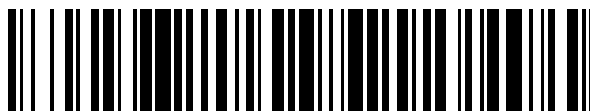


19



OFICINA ESPAÑOLA DE  
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 647 866**

51 Int. Cl.:

**H04N 19/70** (2014.01)

**H04N 19/119** (2014.01)

**H04N 19/196** (2014.01)

**H04N 19/122** (2014.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

86 Fecha de presentación y número de la solicitud internacional: **17.08.2011 PCT/KR2011/006037**

87 Fecha y número de publicación internacional: **23.02.2012 WO12023796**

96 Fecha de presentación y número de la solicitud europea: **17.08.2011 E 11818396 (1)**

97 Fecha y número de publicación de la concesión europea: **15.11.2017 EP 2608539**

54 Título: **Procedimiento de codificación de vídeo y aparato que usa unidad de transformación de estructura de árbol variable y procedimiento y aparato de decodificación de vídeo**

30 Prioridad:

**17.08.2010 US 374348 P**

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

**27.12.2017**

73 Titular/es:

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

72 Inventor/es:

**LEE, TAMMY y  
HAN, WOO-JIN**

74 Agente/Representante:

**CARPINTERO LÓPEZ, Mario**

ES 2 647 866 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 de codificación de vídeo y aparato que usa unidad de transformación de estructura de árbol variable y procedimiento y aparato de decodificación de vídeo

### **Campo técnico**

- 5 La presente invención se refiere a codificación y decodificación de vídeo en el que se realiza transformación entre un dominio espacial y un dominio de la transformada.

### **Antecedentes de la técnica**

- 10 A medida que se está desarrollando y suministrando hardware para reproducir y almacenar contenido de vídeo de alta resolución o de alta calidad, aumenta una necesidad de un códec de vídeo para codificar o decodificar eficazmente el contenido de vídeo de alta resolución o de alta calidad. En un códec de vídeo de la técnica relacionada, un vídeo se codifica de acuerdo con un procedimiento de codificación limitado basándose en un macrobloque que tiene un tamaño predeterminado. También, en el códec de vídeo de la técnica relacionada, los datos de vídeo se codifican y decodifican realizando transformación y transformación inversa usando macrobloques que tienen el mismo tamaño.

### **Descripción detallada de la invención**

#### **Problema técnico**

La presente invención proporciona un procedimiento y aparato para codificar y decodificar vídeo realizando transformación y transformación inversa usando una unidad de datos en una estructura jerárquica variable.

#### **Solución técnica**

- 20 De acuerdo con un aspecto de la presente invención, se proporciona un procedimiento de codificación de vídeo, para codificar una región actual de un vídeo, que incluye operaciones de realización de transformación en la región actual usando unidades de transformación en una estructura de árbol variable que se determinan de entre las unidades de transformación que se dividen jerárquicamente desde una unidad de transformación de base con respecto a la región actual y que se generan basándose en un nivel de división máxima de una unidad de transformación, y determinando unidades de transformación de entre las unidades de transformación en la estructura de árbol variable con respecto a la región actual; y emitiendo datos codificados de la región actual que se generan mediante codificación incluyendo la transformación basándose en las unidades de transformación determinadas, información acerca de un modo de codificación determinado en la codificación con respecto a la región actual, e información de estructura jerárquica de unidad de transformación que incluye información de tamaño máximo e información de tamaño mínimo de la unidad de transformación con respecto al vídeo.

- 30 El documento ANÓNIMO: "Test Model under Consideration", 2. JCT-VC MEETING; 21-7-2010 - 28-7-2010; GINEBRA; (JOINT COLLABORATIVETEAM ON VIDEO CODING OF ISO/IEC JTC1/SC29/WG11 AND ITU-T SG.16); URL:HTTP://WFTP3.ITU.INT/AV-ARCH/JCTVC-SITE/, n.º JCTVC-B205, 28 de julio de 2010 (2010-07-28), ISSN: 0000-0046 desvela un procedimiento de codificación que usa una estructura jerárquica de unidad de transformación.

#### **Efectos ventajosos**

En un procedimiento de codificación y decodificación de un vídeo, se realiza transformación y transformación inversa usando unidades de transformación que tienen diversos tamaños y formas en una estructura de árbol, mediante la cual el vídeo puede codificarse y decodificarse de manera eficaz, teniendo en cuenta una característica de imagen.

### **Descripción de los dibujos**

La Figura 1 es un diagrama de bloques de un aparato de codificación de vídeo que usa unidad de transformación en estructura de árbol variable de acuerdo con una realización de la presente invención.

La Figura 2 es un diagrama de bloques de un aparato de decodificación de vídeo que usa unidad de transformación en estructura de árbol variable, de acuerdo con una realización de la presente invención.

- 45 La Figura 3 ilustra un modelo de jerarquía de una unidad de transformación en una estructura de árbol, de acuerdo con una realización de la presente invención.

La Figura 4 ilustra un modelo de jerarquía de una unidad de transformación en una estructura de árbol variable, de acuerdo con una realización de la presente invención.

- 50 Las Figuras 5 a 8 ilustran unidades de transformación de base, de acuerdo con las realizaciones de la presente invención.

Las Figuras 9 y 10 ilustran unidades de transformación en una estructura de árbol variable, de acuerdo con las realizaciones de la presente invención.

La Figura 11 es un diagrama de flujo de un procedimiento de codificación de vídeo que usa unidades de transformación en una estructura de árbol variable, de acuerdo con una realización de la presente invención.

5 La Figura 12 es un diagrama de flujo de un procedimiento de decodificación de vídeo que usa unidades de transformación en una estructura de árbol variable, de acuerdo con una realización de la presente invención.

La Figura 13 es un diagrama de bloques de un aparato de codificación de vídeo que usa unidad de transformación en estructura de árbol y unidad de transformación en estructura de árbol variable, de acuerdo con una realización de la presente invención.

10 La Figura 14 es un diagrama de bloques de un aparato de decodificación de vídeo que usa unidad de transformación en estructura de árbol y unidad de transformación en estructura de árbol variable, de acuerdo con una realización de la presente invención.

La Figura 15 es un diagrama para describir un concepto de unidades de codificación de acuerdo con una realización de la presente invención.

15 La Figura 16 es un diagrama de bloques de un codificador de imagen basado en unidades de codificación, de acuerdo con una realización de la presente invención.

La Figura 17 es un diagrama de bloques de un decodificador de imagen basado en unidades de codificación, de acuerdo con una realización de la presente invención.

20 La Figura 18 es un diagrama que ilustra unidades de codificación más profundas de acuerdo con las profundidades, y particiones, de acuerdo con una realización de la presente invención.

La Figura 19 es un diagrama para describir una relación entre una unidad de codificación y unidades de transformación, de acuerdo con una realización de la presente invención.

La Figura 20 es un diagrama para describir información de codificación de unidades de codificación que corresponden a una profundidad codificada, de acuerdo con una realización de la presente invención.

25 La Figura 21 es un diagrama de unidades de codificación más profundas de acuerdo con las profundidades, de acuerdo con una realización de la presente invención.

Las Figuras 22 a 24 son diagramas para describir una relación entre unidades de codificación, unidades de predicción, y unidades de transformación, de acuerdo con una realización de la presente invención.

30 La Figura 25 es un diagrama para describir una relación entre una unidad de codificación, una unidad de predicción o una partición, y una unidad de transformación, de acuerdo con información de modo de codificación de la Tabla 1 de acuerdo con una realización de la presente invención.

La Figura 26 es un diagrama de flujo que ilustra un procedimiento de codificación de vídeo que usa una unidad de codificación en una estructura de árbol y una unidad de transformación en una estructura de árbol variable, de acuerdo con una realización de la presente invención.

35 La Figura 27 es un diagrama de flujo que ilustra un procedimiento de decodificación de vídeo que usa una unidad de codificación en una estructura de árbol y una unidad de transformación en una estructura de árbol variable, de acuerdo con una realización de la presente invención.

### **Mejor modo**

40 De acuerdo con un aspecto de la presente invención, se proporciona un procedimiento de codificación de vídeo, para codificar una región actual de un vídeo, que incluye operaciones de realización de transformación en la región actual usando unidades de transformación en una estructura de árbol variable que se determinan de entre las unidades de transformación que se dividen jerárquicamente desde una unidad de transformación de base con respecto a la región actual y que se generan basándose en un nivel de división máxima de una unidad de transformación, y determinando unidades de transformación de entre las unidades de transformación en la estructura de árbol variable con respecto a la región actual; y emitir datos codificados de la región actual que se generan mediante codificación incluyendo la transformación basándose en las unidades de transformación determinadas, información acerca de un modo de codificación determinado en la codificación con respecto a la región actual, e información de estructura jerárquica de unidad de transformación que incluye información de tamaño máximo e información de tamaño mínimo de la unidad de transformación con respecto al vídeo.

50 De acuerdo con otro aspecto de la presente invención, se proporciona un procedimiento de decodificación de vídeo que incluye operaciones de recepción de una secuencia de bits que incluye datos obtenidos codificando un vídeo;

- 5 analizar la secuencia de bits y extraer, desde la secuencia de bits, datos codificados de una región actual, información acerca de un modo de codificación determinado en la codificación con respecto a la región actual, e información de estructura jerárquica de unidad de transformación que incluye información de tamaño máximo e información de tamaño mínimo de la unidad de transformación con respecto al vídeo; y realizar transformación inversa en la región actual usando unidades de transformación en una estructura de árbol variable que se determina de entre las unidades de transformación que se dividen jerárquicamente desde una unidad de transformación de base con respecto a la región actual y que se generan basándose en un nivel de división máxima de la unidad de transformación, decodificar los datos codificados de la región actual, y restaurar el vídeo.
- 10 La información de estructura jerárquica de unidad de transformación puede incluir adicionalmente información de división máxima que indica el nivel de división máxima de la unidad de transformación.
- Un tamaño de la unidad de transformación de base con respecto a la región actual puede ser igual a o menor que un tamaño máximo de la unidad de transformación con respecto al vídeo.
- 15 Una unidad de transformación puede dividirse una vez de modo que pueden generarse las unidades de transformación en un siguiente nivel inferior, un nivel de una unidad de transformación predeterminada puede indicar un número total de divisiones para generar la unidad de transformación predeterminada dividiendo gradualmente la unidad de transformación de base, y la unidad de transformación de base con respecto a la región actual puede ser una unidad de transformación máxima y de nivel más superior que es usable con respecto a la región actual.
- 20 Las unidades de transformación en la estructura de árbol variable que se generan basándose en el nivel de división máxima de la unidad de transformación con respecto a la región actual pueden incluir la unidad de transformación de base y unidades de transformación de acuerdo con niveles que se generan dividiendo gradualmente la unidad de transformación de base al nivel de división máxima.
- 25 Un tamaño mínimo de las unidades de transformación con respecto a la región actual puede determinarse como uno más grande de entre un tamaño mínimo de la unidad de transformación con respecto al vídeo y un tamaño de una unidad de transformación en un nivel más inferior que se obtiene dividiendo la unidad de transformación de base al nivel de división máxima.
- El nivel de división máxima de la unidad de transformación puede ser igual a o menor que un número total de niveles desde una unidad de codificación máxima a una unidad de codificación mínima que corresponde respectivamente a un tamaño máximo y un tamaño mínimo de las unidades de transformación con respecto al vídeo.
- 30 El tamaño de la unidad de transformación de base puede determinarse basándose en al menos uno de un modo de predicción y un tamaño de partición que se usan en la codificación de la región actual.
- 35 El nivel de división máxima de la unidad de transformación con respecto a la región actual puede establecerse para cada uno de los grupos de datos en un nivel de datos de entre los niveles de datos que incluyen una secuencia de instantánea, una instantánea, un corte y una unidad de datos para la codificación del vídeo, y la información de división máxima de la unidad de transformación puede codificarse como un parámetro con respecto a un nivel de datos mediante el cual se determina el nivel de división máxima.
- El nivel de división máxima de la unidad de transformación con respecto a la región actual puede determinarse de manera separada en cada uno de los modos de predicción usados mientras se realiza la codificación en la región actual.
- 40 El nivel de división máxima de la unidad de transformación con respecto a la región actual puede determinarse de manera separada en cada uno de los tipos de corte usados mientras se realiza la codificación en la región actual.
- El tamaño de la unidad de transformación de base con respecto a la región actual puede ser constante.
- La unidad de transformación de base con respecto a la región actual puede determinarse como unidades de datos que tienen formas que se incluyen en particiones para no abarcar un límite entre las particiones que son unidades de datos para codificación de predicción con respecto a la región actual.
- 45 Las unidades de transformación en niveles inferiores que se obtienen dividiendo la unidad de transformación de base con respecto a la región actual puede determinarse como unidades de datos que tienen formas que se incluyen en las particiones para no abarcar el límite entre las particiones que son las unidades de datos para codificación de predicción con respecto a la región actual.
- 50 De acuerdo con otro aspecto de la presente invención, se proporciona un aparato de codificación de vídeo que incluye una unidad de determinación de unidad de transformación, para codificar una región actual del vídeo, realizar la transformación en la región actual usando unidades de transformación en una estructura de árbol variable que se determinan de entre las unidades de transformación que se dividen jerárquicamente desde una unidad de transformación de base con respecto a la región actual y que se generan basándose en un nivel de división máxima de una unidad de transformación, y determinar unidades de transformación de entre las unidades de transformación

5 en la estructura de árbol variable con respecto a la región actual; y una unidad de salida que emite datos codificados de la región actual que se generan mediante codificación incluyendo la transformación basándose en las unidades de transformación determinadas, información acerca de un modo de codificación determinado en la codificación con respecto a la región actual, e información de estructura jerárquica de unidad de transformación que incluye información de tamaño máximo e información de tamaño mínimo de la unidad de transformación con respecto al vídeo.

10 De acuerdo con otro aspecto de la presente invención, se proporciona un aparato de decodificación de vídeo que incluye una unidad de recepción que recibe una secuencia de bits que incluye datos obtenidos codificando el vídeo; una unidad de extracción que analiza la secuencia de bits y extrae, desde la secuencia de bits, datos codificados de una región actual, información acerca de un modo de codificación determinado en la codificación con respecto a la región actual, e información de estructura jerárquica de unidad de transformación que incluye información de tamaño máximo e información de tamaño mínimo de la unidad de transformación con respecto al vídeo; y una unidad de decodificación que realiza transformación inversa en la región actual usando unidades de transformación en una estructura de árbol variable que se determinan de entre las unidades de transformación que se dividen jerárquicamente desde una unidad de transformación de base con respecto a la región actual y que se generan basándose en un nivel de división máxima de la unidad de transformación, decodifica los datos codificados de la región actual y restaura el vídeo.

20 De acuerdo con otro aspecto de la presente invención, se proporciona un medio de grabación legible por ordenador que tiene grabado en el mismo un programa para ejecutar el procedimiento de codificación de vídeo. De acuerdo con otro aspecto de la presente invención, se proporciona un medio de grabación legible por ordenador que tiene grabado en el mismo un programa para ejecutar el procedimiento de decodificación de vídeo.

### **Modo de la invención**

En lo sucesivo, en una o más realizaciones descritas a continuación, una 'imagen' puede indicar de manera colectiva no únicamente una imagen fija sino también una instantánea en movimiento que incluye un vídeo.

25 Cuando se realizan diversas operaciones relacionadas con la imagen, los datos relacionados con la imagen pueden dividirse en grupos de datos, y puede realizarse la misma operación en una pluralidad de piezas de datos incluidas en el mismo grupo de datos. En lo sucesivo, a lo largo de toda la memoria descriptiva, un grupo de datos que se forma de acuerdo con una norma predeterminada se denomina como 'unidad de datos'. En lo sucesivo, a lo largo de toda la memoria descriptiva, una operación que se realiza en cada 'unidad de datos' significa que la operación se realiza usando una pluralidad de piezas de datos incluidas en la unidad de datos.

30 En lo sucesivo, con referencia a las Figuras 1 a 12, se describirá la codificación y decodificación de vídeo que usa una unidad de transformación en una estructura de árbol variable. En lo sucesivo, con referencia a las Figuras 13 a 27, se describirá codificación y decodificación de vídeo que usa una unidad de codificación en una estructura de árbol y una unidad de transformación en una estructura de árbol variable.

35 En lo sucesivo, con referencia a las Figuras 1 a 12, se describirá un aparato de codificación de vídeo y un aparato de decodificación de vídeo, y un procedimiento de codificación de vídeo y un procedimiento de decodificación de vídeo, que usan una unidad de transformación en una estructura de árbol variable.

La Figura 1 es un diagrama de bloques de un aparato de codificación de vídeo que usa la unidad de transformación en estructura 10 de árbol variable de acuerdo con una realización de la presente invención.

40 El aparato de codificación de vídeo que usa unidad de transformación en estructura de árbol variable incluye una unidad 11 de determinación de unidad de transformación y una unidad 13 de salida. En lo sucesivo, por conveniencia de descripción, el aparato de codificación de vídeo que usa unidad de transformación en estructura 10 de árbol variable se denomina como el 'aparato 10 de codificación de vídeo'. Las operaciones por la unidad 11 de determinación de unidad de transformación y la unidad 13 de salida del aparato 10 de codificación de vídeo pueden controlarse de manera cooperativa por un procesador de codificación de vídeo, un procesador central, un procesador de gráficos o similares.

45 Para codificar una instantánea actual de un vídeo de entrada, el aparato 10 de codificación de vídeo divide la instantánea actual en unidades de datos que tienen un tamaño predeterminado y a continuación realiza codificación para cada una de las unidades de datos.

50 Por ejemplo, la instantánea actual consiste en píxeles de un dominio espacial. Para permitir que los píxeles adyacentes en un rango predeterminado formen un grupo para permitir que se codifiquen juntos píxeles que están espacialmente adyacentes entre sí en la instantánea actual, la instantánea actual puede dividirse en grupos de píxeles que tienen un tamaño predeterminado. Mediante una serie de operaciones de codificación que se realizan en píxeles de grupos de píxeles de división que tienen el tamaño predeterminado, puede realizarse codificación con respecto a la instantánea actual.

55

- 5 Puesto que los datos iniciales que son un objetivo de codificación de una instantánea son un valor de píxel del dominio espacial, cada uno de los grupos de píxeles que tienen el tamaño predeterminado puede usarse como una unidad de datos que es un objetivo de codificación. También, se realiza transformación para codificación de vídeo en valores de píxel del grupo de píxeles del dominio espacial, de modo que se generan coeficientes de transformada de un dominio de la transformada, y en este sentido, los coeficientes de transformada mantienen un grupo de coeficientes que tiene el mismo tamaño que el grupo de píxeles del dominio espacial. Por lo tanto, el grupo de coeficientes de los coeficientes de transformada del dominio de la transformada puede usarse también como una unidad de datos para codificación de una instantánea.
- 10 Por lo tanto, a través de todo el dominio espacial y del dominio de la transformada, un grupo de datos que tiene un tamaño predeterminado puede usarse como una unidad de datos para codificar. En este punto, un tamaño de la unidad de datos puede definirse como un número total de piezas de datos incluidas en la unidad de datos. Por ejemplo, un número total de los píxeles del dominio espacial o un número total de los coeficientes de transformada del dominio de la transformada pueden indicar el tamaño de la unidad de datos.
- 15 En lo sucesivo, una 'región actual' puede indicar un grupo de datos que tiene un nivel de datos que es uno de una unidad de datos, un corte, una instantánea y una secuencia de instantáneas que son objetivos de codificación en un vídeo.
- El aparato 10 de codificación de vídeo puede realizar codificación de predicción incluyendo inter predicción e intra predicción, transformada y cuantificación, y codificación por entropía en cada una de las regiones, realizando de esta manera la codificación en la instantánea actual.
- 20 La unidad 11 de determinación de unidad de transformación determina una unidad de transformación que es una unidad de datos para realizar transformación en una región actual de la instantánea actual. La unidad de transformación puede determinarse como la unidad de datos que tiene un tamaño igual a o menor que la región actual para incluirse en la región actual.
- 25 Por ejemplo, la unidad 11 de determinación de unidad de transformación puede generar la unidad de transformación dividiendo por la mitad una altura y una anchura de la región actual. También, la unidad 11 de determinación de unidad de transformación puede generar la unidad de transformación dividiendo asimétricamente al menos una de la altura y la anchura de la región actual. La unidad 11 de determinación de unidad de transformación puede generar la unidad de transformación dividiendo al menos una de la altura y la anchura de la región actual de acuerdo con una relación arbitraria. La unidad de transformación puede ser una unidad de datos de forma poligonal incluida en la
- 30 región actual.
- La unidad 11 de determinación de unidad de transformación puede generar unidades de transformación en un nivel inferior dividiendo de nuevo al menos una de una altura y una anchura de la unidad de transformación.
- 35 La unidad 11 de determinación de unidad de transformación puede determinar unidades de transformación en una estructura de árbol para transformación de la región actual. Las unidades de transformación en una estructura de árbol incluyen unidades de transformación final que se determinan para emitir un resultado de transformación y que se determinan de entre las unidades de transformación en la región actual.
- Para determinar las unidades de transformación en una estructura de árbol, la unidad 11 de determinación de unidad de transformación puede generar unidades de transformación en un nivel inferior dividiendo de manera repetitiva al menos una de una altura y una anchura de una unidad de transformación predeterminada de entre las unidades de transformación en la región actual.
- 40 También, la unidad 11 de determinación de unidad de transformación puede determinar si dividir cada una de las unidades de transformación en unidades de transformación en un nivel inferior, y la determinación puede realizarse independientemente en cada una de las unidades de transformación que tienen el mismo nivel y que se generan dividiéndose desde una unidad de transformación de nivel superior.
- 45 En una realización, una operación en la que una unidad de transformación que tiene un nivel predeterminado se divide en unidades de transformación en un nivel inferior puede incluir una operación en la que se divide al menos una de las unidades de transformación que tiene el nivel predeterminado.
- Por ejemplo, desde una unidad de transformación más superior de la región actual a un nivel predeterminado, todas las unidades de transformación pueden dividirse o pueden no dividirse en cada nivel de división.
- 50 En otra realización, cuando una unidad de transformación se divide desde un nivel predeterminado a un nivel inferior, puede determinarse si dividir unidades de transformación en un cierto nivel independientemente en cada nivel predeterminado, de modo que los tamaños de unidades de transformación en un nivel inferior pueden no ser constantes.
- 55 La unidad 11 de determinación de unidad de transformación genera datos residuales realizando la inter predicción o la intra predicción en la región actual, y realiza transformación con respecto a los datos residuales basándose en una

unidad de transformación determinada por la unidad 11 de determinación de unidad de transformación, de modo que la unidad 11 de determinación de unidad de transformación codifica la región actual. Es decir, los datos residuales para cada partición para predicción de la región actual se vuelven a dividir usando la unidad de transformación determinada por la unidad 11 de determinación de unidad de transformación, de modo que la transformación puede realizarse en los datos residuales para cada una de las unidades de transformación.

En la presente realización, 'transformación' para codificar vídeo indica una técnica de procesamiento de datos usada para transformar datos de vídeo de un dominio espacial en datos de un dominio de la transformada. La transformación para codificar vídeo, que se realiza por la unidad 11 de determinación de unidad de transformación, puede incluir transformación de frecuencia, transformación ortogonal, transformación de números enteros y similares.

La unidad 11 de determinación de unidad de transformación puede realizar de manera repetitiva la transformación en unidades de transformación que se incluyen en la región actual y que están a todos los niveles de acuerdo con una estructura jerárquica, usando unidades de transformación de acuerdo con los niveles, puede comparar un error de la transformación por las unidades de transformación de acuerdo con los niveles, y por lo tanto puede seleccionar las unidades de transformación en un nivel que provoca un error mínimo. La unidad de transformación en el nivel que genera un coeficiente de transformada que provoca un error mínimo, puede determinarse como una unidad de transformación con una profundidad de transformación que está en un nivel en el que se emite un resultado de transformación.

Por consiguiente, la unidad 11 de determinación de unidad de transformación puede determinar las unidades de transformación en una estructura de árbol que incluyen las unidades de transformación que se determinan para emitir el resultado de transformación.

De acuerdo con la presente realización, un tamaño máximo y un tamaño mínimo de una unidad de transformación con respecto al vídeo pueden establecerse con antelación. La unidad 11 de determinación de unidad de transformación puede determinar una unidad de transformación de base que es igual a o menor que el tamaño máximo de la unidad de transformación con respecto al vídeo, para cada una de las regiones en la instantánea actual. La unidad de transformación de base es una unidad de transformación máxima y de nivel más superior que es usable en la región actual.

