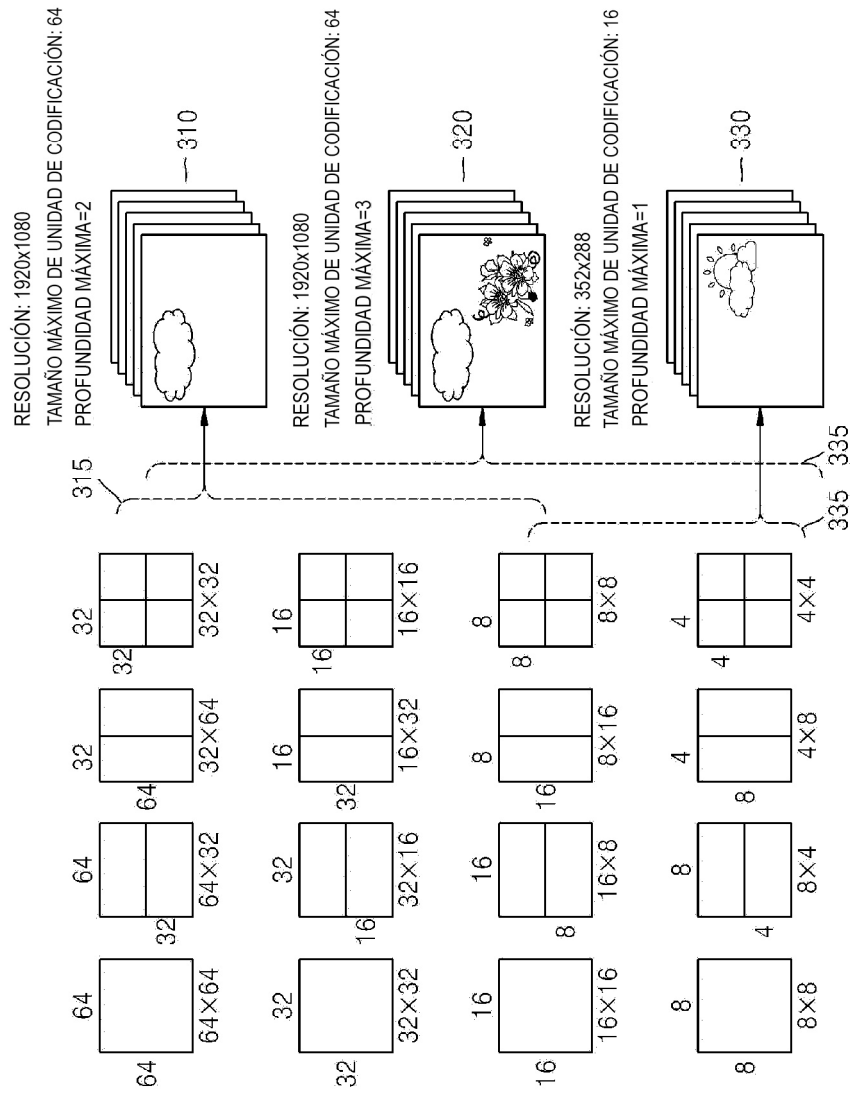
[Figura 1]



