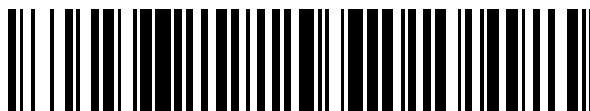


19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 701 979**

51 Int. Cl.:

H04N 7/24	(2011.01) H04N 19/169	(2014.01)
H04N 19/96	(2014.01)	
H04N 19/70	(2014.01)	
H04N 19/61	(2014.01)	
H04N 19/176	(2014.01)	
H04N 19/122	(2014.01)	
H04N 19/119	(2014.01)	
H04N 19/107	(2014.01)	
G06T 9/00	(2006.01)	
H04N 19/625	(2014.01)	

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

96 Fecha de presentación y número de la solicitud europea: **13.08.2010** **E 17193659 (4)**

97 Fecha y número de publicación de la concesión europea: **05.12.2018** **EP 3282696**

54 Título: **Procedimiento para decodificar y aparato para codificar imágenes usando unidad de transformación grande**

30 Prioridad:

13.08.2009 KR 20090074895

45 Fecha de publicación y mención en BOPI de la traducción de la patente:
26.02.2019

73 Titular/es:

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

72 Inventor/es:

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

74 Agente/Representante:

CARPINTERO LÓPEZ, Mario

ES 2 701 979 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 para decodificar y aparato para codificar imágenes usando unidad de transformación grande

Campo técnico

5 Las realizaciones se refieren a un procedimiento para decodificar una imagen, un aparato para codificar una imagen y un medio de almacenamiento legible por ordenador.

Antecedentes de la técnica

10 Para realizar compresión de imagen, la mayoría de procedimientos y aparatos de codificación y decodificación 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 usando la DCT implica realizar la DCT en una imagen de un dominio de píxel, generando coeficientes de coseno discretos, cuantificando los coeficientes de coseno discretos generados y realizando codificación de entropía en los coeficientes de coseno discretos generados.

15 En la norma de codificación de video H.264/AVC (véase Wiegand et al: "Overview of the H.264/ AVC video coding standard", Transacciones en Circuitos y Sistema para Tecnología de Video de IEEE, vol. 13, n.º 7, 1 de julio de 2003, páginas 560-576) una imagen se reparte en macrobloques de tamaño de 16×16. Para predicción de movimiento compensado se soportan particiones de macrobloques de tamaño de 16×16, 16×8, 8×16 y 8×8 y adicionalmente pueden separarse particiones de 8×8 en particiones de 8×4, 4×8 o 4×4 si se indica mediante un elemento de sintaxis. Una transformada entera de tamaño fijo de 4×4 se aplica independientemente del tamaño de las particiones usadas para la predicción.

Divulgación

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

Solución técnica

25 Las realizaciones proporcionan un procedimiento para la decodificación una imagen y aparato para codificar una imagen mediante el uso de una transformada de coseno discreto (DCT) más eficiente, y también proporciona un medio de grabación legible por ordenador.

Efectos ventajosos

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

Descripción de los dibujos

Las anteriores y otras características de las realizaciones serán más evidentes describiendo en detalle realizaciones ilustrativas de las mismas con referencia a los dibujos adjuntos en los que:

35 la Figura 1 es un diagrama de bloques de un aparato de codificación de imagen de acuerdo con una realización ilustrativa;
la Figura 2 es un diagrama de un aparato de decodificación de imagen de acuerdo con otra realización ilustrativa;
la Figura 3 es un diagrama de una unidad de codificación jerárquica de acuerdo con otra realización ilustrativa;
la Figura 4 es un diagrama de bloques de un codificador de imagen a base de una unidad de codificación de acuerdo con otra realización ilustrativa;
40 la Figura 5 es un diagrama de bloques de un decodificador de imagen a base de una unidad de codificación de acuerdo con otra realización ilustrativa;
la Figura 6 ilustra una unidad de codificación máxima, unidades de sub-codificación y unidades de predicción de acuerdo con otra realización ilustrativa;
la Figura 7 es un diagrama de una unidad de codificación y una unidad de transformación de acuerdo con otra realización ilustrativa;
45 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 ilustrativa;
la Figura 9 es un diagrama de bloques de un aparato de codificación de imagen de acuerdo con otra realización ilustrativa;
50 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 ilustrativa;
la Figura 12 ilustra unidades de transformación diferentes de acuerdo con otra realización ilustrativa;
la Figura 13 es un diagrama de bloques de un aparato de decodificación de imagen de acuerdo con otra

realización ilustrativa;

la Figura 14 es un diagrama de flujo de un procedimiento de codificación de imagen, de acuerdo con una realización ilustrativa; y

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

Mejor modo

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

De acuerdo con otra realización, se proporciona un aparato de codificación de imagen según lo establecido en la reivindicación 2 adjunta. De acuerdo con otra realización, se proporciona un medio de grabación legible por ordenador según lo establecido en la reivindicación 3 adjunta.

Modo para la invención

En lo sucesivo, las realizaciones ilustrativas se describirán en detalle con referencia a los dibujos adjuntos. En las realizaciones ilustrativas, "unidad" puede referirse o no a una unidad de tamaño, dependiendo de su contexto, e "imagen" puede representar una imagen fija para un video o una imagen en movimiento, es decir, el propio video.

La Figura 1 es un diagrama de bloques de un aparato 100 para la codificación de una imagen, de acuerdo con una realización ilustrativa.

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.

La unidad 110 de división de unidad de codificación máxima puede dividir una instantánea o corte actual a base de 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.

De acuerdo con una realización ilustrativa, 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 de entre unidades de codificación de la instantánea actual y la profundidad indica el tamaño de una unidad de sub-codificación obtenida disminuyendo jerárquicamente la unidad de codificación. A medida que una profundidad aumenta, una unidad de codificación puede disminuir en tamaño desde una unidad de codificación máxima a una unidad de codificación mínima, en la 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. Ya que el tamaño de una unidad de codificación disminuye desde una unidad de codificación máxima a medida que una profundidad aumenta, una unidad de sub-codificación de una k -ésima profundidad puede incluir una pluralidad de unidades de sub-codificación de una $(k+n)$ -ésima profundidad (k y n son enteros iguales o mayores de 1).

De acuerdo con un aumento del tamaño de una fotografía a codificar, codificar una imagen en una unidad de codificación mayor puede resultar en una relación de compresión de imagen más alta. Sin embargo, si se fija una unidad de codificación mayor, una imagen no puede codificarse eficientemente teniendo en cuenta las características de imagen continuamente cambiantes.

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

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

La unidad 120 de determinación de profundidad de codificación determina una profundidad máxima. La profundidad máxima puede determinarse a base de cálculo de coste de Tasa-Distorsión (R-D). La profundidad máxima puede determinarse de forma 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 datos de imagen de acuerdo con 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 unidades de sub-codificación que tienen diferentes tamaños de acuerdo con profundidades diferentes. Esto se describe en detalle más adelante con referencia a las Figuras 8A y 8B. Además, las unidades de sub-codificación que tienen tamaños diferentes, que se incluyen en la unidad de codificación máxima, pueden predecirse o transformarse a base de unidades de procesamiento que tienen tamaños diferentes. En otras palabras, el aparato 100 puede realizar una pluralidad de operaciones de procesamiento de codificación de imagen a base de 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 de entropía, en las que unidades de procesamiento que tienen el mismo tamaño pueden usarse para cada operación o 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 $2N \times 2N$ (en el que N es un entero positivo), unidades de procesamiento para predicción pueden ser $2N \times 2N$, $2N \times N$, $N \times 2N$ y $N \times N$. En otras palabras, predicción de movimiento puede realizarse a base de una unidad de procesamiento que tiene una forma con lo que al menos una de altura y anchura de una unidad de codificación se divide equitativamente por dos. En lo sucesivo, una unidad de procesamiento, que es la base de 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 únicamente para 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 últimos errores de codificación puede seleccionarse tras realizar predicción para cada unidad de predicción.