La unidad 11 de determinación de unidad de transformación puede limitar un número total de niveles de una unidad de transformación en la región actual. La unidad de transformación se divide una vez de modo que se generan unidades de transformación en un nivel inferior, y un nivel de una unidad de transformación predeterminada puede indicar un número total de divisiones realizadas hasta que se genera la unidad de transformación predeterminada después de que se divide la unidad de transformación de base de acuerdo con los niveles. Por lo tanto, un nivel de división máxima de la unidad de transformación en la región actual puede relacionarse con un número total máximo de divisiones de la unidad de transformación de base de la región actual a una unidad de transformación en un nivel más inferior.

De acuerdo con la unidad de transformación de base y el nivel de división máxima de la unidad de transformación que puede establecerse de manera variable con respecto a la región actual, puede cambiarse un número total de niveles y una estructura jerárquica de la unidad de transformación. Por consiguiente, la unidad 11 de determinación de unidad de transformación puede usar unidades de transformación en una estructura de árbol variable, basándose en el nivel de división máxima de la unidad de transformación. Las unidades de transformación en la estructura de árbol variable pueden incluir la unidad de transformación de base, y unidades de transformación de acuerdo con niveles que se generan dividiéndose gradualmente desde la unidad de transformación de base al nivel de división máxima de la unidad de transformación.

La unidad 11 de determinación de unidad de transformación puede realizar transformación en la región actual basándose en unidades de transformación en una estructura de árbol variable que se generan basándose en el nivel de división máxima de la unidad de transformación y que se determinan de entre las unidades de transformación que se dividen gradualmente desde la unidad de transformación de base con respecto a la región actual. La unidad 11 de determinación de unidad de transformación puede determinar finalmente unidades de transformación que se encuentran entre las unidades de transformación en la estructura de árbol variable y que se han de usar al emitir un coeficiente de transformada de la región actual.

Un tamaño mínimo de las unidades de transformación con respecto a la región actual puede determinarse como uno más mayor de entre un tamaño mínimo de una unidad de transformación con respecto al vídeo y un tamaño de la unidad de transformación en el nivel más inferior que se obtiene dividiendo la unidad de transformación de base al nivel de división máxima de la unidad de transformación.

Un número total máximo de divisiones de una unidad de transformación puede ser igual a o menor que un número total de divisiones desde una unidad de transformación máxima a una unidad de transformación mínima que corresponde respectivamente a un tamaño máximo y el tamaño mínimo de las unidades de transformación con respecto a la región actual. Por lo tanto, el nivel de división máxima de la unidad de transformación puede ser igual a

o menor que el número total de divisiones desde la unidad de transformación máxima a la unidad de transformación mínima que corresponde respectivamente al tamaño máximo y al tamaño mínimo de las unidades de transformación con respecto a la región actual.

5 Un tamaño de las unidades de transformación de base con respecto a la región actual puede ser constante. También, las unidades de transformación de base que tienen diferentes tamaños pueden establecerse de acuerdo con características de regiones. Por ejemplo, un tamaño de la unidad de transformación de base puede determinarse basándose en al menos uno de un modo de predicción y un tamaño de partición que se usan en la codificación de la región actual.

10 El nivel de división máxima de la unidad de transformación puede establecerse para cada grupo de datos en un nivel predeterminado de entre los niveles de datos que incluyen una secuencia de instantánea, una instantánea, un corte y una unidad de datos de codificación que son del vídeo. Es decir, por ejemplo, puede establecerse el nivel de división máxima con respecto a una secuencia de instantánea actual o puede establecerse el nivel de división máxima para cada instantánea, cada corte o cada unidad de datos. En otro ejemplo, el nivel de división máxima de la unidad de transformación puede establecerse implícitamente con antelación entre sistemas de codificación y  
15 decodificación.

El nivel de división máxima de la unidad de transformación con respecto a la región actual puede determinarse de manera separada para cada tipo de corte que se usa en la codificación de la región actual.

20 La unidad de transformación de base con respecto a la región actual puede determinarse como una unidad de datos que tiene una forma incluida en particiones que son una unidad de datos para la codificación por predicción de la región actual, mediante la cual la unidad de datos puede no abarcar un límite entre las particiones.

También, las unidades de transformación en un nivel inferior que se obtienen dividiendo la unidad de transformación de base con respecto a la región actual pueden determinarse como unidades de datos que tienen una forma incluida en las particiones que son la unidad de datos para la codificación por predicción de la región actual, mediante la cual las unidades de datos pueden no abarcar el límite entre las particiones.

25 Ejemplos relacionados con la decisión de la unidad de transformación de base y las unidades de transformación en el nivel inferior se describen más adelante con referencia a las Figuras 5 a 10.

La unidad 13 de salida puede emitir una secuencia de bits que incluye datos codificados de la región actual, información acerca del modo de codificación y diversos tipos de información acerca de una unidad de transformación.

30 La unidad 13 de salida puede emitir los datos codificados de la región actual que se generan después de diversas operaciones de codificación que incluyen la transformación basándose en la unidad de transformación determinada por la unidad 11 de determinación de unidad de transformación, y cuantificación, inter predicción, intra predicción, codificación por entropía o similares.

35 La unidad 13 de salida puede emitir diversos tipos de información acerca de los modos de codificación relacionados con esquemas de codificación usados en las diversas operaciones de codificación que incluyen la transformación basándose en la unidad de transformación determinada por la unidad 11 de determinación de unidad de transformación, y cuantificación, inter predicción, intra predicción, codificación por entropía o similares.

40 La unidad 13 de salida puede emitir información de estructura jerárquica de unidad de transformación que indica la estructura jerárquica de unidades de transformación con respecto al vídeo. La información de estructura jerárquica de unidad de transformación puede incluir información acerca del tamaño máximo y el tamaño mínimo de la unidad de transformación con respecto al vídeo, e información de índice de transformación.

45 En una realización, la información de índice de transformación puede indicar información acerca de una estructura de la unidad de transformación que se usa en transformación de la región actual. Por ejemplo, la información de índice de transformación puede incluir un número total de divisiones desde la región actual a una unidad de transformación en un nivel más inferior, un tamaño de la unidad de transformación y una forma de la unidad de transformación.

En otra realización de acuerdo con un caso que una unidad de transformación en un nivel superior se divide en unidades de transformación que tienen el mismo tamaño en cada nivel, la información de índice de transformación puede indicar un número total de divisiones de la región actual a la unidad de transformación en el nivel más inferior.

50 La unidad 13 de salida puede emitir la información acerca del tamaño máximo y el tamaño mínimo de la unidad de transformación con respecto al vídeo. En una realización, la información acerca del tamaño máximo y el tamaño mínimo de la unidad de transformación con respecto al vídeo puede emitirse mientras se incluye en un conjunto de parámetros de secuencia o un conjunto de parámetros de instantánea de un flujo de vídeo.



La información de estructura jerárquica de unidad de transformación puede incluir información de división máxima de unidad de transformación que indica el nivel de división máxima de la unidad de transformación. Por lo tanto, la unidad 13 de salida puede codificar y emitir la información de división máxima de unidad de transformación. En una realización, la información de división máxima de unidad de transformación puede emitirse mientras se incluye en el conjunto de parámetros de secuencia o el conjunto de parámetros de instantánea, o puede establecerse para cada corte o cada región que tiene un tamaño predeterminado.

En otra realización, cuando el nivel de división máxima de la unidad de transformación se establece implícitamente con antelación entre los sistemas de codificación y de decodificación, no se requiere codificar y emitir la información de división máxima de unidad de transformación.

La unidad 13 de salida puede determinar información de subdivisión de unidad de transformación para cada unidad de transformación en la estructura de árbol variable con respecto a la región actual, y a continuación puede emitir la información de subdivisión de unidad de transformación que indica si una cierta unidad de transformación se divide en unidades de transformación en un siguiente nivel inferior. La unidad 13 de salida puede omitir la información de subdivisión de unidad de transformación con respecto a una unidad de transformación que tiene un tamaño mínimo que se encuentra entre unidades de transformación determinadas con respecto a la región actual y que se permite con respecto a la región actual.

La unidad 13 de salida puede determinar y emitir información de patrón de unidades de transformación jerárquica para cada una de las unidades de transformación que se determinan con respecto a la región actual, y la información de patrón de unidades de transformación jerárquica puede indicar la existencia de una unidad de transformación que incluye un coeficiente distinto de cero y que se encuentra entre unidades de transformación en un nivel inferior.

La Figura 2 es un diagrama de bloques de un aparato de decodificación de vídeo que usa la unidad de transformación en estructura 20 de árbol variable, de acuerdo con una realización de la presente invención.

El aparato de decodificación de vídeo que usa la unidad de transformación en estructura 20 de árbol variable incluye una unidad 21 de recepción, una unidad 23 de extracción y una unidad 25 de decodificación. En lo sucesivo, por conveniencia de descripción, el aparato de decodificación de vídeo que usa la unidad de transformación en estructura 20 de árbol variable se denomina como el 'aparato 20 de decodificación de vídeo'. Las operaciones por la unidad 21 de recepción, la unidad 23 de extracción y la unidad 25 de decodificación del aparato 20 de decodificación de vídeo pueden controlarse de manera cooperativa por un procesador de codificación de vídeo, un procesador central, un procesador de gráficos o similares.

Para restaurar una imagen desde una secuencia de bits, el aparato 20 de decodificación de vídeo puede realizar operaciones que incluyen decodificación por entropía, cuantificación inversa, transformación inversa, inter predicción/compensación, e intra predicción/compensación, decodificando de esta manera datos de imagen codificados de la secuencia de bits.

La unidad 21 de recepción recibe y analiza una secuencia de bits con respecto a un vídeo codificado. La unidad 23 de extracción extrae datos codificados con respecto a cada una de las regiones de una instantánea actual, información acerca de un modo de codificación, y diversos tipos de información acerca de una unidad de transformación desde la secuencia de bits que se analiza por la unidad 21 de recepción.

La unidad 23 de extracción puede entregar datos codificados de una región actual, la información acerca del modo de codificación, y los diversos tipos de información acerca de la unidad de transformación a la unidad 25 de decodificación.

La unidad 25 de decodificación puede realizar diversas operaciones de decodificación que incluyen decodificación por entropía, cuantificación inversa, transformación inversa, inter predicción/compensación e intra predicción/compensación en los datos codificados, de acuerdo con diversos esquemas de decodificación que se determinan basándose en la información acerca del modo de codificación, puede restaurar un valor de píxel de la región actual y a continuación puede restaurar la instantánea actual.

La unidad 23 de extracción puede extraer información de estructura jerárquica de unidad de transformación tal como información de tamaño máximo e información de tamaño mínimo de la unidad de transformación, información de tamaño máximo e información de tamaño mínimo de la unidad de transformación, o similares que están relacionadas con una estructura jerárquica de la unidad de transformación desde la secuencia de bits.

La unidad 25 de decodificación puede determinar una unidad de transformación requerida para la transformación inversa con respecto a la región actual, basándose en los diversos tipos de información acerca de la unidad de transformación que se extraen por la unidad 23 de extracción, y puede realizar la transformación inversa con respecto a la región actual basándose en la unidad de transformación. La transformación inversa que se realiza para la decodificación del vídeo, por la unidad 25 de decodificación, puede indicar un procedimiento de transformación de datos de un dominio de la transformada a datos de un dominio espacial. La transformación inversa por la unidad 25 de decodificación puede incluir transformación de frecuencia inversa, transformación ortogonal inversa,

transformación de números enteros inversa y similares.

5 Los conceptos de una unidad de transformación, una unidad de transformación de base y una estructura jerárquica de la unidad de transformación son los mismos que las descripciones anteriormente mencionadas con referencia a la Figura 1 y al aparato 10 de codificación de vídeo. Es decir, en la presente realización, la unidad de transformación se genera dividiendo al menos una de una altura y una anchura de la región actual o una unidad de transformación en un nivel superior de acuerdo con una relación arbitraria.

10 En particular, las unidades de transformación en una estructura de árbol variable pueden determinarse basándose en un nivel de división máxima de una unidad de transformación o un número total máximo de divisiones de la unidad de transformación con respecto a la región actual. Es decir, las unidades de transformación en una estructura de árbol variable pueden incluir una unidad de transformación de base, y unidades de transformación en un nivel inferior que se dividen desde la unidad de transformación de base al nivel de división máxima que se permite en el vídeo actual.

15 La unidad 23 de extracción puede extraer la información de tamaño máximo y la información de tamaño mínimo de la unidad de transformación con respecto al vídeo, desde la información de estructura jerárquica de unidad de transformación. En una realización, la información de tamaño máximo y la información de tamaño mínimo de la unidad de transformación con respecto al vídeo pueden extraerse desde un conjunto de parámetros de secuencia o un conjunto de parámetros de instantánea de un flujo de vídeo.

20 La unidad 23 de extracción puede extraer información de división máxima de unidad de transformación desde la información de estructura jerárquica de unidad de transformación. En una realización, la información de división máxima de unidad de transformación puede extraerse desde el conjunto de parámetros de secuencia o el conjunto de parámetros de instantánea, o puede extraerse desde un corte o parámetros que se establecen para cada una de las regiones.

25 En otra realización, cuando el nivel de división máxima de la unidad de transformación se establece implícitamente con antelación entre sistemas de codificación y decodificación, no se requiere extraer de manera separada la información de división máxima de unidad de transformación.

La unidad 25 de decodificación puede analizar un número total de niveles permisibles o un número total de divisiones permisibles desde la unidad de transformación de base a una unidad de transformación de nivel más inferior de la región actual.

30 La unidad 25 de decodificación puede leer un nivel de división máxima de una unidad de transformación que se establece con respecto a una secuencia de instantánea del vídeo. Como alternativa, el nivel de división máxima de la unidad de transformación puede leerse de acuerdo con diversos niveles de datos tales como instantáneas, cortes o unidades de datos.

35 En otra realización, las unidades de transformación en una estructura de árbol variable pueden determinarse basándose en el nivel de división máxima de la unidad de transformación que se establece implícitamente con antelación entre los sistemas de codificación y de decodificación.

La unidad 23 de extracción puede extraer la información de índice de transformación desde la información de estructura jerárquica de unidad de transformación. La unidad 23 de extracción puede analizar una estructura de la unidad de transformación, que se usa en la transformación de la región actual, desde la información de índice de transformación.

40 Por ejemplo, un número total de divisiones desde la región actual a la unidad de transformación de nivel más inferior, un tamaño de la unidad de transformación y una forma de la unidad de transformación pueden extraerse desde la información de índice de transformación. Cuando una unidad de transformación en un nivel superior se divide en unidades de transformación que tienen el mismo tamaño en cada nivel, de acuerdo con la información de índice de transformación, puede leerse el número total de divisiones desde la región actual a la unidad de transformación de nivel más inferior.

45 La unidad 23 de extracción puede extraer información de subdivisión de unidad de transformación para cada una de las unidades de transformación que se determinan con respecto a la región actual. La unidad 25 de decodificación puede determinar si dividir una unidad de transformación en un nivel actual en unidades de transformación en un siguiente nivel inferior, basándose en la información de subdivisión de unidad de transformación. Cuando la información de subdivisión de unidad de transformación ya no se extrae con respecto a una unidad de transformación en un nivel predeterminado, la unidad de transformación en el nivel predeterminado puede analizarse como una unidad de transformación de tamaño mínimo que se permite con respecto a la región actual. La unidad 23 de extracción puede extraer información de patrón de unidades de transformación jerárquica desde cada una de las unidades de transformación que se determinan con respecto a la región actual. La unidad 25 de decodificación puede analizar la existencia de una unidad de transformación que incluye un coeficiente distinto de cero y que se encuentra entre unidades de transformación en un nivel inferior de una unidad de transformación actual, desde la información de patrón de unidades de transformación jerárquica.

- 5 En una realización, las unidades de transformación requeridas para la transformación inversa pueden analizarse desde unidades de transformación en una estructura jerárquica que se dividen uniformemente en unidades de transformación que tienen un tamaño constante en cada nivel desde una unidad de transformación de nivel más superior de la región actual a un nivel de división predeterminado, basándose en la información de estructura jerárquica de unidad de transformación. En otra realización, se determina de manera separada si dividir una unidad de transformación en un nivel superior en unidades de transformación en un nivel inferior, las unidades de transformación en un nivel de transformación que se determina que se requieren para la transformación inversa basándose en la información de estructura jerárquica de unidad de transformación no están limitadas a unidades de transformación que tienen un tamaño constante.
- 10 Por lo tanto, la unidad 25 de decodificación puede realizar la transformación inversa en la región actual usando las unidades de transformación que se determinan basándose en información relacionada con la unidad de transformación extraída por la unidad 23 de extracción.
- 15 En particular, la unidad 25 de decodificación puede analizar las unidades de transformación en una estructura de árbol variable que se permiten con respecto a la región actual, basándose en el nivel de división máxima de la unidad de transformación con respecto a la región actual que se determina basándose en la información de división máxima de unidad de transformación, y puede detectar unidades de transformación que se han de usar en transformación inversa de un coeficiente de transformada y que se determinan de entre las unidades de transformación en una estructura de árbol variable. La unidad 25 de decodificación puede realizar la transformación inversa en la región actual usando las unidades de transformación que se detectan basándose en el nivel de división
- 20 máxima.
- La unidad 25 de decodificación puede realizar decodificación en cada una de las imágenes configurando una región predeterminada del vídeo, basándose en información acerca de un modo de codificación correspondiente y la información relacionada con la unidad de transformación, y por lo tanto puede decodificar el vídeo.
- 25 En un procedimiento de codificación y decodificación de vídeo, el aparato 10 de codificación de vídeo y el aparato 20 de decodificación de vídeo pueden realizar la transformación y la transformación inversa usando unidades de transformación de estructura de árbol que tienen diversos tamaños y formas, de modo que el aparato 10 de codificación de vídeo y el aparato 20 de decodificación de vídeo pueden codificar y decodificar de manera eficaz un vídeo, teniendo en cuenta una característica de imagen.
- 30 También, puesto que se realiza la codificación y la decodificación usando la transformación basándose en la unidad de transformación en una estructura de árbol variable de entre unidades de transformación en una estructura de árbol, puede omitirse un procedimiento de codificación y decodificación que usa unidades de transformación en un nivel que no se requiere de acuerdo con la característica de la imagen y una característica de codificación, de modo que puede reducirse una cantidad de cálculo.
- 35 También, el número total máximo de divisiones o el número total de niveles de la unidad de transformación, que se permite con respecto a la región actual, puede predecirse basándose en el nivel de división máxima, de modo que es posible reducir una cantidad de transmisión con respecto a información que incluye la información de subdivisión de unidad de transformación que se requiere en la determinación de las unidades de transformación a usarse en la decodificación de vídeo.
- 40 La Figura 3 ilustra un modelo de jerarquía de una unidad de transformación en una estructura de árbol, de acuerdo con una realización de la presente invención.
- La unidad de transformación en una estructura de árbol que se ha de usar en la codificación de un vídeo actual, por el aparato 10 de codificación de vídeo y el aparato 20 de decodificación de vídeo, puede determinarse basándose en un tamaño máximo y un tamaño mínimo de una unidad de transformación que se permiten en el vídeo actual.
- 45 El aparato 10 de codificación de vídeo puede incluir información de tamaño máximo 'MaxTransformSize' e información de tamaño mínimo 'MinTransformSize' de la unidad de transformación que se permite en el vídeo actual en un conjunto de parámetros de secuencia o un conjunto de parámetros de instantánea y que puede emitirlos.
- 50 Por ejemplo, cuando la información de tamaño máximo 'MaxTransformSize' y la información de tamaño mínimo 'MinTransformSize' de la unidad de transformación indican 32x32 y 4x4, respectivamente, las unidades de transformación con respecto a una región de tamaño 64x64 (es decir, una unidad de codificación (CU) 30) pueden incluir las unidades 35a, 35b, 35c y 35d de transformación de tamaño 32x32 a las unidades 38a y 38b de transformación de tamaño 4x4.
- 55 Por conveniencia de descripción, se supone una relación jerárquica de unidades de transformación, en la que una altura y una anchura de una unidad de transformación actual se dividen por la mitad de modo que se generan cuatro unidades de transformación en un siguiente nivel inferior desde la unidad de transformación actual. Puesto que el tamaño máximo de la unidad de transformación que se permite en el vídeo actual es 32x32, un tamaño de una unidad de transformación en un nivel cero que es un nivel más superior puede ser 32x32, un tamaño de una unidad de transformación en un primer nivel puede ser 16x16, un tamaño de una unidad de transformación en un segundo

nivel puede ser 8x8, y un tamaño de una unidad de transformación en un tercer nivel puede ser 4x4.

En más detalle, las unidades 35a, 35b, 35c y 35d de transformación de tamaño 32x32 en el nivel cero pueden dividirse en las unidades 36a, 36b, 36c, 36d, 36e, 36h, 36i, 36l, 36m y 36p de transformación de tamaño 16x16 en el primer nivel. También, las unidades 36a, 36b, 36c, 36d, 36e, 36h, 36i, 36l, 36m y 36p de transformación de tamaño 16x16 en el primer nivel pueden dividirse en las unidades 37a, 37b, 37c, 37d, 37e y 37f de transformación de tamaño 8x8 en el segundo nivel. También, las unidades 37a, 37b, 37c, 37d, 37e y 37f de transformación de tamaño 8x8 en el segundo nivel pueden dividirse en las unidades 38a y 38b de transformación de tamaño 4x4 en el tercer nivel.

Debido a la limitación en espacio, todas las unidades de transformación disponibles con respecto a las unidades 36a, 36b, 36c, 36d, 36e, 36h, 36i, 36l, 36m y 36p de transformación en el primer nivel, las unidades 37a, 37b, 37c, 37d, 37e y 37f de transformación en el segundo nivel, y las unidades 38a y 38b de transformación en el tercer nivel no se ilustran, pero, pueden generarse cuatro unidades de transformación en un siguiente nivel inferior desde la unidad de transformación actual.

En más detalle, un número total de unidades de transformación de acuerdo con niveles que pueden usarse en la transformación de una región actual (es decir, la CU 30) puede ser 4 unidades de transformación que incluyen las unidades 35a, 35b, 35c y 35d de transformación en el nivel cero, 16 unidades de transformación que incluyen las unidades 36a, 36b, 36c, 36d, 36e, 36h, 36i, 36l, 36m y 36p de transformación en el primer nivel, 64 unidades de transformación que incluyen las unidades 37a, 37b, 37c, 37d, 37e y 37f de transformación en el segundo nivel y 256 unidades de transformación que incluyen las unidades 38a y 38b de transformación en el tercer nivel.

Para determinar las unidades de transformación en una estructura de árbol con respecto a la región 30 actual, el aparato 10 de codificación de vídeo puede realizar de manera repetitiva la transformación en la región 30 actual usando las unidades 35a, 35b, 35c y 35d de transformación de tamaño 32x32, las unidades 36a, 36b, 36c, 36d, 36e, 36h, 36i, 36l, 36m y 36p de transformación de tamaño 16x16, las unidades 37a, 37b, 37c, 37d, 37e y 37f de transformación de tamaño 8x8, y las unidades 38a y 38b de transformación de tamaño 4x4 que se permiten en el vídeo actual.

Después de que el aparato 10 de codificación de vídeo realiza la transformación con respecto a todas las unidades de transformación en los niveles cero, primero, segundo y tercero, el aparato 10 de codificación de vídeo puede seleccionar independientemente una unidad de transformación para cada una de las regiones interiores de la región 30 actual, en la que la unidad de transformación está a un nivel desde el cual se emite un coeficiente de transformada con un error mínimo. En una realización, las unidades de transformación en una estructura de árbol pueden incluir las unidades de transformación en el nivel seleccionado.

Para codificar información acerca de las unidades de transformación en una estructura de árbol de la región 30 actual, el aparato 10 de codificación de vídeo puede codificar y emitir información de profundidad de transformación acerca de las profundidades de transformación que indican niveles de unidades de transformación desde las cuales se generan coeficientes de transformación con el mínimo error.

El aparato 20 de decodificación de vídeo puede leer los niveles de las unidades de transformación usadas para emitir los coeficientes de transformada de la región 30 actual, usando la información de profundidad de transformación extraída desde una secuencia de bits, y puede determinar unidades de transformación en una estructura de árbol para transformación inversa de los coeficientes de transformada de la región 30 actual.

