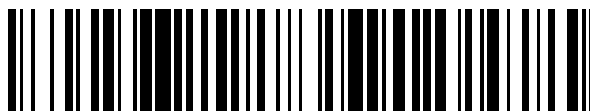


19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 711 671**

51 Int. Cl.:

H04N 19/91 (2014.01)
H04N 19/169 (2014.01)
H04N 19/176 (2014.01)
H04N 19/13 (2014.01)
H04N 19/60 (2014.01)
H04N 19/593 (2014.01)
H04N 19/157 (2014.01)
H04N 19/44 (2014.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

- 86 Fecha de presentación y número de la solicitud internacional: **27.06.2012** **PCT/KR2012/005087**
87 Fecha y número de publicación internacional: **03.01.2013** **WO13002555**
96 Fecha de presentación y número de la solicitud europea: **27.06.2012** **E 12804190 (2)**
97 Fecha y número de publicación de la concesión europea: **13.02.2019** **EP 2728866**

54 Título: **Procedimiento de decodificación de vídeo usando decodificación**

30 Prioridad:

28.06.2011 US 201161502038 P

45 Fecha de publicación y mención en BOPI de la traducción de la patente:
06.05.2019

73 Titular/es:

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

72 Inventor/es:

KIM, IL-KOO y
SEREGIN, VADIM

74 Agente/Representante:

CARPINTERO LÓPEZ, Mario

ES 2 711 671 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 decodificación de vídeo usando decodificación

[Campo técnico]

5 La presente invención se refiere a codificación de vídeo y decodificación de vídeo que implican codificación aritmética y decodificación aritmética, respectivamente.

[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 convencional, se codifica un vídeo de acuerdo con un procedimiento de codificación limitado basado en un macrobloque que tiene un tamaño predeterminado.

15 Los datos de imagen de un dominio espacial se convierten a coeficientes de una región de frecuencia usando un procedimiento de conversión de frecuencia. Un códec de vídeo codifica coeficientes de frecuencia en unidades de bloques dividiendo una imagen en una pluralidad de bloques que tienen un tamaño predeterminado y realizando conversión de Transformación de Coseno Discreta (DCT) para rápida operación de conversión de frecuencia. Los coeficientes de la región de frecuencia se comprimen fácilmente en comparación con los datos de imagen del dominio espacial. En particular, un valor de píxel de una imagen en el dominio espacial se representa como un error de predicción y por lo tanto si se realiza conversión de frecuencia en el error de predicción, una gran cantidad de datos puede convertirse a 0. Un códec de vídeo convierte datos que se generan de manera continua y repetitiva en pequeños datos para reducir una cantidad de datos.

[Divulgación]

[Problema técnico]

La presente invención proporciona un procedimiento y aparato para realizar codificación aritmética y decodificación aritmética de un vídeo clasificando un símbolo en cadenas de bits de prefijo y sufijo.

[Solución técnica]

25 De acuerdo con un ejemplo se proporciona un procedimiento de decodificación de un vídeo a través de decodificación de símbolo, incluyendo el procedimiento: analizar símbolos de bloques de imagen desde una secuencia de bits recibida; clasificar un símbolo actual en una cadena de bits de prefijo y una cadena de bits de sufijo basándose en un valor umbral determinado de acuerdo con un tamaño de un bloque actual; realizar decodificación aritmética usando un procedimiento de decodificación aritmética determinado para cada una de la cadena de bits de prefijo y la cadena de bits de sufijo; realizar conversión a binario inversa usando un procedimiento de conversión a binario determinado para cada una de la cadena de bits de prefijo y la cadena de bits de sufijo; y restaurar los bloques de imagen realizando transformación inversa y predicción en el bloque actual usando el símbolo actual restaurado a través de la decodificación aritmética y la conversión a binario inversa.

[Efectos ventajosos]

35 La eficacia de un procedimiento de codificación/decodificación de símbolo se mejora realizando un procedimiento de conversión a binario que tiene una cantidad relativamente pequeña de carga de operación en la región de sufijo o la cadena de bits de sufijo u omitiendo el modelado de contexto durante la codificación/decodificación aritmética basada en contexto para codificación/decodificación de símbolo.

[Descripción de los dibujos]

40 La Figura 1 es un diagrama de bloques de un aparato de codificación de vídeo, de acuerdo con una realización de la presente invención;

La Figura 2 es un diagrama de bloques de un aparato de decodificación de vídeo, de acuerdo con una realización de la presente invención;

45 Las Figuras 3 y 4 son diagramas para describir codificación aritmética clasificando un símbolo en una cadena de bits de prefijo y una cadena de bits de sufijo de acuerdo con un valor umbral predeterminado, de acuerdo con una realización de la presente invención;

La Figura 5 es un diagrama de flujo para describir un procedimiento de codificación de vídeo, de acuerdo con una realización de la presente invención;

50 La Figura 6 es un diagrama de flujo para describir un procedimiento de decodificación de vídeo, de acuerdo con una

realización de la presente invención;

La Figura 7 es un diagrama de bloques de un aparato de codificación de vídeo basado en unidades de codificación que tienen una estructura de árbol, de acuerdo con una realización de la presente invención;

5 La Figura 8 es un diagrama de bloques de un aparato de decodificación de vídeo basado en una unidad de codificación que tiene una estructura de árbol, de acuerdo con una realización de la presente invención;

La Figura 9 es un diagrama conceptual de unidades de codificación, de acuerdo con una realización de la presente invención;

La Figura 10 es un diagrama de bloques de un codificador de imagen basado en unidades de codificación, de acuerdo con una realización de la presente invención;

10 La Figura 11 es un diagrama de bloques de un decodificador de imagen basado en unidades de codificación, de acuerdo con una realización de la presente invención;

La Figura 12 es un diagrama que muestra unidades de codificación de acuerdo con las profundidades y particiones, de acuerdo con una realización de la presente invención;

15 La Figura 13 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 14 es un diagrama para describir información de codificación de unidades de codificación de acuerdo con las profundidades, de acuerdo con una realización de la presente invención;

La Figura 15 es un diagrama que muestra unidades de codificación de acuerdo con las profundidades, de acuerdo con una realización de la presente invención;

20 Las Figuras 16 a 18 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; y

La Figura 19 es un diagrama para describir una relación entre una unidad de codificación, una unidad de predicción, y una unidad de transformación de acuerdo con la información de modo de codificación de la Tabla 1.

[Mejor modo]

25 De acuerdo con un aspecto de la presente invención, se proporciona un procedimiento de decodificación de un vídeo como se define mediante la reivindicación 1.

[Modo para la invención]

30 En lo sucesivo, la presente invención se describirá más completamente con referencia a los dibujos adjuntos, en los que se muestran realizaciones ejemplares de la invención. Expresiones tales como "al menos uno de", cuando preceden una lista de elementos, modifican la lista completa de elementos y no modifican los elementos individuales de la lista.

35 Un procedimiento de codificación de vídeo que implica codificación aritmética y un procedimiento de decodificación de vídeo que implica decodificación aritmética de acuerdo con una realización de la presente invención se describirán con referencia a las Figuras 1 a 6. También, un procedimiento de codificación de vídeo que implica codificación aritmética y un procedimiento de decodificación de vídeo que implica decodificación aritmética basándose en unidades de codificación que tienen una estructura de árbol de acuerdo con una realización de la presente invención, se describirán con referencia a las Figuras 7 a 19. En lo sucesivo, una 'imagen' puede hacer referencia a una imagen fija de un vídeo o una película, es decir, un propio vídeo.

40 En lo sucesivo, un procedimiento de codificación de vídeo y un procedimiento de decodificación de vídeo, de acuerdo con una realización de la presente invención, basándose en un procedimiento de predicción en un modo de intra predicción se describirán con referencia a las Figuras 1 a 6.

La Figura 1 es un diagrama de bloques de un aparato 10 de codificación de vídeo, de acuerdo con una realización de la presente invención.

45 El aparato 10 de codificación de vídeo puede codificar datos de vídeo de un dominio espacial a través de intra predicción/inter predicción, transformación, cuantificación y codificación de símbolo. En lo sucesivo, se describirán en detalle las operaciones que tienen lugar cuando el aparato 10 de codificación de vídeo codifica símbolos generados por la intra predicción/inter predicción, la transformación, y la cuantificación a través de codificación aritmética.

50 El aparato 10 de codificación de vídeo incluye un codificador 12 de imagen, un codificador 14 de símbolo, y una unidad 16 de salida de secuencia de bits.

- El aparato 10 de codificación de vídeo puede dividir datos de imagen de un vídeo en una pluralidad de unidades de datos y codificar los datos de imagen de acuerdo con las unidades de datos. La unidad de datos puede tener una forma cuadrada o una forma rectangular, o puede ser una forma geométrica arbitraria, pero la unidad de datos no está limitada a una unidad de datos que tiene un tamaño predeterminado. De acuerdo con el procedimiento de codificación de vídeo basado en las unidades de codificación que tienen una estructura de árbol, una unidad de datos puede ser una unidad de codificación máxima, una unidad de codificación, una unidad de predicción, una unidad de transformación o similares. Un ejemplo donde un procedimiento de codificación/decodificación aritmética de acuerdo con una realización de la presente invención se usa en el procedimiento de codificación/decodificación de vídeo basado en las unidades de codificación que tienen una estructura de árbol se describirá con referencia a las Figuras 7 a 19.
- Por conveniencia de descripción, se describirá en detalle un procedimiento de codificación de vídeo para un 'bloque' que es una clase de unidad de datos. Sin embargo, el procedimiento de codificación de vídeo de acuerdo con diversas realizaciones de la presente invención no está limitado al procedimiento de codificación de vídeo para el 'bloque', y puede usarse para diversas unidades de datos.
- El codificador 12 de imagen realiza operaciones, tales como intra predicción/inter predicción, transformación, o cuantificación, en bloques de imagen para generar símbolos.
- El codificador 14 de símbolo clasifica un símbolo actual en una región de prefijo y una región de sufijo basándose en un valor umbral determinado de acuerdo con un tamaño de un bloque actual para codificar el símbolo actual de entre los símbolos generados de acuerdo con los bloques. El codificador 14 de símbolo puede determinar el valor umbral para clasificar el símbolo actual en la región de prefijo y la región de sufijo basándose en al menos una de una anchura y una altura del bloque actual.
- El codificador 14 de símbolo puede determinar un procedimiento de codificación de símbolo para cada una de la región de prefijo y la región de sufijo y codificar cada una de la región de prefijo y la región de sufijo de acuerdo con el procedimiento de codificación de símbolo.
- La codificación de símbolo puede dividirse en un procedimiento de conversión a binario para transformar un símbolo en cadenas de bits y un procedimiento de codificación aritmética para realizar codificación aritmética basada en contexto en las cadenas de bits. El codificador 14 de símbolo puede determinar un procedimiento de conversión a binario para cada una de la región de prefijo y la región de sufijo del símbolo y realizar conversión a binario en cada una de la región de prefijo y la región de sufijo de acuerdo con el procedimiento de conversión a binario. Una cadena de bits de prefijo y una cadena de bits de sufijo pueden generarse desde la región de prefijo y la región de sufijo, respectivamente.
- Como alternativa, el codificador 14 de símbolo puede determinar un procedimiento de codificación aritmética para cada una de la cadena de bits de prefijo y la cadena de bits de sufijo del símbolo y realizar codificación aritmética en cada una de la cadena de bits de prefijo y la cadena de bits de sufijo de acuerdo con el procedimiento de codificación aritmética.
- También, el codificador 14 de símbolo puede determinar un procedimiento de conversión a binario para cada una de la región de prefijo y la región de sufijo del símbolo y realizar conversión a binario en cada una de la región de prefijo y la región de sufijo de acuerdo con el procedimiento de conversión a binario, y puede determinar un procedimiento de codificación aritmética para cada una de la cadena de bits de prefijo y la cadena de bits de sufijo del símbolo y realizar codificación aritmética en la cadena de bits de prefijo y la cadena de bits de sufijo de acuerdo con el procedimiento de codificación aritmética.
- El codificador 14 de símbolo de acuerdo con una realización de la presente invención puede determinar un procedimiento de conversión a binario para cada una de la región de prefijo y la región de sufijo. Los procedimientos de conversión a binario determinados para la región de prefijo y la región de sufijo pueden ser diferentes unos de los otros.
- El codificador 14 de símbolo puede determinar un procedimiento de codificación aritmética para cada una de la cadena de bits de prefijo y la cadena de bits de sufijo. Los procedimientos de codificación aritmética determinados para la cadena de bits de prefijo y la cadena de bits de sufijo pueden ser diferentes unos de los otros.
- Por consiguiente, el codificador 14 de símbolo puede convertir a binario la región de prefijo y la región de sufijo usando diferentes procedimientos únicamente en un procedimiento de conversión a binario de un procedimiento de decodificación de símbolo, o puede codificar la cadena de bits de prefijo y la cadena de bits de sufijo usando diferentes procedimientos únicamente en un procedimiento de codificación aritmética. También, el codificador 14 de símbolo puede codificar la región de prefijo (cadena de bits de prefijo) y la región de sufijo (cadena de bits de sufijo) usando diferentes procedimientos en ambos de los procedimientos de conversión a binario y codificación aritmética.
- El procedimiento de conversión a binario seleccionado puede ser al menos uno de los procedimientos de conversión a binario general, conversión a binario unaria, conversión a binario unaria truncada, conversión a binario de Golomb exponencial, y conversión a binario de longitud fija.

El codificador 14 de símbolo puede realizar codificación de símbolo realizando codificación aritmética para realizar modelado de contexto en la cadena de bits de prefijo de acuerdo con localizaciones de bits y realizando codificación aritmética para omitir modelado de contexto en la cadena de bits de sufijo en un modo de desvío.

5 El codificador 14 de símbolo puede realizar individualmente la codificación de símbolo en la región de prefijo y la región de sufijo con respecto a símbolos que incluyen al menos uno de un modo de intra predicción e información de posición de coeficiente final de un coeficiente de transformación.

10 El codificador 14 de símbolo puede realizar también la codificación aritmética usando un contexto de un índice predeterminado que se asignó previamente a la cadena de bits de prefijo. Por ejemplo, el codificador 14 de símbolo puede realizar codificación aritmética usando un contexto de un índice predeterminado que se asignó previamente a cada localización de los bits de la cadena de bits de prefijo cuando el símbolo es información de posición de coeficiente final del coeficiente de transformación.

La unidad 16 de salida de secuencia de bits emite cadenas de bits generadas a través de la codificación de símbolo en forma de secuencias de bits.