Como alternativa, el aparato 100 puede realizar transformación de frecuencia en datos de imagen a base de 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 a base de una unidad de procesamiento que tiene un tamaño igual a o más pequeño que el de la unidad de codificación. En lo sucesivo, una unidad de procesamiento, que es la base de 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 unidades de sub-codificación incluidas en una unidad de codificación máxima usando optimización RD a base de un multiplicador de Lagrangian. En otras palabras, la unidad 120 de determinación de profundidad de codificación puede determinar las formas de una pluralidad de unidades de sub-codificación divididas de la unidad de codificación máxima, en las que la pluralidad de unidades de sub-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 a base de las formas de división, es decir, las formas que dividen la unidad de codificación máxima, como se determina mediante 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 mediante la unidad 120 de determinación de profundidad de codificación. En otras palabras, el codificador 140 de información de codificación emite una secuencia de bits mediante información de codificación a cerca 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 unidad de sub-codificación para cada profundidad. La información acerca del modo de codificación de la unidad de sub-codificación puede incluir información acerca de una unidad de predicción de la unidad de sub-codificación, información acerca de un modo de predicción para cada unidad de predicción e información acerca de una unidad de transformación de la unidad de sub-codificación.

Ya que existen unidades de sub-codificación que tienen diferentes tamaños para cada unidad de codificación máxima y debe determinarse información acerca de un modo de codificación para cada unidad de sub-codificación, puede determinarse información acerca de al menos un modo de codificación para una unidad de codificación máxima.

El aparato 100 puede generar unidades de sub-codificación dividiendo equitativamente en dos tanto altura como anchura de una unidad de codificación máxima de acuerdo con un aumento de profundidad. Es decir, cuando el tamaño de una unidad de codificación de una k -ésima profundidad es $2N \times 2N$, el tamaño de una unidad de codificación de una $(k+1)$ -ésima profundidad es $N \times N$.

Por consiguiente, el aparato 100 de acuerdo con una realización ilustrativa puede determinar una forma de división óptima para cada unidad de codificación máxima a base de tamaños de unidades de codificación máxima y una profundidad máxima en consideración de las características de imagen. Ajustando de forma 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 unidades de sub-codificación de profundidades diferentes, imágenes que tienen diversas resoluciones pueden codificarse más eficientemente.

La Figura 2 es un diagrama de bloques de un aparato 200 para la decodificación una imagen de acuerdo con una realización ilustrativa.

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 adquiere datos de imagen de acuerdo con unidades de codificación máxima analizando una secuencia de bits recibido 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 de 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 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 unidades de sub-codificación del encabezamiento de la instantánea actual analizando la secuencia de bits recibido 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 unidades de sub-codificación que tienen diferentes tamaños de acuerdo con 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 unidad de sub-codificación, información acerca de un modo de predicción e información acerca de una unidad de transformación.

El decodificador 230 de datos de imagen restablece la instantánea actual decodificando datos de imagen de cada unidad de codificación máxima a base de la información extraída por la unidad 220 de extracción de información de codificación. El decodificador 230 de datos de imagen puede decodificar unidades de sub-codificación incluidas en una unidad de codificación máxima a base de la información acerca de una forma de división de la unidad de codificación máxima. Un procedimiento de decodificación puede incluir una predicción procedimiento 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 a base de 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 también puede realizar transformación inversa para cada unidad de sub-codificación a base de información acerca de una unidad de transformación de una unidad de sub-codificación.

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

Haciendo referencia a la Figura 3, las unidades de codificación jerárquica de acuerdo con una realización ilustrativa pueden incluir unidades de codificación cuya anchura×altura son 64×64, 32×32, 16×16, 8×8 y 4×4. Además de estas unidades de codificación que tienen formas cuadradas perfectas, también pueden existir unidades de codificación cuya anchura×altura son 64×32, 32×64, 32×16, 16×32, 16×8, 8×16, 8×4 y 4×8.

Haciendo referencia a la Figura 3, para datos 310 de imagen cuya resolución es de 1920×1080, el tamaño de una unidad de codificación máxima se establece a 64×64, y una profundidad máxima se establece a 2.

Para datos 320 de imagen cuya resolución es de 1920×1080, el tamaño de una unidad de codificación máxima se establece a 64×64 y una profundidad máxima se establece a 4. Para datos 330 de imagen cuya resolución es de 352×288, el tamaño de una unidad de codificación máxima se establece a 16×16 y una profundidad máxima se establece a 1.

Cuando la resolución es alta o la cantidad de datos es alta, es preferible, pero no necesariamente, que un tamaño máximo de una unidad de codificación sea relativamente grande para aumentar una relación de compresión y reflejar exactamente características de imagen. Por consiguiente, para los datos 310 y 320 de imagen que tienen resolución más alta que los datos 330 de imagen, puede seleccionarse 64×64 como el tamaño de una unidad de codificación máxima.

Una profundidad máxima indica el número total de capas en las unidades de codificación jerárquica. Ya 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 unidades de sub-codificación cuyos tamaños de eje más largo son 32 y 16, de acuerdo con un aumento de una profundidad.

5 Por otra parte, ya 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, ya 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 unidades de sub-codificación cuyos tamaños de eje más largo son 32, 16, 8 y 4 de acuerdo con un aumento de una profundidad. Ya que una imagen se codifica a base de una unidad de sub-codificación más pequeña a medida que una profundidad aumenta, la realización ilustrativa es adecuada para la codificación de una imagen que incluye más detalles minuciosos en escenas.

La Figura 4 es un diagrama de bloques de un codificador de imagen 400 a base de una unidad de codificación, de acuerdo con una realización ilustrativa.

15 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 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 Valores residuales se generan a base de las unidades de predicción emitidas desde el intra predictor 410, la unidad 420 de estimación de movimiento y la unidad 425 compensación de movimiento, y los valores residuales generados se emiten como coeficientes de transformación 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 post-procesan pasando a través de una unidad 480 de desbloqueo y una unidad 490 de filtración de bucle y emitidos como el fotograma 495 de referencia. Los coeficientes de transformada cuantificados pueden emitirse como una secuencia 455 de bits pasando a través de un codificador 450 de entropía.

30 Para realizar codificación a base de un procedimiento de codificación de acuerdo con una realización ilustrativa, componentes del codificador de imagen 400, es decir, el intra predictor 410, la unidad 420 de estimación de movimiento, la unidad 425 compensación de movimiento, el transformador 430, la unidad 440 de cuantificación, el codificador 450 de entropía, la unidad 460 de cuantificación inversa, el transformador 470 inverso de frecuencia, la unidad 480 de desbloqueo y la unidad 490 de filtración de bucle, realizan procedimientos de codificación de imagen a base de una unidad de codificación máxima, una unidad de sub-codificación de acuerdo con 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 a base de una unidad de codificación, de acuerdo con una realización ilustrativa.

40 Una secuencia 505 de bits pasa a través de una unidad 510 de análisis de modo que se analizan datos de imagen codificados a decodificar e información de codificación necesaria para la decodificación. Los datos de imagen codificados se emiten como datos cuantificados inversamente pasando a través de un decodificador 520 de entropía y una unidad 530 de cuantificación inversa y 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 unidades de codificación siguientes o una fotografía siguiente pasando a través de una unidad 570 de desbloqueo y una unidad 580 de filtración de bucle.

50 Para realizar decodificación a base de un procedimiento de decodificación de acuerdo con una realización ilustrativa, componentes del decodificador 500 de imagen, es decir, la unidad 510 de análisis, el decodificador 520 de 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 580 de filtración de bucle, realizan procedimientos de decodificación de imagen a base de una unidad de codificación máxima, una unidad de sub-codificación de acuerdo con 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 determina una unidad de predicción y un modo de predicción en una unidad de sub-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 unidad de sub-codificación y una unidad de predicción,

de acuerdo con una realización ilustrativa.

El aparato 100 y el aparato 200 de acuerdo con una realización ilustrativa usa 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 adaptativamente de acuerdo con la imagen características o establecerse variablemente de acuerdo con requisitos de un usuario.

Una estructura 600 de unidad de codificación jerárquica de acuerdo con una realización ilustrativa ilustra una unidad 610 de codificación máxima cuya altura y anchura son 64 y profundidad máxima es 4. Una profundidad aumenta a lo largo de un eje vertical de la estructura 600 de unidad de codificación jerárquica y a medida que aumenta una profundidad, disminuyen alturas y anchuras de las unidades 620 a 650 de sub-codificación. Unidades de predicción de la unidad 610 de codificación máxima y las unidades 620 a 650 de sub-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 64×64 . Una profundidad aumenta a lo largo del eje vertical y existe una unidad 620 de sub-codificación cuyo tamaño es 32×32 y profundidad es 1, una unidad 630 de sub-codificación cuyo tamaño es 16×16 y profundidad es 2, una unidad 640 de sub-codificación cuyo tamaño es 8×8 y profundidad es 3 y una unidad 650 de sub-codificación cuyo tamaño es 4×4 y profundidad es 4. La unidad 650 de sub-codificación cuyo tamaño es 4×4 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, 64×64 o una unidad 612 de predicción cuyo tamaño es 64×32 , una unidad de 614 predicción cuyo tamaño es 32×64 o una unidad 616 de predicción cuyo tamaño es 32×32 , que tiene un tamaño más pequeño que la unidad 610 de codificación cuyo tamaño es 64×64 .