La Figura 4 ilustra un modelo de jerarquía de una unidad de transformación en una estructura de árbol variable, de acuerdo con una realización de la presente invención.

El aparato 10 de codificación de vídeo puede emitir información de división máxima 'MaxTuDepth' de una unidad de transformación de una región actual junto con información de tamaño máximo 'MaxTransformSize' e información de tamaño mínimo 'MinTransformSize' de la unidad de transformación que se permite en el vídeo actual.

La información de división máxima de la unidad de transformación de la región actual puede indicar un nivel de transformación máximo, es decir, un número total de niveles permisibles de unidades de transformación con respecto a la región actual. Con respecto a una región 40 actual de la región actual, pueden permitirse las unidades de transformación desde una unidad de transformación de base en un nivel más superior a unidades de transformación en un nivel más inferior que es el nivel de transformación máximo.

Por ejemplo, el nivel de transformación máximo con respecto a la región actual puede establecerse como 2.

La unidad de transformación de base puede incluir las unidades 45a, 45b, 45c y 45d de transformación de tamaño 32x32 en un nivel cero.

Por consiguiente, las unidades de transformación con respecto a la región 40 actual que tienen un tamaño de 64x64 pueden incluir las unidades 45a, 45b, 45c y 45d de transformación de tamaño 32x32 en el nivel cero y las unidades 46a, 46b, 46c, 46d, 46e, 46h, 46i, 46l, 46m y 46p de transformación de tamaño 16x16 en un primer nivel.

Basándose en la información de tamaño máximo 'MaxTransformSize' y la información de tamaño mínimo 'MinTransformSize' de la unidad de transformación que se permite en el vídeo actual, y la información de división máxima 'MaxTuDepth' de la unidad de transformación de la región actual, un tamaño mínimo de una unidad de transformación que pueden usarse en la región actual puede calcularse usando la ecuación 1.

5 [Ecuación 1]

Mínimo tamaño TU de hoja posible  
 $= \max(\text{MinTransformSize}, \text{RootTUSize}/(2^{(\text{MaxTuDepth} - 1))})$

10 Es decir, el tamaño mínimo 'Mínimo tamaño TU de hoja posible' de la unidad de transformación de la región actual puede ser mayor que uno de entre la información de tamaño mínimo 'MinTransformSize' de la unidad de transformación que se permite en el vídeo actual, y un tamaño  $\text{RootTUSize}/(2^{(\text{MaxTuDepth} - 1)})$  de una unidad de transformación en un nivel más inferior que se obtiene dividiendo la unidad de transformación de base por un número total máximo de divisiones.

15 En la ecuación 1, un rango de 'MaxTuDepth -1' que corresponde al número total máximo de divisiones satisface la ecuación 2.

[Ecuación 2]

$\text{MaxTuDepth} - 1 \leq \text{Log2}(\text{MaxTransformSize}) - \text{Log2}(\text{MinTransformSize})$

20 Es decir, el número total máximo de divisiones puede ser igual a o menor que un número total de divisiones desde una unidad de transformación máxima a una unidad de transformación mínima que se determinan basándose en la información de tamaño máximo 'MaxTransformSize' y la información de tamaño mínimo 'MinTransformSize' de la unidad de transformación que se permite en el vídeo actual.

Un número total de unidades de transformación de acuerdo con niveles que pueden usarse para realizar transformación en la región 40 actual puede ser cuatro unidades 45a, 45b, 45c y 45d de transformación en el nivel cero y 16 unidades 46a, 46b, 46c, 46d, 46e, 46h, 46i, 46l, 46m y 46p de transformación en el primer nivel.

25 Por lo tanto, aunque todas las unidades de transformación en los niveles cero, primero, segundo y tercero pueden usarse de acuerdo con la información de tamaño máximo y la información de tamaño mínimo de la unidad de transformación con respecto al vídeo actual, el aparato 10 de codificación de vídeo puede realizar transformación en la región 40 actual usando únicamente las unidades de transformación en el primer y segundo niveles, basándose en el nivel de división máxima o el número total máximo de divisiones que se establece con respecto a la región 40 actual.

30 Es decir, como se ha descrito anteriormente, con respecto a la región 30 actual de la Figura 3, la transformación se realiza usando las 4 unidades de transformación en el nivel cero, las 16 unidades de transformación en el primer nivel, las 64 unidades de transformación en el segundo nivel, y las 256 unidades de transformación en el tercer nivel, basándose en la información de tamaño máximo y la información de tamaño mínimo de la unidad de transformación con respecto al vídeo actual. En contraste, con respecto a la región 40 actual de la Figura 4, la transformación puede realizarse usando únicamente 4 unidades de transformación en el nivel cero y 16 unidades de transformación en el primer nivel, basándose en el número total máximo de divisiones o el nivel de división máxima de la unidad de transformación.

35 También, el aparato 10 de codificación de vídeo puede codificar información de subdivisión de unidad de transformación para cada unidad de transformación que indica si una correspondiente unidad de transformación se divide en unidades de transformación en un siguiente nivel inferior y a continuación puede emitirla. Puesto que una unidad de transformación que tiene un tamaño mínimo ya no se divide en unidades de transformación en un siguiente nivel inferior, la información de subdivisión de unidad de transformación no se requiere que se codifique con respecto a la unidad de transformación en un nivel más inferior.

40 Por lo tanto, con respecto a la región 30 actual de la Figura 3, la información de subdivisión de unidad de transformación puede codificarse con respecto a las 4 unidades de transformación en el nivel cero, las 16 unidades de transformación en el primer nivel, y las 64 unidades de transformación en el segundo nivel. En contraste, con respecto a la región 40 actual de la Figura 4, la información de subdivisión de unidad de transformación se codifica con respecto a únicamente las 4 unidades de transformación en el nivel cero, basándose en el número total máximo de divisiones o el nivel de división máxima de la unidad de transformación.

45 También, como se ha descrito anteriormente, el número total máximo de divisiones de la unidad de transformación puede ser igual a o menor que un número total de divisiones desde una unidad de transformación máxima a una unidad de transformación mínima que corresponde respectivamente a un tamaño máximo y un tamaño mínimo que se permiten para un vídeo, y el nivel de división máxima de la unidad de transformación puede predecirse de acuerdo con un resultado del mismo. De esta manera, basándose en la capacidad de predicción del nivel de división máxima, pueden reducirse los bits de información de división máxima.

Por ejemplo, cuando la información de división máxima se codifica de acuerdo con un esquema de Codificación Máxima Unaria Truncada, el número total de divisiones desde la unidad de transformación máxima a la unidad de transformación mínima que se permiten para el vídeo se establece como un valor máximo del nivel de división máxima, de modo que, cuando se codifica el nivel de división máxima que corresponde al valor máximo, puede reducirse un bit.

Como resultado de la comparación entre las unidades de transformación en una estructura de árbol de la Figura 3 y las unidades de transformación en una estructura de árbol variable de la Figura 4, puesto que se cambia un número total de niveles de unidad de transformación que son usables para la transformación de acuerdo con el nivel de división máxima o el número total máximo de divisiones, se cambia una estructura jerárquica de la unidad de transformación. El aparato 10 de codificación de vídeo puede realizar transformación en cada una de las regiones de la región 40 actual usando las unidades de transformación en una estructura de árbol variable, y de acuerdo con un resultado de la transformación, el aparato 10 de codificación de vídeo puede seleccionar independientemente una unidad de transformación para cada una de las regiones, en las que la unidad de transformación se establece a un nivel desde el cual se emiten los coeficientes de transformada con un error mínimo.

El aparato 10 de codificación de vídeo puede codificar y emitir información de profundidad de transformación e información de división máxima de unidad de transformación que se requieren para que el aparato 10 de codificación de vídeo determine las unidades de transformación en una estructura de árbol variable de la región 40 actual.

El aparato 20 de decodificación de vídeo puede leer una profundidad de transformación de las unidades de transformación y el nivel de división máxima de la unidad de transformación que se usan en la codificación de la región 40 actual, usando la información de profundidad de transformación y la información de división máxima de unidad de transformación extraídas desde la secuencia de bits. Basándose en la profundidad o nivel de transformación leídos, pueden determinarse las unidades de transformación en una estructura de árbol variable que se usan para transformar a la inversa los coeficientes de transformada de la región 40 actual.

El número total máximo de divisiones o el nivel de división máxima de la unidad de transformación puede determinarse de acuerdo con una característica de la región actual. Por lo tanto, de acuerdo con una característica de imagen, si no se requiere realizar la transformación usando diversos tipos de una unidad de transformación, se realiza codificación y decodificación de vídeo usando las unidades de transformación en una estructura de árbol variable, de modo que es posible reducir una cantidad de cálculo provocado por realizar la transformación usando unidades de transformación con diversos niveles y tamaños. También, puesto que una estructura jerárquica de unidades de transformación disponibles puede predecirse basándose en el nivel de división máxima de la unidad de transformación, pueden reducirse las tasas de bits para codificar información relacionada con la estructura jerárquica de las unidades de transformación, de modo que puede mejorarse una eficacia de transmisión de un resultado de codificación.

Con referencia a las Figuras 3 y 4, se describen las realizaciones en las que se ha dividido por la mitad y partido una altura y una anchura de una unidad de transformación en unidades de transformación en un nivel inferior. Sin embargo, las unidades de transformación en una estructura jerárquica no están limitadas a las realizaciones de las Figuras 3 y 4, y por lo tanto, en otras realizaciones, al menos una de la altura y la anchura de la unidad de transformación pueden dividirse de acuerdo con una relación arbitraria, de modo que pueden generarse las unidades de transformación en un nivel inferior.

Las Figuras 5 a 8 ilustran unidades de transformación de base, de acuerdo con las realizaciones de la presente invención.

En una realización, las unidades de transformación en una estructura de árbol variable incluyen unidades de transformación en niveles inferiores que se dividen desde una unidad de transformación de base, de modo que una forma y un tamaño de las unidades de transformación en la estructura de árbol variable pueden determinarse de acuerdo con una forma o un tamaño de la unidad de transformación de base.

Básicamente, el tamaño de la unidad de transformación de base es igual a o menor que una unidad de transformación máxima con respecto a un vídeo actual. Basándose en uno de los modos de codificación de una región, puede determinarse una forma de una unidad de transformación de base de la región.

Por ejemplo, una forma de una unidad de transformación de base de una región actual puede determinarse de acuerdo con un modo de predicción de entre los modos de codificación de la región. Por ejemplo, un tamaño de la unidad de transformación de base puede determinarse basándose en si el modo de predicción de la región actual es un inter modo o un intra modo.

Por ejemplo, la forma de la unidad de transformación de base de la región actual puede determinarse de acuerdo con un tamaño de una partición de entre los modos de codificación de la región. La partición indica una unidad de datos obtenida dividiendo la región para realizar una inter predicción o una intra predicción en la región, y el tamaño de la partición puede indicar una forma o el tamaño de la partición.

- Haciendo referencia a la Figura 5, puede determinarse una unidad de transformación de base que tiene la misma forma que una unidad de datos para predicción de una región. Por ejemplo, una partición 51 de tamaño  $2N \times 2N$  con respecto a una región 50 de tamaño  $2N \times 2N$  es una unidad de datos para una inter predicción o una intra predicción, y una unidad de transformación de base de la región 50 de tamaño  $2N \times 2N$  puede determinarse como una unidad de transformación de tamaño  $2N \times 2N$ .
- En otra realización, una unidad de transformación de base puede determinarse como una unidad de transformación que tiene un tamaño igual a o menor que un tamaño de particiones de una región. En este caso, puesto que la unidad de transformación de base se incluye en una partición situada en una región correspondiente, la unidad de transformación de base puede determinarse que no abarca un límite entre particiones de una región actual.
- Haciendo referencia a la Figura 6, cuando se determinan las particiones 61a y 61b de tamaño  $N \times 2N$  con respecto a la región 60 de tamaño  $2N \times 2N$ , pueden determinarse las unidades 65a, 65b, 65c y 65d de transformación de base de tamaño  $2N \times 2N$  con respecto a la región 60 de tamaño  $2N \times 2N$ . Las unidades 65a, 65b, 65c y 65d de transformación de base de tamaño  $2N \times 2N$  son menores que las particiones 61a y 61b de tamaño  $N \times 2N$ , de modo que las unidades 65a, 65b, 65c y 65d de transformación de base de tamaño  $2N \times 2N$  se incluyen en las particiones 61a y 61b de tamaño  $N \times 2N$  y no cruzan un límite entre las particiones 61a y 61b de tamaño  $N \times 2N$ .
- Haciendo referencia a la Figura 7, las particiones 71a y 71b de las cuales las anchuras se dividen asimétricamente pueden determinarse con respecto a una región 70 de tamaño  $4M \times 4M$ . Es decir, la partición 71a de tamaño  $M \times 4M$  y la partición 71b de tamaño  $3M \times 4M$  pueden determinarse con respecto a la región 70 de tamaño  $4M \times 4M$ . En este caso, pueden determinarse las unidades de transformación de base con respecto a la región 70 de tamaño  $4M \times 4M$  como las unidades 75a, 75b, 75c y 75d de transformación de tamaño  $M \times M$  y las unidades 75i y 75j de transformación de tamaño  $2M \times 2M$  por lo que no abarcan un límite entre las particiones 71a y 71b. Las unidades 75a, 75b, 75c y 75d de transformación de tamaño  $M \times M$  y las unidades 75i y 75j de transformación de tamaño  $2M \times 2M$  pueden incluirse en la partición 71a o 71b de una correspondiente región.
- También, es posible limitar un tamaño de unidades de transformación de base para que sea constante con respecto a una región actual. Haciendo referencia a la Figura 8, una partición 81a de tamaño  $M \times 4M$  y una partición 81b de tamaño  $3M \times 4M$  pueden determinarse con respecto a una región 80 de tamaño  $4M \times 4M$ . En este caso, las unidades 85a, 85b, 85c, 85d, 85e, 85f, 85g, 85h, 85i, 85j, 85k, 85l, 85m, 85n, 85o y 85p de transformación de tamaño  $M \times M$  que no abarcan un límite entre las particiones 81a y 81b y que tienen un tamaño constante pueden determinarse con respecto a la región 80 de tamaño  $4M \times 4M$ .
- En una realización, un nivel de división máxima puede determinarse de manera separada para cada uno de los modos de codificación de la región.
- El nivel de división máxima puede determinarse de manera separada para cada modo de predicción de la región. Por ejemplo, el nivel de división máxima puede determinarse de manera separada para cada modo de predicción de tal manera que la información de división máxima 'MaxTUDepthOfInter' se determina para una región con un inter modo o la información de división máxima 'MaxTUDepthOfIntra' se determina para una región con un intra modo o similar.
- El nivel de división máxima puede determinarse de manera separada para cada tipo de corte. Por ejemplo, el nivel de división máxima puede determinarse de manera separada para cada tipo de corte de tal manera que un valor de nivel de división máxima 'MaxTUDepthOfIntraSlice' se determina para un corte de tipo intra, un valor de nivel de división máxima 'MaxTUDepthOfInterP' se determina para un corte de tipo inter P, un valor de nivel de división máxima 'MaxTUDepthOfInterB' se determina para un corte de tipo inter B o similares. En este caso, un encabezamiento de corte puede codificarse mientras se incluye información de división máxima para cada tipo de corte.
- Las Figuras 9 y 10 ilustran unidades de transformación en una estructura de árbol variable, de acuerdo con las realizaciones de la presente invención.
- Las unidades de transformación en una estructura de árbol variable pueden incluir una unidad de transformación de base y unidades de transformación que se dividen desde la unidad de transformación de base y que tienen niveles inferiores que la unidad de transformación de base en al menos un nivel. Por ejemplo, la unidad de transformación de base es irrelevante para un tamaño de partición, pero, las unidades de transformación en niveles inferiores que se dividen desde la unidad de transformación de base puede determinarse basándose en el tamaño de partición.
- Por ejemplo, haciendo referencia a la Figura 9, un tipo de partición de una región 90 de tamaño  $2N \times 2N$  puede determinarse como las particiones 91a y 91b de tamaño  $N \times 2N$ . Independientemente de un tamaño de las particiones 91a y 91b de tamaño  $N \times 2N$ , un tamaño de una unidad 95 de transformación de base en un nivel cero puede determinarse como un tamaño  $2N \times 2N$  que es el mismo que la región 90 de tamaño  $2N \times 2N$ . Las unidades 97a, 97b, 97c y 97d de transformación en un primer nivel que es un nivel inferior desde la unidad 95 de transformación de base pueden determinarse como unidades de transformación de tamaño  $N \times N$  que no abarcan un límite entre las particiones 91a y 91b de tamaño  $N \times 2N$  y que son menores que las particiones 91a y 91b de tamaño  $N \times 2N$ .

Haciendo referencia a la Figura 10, un tipo de partición de una región 92 de tamaño 4Mx4M puede determinarse como las particiones 93a y 93b que son un tipo de partición asimétrica. Independientemente de los tamaños de las particiones 93a y 93b, un tamaño de una unidad 94 de transformación de base en un nivel cero puede determinarse como 4Mx4M que es igual a la región 92 de tamaño 4Mx4M.

5 En una realización, las unidades de transformación en un primer nivel que es un nivel inferior desde la unidad 94 de transformación de base en el nivel cero pueden determinarse como las unidades 96a, 96b, 96c, 96d, 96e, 96f, 96g y 96h de transformación de tamaño MxM, y las unidades 96i y 96j de transformación de tamaño 2Mx2M por lo que no abarcan un límite entre las particiones 93a y 93b.

10 En otra realización, las unidades de transformación en el primer nivel que es un nivel inferior desde la unidad 94 de transformación de base en el nivel cero pueden determinarse como las unidades 98a, 98b, 98c, 98d, 98e, 98f, 98g, 98h, 98i, 98j, 98k, 98l, 98m, 98n, 98o y 98p de transformación de tamaño MxM para que tengan un tamaño constante mientras no crucen el límite entre las particiones 93a y 93b.

15 Con referencia a las Figuras 5 a 8, las unidades de transformación de base de acuerdo con las realizaciones de la presente invención se han descrito anteriormente, y con referencia a las Figuras 9 y 10, las unidades de transformación en una estructura de árbol variable de acuerdo con las realizaciones de la presente invención se han descrito anteriormente. Aunque las unidades de transformación anteriormente mencionadas son unidades de datos de forma cuadrada obtenidas dividiendo por la mitad una altura y una anchura de una unidad de transformación en un nivel superior, una forma de las unidades de transformación no está limitada a la unidad de datos de forma cuadrada.

20 La Figura 11 es un diagrama de flujo de un procedimiento de codificación de vídeo que usa unidades de transformación en una estructura de árbol variable, de acuerdo con una realización de la presente invención.

25 En la operación 111, un vídeo se divide en una pluralidad de regiones y a continuación cada una de las regiones se codifica. Para codificar una región actual que tiene un tamaño predeterminado en el vídeo, la transformación puede realizarse en la región actual basándose en unidades de transformación en una estructura de árbol variable que se generan basándose en un nivel de división máxima de una unidad de transformación de entre unidades de transformación que se dividen jerárquicamente desde una unidad de transformación de base de la región actual. La región actual puede ser una unidad de datos, un macrobloque, una instantánea, un corte, o similares que son para una operación de codificación.

30 La unidad de transformación de base de la unidad de región actual puede ser una unidad de transformación de tamaño máximo y de nivel más superior que es usable con respecto a la región actual. Un tamaño de la unidad de transformación de base puede ser igual a o menor que un tamaño máximo de una unidad de transformación con respecto al vídeo.

35 El nivel de división máxima de la unidad de transformación puede indicar un número total de niveles de la unidad de transformación que se permiten con respecto a la región actual y puede corresponder a un número total de divisiones desde la unidad de transformación de base de la región actual a una unidad de transformación en un nivel más inferior que se permite con respecto a la región actual.

40 Las unidades de transformación en una estructura de árbol variable pueden incluir la unidad de transformación de base de acuerdo con el nivel de división máxima de la unidad de transformación, y unidades de transformación de acuerdo con niveles que se generan dividiéndose gradualmente desde la unidad de transformación de base al nivel de división máxima de la unidad de transformación.

El nivel de división máxima de la unidad de transformación puede determinarse de manera separada para cada uno de los modos de codificación que incluyen un modo de predicción, un tipo de corte, y similares que se usan en la codificación de la región actual.

45 La unidad de transformación de base con respecto a la región actual puede establecerse de manera variable de acuerdo con una característica de imagen.

50 Basándose en la unidad de transformación de base que puede establecerse de manera variable o el nivel de división máxima que puede establecerse de manera variable, las unidades de transformación en una estructura de árbol variable pueden usarse en la transformación con respecto a la región actual. Las unidades de transformación que provocan un error mínimo y que están de acuerdo con un resultado de una operación en la que se realiza la transformación en la región actual usando las unidades de transformación en una estructura de árbol variable pueden determinarse como unidades de transformación con una profundidad de transformación con respecto a la región actual, de modo que puede emitirse un coeficiente de transformada.

55 En la operación 112, los datos codificados de la región actual y la información acerca de un modo de codificación que se determina en la codificación de la región actual se emiten en la forma de una secuencia de bits, en la que los datos codificados se generan mediante la codificación que incluye la transformación basándose en la unidad de transformación determinada en la operación 111. También, la información de tamaño máximo y la información de



tamaño mínimo de la unidad de transformación con respecto al vídeo, y la información de división máxima de unidad de transformación que indica el nivel de división máxima de la unidad de transformación se emiten mientras se incluyen en la secuencia de bits.

5 Para cada una de las unidades de transformación que se encuentran entre las unidades de transformación determinadas como una unidad de transformación final con respecto a la región actual y que excluyen una unidad de transformación que tiene un tamaño mínimo de la región actual, la información de subdivisión de unidad de transformación que indica si cada una de las unidades de transformación se divide en unidades de transformación en un siguiente nivel inferior puede codificarse y puede emitirse.

10 También, para cada una de las unidades de transformación, la información de patrón de unidades de transformación jerárquica que indica la existencia de una unidad de transformación que incluye un coeficiente distinto de cero y que se encuentra entre las unidades de transformación en un nivel inferior puede codificarse y puede emitirse.

La Figura 12 es un diagrama de flujo de un procedimiento de decodificación de vídeo que usa unidades de transformación en una estructura de árbol variable, de acuerdo con una realización de la presente invención.

En la operación 121, se recibe una secuencia de bits que incluye datos codificados de un vídeo.

15 En la operación 122, se analiza la secuencia de bits recibida en la operación 121, de modo que los datos codificados de una región actual y la información acerca de un modo de codificación que se determina en la codificación de la región actual se extraen desde la secuencia de bits.

20 También, la información de tamaño máximo y la información de tamaño mínimo de una unidad de transformación con respecto al vídeo, y la información de división máxima de unidad de transformación se extraen desde la secuencia de bits. La información de división máxima de unidad de transformación acerca de una unidad de transformación con respecto a la región actual puede leerse de manera separada para cada modo de predicción o cada tipo de corte que se usa en la codificación de la región actual.

La información de subdivisión de unidad de transformación o información de patrón de unidades de transformación jerárquica puede extraerse desde la secuencia de bits.

25 En la operación 123, el modo de codificación de la región actual puede leerse desde la información extraída acerca del modo de codificación, y los datos codificados de la región actual pueden decodificarse basándose en el modo de codificación.

30 En particular, un nivel de división máxima de la unidad de transformación con respecto a la región actual puede leerse basándose en la información de división máxima de unidad de transformación. Las unidades de transformación en una estructura de árbol variable pueden determinarse basándose en el nivel de división máxima, en el que las unidades de transformación en una estructura de árbol variable se generan basándose en el nivel de división máxima de la unidad de transformación y se determinan de entre unidades de transformación que se dividen jerárquicamente desde una unidad de transformación de base con respecto a la región actual. Una unidad de transformación con una profundidad de transformación puede determinarse de entre las unidades de transformación en una estructura de árbol variable, y puede realizarse transformación inversa en un coeficiente de transformada de la región actual usando la unidad de transformación con la profundidad de la transformación. Posteriormente, el vídeo puede restaurarse combinando resultados codificados con respecto a las imágenes.

40 Un número total máximo de divisiones o el nivel de división máxima de la unidad de transformación puede determinarse de manera separada de acuerdo con una característica de dominio espacial de una imagen. El número total máximo de divisiones o el nivel de división máxima de la unidad de transformación puede determinarse basándose en una capacidad o un entorno de comunicación de datos de un sistema de codificación o un sistema de decodificación. Puesto que el número total máximo de divisiones o el nivel de división máxima se limita de manera selectiva desde la unidad de transformación de base, puede reducirse una cantidad de cálculo de codificación y una cantidad de bits de transmisión.

