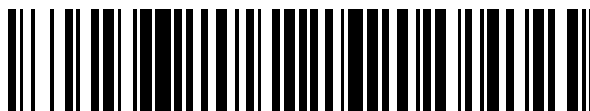


19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 644 002**

51 Int. Cl.:

H04N 19/96	(2014.01)
H04N 19/61	(2014.01)
H04N 19/50	(2014.01)
H04N 19/176	(2014.01)
H04N 19/147	(2014.01)
H04N 19/122	(2014.01)
H04N 19/119	(2014.01)
H04N 19/107	(2014.01)
G06T 9/00	(2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

96 Fecha de presentación y número de la solicitud europea: **14.01.2011 E 15183034 (6)**

97 Fecha y número de publicación de la concesión europea: **11.10.2017 EP 2996340**

54 Título: **Procedimiento y aparato de codificación y decodificación de imágenes usando unidades de transformada grandes**

30 Prioridad:

14.01.2010 KR 2010003558

45 Fecha de publicación y mención en BOPI de la traducción de la patente:
27.11.2017

73 Titular/es:

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

72 Inventor/es:

**CHEON, MIN-SU;
HAN, WOO-JIN;
JUNG, HAE-KYUNG;
KIM, IL-KOO;
LEE, TAMMY y
CHEN, JIANLE**

74 Agente/Representante:

CARPINTERO LÓPEZ, Mario

ES 2 644 002 T3

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

DESCRIPCIÓN

Procedimiento y aparato de codificación y decodificación de imágenes usando unidades de transformada grandes

[Campo técnico]

Las realizaciones ejemplares se refieren a aparatos para decodificar una imagen.

5 **[Antecedentes de la técnica]**

En la mayoría de los procedimientos y aparatos para codificar y decodificar una imagen, una imagen de un dominio de píxel se transforma a un dominio de frecuencia y la imagen transformada se codifica para comprimir la imagen. La transformada de coseno discreta (DCT) es una tecnología bien conocida usada para comprimir datos de audio/video (AV). En los últimos años, se han realizado muchos intentos para hallar procedimientos de codificación más eficaces. En codificación de audio, la codificación paramétrica rinde mejor que la DCT y, en datos bidimensionales, la transformada de Karhunen Loeve (KLT) tiene un tamaño de bits mínimo pero tiene un tamaño de tara grande.

10

El artículo WIEGAND T Y COL, "Overview of the H.264/AVC video coding standard", IEEE TRANSACTIONS ON CIRCUITS AND SYSTEMS FOR VIDEO TECHNOLOGY, IEEE SERVICE CENTER, PISCATAWAY, NJ, US, vol. 13, n.º 7, páginas 560 - 576 proporciona una vista general de las características técnicas de H.264/AVC, describe perfiles y aplicaciones para la norma, y esboza el historial del procedimiento de normalización. Los objetivos principales del esfuerzo de normalización de H.264/AVC han sido potenciar el rendimiento de compresión y provisión de una representación de vídeo "amigable para la red" que trata las aplicaciones "conversacionales" (telefonía de vídeo) y "no conversacionales" (almacenamiento, difusión o flujo continuo).

15

20 **[Divulgación de la invención]**

[Solución al problema]

Las realizaciones ejemplares proporcionan un aparato para decodificar una imagen usando transformada eficaz.

[Efectos ventajosos de la invención]

De acuerdo con las realizaciones ejemplares, una imagen se comprime y codifica más eficazmente puesto que una unidad de transformada puede establecerse para que tenga un tamaño mayor que una unidad de predicción, y la transformada puede realizarse en la unidad de transformada.

25

[Breve descripción de los dibujos]

Los anteriores y/u otros aspectos se harán más evidentes describiendo ciertas realizaciones ejemplares, con referencia a los dibujos adjuntos, en los que:

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

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

La Figura 3 ilustra unidades de codificación jerárquica;

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

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

La Figura 6 ilustra una unidad de codificación máxima, una sub unidad de codificación y una unidad de predicción;

La Figura 7 ilustra una unidad de codificación y una unidad de transformada;

40 Las Figuras 8A, 8B, 8C y 8D ilustran formas de división de una unidad de codificación, una unidad de predicción y una unidad de transformada;

La Figura 9 es un diagrama de bloques de un aparato para codificar una imagen;

La Figura 10 es un diagrama para describir un procedimiento de predicción;

La Figura 11 es un diagrama de bloques de un transformador;

Las Figuras 12A a 12C son diagramas de tipos de unidades de transformada;

45 Las Figuras 13A a 13D son diagramas de tipos de unidades de transformada;

La Figura 14 es un diagrama de diferentes unidades de transformada;

La Figura 15 es un diagrama de bloques de un aparato para decodificar una imagen, de acuerdo con otra realización ejemplar;

La Figura 16 es un diagrama de flujo que ilustra un procedimiento de codificación de una imagen; y

5 La Figura 17 es un diagrama de flujo que ilustra un procedimiento de decodificación de una imagen.

[Mejor modo para llevar a cabo la invención]

De acuerdo con otro aspecto de una realización ejemplar, se proporciona un aparato para decodificar una imagen, el aparato como se expone en la reivindicación adjunta 1.

[Modo para la invención]

10 Ciertas realizaciones ejemplares se describen en mayor detalle a continuación con referencia a los dibujos adjuntos. Expresiones tales como “al menos uno de”, cuando preceden una lista de elementos, modifican toda la lista de elementos y no modifican los elementos individuales de la lista. En la presente memoria descriptiva, una “imagen” puede indicar una imagen fija para un vídeo o una imagen en movimiento, es decir, el mismo vídeo.

15 En la siguiente descripción, se usan números de referencia similares para los elementos similares, incluso en diferentes dibujos. Las materias definidas en la descripción, tales como la construcción y elementos detallados, se proporcionan para ayudar a un entendimiento comprensivo de las realizaciones ejemplares. Sin embargo, las realizaciones ejemplares pueden ponerse en práctica sin estas materias específicamente definidas.

20 La Figura 1 es un diagrama de bloques de un aparato 100 de codificación de imagen para codificar una imagen. El aparato 100 de codificación de imagen puede implementarse como un aparato de hardware tal como, por ejemplo, un procesador de un ordenador o un sistema informático. El aparato 100 de codificación de imagen puede implementarse también como un módulo de software que reside en el sistema informático.

25 Haciendo referencia a la Figura 1, el aparato 100 de codificación de imagen incluye un divisor 110 de unidad de codificación máxima, un determinador 120 de profundidad de codificación, un codificador 130 de datos de imagen y un codificador 140 de información de codificación que pueden implementarse, por ejemplo, como módulos de hardware o software integrados dentro del aparato 100 de codificación de imagen o de manera separada del aparato 100 de codificación de imagen.

El divisor 110 de unidad de codificación máxima puede dividir un fotograma actual o corte basándose en una unidad de codificación máxima que es una unidad de codificación del tamaño más grande. Es decir, el divisor 110 de unidad de codificación máxima puede dividir el fotograma o corte actual en al menos una unidad de codificación máxima.

30 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 las unidades de codificación del fotograma actual, y la profundidad indica un grado de jerarquía que reduce la unidad de codificación. A medida que una profundidad se hace profunda, una unidad de codificación puede reducirse de una unidad de codificación máxima a una unidad de codificación mínima, en el que
35 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 de 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 o mayores que 1).

40 De acuerdo con un aumento del tamaño de un fotograma a codificarse, codificar una imagen en una unidad de codificación mayor puede provocar una tasa de compresión de imagen superior. Sin embargo, si se fija una unidad de codificación mayor, una imagen puede no codificarse eficazmente reflejando de manera continua las características de la imagen cambiantes.

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

50 Por consiguiente, se establece una unidad de codificación de imagen máxima diferente y una profundidad máxima diferente para cada fotograma o corte. Puesto que una profundidad máxima indica el número máximo de veces en las que se puede reducir 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. La profundidad máxima puede determinarse de manera diferente para cada fotograma o corte o para cada unidad de codificación máxima.

El determinador 120 de profundidad de codificación determina una forma de división de la unidad de codificación máxima. La forma de división puede determinarse basándose en el cálculo de costes de tasa-distorsión (RD). La forma de división determinada de la unidad de codificación máxima 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.

Una unidad de codificación máxima puede dividirse en sub unidades de codificación que tienen diferentes tamaños de acuerdo con diferentes profundidades, y las sub unidades de codificación que tienen diferentes tamaños, que se incluyen en la unidad de codificación máxima, pueden predecirse o transformarse en frecuencia basándose en unidades de procesamiento que tienen diferentes tamaños. En otras palabras, el aparato 100 de codificación de imagen 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 o diferentes tamaños pueden usarse para cada operación.

Por ejemplo, el aparato 100 de codificación de imagen 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, la predicción de movimiento puede realizarse basándose en una unidad de procesamiento que tiene una forma, mediante la cual al menos una de una altura y una 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 específico o una forma específica. Por ejemplo, el intra modo puede realizarse para únicamente unidades de predicción que tienen los tamaños de $2N \times 2N$ o $N \times N$ y la forma de 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 menores errores de codificación puede seleccionarse después de realizar predicción para cada unidad de predicción.

Como alternativa, el aparato 100 de codificación de imagen puede realizar transformada de frecuencia en datos de imagen basándose en una unidad de procesamiento que tiene un tamaño diferente de un tamaño de la unidad de codificación. Para la transformada de frecuencia en la unidad de codificación, la transformada 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 transformada de frecuencia, se define como una unidad de transformada. La transformada de frecuencia puede ser la transformada de coseno discreta (DCT) o la transformada de Karhunen Loeve (KLT).

El determinador 120 de profundidad de codificación puede determinar sub unidades de codificación incluidas en una unidad de codificación máxima que usa optimización de RD basándose en un multiplicador de Lagrange. En otras palabras, el determinador 120 de profundidad de codificación puede determinar una forma de una pluralidad de sub unidades de codificación divididas a partir de la unidad de codificación máxima, en el que las sub unidades de codificación tienen diferentes tamaños de acuerdo con las profundidades de las sub unidades de codificación. El codificador 130 de datos de imagen emite una secuencia de bits codificando la unidad de codificación máxima basándose en la forma de las divisiones determinadas por el determinador 120 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 determinado por el determinador 120 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 sub unidad de codificación para cada profundidad. La información acerca del modo de codificación de la sub unidad de codificación puede incluir información acerca de una unidad de predicción de la sub unidad de codificación, información acerca de un modo de predicción para cada unidad de predicción e información acerca de una unidad de transformada de la sub unidad de codificación.

La información acerca de la forma de división de la unidad de codificación máxima puede ser información de bandera, que indica si se divide cada unidad de codificación. Por ejemplo, cuando se divide y codifica la unidad de codificación máxima, se codifica información que indica si se divide la unidad de codificación máxima. También, cuando se dividen y codifican unas sub unidades de codificación divididas a partir de la unidad de codificación máxima, se codifica la información que indica si se divide la sub unidad de codificación.

Puesto que existen sub unidades de codificación que tienen diferentes tamaños para cada unidad de codificación máxima y se determina la información acerca de un modo de codificación para cada sub unidad de codificación, puede determinarse información acerca de al menos un modo de codificación para una unidad de codificación

máxima.

El aparato 100 de codificación de imagen puede generar sub unidades de codificación dividiendo igualmente la altura y anchura de una unidad de codificación máxima en dos de acuerdo con un aumento de 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 $N \times N$.

Por consiguiente, el aparato 100 de codificación de imagen 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 la imagen. Ajustando variablemente el tamaño de una unidad de codificación máxima en consideración de las características de la imagen y codificando una imagen a través de división de una unidad de codificación máxima en sub unidades de codificación de diferentes profundidades, las imágenes que tienen diversas resoluciones pueden codificarse más eficazmente.