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

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

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

Finalmente, la unidad 650 de codificación cuya profundidad es 4 y tamaño es 4×4 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 4×4 , una unidad 652 de predicción que tiene un tamaño de 4×2 , una unidad 654 de predicción que tiene un tamaño de 2×4 o una unidad 656 de predicción que tiene un tamaño de 2×2 .

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

El aparato 100 y el aparato 200, de acuerdo con una realización ilustrativa, realizan codificación con una propia unidad de codificación máxima o con unidades de sub-codificación, que son iguales a o más pequeñas que la unidad de codificación máxima, divididas de la unidad de codificación máxima.

En el procedimiento de codificación, se selecciona el tamaño de una unidad de transformación para la transformación de frecuencia para no ser más grande que el de una unidad de codificación correspondiente. Por ejemplo, cuando una unidad 710 de codificación tiene el tamaño de 64×64 , puede realizarse transformación de frecuencia usando una unidad 720 de transformación que tiene el tamaño de 32×32 .

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

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

Un lado izquierdo de la Figura 8A muestra una forma de división seleccionada por el aparato 100, de acuerdo con una realización ilustrativa, 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 a base de coste R-D. Cuando es óptimo codificar la unidad 810 de codificación máxima tal 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 dividiendo la misma en unidades de sub-codificación cuyas profundidades son iguales a o mayores que 1. Es decir, la unidad 810 de codificación máxima se divide en 4 unidades de sub-codificación cuyas profundidades son 1, y todas o algunas de las unidades de sub-codificación cuyas profundidades son 1 se dividen en unidades de sub-codificación cuyas profundidades son 2.

Una unidad de sub-codificación ubicada en un lado superior derecho y una unidad de sub-codificación ubicada en un lado inferior izquierdo entre las unidades de sub-codificación cuyas profundidades son 1 se dividen en unidades de sub-codificación cuyas profundidades son iguales a o mayores que 2. Algunas de las unidades de sub-codificación cuyas profundidades son iguales a o mayores que 2 pueden dividirse en unidades de sub-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 forma diferente que la unidad 810 de codificación máxima. En otras palabras, una unidad de predicción para cada una de las unidades de sub-codificación puede ser más pequeña que una correspondiente unidad de sub-codificación.

Por ejemplo, una unidad de predicción para una unidad 854 de sub-codificación ubicada en un lado inferior derecho entre las unidades de sub-codificación cuyas profundidades son 1 puede ser más pequeña que la unidad 854 de sub-codificación. Además, unidades de predicción para algunas (814, 816, 850 y 852) unidades 814, 816, 818, 828, 850 y 852 de sub-codificación cuyas profundidades son 2 pueden ser más pequeñas que las unidades 814, 816, 850 y 852 de sub-codificación, respectivamente. Además, unidades de predicción para unidades 822, 832 y 848 de sub-codificación cuyas profundidades son 3 pueden ser más pequeñas que las unidades 822, 832 y 848 de sub-codificación, respectivamente. Las unidades de predicción pueden tener una forma con lo que respectivas unidades de sub-codificación se dividen equitativamente por dos en una dirección de altura o anchura o tienen una forma con lo que respectivas unidades de sub-codificación se dividen equitativamente por 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 ilustrativa.

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 forma diferente que la unidad 860 de predicción.

Por ejemplo, incluso aunque una unidad de predicción para la unidad 854 de codificación cuya profundidad es 1 se selecciona con una forma con lo que la altura de la unidad 854 de codificación se divide equitativamente por dos, puede seleccionarse una unidad de transformación con el mismo tamaño que la unidad 854 de codificación.

Análogamente, incluso aunque se seleccionan unidades de predicción para unidades 814 y 850 de codificación cuyas profundidades son 2 con una forma con lo que la altura de cada una de las unidades 814 y 850 de codificación se divide equitativamente por dos, puede seleccionarse una unidad de transformación 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 más pequeño que una unidad de predicción. Por ejemplo, cuando se selecciona una unidad de predicción para la unidad 852 de codificación cuya profundidad es 2 con una forma con lo que la anchura de la unidad 852 de codificación se divide equitativamente por dos, puede seleccionarse una unidad de transformación con una forma con lo que la unidad 852 de codificación se divide equitativamente por cuatro en direcciones de altura y anchura, que tiene un tamaño más pequeño 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 ilustrativa.

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

El transformador 910 recibe una imagen unidad de procesamiento de un dominio de píxel y transforma la imagen unidad de procesamiento 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 ilustrativa, la transformada al dominio de la frecuencia puede producirse a través de una Transformada de Coseno Discreta (DCT) o Transformada de Karhunen-Loeve (KLT), y como resultado de la DCT o KLT, se generan coeficientes de dominio de frecuencia. En lo sucesivo, la transformada al dominio de la frecuencia puede ser la DCT, sin embargo, es obvio para un experto en la técnica que la transformada al dominio de la frecuencia puede ser cualquier transformada que implica transformación de una imagen de un dominio de píxel en un dominio de frecuencia.

También, de acuerdo con la presente realización ilustrativa, 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 transformación.

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 a base de 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 sobrecargas añadidas a medida que se disminuyen las unidades de transformación, de tal forma 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 ilustrativa agrupa una pluralidad de unidades de predicción adyacentes en una unidad de transformación y realiza transformación de acuerdo con la unidad de transformación que se genera mediante 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 se realiza la transformación en la misma, puede aumentarse altamente una tasa de compresión de una operación de codificación.

Para este aumento, la unidad 1010 de selección selecciona las unidades de predicción adyacentes a agrupar 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 ilustrativa.

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 teniendo una anchura de la unidad 1110 de codificación. La unidad 1110 de codificación puede ser una unidad de codificación máxima o puede ser una unidad de sub-codificación que tiene un tamaño más pequeño 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 puede ser mayor que la unidad 1120 de predicción. Es decir, las unidades 1130 a 1150 de transformación pueden establecerse mientras no tienen 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 ser mayor que una unidad de codificación de una manera que una pluralidad de unidades de predicción, que se incluyen no en una unidad de codificación sino en una pluralidad de unidades de codificación, se establecen 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 ser igual a o menor que un tamaño de una unidad de codificación o para ser mayor que el tamaño de la unidad de codificación. Es decir, la unidad de transformación puede establecerse mientras no tiene 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 agrupamiento 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 tener formas rectangulares como se ilustra en las Figuras 11A a 11C, pero se establece para tener

cuatro formas cuadradas obtenidas seccionando en cuatro la unidad 1110 de codificación, unidades de predicción superiores e inferiores o unidades de predicción izquierdas y derechas se agrupan 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 existe límite en un criterio por el que la unidad 1010 de selección selecciona las unidades de predicción adyacentes. Sin embargo, de acuerdo con la realización ilustrativa, 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 reducción de tamaño que se realiza gradualmente desde una unidad de codificación máxima de un corte actual o una instantánea actual a una unidad de sub-codificación. Como se ha descrito anteriormente con referencia a las Figuras 3 y 6, a medida que se aumenta la profundidad, se disminuye un tamaño de una unidad de sub-codificación, y por lo tanto también se disminuye una unidad de predicción incluida en la unidad de sub-codificación. 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 ya que información de encabezamiento se añade a cada unidad de transformación.

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

También, de acuerdo con otra realización ilustrativa, 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 establecer las mismas 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 establece las mismas como una unidad de transformación. Ya que existe 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 transformación transforma las unidades de predicción adyacentes en un dominio de frecuencia, de acuerdo con la unidad de transformación. La unidad 1020 de realización de transformación 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 de entropía realiza codificación de entropía en los coeficientes de componente de frecuencia que se cuantifican mediante la unidad 920 de cuantificación. El codificador 930 de entropía puede realizar la codificación de 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 repetitivamente la DCT, la cuantificación y la codificación de entropía en diferentes unidades de transformación. Un procedimiento para la selección de 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 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 ilustrativa.

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

Como se ilustra en la Figura 12, puede predecirse y codificarse una unidad 1210 de codificación a base de una unidad 1220 de predicción que tiene un tamaño más pequeño que la unidad 1210 de codificación. Se realiza una transformación en valores residuales que se generan mediante un resultado de la predicción y en este documento, como se ilustra en la Figura 12, la DCT puede realizarse en los valores residuales a base de las diferentes unidades de transformación.