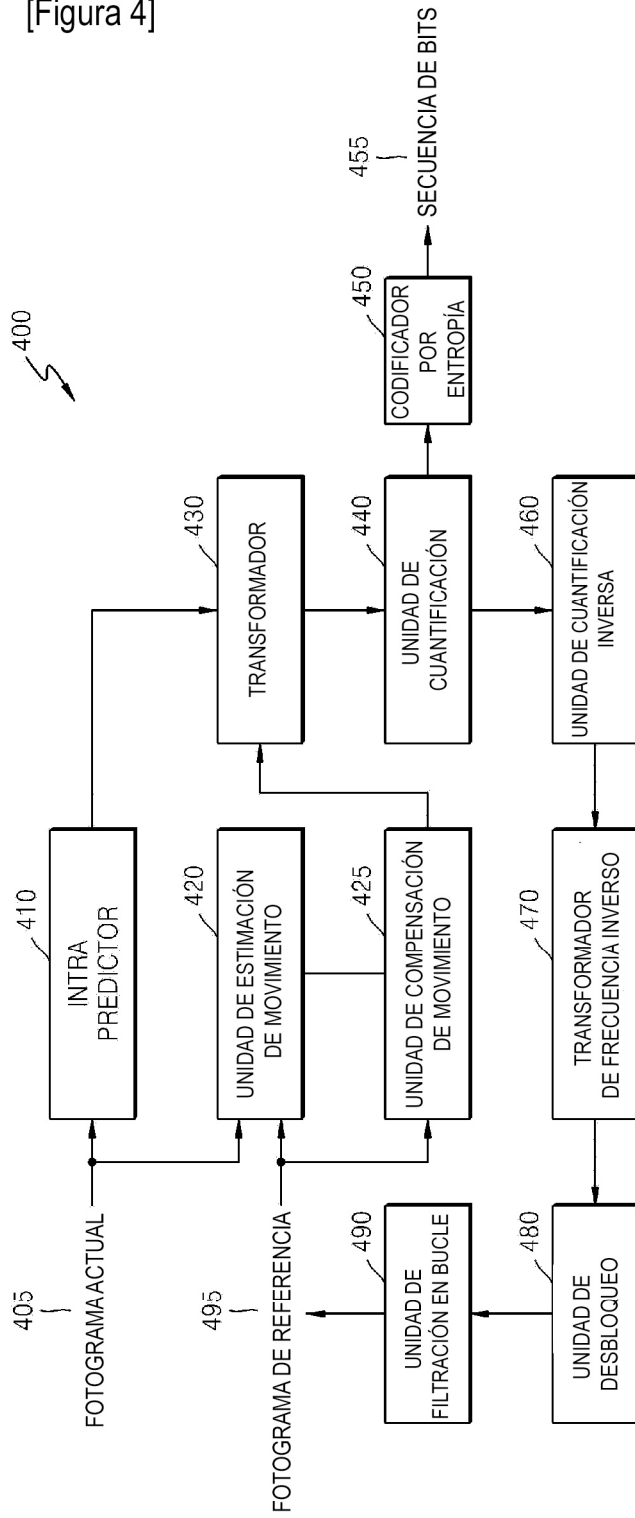
[Figura 2]



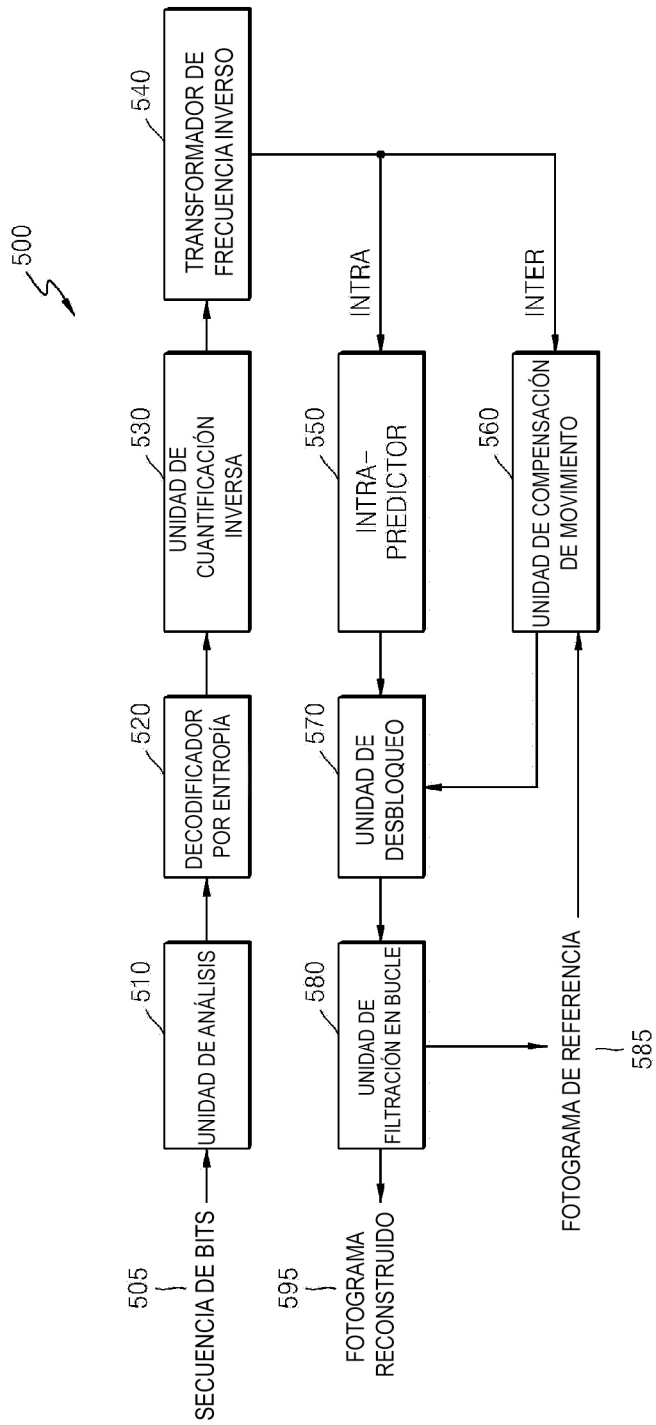
[Figura 3]