La Figura 2 es un diagrama de bloques de un aparato 200 de decodificación de imagen para decodificar una imagen, de acuerdo con una realización ejemplar. El aparato 200 de decodificación de imagen puede implementarse como un aparato de hardware tal como, por ejemplo, un procesador de un ordenador, o un sistema informático. El aparato 200 de decodificación de imagen puede implementarse también como un módulo de software que reside en el sistema informático.

Haciendo referencia a la Figura 2, el aparato 200 de decodificación de imagen incluye una unidad 210 de adquisición de datos de imagen, un extractor 220 de información de codificación y un decodificador 230 de datos de imagen que pueden implementarse, por ejemplo, como módulos de hardware o software integrados dentro del aparato 200 de decodificación de imagen o de manera separada del aparato 200 de codificación de imagen.

La unidad 210 de adquisición 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 de decodificación de imagen y emite los datos de imagen al decodificador 230 de datos de imagen. La unidad 210 de adquisición de datos de imagen puede extraer información acerca de una unidad de codificación máxima de un fotograma o corte actual desde un encabezamiento del fotograma o corte actual. En otras palabras, la unidad 210 de adquisición 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.

El extractor 220 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 sub unidades de codificación desde el encabezamiento del fotograma actual analizando la secuencia de bits recibida por el aparato 200 de decodificación de imagen. 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 sub unidades de codificación que tienen diferentes tamaños de acuerdo con las profundidades e incluirse en la unidad de codificación máxima, y puede ser información de bandera que indica si se divide cada unidad de codificación.

La información acerca de un modo de codificación puede incluir información acerca de una unidad de predicción de acuerdo con sub unidades de codificación, información acerca de un modo de predicción e información acerca de una unidad de transformada.

El decodificador 230 de datos de imagen restaura el fotograma actual decodificando datos de imagen de cada unidad de codificación máxima basándose en la información extraída por el extractor 220 de información de codificación.

El decodificador 230 de datos de imagen puede decodificar sub unidades 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 transformada 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 transformada inversa para cada sub unidad de codificación basándose en información acerca de una unidad de transformada de una sub unidad de codificación.

La Figura 3 ilustra unidades de codificación jerárquica.

Haciendo referencia a la Figura 3, las unidades de codificación jerárquica pueden incluir unidades de codificación cuyas anchuras y alturas son 64×64 , 32×32 , 16×16 , 8×8 y 4×4 . Además de estas unidades de codificación que tienen formas cuadradas perfectas, pueden existir también las unidades de codificación cuyas anchuras y alturas son

64x32, 32x64, 32x16, 16x32, 16x8, 8x16, 8x4 y 4x8.

Haciendo referencia a la Figura 3, para el conjunto 310 de datos 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.

5 Para el conjunto 320 de datos 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 3. Para el conjunto 330 de datos 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.

10 Cuando la resolución es alta o la cantidad de datos es grande, un tamaño máximo de una unidad de codificación puede establecerse relativamente grande para aumentar una relación de compresión y reflejar características de imagen con mayor precisión. Por consiguiente, para los conjuntos 310 y 320 de datos de imágenes que tienen resolución superior que el conjunto 330 de datos de imagen, puede seleccionarse 64x64 como el tamaño de una unidad de codificación máxima.

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

20 Por otra parte, puesto que la profundidad máxima del conjunto 330 de datos de imagen es 1, una unidad 335 de codificación del conjunto 330 de datos de imagen puede incluir una unidad de codificación máxima cuyo tamaño de eje más largo es 16 y unidades de codificación cuyo tamaño de eje más largo es 8, de acuerdo con un aumento de una profundidad.

25 Sin embargo, puesto que la profundidad máxima de los datos 320 de imagen es 3, una unidad 325 de codificación del conjunto 320 de datos de imagen puede incluir una unidad de codificación máxima cuyo tamaño de eje más largo es 64 y sub unidades de codificación cuyos tamaños de eje más largos son 32, 16, 8 y 4 de acuerdo con un aumento de una profundidad. Puesto que una imagen se codifica basándose en una sub unidad de codificación más pequeña a medida que una profundidad se hace profunda, hay ejemplos adecuados para codificar una imagen que incluye escenas de más minutos.

30 La Figura 4 es un diagrama de bloques de un codificador 400 de imagen basándose en una unidad de codificación. El codificador 400 de imagen puede implementarse como un dispositivo de hardware tal como, por ejemplo, un procesador de un ordenador o como un módulo de software que reside en el sistema informático.

35 Un intra predictor 410 realiza intra predicción en unidades de predicción del intra modo en un fotograma 405 actual, y un estimador 420 de movimiento y un compensador 425 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. El intra predictor 410, el estimador 420 de movimiento, el compensador 425 de movimiento y el fotograma 495 de referencia pueden implementarse, por ejemplo, como módulos de hardware o software integrados dentro del codificador 400 de imagen o de manera separada del codificador 400 de imagen.

Se generan valores residuales basándose en las unidades de predicción emitidas desde el intra predictor 410, el estimador 420 de movimiento y el compensador 425 de movimiento. Los valores residuales generados se emiten como coeficientes de transformada cuantificados pasando a través de un transformador 430 y un cuantificador 440.

40 Los coeficientes de transformada cuantificados se restauran a valores residuales pasando a través de un cuantificador 460 inverso y un transformador 470 inverso, y los valores residuales restaurados se post-procesan pasando a través de una unidad 480 de desbloqueo y una unidad 490 de filtración en 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.

45 Para realizar la codificación basándose en un procedimiento de codificación, el intra predictor 410, el estimador 420 de movimiento, el compensador 425 de movimiento, el transformador 430, el cuantificador 440, el codificador 450 por entropía, el cuantificador 460 inverso, el transformador 470 inverso, la unidad 480 de desbloqueo, y la unidad 490 de filtración en bucle del codificador 400 de imagen realizan procedimientos de codificación de imagen basándose en una unidad de codificación máxima, una sub unidad de codificación de acuerdo con las profundidades, una unidad de predicción y una unidad de transformada.

50 La Figura 5 es un diagrama de bloques de un decodificador 500 de imagen basándose en una unidad de codificación, de acuerdo con una realización ejemplar. El decodificador 500 de imagen puede implementarse como un dispositivo de hardware tal como, por ejemplo, un procesador de un ordenador o como un módulo de software que reside en el sistema informático.

55

- Una secuencia de bits 505 pasa a través de un analizador 510 de modo que se analizan los datos de imagen codificados a decodificarse y la información de codificación necesaria para la decodificación. Los datos de imagen codificados se emiten como datos cuantificados a la inversa pasando a través de un decodificador 520 por entropía y un cuantificador 530 inverso y se restauran a valores residuales pasando a través de un transformador 540 inverso.
- 5 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 un compensador 560 de movimiento. Las unidades 585, 595 de codificación restauradas se usan para predicción de las siguientes unidades de codificación o de un siguiente fotograma pasando a través de una unidad 570 de desbloqueo y una unidad 580 de filtración en bucle.
- 10 El analizador 510, el decodificador 520 por entropía, el cuantificador 530 inverso, el transformador 540 inverso, el intra predictor 550, el compensador 560, la unidad 570 de desbloqueo y la unidad 580 de filtrado en bucle pueden implementarse, por ejemplo, como módulos de hardware o software integrados dentro del decodificador 500 de imagen o de manera separada del decodificador 500 de imagen.
- 15 Para realizar la decodificación basándose en un procedimiento de decodificación de acuerdo con una realización ejemplar, el analizador 510, el decodificador 520 por entropía, el cuantificador 530 inverso, el transformador 540 inverso, el intra predictor 550, el compensador 560 de movimiento, la unidad 570 de desbloqueo y la unidad 580 de filtrado en bucle del decodificador 500 de imagen realizan procedimientos de decodificación de imagen basándose en una unidad de codificación máxima, una sub unidad de codificación de acuerdo con las profundidades, una unidad de predicción y una unidad de transformada.
- 20 En particular, el intra predictor 550 y el compensador 560 de movimiento determinan una unidad de predicción y un modo de predicción en una sub unidad de codificación considerando una unidad de codificación máxima y una profundidad, y el transformador 540 inverso realiza transformada inversa considerando el tamaño de una unidad de transformada.
- La Figura 6 ilustra una unidad de codificación máxima, una sub unidad de codificación y una unidad de predicción.
- 25 El aparato 100 de codificación de imagen ilustrado en la Figura 1 y el aparato 200 de decodificación de imagen ilustrado en la Figura 2 usan unidades de codificación jerárquica para realizar codificación y decodificación en consideración de las características de la imagen. Una unidad de codificación máxima y una profundidad máxima pueden establecerse de manera adaptativa de acuerdo con las características de la imagen o establecerse de manera variable de acuerdo con los requisitos de un usuario.
- 30 En la Figura 6, una estructura 600 de unidad de codificación jerárquica tiene una unidad 610 de codificación máxima que es una unidad 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 una profundidad se hace profunda, se reducen las alturas y anchuras de las sub unidades 620 a 650 de codificación. Las unidades de predicción de la unidad 610 de codificación máxima y las sub unidades 620 a 650 de codificación se muestran a lo largo de un eje horizontal de la estructura 600 de unidad de codificación jerárquica.
- 35 La unidad 610 de codificación máxima tiene una profundidad de 0 y el tamaño de una unidad de codificación, o una altura y una anchura, de 64x64. Una profundidad aumenta a lo largo del eje vertical, y existe una primera sub unidad 620 de codificación cuyo tamaño es 32x32 y profundidad es 1, una segunda sub unidad 630 de codificación cuyo tamaño es 16x16 y profundidad es 2, una tercera sub unidad 640 de codificación cuyo tamaño es 8x8 y profundidad es 3, y una unidad 650 de codificación mínima cuyo tamaño es 4x4 y profundidad es 4. La unidad 650 de codificación mínima 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 tiene un tamaño menor que la unidad de codificación mínima.
- 40 Haciendo referencia a la Figura 6, se muestran ejemplos de unidades 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 al tamaño 64x64 de la unidad de codificación máxima, 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 el de la unidad de codificación máxima cuyo tamaño es 64x64.
- 45 Una unidad de predicción de la primera sub unidad 620 de codificación cuya profundidad es 1 y tamaño es 32x32 puede ser una unidad de predicción cuyo tamaño es igual al tamaño 32x32 de la primera sub unidad de codificación, 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 el de la primera sub unidad 620 de codificación cuyo tamaño es 32x32.
- 50 Una unidad de predicción de la segunda sub unidad 630 de codificación cuya profundidad es 2 y tamaño es 16x16 puede ser una unidad de predicción cuyo tamaño es igual al tamaño 16x16 de la segunda sub unidad 630 de codificación, 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 el de la segunda sub
- 55

unidad 630 de codificación cuyo tamaño es 16x16.

5 Una unidad de predicción de la tercera sub unidad 640 de codificación cuya profundidad es 3 y tamaño es 8x8 puede ser una unidad de predicción cuyo tamaño es igual al tamaño 8x8 de la tercera sub unidad 640 de codificación 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 el de la tercera sub unidad 640 de codificación cuyo tamaño es 8x8.

10 La unidad 650 de codificación cuya profundidad es 4 y tamaño es 4x4 es una unidad de codificación mínima y una unidad de codificación de una profundidad máxima. Una unidad de predicción de la unidad 650 de codificación mínima 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 transformada.

15 El aparato 100 de codificación de imagen ilustrado en la Figura 1 y el aparato 200 de decodificación de imagen ilustrado en la Figura 2 realizan codificación y decodificación con una unidad de codificación máxima o con sub unidades de codificación, que tienen tamaño igual a o menor que la unidad de codificación máxima, divididas a partir de la unidad de codificación máxima. En el procedimiento de codificación y decodificación, el tamaño de una unidad de transformada para transformada de frecuencia se selecciona para que no sea mayor que el de una unidad de codificación correspondiente. Por ejemplo, si una unidad 710 de codificación actual tiene el tamaño de 64x64, la transformada de frecuencia puede realizarse usando una unidad 720 de transformada que tiene el tamaño de 32x32.