15 El aparato 10 de codificación de vídeo puede realizar codificación aritmética en símbolos de bloques de un vídeo y emitir los símbolos.

20 El aparato 10 de codificación de vídeo puede incluir un procesador central (no mostrado) para controlar todo el codificador 12 de imagen, el codificador 14 de símbolo y la unidad 16 de salida de secuencia de bits. Como alternativa, el codificador 12 de imagen, el codificador 14 de símbolo y la unidad 16 de salida de secuencia de bits pueden operarse por procesadores (no mostrados) respectivamente instalados en el mismo, y todo el aparato 10 de codificación de vídeo puede operarse operando sistemáticamente el procesador (no mostrado). Como alternativa, el codificador 12 de imagen, el codificador 14 de símbolo, y la unidad 16 de salida de secuencia de bits pueden controlarse por un procesador externo (no mostrado) del aparato 10 de codificación de vídeo.

25 El aparato 10 de codificación de vídeo puede incluir al menos una unidad de almacenamiento de datos (no mostrada) para almacenar datos que se introducen/emiten a/desde el codificador 12 de imagen, el codificador 14 de símbolo y la unidad 16 de salida de secuencia de bits. El aparato 10 de codificación de vídeo puede incluir un controlador de memoria (no mostrado) para controlar entrada/salida de datos almacenados en la unidad de almacenamiento de datos (no mostrada).

30 El aparato 10 de codificación de vídeo se opera estando enlazado con un procesador de codificación de vídeo interno o un procesador de codificación de vídeo externo para realizar la codificación de vídeo incluyendo predicción y transformación, emitiendo de esta manera un resultado de la codificación de vídeo. El procesador de codificación de vídeo interno del aparato 10 de codificación de vídeo puede realizar una operación de codificación de vídeo básica no únicamente usando un procesador separado, sino también incluyendo un módulo de procesamiento de codificación de vídeo en el aparato 10 de codificación de vídeo, un aparato de operación central, o un aparato de operación de gráficos.

35 La Figura 2 es un diagrama de bloques de un aparato 20 de decodificación de vídeo, de acuerdo con una realización de la presente invención.

40 El aparato 20 de decodificación de vídeo puede decodificar los datos de vídeo codificados por el aparato 10 de codificación de vídeo a través de análisis, decodificación de símbolo, cuantificación inversa, transformación inversa, intra predicción/compensación de movimiento, etc., y restaurar los datos de vídeo cerca de los datos de vídeo originales del dominio espacial. En lo sucesivo, se describirá un procedimiento en el que el aparato 20 de decodificación de vídeo realiza decodificación aritmética en los símbolos analizados de una secuencia de bits para restaurar los símbolos.

El aparato 20 de decodificación de vídeo incluye un analizador 22, un decodificador 24 de símbolo y una unidad 26 de restauración de imagen.

45 El aparato 20 de decodificación de vídeo puede recibir una secuencia de bits que incluye datos codificados de un vídeo. El analizador 22 puede analizar símbolos de bloques de imagen desde la secuencia de bits.

El analizador 22 puede analizar los símbolos codificados a través de codificación aritmética con respecto a los bloques del vídeo desde la secuencia de bits.

50 El analizador 22 puede analizar símbolos que incluyen un modo de intra predicción del bloque del vídeo, información de posición de coeficiente final de un coeficiente de transformación, etc., desde la secuencia de bits recibida.

El decodificador 24 de símbolo determina un valor umbral para clasificar un símbolo actual en una cadena de bits de prefijo y una cadena de bits de sufijo. El decodificador 24 de símbolo puede determinar el valor umbral para clasificar el símbolo actual en la cadena de bits de prefijo y la cadena de bits de sufijo basándose en un tamaño de un bloque actual, es decir, al menos una de una anchura y una altura del bloque actual. El decodificador 24 de símbolo

determina un procedimiento de decodificación aritmética para cada una de la cadena de bits de prefijo y la cadena de bits de sufijo. El decodificador 24 de símbolo realiza decodificación de símbolo usando el procedimiento de decodificación aritmética determinado para cada una de la cadena de bits de prefijo y la cadena de bits de sufijo.

5 Los procedimientos de decodificación aritmética determinados para la cadena de bits de prefijo y la cadena de bits de sufijo pueden ser diferentes unos de los otros.

El decodificador 24 de símbolo puede determinar un procedimiento de conversión a binario para cada una de la cadena de bits de prefijo y la cadena de bits de sufijo del símbolo. Por consiguiente, el decodificador 24 de símbolo puede realizar conversión a binario inversa en la cadena de bits de prefijo del símbolo usando el procedimiento de conversión a binario. Los procedimientos de conversión a binario determinados para la cadena de bits de prefijo y la
10 cadena de bits de sufijo pueden ser diferentes unos de los otros.

También, el decodificador 24 de símbolo puede realizar decodificación aritmética usando el procedimiento de decodificación aritmética determinado para cada una de la cadena de bits de prefijo y la cadena de bits de sufijo del símbolo, y puede realizar conversión a binario inversa usando el procedimiento de conversión a binario determinado para cada una de la cadena de bits de prefijo y la cadena de bits de sufijo generada a través de la decodificación aritmética.
15

Por consiguiente, el decodificador 24 de símbolo puede decodificar la cadena de bits de prefijo y la cadena de bits de sufijo usando diferentes procedimientos únicamente en un procedimiento de decodificación aritmética de un procedimiento de decodificación de símbolo, o puede realizar conversión a binario inversa usando diferentes procedimientos únicamente en un procedimiento de conversión a binario inversa. También, el decodificador 24 de
20 símbolo puede decodificar la cadena de bits de prefijo y la cadena de bits de sufijo usando diferentes procedimientos en ambos procedimientos de decodificación aritmética y conversión a binario inversa.

El procedimiento de conversión a binario determinado para cada una de la cadena de bits de prefijo y la cadena de bits de sufijo del símbolo puede no únicamente ser un procedimiento de conversión a binario general, sino que puede ser también al menos un procedimiento de conversión a binario unaria, conversión a binario unaria truncada, conversión a binario de Golomb exponencial, y procedimientos de conversión a binario de longitud fija.
25

El decodificador 24 de símbolo puede realizar decodificación aritmética para realizar modelado de contexto en la cadena de bits de prefijo de acuerdo con localizaciones de bits. El decodificador 24 de símbolo puede usar un procedimiento de decodificación aritmética para omitir modelado de contexto en la cadena de bits de sufijo en un modo de desvío. Por consiguiente, el decodificador 24 de símbolo puede realizar decodificación de símbolo a través de la decodificación aritmética realizada en cada una de la cadena de bits de prefijo y la cadena de bits de sufijo del símbolo.
30

El decodificador 24 de símbolo puede realizar la decodificación aritmética en la cadena de bits de prefijo y la cadena de bits de sufijo de símbolos que incluyen al menos uno de un modo de intra predicción e información de posición de coeficiente final de un coeficiente de transformación.

35 El decodificador 24 de símbolo puede realizar decodificación aritmética usando un contexto de un índice predeterminado que se asignó previamente de acuerdo con localizaciones de los bits de la cadena de bits de prefijo cuando el símbolo es información acerca de la posición de coeficiente final del coeficiente de transformación.

La unidad 26 de restauración de imagen puede restaurar una región de prefijo y una región de sufijo de un símbolo realizando la decodificación aritmética y la conversión a binario inversa en cada una de la cadena de bits de prefijo y la cadena de bits de sufijo. La unidad 26 de restauración de imagen puede restaurar el símbolo sintetizando la región de prefijo y la región de sufijo del símbolo.
40

La unidad 26 de restauración de imagen realiza transformación inversa y predicción en el bloque actual usando el símbolo actual restaurado a través de la decodificación aritmética y la conversión a binario inversa. La unidad 26 de restauración de imagen puede restaurar bloques de imagen realizando operaciones, tales como cuantificación inversa, transformación inversa, o intra predicción/compensación de movimiento, usando los correspondientes símbolos para cada uno de los bloques de imagen.
45

El aparato 20 de decodificación de vídeo de acuerdo con una realización de la presente invención puede incluir un procesador central (no mostrado) para controlar todos del analizador 22, el decodificador 24 de símbolo, y la unidad 26 de restauración de imagen. Como alternativa, el analizador 22, el decodificador 24 de símbolo, y la unidad 26 de restauración de imagen puede operarse por procesadores (no mostrados) respectivamente instalados en el mismo, y todo el aparato 20 de decodificación de vídeo puede operarse operando sistemáticamente el procesador (no mostrado). Como alternativa, el analizador 22, el decodificador 24 de símbolo, y la unidad 26 de restauración de imagen pueden controlarse por un procesador externo (no mostrado) del aparato 20 de decodificación de vídeo.
50

El aparato 20 de decodificación de vídeo puede incluir al menos una unidad de almacenamiento de datos (no mostrada) para almacenar datos que se introducen/emiten a/desde el analizador 22, el decodificador 24 de símbolo, y la unidad 26 de restauración de imagen. El aparato 20 de decodificación de vídeo puede incluir un controlador de
55

memoria (no mostrado) para controlar entrada/salida de datos almacenados en la unidad de almacenamiento de datos (no mostrada).

El aparato 20 de decodificación de vídeo se opera enlazándose con un procesador de decodificación de vídeo interno o un procesador de decodificación de vídeo externo para realizar decodificación de vídeo incluyendo transformación inversa. El procesador de decodificación de vídeo interno del aparato 20 de decodificación de vídeo puede realizar una operación de decodificación de vídeo básica no únicamente usando un procesador separado, sino también incluyendo un módulo de procesamiento de decodificación de vídeo en el aparato 20 de decodificación de vídeo, un aparato de operación central o un aparato de operación de gráficos.

La codificación aritmética binaria adaptativa basada en contexto (CABAC) se usa ampliamente como un procedimiento de codificación/decodificación aritmética basándose en un contexto para codificación/decodificación de símbolo. De acuerdo con la codificación/decodificación aritmética basada en contexto, cada bit de una cadena de bits de símbolo puede ser un binario de un contexto, y una localización de cada bit puede mapearse en un índice binario. Una longitud de la cadena de bits, es decir, una longitud del binario, puede variar de acuerdo con un tamaño de un valor de símbolo. El modelado de contexto para determinar un contexto de un símbolo se requiere para realizar la codificación/decodificación aritmética basada en contexto. El contexto se renueva de acuerdo con localizaciones de bits de la cadena de bits de símbolo, es decir, en cada índice binario, para realizar el modelado de contexto, y por lo tanto se requiere un procedimiento de operación complicado.

De acuerdo con el aparato 10 de codificación de vídeo y el aparato 20 de decodificación de vídeo descritos con referencia a las Figuras 1 y 2, el símbolo se clasifica en la región de prefijo y la región de sufijo, y un procedimiento de conversión a binario relativamente sencillo puede usarse para la región de sufijo en comparación con la región de prefijo. También, la codificación/decodificación aritmética a través del modelado de contexto se realiza en la cadena de bits de prefijo, y el modelado de contexto no se realiza en la cadena de bits de sufijo, y por lo tanto puede reducirse una carga de una cantidad de operación para la codificación/decodificación aritmética basada en contexto. Por consiguiente, el aparato 10 de codificación de vídeo y el aparato 20 de decodificación de vídeo pueden mejorar la eficacia de un procedimiento de codificación/decodificación de símbolo realizando un procedimiento de conversión a binario que tiene una cantidad relativamente pequeña de carga de operación en la región de sufijo o la cadena de bits de sufijo u omitiendo el modelado de contexto durante la codificación/decodificación aritmética basada en contexto para codificación/decodificación de símbolo.

En lo sucesivo, se describirán diversas realizaciones para codificación aritmética que pueden realizarse por el aparato 10 de codificación de vídeo y el aparato 20 de decodificación de vídeo.

Las Figuras 3 y 4 son diagramas para describir codificación aritmética clasificando un símbolo en una cadena de bits de prefijo y una cadena de bits de sufijo de acuerdo con un valor umbral predeterminado, de acuerdo con una realización de la presente invención.

Haciendo referencia a la Figura 3, se describirá en detalle un procedimiento de realización de codificación de símbolo, de acuerdo con una realización de la presente invención, en información de posición de coeficiente final de un símbolo. La información de posición de coeficiente final es un símbolo que representa una localización de un coeficiente final, no 0, de entre coeficientes de transformación de un bloque. Puesto que un tamaño del bloque se define como una anchura y una altura, la información de posición de coeficiente final puede representarse por coordenadas bidimensionales, es decir, una coordenada x en una dirección de anchura y una coordenada y en una dirección de altura. Por conveniencia de descripción, la Figura 3 muestra un caso donde la codificación de símbolo se realiza en la coordenada x en la dirección de anchura de entre la información de posición de coeficiente final cuando una anchura de un bloque es w.

Un rango de la coordenada x de la información de posición de coeficiente final está dentro de la anchura del bloque, y por lo tanto la coordenada x de la información de posición de coeficiente final es igual o mayor que 0 e igual o menor que w-1. Para la codificación aritmética del símbolo, el símbolo puede clasificarse en una región de prefijo y una región de sufijo basándose en un valor umbral predeterminado th . Por lo tanto, la codificación aritmética puede realizarse en la cadena de bits de prefijo en la que se convierte a binario la región de prefijo, basándose en el contexto determinado a través del modelado de contexto. También, la codificación aritmética puede realizarse en la cadena de bits de sufijo en la que se convierte a binario la región de sufijo, en un modo de desvío en el que se omite el modelado de contexto.

En este punto, el valor umbral th para clasificar el símbolo en la región de prefijo y la región de sufijo puede determinarse basándose en la anchura w del bloque. Por ejemplo, el valor umbral th puede determinarse para que sea $(w/2)-1$ para dividir la cadena de bits por dos (fórmula de determinación de valor umbral 1). Como alternativa, la anchura w del bloque generalmente tiene un cuadrado de 2, y por lo tanto el valor umbral th puede determinarse basándose en un valor logarítmico de la anchura w (fórmula de determinación de valor umbral 2).

< fórmula de determinación de valor umbral 1 > $th = (w/2) - 1$;

< fórmula de determinación de valor umbral 2 > $th = (\log_2 w \ll 1) - 1$;

En la Figura 3, de acuerdo con la fórmula de determinación de valor umbral 1, cuando la anchura w del bloque es 8, la fórmula proporciona el valor umbral $th = (8/2) - 1 = 3$. Por lo tanto, en la coordenada x de la información de posición de coeficiente final, 3 puede clasificarse como la región de prefijo, y el resto de valores distintos de 3 pueden clasificarse como la región de sufijo. La región de prefijo y la región de sufijo pueden convertirse a binario de acuerdo con el procedimiento de conversión a binario determinado para cada una de la región de prefijo y la región de sufijo.