45 En lo sucesivo, con referencia a las Figuras 13 a 27, se describirá un aparato de codificación de vídeo y un aparato de decodificación de vídeo, y un procedimiento de codificación de vídeo y un procedimiento de decodificación de vídeo, que usan una unidad de transformación en una estructura de árbol y una unidad de transformación en una estructura de árbol variable.

50 La Figura 13 es un diagrama de bloques de un aparato de codificación de vídeo que usa unidad de transformación en estructura de árbol y unidad de transformación en estructura 100 de árbol variable, de acuerdo con una realización de la presente invención.

55 El aparato de codificación de vídeo que usa unidad de transformación en estructura de árbol y unidad de transformación en estructura 100 de árbol variable incluye un divisor 110 de unidad de codificación máxima, un determinador 120 de unidad de codificación, y una unidad 130 de salida. En lo sucesivo, por conveniencia de descripción, el aparato de codificación de vídeo que usa unidad de transformación en estructura de árbol variable y

unidad de transformación en estructura 100 de árbol variable se denomina como el 'aparato 100 de codificación de vídeo'.

5 El divisor 110 de unidad de codificación máxima puede dividir una instantánea actual basándose en una unidad de codificación máxima para la instantánea actual de una imagen. Si la instantánea actual es mayor que la unidad de codificación máxima, los datos de imagen de la instantánea actual pueden dividirse en la al menos una unidad de codificación máxima. La unidad de codificación máxima puede ser una unidad de datos que tiene un tamaño de 32x32, 64x64, 128x128, 256x256, etc., en la que una forma de la unidad de datos es un cuadrado que tiene una anchura y altura en cuadrados de 2. Los datos de imagen pueden emitirse al determinador 120 de unidad de codificación de acuerdo con la al menos una unidad de codificación máxima.

10 Una unidad de codificación puede caracterizarse por un tamaño máximo y una profundidad. La profundidad indica un número de veces que la unidad de codificación se divide espacialmente a partir de la unidad de codificación máxima. Por consiguiente, a medida que la profundidad se hace más profunda, las unidades de codificación más profundas de acuerdo con las profundidades pueden dividirse desde la unidad de codificación máxima a una unidad de codificación mínima. Una profundidad de la unidad de codificación máxima es una profundidad más superior y una  
15 profundidad de la unidad de codificación mínima es una profundidad más inferior. Puesto que un tamaño de una unidad de codificación que corresponde a cada profundidad se reduce a medida que la profundidad de la unidad de codificación máxima se hace profunda, una unidad de codificación que corresponde a una profundidad superior puede incluir una pluralidad de unidades de codificación que corresponden a profundidades inferiores.

20 Como se ha descrito anteriormente, los datos de imagen de la instantánea actual se dividen en una o más unidades de codificación máxima de acuerdo con un tamaño máximo de la unidad de codificación, y cada una de las unidades de codificación máxima puede incluir unidades de codificación más profundas que se dividen de acuerdo con las profundidades. Puesto que la unidad de codificación máxima se divide de acuerdo con las profundidades, los datos de imagen de un dominio espacial incluidos en la unidad de codificación máxima pueden clasificarse jerárquicamente de acuerdo con las profundidades.

25 Puede predeterminarse una profundidad máxima y un tamaño máximo de una unidad de codificación, que limita el número total de veces que se divide jerárquicamente una altura y una anchura de la unidad de codificación máxima.

El determinador 120 de unidad de codificación codifica al menos una región de división obtenida dividiendo una región de la unidad de codificación máxima de acuerdo con las profundidades, y determina una profundidad para emitir unos datos de imagen finalmente codificados de acuerdo con la al menos una región de división. El  
30 determinador 120 de unidad de codificación codifica al menos una región de división obtenida dividiendo una región de la unidad de codificación máxima de acuerdo con las profundidades, y determina una profundidad para emitir unos datos de imagen finalmente codificados de acuerdo con la al menos una región de división. Por ejemplo, el determinador 120 de unidad de codificación determina una profundidad codificada codificando los datos de imagen en las unidades de codificación más profundas de acuerdo con las profundidades, de acuerdo con la unidad de  
35 codificación máxima de la instantánea actual, y selecciona una profundidad que tiene los mínimos errores de codificación. Por ejemplo, el determinador 120 de unidad de codificación determina una profundidad codificada codificando los datos de imagen en las unidades de codificación más profundas de acuerdo con las profundidades, de acuerdo con la unidad de codificación máxima de la instantánea actual, y selecciona una profundidad que tiene los mínimos errores de codificación. Por lo tanto, los datos de imagen codificados de la unidad de codificación que corresponden a la profundidad codificada determinada se emiten por el determinador 120 de unidad de codificación. Por lo tanto, los datos de imagen codificados de la unidad de codificación que corresponden a la profundidad codificada determinada se emiten por el determinador 120 de unidad de codificación. También, las unidades de codificación que corresponden a la profundidad codificada pueden considerarse como unidades de codificación codificadas. También, las unidades de codificación que corresponden a la profundidad codificada pueden  
40 considerarse como unidades de codificación codificadas.

La profundidad codificada determinada y los datos de imagen codificados de acuerdo con la profundidad codificada determinada se emiten a la unidad 130 de salida.

Los datos de imagen en la unidad de codificación máxima se codifican basándose en las unidades de codificación más profundas que corresponden a al menos una profundidad igual a o por debajo de la profundidad máxima, y los  
50 resultados de la codificación de los datos de imagen se comparan basándose en cada una de las unidades de codificación más profundas. Una profundidad que tiene los mínimos errores de codificación puede seleccionarse después de comparar errores de codificación de las unidades de codificación más profundas. Al menos una profundidad codificada puede seleccionarse para cada unidad de codificación máxima.

El tamaño de la unidad de codificación máxima se divide como una unidad de codificación que se divide jerárquicamente de acuerdo con las profundidades, y a medida que el número de unidades de codificación aumenta. También, incluso si las unidades de codificación corresponden a la misma profundidad en una unidad de codificación máxima, se determina si dividir cada una de las unidades de codificación que corresponden a la misma profundidad a una profundidad inferior midiendo un error de codificación de los datos de imagen de cada unidad de codificación de manera separada. Por consiguiente, incluso cuando se incluyen los datos de imagen en una unidad de  
55

codificación máxima, los datos de imagen se dividen en regiones de acuerdo con las profundidades y los errores de codificación pueden diferenciarse de acuerdo con las regiones en una unidad de codificación máxima. Por lo tanto, las profundidades candidatas pueden diferenciarse de acuerdo con las regiones en los datos de imagen. Por lo tanto, una o más profundidades codificadas pueden determinarse en una unidad de codificación máxima, y los datos de imagen de la unidad de codificación máxima pueden dividirse de acuerdo con unidades de codificación de al menos una profundidad codificada.

Por consiguiente, el determinador 120 de unidad de codificación puede determinar unidades de codificación que tienen una estructura de árbol incluida en la unidad de codificación máxima. Las unidades de codificación que tienen una estructura de árbol incluyen unidades de codificación que corresponden a una profundidad determinada para que sea la profundidad codificada, de entre todas las unidades de codificación más profundas incluidas en la unidad de codificación máxima. Una unidad de codificación de una profundidad codificada puede determinarse jerárquicamente de acuerdo con las profundidades en la misma región de la unidad de codificación máxima, y puede determinarse independientemente en diferentes regiones. De manera similar, una profundidad codificada en una región actual puede determinarse independientemente desde una profundidad codificada en otra región.

Una profundidad máxima es un índice relacionado con un número de veces de división desde una unidad de codificación máxima a una unidad de codificación mínima. Una primera profundidad máxima puede indicar un número total de veces de división desde la unidad de codificación máxima a la unidad de codificación mínima. Una segunda profundidad máxima puede indicar un número total de niveles de profundidad desde la unidad de codificación máxima a la unidad de codificación mínima. Por ejemplo, cuando una profundidad de la unidad de codificación máxima es 0, una profundidad de una unidad de codificación, en la que se divide la unidad de codificación máxima una vez, puede establecerse a 1, y una profundidad de una unidad de codificación, en la que se divide la unidad de codificación máxima dos veces, puede establecerse a 2. En este punto, si la unidad de codificación mínima es una unidad de codificación en la que la unidad de codificación máxima se divide cuatro veces, existen 5 niveles de profundidad de profundidades 0, 1, 2, 3 y 4. En este caso, la primera profundidad máxima puede establecerse a 4, y la segunda profundidad máxima puede establecerse a 5.

Puede realizarse la codificación de predicción y transformación de acuerdo con la unidad de codificación máxima. La codificación por predicción y la transformación pueden realizarse también basándose en las unidades de codificación más profundas de acuerdo con una profundidad igual a, o profundidades menores que, la profundidad máxima, de acuerdo con la unidad de codificación máxima. La transformación puede realizarse de acuerdo con un procedimiento de transformación ortogonal o transformación de números enteros.

Puesto que el número de unidades de codificación más profundas aumenta cada vez que se divide la unidad de codificación máxima de acuerdo con las profundidades, la codificación que incluye la codificación por predicción y la transformación puede realizarse en todas las unidades de codificación más profundas generadas a medida que la profundidad se hace más profunda. Por conveniencia de descripción, la codificación por predicción y la transformación se describirán ahora basándose en una unidad de codificación de una profundidad actual, en una unidad de codificación máxima.

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

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

Para realizar la codificación por predicción en la unidad de codificación máxima, la codificación por predicción puede realizarse basándose en una unidad de codificación que corresponde a una profundidad codificada, es decir, basándose en una unidad de codificación que ya no se divide en unidades de codificación que corresponden a una profundidad inferior. En lo sucesivo, la unidad de codificación que ya no se divide y se vuelve una unidad de base para la codificación por predicción se denominará ahora como una unidad de predicción. Una partición obtenida dividiendo la unidad de predicción puede incluir una unidad de predicción o una unidad de datos obtenida dividiendo al menos una de una altura y una anchura de la unidad de predicción.

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

Un modo de predicción de la unidad de predicción puede ser al menos uno de un intra modo, un inter modo y un modo de salto. Por ejemplo, el intra modo o el inter modo pueden realizarse en la partición de  $2N \times 2N$ ,  $2N \times N$ ,  $N \times 2N$ ,

o NxN. También, el modo de salto puede realizarse únicamente en la partición de  $2N \times 2N$ . La codificación se realiza independientemente en unidades de predicción en una unidad de codificación, seleccionando de esta manera un modo de predicción que tiene un mínimo error de codificación.

5 El aparato 100 de codificación de vídeo puede realizar también la transformación en los datos de imagen en una unidad de codificación basándose no únicamente en la unidad de codificación para codificar los datos de imagen, sino también basándose en una unidad de datos que es diferente de la unidad de codificación.

10 Como se ha descrito con referencia a las Figuras 1 a 12, para realizar la transformación en la unidad de codificación, la transformación puede realizarse basándose en una unidad de datos que tiene un tamaño igual a o menor que la unidad de transformación. Por ejemplo, la unidad de transformación para la transformación puede incluir una unidad de transformación para un intra modo y una unidad de transformación para un inter modo.

15 De manera similar a las unidades de transformación en una estructura de árbol de acuerdo con las realizaciones anteriores, la unidad de transformación en la unidad de codificación puede dividirse de manera recursiva en unidades de transformación de tamaño más pequeño, de modo que los datos residuales en la unidad de codificación pueden dividirse de acuerdo con las unidades de transformación en la estructura de árbol de acuerdo con las profundidades de transformación.

20 Una profundidad de transformación que indica un número de veces de división para alcanzar la unidad de transformación dividiendo una altura y una anchura de la unidad de codificación puede establecerse también para la unidad de transformación. Por ejemplo, en una unidad de codificación actual de  $2N \times 2N$ , una profundidad de transformación puede ser 0 cuando un tamaño de una unidad de transformación es  $2N \times 2N$ , la profundidad de transformación puede ser 1 cuando el tamaño de la unidad de transformación es NxN, y la profundidad de transformación puede ser 2 cuando el tamaño de la unidad de transformación es  $N/2 \times N/2$ . En otras palabras, la unidad de transformación en la estructura de árbol puede establecerse para la unidad de transformación de acuerdo con una profundidad de transformación.

25 La información de codificación de acuerdo con unidades de codificación que corresponden a una profundidad codificada no usa únicamente información acerca de la profundidad codificada, sino también información acerca de información relacionada con la codificación de predicción y transformación. Por consiguiente, el determinador 120 de unidad de codificación puede determinar una profundidad codificada que tiene un error de codificación mínimo y puede determinar también un tipo de partición en una unidad de predicción, un modo de predicción de acuerdo con unidades de predicción, y un tamaño de una unidad de transformación para transformación.

30 También, el determinador 120 de unidad de codificación puede realizar la transformación usando las unidades de transformación en una estructura de árbol variable en un procedimiento de codificación con respecto a la unidad de codificación, basándose en un nivel de división máxima de la unidad de transformación que se establece de manera limitada con antelación para cada una de una unidad de codificación máxima o la unidad de codificación actual.

35 La unidad de transformación en la estructura de árbol variable basada en el nivel de división máxima de la unidad de transformación corresponde a la descripción anteriormente mencionada con referencia a las Figuras 1 a 12. Es decir, la unidad de transformación en la estructura de árbol variable puede incluir una unidad de transformación de base, y unidades de transformación en niveles inferiores que se dividen desde la unidad de transformación de base al nivel de división máxima que se permite para la unidad de codificación.

40 La unidad de transformación de base y el nivel de división máxima pueden establecerse de manera variable de acuerdo con modos de codificación. Por ejemplo, una forma de una unidad de transformación de base de una imagen actual puede determinarse de acuerdo con un tamaño de una partición o un modo de predicción de entre los modos de codificación de la unidad de codificación.

45 La partición puede indicar una unidad de datos obtenida dividiendo una unidad de predicción de la unidad de codificación, y la unidad de predicción puede ser una partición que tiene el mismo tamaño que la unidad de codificación. En una realización, la unidad de transformación de base puede determinarse que tiene la misma forma que la unidad de predicción de la unidad de codificación. En otra realización, la unidad de transformación de base puede determinarse que tiene un tamaño igual a o menor que un tamaño de las particiones para no abarcar un límite entre las particiones.

50 También, mientras que la unidad de transformación de base es mayor que la partición, las unidades de transformación en niveles inferiores que la unidad de transformación de base puede determinarse que tienen un tamaño menor que el tamaño de las particiones para no abarcar el límite entre las particiones.

El determinador 120 de unidad de codificación puede realizar la transformación en cada una de las unidades de codificación usando las unidades de transformación en la estructura de árbol variable y por lo tanto puede determinar las unidades de transformación en la estructura de árbol.

55 Las unidades de codificación de acuerdo con una estructura de árbol en una unidad de codificación máxima y un procedimiento de determinación de una partición y una unidad de transformación en una estructura de árbol, de

acuerdo con una o más realizaciones, se describirán en detalle más adelante con referencia a las Figuras 15 a 25.

El determinador 120 de unidad de codificación puede medir un error de codificación de unidades de codificación más profundas de acuerdo con las profundidades usando Optimización de Tasa-Distorsión en multiplicadores Langranianos.

5 La unidad 130 de salida emite los datos de imagen de la unidad de codificación máxima, que se codifican basándose en la al menos una profundidad codificada determinada por el determinador 120 de unidad de codificación, y la información acerca del modo de codificación de acuerdo con la profundidad codificada, en secuencias de bits.

10 Los datos de imagen codificados pueden obtenerse codificando datos residuales de una imagen. La información acerca del modo de codificación de acuerdo con la profundidad codificada puede incluir al menos uno de información acerca de la profundidad codificada, información acerca del tipo de partición en la unidad de predicción, el modo de predicción y el tamaño de la unidad de transformación.

15 La información acerca de la profundidad codificada puede definirse usando información de división de acuerdo con las profundidades, que indica si la codificación se realiza en unidades de codificación de una profundidad inferior en lugar de una profundidad actual. Si la profundidad actual de la unidad de codificación actual es la profundidad codificada, los datos de imagen en la unidad de codificación actual se codifican y emiten, y por lo tanto la información de división puede definirse para no dividir la unidad de codificación actual a una profundidad inferior. Como alternativa, si la profundidad actual de la unidad de codificación actual no es la profundidad codificada, la codificación se realiza en la unidad de codificación de la profundidad inferior. Por lo tanto, la información de división puede definirse para dividir la unidad de codificación actual para obtener las unidades de codificación de la profundidad inferior.

20 Si la profundidad actual no es la profundidad codificada, la codificación se realiza en la unidad de codificación que se divide en la unidad de codificación de la profundidad inferior. Puesto que la al menos una unidad de codificación de la profundidad inferior existe en una unidad de codificación de la profundidad actual, la codificación se realiza de manera repetitiva en cada unidad de codificación de la profundidad inferior. Por lo tanto, la codificación puede realizarse de manera recursiva para las unidades de codificación que tienen la misma profundidad.

25 Puesto que las unidades de codificación que tienen una estructura de árbol se determinan para una unidad de codificación máxima, y la información acerca de al menos uno modo de codificación se determina para una unidad de codificación de una profundidad codificada, la información acerca de al menos uno modo de codificación puede determinarse para una unidad de codificación máxima. También, una profundidad codificada de los datos de imagen de la unidad de codificación máxima puede ser diferente de acuerdo con las localizaciones puesto que los datos de imagen se dividen jerárquicamente de acuerdo con las profundidades. Por lo tanto, la información acerca de la profundidad codificada y el modo de codificación puede establecerse para los datos de imagen.

30 Por consiguiente, la unidad 130 de salida puede asignar información de codificación acerca de una profundidad codificada correspondiente y un modo de codificación a al menos una de la unidad de codificación, la unidad de predicción, y una unidad mínima incluida en la unidad de codificación máxima.

35 La unidad mínima puede ser una unidad de datos rectangular obtenida dividiendo la unidad de codificación mínima que tiene la profundidad más inferior en 4. Como alternativa, la unidad mínima puede ser una unidad de datos rectangular de tamaño máximo que puede incluirse en todas las unidades de codificación, unidades de predicción, unidades de partición y unidades de transformación incluidas en la unidad de codificación máxima.

40 Por ejemplo, la información de codificación emitida a través de la unidad 130 de salida puede clasificarse en información de codificación de acuerdo con unidades de codificación, e información de codificación de acuerdo con unidades de predicción. La información de codificación de acuerdo con las unidades de codificación puede incluir al menos una de información acerca del modo de predicción e información acerca de un tamaño de las particiones. La información de codificación de acuerdo con las unidades de predicción puede incluir al menos una de información acerca de una dirección estimada de un inter modo, información acerca de un índice de imagen de referencia del inter modo, información acerca de un vector de movimiento, información acerca de un componente de prominancia de un intra modo, e información acerca de un procedimiento de interpolación del intra modo.

45 La información acerca de un tamaño máximo de la unidad de codificación definida de acuerdo con instantáneas, cortes, o grupos de instantáneas (GOP), y la información acerca de una profundidad máxima puede insertarse en un encabezamiento de una secuencia de bits, un conjunto de parámetros de secuencia (SPS) o un conjunto de parámetros de instantánea.

50 También, la información de tamaño máximo de la unidad de transformación y la información de tamaño mínimo de la unidad de transformación que se permiten para el vídeo actual pueden emitirse mediante el encabezamiento de la secuencia de bits, el SPS o el conjunto de parámetros de instantánea. La unidad 130 de salida puede emitir la información de división máxima de unidad de transformación, la información de índice de transformación, la información de subdivisión de unidad de transformación, la información de patrón de unidades de transformación jerárquica, y similares que se han descrito anteriormente con referencia a las Figuras 1 a 12.

5 En el aparato 100 de codificación de vídeo, la unidad de codificación más profunda puede ser una unidad de codificación obtenida dividiendo en dos al menos una de una altura y una anchura de una unidad de codificación de una profundidad superior, que se encuentra un nivel por encima. En otras palabras, cuando el tamaño de la unidad de codificación de la profundidad actual es  $2N \times 2N$ , el tamaño de la unidad de codificación de la profundidad inferior puede ser  $N \times N$ . También, la unidad de codificación de la profundidad actual que tiene el tamaño de  $2N \times 2N$  puede incluir 4 de las unidades de codificación de la profundidad inferior.

10 Por consiguiente, el aparato 100 de codificación de vídeo puede formar las unidades de codificación que tienen la estructura de árbol determinando unidades de codificación que tienen una forma óptima y un tamaño óptimo para cada unidad de codificación máxima, basándose en el tamaño de la unidad de codificación máxima y la profundidad máxima ambas determinadas considerando las características de la instantánea actual. También, puesto que la codificación puede realizarse en cada unidad de codificación máxima usando cualquiera de diversos modos de predicción y transformaciones, puede determinarse un modo de codificación óptimo considerando las características de la unidad de codificación de diversos tamaños de imagen.

15 Por lo tanto, si una imagen que tiene una alta resolución o una gran cantidad de datos se codifica en unidades de macrobloques de la técnica relacionada, un número de macrobloques por instantánea aumenta excesivamente. Por consiguiente, un número de piezas de información comprimida generadas para cada macrobloque aumenta, y por lo tanto, es difícil transmitir la información comprimida y la eficacia de compresión de datos se reduce. Sin embargo, usando el aparato 100 de codificación de vídeo, la eficacia de compresión de imagen puede aumentarse puesto que una unidad de codificación se ajusta mientras se consideran características de una imagen mientras se aumenta un tamaño máximo de una unidad de codificación mientras se considera un tamaño de la imagen.

20 El aparato 100 de codificación de vídeo de la Figura 13 puede corresponder al aparato 10 de codificación de vídeo anteriormente descrito con referencia a la Figura 1.

25 Es decir, en el aparato 10 de codificación de vídeo, una región actual puede indicar una unidad de codificación actual que es una de las unidades de codificación obtenidas dividiendo una instantánea actual de un vídeo para codificar el vídeo.

30 La unidad 11 de determinación de unidad de transformación del aparato 10 de codificación de vídeo puede dividir la instantánea actual en unidades de codificación máxima, puede realizar transformación en cada una de las unidades de codificación máxima basándose en unidades de codificación de acuerdo con las profundidades, puede seleccionar unidades de codificación con una profundidad codificada a partir de la cual se emite un resultado de codificación que genera un error mínimo, y a continuación puede determinar unidades de codificación en una estructura de árbol.

35 Cuando la unidad 11 de determinación de unidad de transformación del aparato 10 de codificación de vídeo realiza la transformación basándose en la unidad de codificación, la unidad 11 de determinación de unidad de transformación puede realizar la transformación basándose en una unidad de transformación. En particular, las unidades de transformación en una estructura de árbol variable pueden configurarse basándose en un nivel de división máxima de la unidad de transformación que establece la unidad de codificación máxima o cada una de las unidades de codificación actual.

40 La unidad 11 de determinación de unidad de transformación del aparato 10 de codificación de vídeo puede realizar la transformación en cada una de las unidades de codificación, basándose en la unidad de transformación en la estructura de árbol variable, y por lo tanto puede determinar una profundidad de transformación que está a un nivel de una unidad de transformación, y una profundidad codificada de una unidad de codificación, que puede provocar un resultado de codificación óptimo. Por consiguiente, la unidad 11 de determinación de unidad de transformación puede determinar unidades de codificación en una estructura de árbol y unidades de transformación en una estructura de árbol variable con respecto a cada una de las unidades de codificación máxima.

45 La unidad 13 de salida del aparato 10 de codificación de vídeo puede emitir datos codificados de una instantánea que se codifica de acuerdo con las unidades de codificación en la estructura de árbol con respecto a cada una de las unidades de codificación máxima, puede codificar información acerca de profundidades codificadas y modos de codificación de las unidades de codificación en la estructura de árbol, puede codificar información de división máxima de unidad de transformación y puede emitir las.

50 La Figura 14 es un diagrama de bloques de un aparato de decodificación de vídeo que usa unidad de transformación en estructura de árbol y unidad de transformación en estructura 200 de árbol variable, de acuerdo con una realización de la presente invención.

55 El aparato de decodificación de vídeo que usa unidad de transformación en estructura de árbol y unidad de transformación en estructura 200 de árbol variable incluye un receptor 210, un extractor 220 de datos de imagen e información de codificación, y un decodificador 230 de datos de imagen. En lo sucesivo, por conveniencia de descripción, el aparato de decodificación de vídeo que usa unidad de transformación en estructura de árbol y unidad de transformación en estructura 200 de árbol variable se denomina como el 'aparato 200 de decodificación de vídeo'.