20 Las Figuras 8A, 8B, 8C y 8D ilustran formas de división de una unidad de codificación, una unidad de predicción y una unidad de transformada.

Las Figuras 8A y 8B ilustran respectivamente una unidad de codificación y una unidad de predicción.

25 La Figura 8A muestra una forma de división seleccionada por el aparato 100 de codificación de imagen ilustrado en la Figura 1, para codificar una unidad 810 de codificación máxima. El aparato 100 de codificación de imagen 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 los costes de RD. Cuando es óptimo que la unidad 810 de codificación máxima se codifique, la unidad 810 de codificación máxima puede codificarse sin dividir la unidad 810 de codificación máxima, como se ilustra en las Figuras 8A a 8D.

30 Haciendo referencia a la Figura 8A, la unidad 810 de codificación máxima cuya profundidad es 0 se codifica dividiendo la unidad 810 de codificación máxima en sub unidades 812, 854 de codificación cuyas profundidades son iguales o mayores a 1. Es decir, la unidad 810 de codificación máxima se divide en 4 sub unidades de codificación cuyas profundidades son 1, y todas o algunas de las sub unidades de codificación cuyas profundidades son 1 se dividen en sub unidades 814, 816, 818, 828, 850 y 852 de codificación cuyas profundidades son 2.

35 Una sub unidad de codificación localizada en un lado superior derecha y una sub unidad de codificación localizada en un lado inferior izquierda entre las sub unidades de codificación cuyas profundidades son 1 se dividen en sub unidades de codificación cuyas profundidades son iguales o mayores que 2. Algunas de las sub unidades de codificación cuyas profundidades son iguales o mayores que 2 pueden dividirse adicionalmente en sub unidades 820, 822, 824, 826, 830, 832, 840, 842, 844, 846 y 848 de codificación cuyas profundidades son iguales o mayores que 3.

40 La Figura 8B muestra una forma de división de una unidad de predicción para la unidad 810 de codificación máxima.

Haciendo referencia a la Figura 8B, 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 sub unidades de codificación puede ser menor que una correspondiente sub unidad de codificación.

45 Por ejemplo, una unidad de predicción para una sub unidad 854 de codificación localizada en un lado derecho inferior entre las sub unidades 812, 854 de codificación cuyas profundidades son 1 puede ser menor que la sub unidad 854 de codificación. Además, las unidades de predicción para sub unidades 814, 816, 850 y 852 de codificación de las sub unidades 814, 816, 818, 828, 850 y 852 de codificación cuyas profundidades son 2 pueden ser menores que las sub unidades 814, 816, 850 y 852 de codificación, respectivamente.

50 Además, las unidades de predicción para las sub unidades 822, 832 y 848 de codificación cuyas profundidades son 3 pueden ser menores que las sub unidades 822, 832 y 848 de codificación, respectivamente. Las unidades de predicción pueden tener una forma mediante la cual las respectivas sub unidades de codificación se dividen igualmente en dos en una dirección de altura o anchura o tienen una forma mediante la cual las respectivas sub unidades de codificación se dividen igualmente en cuatro en las direcciones de altura y anchura.

Las Figuras 8C y 8D ilustran una unidad de predicción y una unidad de transformada.

La Figura 8C muestra una forma de división de una unidad de predicción para la unidad 810 de codificación máxima mostrada en la Figura 8B, y la Figura 8D muestra una forma de división de una unidad de transformada de la unidad 810 de codificación máxima.

- 5 Haciendo referencia a la Figura 8D, una forma de división de una unidad 870 de transformada puede establecerse de manera diferente de la unidad 860 de predicción.

Por ejemplo, incluso aunque una unidad de predicción para la sub unidad 854 de codificación cuya profundidad es 1 se seleccione con una forma mediante la cual la altura de la sub unidad 854 de codificación se divide igualmente en dos, una unidad de transformada puede seleccionarse con el tamaño original de la sub unidad 854 de codificación.

- 10 Análogamente, incluso aunque las unidades de predicción para las sub 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 sub unidades 814 y 850 de codificación se dividen igualmente en dos, una unidad de transformada puede seleccionarse con el mismo tamaño que el tamaño original de cada una de las sub unidades 814 y 850 de codificación.

- 15 Una unidad de transformada puede seleccionarse con un tamaño menor que una unidad de predicción. Por ejemplo, cuando se selecciona una unidad de predicción para la sub unidad 852 de codificación cuya profundidad es 2 con una forma mediante la cual la anchura de la sub unidad 852 de codificación se divide igualmente en dos, una unidad de transformada puede seleccionarse con una forma mediante la cual la sub 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.

- 20 Como alternativa, como se describirá con referencia a las Figuras 13A a 13D, una unidad de transformada puede establecerse para que tenga un tamaño mayor que una unidad de codificación, independientemente de la unidad de codificación.

La Figura 9 es un diagrama de bloques de un aparato 900 para codificar una imagen.

- 25 Haciendo referencia a la Figura 9, el aparato 900 de codificación de imagen incluye un predictor 910, un transformador 920, un cuantificador 930 y un codificador 940 por entropía.

- 30 El predictor 910 genera valores residuales realizando intra predicción o inter predicción en una o más unidades de codificación. Como se describirá más adelante, los valores residuales incluidos en una pluralidad de unidades de predicción pueden agruparse en una unidad de transformada y a continuación transformarse a un dominio de frecuencia, y por lo tanto los valores residuales se generan prediciendo la una o más unidades de codificación basándose en la pluralidad de unidades de predicción. La transformada al dominio de la frecuencia puede ser DCT o KLT.

- 35 Como se ha descrito con referencia a la Figura 8A, en el procedimiento de codificación de imagen, una unidad de codificación puede incluir una pluralidad de unidades de predicción. Por lo tanto, el predictor 910 puede predecir cada una de las unidades de predicción, y generar los valores residuales de las unidades de predicción incluidos en la unidad de codificación.

- 40 Como alternativa, la unidad 910 de predicción puede predecir la pluralidad de unidades de codificación todas a la vez. Como se describirá más adelante, una pluralidad de unidades de predicción incluidas en una pluralidad de unidades de codificación pueden agruparse en una unidad de transformada, y por lo tanto se generan valores residuales prediciendo cada una de las unidades de predicción incluidas en las unidades de codificación. Por ejemplo, todas las sub unidades de codificación incluidas en una unidad de codificación máxima pueden predecirse para generar los valores residuales de las unidades de codificación.

- 45 De acuerdo con la tecnología convencional, puesto que se realiza la transformada (por ejemplo DCT o KLT) con un tamaño menor que o igual a una unidad de predicción, se codifica, restaura y a continuación se usa de manera independiente una unidad de predicción predeterminada para predecir una siguiente unidad de predicción. Sin embargo, de acuerdo con un procedimiento de codificación de una imagen, que se describirá más adelante, puesto que la transformada se realiza agrupando unidades de predicción incluidas en una o más unidades de codificación en una unidad de transformada, una unidad de predicción predeterminada no puede codificarse y restaurarse de manera independiente. Esto se describirá en detalle con referencia a la Figura 10.

La Figura 10 es un diagrama para describir un procedimiento de predicción.

- 50 Haciendo referencia a la Figura 10, una unidad 1000 de codificación puede incluir una pluralidad de unidades 1010 a 1040 de predicción. Si se realiza la transformada con un tamaño menor que o igual a una unidad de predicción, como en la tecnología convencional, las unidades 1010 a 1030 de predicción pueden codificarse y restaurarse antes de codificar la unidad 1040 de predicción en un lado inferior derecho.

Por consiguiente, si la unidad 1040 de predicción se ha de predecir mediante intra predicción de acuerdo con la tecnología convencional, la unidad 1040 de predicción se intra predice usando píxeles adyacentes a la unidad 1040 de predicción, de entre los píxeles generados codificando y a continuación restaurando las unidades 1010 a 1030 de predicción.

5 Por otra parte, en este punto, una pluralidad de unidades de predicción se agrupan en una unidad de transformada, y a continuación se realiza la transformada. En este punto, si las unidades 1010 a 1040 de predicción de la Figura 10 se agrupan en una unidad de transformada, la unidad 1040 de predicción en el lado inferior derecha se codifica con las otras unidades 1010 a 1030 de predicción, y por lo tanto las unidades 1010 a 1030 de predicción no se codifican antes de codificar la unidad 1040 de predicción. Por consiguiente, la unidad 1040 de predicción no puede intra
10 predecirse usando los píxeles generados codificando y a continuación restaurando las unidades 1010 a 1030 de predicción.

En consecuencia, la unidad 910 de predicción de la Figura 9 puede predecir la unidad 1040 de predicción usando valores de predicción de las unidades 1010 a 1030 de predicción. La unidad 1040 de predicción en el lado inferior derecha se predice usando los valores de predicción de las unidades 1010 a 1030 de predicción, en lugar de los
15 píxeles generados codificando y a continuación restaurando las unidades 1010 a 1030 de predicción.

En otras palabras, si hay una primera unidad de predicción prevista mediante intra predicción, de entre las unidades de predicción agrupadas en una unidad de transformada, la primera unidad de predicción puede intra predecirse usando valores de predicción de al menos una unidad de predicción adyacente.

20 Como alternativa, las unidades de predicción agrupadas en una unidad de transformada pueden predecirse todas mediante inter predicción. Como se describe con referencia a la Figura 10, puesto que una unidad de predicción que se predice mediante intra predicción está en cuestión mientras se agrupa una pluralidad de unidades de predicción en una unidad de transformada, todas las unidades de predicción agrupadas en la unidad de transformada pueden predecirse usando únicamente inter predicción.

25 Haciendo referencia de vuelta a la Figura 9, el transformador 920 recibe una unidad de procesamiento de imagen en un dominio de píxel, y transforma la unidad de procesamiento de imagen en un dominio de frecuencia. El transformador 920 transforma los valores residuales generados por la unidad 910 de predicción en el dominio de la frecuencia.

30 Como se ha descrito anteriormente, el transformador 920 agrupa las unidades de predicción en una unidad de transformada, y realiza DCT o KLT de acuerdo con la unidad de transformada. Los valores residuales pueden ser valores residuales de una pluralidad de unidades de predicción incluidas en una o más unidades de codificación. Los coeficientes de los componentes de frecuencia se generan como resultado de transformar el dominio de píxel al dominio de la frecuencia.

35 La transformada al dominio de la frecuencia puede realizarse mediante DCT o KLT, y se generan coeficientes de coseno discretos como resultado de la DCT o KLT. Sin embargo, puede usarse cualquier transformada para transformar una imagen en un dominio de píxel al dominio de la frecuencia.

La Figura 11 es un diagrama de bloques del transformador 920.

Haciendo referencia a la Figura 11, el transformador 920 incluye un selector 1110 y un realizador 1120 de transformada.

40 El selector 1110 establece una unidad de transformada seleccionando una pluralidad de unidades de predicción adyacentes. De acuerdo con los aparatos de codificación de imagen convencionales anteriormente descritos, la intra predicción o inter predicción se realiza de acuerdo con una unidad de predicción predeterminada y la DCT o KLT se realiza con un tamaño menor que o igual a la unidad de predicción predeterminada. En otras palabras, los aparatos de codificación de imagen convencionales realizan la DCT o KLT basándose en una unidad de transformada que tiene un tamaño menor que o igual a una unidad de predicción.

45 Sin embargo, se deteriora una relación de compresión de codificación de imagen puesto que aumenta una sobrecarga añadida a medida que se reduce un tamaño de una unidad de transformada debido a la información de encabezamiento para cada unidad de transformada. Por consiguiente, el aparato 900 de codificación de imagen agrupa las unidades de predicción adyacentes en una unidad de transformada, y a continuación realiza la DCT o KLT de acuerdo con la unidad de transformada. Específicamente, puesto que es altamente probable que las
50 unidades de predicción adyacentes tengan valores residuales similares, una relación de compresión de la codificación puede aumentarse de manera destacable cuando se realiza la DCT o KLT de acuerdo con la unidad de transformada generada agrupando las unidades de predicción adyacentes.