Cuando una coordenada x N de información de posición de coeficiente final actual es 5, la coordenada x de la información de posición de coeficiente final puede clasificarse como $N = th + 2 = 3 + 2$. En otras palabras, en la coordenada x de la información de posición de coeficiente final, 3 puede clasificarse como la región de prefijo, y 2 puede clasificarse como la región de sufijo.

De acuerdo con una realización de la presente invención, la región de prefijo y la región de sufijo pueden convertirse a binario de acuerdo con diferentes procedimientos de conversión a binario determinados para la región de prefijo y la región de sufijo, respectivamente. Por ejemplo, la región de prefijo puede convertirse a binario de acuerdo con un procedimiento de conversión a binario unaria, y la región de sufijo puede convertirse a binario de acuerdo con un procedimiento de conversión a binario general.

Por consiguiente, después de que 3 se convierte a binario de acuerdo con el procedimiento de conversión a binario unaria, una cadena 32 de bits de prefijo '0001' puede generarse a partir de la región de prefijo, y después de que 2 se convierte a binario de acuerdo con el procedimiento de conversión a binario general, una cadena 34 de bits de sufijo '010' puede generarse a partir de la región de sufijo.

También, la codificación aritmética basada en contexto puede realizarse en la cadena 32 de bits de prefijo '0001' a través de modelado de contexto. Por lo tanto, un índice de contexto puede determinarse para cada binario de la cadena 32 de bits de prefijo '0001'.

La codificación aritmética puede realizarse en la cadena 34 de bits de sufijo '010' en un modo de desvío sin realizar modelado de contexto. La codificación aritmética puede realizarse sin realizar el modelado de contexto suponiendo que en el modo de desvío cada binario tiene un contexto de un estado de probabilidad igual, es decir, el contexto del 50 %.

Por consiguiente, la codificación aritmética basada en contexto puede realizarse en cada una de la cadena 32 de bits de prefijo '0001' y la cadena 34 de bits de sufijo '010' para completar la codificación de símbolo con respecto a la coordenada x N de la información de posición de coeficiente final actual.

Aunque se ha descrito la realización en la que se realiza la codificación de símbolo mediante la conversión a binario y la codificación aritmética, la decodificación de símbolo puede realizarse de la misma manera. En otras palabras, una cadena de bits de símbolos analizados puede clasificarse en una cadena de bits de prefijo y una cadena de bits de sufijo basándose en la anchura w del bloque, la decodificación aritmética puede realizarse en la cadena 32 de bits de prefijo a través del modelado de contexto, y la decodificación aritmética puede realizarse en la cadena 34 de bits de sufijo sin realizar el modelado de contexto. La conversión a binario inversa puede realizarse en la cadena 32 de bits de prefijo después de la decodificación aritmética usando el procedimiento de conversión a binario unaria, y la región de prefijo puede restaurarse. También, la conversión a binario inversa puede realizarse en la cadena 34 de bits de sufijo después de la codificación aritmética usando el procedimiento de conversión a binario general, y por lo tanto la región de sufijo puede restaurarse. El símbolo puede restaurarse sintetizando la región de prefijo y la región de sufijo restauradas.

Aunque se ha descrito la realización en la que se usa el procedimiento de conversión a binario unaria para la región de prefijo (cadena de bits de prefijo) y se usa el procedimiento de conversión a binario general para la región de sufijo (cadena de bits de sufijo), el procedimiento de conversión a binario no está limitado a lo mismo. Como alternativa, puede usarse procedimiento de conversión a binario unaria truncada para la región de prefijo (cadena de bits de prefijo), y puede usarse un procedimiento de conversión a binario de longitud fija para la región de sufijo (cadena de bits de sufijo).

Aunque únicamente se ha descrito la realización con respecto a la información de posición de coeficiente final en una dirección de anchura del bloque, puede usarse también una realización con respecto a una información de posición de coeficiente final en una dirección de altura del bloque.

También, no hay necesidad de realizar modelado de contexto en la cadena de bits de sufijo para realizar la codificación aritmética usando un contexto que tiene una probabilidad fija, pero es necesario realizar modelado de contexto variable en la cadena de bits de prefijo. El modelado de contexto a realizarse en la cadena de bits de prefijo puede determinarse de acuerdo con un tamaño del bloque.

<Tabla de mapeo de contexto>

Tamaño de bloque	N.º de índice binario de contexto seleccionado
4x4	0, 1, 2, 2

8x8	3, 4, 5, 5
16x16	6, 7, 8, 9, 10, 10, 11, 11
32x32	12, 13, 14, 15, 16, 16, 16, 16, 17, 17, 17, 17, 18, 18, 18, 18

En la tabla de mapeo de contexto, una localización de cada número corresponde al índice binario de la cadena de bits de prefijo, y el número indica un índice de contexto a usarse en una localización del bit correspondiente. Por conveniencia de descripción, por ejemplo, en un bloque de 4x4, la cadena de bits de prefijo está comprendida de un total de cuatro bits, y cuando k es 0, 1, 2 y 3 de acuerdo con la tabla de mapeo de contexto, los índices de contexto 0, 1, 2 y 2 se determinan para un índice binario de orden k, y por lo tanto puede realizarse la codificación aritmética basándose en el modelado de contexto.

La Figura 4 muestra una realización en la que un modo de intra predicción incluye un intra modo de luminancia y un intra modo de crominancia que indica una dirección de intra predicción de un bloque de luminancia y un bloque de crominancia, respectivamente. Cuando el modo de intra predicción es 6, una cadena 40 de bits de símbolo '0000001' se genera de acuerdo con un procedimiento de conversión a binario unaria. En este caso, la codificación aritmética puede realizarse en un primer bit 41 '0' de la cadena 40 de bits de símbolo del modo de intra predicción, a través de modelado de contexto, y la codificación aritmética puede realizarse en el resto de bits 45 '000001' de la cadena 40 de bits de símbolo, en un modo de desvío. En otras palabras, el primer bit 41 de la cadena 40 de bits de símbolo corresponde a una cadena de bits de prefijo, y el resto de bits 45 de la cadena 40 de bits de símbolo corresponde a una cadena de bits de sufijo.

Cuántos bits de la cadena 40 de bits de símbolo se codifican en codificación aritmética como la cadena de bits de prefijo a través del modelado de contexto y cuántos bits de la cadena 40 de bits de símbolo se codifican en codificación aritmética como la cadena de bits de sufijo en el modo de desvío puede determinarse de acuerdo con un tamaño de un bloque o un tamaño de un conjunto de bloques. Por ejemplo, con respecto a un bloque de 64x64, la codificación aritmética puede realizarse únicamente en un primer bit de entre las cadenas de bits de un modo de intra predicción, y la codificación aritmética puede realizarse en el resto de bits en un modo de desvío. Con respecto a los bloques que tienen otros tamaños, la codificación aritmética puede realizarse en todos los bits de las cadenas de bits del modo de intra predicción en el modo de desvío.

En general, la información acerca de los bits cercanos a un bit menos significativo (LSB) es relativamente menos importante que la información acerca de los bits cercanos a un bit más significativo (MSB) de una cadena de bits de símbolo. Por consiguiente, el aparato 10 de codificación de vídeo y el aparato 20 de decodificación de vídeo pueden seleccionar un procedimiento de codificación aritmética de acuerdo con un procedimiento de conversión a binario que tiene una precisión relativamente alta con respecto a la cadena de bits de prefijo cercana al MSB incluso aunque haya una carga de una cantidad de operación, y puede seleccionar un procedimiento de codificación aritmética de acuerdo con un procedimiento de conversión a binario que puede realizar una operación sencilla con respecto a la cadena de bits de sufijo cercana al LSB. También, el aparato 10 de codificación de vídeo y el aparato 20 de decodificación de vídeo pueden seleccionar un procedimiento de codificación aritmética basándose en modelado de contexto con respecto al modelado de contexto y puede seleccionar un procedimiento de codificación aritmética, no realizando modelado de contexto, con respecto a la cadena de bits de sufijo cercana al LSB.

En la descripción anterior, se ha descrito la realización en la que se realiza conversión a binario en la cadena de bits de prefijo y la cadena de bits de sufijo de la información de posición de coeficiente final del coeficiente de transformación usando diferentes procedimientos con referencia a la Figura 3. También, se ha descrito la realización en la que se realiza codificación aritmética en la cadena de bits de prefijo y la cadena de bits de sufijo de entre las cadenas de bits del modo de intra predicción usando diferentes procedimientos con referencia a la Figura 4.

Sin embargo, de acuerdo con diversas realizaciones de la presente invención, un procedimiento de codificación de símbolo en el que se usan procedimientos de codificación aritmética/conversión a binario individualmente determinados para la cadena de bits de prefijo y la cadena de bits de sufijo o se usan diferentes procedimientos de codificación aritmética/conversión a binario no están limitados a las realizaciones descritas con referencia a las Figuras 3 y 4, y pueden usarse diversos procedimientos de codificación aritmética/conversión a binario para diversos símbolos.

La Figura 5 es un diagrama de flujo para describir un procedimiento de codificación de vídeo, de acuerdo con una realización de la presente invención.

En la operación 51, se generan símbolos realizando predicción y transformación en bloques de imagen.

En las operaciones 53, un símbolo actual se clasifica en una región de prefijo y una región de sufijo basándose en un valor umbral determinado de acuerdo con un tamaño de un bloque actual.

En la operación 55, una cadena de bits de prefijo y una cadena de bits de sufijo se generan usando procedimientos de conversión a binario determinados individualmente para la región de prefijo y la región de sufijo del símbolo.

En la operación 57, se realiza codificación de símbolo usando procedimientos de codificación aritmética

determinados individualmente para la cadena de bits de prefijo y la cadena de bits de sufijo.

En la operación 59, las cadenas de bits generadas a través de la codificación de símbolo se emiten en forma de secuencias de bits.

5 En la operación 57, la codificación de símbolo puede realizarse en la cadena de bits de prefijo usando un procedimiento de codificación aritmética para realizar modelado de contexto de acuerdo con localizaciones de bits, y la codificación de símbolo puede realizarse también en la cadena de bits de sufijo usando un procedimiento de codificación aritmética para omitir el modelado de contexto en un modo de desvío.

10 En la operación 57, cuando el símbolo es información de posición de coeficiente final de un coeficiente de transformación, la codificación aritmética puede realizarse usando un contexto de un índice predeterminado que se asignó previamente a las localizaciones de los bits de la cadena de bits de prefijo.

La Figura 6 es un diagrama de flujo para describir un procedimiento de decodificación de vídeo, de acuerdo con una realización de la presente invención.

En la operación 61, se analizan símbolos de bloques de imagen desde una secuencia de bits recibida.

15 En la operación 63, un símbolo actual se clasifica en una cadena de bits de prefijo y una cadena de bits de sufijo basándose en un valor umbral determinado de acuerdo con un tamaño de un bloque actual.

En la operación 65, la decodificación aritmética se realiza usando un procedimiento de decodificación aritmética determinado para cada una de la cadena de bits de prefijo y la cadena de bits de sufijo del símbolo actual.

20 En la operación 67, después de la decodificación aritmética, se realiza conversión a binario inversa usando un procedimiento de conversión a binario determinado para cada una de la cadena de bits de prefijo y la cadena de bits de sufijo.

La región de prefijo y la región de sufijo del símbolo pueden restaurarse realizando la conversión a binario inversa usando el procedimiento de conversión a binario determinado para cada una de la cadena de bits de prefijo y la cadena de bits de sufijo.

25 En la operación 69, los bloques de imagen pueden restaurarse realizando transformación inversa y predicción en el bloque actual usando el símbolo actual restaurado a través de la decodificación aritmética y la conversión a binario inversa.

En la operación 65, la decodificación aritmética para determinar el modelado de contexto de acuerdo con las localizaciones de los bits puede realizarse en la cadena de bits de prefijo, y la decodificación aritmética para omitir el modelado de contexto puede realizarse en la cadena de bits de sufijo en un modo de desvío.

30 En la operación 65, cuando el símbolo es la información de posición de coeficiente final del coeficiente de transformación, la decodificación aritmética puede realizarse usando el contexto del índice predeterminado que se asignó previamente a las localizaciones de los bits de la cadena de bits de prefijo.

35 En el aparato 10 de codificación de vídeo de acuerdo con una realización de la presente invención y el aparato 20 de decodificación de vídeo de acuerdo con otra realización de la presente invención, los bloques en los que se dividen los datos de vídeo se dividen en unidades de codificación que tienen una estructura de árbol, se usan unidades de predicción para realizar intra predicción en las unidades de codificación, y se usa una unidad de transformación para transformar las unidades de codificación.

40 En lo sucesivo, se describirá un procedimiento y aparato para codificar un vídeo y un procedimiento y aparato para decodificar un vídeo basándose en una unidad de codificación que tiene una estructura de árbol, una unidad de predicción, y una unidad de transformación.

La Figura 7 es un diagrama de bloques de un aparato 100 de codificación de vídeo basado en unidades de codificación que tienen una estructura de árbol, de acuerdo con una realización de la presente invención.

45 El aparato 100 de codificación de vídeo que implica predicción de vídeo basándose en la unidad de codificación que tiene una estructura de árbol incluye un divisor 110 de unidad de codificación máxima, un determinador 120 de unidad de codificación, y una unidad 130 de salida. Por conveniencia de descripción, el aparato 100 de codificación de vídeo que implica predicción de vídeo basándose en la unidad de codificación que tiene una estructura de árbol se denominará como un aparato 100 de codificación de vídeo.

50 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 de acuerdo con una realización de la presente invención 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 longitud 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.

5 Una unidad de codificación de acuerdo con una realización de la presente invención puede estar caracterizada por un tamaño máximo y una profundidad. La profundidad indica un número de veces que la unidad de codificación se divide espacialmente de la unidad de codificación máxima, y a medida que la profundidad se 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 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.

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

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

25 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 finalmente datos de imagen codificados de acuerdo con la al menos una región de división. En otras palabras, 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 seleccionando una profundidad que tiene el mínimo error de codificación. 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.

30 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 resultados de codificar los datos de imagen se comparan basándose en cada una de las unidades de codificación más profundas. Una profundidad que tiene el mínimo error 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.

35 El tamaño de la unidad de codificación máxima se divide a medida que una unidad de codificación se divide jerárquicamente de acuerdo con las profundidades, y a medida que aumenta el número de unidades de codificación. También, incluso si las unidades de codificación corresponden a una 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 los datos de imagen se incluyen en una unidad de 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 la unidad de codificación máxima, y por lo tanto las profundidades codificadas 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.