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 reduciendo a 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.

Una tercera unidad 1250 de transformación ilustrada tiene tamaños obtenidos reduciendo a 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 a base de la cuarta unidad 1260 de transformación ilustrada que tienen el mismo tamaño que la unidad 1220 de predicción.

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

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

El decodificador 1310 de entropía realiza decodificación de entropía en coeficientes de componente de frecuencia con respecto a una unidad de transformación predeterminada. Como se ha descrito anteriormente 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.

Como se ha descrito anteriormente 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 modo de intra predicción o un modo de inter predicción.

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

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

La unidad 1320 de cuantificación inversa cuantifica inversamente los coeficientes de componente de frecuencia que se decodifican en entropía mediante el decodificador 1310 de entropía.

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

El transformador 1330 inverso transforma inversamente los coeficientes de componente de frecuencia inversamente cuantificados en un dominio de píxel. El transformador inverso puede realizar una DCT inversa en coeficientes de coseno discretos inversamente cuantificados (es decir, los coeficientes de componente de frecuencia inversamente cuantificados) 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 ilustrativa.

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 por lo tanto genera coeficientes de coseno discretos.

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

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

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

Haciendo referencia a la Figura 15, en la operación 1510, un aparato de decodificación de imagen realiza decodificación de 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.

En la operación 1520, el aparato de decodificación de imagen cuantifica inversamente los coeficientes de componente de frecuencia que se decodifican en entropía en la operación 1510. El aparato de decodificación de imagen cuantifica inversamente los coeficientes de coseno discretos usando una etapa de cuantificación usada en una operación de codificación.

En la operación 1530, el aparato de decodificación de imagen transforma inversamente los coeficientes de componente de frecuencia, que se han cuantificado inversamente 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 ilustrativas, es posible establecer la unidad de transformación para ser mayor que la unidad de predicción y para realizar la DCT, de modo que una imagen puede comprimirse y codificarse eficientemente.

Las realizaciones ilustrativas también pueden incorporarse 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 posteriormente pueden leerse mediante 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 también puede distribuirse a través de sistemas informáticos acoplados a red de modo que el código legible por ordenador se almacena y ejecuta de una manera distribuida.

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 ejemplares 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 ejemplares pueden incluir una memoria acoplada al al menos un procesador que se acopla al bus para almacenar órdenes, mensajes recibidos o mensajes generados y para ejecutar las órdenes.

La invención queda definida en las reivindicaciones anexas.

REIVINDICACIONES

1. Un procedimiento de decodificación de imágenes que comprende:

recibir información acerca de un tamaño máximo de una unidad de codificación;

5 dividir una imagen en una pluralidad de unidades de codificación máximas cuadradas usando la información acerca del tamaño máximo de la unidad de codificación;

determinar la unidad de codificación dentro de una estructura jerárquica de una unidad de codificación máxima, entre la pluralidad de unidades de codificación máximas usando información acerca de las unidades de codificación analizadas por un flujo de bits;

10 analizar, a partir del flujo de bits, coeficientes de transformación cuantificados de al menos una unidad de transformación en la unidad de codificación;

cuantificar de manera inversa los coeficientes de transformación cuantificados para generar coeficientes de transformación de la al menos una unidad de transformación; y

transformar de manera inversa los coeficientes de transformación para generar componentes residuales de la al menos una unidad de transformación,

15 en el que, la unidad de codificación máxima se divide jerárquicamente en una o varias unidades de codificación cuadradas de profundidad que incluyen al menos una de una unidad de codificación de una profundidad actual y una unidad de codificación de una profundidad superior a la profundidad actual,

20 una unidad de codificación de la profundidad actual se divide en cuatro unidades de codificación cuadradas de una profundidad superior a la profundidad actual, independientemente de las unidades de codificación vecinas de la profundidad actual, y

una unidad de codificación de la profundidad superior a la profundidad actual se predice usando al menos una unidad de predicción y se transforma de manera inversa usando al menos una unidad de transformación,

25 en el que la al menos una unidad de predicción es uno entre los bloques que incluyen: un bloque igual en tamaño a la unidad de codificación de la profundidad superior a la profundidad actual; y un bloque entre una pluralidad de bloques generados al dividir de igual manera al menos uno entre la altura y el ancho de la unidad de codificación de la profundidad superior a la profundidad actual,

30 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 superior a la profundidad actual; y un bloque entre una pluralidad de bloques generados al dividir de igual manera al menos uno entre la altura y el ancho de la unidad de codificación de la profundidad superior a la profundidad actual, y **caracterizado porque**

el procedimiento de decodificación soporta 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 entre la al menos una unidad de predicción que tiene un tamaño de $N \times N$.

2. Un aparato de codificación de imagen que comprende:

35 un procesador configurado para generar información acerca de un tamaño máximo de una unidad de codificación, usada para dividir una imagen en una pluralidad de unidades de codificación máximas cuadradas,

generar información acerca de las unidades de codificación usadas para dividir jerárquicamente una unidad de codificación máxima, entre la pluralidad de unidades de codificación máximas, en unidades de codificación cuadradas,

40 generar información acerca de una unidad de predicción usada para determinar al menos una unidad de predicción para predecir una unidad de codificación entre las unidades de codificación,

generar información acerca de una unidad de transformación usada para determinar al menos una unidad de transformación para transformar las unidades de codificación; y

45 un codificador a configurar para realizar intra-predicción o inter-predicción sobre la al menos una unidad de predicción para generar un predictor,

generar coeficientes de transformación basados en el predictor y la imagen, y generar un flujo de bits que incluye información acerca del tamaño máximo de la unidad de codificación, la información acerca de las unidades de codificación, la información acerca de la unidad de predicción, la información acerca de la unidad de transformación y los coeficientes de transformación, en el que:

la unidad de codificación máxima se divide jerárquicamente en las unidades de codificación que incluye una unidad de codificación de una profundidad actual y una unidad de codificación de una profundidad superior a la profundidad actual,

5 la unidad de codificación de la profundidad actual se divide en cuatro unidades de codificación de la profundidad superior a la profundidad actual, independientemente de las unidades de codificación vecinas de la profundidad actual, y

la unidad de codificación de la profundidad superior a la profundidad actual se predice usando la al menos una unidad de predicción y se transforma usando al menos una unidad de transformación,

10 en el que la al menos una unidad de predicción es uno entre los bloques que incluyen: un bloque igual en tamaño a la unidad de codificación de la profundidad superior a la profundidad actual; y un bloque entre una pluralidad de bloques generados al dividir de igual manera al menos uno entre la altura y el ancho de la unidad de codificación de la profundidad superior a la profundidad actual,

15 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 superior a la profundidad actual; y un bloque entre una pluralidad de bloques generados al dividir de igual manera al menos uno entre la altura y el ancho de la profundidad superior a la unidad de codificación de la profundidad actual, y

caracterizado porque el aparato de codificación 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 $N \times N$.

20 3. Un medio de almacenamiento legible por ordenador no transitorio que almacena un flujo bits que comprende: información acerca de un tamaño máximo de una unidad de codificación, usada para dividir una imagen en una pluralidad de unidades de codificación máximas cuadradas;

información acerca de las unidades de codificación usadas para dividir jerárquicamente una unidad de codificación máxima, entre la pluralidad de unidades de codificación máximas, en unidades de codificación cuadradas;

25 información acerca de una unidad de predicción usada para determinar al menos una unidad de predicción para predecir una unidad de codificación entre las unidades de codificación,

información acerca de una unidad de transformación usada para determinar al menos una unidad de transformación para transformar la unidad de codificación; y

30 coeficientes de transformación generados basándose en un predictor obtenido por intra-predicción o inter-predicción sobre la al menos una unidad de predicción, y la imagen, en el que:

la unidad de codificación máxima de las unidades de codificación máximas se divide jerárquicamente en las unidades de codificación que incluye una unidad de codificación de una profundidad actual y una unidad de codificación de una profundidad superior a la profundidad actual,

35 la unidad de codificación de la profundidad actual se divide en cuatro unidades de codificación de la profundidad superior a la profundidad actual, independientemente de las unidades de codificación vecinas de la profundidad actual,

y la unidad de codificación de la profundidad superior a la profundidad actual se predice usando la al menos una unidad de predicción y se transforma usando al menos una unidad de transformación,