Por consiguiente, el selector 1110 selecciona las unidades de predicción para que se agrupen en una unidad de transformada y en las que se ha de realizar DCT o KLT. Las unidades de predicción pueden estar adyacentes entre
55 sí. Esto se describirá en detalle con referencia a las Figuras 12A a 12C y 13A a 13D.

Las Figuras 12A a 12C son diagramas de tipos de unidades 1230 a 1250 de transformada.

Haciendo referencia a las Figuras 12A a 12C, una unidad 1220 de predicción puede tener una forma mediante la cual una unidad 1210 de codificación se divide igualmente en dos en una dirección de anchura. La unidad 1210 de codificación puede ser una unidad de codificación máxima como se ha descrito anteriormente, o una sub unidad de codificación que tiene un tamaño menor que la unidad de codificación máxima.

Incluso cuando la unidad 1210 de codificación y la unidad 1220 de predicción son idénticas, las unidades 1230 a 1250 de transformada pueden ser diferentes. Un tamaño de la unidad 1230 de transformada puede ser menor que el de la unidad 1220 de predicción como se muestra en la Figura 12A, o un tamaño de la unidad 1240 de transformada puede ser idéntico al de la unidad 1220 de predicción como se muestra en la Figura 12B. Como alternativa, un tamaño de la unidad 1250 de transformada puede ser mayor que el de la unidad 1220 de predicción como se muestra en la Figura 12C.

Las unidades de predicción agrupadas en una unidad de transformada pueden ser una pluralidad de unidades de predicción incluidas en una unidad de codificación como se muestra en las Figuras 12A a 12C, o pueden ser una pluralidad de unidades de predicción incluidas en diferentes unidades de codificación. En otras palabras, una pluralidad de unidades de predicción incluidas en al menos una unidad de codificación pueden agruparse en una unidad de transformada y a continuación transformarse al dominio de la frecuencia.

Las Figuras 13A a 13D son diagramas de tipos de unidades de transformada.

Una unidad 1300 de codificación máxima puede dividirse en sub unidades 1302 a 1308 de codificación que tienen diferentes tamaños y a continuación codificarse como se muestra en la Figura 13A, y cada una de las sub unidades 1302 a 1308 de codificación puede incluir al menos una unidad 1310 a 1340 de predicción, como se muestra en la Figura 13B.

El selector 1110 puede agrupar las unidades 1310 a 1340 de predicción mostradas en la Figura 13B en una unidad 1350 de transformada mostrada en la Figura 13C, y a continuación transformar la unidad 1350 de transformada en el dominio de la frecuencia.

Como alternativa, el selector 1110 puede agrupar las unidades 1310 y 1330 a 1339 de predicción de las sub unidades 1302 y 1306 de codificación a la izquierda en una unidad 1360 de transformada, y agrupar las unidades 1320 a 1328 y 1340 de predicción de las sub unidades 1304 y 1308 de codificación a la derecha en una unidad 1362 de transformada, como se muestra en la Figura 13D.

Haciendo referencia de vuelta a la Figura 11, un criterio para que el selector 1110 seleccione una pluralidad de unidades de predicción adyacentes no está limitado. Sin embargo, en un ejemplo el selector 1110 puede seleccionar una unidad de transformada basándose en una profundidad. Como se ha descrito anteriormente, la profundidad indica un grado de jerarquía que reduce una unidad de codificación desde una unidad de codificación máxima de un corte o fotograma actual a sub unidades de codificación. Como se ha descrito con referencia a las Figuras 3 y 6, a medida que una profundidad se hace profunda, se reduce un tamaño de una sub unidad de codificación, y por lo tanto se reduce un tamaño de una unidad de predicción incluido en la sub unidad de codificación. En este punto, cuando se realiza la DCT o KLT de acuerdo con una unidad de transformada que tiene un tamaño menor que o igual a una unidad de predicción, se reduce una relación de compresión de la codificación de imagen puesto que se añade información de encabezamiento para cada unidad de transformada como se ha descrito anteriormente.

Por consiguiente, las unidades de predicción incluidas en una sub unidad de codificación cuya profundidad es igual o por encima de un valor predeterminado pueden agruparse en una unidad de transformada, y a continuación puede realizarse la DCT o KLT en la unidad de transformada. Por lo tanto, el selector 1110 puede establecer la unidad de transformada basándose en la profundidad de la sub unidad de codificación. Por ejemplo, cuando una profundidad de la unidad 1210 de codificación de la Figura 12C es superior a k , el selector 1110 agrupa las unidades 1220 de predicción en una unidad 1250 de transformada.

Como alternativa, cuando una unidad de codificación máxima incluye una pluralidad de sub unidades de codificación cuyas profundidades son iguales o por encima de un valor predeterminado, el selector 1110 puede agrupar unidades de predicción de las sub unidades de codificación en una unidad de transformada. La Figura 13C ilustra un ejemplo de agrupación de unidades de predicción de sub unidades de codificación cuya profundidad es mayor que una unidad de codificación máxima, es decir, cuya profundidad es mayor que 1, en una unidad de transformada.

De acuerdo con otro ejemplo, el selector 1110 puede establecer una pluralidad de unidades de predicción adyacentes, en las que se realiza predicción de acuerdo con un mismo tipo de modo de predicción, en una unidad de transformada. Las unidades de predicción adyacentes que se predicen usando intra predicción o inter predicción se agrupan en una unidad de transformada. Puesto que es altamente probable que las unidades de predicción adyacentes que se predicen de acuerdo con el mismo tipo de modo de predicción tengan valores residuales similares, puede realizarse la DCT o KLT agrupando las unidades de predicción adyacentes en una unidad de transformada.

Cuando el selector 1110 establece la unidad de transformada, el realizador 1120 de transformada transforma las unidades de predicción adyacentes en un dominio de frecuencia de acuerdo con la unidad de transformada establecida. Los coeficientes del dominio de frecuencia (por ejemplo coeficientes de coseno discreto) se generan transformando las unidades de predicción seleccionadas en una unidad de transformada.

- 5 Haciendo referencia de vuelta a la Figura 9, el cuantificador 930 cuantifica los coeficientes de componente de frecuencia generados por el transformador 920. El cuantificador 930 puede cuantificar la entrada de los coeficientes de acuerdo con un procedimiento de cuantificación predeterminado.

10 El codificador 940 por entropía codifica por entropía los coeficientes cuantificados por el cuantificador 930. En este punto, los coeficientes de coseno discreto pueden codificarse por entropía usando codificación binaria aritmética adaptativa según contexto (CABAC) o codificación de longitud variable adaptativa según contexto (CAVLC).

15 El aparato 900 de codificación de imagen puede codificar información de bandera que indica si la unidad de transformada generada agrupando las unidades de predicción incluye los coeficientes. Si no hay coeficientes a codificar por entropía, es decir, cuando los coeficientes cuantificados son todos '0', indicando la información de bandera que se codifica la unidad de transformada que no incluye los coeficientes, y los coeficientes cuantificados no se codifican por entropía de manera separada.

20 El aparato 900 de codificación de imagen de acuerdo con el ejemplo actual puede determinar una unidad de transformada óptima realizando repetitivamente transformada, cuantificación y codificación por entropía en diferentes unidades de transformada. La unidad de transformada óptima puede determinarse repitiendo mecánicamente un procedimiento de selección de una pluralidad de unidades de predicción usando diversos procedimientos, en lugar de seleccionar las unidades de predicción basándose en un criterio predeterminado, tal como una profundidad o un mismo tipo de modo de predicción. La unidad de transformada óptima puede determinarse basándose en el cálculo de costes de RD, y esto se describirá en detalle con referencia a la Figura 14.

La Figura 14 es un diagrama de diferentes unidades 1430 a 1460 de transformada de ejemplo.

25 Haciendo referencia a la Figura 14, el aparato 900 de codificación de imagen codifica de manera repetitiva diferentes unidades 1430 a 1460 de transformada.

Como se muestra en la Figura 14, una unidad 1410 de codificación puede predecirse y codificarse basándose en una unidad 1420 de predicción que tiene un tamaño menor que la unidad 1410 de codificación. La DCT o KLT se realiza en valores residuales generados como resultado de la predicción, y en este punto, la DCT o KLT puede realizarse basándose en las diferentes unidades 1430 a 1460 de transformada como se muestra en la Figura 14.

30 La unidad 1430 de transformada tiene el mismo tamaño que la unidad 1410 de codificación, y se genera agrupando todas las unidades de predicción incluidas en la unidad 1410 de codificación.

Las unidades 1440 de transformada tienen un tamaño mediante el cual la unidad 1410 de codificación se divide igualmente en dos en una dirección de anchura, y se generan agrupando las unidades de predicción que son adyacentes en una dirección vertical.

35 Las unidades 1450 de transformada tienen un tamaño mediante el cual la unidad 1410 de codificación se divide igualmente en dos en una dirección de altura, y se generan agrupando las unidades de predicción que son adyacentes en una dirección horizontal.

Las unidades 1460 de transformada tienen los mismos tamaños que las unidades 1420 de predicción.

40 El aparato 900 de codificación de imagen puede determinar la unidad de transformada óptima realizando de manera repetitiva transformada, cuantificación y codificación por entropía en las unidades 1430 a 1460 de transformada.

45 Como alternativa, el aparato 900 de codificación de imagen puede codificar información de bandera que indica si la unidad de transformada se genera agrupando una pluralidad de unidades de predicción incluidas en una o más unidades de codificación. Por ejemplo, cuando una unidad de transformada se establece agrupando una pluralidad de unidades de predicción incluidas en una unidad de codificación como se muestra en las Figuras 12A a 12C, la información de bandera se establece a '0', y cuando una unidad de transformada se establece agrupando una pluralidad de unidades de predicción incluidas en una pluralidad de unidades de codificación como se muestra en las Figuras 13A a 13D, la información de bandera se establece a '1'.

50 La Figura 14 ilustra un ejemplo de determinación de la unidad de transformada óptima cuando una unidad de transformada se establece agrupando unidades de predicción incluidas en una unidad de codificación. Sin embargo, la unidad de transformada óptima puede determinarse realizando de manera repetitiva la DCT, cuantificación y codificación por entropía en diferentes unidades de transformada, como se muestra en la Figura 14, incluso cuando una unidad de transformada se establece agrupando unidades de predicción incluidas en una pluralidad de unidades de codificación.

La Figura 15 es un diagrama de bloques de un aparato 1500 para decodificar una imagen, de acuerdo con otra realización ejemplar.

Haciendo referencia a la Figura 15, el aparato 1500 de decodificación de imagen incluye un decodificador 1510 por entropía, un cuantificador 1520 inverso, un transformador 1530 inverso y un restaurador 1540.

5 El decodificador 1510 por entropía decodifica por entropía coeficientes de componente de frecuencia de una unidad de transformada predeterminada. Como se ha descrito con referencia a las Figuras 12A a 12C y 13A a 13D, la unidad de transformada puede generarse agrupando una pluralidad de unidades de predicción. Como se ha descrito anteriormente, las unidades de predicción pueden ser adyacentes entre sí, y pueden incluirse en una unidad de codificación o en una pluralidad de diferentes unidades de codificación.

10 Como se ha descrito con referencia al aparato 900 de codificación de imagen, la unidad de transformada puede generarse agrupando una pluralidad de unidades de predicción adyacentes basándose en una profundidad, o agrupando una pluralidad de unidades de predicción adyacentes en las que se realiza predicción de acuerdo con un mismo tipo de modo de predicción, es decir, de acuerdo con un modo de intra predicción o un modo de inter predicción. Como alternativa, como se describe con referencia a la Figura 14, una unidad de transformada óptima puede seleccionarse realizando de manera repetitiva transformada, cuantificación y decodificación por entropía en diferentes unidades de transformada repitiendo mecánicamente un procedimiento de agrupación de una pluralidad de unidades de predicción.