40 Por consiguiente, el determinador 120 de unidad de codificación puede determinar unidades de codificación que tienen una estructura de árbol incluidas en la unidad de codificación máxima. Las 'unidades de codificación que tienen una estructura de árbol' de acuerdo con una realización de la presente invención incluyen unidades de codificación 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 de una profundidad codificada en otra región.

Una profundidad máxima de acuerdo con una realización de la presente invención es un índice relacionado con el número de veces que se realiza división desde una unidad de codificación máxima a una unidad de codificación

mínima. Una primera profundidad máxima de acuerdo con una realización de la presente invención puede indicar el número total de veces que se realiza división desde la unidad de codificación máxima a la unidad de codificación mínima. Una segunda profundidad máxima de acuerdo con una realización de la presente invención puede indicar el número total de niveles de profundidad desde la unidad de codificación máxima 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 la unidad de codificación máxima se divide una vez, puede establecerse a 1, y una profundidad de una unidad de codificación, en la que la unidad de codificación máxima se divide 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, y por lo tanto la primera profundidad máxima puede establecerse a 4, y la segunda profundidad máxima puede establecerse a 5.

La codificación y transformación de predicción puede realizarse de acuerdo con la unidad de codificación máxima. La codificación por predicción y la transformación se realizan 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.

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, incluyendo la codificación la codificación por predicción y la transformación se realizan 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, tal como codificación por 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 codificación de 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 más en unidades de codificación que corresponden a una profundidad inferior. En lo sucesivo, la unidad de codificación que ya no se divide más y se convierte en una unidad de base para codificación de 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. Una partición es una unidad de datos que tiene una forma en la que se divide la 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.

Por ejemplo, cuando una unidad de codificación de $2N \times 2N$ (donde N es un entero positivo) ya no se divide y se vuelve una unidad de predicción de $2N \times 2N$, y 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 anchura de la unidad de predicción, particiones obtenidas dividiendo asimétricamente la altura o anchura de la unidad de predicción, tal como $1:n$ o $n:1$, particiones que se obtienen dividiendo geométricamente 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 puede realizarse en la partición de $2N \times 2N$, $2N \times N$, $N \times 2N$, o $N \times N$. También, el modo de salto puede realizarse únicamente en la partición de $2N \times 2N$. La codificación se realiza de manera independiente en una unidad de predicción en una unidad de codificación, seleccionando de esta manera un modo de predicción que tiene un mínimo error de codificación.

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

De manera similar a la unidad de codificación, la unidad de transformación en la unidad de codificación puede dividirse de manera recursiva en regiones con tamaño más pequeño. Por lo tanto, los datos residuales en la unidad de codificación pueden dividirse de acuerdo con la transformación que tiene una estructura de árbol de acuerdo con

profundidades de transformación.

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

La información de codificación de acuerdo con unidades de codificación que corresponden a una profundidad codificada requiere no únicamente información acerca de la profundidad codificada, sino también información relacionada con codificación de predicción y transformación. Por consiguiente, el determinador 120 de unidad de codificación no únicamente determina una profundidad codificada que tiene un mínimo error de codificación, sino también determina 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.

Se describirá en detalle más adelante las unidades de codificación de acuerdo con una estructura de árbol en una unidad de codificación máxima y un procedimiento de determinación de una unidad de predicción/partición y una unidad de transformación, de acuerdo con las realizaciones de la presente invención con referencia a las Figuras 7 a 19.

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 basada en multiplicadores de Lagrange.

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.

Los datos de imagen codificados pueden obtenerse codificando datos residuales de una imagen.

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

La información acerca de la profundidad codificada puede definirse 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, y 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.

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 al menos existe una unidad de codificación de la profundidad inferior 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, y por lo tanto puede realizarse de manera recursiva la codificación para las unidades de codificación que tienen la misma profundidad.

Puesto que las unidades de codificación que tienen una estructura de árbol se determinan para una unidad de codificación máxima, y se determina la información acerca de al menos un modo de codificación para una unidad de codificación de una profundidad codificada, la información acerca de al menos un 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, y por lo tanto la información acerca de la profundidad codificada y el modo de codificación puede establecerse para los datos de imagen.

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

La unidad mínima de acuerdo con una realización de la presente invención es una unidad de datos rectangular obtenida dividiendo la unidad de codificación mínima que constituye la profundidad más inferior por 4. Como alternativa, la unidad mínima puede ser una unidad de datos rectangular máxima que puede incluirse en todas las unidades de codificación, unidades de predicción, unidades de partición y unidades de transformación incluidas en la

unidad de codificación máxima.

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 la información acerca del modo de predicción y acerca del tamaño de las particiones. La información de codificación de acuerdo con las unidades de predicción puede incluir información acerca de una dirección estimada de un inter modo, acerca de un índice de imagen de referencia del inter modo, acerca de un vector de movimiento, acerca de un componente de crominancia de un intra modo, y acerca de un procedimiento de interpolación del intra modo. También, 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 conjunto de parámetros de secuencia (SPS) o un conjunto de parámetros de instantánea (PPS).

También, la información acerca del tamaño máximo de la unidad de transformación permitida para el vídeo actual y la información acerca del tamaño mínimo de la unidad de transformación pueden emitirse mediante un encabezamiento de una secuencia de bits, un SPS o un PPS. La unidad 130 de salida puede codificar y emitir información de referencia, información de predicción de única dirección, información de tipo de corte que incluye un cuarto tipo de corte, etc., relacionada con la predicción anteriormente descrita con referencia a las Figuras 1 a 6.

En el aparato 100 de codificación de vídeo, la unidad de codificación más profunda puede ser una unidad de codificación obtenida dividiendo por dos una altura o anchura de una unidad de codificación de una profundidad superior, que se encuentra una capa 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 es $N \times N$. También, la unidad de codificación de la profundidad actual que tiene el tamaño de $2N \times 2N$ puede incluir como máximo 4 unidades de codificación de la profundidad inferior.

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 determinadas considerando características de la instantánea actual. También, puesto que la codificación puede realizarse en cada unidad de codificación máxima usando uno cualquiera de diversos modos de predicción y transformaciones, puede determinarse un modo de codificación óptimo considerando características de la unidad de codificación de diversos tamaños de imagen.

Por lo tanto, si una imagen que tiene alta resolución o una gran cantidad de datos se codifica en un macrobloque convencional, 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 disminuye la eficacia de compresión de datos. Sin embargo, usando el aparato 100 de codificación de vídeo, la eficacia de compresión de imagen puede aumentarse puesto que se ajusta una unidad de codificación 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.

El aparato 100 de codificación de vídeo de la Figura 7 puede realizar operaciones del aparato 10 de codificación de vídeo descritas con referencia a la Figura 1.

El determinador 120 de unidad de codificación puede realizar operaciones del codificador 12 de imagen del aparato 10 de codificación de vídeo. El determinador 120 de unidad de codificación puede determinar una unidad de predicción para intra predicción de acuerdo con unidades de codificación que tienen una estructura de árbol para cada unidad de codificación máxima, realizar la intra predicción en cada unidad de predicción, determinar una unidad de transformación para transformación, y realizar transformación en cada unidad de transformación.

La unidad 130 de salida puede realizar operaciones de una unidad 14 de codificación de símbolo y una unidad 16 de salida de secuencia de bits del aparato 10 de codificación de vídeo. Se generan los símbolos para diversas unidades de datos, tales como una instantánea, un corte, una unidad de codificación máxima, una unidad de codificación, una unidad de predicción, y una unidad de transformación, y cada uno de los símbolos se clasifica en una región de prefijo y una región de sufijo de acuerdo con un valor umbral determinado basándose en un tamaño de la correspondiente unidad de datos. La unidad 130 de salida puede generar una cadena de bits de prefijo y una cadena de bits de sufijo usando un procedimiento de conversión a binario determinado para cada una de la región de prefijo y la región de sufijo del símbolo. Una cualquiera de una conversión a binario general, una conversión a binario unaria, una conversión a binario unaria truncada, una conversión a binario de Golomb exponencial, y una conversión a binario de longitud fija se selecciona para convertir a binario la región de prefijo y la región de sufijo, generando de esta manera la cadena de bits de prefijo y la cadena de bits de sufijo.

La unidad 130 de salida puede realizar codificación de símbolo realizando codificación aritmética determinada para cada una de la cadena de bits de prefijo y la cadena de bits de sufijo. La unidad 130 de salida puede realizar codificación de símbolo realizando codificación aritmética para realizar modelado de contexto de acuerdo con localizaciones de bits en la cadena de bits de prefijo y realizando codificación aritmética para omitir modelado de

contexto en la cadena de bits de sufijo en un modo de desvío.

Por ejemplo, cuando se codifica información de posición de coeficiente final de un coeficiente de transformación de la unidad de transformación, el valor umbral para clasificar la cadena de bits de prefijo y la cadena de bits de sufijo puede determinarse de acuerdo con un tamaño (anchura o altura) de la unidad de transformación. Como alternativa, el valor umbral puede determinarse de acuerdo con tamaños de un corte que incluye la unidad de transformación actual, una unidad de codificación máxima, una unidad de codificación, una unidad de predicción, etc.

Como alternativa, puede determinarse por un índice máximo de un modo de intra predicción cuántos bits de una cadena de bits de símbolo se codifican en codificación aritmética como la cadena de bits de prefijo a través de modelado de contexto en el modo de intra predicción y cuántos bits de la cadena de bits de símbolo se codifican en codificación aritmética como la cadena de bits de sufijo en un modo de desvío. Por ejemplo, un total de 34 modos de intra predicción pueden usarse para unidades de predicción que tienen tamaños de 8x8, 16x16, y 32x32, un total de 17 modos de intra predicción pueden usarse para una unidad de predicción que tiene un tamaño de 4x4, y un número total de modos de intra predicción pueden usarse para una unidad de predicción que tiene un tamaño de 64x64. En este caso, puesto que las unidades de predicción que pueden usar el mismo número de los modos de intra predicción se consideran que tienen características estadísticas similares, un primer bit de entre las cadenas de bits en el modo de intra predicción puede codificarse a través del modelado de contexto para codificación aritmética con respecto a las unidades de predicción que tienen tamaños de 8x8, 16x16, y 32x32. También, todos los bits de entre las cadenas de bits en el modo de intra predicción pueden codificarse en el modo de desvío para codificación aritmética con respecto al resto de unidades de predicción, es decir, las unidades de predicción que tienen tamaños de 4x4 y 64x64.

La unidad 130 de salida puede emitir las cadenas de bits generadas a través de la codificación de símbolo en forma de secuencias de bits.

La Figura 8 es un diagrama de bloques de un aparato 200 de decodificación de vídeo basado en una unidad de codificación que tiene una estructura de árbol, de acuerdo con una realización de la presente invención.

El aparato 200 de decodificación de vídeo que realiza predicción de vídeo basándose en la unidad de codificación que tiene una estructura de árbol incluye un receptor 210, un extractor 220 de datos de imagen e información de codificación, y un decodificador 230 de datos de imagen.

Las definiciones de diversos términos, tales como una unidad de codificación, una profundidad, una unidad de predicción, una unidad de transformación, e información acerca de diversos modos de codificación, para diversas operaciones del aparato 200 de decodificación de vídeo son idénticas a aquellas descritas con referencia a la Figura 7 y el aparato 100 de codificación de vídeo.

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 el 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 una unidad de codificación de una instantánea actual, desde un encabezamiento acerca de la instantánea actual, o un SPS, o un PPS.

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. En otras palabras, los datos de imagen en una cadena de bits se dividen en la unidad de codificación máxima de modo que el decodificador 230 de datos de imagen decodifica los datos de imagen para cada unidad de codificación máxima.

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, y la información acerca de un modo de codificación puede incluir información acerca de un tipo de partición de una unidad de codificación correspondiente que corresponde a la profundidad codificada, 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.

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 la 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.

Puesto que la información de codificación acerca de la profundidad codificada y el modo de codificación pueden 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 pueden inferirse para que sean las unidades de datos incluidas en la misma unidad de codificación máxima.

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 incluidas en cada unidad de codificación máxima. Un procedimiento de decodificación puede incluir predicción que incluye intra predicción y compensación de movimiento, y transformación inversa.

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 profundidades codificadas.

También, el decodificador 230 de datos de imagen puede realizar transformación inversa de acuerdo con cada unidad de transformación en la unidad de codificación, basándose en la información acerca de la unidad de transformación de acuerdo con las unidades de codificación que tienen una estructura de árbol, para realizar la transformación inversa de acuerdo con unidades de codificación máxima. Un valor de píxel de un dominio espacial en la unidad de codificación puede restaurarse a través de la transformación inversa.

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 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.

En otras palabras, las unidades de datos que contienen la información de codificación que incluye la misma información de división pueden recogerse 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, y las unidades de datos recogidas pueden considerarse que son una unidad de datos a decodificarse por el decodificador 230 de datos de imagen en el mismo modo de codificación. La decodificación de la unidad de codificación actual puede realizarse obteniendo información acerca del modo de codificación para cada unidad de codificación determinada de tal manera.

También, el aparato 200 de decodificación de vídeo de la Figura 8 puede realizar operaciones del aparato 20 de decodificación de vídeo anteriormente descrito con referencia a la Figura 2.

El receptor 210 y el extractor 220 de datos de imagen e información de codificación pueden realizar operaciones del analizador 22 y el decodificador 24 de símbolo del aparato 20 de decodificación de vídeo. El decodificador 230 de datos de imagen puede realizar operaciones del decodificador 24 de símbolo del aparato 20 de decodificación de vídeo.

El receptor 210 recibe una secuencia de bits de una imagen, y el extractor 220 de datos de imagen e información de codificación analiza símbolos de bloques de imagen desde la secuencia de bits recibida.

El extractor 220 de datos de imagen e información de codificación puede clasificar un símbolo actual en una cadena de bits de prefijo y una cadena de bits de sufijo basándose en un valor umbral determinado de acuerdo con un tamaño de un bloque actual. Por ejemplo, cuando se decodifica la información de posición de coeficiente final del coeficiente de transformación de la unidad de transformación, el valor umbral para clasificar la cadena de bits de prefijo y la cadena de bits de sufijo puede determinarse de acuerdo con un tamaño (anchura o altura) de la unidad de transformación. Como alternativa, el valor umbral puede determinarse de acuerdo con tamaños del recorte que incluye la unidad de transformación actual, la unidad de codificación máxima, la unidad de codificación, la unidad de predicción, etc. Como alternativa, puede determinarse por el índice máximo del modo de intra predicción cuántos bits de la cadena de bits de símbolo se codifican en codificación aritmética como la cadena de bits de prefijo a través

de modelado de contexto en el modo de intra predicción y cuántos bits de la cadena de bits de símbolo se codifican en codificación aritmética como la cadena de bits de sufijo en el modo de desvío.