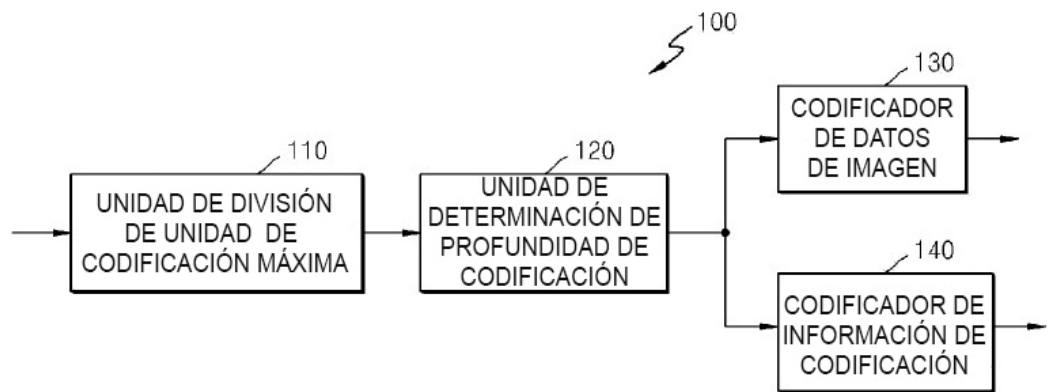
40 en el que la al menos una unidad de predicción es uno entre los bloques que incluyen: un bloque igual en tamaño a la unidad de codificación de la profundidad superior a la profundidad actual; y un bloque entre una pluralidad de bloques generados al dividir de igual manera al menos uno entre la altura y el ancho de la unidad de codificación de la profundidad superior a la profundidad actual,

45 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 superior a la profundidad actual; y un bloque entre una pluralidad de bloques generados al dividir de igual manera al menos uno entre la altura y el ancho de la profundidad superior a la unidad de codificación de la profundidad actual, y

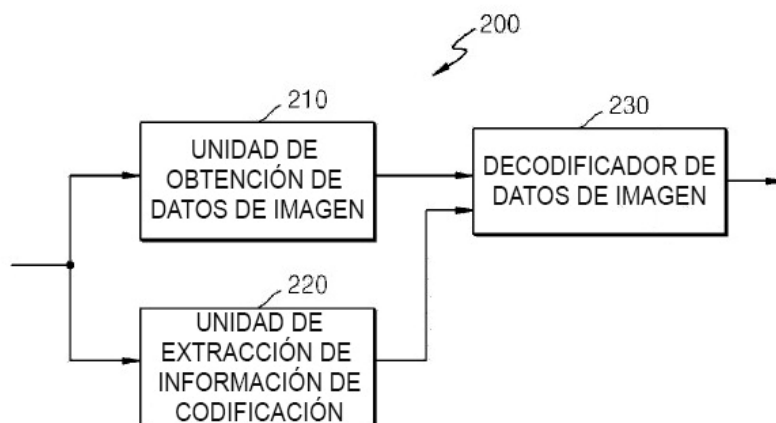
caracterizado porque el flujo de bits soporta una unidad de transformación, entre la al menos una unidad de transformación, que tiene un tamaño de $2N \times 2N$ que incluye cuatro unidades de predicción que tienen un tamaño de $N \times N$.