15 Si una unidad de transformada no incluye coeficientes (por ejemplo coeficientes de coseno discretos), el decodificador 1510 por entropía puede no decodificar por entropía de manera separada coeficientes cuantificados. Si la unidad de transformada no incluye los coeficientes cuantificados, los coeficientes cuantificados no se codifican por entropía de manera separada haciendo referencia una información de bandera predeterminada.

20 El cuantificador 1520 inverso cuantifica a la inversa los coeficientes de componente de frecuencia que se decodifican por entropía por el decodificador 1510 por entropía. Los coeficientes de componente de frecuencia que se decodifican por entropía de acuerdo con una etapa de cuantificación usada mientras se codifica la unidad de transformada se cuantifican a la inversa.

25 El transformador 1530 inverso transforma a la inversa los coeficientes de componente de frecuencia cuantificados a la inversa en un dominio de píxel. La DCT inversa o KLT inversa se realiza en los coeficientes de coseno discretos cuantificados a la inversa para restaurar una unidad de transformada en un dominio de píxel. Como resultado de la transformada inversa, se restauran valores residuales de la unidad de transformada.

30 La unidad de transformada restaurada incluye una pluralidad de unidades de predicción, y como se ha descrito anteriormente, las unidades de predicción pueden incluirse en una unidad de codificación o en una pluralidad de diferentes unidades de codificación.

35 El restaurador 1540 genera valores de predicción prediciendo una pluralidad de unidades de predicción incluidas en la unidad de transformada restaurada. Se generan valores de predicción de una unidad de codificación si las unidades de predicción agrupadas en una unidad de transformada se incluyen en una unidad de codificación, y se generan los valores de predicción de una pluralidad de unidades de codificación si las unidades de predicción agrupadas en una unidad de transformada se incluyen en una pluralidad de unidades de codificación. Una unidad de codificación o una pluralidad de unidades de codificación se restauran añadiendo los valores de predicción generados y los valores residuales restaurados por el transformador 1530 inverso.

40 Si los valores de predicción se generan para una unidad de codificación o una pluralidad de unidades de codificación puede determinarse basándose en información de bandera que indica si el aparato 900 de codificación de imagen generó una unidad de transformada agrupando una pluralidad de unidades de predicción incluidas en una unidad de codificación o en una pluralidad de unidades de codificación.

45 De acuerdo con un ejemplo, si las unidades de predicción agrupadas en una unidad de transformada incluyen una unidad de predicción que está intra-prevista, puede realizarse intra predicción basándose en valores de predicción de al menos una unidad de predicción adyacente, como se describe con referencia a la Figura 10. Como alternativa, una pluralidad de unidades de predicción agrupadas en una unidad de transformada pueden predecirse todas usando inter predicción.

La Figura 16 es un diagrama de flujo que ilustra un procedimiento de codificación de una imagen.

50 Haciendo referencia a la Figura 16, un aparato para codificar una imagen genera valores residuales realizando predicción en una o más unidades de codificación en la operación 1610.

55 Una pluralidad de unidades de predicción agrupadas en una unidad de transformada pueden incluirse en una unidad de codificación o en una pluralidad de unidades de codificación. Por consiguiente, cuando las unidades de predicción se incluyen en una unidad de codificación, los valores residuales se generan realizando predicción en una unidad de codificación, y cuando las unidades de predicción se incluyen en una pluralidad de unidades de

codificación, los valores residuales se generan realizando predicción en la pluralidad de unidades de codificación.

Un procedimiento de generación de los valores residuales prediciendo las unidades de predicción todas a la vez se han descrito anteriormente con referencia a la Figura 10.

5 En la operación 1620, el aparato establece una unidad de transformada seleccionando una pluralidad de unidades de predicción. Las unidades de predicción pueden incluirse en una unidad de codificación o en una pluralidad de unidades de codificación. Las unidades de predicción adyacentes pueden seleccionarse basándose en la profundidad, o pueden seleccionarse unidades de predicción adyacentes en las que se realiza predicción en un mismo tipo de modo de predicción.

10 En la operación 1630, el aparato transforma las unidades de predicción en un dominio de frecuencia de acuerdo con la unidad de transformada establecida en la operación 1620. Los coeficientes de dominio de frecuencia se generan realizando transformada en el conjunto de unidad de transformada agrupando las unidades de predicción.

En la operación 1640, el aparato cuantifica coeficientes de componente de frecuencia, por ejemplo los coeficientes de coseno discreto generados en la operación 1630, de acuerdo con un procedimiento de cuantificación predeterminado.

15 En la operación 1650, el aparato codifica por entropía los coeficientes de componente de frecuencia cuantificados en la operación 1640. La codificación por entropía se realiza mediante CABAC o CAVLC.

20 Como se describe con referencia a la Figura 14, el procedimiento puede incluir adicionalmente ajustar una unidad de transformada óptima repitiendo las operaciones 1610 a 1640 en diferentes unidades de transformada. La unidad de transformada óptima puede establecerse realizando de manera repetitiva transformada, cuantificación y codificación por entropía en las diferentes unidades de transformada como se muestra en la Figura 14.

La Figura 17 es un diagrama de flujo que ilustra un procedimiento de decodificación de una imagen.

25 Haciendo referencia a la Figura 17, el aparato decodifica por entropía coeficientes de componente de frecuencia de una unidad de transformada predeterminada, en la operación 1710. Los coeficientes de componente de frecuencia pueden ser coeficientes de coseno discreto. La unidad de transformada puede establecerse agrupando una pluralidad de unidades de predicción. Como se ha descrito anteriormente, las unidades de predicción pueden ser adyacentes entre sí, y pueden incluirse en una unidad de codificación o en una pluralidad de diferentes unidades de codificación.

30 En la operación 1720, el aparato cuantifica a la inversa los coeficientes de componente de frecuencia que se decodifican por entropía en la operación 1710. Los coeficientes de coseno discreto se cuantifican a la inversa usando una etapa de cuantificación usada durante la codificación.

35 En la operación 1730, el aparato transforma a la inversa los coeficientes de componente de frecuencia que se cuantifican a la inversa en la operación 1720 en un dominio de píxel para restaurar una unidad de transformada. La unidad de transformada restaurada se establece agrupando una pluralidad de unidades de predicción. Los valores residuales incluidos en la unidad de transformada se restauran. Los valores residuales de una unidad de codificación se restauran si las unidades de predicción se incluyen en una unidad de codificación, y se restauran valores residuales de una pluralidad de unidades de codificación si se incluyen las unidades de predicción en las unidades de codificación.

40 Como se ha descrito anteriormente, la unidad de transformada puede establecerse agrupando unidades de predicción adyacentes basándose en una profundidad, o agrupando unidades de predicción adyacentes en las que se realiza predicción de acuerdo con un mismo tipo de modo de predicción.

45 En la operación 1740, el aparato restaura la una o más unidades de codificación basándose en los valores residuales incluidos en la unidad de transformada restaurada en la operación 1730. Los valores de predicción se generan prediciendo la una o más unidades de codificación, y la una o más unidades de codificación se restauran añadiendo los valores de predicción generados y los valores residuales restaurados en la operación 1730. Un procedimiento de predicción de los valores de predicción incluidos en una o más unidades de codificación se ha descrito anteriormente con referencia a la Figura 10.

50 Si la unidad de transformada se establece agrupando las unidades de predicción incluidas en una unidad de codificación, se restaura una unidad de codificación, y si la unidad de transformada se establece agrupando las unidades de predicción incluidas en una pluralidad de unidades de codificación, se restaura la pluralidad de unidades de codificación.

De acuerdo con las realizaciones ejemplares, una imagen se comprime más eficazmente y se codifica puesto que una unidad de transformada puede establecerse para que tenga un tamaño mayor que una unidad de predicción, y puede realizarse la transformada en la unidad de transformada.

Mientras que la presente invención se ha mostrado y descrito particularmente con referencia a realizaciones ejemplares de la misma, se entenderá por un experto en la materia que diversos cambios en forma y detalles pueden realizarse en la misma sin alejarse del alcance de la invención como se define mediante las siguientes reivindicaciones.

- 5 El aparato de codificación o decodificación de imagen o el codificador o decodificador de imagen ilustrado en las Figuras 1, 2, 4, 5, 9, 11 o 15 puede incluir un bus acoplado a cada unidad del aparato o codificador o decodificador, al menos un procesador que se conecta al bus y es para ejecutar comandos, y memoria conectada al bus para almacenar los comandos, mensajes recibidos y mensajes generados.

REIVINDICACIONES

1. Un aparato (500, 1500) de decodificación de una imagen, comprendiendo el aparato:

un decodificador (520, 1510) por entropía que realiza decodificación por entropía para obtener coeficientes de transformación cuantificados de al menos una unidad (870) de transformación en una unidad (810) de codificación;

5 un cuantificador (530, 1520) inverso y un transformador (540, 1530) inverso que realizan cuantificación inversa y transformación inversa en los coeficientes de transformación cuantificados de la al menos una unidad de transformación para obtener los residuos, recibe información acerca de un tamaño máximo de la unidad de codificación, divide una imagen en una pluralidad de unidades de codificación máxima cuadradas usando la información acerca del tamaño máximo de la unidad de codificación y determina la unidad de codificación de una estructura jerárquica incluida en una unidad de codificación máxima, entre la pluralidad de unidades de codificación máxima, usando información acerca de una forma de división analizada desde una secuencia de bits; y

10 un restaurador (550-580, 1540) que realiza inter predicción para al menos una unidad (860) de predicción en la unidad (810) de codificación para generar un predictor y restaurar la imagen usando los residuos y el predictor,

15 en el que, la unidad de codificación máxima se divide jerárquicamente en una o más unidades de codificación de profundidad 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, de acuerdo con la información acerca de una forma de división,

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

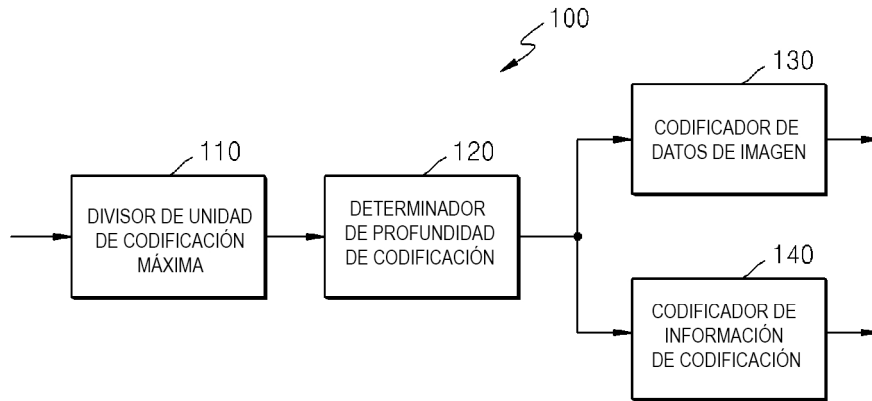
la unidad de codificación de la profundidad mayor que la profundidad actual se predice usando al menos una unidad de predicción y se transforma a la inversa usando al menos una unidad de transformación,

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 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,

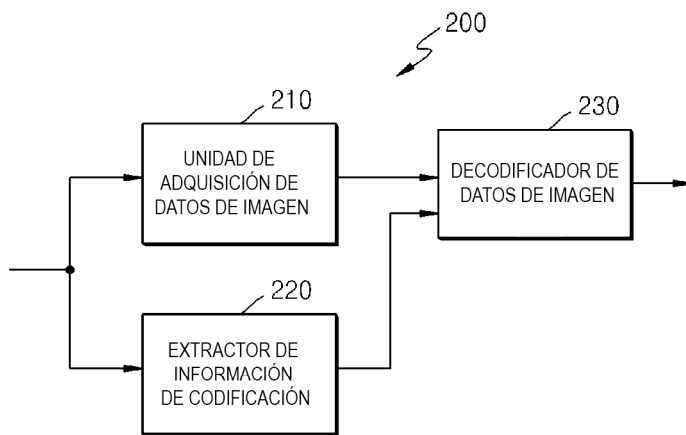
30 en el que la al menos una unidad de transformación es uno entre los bloques que incluyen: un bloque igual en tamaño a la unidad de codificación de la profundidad mayor que la profundidad actual; y un bloque entre una pluralidad de bloques generados dividiendo igualmente la altura y anchura de la unidad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

35 en el que, cuando se determina que un modo de predicción es un modo de inter predicción, no un modo de intra predicción, el aparato (500, 1500) de decodificación de la imagen está configurado para soportar una unidad de transformación, incluida entre la al menos una unidad de transformación que tiene un tamaño de $2N \times 2N$, que incluye cuatro unidades de predicción incluidas entre la al menos una unidad de predicción y que tiene un tamaño de $N \times N$.