La decodificación aritmética se realiza usando un procedimiento de decodificación aritmética determinado para cada una de la cadena de bits de prefijo y la cadena de bits de sufijo del símbolo actual. La decodificación aritmética para determinar el modelado de contexto de acuerdo con las posiciones de bits puede realizarse en la cadena de bits de prefijo, y la decodificación aritmética para omitir el modelado de contexto puede realizarse en la cadena de bits de sufijo usando el modo de desvío.

Después de la decodificación aritmética, se realiza conversión a binario inversa de acuerdo con un procedimiento de conversión a binario determinado para cada una de la cadena de bits de prefijo y la cadena de bits de sufijo. La región de prefijo y la región de sufijo del símbolo pueden restaurarse realizando la conversión a binario inversa de acuerdo con el procedimiento de conversión a binario determinado para cada una de la cadena de bits de prefijo y la cadena de bits de sufijo.

El decodificador 230 de datos de imagen puede restaurar bloques de imagen realizando transformación inversa y predicción en el bloque actual usando el símbolo actual restaurado a través de la decodificación aritmética y la conversión a binario inversa.

En consecuencia, 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 la codificación de manera recursiva para cada unidad de codificación máxima, y puede usar la información para decodificar la instantánea actual. En otras palabras, pueden codificarse las unidades de codificación que tienen la estructura de árbol determinadas para que sean las unidades de codificación óptimas en cada unidad de codificación máxima.

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

La Figura 9 es un diagrama conceptual 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. 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, o 8x8, y una unidad de codificación de 8x8 puede dividirse en particiones de 8x8, 8x4, 4x8, o 4x4.

En los datos 310 de vídeo, una resolución es 1920x1080, un tamaño máximo de una unidad de codificación es 64, y una profundidad máxima es 2. En los datos 320 de vídeo, una resolución es 1920x1080, un tamaño máximo de una unidad de codificación es 64, y una profundidad máxima es 3. En los datos 330 de vídeo, una resolución es 352x288, un tamaño máximo de una unidad de codificación es 16, y una profundidad máxima es 1. La profundidad máxima mostrada en la Figura 9 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 únicamente aumentar la eficacia de codificación sino también para reflejar de manera precisa características de una imagen. Por consiguiente, el tamaño máximo de la unidad de codificación de los datos 310 y 320 de vídeo que tienen una resolución superior a los datos 330 de vídeo puede ser 64.

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

Puesto que la profundidad máxima de los datos 320 de vídeo es 3, las unidades 325 de codificación de los datos 320 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, 16, y 8 puesto que las profundidades se hacen profundas a 3 capas dividiendo la unidad de codificación máxima tres veces. A medida que una profundidad se hace profunda, puede expresarse de manera precisa la información detallada.

La Figura 10 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.

- 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. En otras palabras, 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 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, y los datos restaurados en el dominio espacial se emiten como el fotograma 495 de referencia después de post-procesarse 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, todos 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 en bucle realizan operaciones basándose en cada unidad de codificación de entre las 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 el tamaño máximo y la profundidad máxima de una unidad de codificación máxima actual, y el transformador 430 determina el tamaño de unidad de transformación en cada unidad de codificación de entre las unidades de codificación que tienen una estructura de árbol.
- En particular, el codificador 450 por entropía puede realizar codificación de símbolo en la región de prefijo y la región de sufijo clasificando un símbolo en la región de prefijo y la región de sufijo de acuerdo con un valor umbral predeterminado y usando diferentes procedimientos de codificación aritmética y conversión a binario con respecto a la región de prefijo y la región de sufijo.
- El valor umbral para clasificar el símbolo en la región de prefijo y la región de sufijo puede determinarse basándose en tamaños de unidades de datos del símbolo, es decir, un corte, una unidad de codificación máxima, una unidad de codificación, una unidad de predicción, una unidad de transformación, etc.
- La Figura 11 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.
- Un analizador 510 analiza datos de imagen codificados a decodificarse e información acerca de la codificación requerida para decodificar desde una secuencia de bits 505. Los datos de imagen codificados se emiten como datos cuantificados inversos a través de un decodificador 520 por entropía y un cuantificador 530 inverso, y los datos cuantificados inversos se restauran en datos de imagen en un dominio espacial a través de un transformador 540 inverso.
- 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.
- Los datos de imagen en el dominio espacial, que se pasan a través del intra predictor 550 y el compensador 560 de movimiento, pueden emitirse como un fotograma 595 restaurado después de post-procesarse a través de una unidad 570 de desbloqueo y una unidad 580 de filtración de bucle. También, los datos de imagen que se post-procesan a través de la unidad 570 de desbloqueo y la unidad 580 de filtración 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 las operaciones que se realizan después del analizador 510.
- Para que el decodificador 500 de imagen se aplique en el aparato 200 de decodificación de vídeo, todos 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 filtración en bucle realizan operaciones basándose en unidades de codificación que tienen una estructura de árbol para cada unidad de codificación máxima.
- Específicamente, el intra predictor 550 y el compensador 560 de movimiento realizan operaciones basándose en particiones y un modo de predicción para cada una de las unidades de codificación que tienen una estructura de

árbol, y el transformador 540 inverso realiza operaciones basándose en un tamaño de una unidad de transformación para cada unidad de codificación.

En particular, el decodificador 520 por entropía puede realizar decodificación de símbolo para cada una de una cadena de bits de prefijo y una cadena de bits de sufijo clasificando la cadena de bits de símbolo analizada en la cadena de bits de prefijo y la cadena de bits de sufijo de acuerdo con un valor umbral predeterminado y usando diferentes procedimientos de decodificación aritmética y conversión a binario con respecto a la cadena de bits de prefijo y la cadena de bits de sufijo.

El valor umbral para clasificar la cadena de bits de símbolo en la cadena de bits de prefijo y la cadena de bits de sufijo puede determinarse basándose en tamaños de unidades de datos del símbolo, es decir, un corte, una unidad de codificación máxima, una unidad de codificación, una unidad de predicción, una unidad de transformación, etc.

La Figura 12 es un diagrama que muestra unidades de codificación más profundas de acuerdo con las profundidades y particiones, de acuerdo con una realización de la presente invención.

El aparato 100 de codificación de vídeo y el 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 el tamaño máximo predeterminado de la unidad de codificación.

En una estructura 600 jerárquica de unidades de codificación, de acuerdo con una realización de la presente invención, la altura máxima y la anchura máxima de las unidades de codificación son cada una 64, y la profundidad máxima es 4. En este punto, la profundidad máxima indica un número total de veces que se realiza la división desde la unidad de codificación máxima a la unidad de codificación mínima. Puesto que una profundidad se hace profunda a lo largo de un eje vertical de la estructura 600 jerárquica, se divide cada una de una altura y una anchura de la unidad de codificación más profunda. También, una unidad de predicción y 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.

En otras palabras, una unidad 610 de codificación es una unidad de codificación máxima en la estructura 600 jerárquica, en el que una profundidad es 0 y un tamaño, es decir, una altura por anchura, es 64x64. La profundidad se hace profunda a lo largo del eje vertical, y existe una unidad 620 de codificación que tiene un tamaño de 32x32 y una profundidad de 1, una unidad 630 de codificación que tiene un tamaño de 16x16 y una profundidad de 2, una unidad 640 de codificación que tiene un tamaño de 8x8 y una profundidad de 3, y una unidad 650 de codificación que tiene un tamaño de 4x4 y una profundidad de 4. La unidad 650 de codificación que tiene el tamaño de 4x4 y la profundidad de 4 es una unidad de codificación mínima.

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

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

De manera similar, una unidad de predicción de la unidad 630 de codificación que tiene el tamaño de 16x16 y la profundidad de 2 puede dividirse en particiones incluidas en la unidad 630 de codificación, es decir una partición que tiene un tamaño de 16x16 incluida en la 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.

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

La unidad 650 de codificación que tiene el tamaño de 4x4 y la profundidad de 4 es la unidad de codificación mínima y una unidad de codificación de la profundidad más inferior. Una unidad de predicción de la unidad 650 de codificación se asigna únicamente a una partición que tiene un tamaño de 4x4.

Para determinar la al menos una profundidad codificada de las unidades de codificación que constituyen la unidad 610 de codificación máxima, el determinador 120 de unidad de codificación del aparato 100 de codificación de vídeo

realiza codificación para unidades de codificación que corresponden a cada profundidad incluida en la unidad 610 de codificación máxima.

Un número de unidades de codificación más profundas de acuerdo con las profundidades que incluyen datos en el mismo intervalo y el mismo tamaño aumentan 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 están incluidos en una unidad de codificación que corresponde a una profundidad de 1. Por consiguiente, para comparar 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.

Para realizar codificación para una profundidad actual de entre las profundidades, puede seleccionarse un mínimo error de codificación para la profundidad actual realizando 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 mínimos errores de codificación de acuerdo con las profundidades, realizando 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 unidad 610 de codificación pueden seleccionarse como la profundidad codificada y un tipo de partición de la unidad 610 de codificación.

La Figura 13 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.

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

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

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

La Figura 14 es un diagrama para describir información de codificación de unidades de codificación de acuerdo con las profundidades, de acuerdo con una realización de la presente invención.

La unidad 130 de salida del aparato 100 de codificación de vídeo puede codificar y transmitir información 800 acerca de un tipo de partición, información 810 acerca de un modo de predicción, e 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.

La información 800 indica información acerca de una forma de una partición obtenida dividiendo una unidad de predicción de una unidad de codificación actual, en el que la partición es una unidad de datos para codificar por predicción 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 NxN. En este punto, la información 800 acerca de un tipo de partición se establece para indicar una de la partición 804 que tiene un tamaño de 2NxN, la partición 806 que tiene un tamaño de Nx2N, y la partición 808 que tiene un tamaño de NxN.

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

La información 820 indica una unidad de transformación a basarse cuando la transformación se realiza en una unidad de codificación actual. Por ejemplo, la unidad de transformación puede ser una primera unidad 822 de intra transformación, una segunda unidad 824 de intra transformación, una primera unidad 826 de inter transformación, o una segunda unidad 828 de inter transformación.

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

La Figura 15 es un diagrama que muestra 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.

Una unidad 910 de predicción para codificar por predicción 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 15 únicamente ilustra los tipos 912 a 918 de partición que se obtienen dividiendo simétricamente la unidad 910 de predicción, aunque un tipo de partición no está limitado a lo mismo, y las particiones de la unidad 910 de predicción pueden incluir particiones asimétricas, particiones que tienen una forma predeterminada, y particiones que tienen una forma geométrica.

La codificación por 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 pueden 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 $2N_0 \times 2N_0$.

Los errores de la codificación que incluyen la codificación por predicción en los tipos 912 a 918 de particiones se comparan, y se determina el mínimo error de codificación 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 puede no dividirse en una profundidad inferior.

Si el error de codificación es el más pequeño en el tipo 918 de partición, una profundidad se cambia de 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 codificar por predicción 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 $N_1 \times N_1$.

Si un error de codificación es el más pequeño en el tipo 948 de partición, una profundidad se cambia de 1 a 2 para dividir el tipo 948 de partición en la operación 950, y se realiza de manera repetitiva la codificación 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.

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 $d-2$. En otras palabras, cuando se realiza la 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 codificar por predicción 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 tiene 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 $N_{(d-1)} \times N_{(d-1)}$.

La codificación por predicción puede realizarse de manera repetitiva en una partición que tiene un tamaño de $2N_{(d-1)} \times 2N_{(d-1)}$, dos particiones que tienen un tamaño de $2N_{(d-1)} \times N_{(d-1)}$, dos particiones que tienen un tamaño de $N_{(d-1)} \times 2N_{(d-1)}$, cuatro particiones que tienen un tamaño de $N_{(d-1)} \times N_{(d-1)}$ 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 $CU_{(d-1)}$ que tiene una profundidad de $d-1$ ya no se divide a una profundidad inferior, y una profundidad codificada para las unidades de codificación que constituyen 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 que es $N_{(d-1)} \times N_{(d-1)}$. 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, la información de división para la unidad 980 de codificación mínima no se establece.

Una unidad 999 de datos puede ser una 'unidad mínima' para la unidad de codificación máxima actual. Una unidad mínima de acuerdo con una realización de la presente invención puede ser una unidad de datos rectangular

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

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 el mínimo error de codificación puede determinarse como una profundidad codificada. La profundidad codificada, el tipo de partición de la unidad de predicción, y el modo de predicción pueden codificarse y transmitirse como información 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 profundidades excluyendo la profundidad codificada se establece a 1.

El extractor 220 de datos de imagen e información de codificación del aparato 200 de decodificación de vídeo pueden 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 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 decodificación.

Las Figuras 16 a 18 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.

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

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

En las unidades 1060 de predicción, algunas unidades 1014, 1016, 1022, 1032, 1048, 1050, 1052, y 1054 de codificación se obtienen dividiendo las unidades de codificación en las unidades 1010 de codificación. En otras palabras, los tipos de partición en las unidades 1014, 1022, 1050, y 1054 de codificación tienen un tamaño de $2N \times N$, los tipos de partición en las unidades 1016, 1048, y 1052 de codificación tienen un tamaño de $N \times 2N$, y un tipo de partición de la unidad 1032 de codificación tiene un tamaño de $N \times N$. Las unidades de predicción y 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. En otras palabras, los aparatos 100 y 200 de codificación y decodificación de vídeo pueden realizar intra predicción, estimación de movimiento, compensación de movimiento, transformación, y transformación inversa individualmente en una unidad de datos en la misma unidad de codificación.

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 las unidades de codificación que tienen una estructura de árbol recursiva. La información de codificación puede incluir 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 la información de codificación 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 $2N \times 2N$ y profundidad actual de d)					Información de división 1
Modo de predicción	Tipo de partición		Tamaño de unidad de transformación		Codificar repetitivamente unidades de codificación que tienen profundidad inferior de $d+1$
Intra Inter	Tipo de partición de simétrica	Tipo de partición de asimétrica	Información de división 0 de unidad transformación	Información de división 1 de unidad transformación	
Salto (únicamente $2N \times 2N$)	$2N \times 2N$ $2N \times N$ $N \times 2N$ $N \times N$	$2N \times nU$ $2N \times nD$ $nL \times 2N$ $nR \times 2N$	$2N \times 2N$	$N \times N$ (tipo simétrico) $N/2 \times N/2$ (tipo asimétrico)	

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.