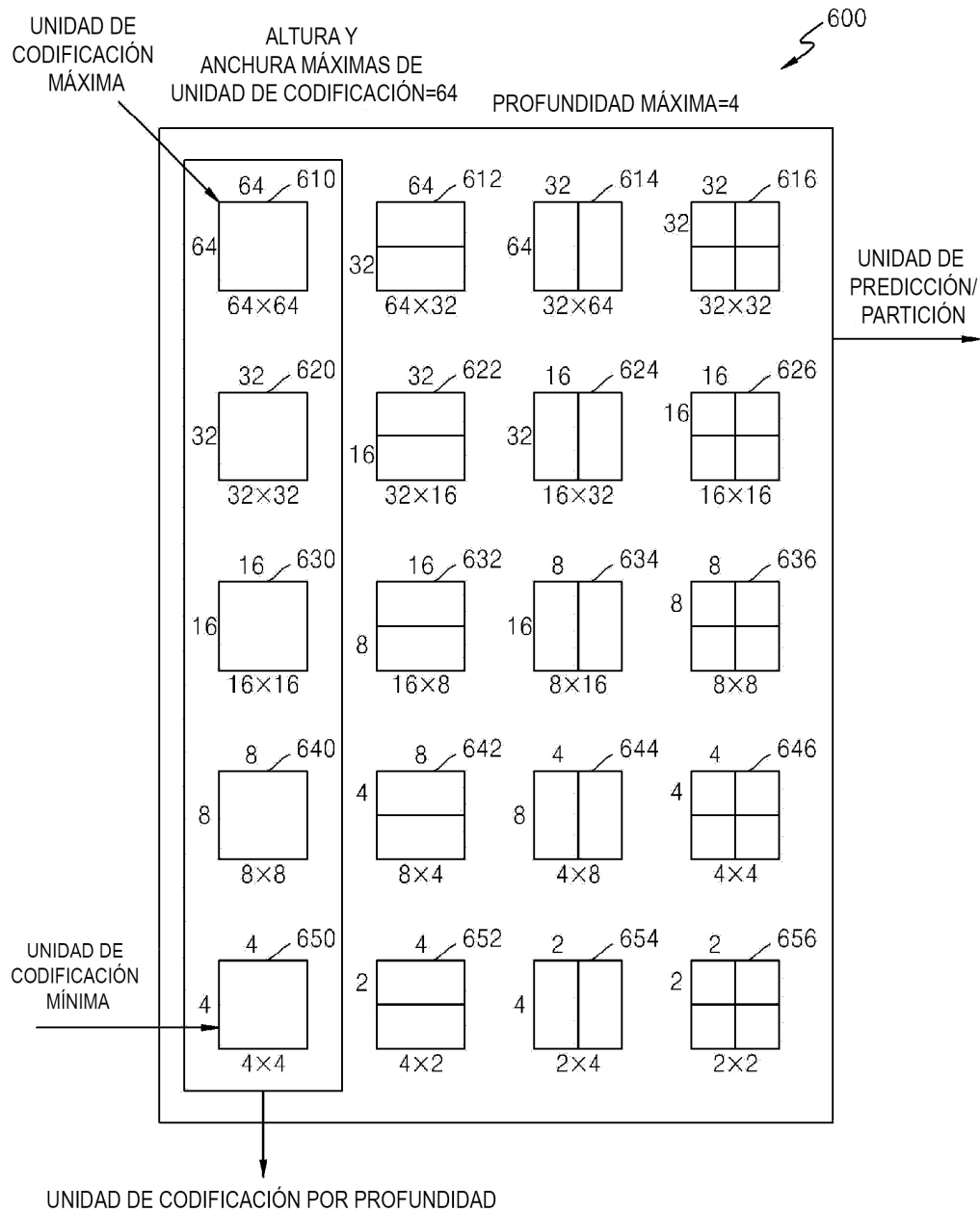
[Figura 4]