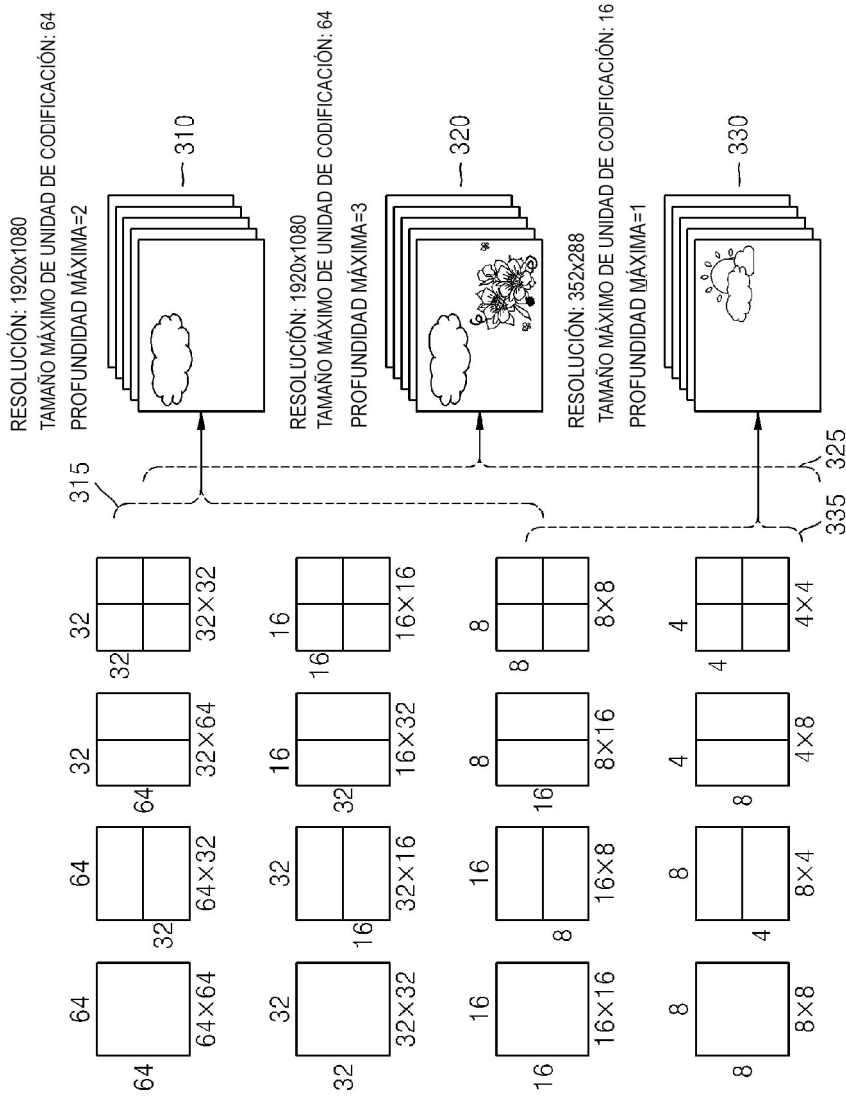
[Figura 1]



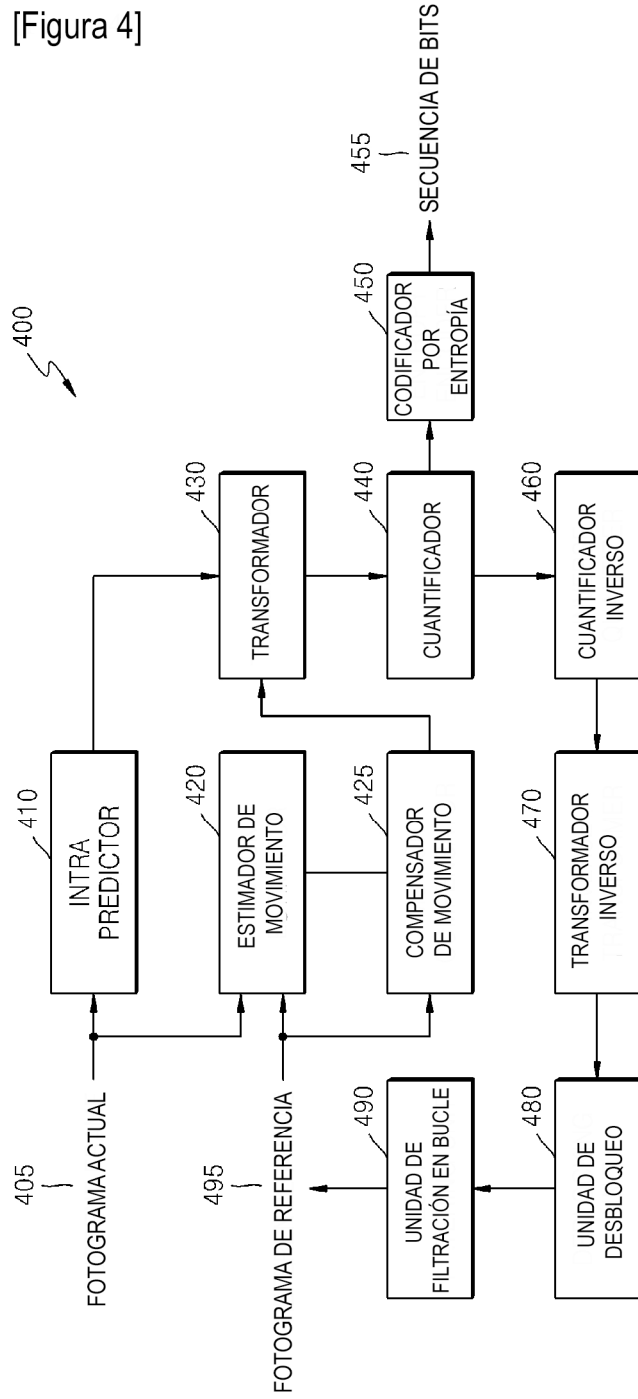
[Figura 2]



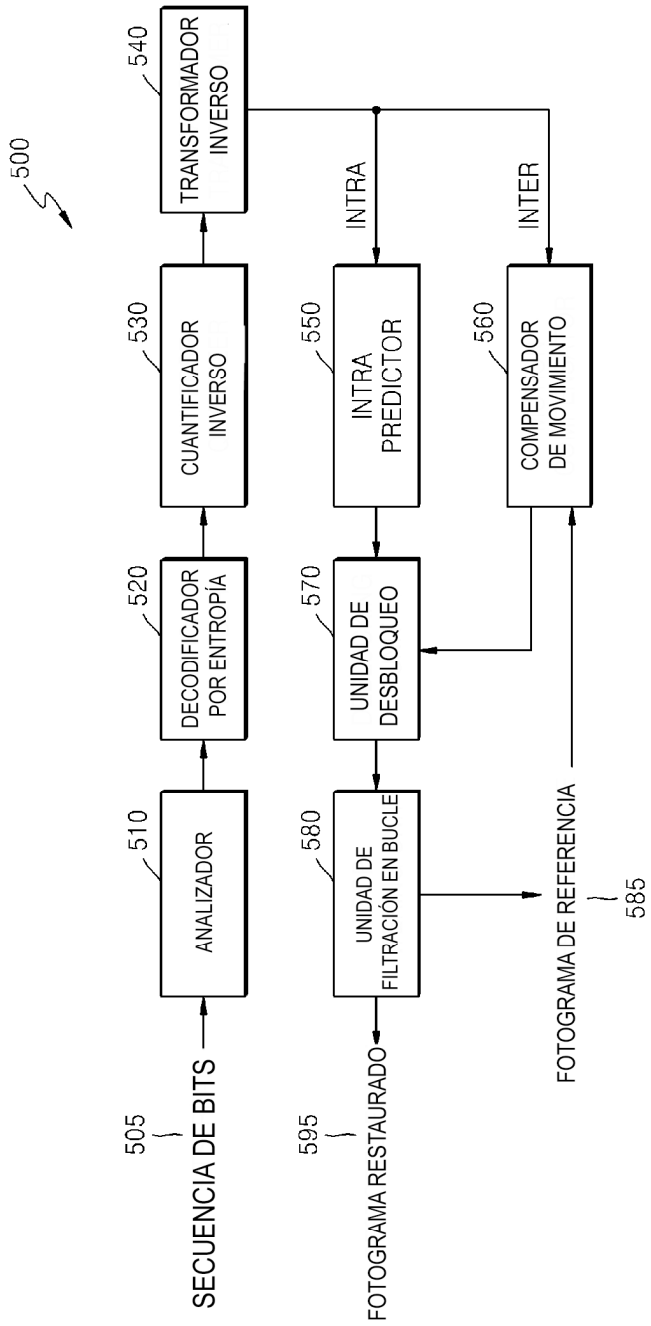
[Figura 3]



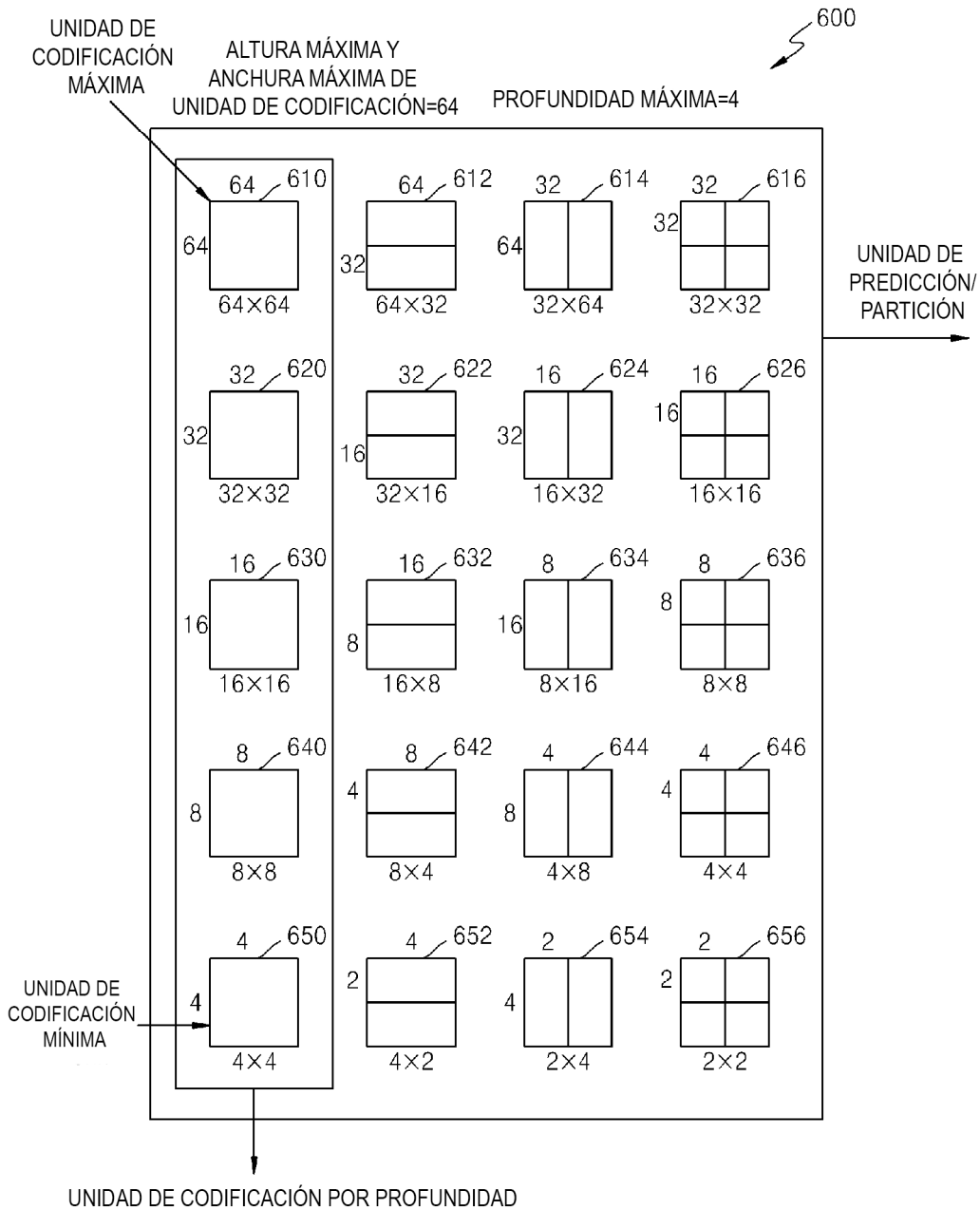
[Figura 4]