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, la codificación se realiza de manera independiente en cuatro unidades de codificación de división de una profundidad inferior.

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 se define únicamente en un tipo de partición que tiene un tamaño de $2N \times 2N$.

La información acerca del tipo de partición puede indicar tipos de partición simétrica que tienen tamaños de $2N \times 2N$, $2N \times N$, $N \times 2N$, y $N \times N$, que se obtienen dividiéndose simétricamente una altura o una anchura de una unidad de predicción, y tipos de partición asimétrica que tienen tamaños de $2N \times nU$, $2N \times nD$, $nL \times 2N$, y $nR \times 2N$, que se obtienen dividiendo asimétricamente la altura o anchura de la unidad de predicción. Los tipos de partición asimétrica que tienen los tamaños de $2N \times nU$ y $2N \times nD$ pueden obtenerse respectivamente dividiendo la altura de la unidad de predicción en 1:3 y 3:1, y los tipos de partición asimétrica que tienen los tamaños de $nL \times 2N$ y $nR \times 2N$ pueden obtenerse respectivamente dividiendo la anchura de la unidad de predicción en 1:3 y 3:1.

El tamaño de unidad de transformación puede establecerse para que sea dos tipos en el intra modo y dos tipos en el inter modo. En otras palabras, si la información de división de la unidad de transformación es 0, el tamaño de unidad de transformación puede ser $2N \times 2N$, 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 $2N \times 2N$ es un tipo de partición simétrica, un tamaño de una unidad de transformación puede ser $N \times N$, y si el tipo de partición de la unidad de codificación actual es un tipo de partición asimétrica, el tamaño de unidad de transformación puede ser $N/2 \times N/2$.

La información de codificación acerca de las 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 contiene la misma información de codificación.

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 profundidades codificadas en una unidad de codificación máxima.

5 Por consiguiente, si una unidad de codificación actual se predice basándose en información de codificación de unidades de datos adyacentes, puede hacerse referencia directamente a la información de codificación de unidades de datos en unidades de codificación más profundas adyacentes a la unidad de codificación actual y usarse.

10 Como alternativa, 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 puede hacerse referencia a las unidades de codificación adyacentes buscadas para predecir la unidad de codificación actual.

La Figura 19 es un diagrama para describir una relación entre una unidad de codificación, una unidad de predicción, y una unidad de transformación de acuerdo con la información de modo de codificación de la Tabla 1.

15 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 uno de 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 $nL \times 2N$, y un tipo 1338 de partición que tiene un tamaño de $nR \times 2N$.

La información de división (bandera de tamaño TU) de una unidad de transformación es una clase de un índice de transformación, y un tamaño de unidad de transformación que corresponde al índice de transformación puede variar de acuerdo con un tipo de la unidad de predicción o la partición de la unidad de codificación.

25 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 la unidad de transformación es 0, y se establece una unidad 1344 de transformación que tiene un tamaño de $N \times N$ si una bandera de tamaño de TU es 1.

30 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 una bandera de tamaño de TU es 0, y se establece una unidad 1354 de transformación que tiene un tamaño de $N/2 \times N/2$ si una bandera de tamaño de TU es 1.

35 Haciendo referencia a la Figura 19, la bandera de tamaño de TU es una bandera que tiene un valor de 0 o 1, pero la bandera de tamaño de TU no está limitada a 1 bit, y una unidad de transformación puede dividirse jerárquicamente teniendo una estructura de árbol mientras la bandera de tamaño de TU aumenta desde 0. La bandera de tamaño de TU puede usarse como una realización del índice de transformación.

40 En este caso, si la información de división de la unidad de transformación se usa junto con el tamaño de unidad de transformación máxima y el tamaño de la unidad de transformación mínima, puede expresarse el tamaño de unidad de transformación que se usa realmente. El aparato 100 de codificación de vídeo puede codificar información de tamaño de unidad de transformación máxima, información de tamaño de unidad de transformación mínima, y la información de división de unidad de transformación máxima. La información de tamaño de unidad de transformación máxima codificada, información de tamaño de unidad de transformación mínima, y la información de división de unidad de transformación máxima pueden insertarse en un SPS. El aparato 200 de decodificación de vídeo puede realizar decodificación de vídeo usando la información de tamaño de unidad de transformación máxima, la información de tamaño de unidad de transformación mínima, y la información de división de unidad de transformación máxima.

45 Por ejemplo, si una unidad de codificación actual tiene un tamaño de 64×64 y el tamaño de unidad de transformación máxima es 32×32 , cuando la información de división de unidad de transformación es 0, un tamaño de unidad de transformación puede establecerse a 32×32 , cuando la información de división de unidad de transformación es 1, el tamaño de unidad de transformación puede establecerse a 16×16 , y cuando la información de división de unidad de transformación es 2, el tamaño de unidad de transformación puede establecerse a 8×8 .

55 Como alternativa, si la unidad de codificación actual tiene un tamaño de 32×32 y el tamaño de la unidad de transformación mínima es 32×32 , cuando la información de división de unidad de transformación es 1, el tamaño de unidad de transformación puede establecerse a 32×32 , y puesto que el tamaño de unidad de transformación es igual o mayor que 32×32 , no puede establecerse más información de división de unidad de transformación.

Como alternativa, si la unidad de codificación actual tiene un tamaño de 64x64 y la información de división de unidad de transformación máxima es 1, la información de división de unidad de transformación puede establecerse a 0 o 1, y puede no establecerse otra información de división de unidad de transformación.

- 5 Por consiguiente, si la información de división de unidad de transformación máxima se define como 'MaxTransformSizeIndex', si un tamaño de unidad de transformación mínima se define como 'MinTransformSize', y si un tamaño de unidad de transformación se define como 'RootTuSize' cuando la información de división de unidad de transformación es 0, 'CurrMinTuSize' que es un tamaño de unidad de transformación mínima disponible en la unidad de codificación actual puede definirse por la fórmula (1) a continuación

CurrMinTuSize

- 10 = max (MinTransformSize, RootTuSize/(2^MaxTransformSizeIndex))... (1)

En comparación con 'CurrMinTuSize' que es el tamaño de la unidad de transformación mínima disponible en la unidad de codificación actual, 'RootTuSize' que es un tamaño de unidad de transformación cuando la información de división de unidad de transformación es 0 puede representar un tamaño de unidad de transformación máxima que puede adoptarse en un sistema. En otras palabras, de acuerdo con la fórmula (1),
15 'RootTuSize/(2^MaxTransformSizeIndex)' es un tamaño de unidad de transformación en el que 'RootTuSize' se divide un número de veces que corresponde a la información de división de unidad de transformación máxima y 'MinTransformSize' es un tamaño de unidad de transformación mínima, y por lo tanto un valor menor de entre 'RootTuSize/(2^MaxTransformSizeIndex)' y 'MinTransformSize' puede ser 'CurrMinTuSize' que es el tamaño de la unidad de transformación mínima disponible en la unidad de codificación actual.

- 20 'RootTuSize' que es el tamaño de unidad de transformación máxima puede variar de acuerdo con un modo de predicción.

Por ejemplo, si un modo de predicción actual es un inter modo, 'RootTuSize' puede determinarse de acuerdo con la fórmula (2) a continuación. En la fórmula (1), 'MaxTransformSize' indica un tamaño de unidad de transformación máxima, y 'PUSize' indica un tamaño de unidad de predicción actual.

- 25 RootTuSize = min(MaxTransformSize, PUSize).....(2)

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

- 30 Si un modo de predicción de una unidad de partición actual es un intra modo, 'RootTuSize' puede determinarse de acuerdo con la fórmula (3) a continuación. 'PartitionSize' indica un tamaño de unidad de partición actual.

RootTuSize = min(MaxTransformSize, PartitionSize)(3)

En otras palabras, si el modo de predicción actual es un intra modo, 'RootTuSize' puede establecerse a un valor menor de entre el tamaño de unidad de transformación máxima y el tamaño de unidad de partición actual.

- 35 Sin embargo, el tamaño de unidad de transformación máxima actual 'RootTuSize' que varía de acuerdo con una unidad de modo de predicción de la partición es solamente un ejemplo, y un factor para determinar el tamaño de unidad de transformación máxima actual no está limitado a lo mismo.

- 40 Los datos de imagen de un dominio espacial se codifican para cada unidad de codificación que tiene una estructura de árbol usando un procedimiento de codificación de vídeo basándose en las unidades de codificación que tienen una estructura de árbol anteriormente descritas con referencia a las Figuras 7 a 19, y la decodificación se realiza en cada unidad de codificación máxima usando un procedimiento de decodificación de vídeo basándose en las unidades de codificación que tienen una estructura de árbol, y por lo tanto los datos de imagen del dominio espacial se restauran, restaurando de esta manera un vídeo que es una instantánea y 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.

- 45 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. Ejemplos del medio de grabación legible por ordenador incluyen medio de almacenamiento magnético (por ejemplo, ROM, discos flexibles, discos duros, etc.) y medio de grabación óptico (por ejemplo, CD-ROM o DVD).

REIVINDICACIONES

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

- dividir una imagen en una pluralidad de bloques de codificación máxima basándose en la información de tamaño de los bloques de codificación máxima analizados desde un flujo de bits;
- 5 dividir jerárquicamente uno de los bloques de codificación máxima en al menos un bloque de codificación usando información de división de bloque de codificación analizado desde un flujo de bits,
- determinar al menos un bloque de transformación jerárquicamente dividido de un bloque de codificación actual usando información de división de bloque de transformación analizada desde el flujo de bits,
- analizar símbolos de un bloque de codificación actual desde el flujo de bits;
- 10 si un símbolo de prefijo y un símbolo de sufijo se clasifican desde un símbolo acerca de al menos una ubicación de coeficiente significativo de un bloque de transformación actual basándose en un valor de umbral predeterminado, realizar la decodificación aritmética basada en el contexto para el símbolo de prefijo para obtener una cadena de bits de prefijo y realizar decodificación aritmética en modo desviación para el símbolo de sufijo para obtener una cadena de bits de sufijo;
- 15 obtener un prefijo binarizado inverso realizando una binarización inversa en la cadena de bits del prefijo de acuerdo con un procedimiento de binarización truncado;
- obtener un sufijo binarizado inverso realizando una binarización inversa en la cadena de bits de sufijo de acuerdo con un procedimiento de longitud fija;
- 20 reconstruir la última ubicación del coeficiente significativo del bloque de transformación actual mediante el uso del prefijo binarizado inverso y el sufijo binarizado inverso; y
- restaurar la imagen realizando una transformación inversa en el bloque de transformación actual basándose en la última ubicación de coeficiente significativo reconstruida,
- 25 en el que la decodificación aritmética basada en el contexto en el símbolo de prefijo se realiza mediante el uso de un índice de contexto determinado en función del tamaño del bloque de transformación actual y la ubicación de un bit de la cadena de bits de prefijo,
- en el que el prefijo binarizado inverso se obtiene al decodificar el símbolo del prefijo,
- el sufijo inverso binarizado se obtiene al decodificar el símbolo del sufijo,
- la última ubicación del coeficiente significativo del bloque de transformación actual se reconstruye mediante la combinación del prefijo binarizado inverso y el sufijo binarizado inverso, y
- 30 la ubicación del último coeficiente significativo reconstruido del bloque de transformación actual indica una coordenada x de la última ubicación del coeficiente significativo en una dirección de anchos del bloque de transformación actual o una coordenada y de la última ubicación del coeficiente significativo en una dirección de alturas del bloque de transformación actual.