50

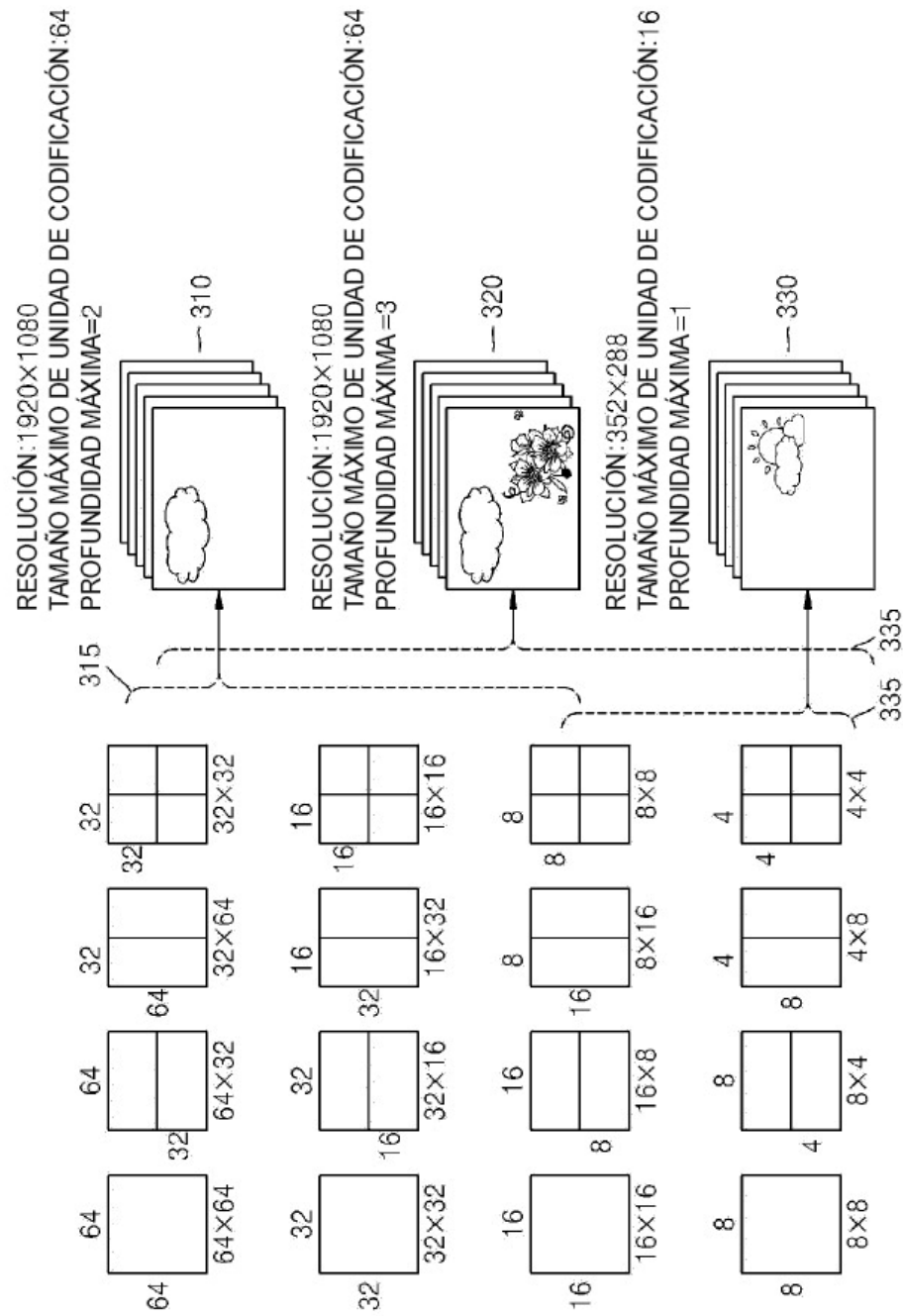
【Figura 1】



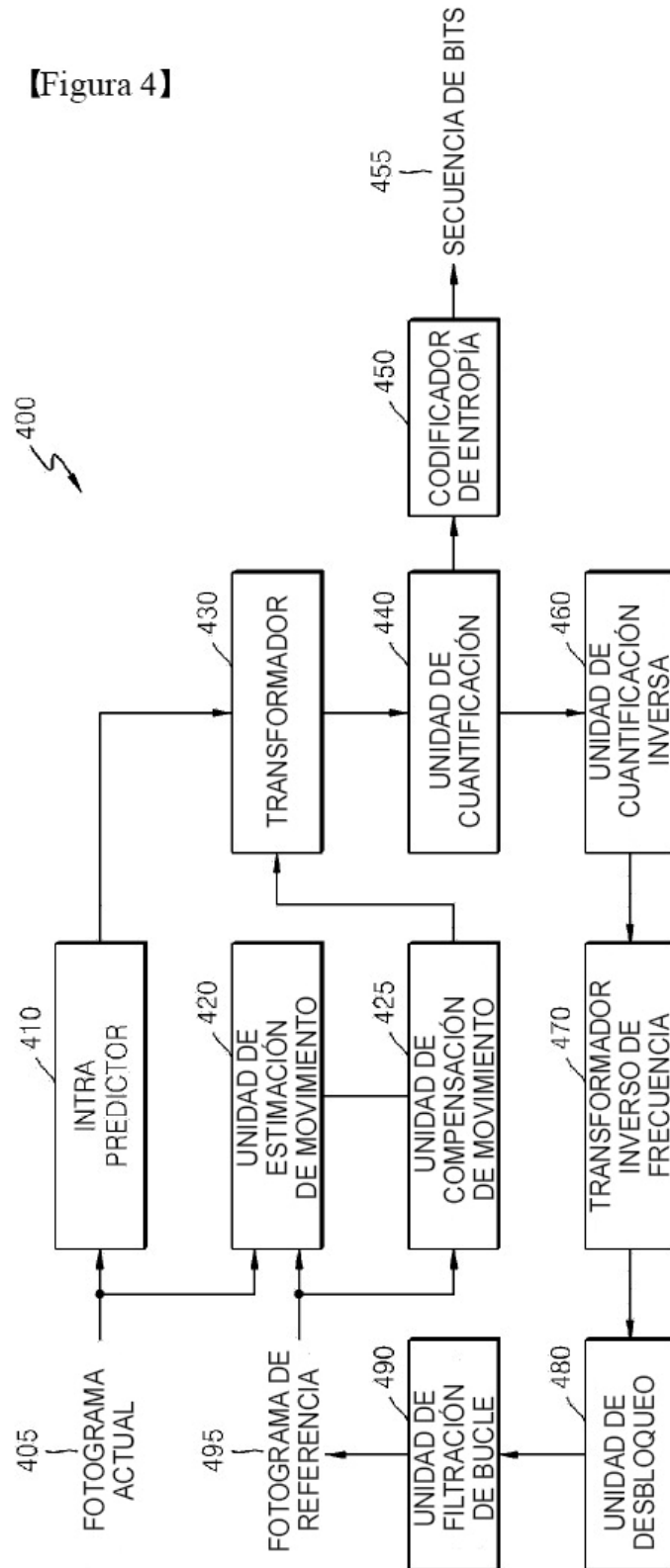
【Figura 2】



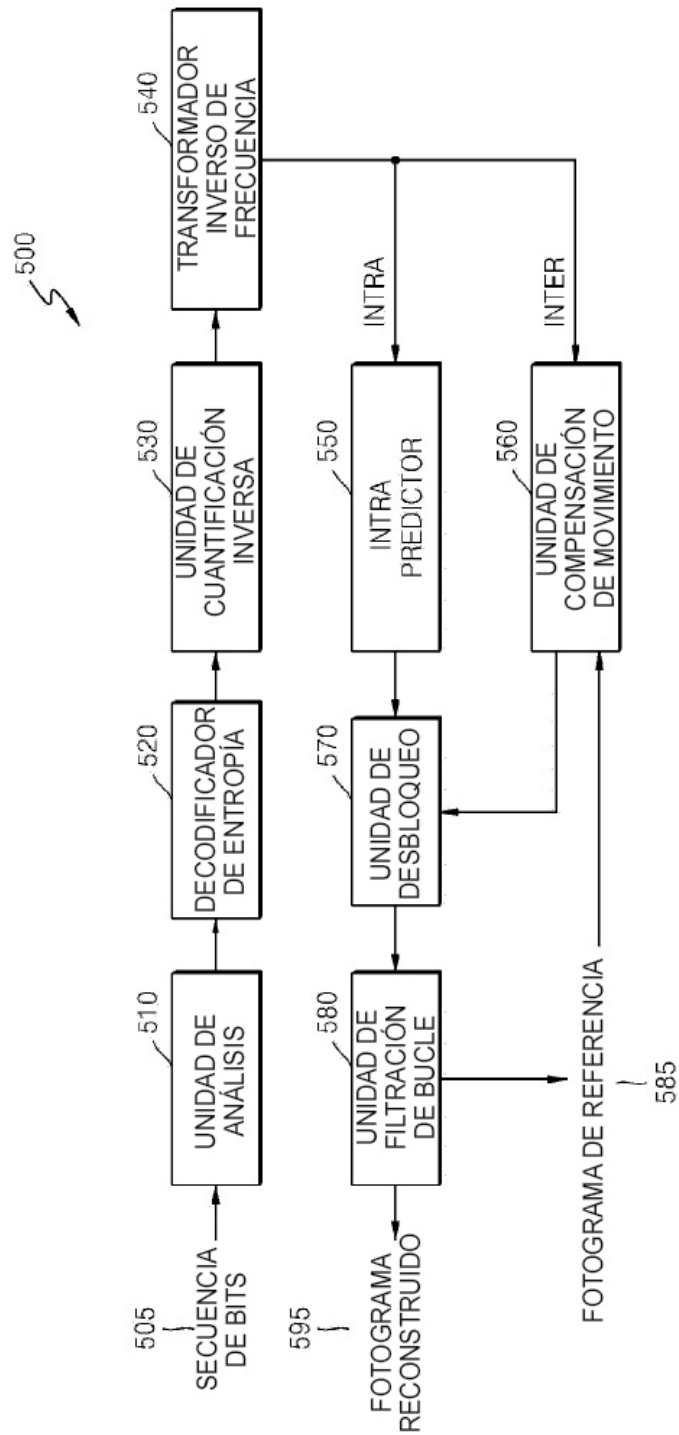
【Figura 3】



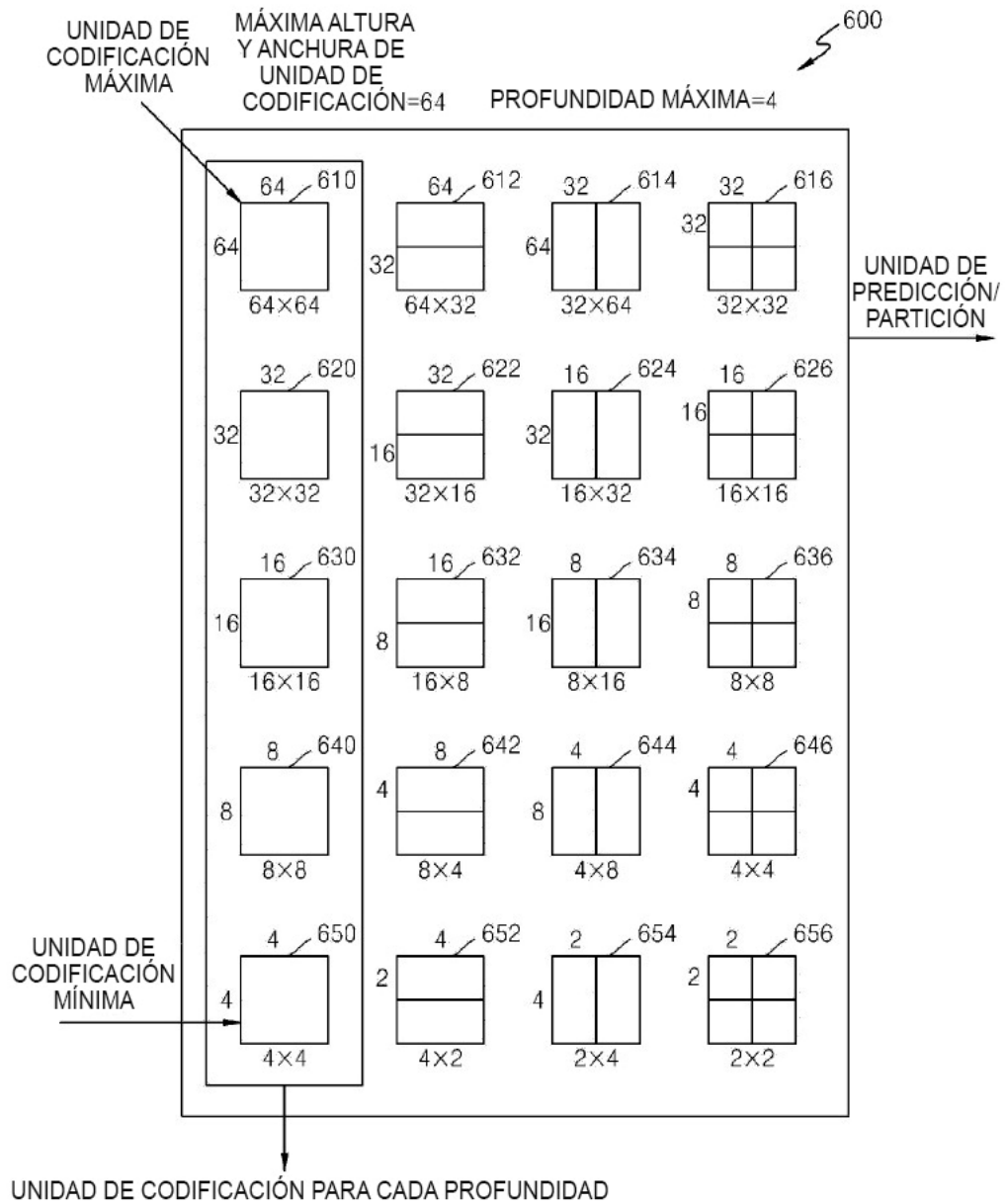
【Figura 4】



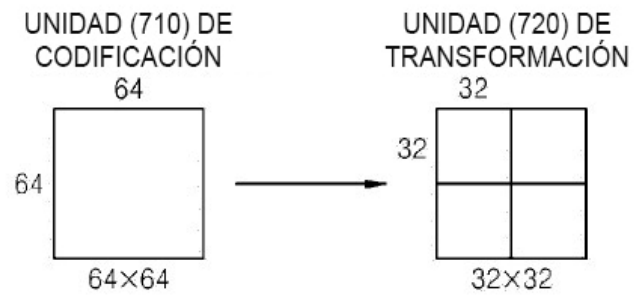
【Figura 5】