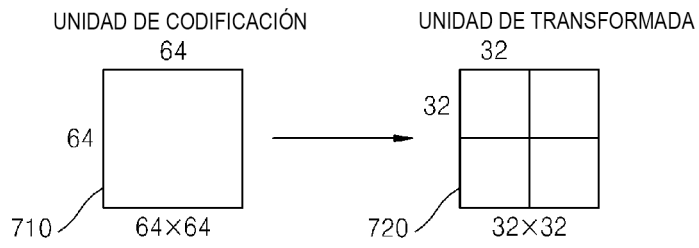
[Figura 5]



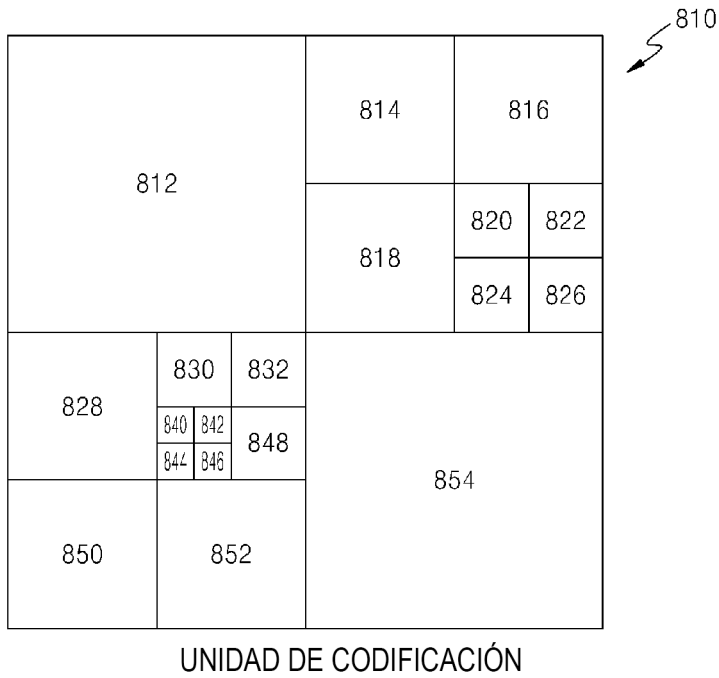
[Figura 6]



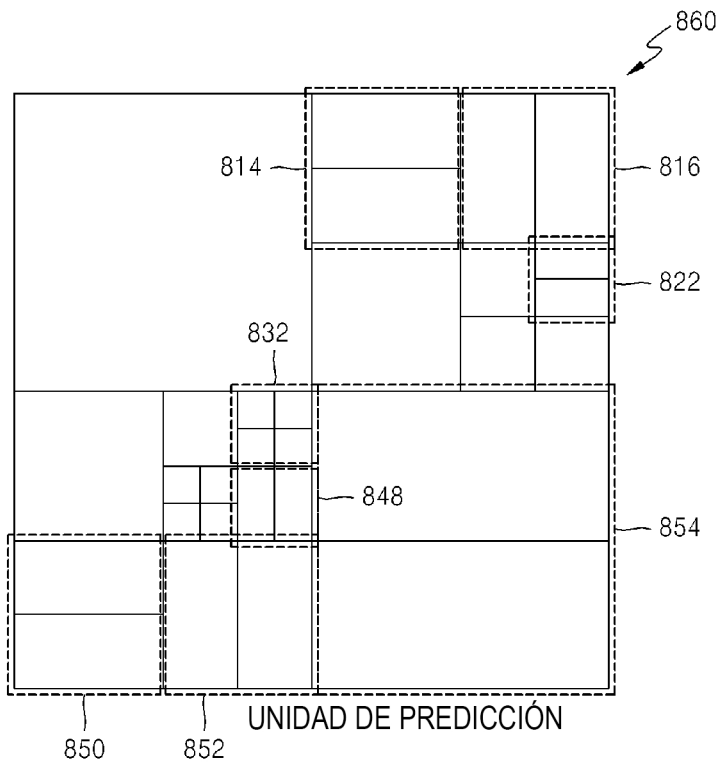
[Figura 7]



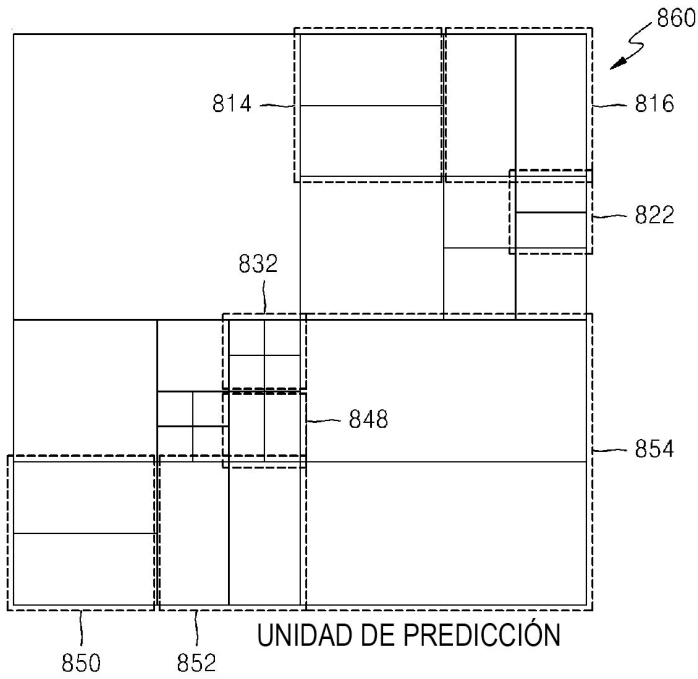
[Figura 8a]



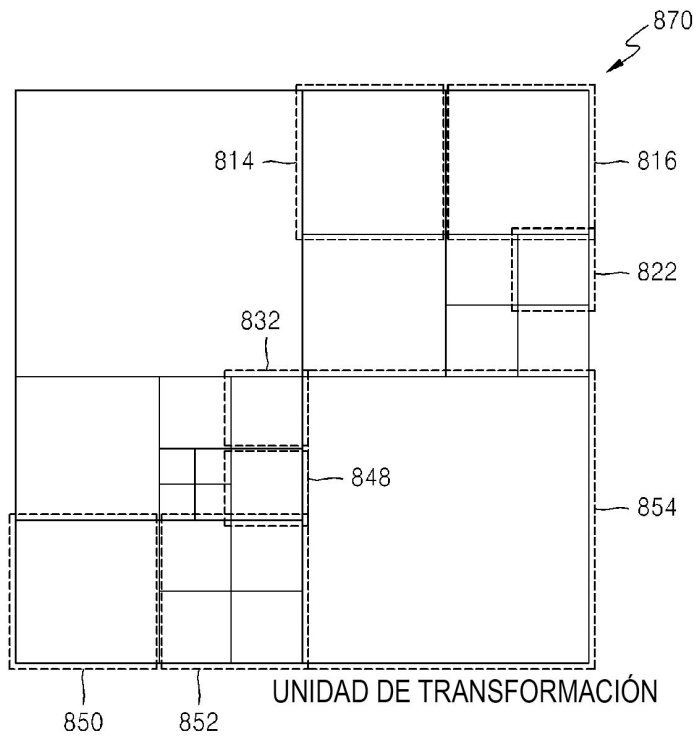
[Figura 8b]



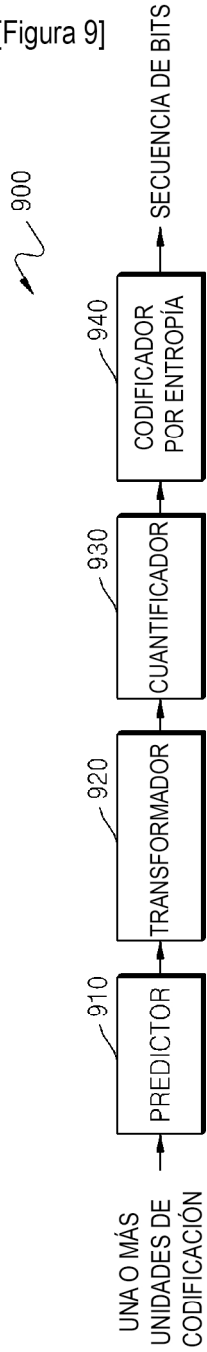
[Figura 8c]



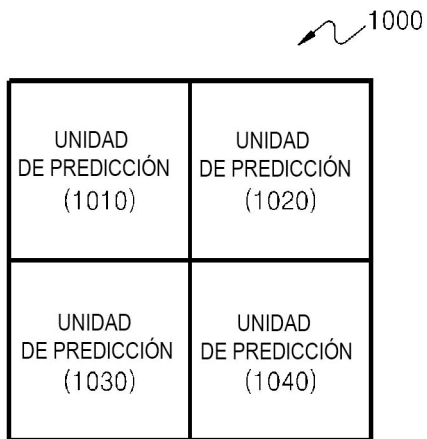
[Figura 8d]



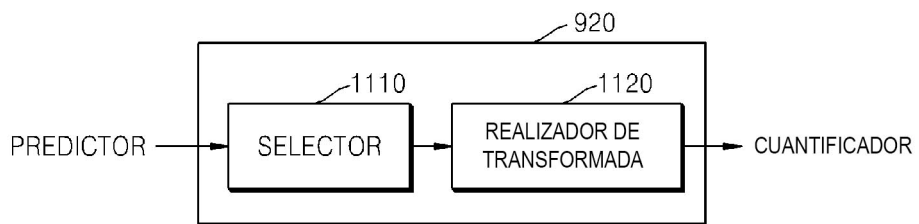
[Figura 9]



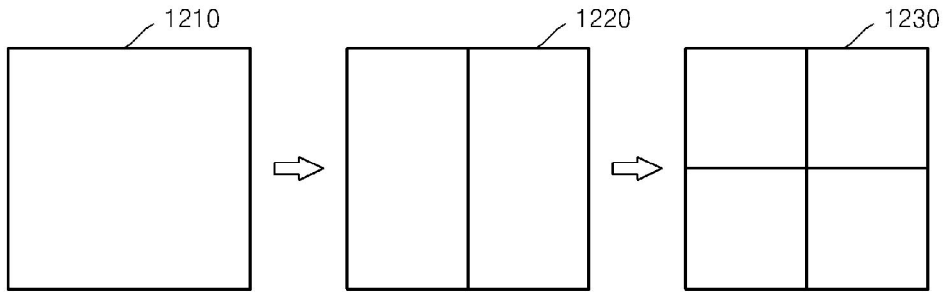
[Figura 10]