Las definiciones de diversos términos, tales como una unidad de codificación, una profundidad, una unidad de predicción, una unidad de transformación e información acerca de diversos modos de codificación, para diversas operaciones del aparato 200 de decodificación de vídeo son las mismas o similares a aquellas anteriormente descritas con referencia a la Figura 13 y al aparato 100 de codificación de vídeo.

5 El receptor 210 recibe y analiza una secuencia de bits de un vídeo codificado. El extractor 220 de datos de imagen e información de codificación extrae datos de imagen codificados para cada unidad de codificación desde la secuencia de bits analizada, en la que las unidades de codificación tienen una estructura de árbol de acuerdo con cada unidad de codificación máxima, y emite los datos de imagen extraídos al decodificador 230 de datos de imagen. El extractor 220 de datos de imagen e información de codificación puede extraer información acerca de un tamaño máximo de  
10 una unidad de codificación de una instantánea actual desde un encabezamiento que corresponde a la instantánea actual, un SPS o un conjunto de parámetros de instantánea.

También, el extractor 220 de datos de imagen e información de codificación extrae información acerca de una profundidad codificada y un modo de codificación para las unidades de codificación que tienen una estructura de árbol de acuerdo con cada unidad de codificación máxima, desde la secuencia de bits analizada. La información extraída acerca de la profundidad codificada y el modo de codificación se emite al decodificador 230 de datos de imagen. Por lo tanto, los datos de imagen en una secuencia de bits se dividen en la unidad de codificación máxima de modo que el decodificador 230 de datos de imagen decodifica los datos de imagen para cada unidad de  
15 codificación máxima.

La información acerca de la profundidad codificada y el modo de codificación de acuerdo con la unidad de codificación máxima puede establecerse para información acerca de al menos una unidad de codificación que corresponde a la profundidad codificada. Adicionalmente, la información acerca del modo de codificación puede incluir al menos uno de información acerca de un tipo de partición de una unidad de codificación correspondiente que corresponde a la profundidad codificada, información acerca de un modo de predicción, y un tamaño de una unidad de transformación. También, la información de división de acuerdo con las profundidades puede extraerse como la información acerca de la profundidad codificada.  
20

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

También, el extractor 220 de datos de imagen e información de codificación puede extraer información de tamaño máximo de una unidad de transformación e información de tamaño mínimo de una unidad de transformación que se permiten para el vídeo actual desde el encabezamiento de la secuencia de bits, el SPS o el conjunto de parámetros de instantánea. El extractor 220 de datos de imagen e información de codificación puede extraer la información de división máxima de unidad de transformación, información de índice de transformación, información de subdivisión de unidad de transformación, información de patrón de unidades de transformación jerárquica, y similares, que están relacionadas con la unidad de transformación anteriormente descrita con referencia a las Figuras 1 a 12, como información de codificación.  
40

Puesto que la información de codificación acerca de la profundidad codificada y el modo de codificación puede asignarse a una unidad de datos predeterminada de entre una unidad de codificación correspondiente, una unidad de predicción, y una unidad mínima, el extractor 220 de datos de imagen e información de codificación puede extraer la información acerca de la profundidad codificada y el modo de codificación de acuerdo con las unidades de datos predeterminadas. Las unidades de datos predeterminadas a las que se asigna la misma información acerca de la profundidad codificada y el modo de codificación puede inferirse para que sean las unidades de datos incluidas en la misma unidad de codificación máxima.  
45

El decodificador 230 de datos de imagen restaura la instantánea actual decodificando los datos de imagen en cada unidad de codificación máxima basándose en la información acerca de la profundidad codificada y el modo de codificación de acuerdo con las unidades de codificación máxima. En otras palabras, el decodificador 230 de datos de imagen puede decodificar los datos de imagen codificados basándose en la información extraída acerca del tipo de partición, el modo de predicción, y la unidad de transformación para cada unidad de codificación de entre las unidades de codificación que tienen la estructura de árbol incluida en cada unidad de codificación máxima. Un procedimiento de decodificación puede incluir al menos uno de una predicción que incluye intra predicción y compensación de movimiento, y una transformación inversa. La transformación inversa puede realizarse de acuerdo con el procedimiento de transformación ortogonal inversa o transformación de números enteros inversa.  
50  
55

El decodificador 230 de datos de imagen puede realizar intra predicción o compensación de movimiento de acuerdo con una partición y un modo de predicción de cada unidad de codificación, basándose en la información acerca del

tipo de partición y el modo de predicción de la unidad de predicción de la unidad de codificación de acuerdo con las profundidades codificadas.

5 También, para realizar transformación inversa de acuerdo con las unidades de codificación máxima, el decodificador 230 de datos de imagen puede leer unidades de transformación en una estructura de árbol que incluyen información de tamaño acerca de las unidades de transformación de las unidades de codificación de acuerdo con las profundidades codificadas, y a continuación puede realizar la transformación inversa en cada una de las unidades de codificación, basándose en la unidad de transformación.

10 El decodificador 230 de datos de imagen puede determinar las unidades de transformación en la estructura de árbol variable que incluye niveles inferiores permisibles desde una unidad de transformación de base, basándose en un nivel de división máxima de una unidad de transformación con respecto a cada una de las unidades de codificación. El decodificador 230 de datos de imagen puede determinar unidades de transformación con una profundidad de transformación que se requiere para realizar transformación inversa en el coeficiente de transformada, en el que las unidades de transformación se determinan de entre las unidades de transformación en la estructura de árbol variable, puede realizar la transformación inversa en el coeficiente de transformada, y por lo tanto puede restaurar un valor de píxel.

15 El decodificador 230 de datos de imagen puede determinar al menos una profundidad codificada de una unidad de codificación máxima actual usando información de división de acuerdo con las profundidades. Si la información de división indica que los datos de imagen ya no se dividen en la profundidad actual, la profundidad actual es una profundidad codificada. Por consiguiente, el decodificador 230 de datos de imagen puede decodificar datos codificados de al menos una unidad de codificación que corresponde a cada profundidad codificada en la unidad de codificación máxima actual usando la información acerca del tipo de partición de la unidad de predicción, el modo de predicción, y el tamaño de la unidad de transformación para cada unidad de codificación que corresponde a la profundidad codificada, y emite los datos de imagen de la unidad de codificación máxima actual.

20 En otras palabras, las unidades de datos que incluyen la información de codificación que incluyen la misma información de división pueden recopilarse observando la información de codificación establecida asignada para la unidad de datos predeterminada de entre la unidad de codificación, la unidad de predicción y la unidad mínima. Además, las unidades de datos recopiladas pueden considerarse que son una unidad de datos a decodificarse por el decodificador 230 de datos de imagen en el mismo modo de codificación.

25 El aparato 200 de decodificación de vídeo puede obtener información acerca de al menos una unidad de codificación que genera el error de codificación mínimo cuando se realiza de manera recursiva la codificación para cada unidad de codificación máxima, y puede usar la información para decodificar la instantánea actual. En otras palabras, pueden decodificarse las unidades de codificación que tienen la estructura de árbol determinada para que sean las unidades de codificación óptimas en cada unidad de codificación máxima. También, un tamaño máximo de la unidad de codificación puede determinarse considerando la resolución y una cantidad de datos de imagen.

30 Por consiguiente, incluso si los datos de imagen tienen una alta resolución y una gran cantidad de datos, los datos de imagen pueden decodificarse de manera eficaz y restaurarse usando un tamaño de una unidad de codificación y un modo de codificación, que se determinan de manera adaptativa de acuerdo con las características de los datos de imagen, usando información acerca de un modo de codificación óptimo recibido desde un codificador.

35 El aparato 200 de decodificación de vídeo de la Figura 14 puede corresponder al aparato 20 de decodificación de vídeo anteriormente descrito con referencia a la Figura 2.

40 Es decir, en el aparato 20 de decodificación de vídeo, una región actual puede indicar una unidad de codificación actual que es una de las unidades de codificación obtenidas dividiendo una instantánea actual de un vídeo para codificar el vídeo.

45 La unidad 23 de extracción del aparato 20 de decodificación de vídeo puede extraer, desde una secuencia de bits analizada, datos codificados de una instantánea que se codifican de acuerdo con las unidades de codificación en la estructura de árbol que se incluyen en cada una de las unidades de codificación máxima, y puede extraer información acerca de profundidades codificadas y modos de codificación para cada una de las unidades de codificación. También, la unidad 23 de extracción del aparato 20 de decodificación de vídeo puede extraer información de división máxima de unidad de transformación que se establece para cada instantánea, cada corte, cada unidad de codificación máxima, o cada unidad de codificación.

50 La unidad 25 de decodificación del aparato 20 de decodificación de vídeo puede realizar la transformación inversa en cada una de las unidades de codificación máxima usando las unidades de transformación en la estructura de árbol. En particular, la unidad 25 de decodificación puede realizar la transformación inversa usando unidades de transformación que se determinan de entre las unidades de transformación en la estructura de árbol variable que están configuradas basándose en la información de división máxima de unidad de transformación, puede decodificar los datos codificados en cada una de las unidades de codificación, y por lo tanto puede restaurar una instantánea.



La Figura 15 es un diagrama para describir un concepto de unidades de codificación de acuerdo con una realización de la presente invención.

Un tamaño de una unidad de codificación puede expresarse en anchura x altura, y puede ser 64x64, 32x32, 16x16 y 8x8, aunque se entiende que otra realización no está limitada a lo mismo. Una unidad de codificación de 64x64 puede dividirse en particiones de 64x64, 64x32, 32x64 o 32x32, una unidad de codificación de 32x32 puede dividirse en particiones de 32x32, 32x16, 16x32 o 16x16, una unidad de codificación de 16x16 puede dividirse en particiones de 16x16, 16x8, 8x16 u 8x8, y una unidad de codificación de 8x8 puede dividirse en particiones de 8x8, 8x4, 4x8 o 4x4.

Haciendo referencia a la Figura 15, los primeros datos 310 de vídeo tienen una resolución de 1920x1080, un tamaño máximo de una unidad de codificación de 64, y una profundidad máxima de 2. Los segundos datos 320 de vídeo tienen una resolución de 1920x1080, un tamaño máximo de una unidad de codificación de 64, y una profundidad máxima de 3. Los terceros datos 330 de vídeo tienen una resolución de 352x288, un tamaño máximo de una unidad de codificación de 16, y una profundidad máxima de 1. La profundidad máxima mostrada en la Figura 15 indica un número total de divisiones desde una unidad de codificación máxima a una unidad de decodificación mínima.

Si una resolución es alta o una cantidad de datos es grande, un tamaño máximo de una unidad de codificación puede ser grande para no aumentar únicamente la eficacia de codificación sino también para reflejar de manera precisa las características de una imagen. Por consiguiente, el tamaño máximo de las unidades de codificación de los primeros y segundos datos 310 de vídeo y 320 que tienen una resolución superior que los terceros datos 330 de vídeo puede ser 64.

Puesto que la profundidad máxima de los primeros datos 310 de vídeo es 2, las unidades 315 de codificación de los primeros datos 310 de vídeo pueden incluir una unidad de codificación máxima que tiene un tamaño de eje largo de 64, y unidades de codificación que tienen tamaños de eje largo de 32 y 16 puesto que las profundidades se hacen profundas a dos niveles dividiendo la unidad de codificación máxima dos veces. Mientras tanto, puesto que la profundidad máxima de los terceros datos 330 de vídeo es 1, las unidades 335 de codificación de los terceros datos 330 de vídeo pueden incluir una unidad de codificación máxima que tiene un tamaño de eje largo de 16, y las unidades de codificación que tienen un tamaño de eje largo de 8 puesto que las profundidades se hacen profundas a un nivel dividiendo la unidad de codificación máxima una vez.

Puesto que la profundidad máxima de los segundos datos 320 de vídeo es 3, las unidades 325 de codificación de los segundos datos 320 de vídeo pueden incluir una unidad de codificación máxima que tienen un tamaño de eje largo de 64, y las unidades de codificación que tienen tamaños de eje largo de 32, 16 y 8 puesto que las profundidades se hacen profundas a 3 niveles dividiendo la unidad de codificación máxima tres veces. A medida que una profundidad se hace profunda (es decir, aumenta), la información detallada puede expresarse de manera precisa.

La Figura 16 es un diagrama de bloques de un codificador 400 de imagen basado en unidades de codificación, de acuerdo con una realización de la presente invención.

Haciendo referencia a la Figura 16, el codificador 400 de imagen realiza operaciones del determinador 120 de unidad de codificación del aparato 100 de codificación de vídeo para codificar datos de imagen. Por ejemplo, un intra predictor 410 realiza intra predicción en unidades de codificación en un intra modo, de entre un fotograma 405 actual, y un estimador 420 de movimiento y un compensador 425 de movimiento realiza inter estimación y compensación de movimiento, respectivamente, en unidades de codificación en un inter modo de entre el fotograma 405 actual usando el fotograma 405 actual y un fotograma 495 de referencia.

Los datos emitidos desde el intra predictor 410, el estimador 420 de movimiento y el compensador 425 de movimiento se emiten como un coeficiente de transformación cuantificado a través de un transformador 430 y un cuantificador 440. El coeficiente de transformación cuantificado se restaura como datos en un dominio espacial a través de un cuantificador 460 inverso y un transformador 470 inverso. Los datos restaurados en el dominio espacial se emiten como el fotograma 495 de referencia después de procesarse posteriormente a través de una unidad 480 de desbloqueo y una unidad 490 de filtración de bucle. El coeficiente de transformación cuantificado puede emitirse como una secuencia de bits 455 a través de un codificador 450 por entropía.

Para que el codificador 400 de imagen se aplique en el aparato 100 de codificación de vídeo, los elementos del codificador 400 de imagen, es decir, el intra predictor 410, el estimador 420 de movimiento, el compensador 425 de movimiento, el transformador 430, el cuantificador 440, el codificador 450 por entropía, el cuantificador 460 inverso, el transformador 470 inverso, la unidad 480 de desbloqueo, y la unidad 490 de filtración de bucle, realizan las operaciones basándose en cada unidad de codificación de entre unidades de codificación que tienen una estructura de árbol mientras se considera la profundidad máxima de cada unidad de codificación máxima.

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

árbol.

La Figura 17 es un diagrama de bloques de un decodificador 500 de imagen basado en unidades de codificación, de acuerdo con una realización de la presente invención.

5 Haciendo referencia a la Figura 17, un analizador 510 analiza datos de imagen codificados para decodificarse e información acerca de la codificación usada para la decodificación desde una secuencia de bits 505. Los datos de imagen codificados se emiten como datos cuantificados invertidos a través de un decodificador 520 por entropía y un cuantificador 530 inverso, y los datos cuantificados invertidos se restauran a datos de imagen en un dominio espacial a través de un transformador 540 inverso.

10 Un intra predictor 550 realiza intra predicción en unidades de codificación en un intra modo con respecto a los datos de imagen en el dominio espacial, y un compensador 560 de movimiento realiza compensación de movimiento en unidades de codificación en un inter modo usando un fotograma 585 de referencia.

15 Los datos de imagen en el dominio espacial, que se pasan a través del intra predictor 550 y del compensador 560 de movimiento, pueden emitirse como un fotograma 595 restaurado después de procesarse posteriormente a través de una unidad 570 de desbloqueo y una unidad 580 de filtración de bucle. También, los datos de imagen que se procesan posteriormente a través de la unidad 570 de desbloqueo y la unidad 580 de filtrado en bucle pueden emitirse como el fotograma 585 de referencia.

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

20 Para que el decodificador 500 de imagen se aplique en el aparato 200 de decodificación de vídeo, los elementos del decodificador 500 de imagen, es decir, 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, realizan operaciones basándose en unidades de codificación que tienen una estructura de árbol para cada unidad de codificación máxima.

25 Específicamente, la intra predicción 550 y el compensador 560 de movimiento necesitan determinar particiones y un modo de predicción para cada una de las unidades de codificación que tienen una estructura de árbol, y el transformador 540 inverso necesita determinar un tamaño de una unidad de transformación para cada unidad de codificación.

La Figura 18 es un diagrama que ilustra unidades de codificación más profundas de acuerdo con las profundidades, y particiones, de acuerdo con una realización de la presente invención.

30 Un aparato 100 de codificación de vídeo y un aparato 200 de decodificación de vídeo usan unidades de codificación jerárquica para considerar características de una imagen. Una altura máxima, una anchura máxima y una profundidad máxima de unidades de codificación pueden determinarse de manera adaptativa de acuerdo con las características de la imagen, o pueden establecerse de manera diferente por un usuario. Los tamaños de unidades de codificación más profundas de acuerdo con las profundidades pueden determinarse de acuerdo con un tamaño máximo predeterminado de la unidad de codificación.

35 En una estructura 600 jerárquica de unidades de codificación, la altura máxima y la anchura máxima de las unidades de codificación son cada una 64 y la profundidad máxima es 4. En este punto, la profundidad máxima indica un número total de divisiones desde una unidad de codificación máxima a una unidad de codificación mínima. Puesto que una profundidad se hace profunda (es decir, aumenta) a lo largo de un eje vertical de la estructura 600 jerárquica, se divide cada una de una altura y una anchura de las unidades de codificación más profundas. También, una unidad de predicción y las particiones, que son las bases para codificación de predicción de cada unidad de codificación más profunda, se muestran a lo largo de un eje horizontal de la estructura 600 jerárquica.

40 Por ejemplo, una primera unidad 610 de codificación es una unidad de codificación máxima en la estructura 600 jerárquica, en la que una profundidad de la misma es 0 y un tamaño, es decir, una altura por anchura, de la misma es 64x64. La profundidad se hace profunda a lo largo del eje vertical de manera que la estructura 600 jerárquica incluye una segunda unidad 620 de codificación que tiene un tamaño de 32x32 y una profundidad de 1, una tercera unidad 630 de codificación que tiene un tamaño de 16x16 y una profundidad de 2, y una cuarta unidad 640 de codificación que tiene un tamaño de 8x8 y una profundidad de 3. La cuarta unidad 640 de codificación que tiene el tamaño de 8x8 y la profundidad de 3 es una unidad de codificación mínima.

45 La unidad de predicción y las particiones de las unidades 610, 620, 630 y 640 de codificación se disponen a lo largo del eje horizontal de acuerdo con cada profundidad. En otras palabras, si la primera unidad 610 de codificación que tiene el tamaño de 64x64 y la profundidad de 0 es una unidad de predicción, la unidad de predicción puede dividirse en particiones incluidas en la primera unidad 610 de codificación, es decir una partición 610 que tiene un tamaño de 64x64, las particiones 612 que tienen un tamaño de 64x32, las particiones 614 que tienen el tamaño de 32x64 o las  
50 particiones 616 que tienen un tamaño de 32x32.

De manera similar, una unidad de predicción de la segunda unidad 620 de codificación que tiene el tamaño de 32x32 y la profundidad de 1 puede dividirse en particiones incluidas en la segunda unidad 620 de codificación, es decir una partición 620 que tiene un tamaño de 32x32, particiones 622 que tienen un tamaño de 32x16, particiones 624 que tienen un tamaño de 16x32, y particiones 626 que tienen un tamaño de 16x16.

5 De manera similar, una unidad de predicción de la tercera unidad 630 de codificación que tiene el tamaño de 16x16 y la profundidad de 2 puede dividirse en particiones incluidas en la tercera unidad 630 de codificación, es decir una partición que tienen un tamaño de 16x16 incluido en la tercera unidad 630 de codificación, particiones 632 que tienen un tamaño de 16x8, particiones 634 que tienen un tamaño de 8x16, y particiones 636 que tienen un tamaño de 8x8.

10 De manera similar, una unidad de predicción de la cuarta unidad 640 de codificación que tiene el tamaño de 8x8 y la profundidad de 3 puede dividirse en particiones incluidas en la cuarta unidad 640 de codificación, es decir una partición que tienen un tamaño de 8x8 incluida en la cuarta unidad 640 de codificación, particiones 642 que tienen un tamaño de 8x4, particiones 644 que tienen un tamaño de 4x8, y particiones 646 que tienen un tamaño de 4x4.

15 Para determinar la al menos una profundidad codificada de las unidades de codificación de la unidad 610 de codificación máxima, el determinador 120 de unidad de codificación del aparato 100 de codificación de vídeo se requiere que realice la codificación para unidades de codificación que corresponden a cada profundidad incluida en la unidad 610 de codificación máxima.

20 Un número total de unidades de codificación más profundas de acuerdo con las profundidades que incluyen datos en el mismo intervalo y el mismo tamaño aumenta a medida que la profundidad se hace más profunda. Por ejemplo, se requieren cuatro unidades de codificación que corresponden a una profundidad de 2 para cubrir datos que se incluyen en una unidad de codificación que corresponde a una profundidad de 1. Por consiguiente, para comparar los resultados de codificación de los mismos datos de acuerdo con las profundidades, se codifica cada una de la unidad de codificación que corresponde a la profundidad de 1 y cuatro unidades de codificación que corresponden a la profundidad de 2.

25 Para realizar la codificación para una profundidad actual de entre las profundidades, puede seleccionarse un error de codificación mínimo para la profundidad actual realizando la codificación para cada unidad de predicción en las unidades de codificación que corresponden a la profundidad actual, a lo largo del eje horizontal de la estructura 600 jerárquica. Como alternativa, puede buscarse el error de codificación mínimo comparando los errores de codificación mínimos de acuerdo con las profundidades, realizando la codificación para cada profundidad a medida que la profundidad se hace más profunda a lo largo del eje vertical de la estructura 600 jerárquica. Una profundidad y una partición que tienen el error de codificación mínimo en la primera unidad 610 de codificación pueden seleccionarse como la profundidad codificada y un tipo de partición de la primera unidad 610 de codificación.

La Figura 19 es un diagrama para describir una relación entre una unidad 710 de codificación y las unidades 720 de transformación, de acuerdo con una realización de la presente invención.

35 El aparato 100 de codificación de vídeo y un aparato 200 de decodificación de vídeo codifican y decodifican, respectivamente, una imagen de acuerdo con unidades de codificación que tienen tamaños menores que o iguales a una unidad de codificación máxima para cada unidad de codificación máxima. Los tamaños de unidades de transformación para transformación durante la codificación pueden seleccionarse basándose en unidades de datos que ya no son mayores que una unidad de codificación correspondiente.

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

45 También, los datos de la unidad 710 de codificación que tienen el tamaño de 64x64 pueden codificarse realizando la transformación en cada una de las unidades de transformación que tienen el tamaño de 32x32, 16x16, 8x8 y 4x4, que son menores que 64x64, y a continuación puede seleccionarse una unidad de transformación que tiene los mínimos errores de codificación.

La Figura 20 es un diagrama para describir información de codificación de unidades de codificación que corresponden a una profundidad codificada, de acuerdo con una realización de la presente invención.

50 La unidad 130 de salida de un aparato 100 de codificación de vídeo puede codificar y transmitir la primera información 800 acerca de un tipo de partición, segunda información 810 acerca de un modo de predicción, y tercera información 820 acerca de un tamaño de una unidad de transformación para cada unidad de codificación que corresponde a una profundidad codificada, como información acerca de un modo de codificación.

55 La primera información 800 indica información acerca de una forma de una partición obtenida dividiendo una unidad de predicción de una unidad de codificación actual, en la que la partición es una unidad de datos para codificación de predicción de la unidad de codificación actual. Por ejemplo, una unidad de codificación actual CU\_0 que tiene un tamaño de 2Nx2N puede dividirse en una cualquiera de una partición 802 que tiene un tamaño de 2Nx2N, una partición 804 que tiene un tamaño de 2NxN, una partición 806 que tiene un tamaño de Nx2N, y una partición 808

que tiene un tamaño de  $N \times N$ . En este punto, la primera información 800 acerca de un tipo de partición se establece para indicar una de la partición 804 que tiene un tamaño de  $2N \times N$ , la partición 806 que tiene un tamaño de  $N \times 2N$ , y la partición 808 que tiene un tamaño de  $N \times N$ .

5 La segunda información 810 indica un modo de predicción de cada partición. Por ejemplo, la segunda información 810 puede indicar un modo de codificación de predicción realizado en una partición indicada por la primera información 800, es decir, un intra modo 812, un inter modo 814 o un modo de salto 816.