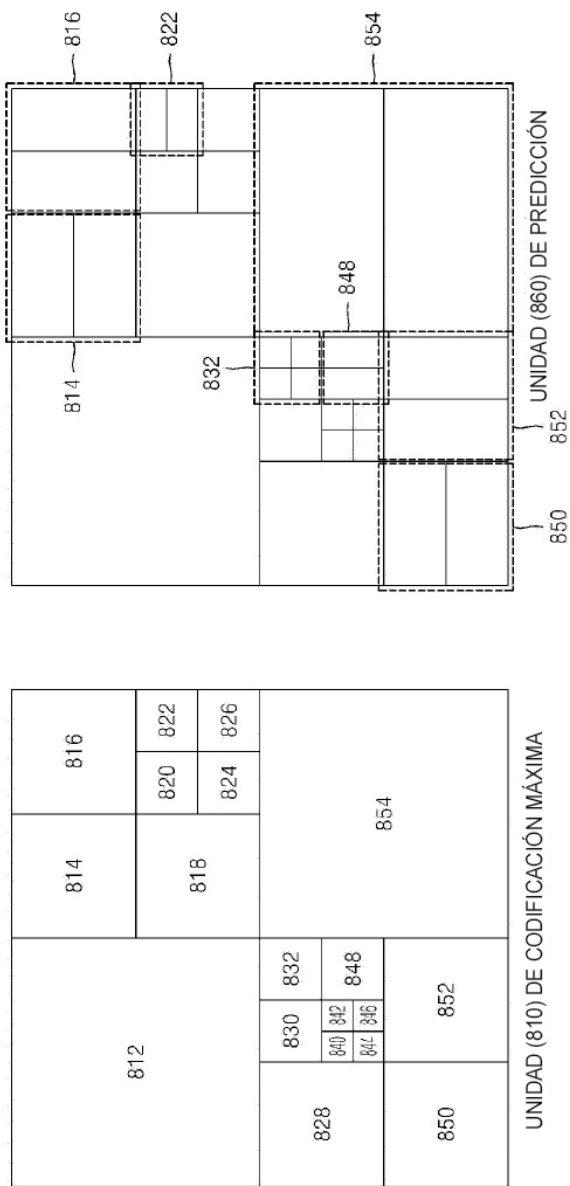
【Figura 6】



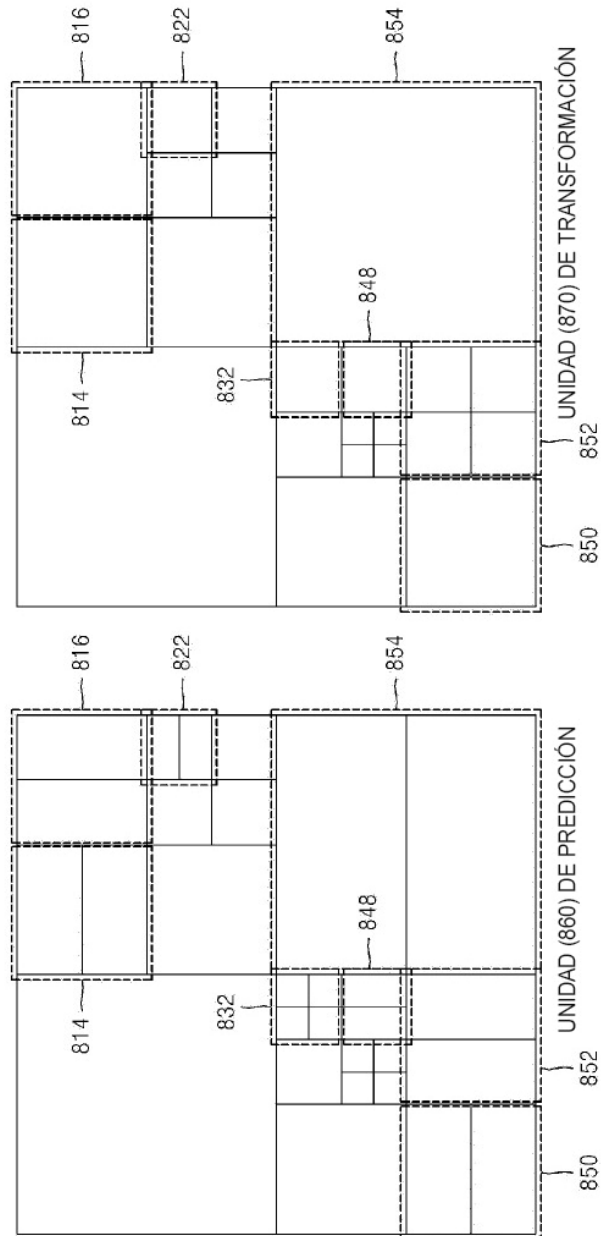
【Figura 7】



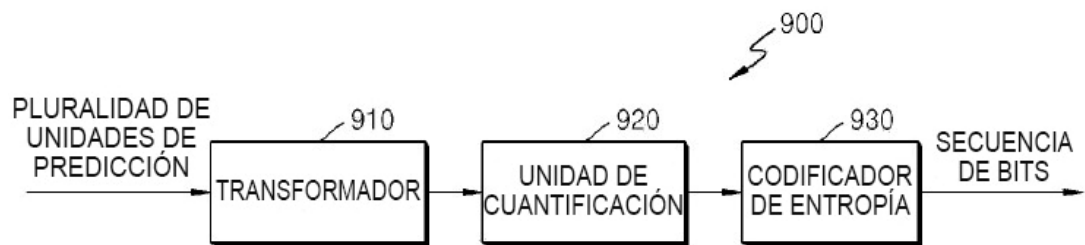
【Figura 8a】



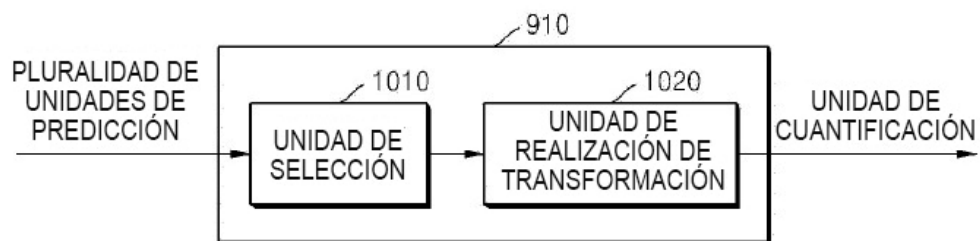
【Figura 8b】



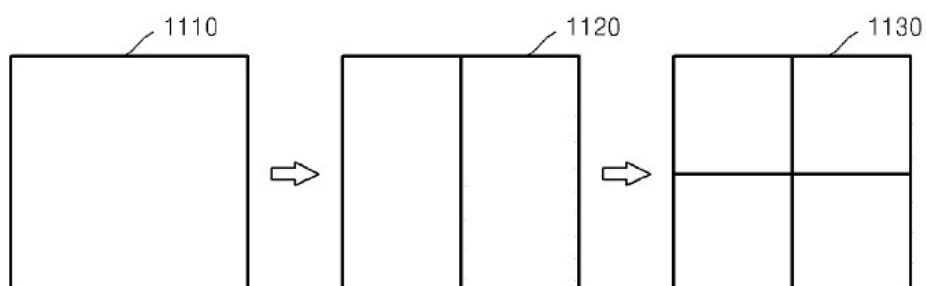
【Figura 9】



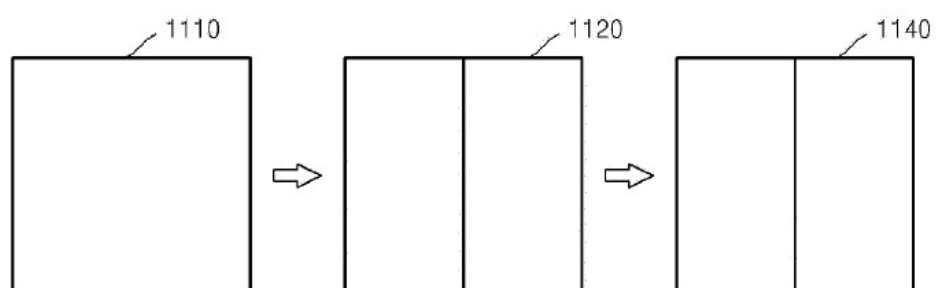