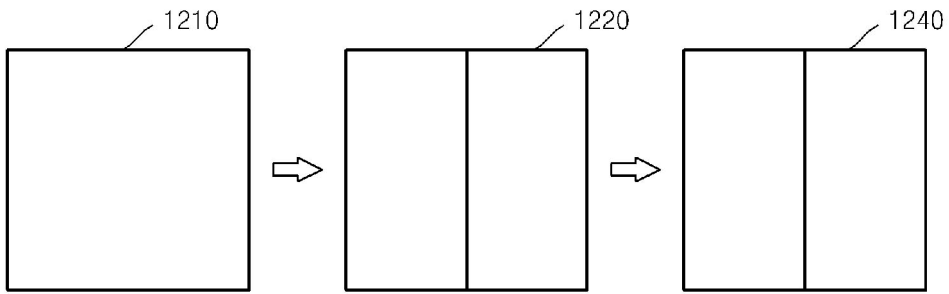
[Figura 11]



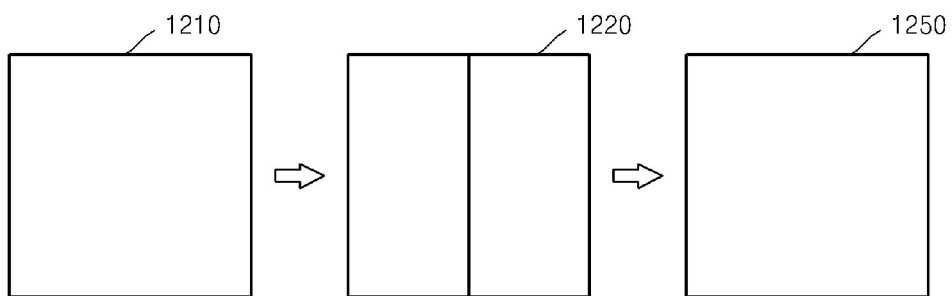
[Figura 12a]



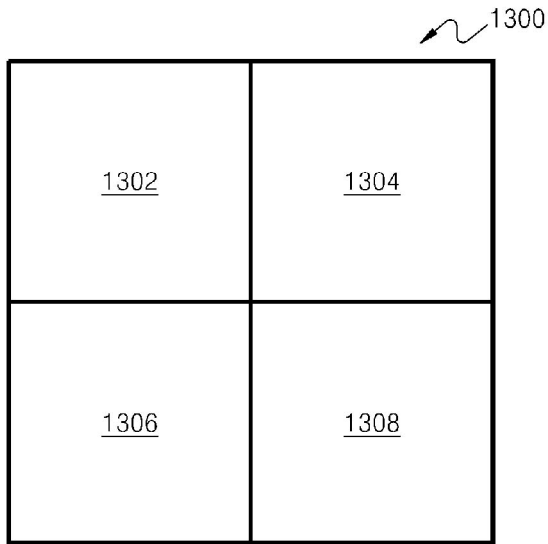
[Figura 12b]



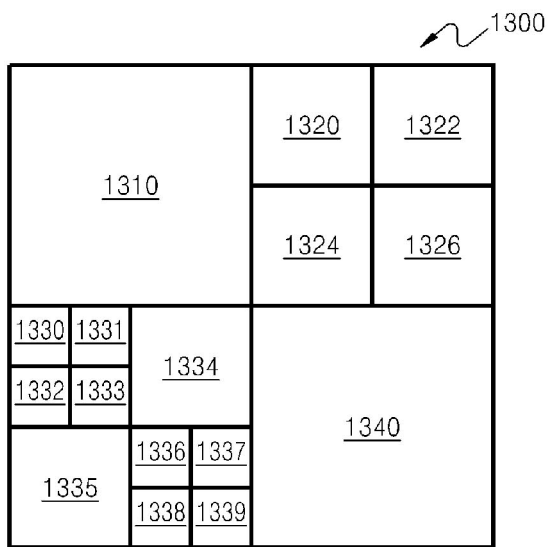
[Figura 12c]



[Figura 13a]



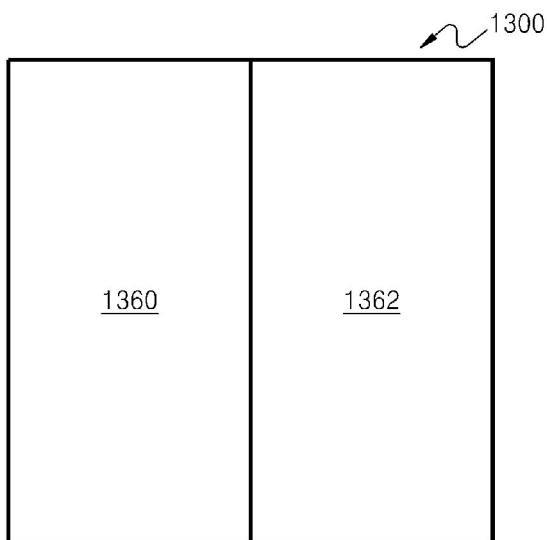
[Figura 13b]



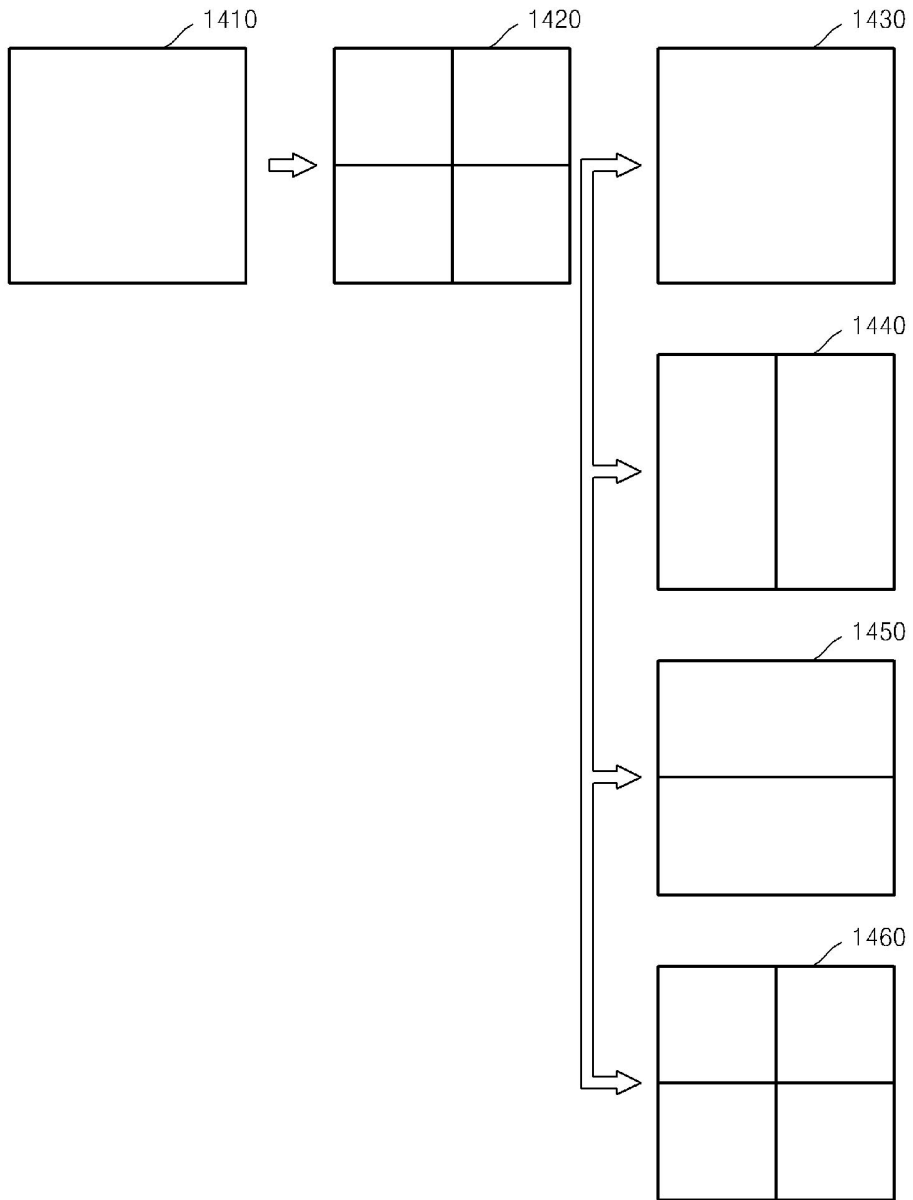
[Figura 13c]



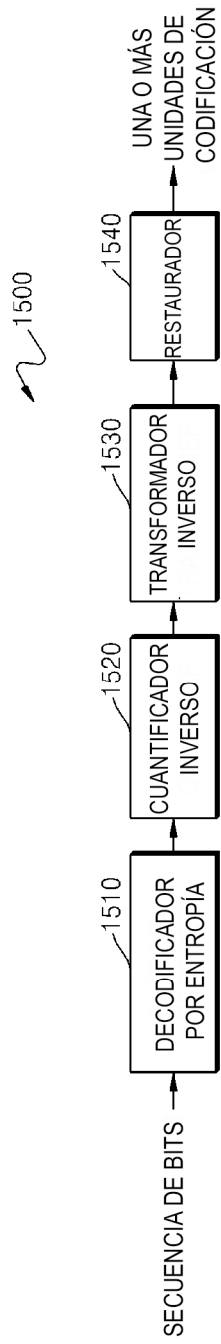
[Figura 13d]



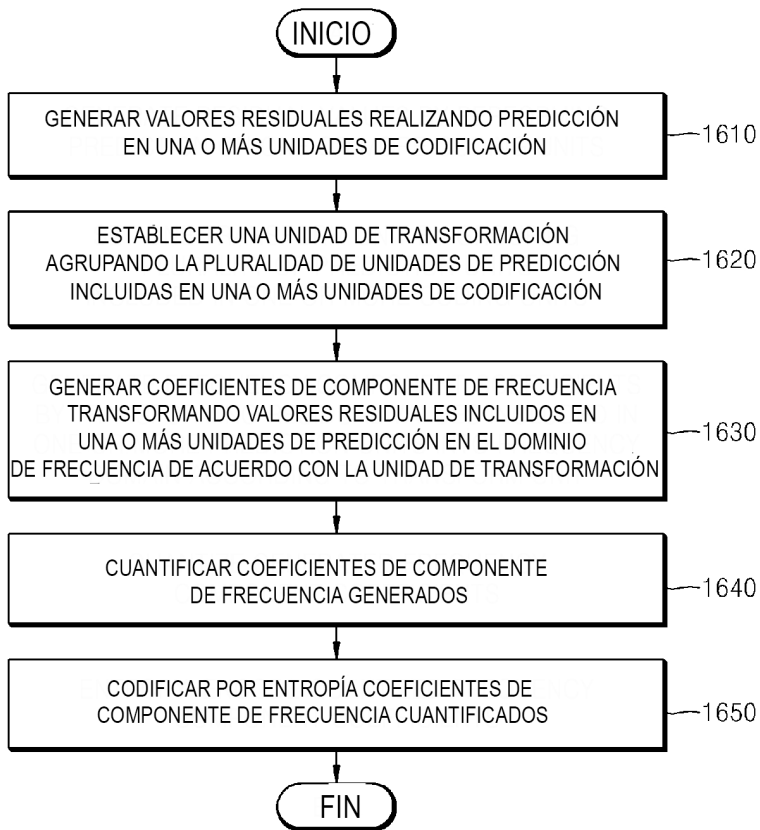
[Figura 14]



[Figura 15]



[Figura 16]



[Figura 17]