10 La tercera información 820 indica una unidad de transformación para que esté basada cuando se realiza la transformación en una unidad de codificación actual. Por ejemplo, la unidad de transformación puede ser una primera unidad 822 de transformación intra, una segunda unidad 824 de transformación intra, una primera unidad 826 de transformación inter, o una segunda unidad 828 de transformación intra.

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

15 La Figura 21 es un diagrama de unidades de codificación más profundas de acuerdo con las profundidades, de acuerdo con una realización de la presente invención.

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

20 Una unidad 910 de predicción para codificación de predicción de una unidad 900 de codificación que tiene una profundidad de 0 y un tamaño de  $2N_0 \times 2N_0$  puede incluir particiones de un tipo 912 de partición que tiene un tamaño de  $2N_0 \times 2N_0$ , un tipo 914 de partición que tiene un tamaño de  $2N_0 \times N_0$ , un tipo 916 de partición que tiene un tamaño de  $N_0 \times 2N_0$ , y un tipo 918 de partición que tiene un tamaño de  $N_0 \times N_0$ . La Figura 9 únicamente ilustra los tipos 912 a 918 de partición que se obtienen dividiendo simétricamente la unidad 910 de predicción, pero se entiende que un tipo de partición no está limitado al mismo en otra realización. Por ejemplo, las particiones de la  
25 unidad 910 de predicción pueden incluir particiones asimétricas, particiones que tienen un tamaño predeterminado y particiones que tienen una forma geométrica.

30 La codificación de predicción se realiza de manera repetitiva en una partición que tiene un tamaño de  $2N_0 \times 2N_0$ , dos particiones que tienen un tamaño de  $2N_0 \times N_0$ , dos particiones que tienen un tamaño de  $N_0 \times 2N_0$ , y cuatro particiones que tienen un tamaño de  $N_0 \times N_0$ , de acuerdo con cada tipo de partición. La codificación por predicción en un intra modo y un inter modo puede realizarse en las particiones que tienen los tamaños de  $2N_0 \times 2N_0$ ,  $N_0 \times 2N_0$ ,  $2N_0 \times N_0$  y  $N_0 \times N_0$ . La codificación por predicción en un modo de salto se realiza únicamente en la partición que tiene el tamaño de  $2N_0 \times 2N_0$ .

35 Los errores de codificación que incluyen la codificación por predicción en los tipos 912 a 918 de particiones se comparan, y se determina el error de codificación mínimo entre los tipos de partición. Si un error de codificación es el más pequeño en uno de los tipos 912 a 916 de partición, la unidad 910 de predicción pueden no dividirse en una profundidad inferior.

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

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

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

50 Cuando una profundidad máxima es  $d$ , la unidad de codificación de acuerdo con cada profundidad puede realizarse hasta cuando una profundidad se hace  $d-1$ , y la información de división puede codificarse hasta cuando una profundidad es una de 0 a  $d-2$ . Por ejemplo, cuando se realiza codificación hasta cuando la profundidad es  $d-1$  después de que una unidad de codificación que corresponde a una profundidad de  $d-2$  se divide en la operación 970, una unidad 990 de predicción para codificación de predicción de una unidad 980 de codificación que tiene una profundidad de  $d-1$  y un tamaño de  $2N_{(d-1)} \times 2N_{(d-1)}$  puede incluir particiones de un tipo 992 de partición que  
55 tienen un tamaño de  $2N_{(d-1)} \times 2N_{(d-1)}$ , un tipo 994 de partición que tiene un tamaño de  $2N_{(d-1)} \times N_{(d-1)}$ , un tipo 996 de partición que tiene un tamaño de  $N_{(d-1)} \times 2N_{(d-1)}$ , y un tipo 998 de partición que tiene un tamaño de  $N_{(d-1)} \times N_{(d-1)}$ .

1)xN<sub>(d-1)</sub>.

La codificación de predicción puede realizarse de manera repetitiva en una partición que tiene un tamaño de 2N<sub>(d-1)</sub>x2N<sub>(d-1)</sub>, dos particiones que tienen un tamaño de 2N<sub>(d-1)</sub>xN<sub>(d-1)</sub>, dos particiones que tienen un tamaño de N<sub>(d-1)</sub>x2N<sub>(d-1)</sub>, cuatro particiones que tienen un tamaño de N<sub>(d-1)</sub>xN<sub>(d-1)</sub> de entre los tipos 992 a 998 de partición para buscar un tipo de partición que tiene un error de codificación mínimo.

Incluso cuando el tipo 998 de partición tiene el error de codificación mínimo, puesto que una profundidad máxima es d, una unidad de codificación CU<sub>(d-1)</sub> que tiene una profundidad de d-1 ya no se divide a una profundidad inferior, y una profundidad codificada para las unidades de codificación de una unidad 900 de codificación máxima actual se determina para que sea d-1 y un tipo de partición de la unidad 900 de codificación máxima actual puede determinarse para que sea N<sub>(d-1)</sub>xN<sub>(d-1)</sub>. También, puesto que la profundidad máxima es d y una unidad 980 de codificación mínima que tiene una profundidad más inferior de d-1 ya no se divide a una profundidad inferior, no se establece la información de división para la unidad 980 de codificación mínima.

Una unidad 999 de datos puede considerarse una unidad mínima para la unidad de codificación máxima actual. Una unidad mínima puede ser una unidad de datos rectangular obtenida dividiendo una unidad 980 de codificación mínima en 4. Realizando la codificación de manera repetitiva, un aparato 100 de codificación de vídeo puede seleccionar una profundidad que tiene el error de codificación mínimo comparando errores de codificación de acuerdo con las profundidades de la unidad 900 de codificación para determinar una profundidad codificada, y establecer un correspondiente tipo de partición y un modo de predicción como un modo de codificación de la profundidad codificada.

Como tal, los errores de codificación mínimos de acuerdo con las profundidades se comparan en todas las profundidades de 1 a d, y una profundidad que tiene los mínimos errores de codificación puede determinarse como una profundidad codificada. Al menos uno de la profundidad codificada, el tipo de partición de la unidad de predicción, y el modo de predicción puede codificarse y transmitirse como información acerca de un modo de codificación. También, puesto que una unidad de codificación se divide desde una profundidad de 0 a una profundidad codificada, únicamente la información de división de la profundidad codificada se establece a 0, y la información de división de las profundidades que excluyen la profundidad codificada se establece a 1.

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

Las Figuras 22 a 24 son diagramas para describir una relación entre las unidades 1010 de codificación, unidades 1060 de predicción y unidades 1070 de transformación, de acuerdo con una realización de la presente invención.

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

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

En las unidades 1060 de predicción, algunas unidades 1014, 1016, 1022, 1032, 1048, 1050, 1052 y 1054 de codificación se obtienen dividiendo las unidades de codificación de las unidades 1010 de codificación. Por ejemplo, los tipos de partición en las unidades 1014, 1022, 1050 y 1054 de codificación tienen un tamaño de 2NxN, los tipos de partición en las unidades 1016, 1048 y 1052 de codificación tienen un tamaño de Nx2N, y un tipo de partición de la unidad 1032 de codificación tiene un tamaño de NxN. Las unidades de predicción y particiones de las unidades 1010 de codificación son menores que o iguales a cada unidad de codificación.

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

5 Por consiguiente, la codificación se realiza de manera recursiva en cada una de las unidades de codificación que tienen una estructura jerárquica en cada región de una unidad de codificación máxima para determinar una unidad de codificación óptima, y por lo tanto pueden obtenerse unidades de codificación que tienen una estructura de árbol recursiva. La información de codificación puede incluir al menos una de información de división acerca de una unidad de codificación, información acerca de un tipo de partición, información acerca de un modo de predicción e información acerca de un tamaño de una unidad de transformación. La Tabla 1 muestra información de codificación ejemplar que puede establecerse por los aparatos 100 y 200 de codificación y decodificación de vídeo.

[Tabla 1]

Información de división 0 (Codificación en unidad de codificación que tiene tamaño de 2Nx2N y profundidad actual de d)				Información de división 1
Modo de predicción	Tipo de partición		Tamaño de unidad de transformación	
Intra	Tipo de partición simétrica	Tipo de partición asimétrica	Información de división 0 de unidad de transformación	Información de división 1 de unidad de transformación
Inter				
Salto (únicamente 2Nx2N)	2Nx2N 2NxN Nx2N NxN	2NxnU 2NxnD nLx2N nRx2N	2Nx2N	NxN (tipo simétrico)  N/2xn/2 (tipo asimétrico)
Codificar repetitivamente unidades de codificación que tienen profundidad inferior de d+1				

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

15 La información de división indica si una unidad de codificación actual se divide en unidades de codificación de una profundidad inferior. Si la información de división de una profundidad actual d es 0, una profundidad en la que una unidad de codificación actual ya no se divide en una profundidad inferior es una profundidad codificada, y por lo tanto la información acerca de un tipo de partición, modo de predicción y un tamaño de una unidad de transformación puede definirse para la profundidad codificada. Si la unidad de codificación actual se divide adicionalmente de acuerdo con la información de división, se realiza la codificación independientemente en cuatro unidades de codificación de división de una profundidad inferior.

20 Un modo de predicción puede ser uno de un intra modo, un inter modo y un modo de salto. El intra modo y el inter modo pueden definirse en todos los tipos de partición, y el modo de salto puede definirse únicamente en un tipo de partición que tiene un tamaño de 2Nx2N.

25 La información acerca del tipo de partición puede indicar tipos de partición simétrica que tienen tamaños de 2Nx2N, 2NxN, Nx2N y NxN, que se obtienen dividiendo simétricamente al menos una de una altura y una anchura de una unidad de predicción, y tipos de partición asimétrica que tienen tamaños de 2NxnU, 2NxnD, nLx2N y nRx2N, que se obtienen dividiendo asimétricamente al menos una de la altura y la anchura de la unidad de predicción. Los tipos de partición asimétrica que tienen los tamaños de 2NxnU y 2NxnD pueden obtenerse respectivamente dividiendo la altura de la unidad de predicción en 1:3 y 3:1, y los tipos de partición asimétrica que tienen los tamaños de nLx2N y nRx2N pueden obtenerse respectivamente dividiendo la anchura de la unidad de predicción en 1:3 y 3:1

30 El tamaño de la unidad de transformación puede establecerse para que sea dos tipos en el intra modo y dos tipos en el inter modo. Por ejemplo, si la información de división de la unidad de transformación es 0, el tamaño de la unidad de transformación puede ser 2Nx2N, que es el tamaño de la unidad de codificación actual. Si la información de división de la unidad de transformación es 1, las unidades de transformación pueden obtenerse dividiendo la unidad de codificación actual. También, si un tipo de partición de la unidad de codificación actual que tiene el tamaño de 2Nx2N es un tipo de partición simétrica, un tamaño de una unidad de transformación puede ser NxN, y si el tipo de partición de la unidad de codificación actual es un tipo de partición asimétrica, el tamaño de la unidad de transformación puede ser N/2xn/2.

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

- 5 Por consiguiente, se determina si las unidades de datos adyacentes están incluidas en la misma unidad de codificación que corresponde a la profundidad codificada comparando información de codificación de las unidades de datos adyacentes. También, una unidad de codificación correspondiente que corresponde a una profundidad codificada se determina usando información de codificación de una unidad de datos, y por lo tanto puede determinarse una distribución de las profundidades codificadas en una unidad de codificación máxima.
- 10 Por lo tanto, si una unidad de codificación actual se predice basándose en información de codificación de unidades de datos adyacentes, puede hacerse referencia directamente y usarse la información de codificación de unidades de datos en unidades de codificación más profundas adyacentes a la unidad de codificación actual.

- En otra realización, si una unidad de codificación actual se predice basándose en información de codificación de unidades de datos adyacentes, se buscan unidades de datos adyacentes a la unidad de codificación actual usando información codificada de las unidades de datos, y las unidades de codificación adyacentes buscadas pueden hacerse referencia para predecir la unidad de codificación actual.
- 15

La Figura 25 es un diagrama para describir una relación entre una unidad de codificación, una unidad de predicción o una partición, y una unidad de transformación, de acuerdo con información de modo de codificación de la Tabla 1 de acuerdo con una realización de la presente invención.

- 20 Una unidad 1300 de codificación máxima incluye las unidades 1302, 1304, 1306, 1312, 1314, 1316 y 1318 de codificación de profundidades codificadas. En este punto, puesto que la unidad 1318 de codificación es una unidad de codificación de una profundidad codificada, la información de división puede establecerse a 0. La información acerca de un tipo de partición de la unidad 1318 de codificación que tiene un tamaño de  $2N \times 2N$  puede establecerse para que sea un tipo 1322 de partición que tiene un tamaño de  $2N \times 2N$ , un tipo 1324 de partición que tiene un tamaño de  $2N \times N$ , un tipo 1326 de partición que tiene un tamaño de  $N \times 2N$ , un tipo 1328 de partición que tiene un tamaño de  $N \times N$ , un tipo 1332 de partición que tiene un tamaño de  $2N \times nU$ , un tipo 1334 de partición que tiene un tamaño de  $2N \times nD$ , un tipo 1336 de partición que tiene un tamaño de  $nL \times 2N$ , o un tipo 1338 de partición que tiene un tamaño de  $nR \times 2N$ .
- 25

- La bandera de tamaño TU de información de división de unidad de transformación puede ser uno de los índices de transformada, y un tamaño de la unidad de transformación que corresponde al índice de transformada puede cambiarse de acuerdo con un tipo de unidad de predicción o un tipo de partición de la unidad de codificación.
- 30

- Por ejemplo, cuando el tipo de partición se establece para que sea simétrica, es decir, el tipo 1322, 1324, 1326 o 1328 de partición, se establece una unidad 1342 de transformación que tiene un tamaño de  $2N \times 2N$  si la información de división de unidad de transformación es 0, y se establece una unidad 1344 de transformación que tiene un tamaño de  $N \times N$  si la información de división de unidad de transformación es 1.
- 35

Cuando el tipo de partición se establece para que sea asimétrica, es decir, el tipo 1332, 1334, 1336 o 1338 de partición, se establece una unidad 1352 de transformación que tiene un tamaño de  $2N \times 2N$  si la información de división de unidad de transformación es 0, y se establece una unidad 1354 de transformación que tiene un tamaño de  $N/2 \times n/2$  si la información de división de unidad de transformación es 1.

- 40 Haciendo referencia a la Figura 21, la bandera de tamaño TU de la información de división de unidad de transformación es una bandera que tiene un valor, 0 o 1, aunque se entiende que otra realización no está limitada a una bandera de 1 bit, y una unidad de transformación puede dividirse jerárquicamente teniendo una estructura de árbol mientras la información de división de unidad de transformación aumenta desde 0 en otra realización. La información de división de unidad de transformación puede usarse como una realización del índice de transformada.

- 45 En este caso, cuando se usa la información de división de unidad de transformación junto con un tamaño máximo de una unidad de transformación y un tamaño mínimo de la misma, puede representarse el tamaño de una unidad de transformación realmente usada. El aparato 100 de codificación de vídeo puede codificar información de tamaño de una unidad de transformación máxima, información de tamaño de una unidad de transformación mínima e información de división máxima de una unidad de transformación. La información de tamaño codificada de la unidad de transformación máxima, la información de tamaño de la unidad de transformación mínima codificada, y la información de división máxima de una unidad de transformación pueden insertarse en el SPS. El aparato 200 de decodificación de vídeo puede realizar decodificación de vídeo usando la información de tamaño de la unidad de transformación máxima, la información de tamaño de la unidad de transformación mínima, y la información de división máxima de una unidad de transformación.
- 50

- 55 Por ejemplo, cuando una unidad de codificación actual tiene un tamaño de  $64 \times 64$ , una unidad de transformación máxima tiene un tamaño de  $32 \times 32$ , y la información de división de unidad de transformación es 0, una unidad de transformación puede establecerse para que tenga un tamaño de  $32 \times 32$ . Cuando la unidad de codificación actual

5 tiene un tamaño de 64x64, la unidad de transformación máxima tiene un tamaño de 32x32, y la información de división de unidad de transformación es 1, la unidad de transformación puede establecerse para que tenga un tamaño de 16x16. Cuando la unidad de codificación actual tiene un tamaño de 64x64, la unidad de transformación máxima tiene un tamaño de 32x32, y la información de división de unidad de transformación es 2, la unidad de transformación puede establecerse para que tenga un tamaño de 8x8.

10 Por ejemplo, cuando la unidad de codificación actual tiene un tamaño de 32x32, una unidad de transformación mínima tiene un tamaño de 32x32, y la información de división de unidad de transformación es 0, la unidad de transformación puede establecerse para que tenga un tamaño de 32x32. Puesto que el tamaño de la unidad de transformación no puede ser menor que 32x32, no puede establecerse información de división de unidad de transformación.

Por ejemplo, cuando la unidad de codificación actual tiene un tamaño de 64x64 y la información de división máxima de una unidad de transformación es 1, la información de división de unidad de transformación puede ser 0 o 1, y otra información de división de unidad de transformación no puede establecerse.

15 Por consiguiente, cuando la información de división máxima de una unidad de transformación se define como 'MaxTransformSizeIndex', el tamaño de la unidad de transformación mínima se define como 'MinTransformSize', y el tamaño de una unidad de transformación cuando la información de división de unidad de transformación es 0 se define como 'RootTuSize', un tamaño de unidad de transformación mínima posible en una unidad de codificación actual, 'CurrMinTuSize', puede definirse como la siguiente ecuación (1) de relación:

20 
$$\text{CurrMinTuSize} = \max(\text{MinTransformSize}, \text{RootTuSize}/(2^{\text{MaxTransformSizeIndex}})) \dots (1)$$

25 En comparación con el tamaño de unidad de transformación mínima posible en una unidad de codificación actual, 'CurrMinTuSize', el tamaño de unidad de transformación cuando la bandera de tamaño TU es 0, 'RootTuSize', puede representar un tamaño de unidad de transformación máxima que puede emplearse por un sistema. En otras palabras, de acuerdo con la Ecuación (1), puesto que 'RootTuSize/(2^MaxTransformSizeIndex)' indica un tamaño de unidad de transformación obtenido dividiendo el tamaño de unidad de transformación cuando la información de división de unidad de transformación es 0, 'RootTuSize', en un número de veces que corresponde a la información de división máxima de una unidad de transformación, y 'MinTransformSize' indica un tamaño de unidad de transformación mínima, un tamaño menor de entre los dos tamaños puede ser el tamaño de unidad de transformación mínima posible en la unidad de codificación actual, 'CurrMinTuSize'.

30 El tamaño de la unidad de transformación máxima actual RootTuSize puede variar dependiendo del tipo de modo de predicción.

Por ejemplo, cuando un modo de predicción actual es un inter modo, RootTuSize puede determinarse de acuerdo con la ecuación (2) de relación a continuación, donde 'MaxTransformSize' indica un tamaño de unidad de transformación máxima y 'PUSize' indica un tamaño de unidad de predicción actual:

35 
$$\text{RootTuSize} = \min(\text{MaxTransformSize}, \text{PUSize}) \dots (2)$$

En otras palabras, cuando el modo de predicción actual es un inter modo, el tamaño de unidad de transformación cuando la información de división de unidad de transformación es 0, 'RootTuSize', puede establecerse para que sea un tamaño menor de entre el tamaño de la unidad de transformación máxima y el tamaño de la unidad de predicción actual.

40 Cuando un modo de predicción de una unidad de partición actual es un intra modo, 'RootTuSize' puede determinarse de acuerdo con la ecuación (3) de relación a continuación, donde 'PartitionSize' indica el tamaño de la unidad de partición actual:

$$\text{RootTuSize} = \min(\text{MaxTransformSize}, \text{PartitionSize}) \dots (3)$$

45 En otras palabras, cuando el modo de predicción actual es un intra modo, el tamaño de unidad de transformación cuando la información de división de unidad de transformación es 0, 'RootTuSize', puede establecerse para que sea un tamaño menor de entre el tamaño de la unidad de transformación máxima y el tamaño de la unidad de partición actual.

50 Sin embargo, debería observarse que un tamaño de unidad de transformación máxima actual 'RootTuSize' que varía de acuerdo con el modo de predicción de una unidad de partición es únicamente una realización, y que un factor para determinar el tamaño de la unidad de transformación máxima actual 'RootTuSize' no está limitado a la realización.

La Figura 26 es un diagrama de flujo que ilustra un procedimiento de codificación de vídeo usando una unidad de codificación en una estructura de árbol y una unidad de transformación en una estructura de árbol variable, de acuerdo con una realización de la presente invención.



En la operación 1210, una instantánea actual se divide en al menos una unidad de codificación máxima. Una profundidad máxima que indica un número total de posibles divisiones de división pueden establecerse con antelación.

5 En la operación 1220, al menos se codifica una región de división obtenida dividiendo una región de la unidad de codificación máxima de acuerdo con las profundidades, de modo que se determina una profundidad codificada para emitir un resultado de codificación final para cada al menos una región de división y se determina la unidad de codificación en la estructura de árbol y la unidad de transformación en la estructura de árbol variable.

10 La unidad de codificación máxima se divide espacialmente cada vez que la profundidad se hace profunda, y por lo tanto se divide en unidades de codificación de una profundidad inferior. Cada unidad de codificación puede dividirse en unidades de codificación de otra profundidad inferior dividiéndose espacialmente de manera independiente de unidades de codificación adyacentes. La codificación se realiza de manera repetitiva en cada unidad de codificación de acuerdo con las profundidades.

15 También, se determinan tipos de partición y unidades de transformación que tienen un error de codificación mínimo para cada unidad de codificación más profunda. Para determinar una profundidad codificada que tiene un error de codificación mínimo en cada unidad de codificación máxima, pueden medirse errores de codificación y compararse en todas las unidades de codificación más profundas de acuerdo con las profundidades.

20 Cuando se determina cada unidad de codificación, puede determinarse una unidad de transformación para transformación de la unidad de codificación. La unidad de transformación puede determinarse como una unidad de datos que minimiza un error producido transformando la unidad de codificación. Como resultado de la realización de la transformación basándose en niveles de acuerdo con las profundidades de transformación en una unidad de codificación actual, puede determinarse una unidad de transformación en una estructura de árbol que es independiente de otra unidad de transformación en una región adyacente y que forma una estructura jerárquica con unidades de transformación de acuerdo con las profundidades en la misma región.

25 También, un nivel de división máxima de la unidad de transformación puede establecerse previamente para cada unidad de codificación máxima o cada unidad de codificación. De acuerdo con un nivel de división máxima de la unidad de codificación actual, la transformación puede realizarse usando unidades de transformación en una estructura de árbol variable que incluye una unidad de transformación de base a través de una unidad de transformación mínima que se permite para la unidad de codificación actual. Las unidades de transformación con una profundidad de transformación que emite un resultado de codificación que tiene un error mínimo pueden determinarse desde las unidades de transformación en la estructura de árbol variable con respecto a la unidad de codificación actual, de modo que pueden determinarse las unidades de transformación en la estructura de árbol.

30 En la operación 1230, se emiten los datos de imagen que son un resultado final de la codificación de al menos una región de división de cada unidad de codificación máxima, con información codificada acerca de una profundidad codificada y un modo de codificación. La información codificada acerca del modo de codificación puede incluir información acerca de una profundidad codificada o información de división, información acerca de un tipo de partición de una unidad de predicción, y un modo de predicción, e información de estructura jerárquica de unidad de transformación. La información codificada acerca del modo de codificación puede transmitirse a un decodificador con los datos de imagen codificados.

40 La Figura 27 es un diagrama de flujo que ilustra un procedimiento de decodificación de vídeo usando una unidad de codificación en una estructura de árbol y una unidad de transformación en una estructura de árbol variable, de acuerdo con una realización de la presente invención.

En la operación 1310, se recibe y analiza una secuencia de bits de un vídeo codificado.

45 En la operación 1320, se obtienen los datos de imagen codificados de una instantánea actual asignados a una unidad de codificación máxima desde la secuencia de bits analizada, y se extrae la información acerca de una profundidad codificada y un modo de codificación de acuerdo con unidades de codificación máxima desde la secuencia de bits analizada. La profundidad codificada de cada unidad de codificación máxima es una profundidad que tiene un error de codificación mínimo en cada unidad de codificación máxima. Al codificar cada unidad de codificación máxima, los datos de imagen se codifican basándose en al menos una unidad de datos obtenida dividiendo jerárquicamente cada unidad de codificación máxima de acuerdo con las profundidades.