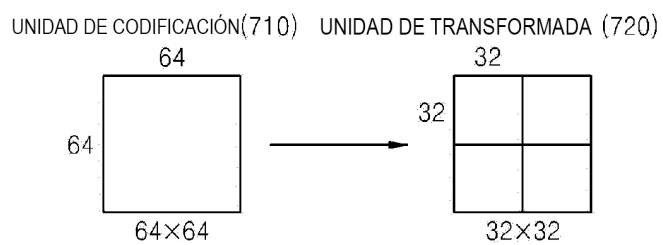
[Figura 5]



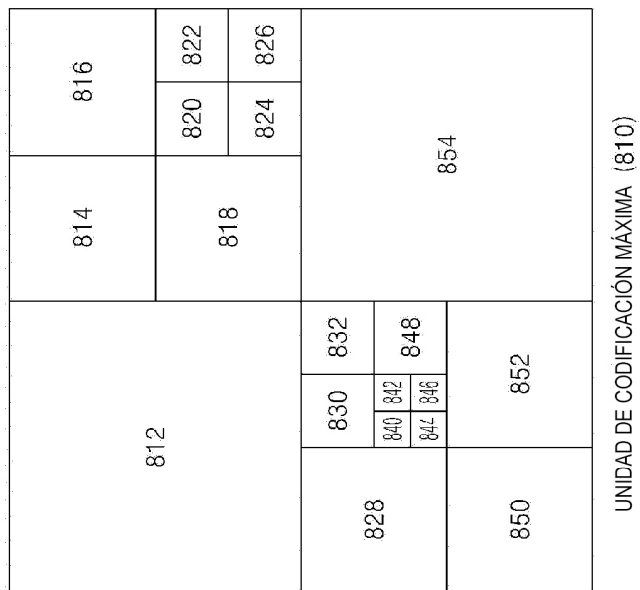
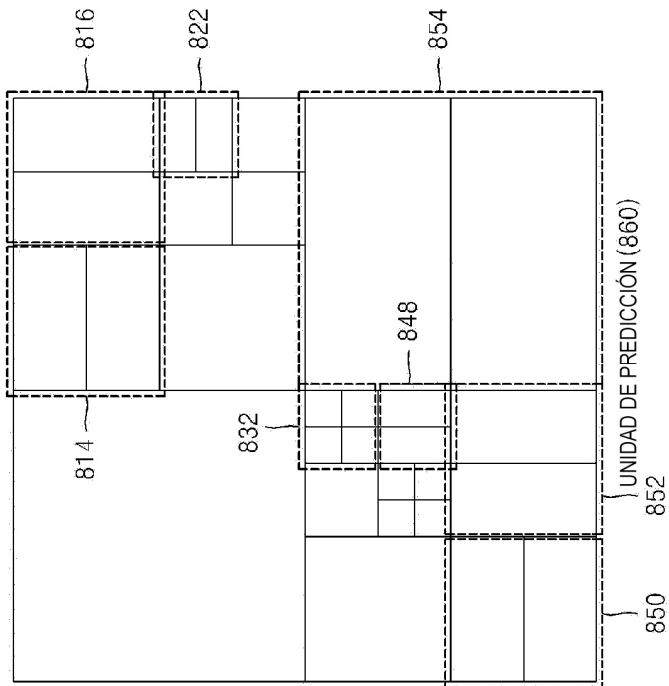
[Figura 6]