FIG. 1

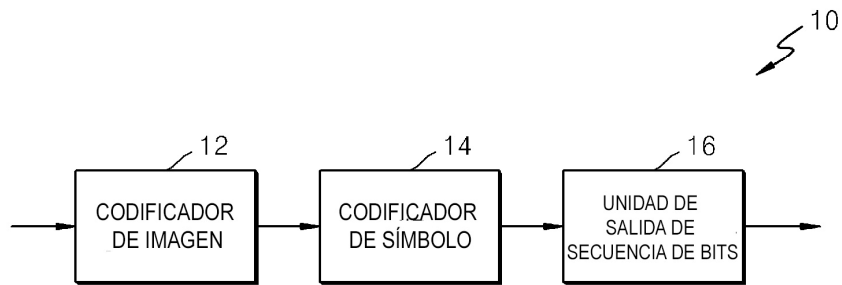


FIG. 2

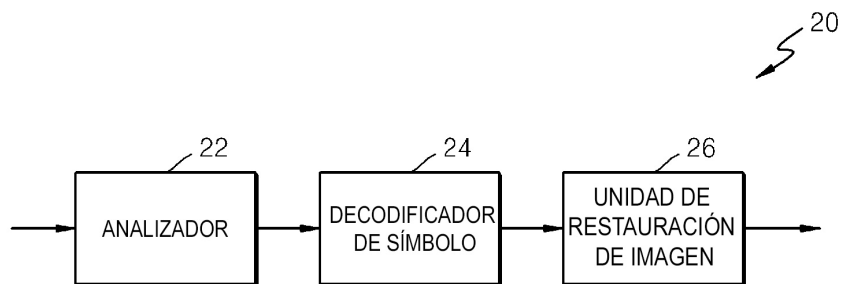


FIG. 3

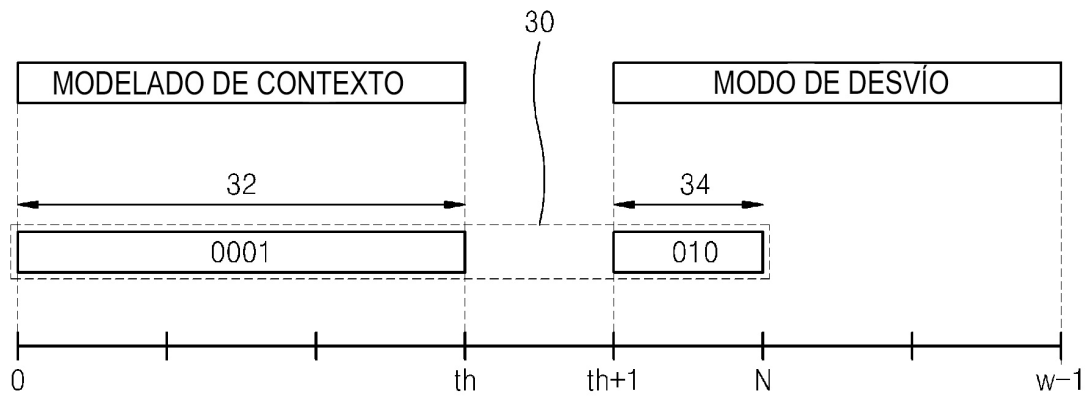


FIG. 4

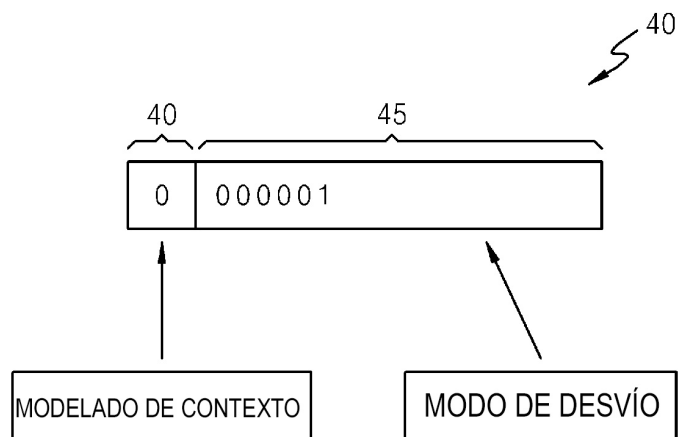


FIG. 5

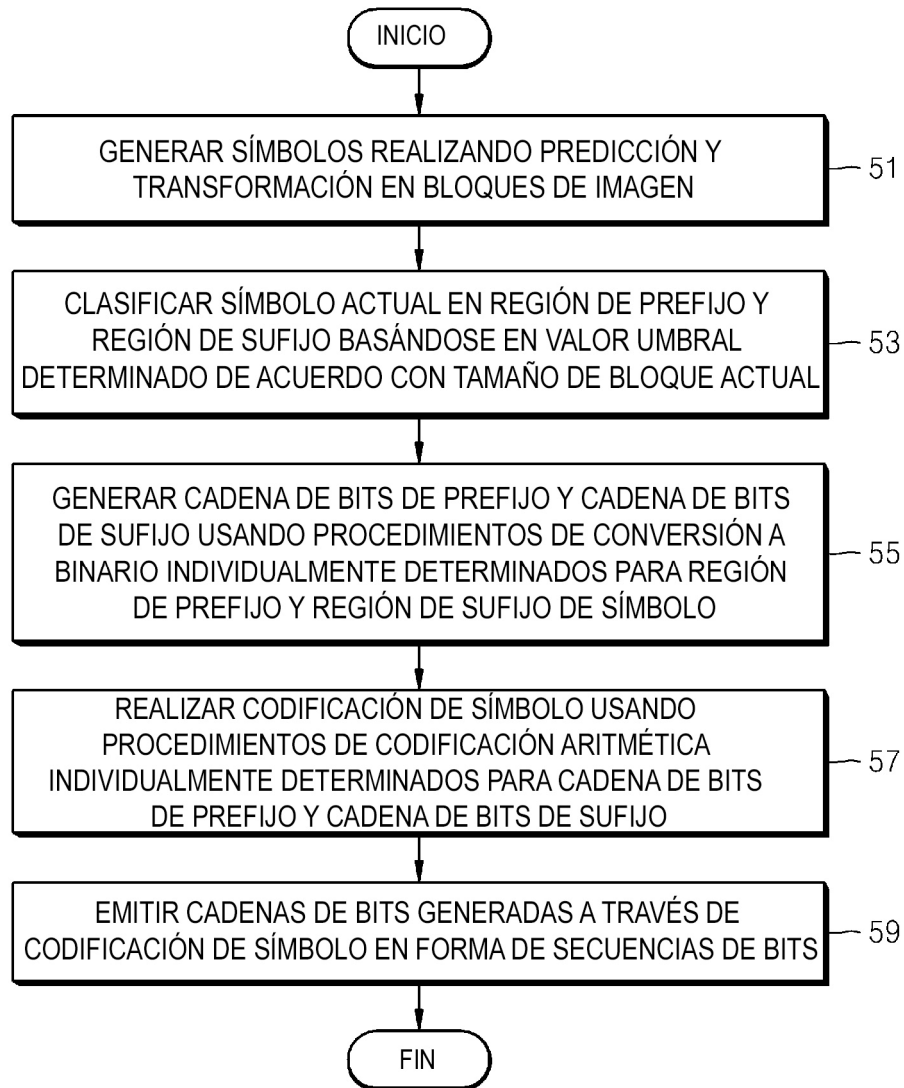


FIG. 6

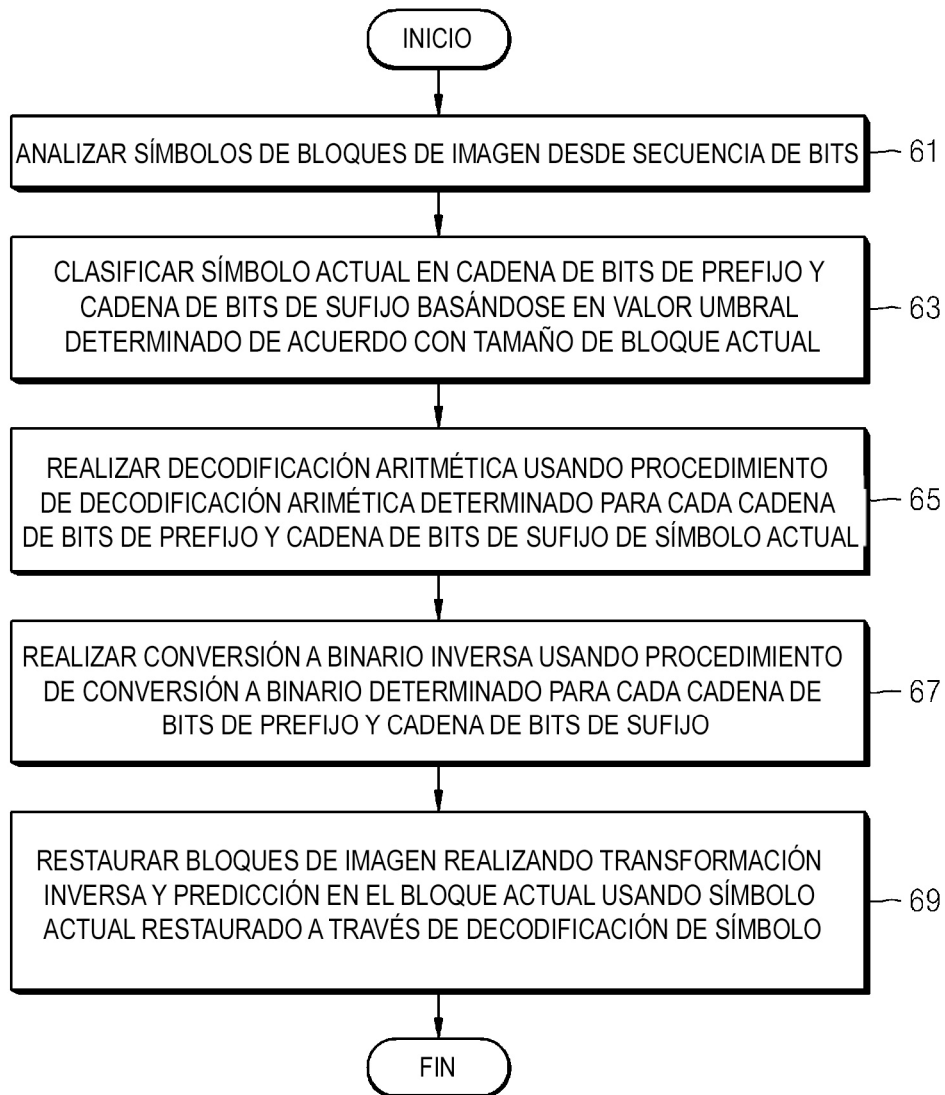


FIG. 7

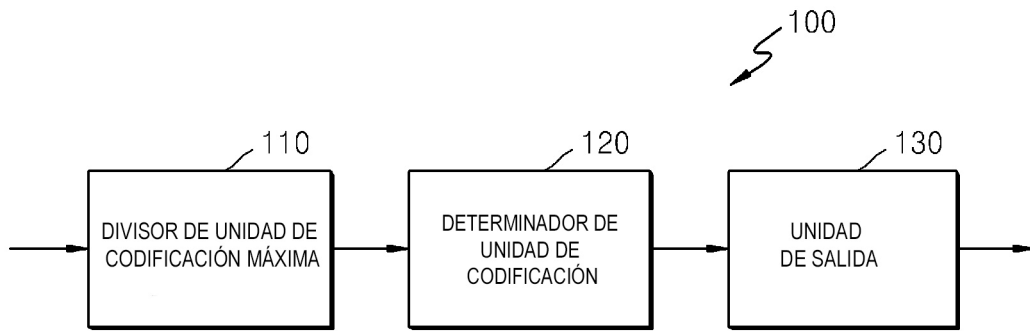


FIG. 8