50 De acuerdo con la información acerca de la profundidad codificada y el modo de codificación, la unidad de codificación máxima puede dividirse en unidades de codificación que tienen una estructura de árbol. Cada una de las unidades de codificación que tiene la estructura de árbol se determina como una unidad de codificación que corresponde a una profundidad codificada, y se codifica de manera óptima para emitir el error de codificación mínimo. Por consiguiente, la eficacia de codificación y decodificación de una imagen puede mejorarse decodificando cada pieza de datos de imagen codificados en las unidades de codificación que tienen una estructura de árbol  
55 después de determinar al menos una profundidad codificada de acuerdo con las unidades de codificación.

5 También, de acuerdo con la información de estructura jerárquica de unidad de transformación extraída, la unidad de transformación en la estructura de árbol puede determinarse en la unidad de codificación. Por ejemplo, un número total de niveles desde una unidad de transformación de base a una unidad de transformación en un nivel más inferior que se permite para una unidad de codificación actual puede leerse basándose en la información de división máxima de unidad de transformación. Como alternativa, el número total de niveles desde la unidad de transformación de base a la unidad de transformación en el nivel más inferior puede determinarse basándose en un nivel de división máxima que se establece con antelación entre sistemas de codificación y decodificación.

10 La unidad de transformación de base puede determinarse de manera variable de acuerdo con un modo de codificación basándose en un esquema preestablecido. Por lo tanto, las unidades de transformación en una estructura de árbol variable con respecto a la unidad de codificación actual pueden determinarse basándose en la información de división máxima de unidad de transformación, y puede determinarse una unidad de transformación con una profundidad de transformación para usarse en transformación inversa con respecto a la unidad de codificación actual de entre las unidades de transformación en la estructura de árbol variable.

15 En la operación 1330, los datos de imagen codificados de cada unidad de codificación máxima se decodifican basándose en la información acerca de una profundidad codificada y un modo de codificación de acuerdo con unidades de codificación máxima. En este sentido, puede realizarse transformación inversa en la unidad de codificación actual usando la unidad de transformación que se determina de entre las unidades de transformación en la estructura de árbol variable mientras se realiza la decodificación en la unidad de codificación actual basándose en la información acerca de una profundidad codificada y un modo de codificación. A medida que se realiza la decodificación en cada unidad de codificación y cada unidad de codificación máxima, pueden restaurarse datos de imagen de una región espacial y a continuación puede restaurarse una instantánea y un vídeo que es una secuencia de instantáneas. El vídeo restaurado puede reproducirse por un aparato de reproducción, puede almacenarse en un medio de almacenamiento, o puede transmitirse mediante una red.

25 Las realizaciones de la presente invención pueden escribirse como programas informáticos y pueden implementarse en ordenadores digitales de uso general que ejecutan los programas usando un medio de grabación legible por ordenador. Además, una estructura de datos usada en las realizaciones de la presente invención puede escribirse en un medio de grabación legible por ordenador a través de diversos medios. Ejemplos del medio de grabación legible por ordenador incluye medio de almacenamiento magnético (por ejemplo, ROM, discos flexibles, discos duros, etc.), medio de grabación óptico (por ejemplo, CD-ROM o DVD), etc.

30

**REIVINDICACIONES**

1. Un procedimiento de decodificación de vídeo que comprende recibir (121) una secuencia de bits que comprende datos obtenidos codificando un vídeo, que comprende analizar (122) la secuencia de bits y extraer, desde la secuencia de bits, datos codificados de una información de estructura jerárquica de unidad de transformación de  
 5 unidad de codificación actual, que comprende información de división máxima para una unidad de transformación con un inter modo, información de división máxima para una unidad de transformación con un intra modo, información de tamaño máximo de una unidad de transformación e información de tamaño mínimo de una unidad de transformación con respecto al vídeo, información acerca de un modo de codificación con respecto a la unidad de  
 10 codificación, en el que la información de división máxima para una unidad de transformación con un inter modo es un número total de niveles permisibles de unidades de transformación con respecto a una unidad de codificación prevista de acuerdo con el inter modo;
- en el que la información de división máxima para una unidad de transformación con el intra modo es un número total de niveles permisibles de unidades de transformación con respecto a una unidad de codificación prevista de acuerdo con el intra modo;
- 15 en el que la información de división máxima para una unidad de transformación con el inter modo y la información de división máxima para una unidad de transformación con el intra modo se determinan de manera separada; en el que cuando la información acerca del modo de codificación indica que la unidad de codificación se predice de acuerdo con el inter modo, determinar un nivel de división máxima de una unidad de transformación usando la información de división máxima para la unidad de transformación con el inter modo;
- 20 realizar transformación inversa en la unidad de codificación actual prevista de acuerdo con el inter modo usando unidades de transformación en una estructura de árbol variable que se determinan de entre las unidades de transformación que se dividen jerárquicamente desde una unidad de transformación de base con respecto a la unidad de codificación actual y que se generan en base al nivel de división máxima determinado de una unidad de transformación.
- 25

FIG. 1

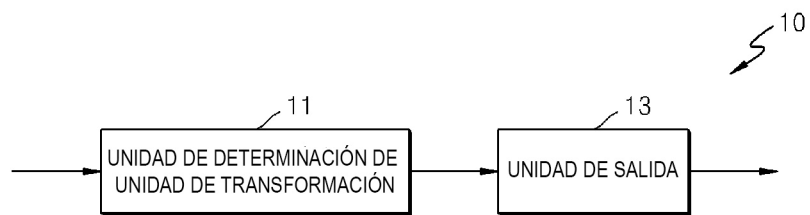
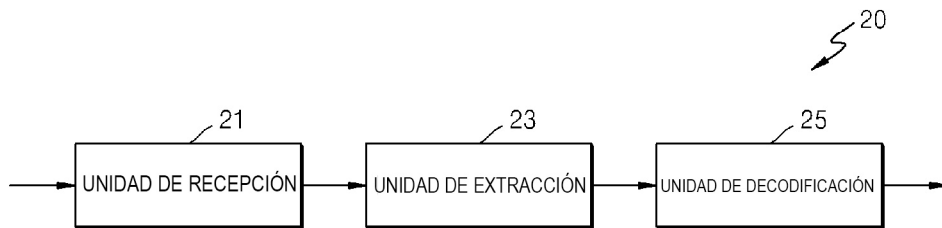


FIG. 2



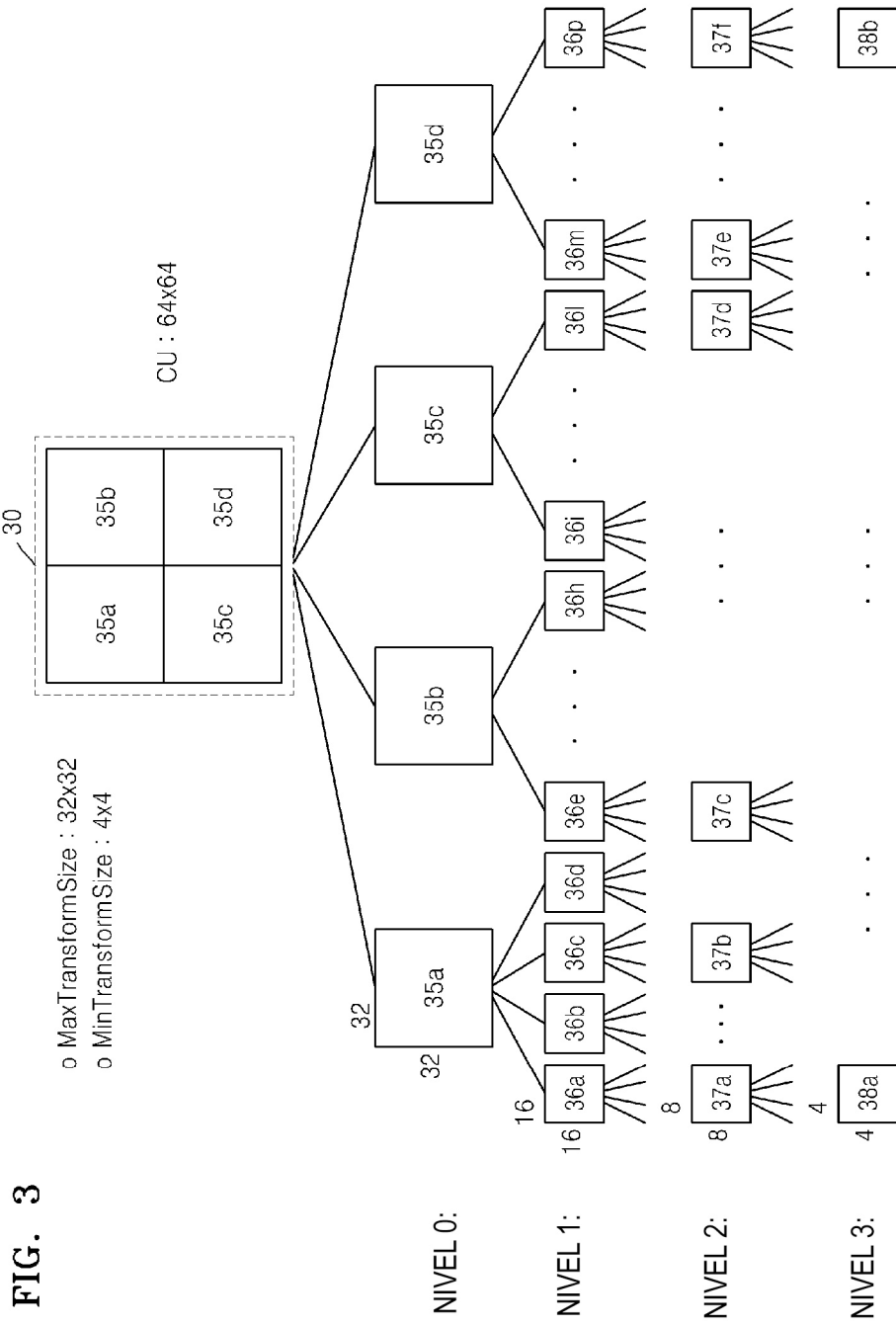


FIG. 4

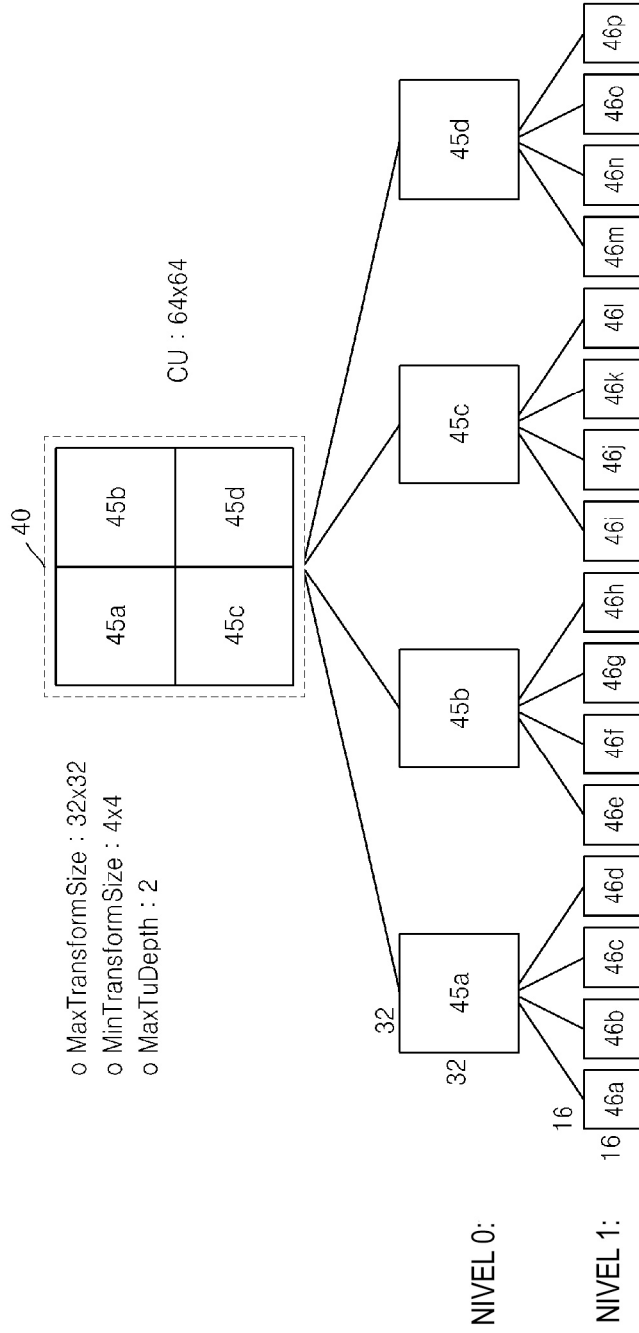


FIG. 5

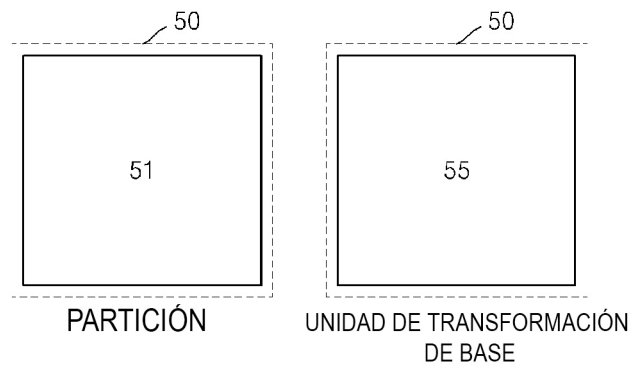


FIG. 6

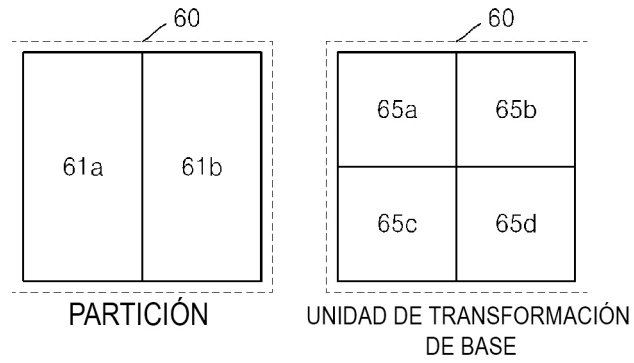
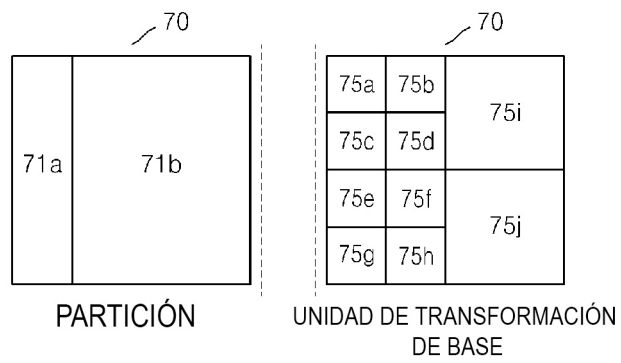
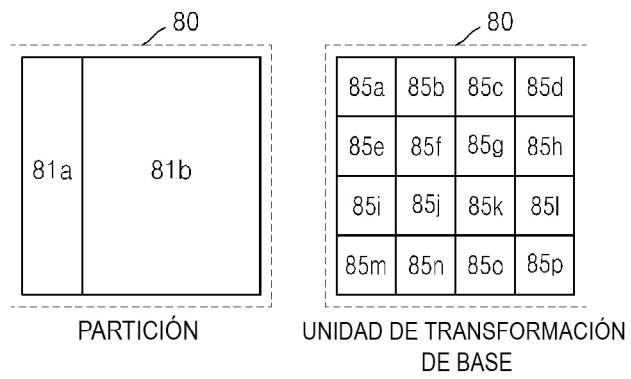