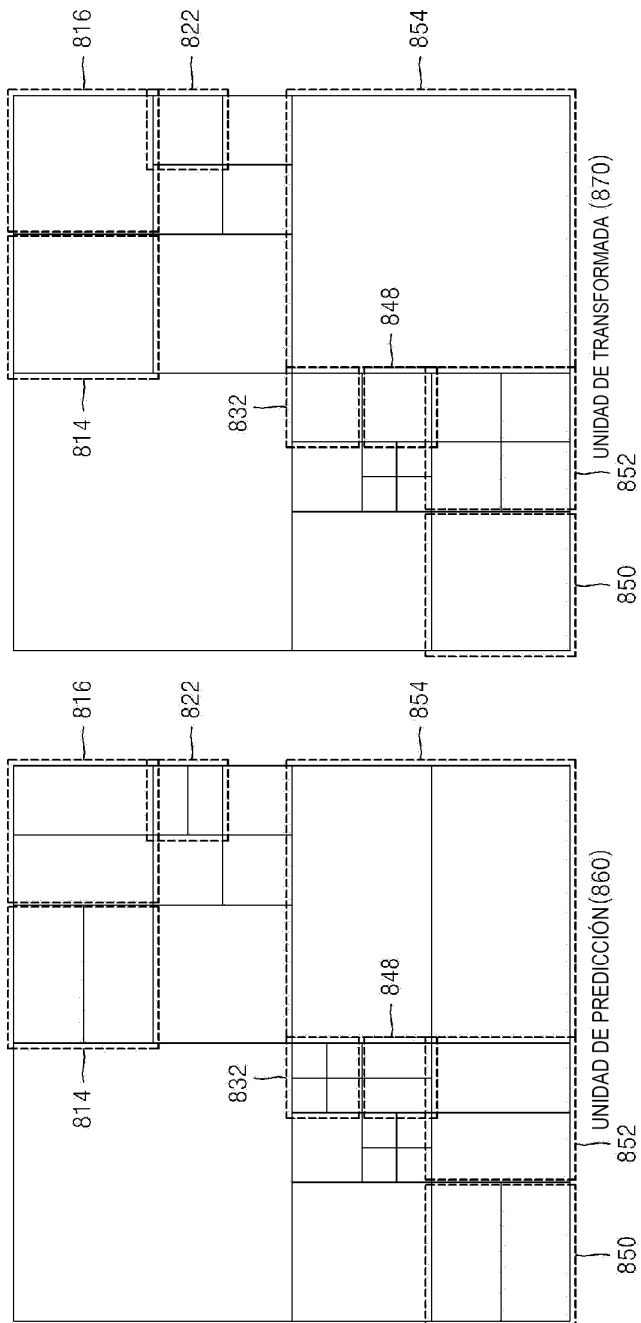
[Figura 7]



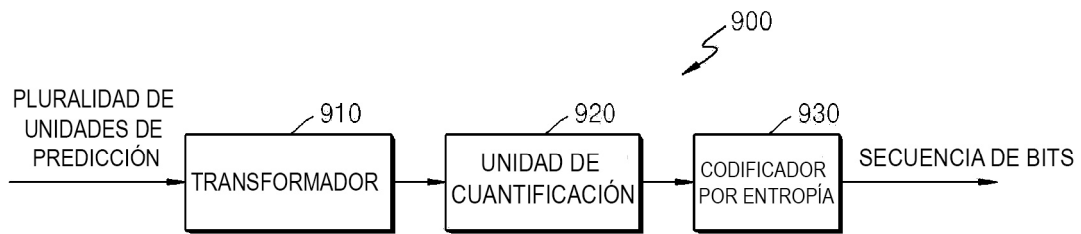
[Figura 8a]



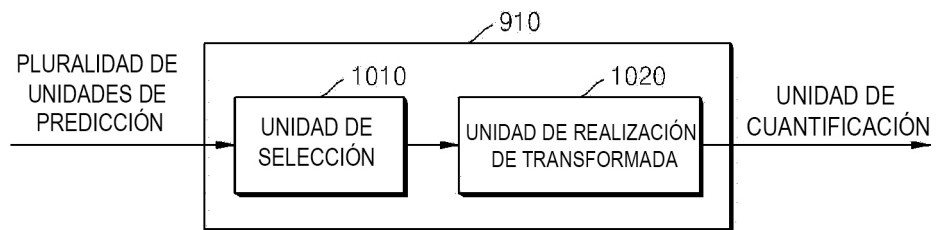
[Figura 8b]