【Figura 10】



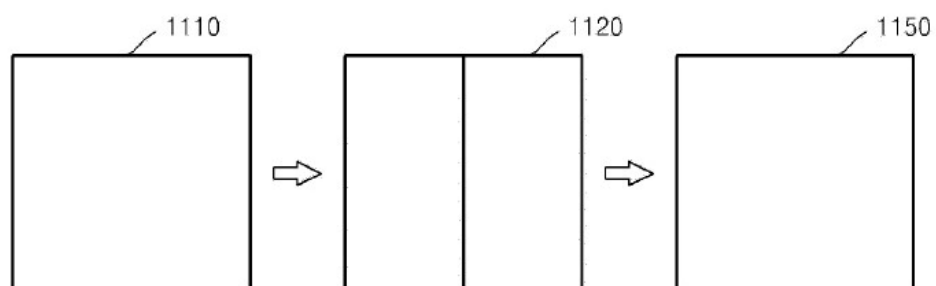
【Figura 11a】



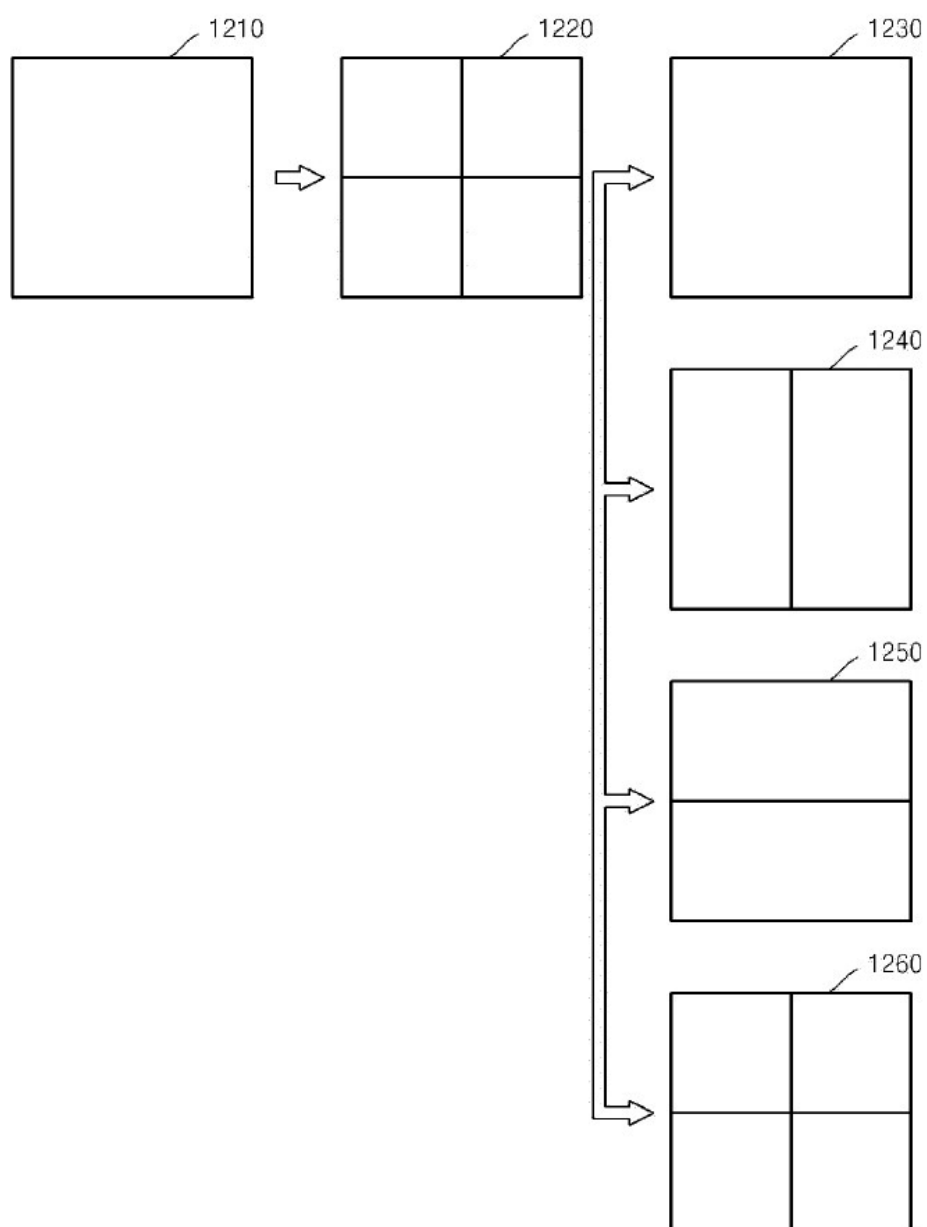
【Figura 11b】



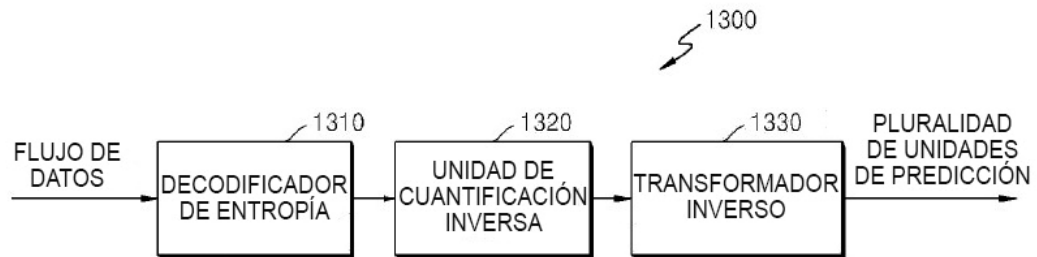
【Figura 11c】



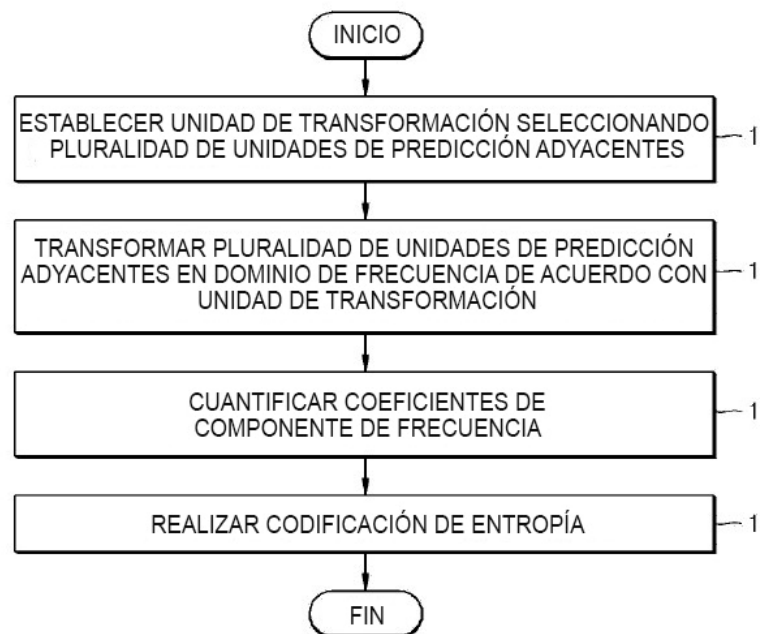
【Figura 12】



【Figura 13】



【Figura 14】



【Figura 15】