FIG. 7



**FIG. 8**



**FIG. 9**

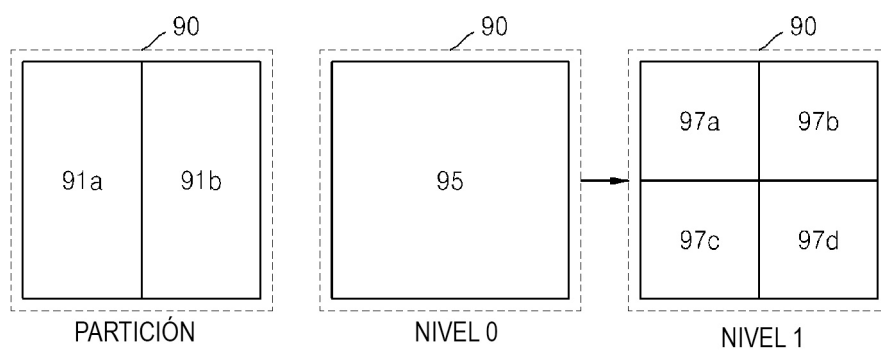




FIG. 10

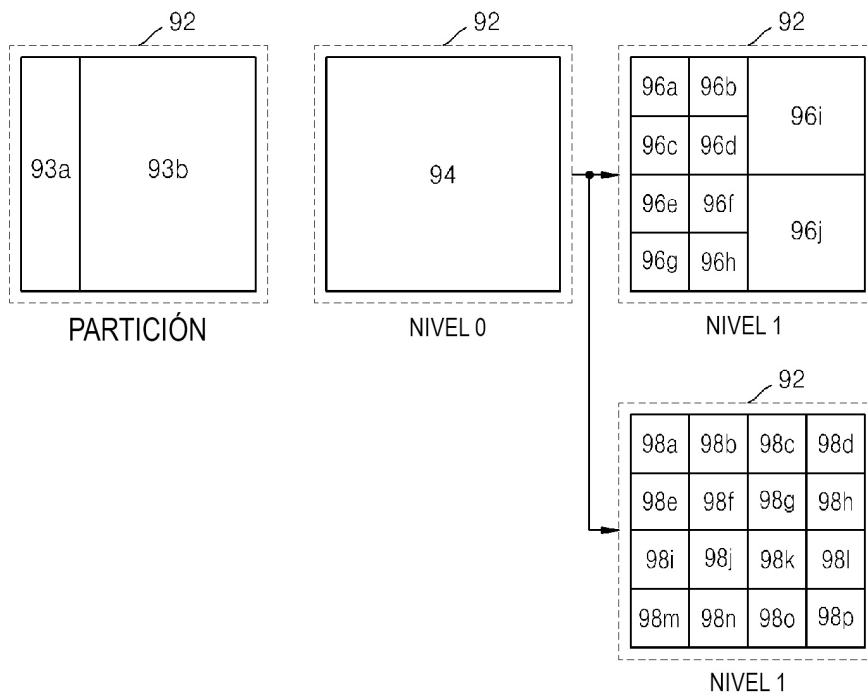


FIG. 11

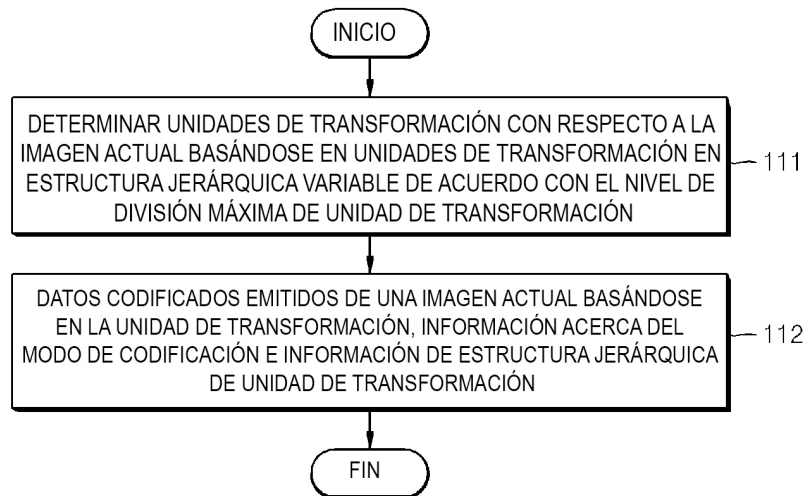


FIG. 12

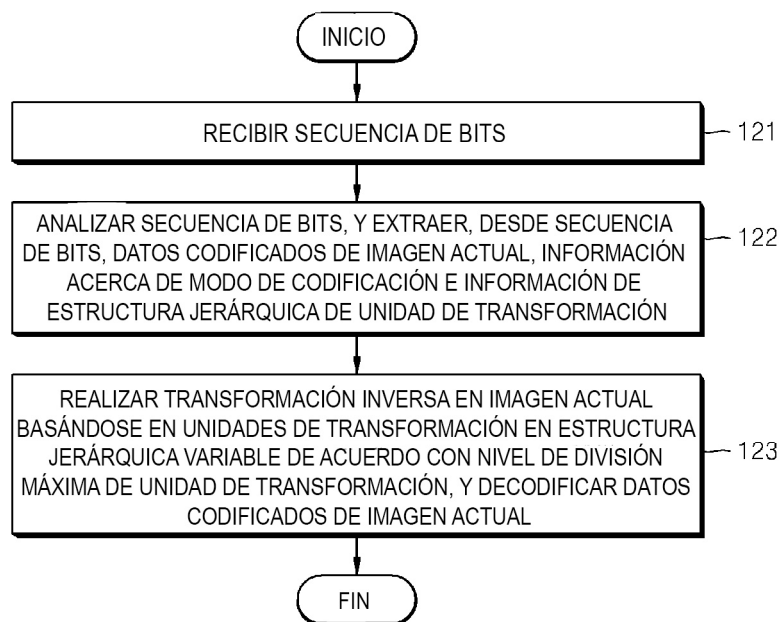


FIG. 13

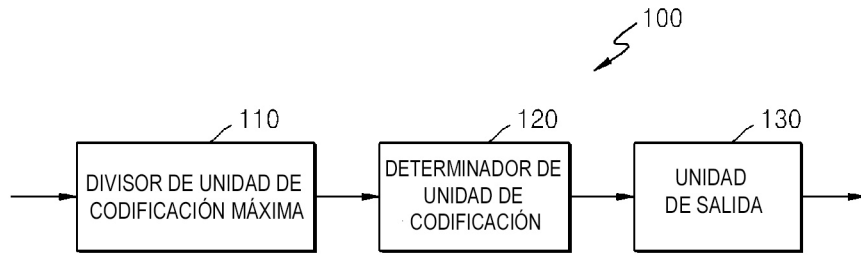


FIG. 14

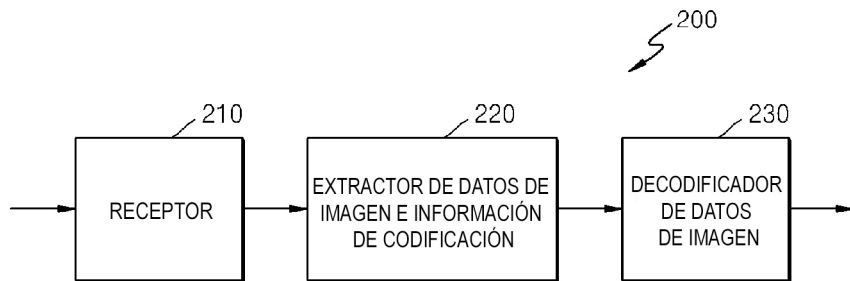


FIG. 15

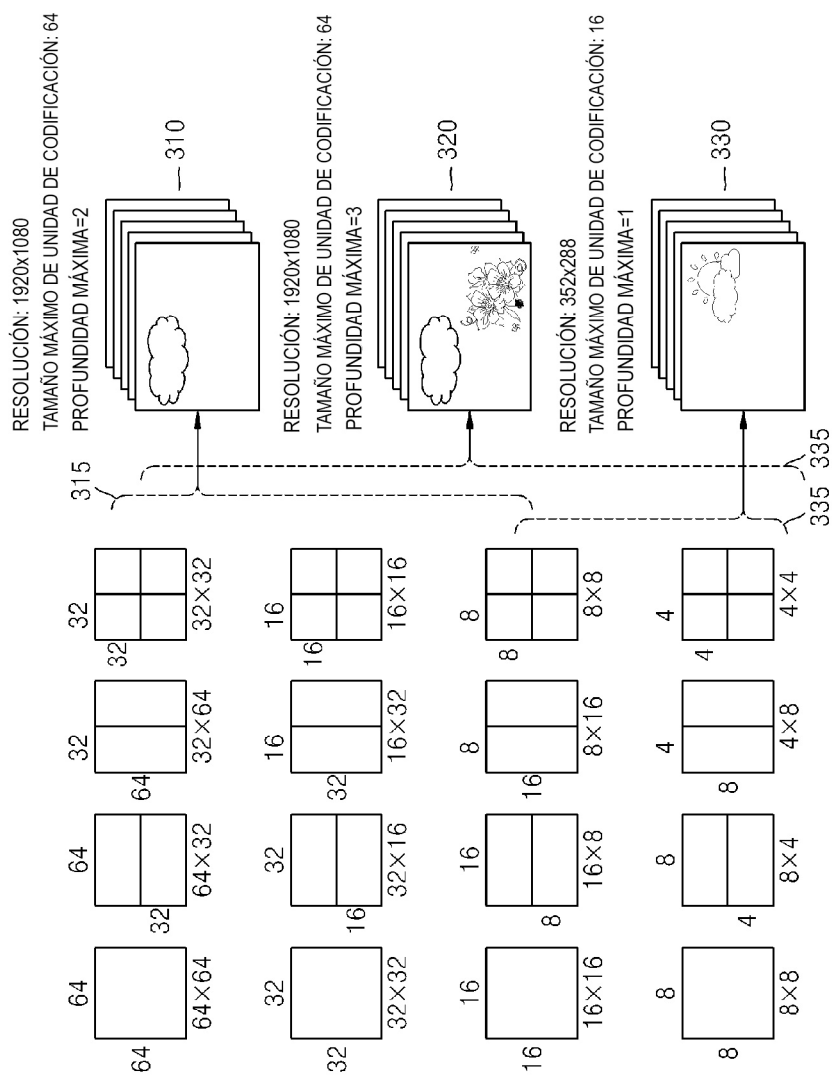


FIG. 16

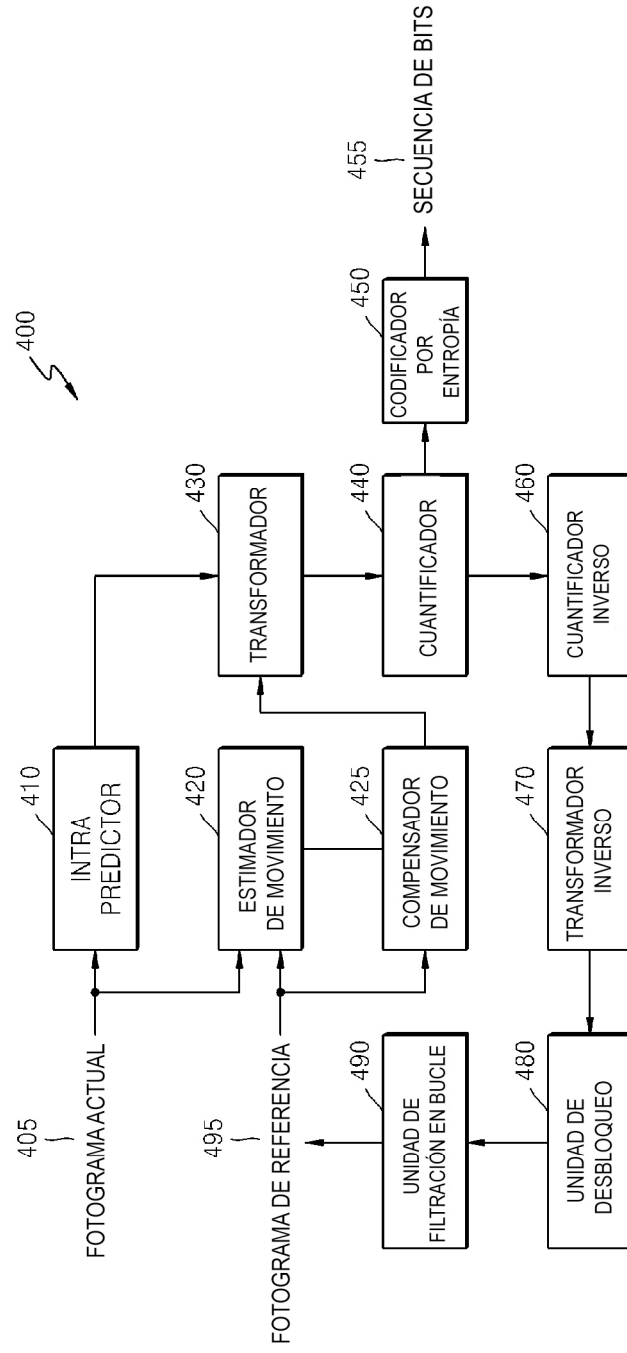


FIG. 17

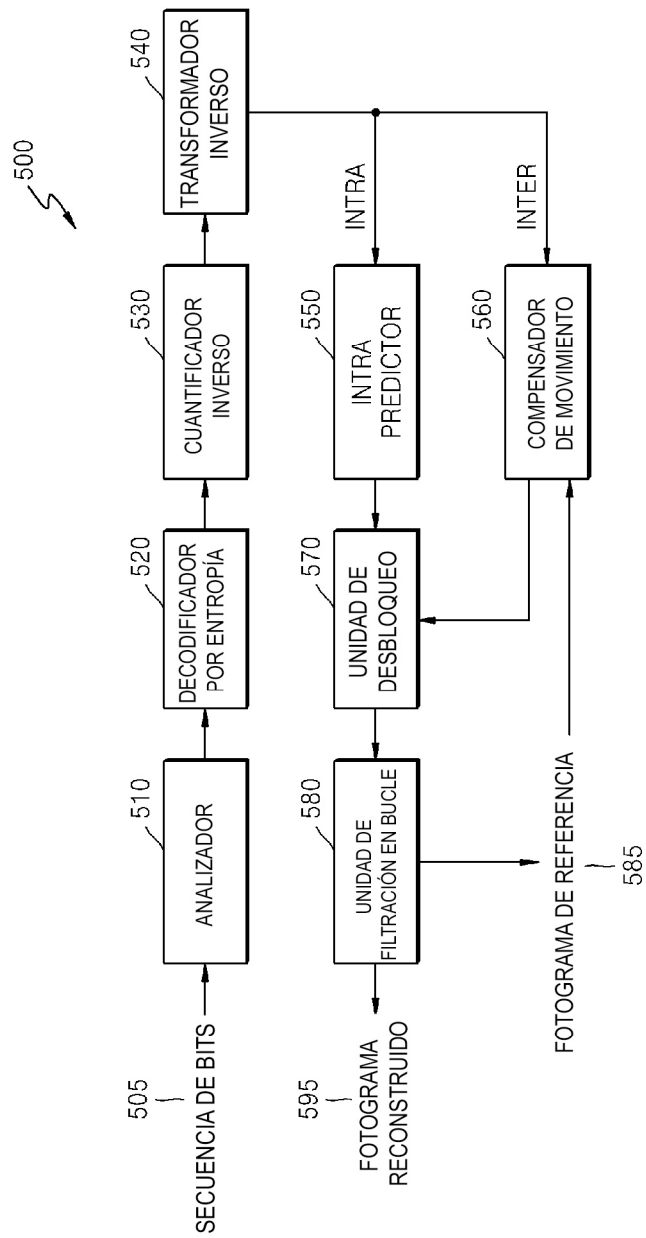
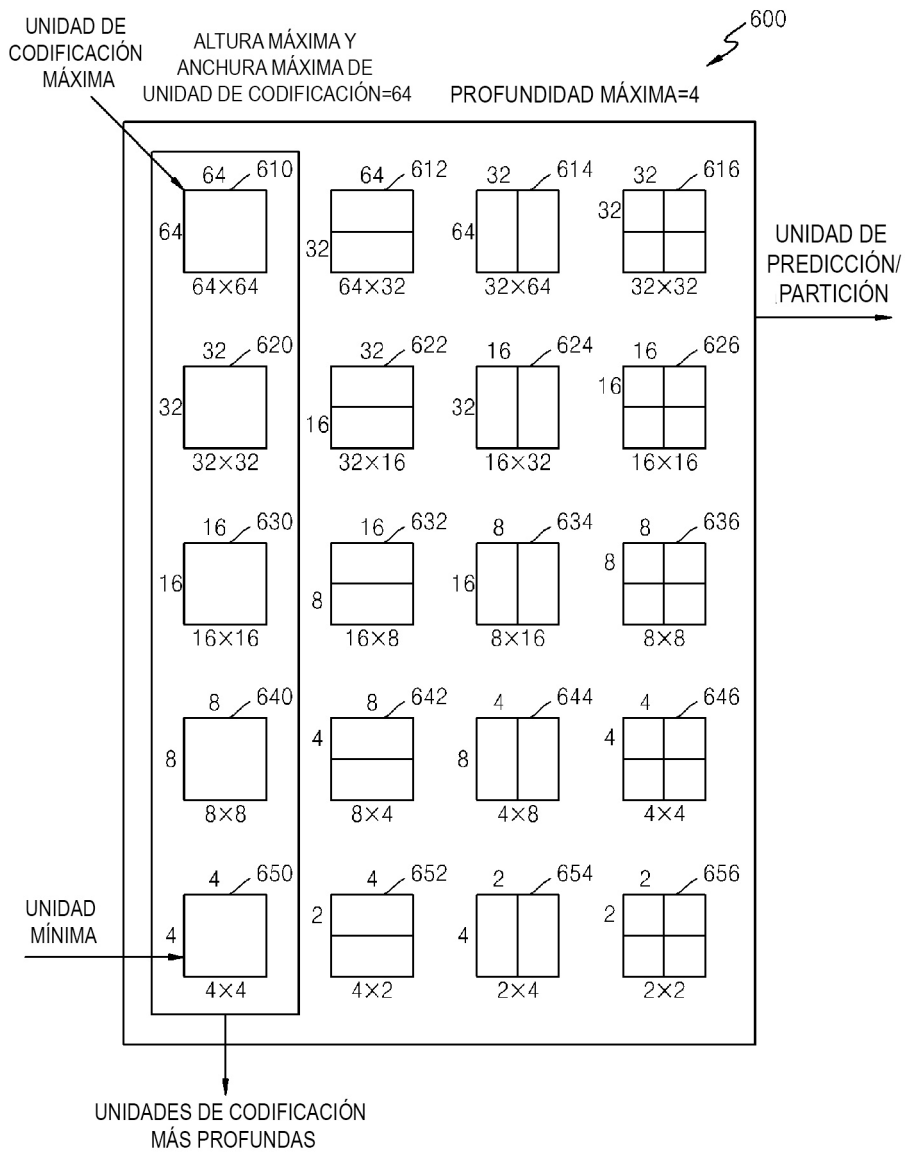
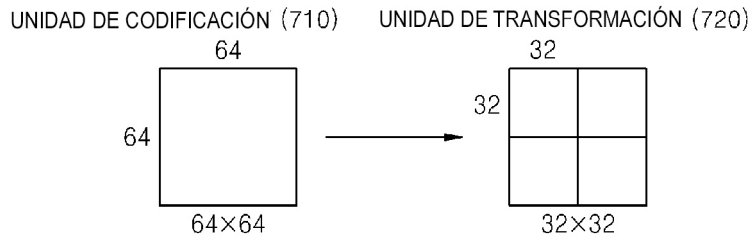


FIG. 18



**FIG. 19**



**FIG. 20**

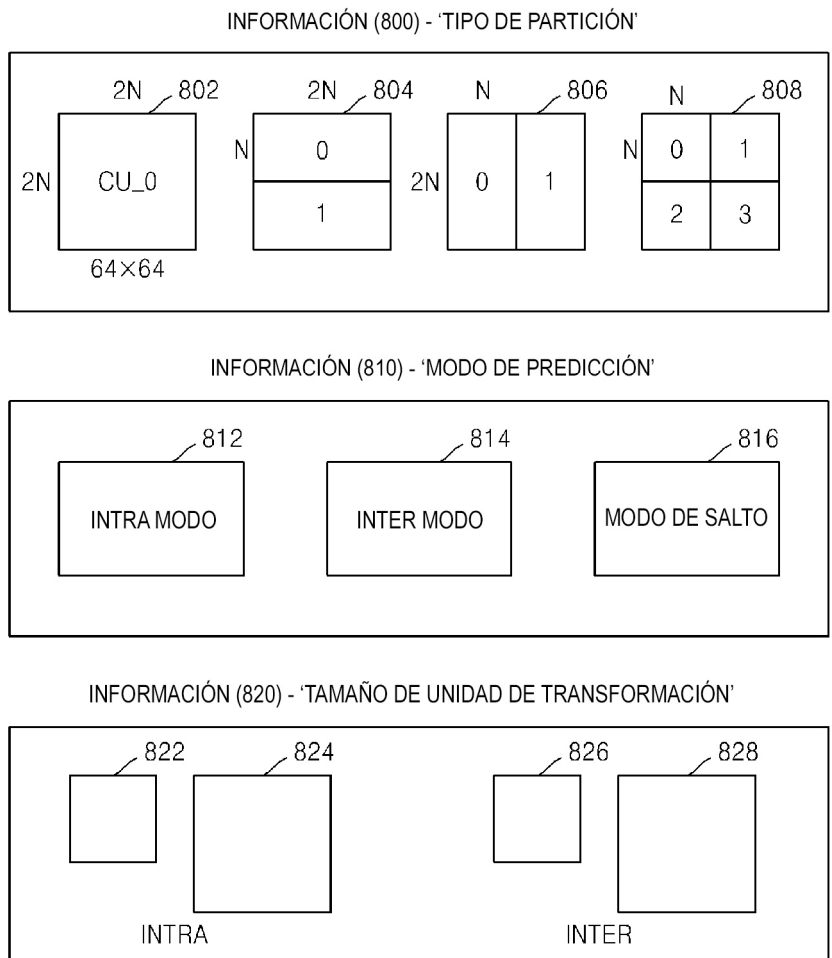




FIG. 21

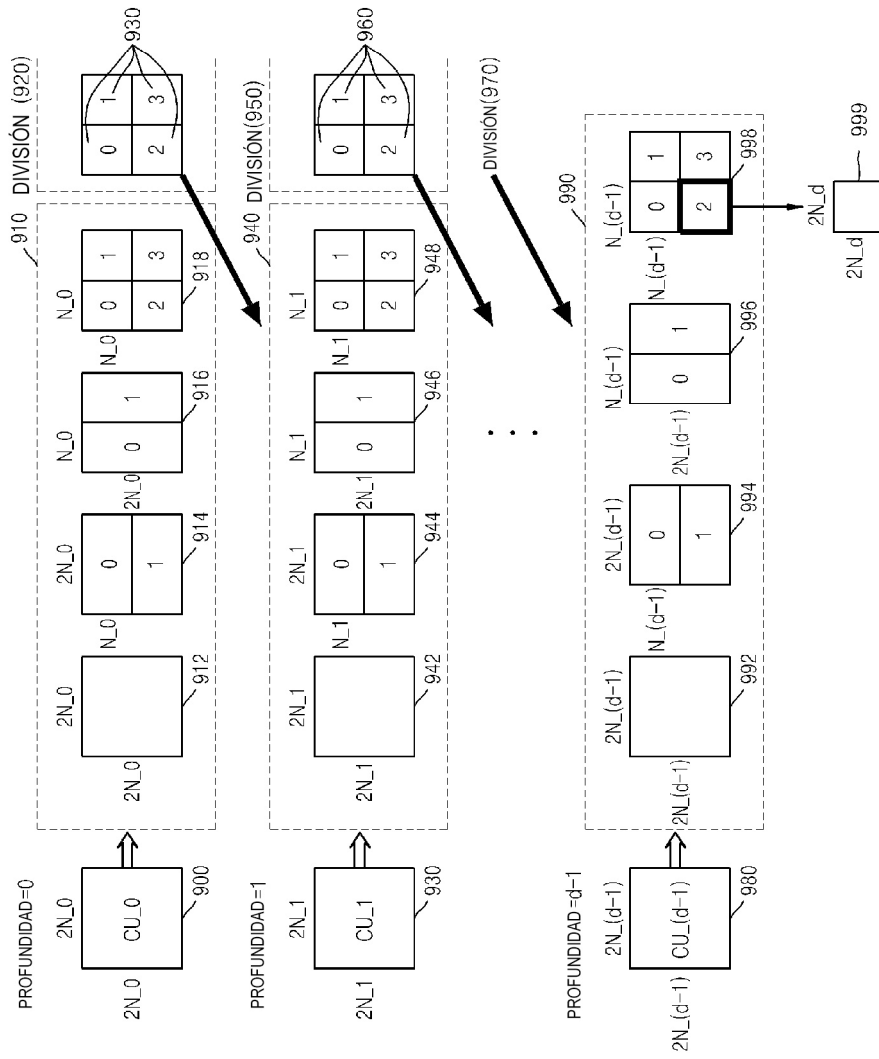


FIG. 22

1012			1014	1016	
			1018	1020	1022
1028	1030	1032		1054	
	1040	1042	1048		
1044	1046				
1050	1052				

UNIDADES DE CODIFICACIÓN (1010)

FIG. 23

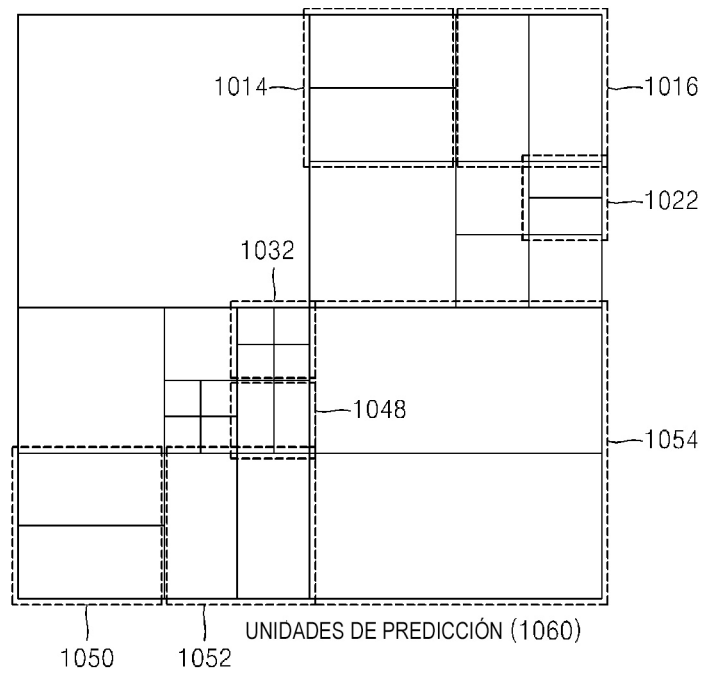


FIG. 24

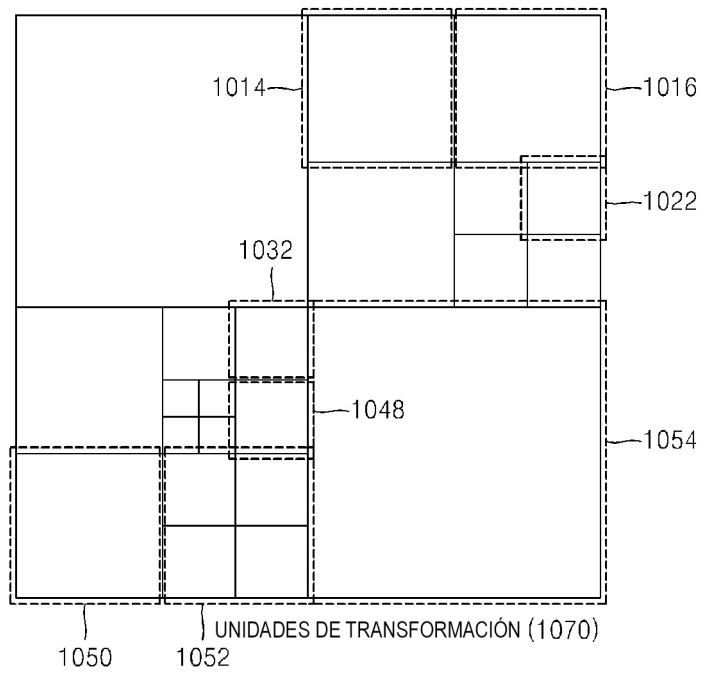


FIG. 25

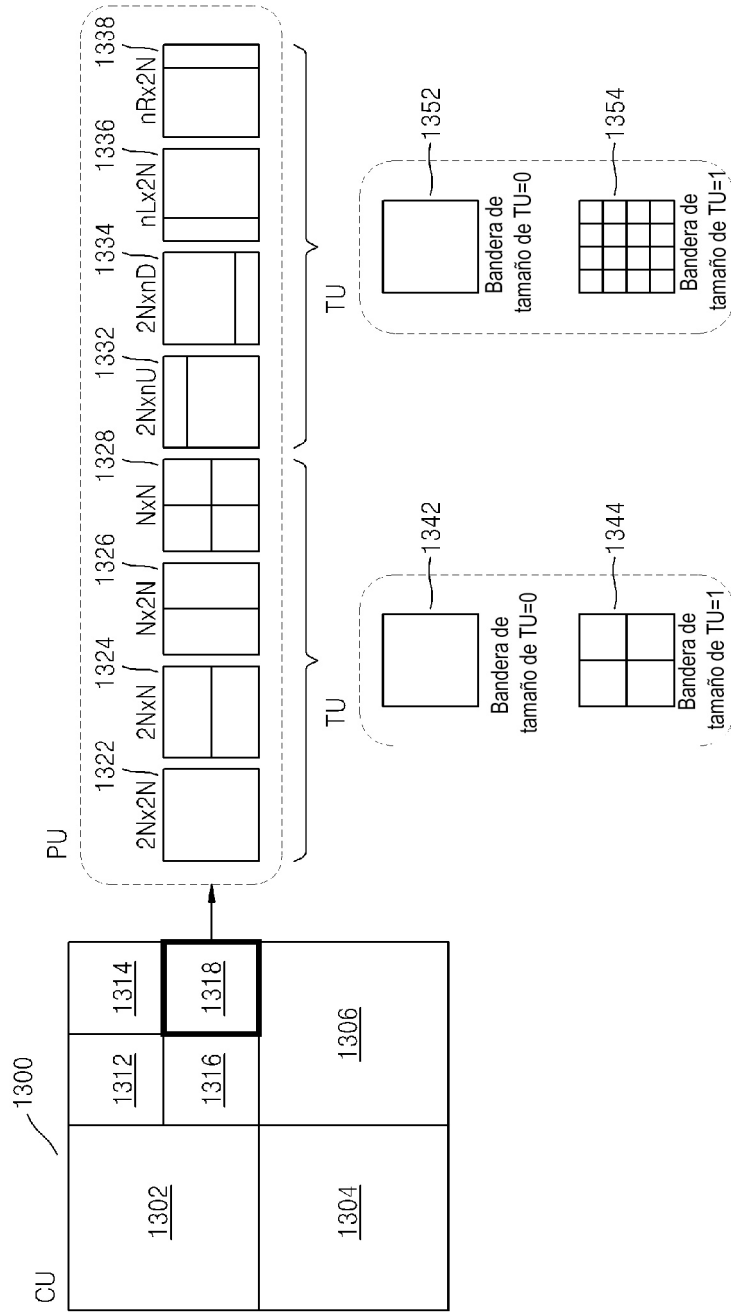


FIG. 26

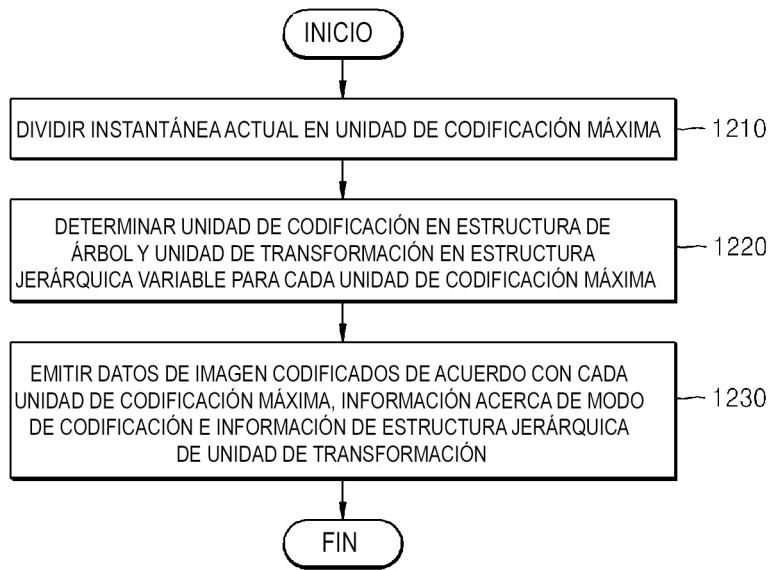


FIG. 27