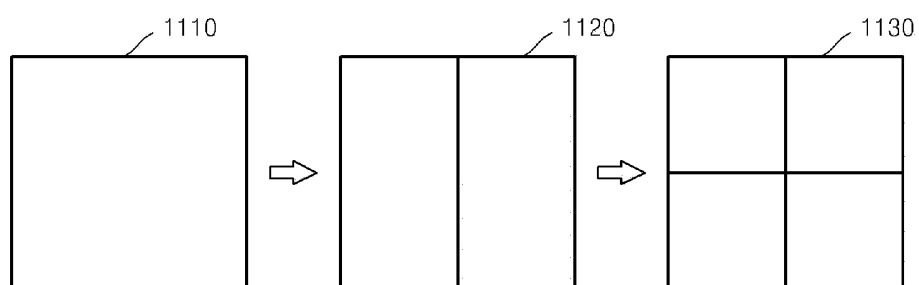
[Figura 9]



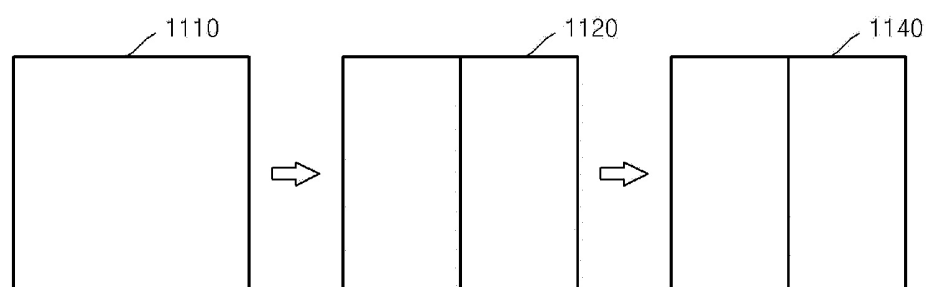
[Figura 10]



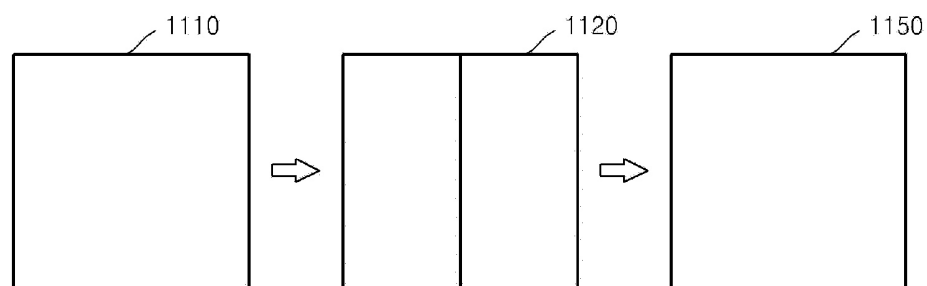
[Figura 11a]



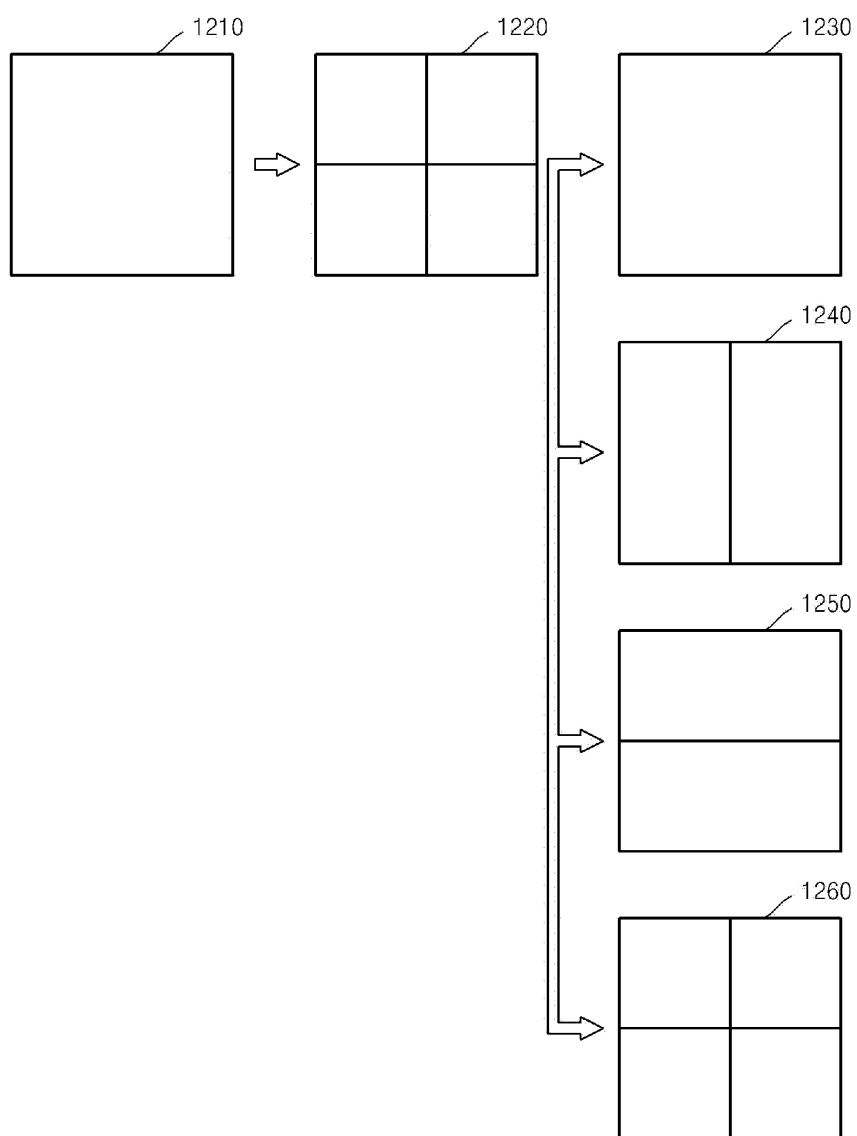
[Figura 11b]



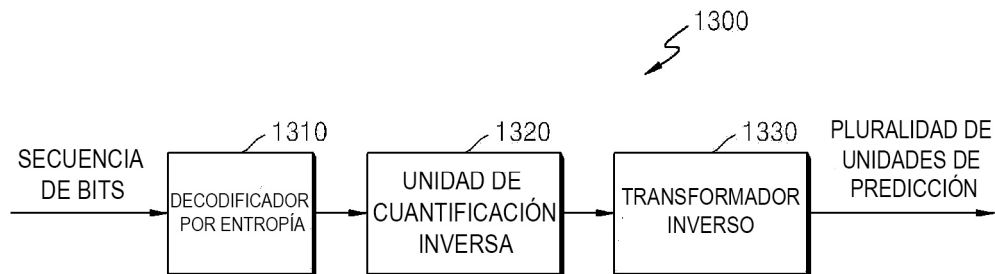
[Figura 11c]



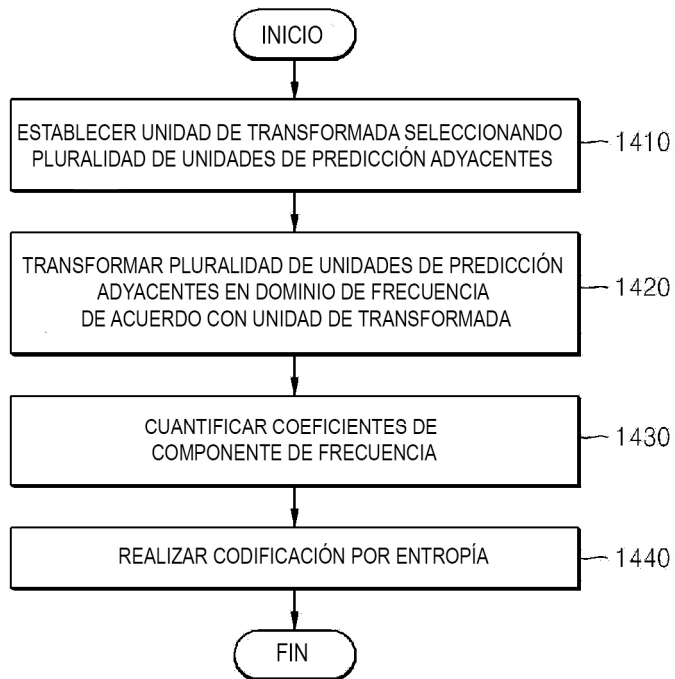
[Figura 12]



[Figura 13]



[Figura 14]



[Figura 15]