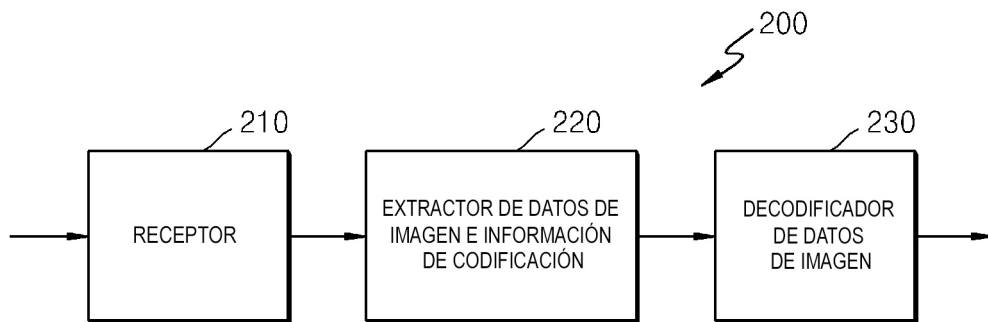


FIG. 9

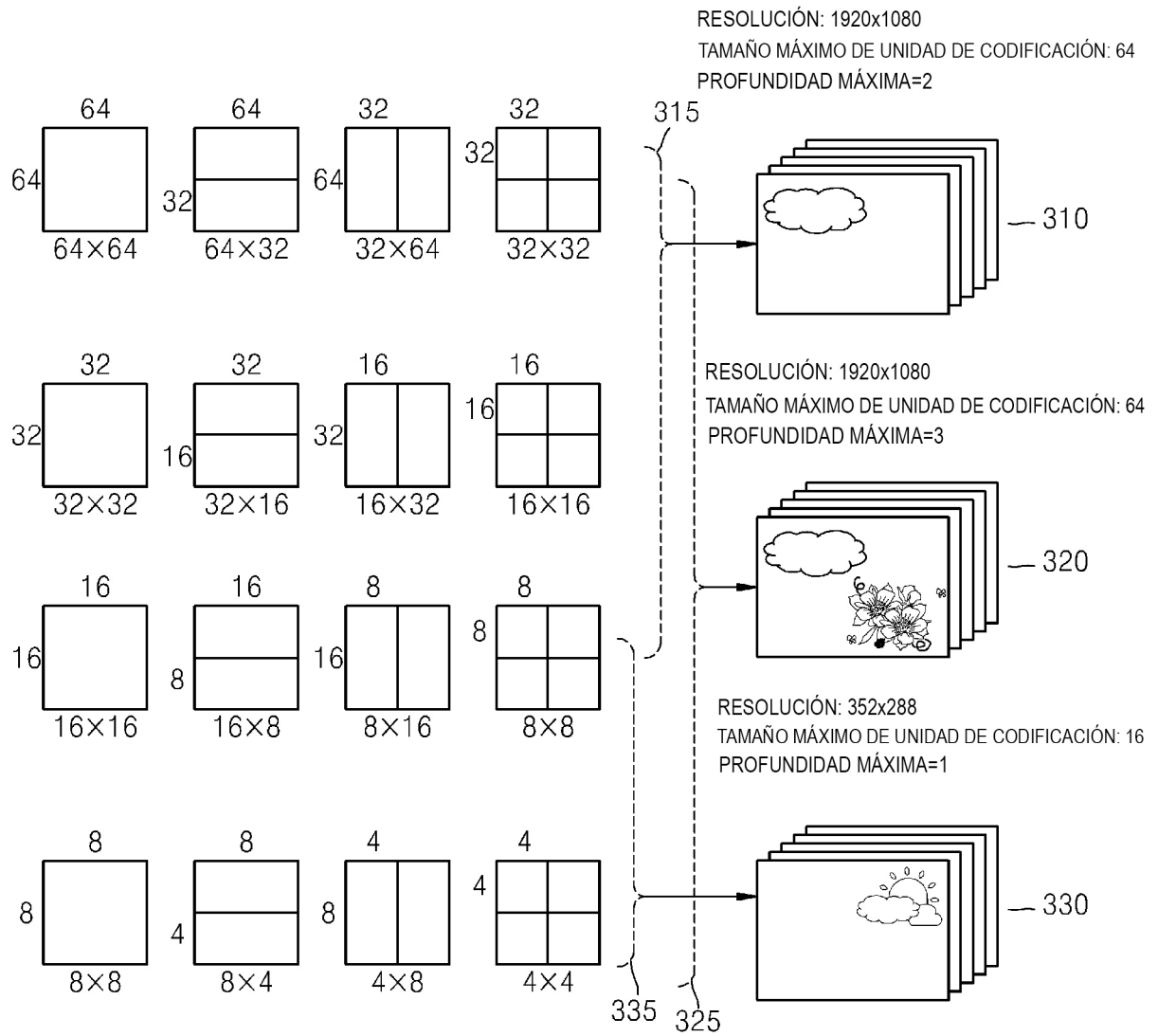


FIG. 10

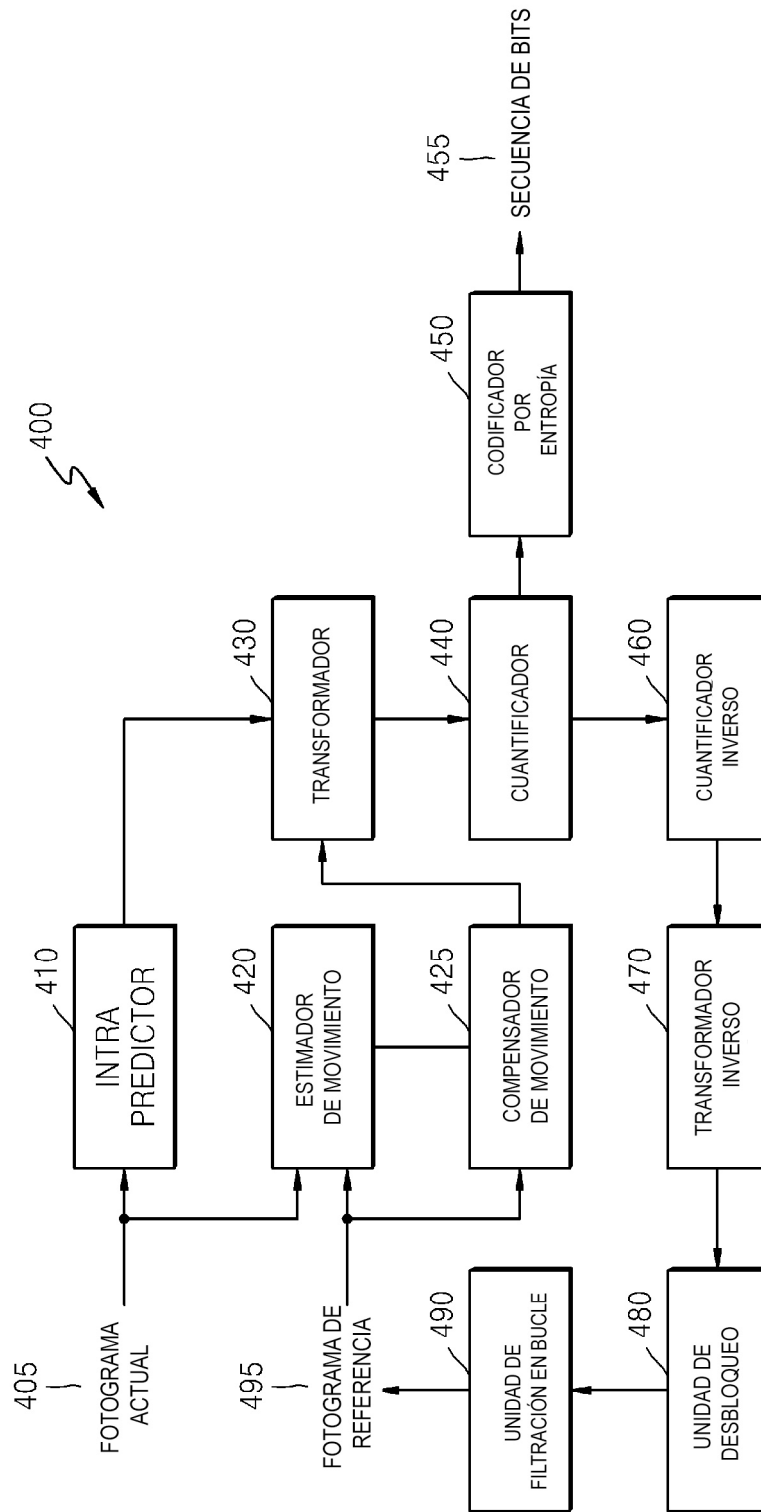


FIG. 11

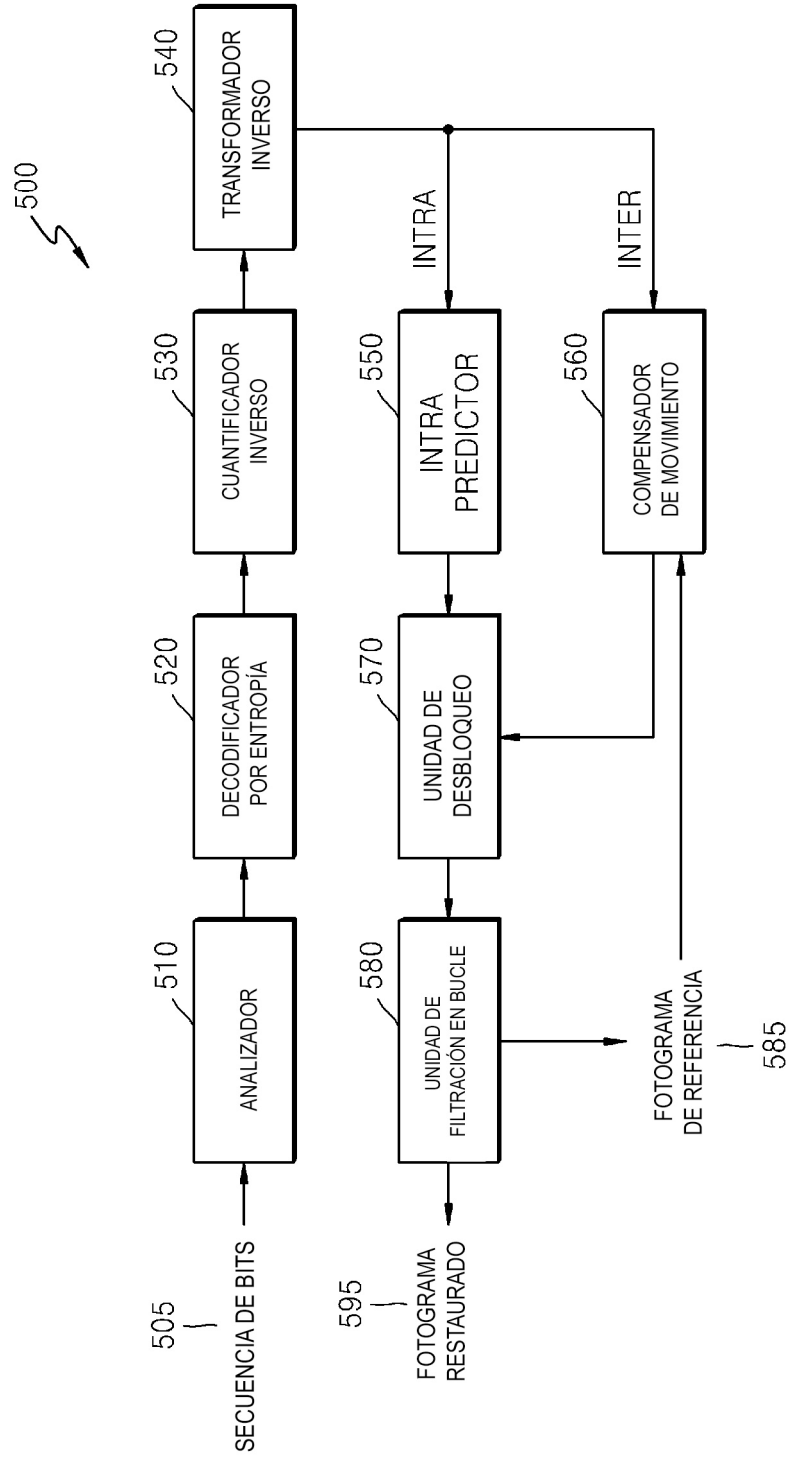


FIG. 12

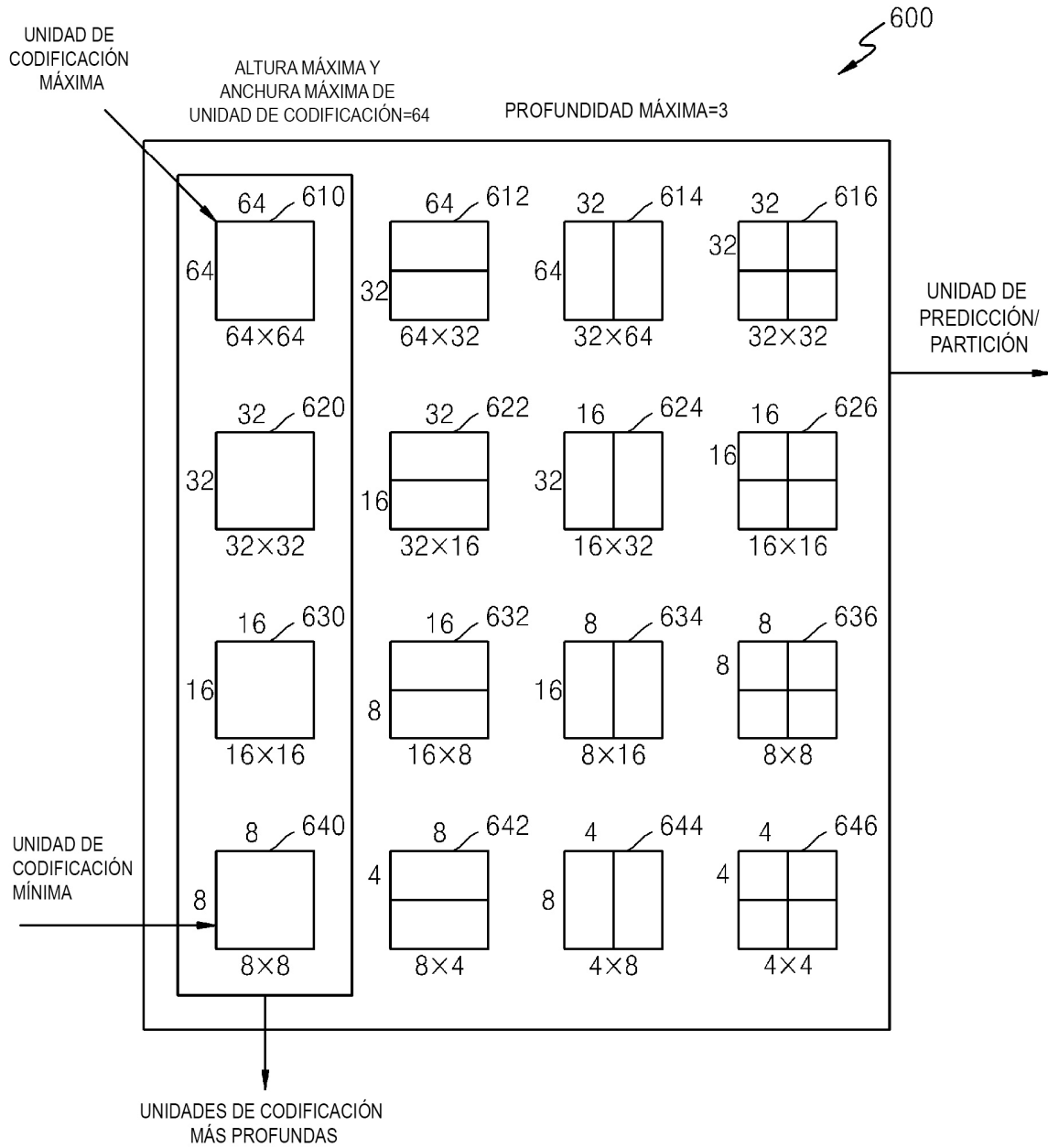


FIG. 13

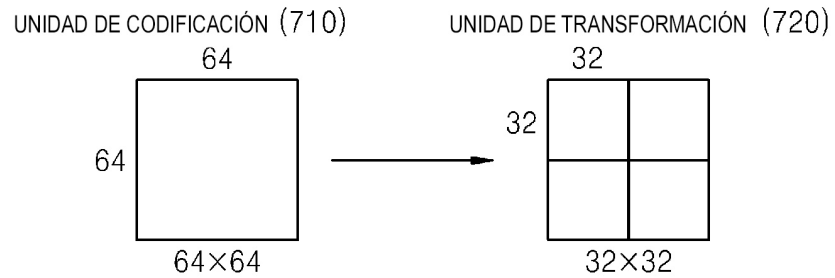


FIG. 14

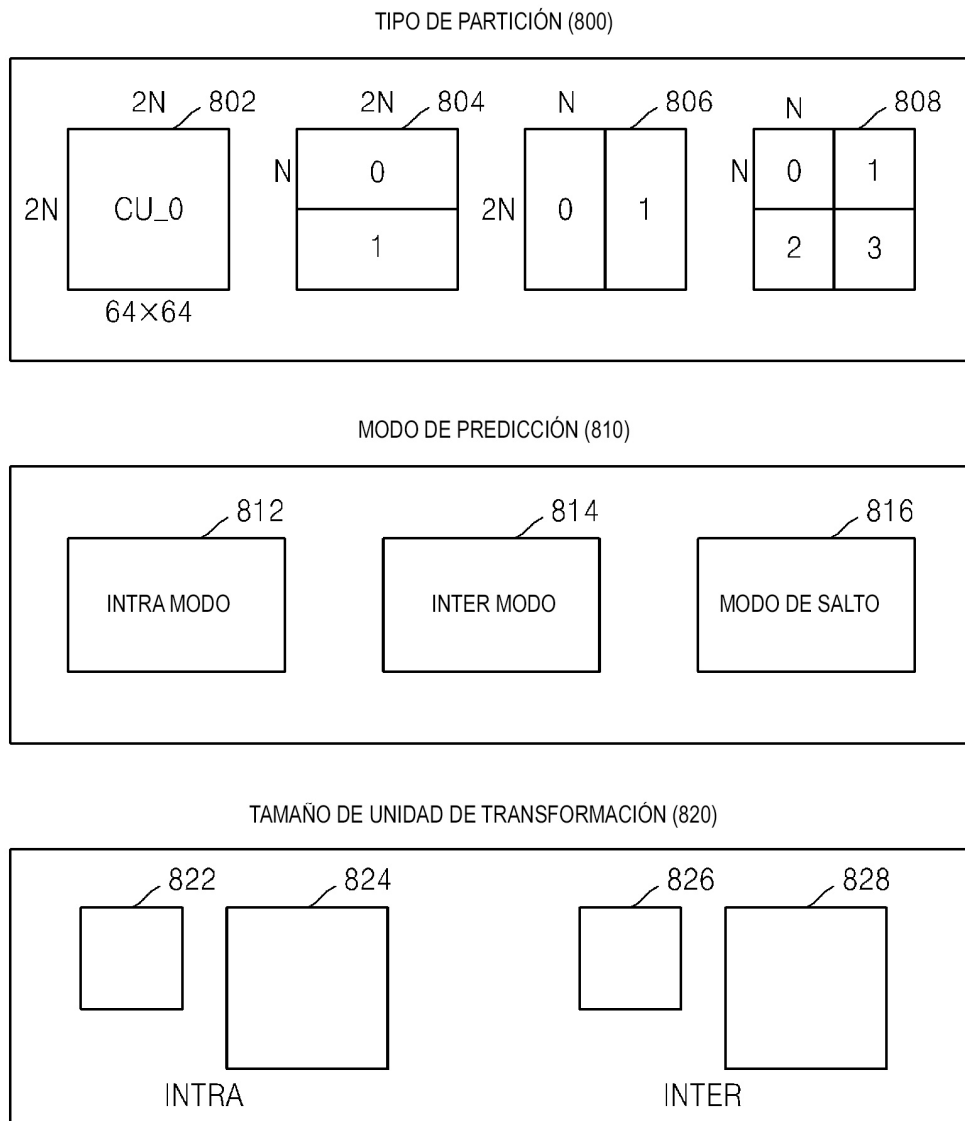


FIG. 15

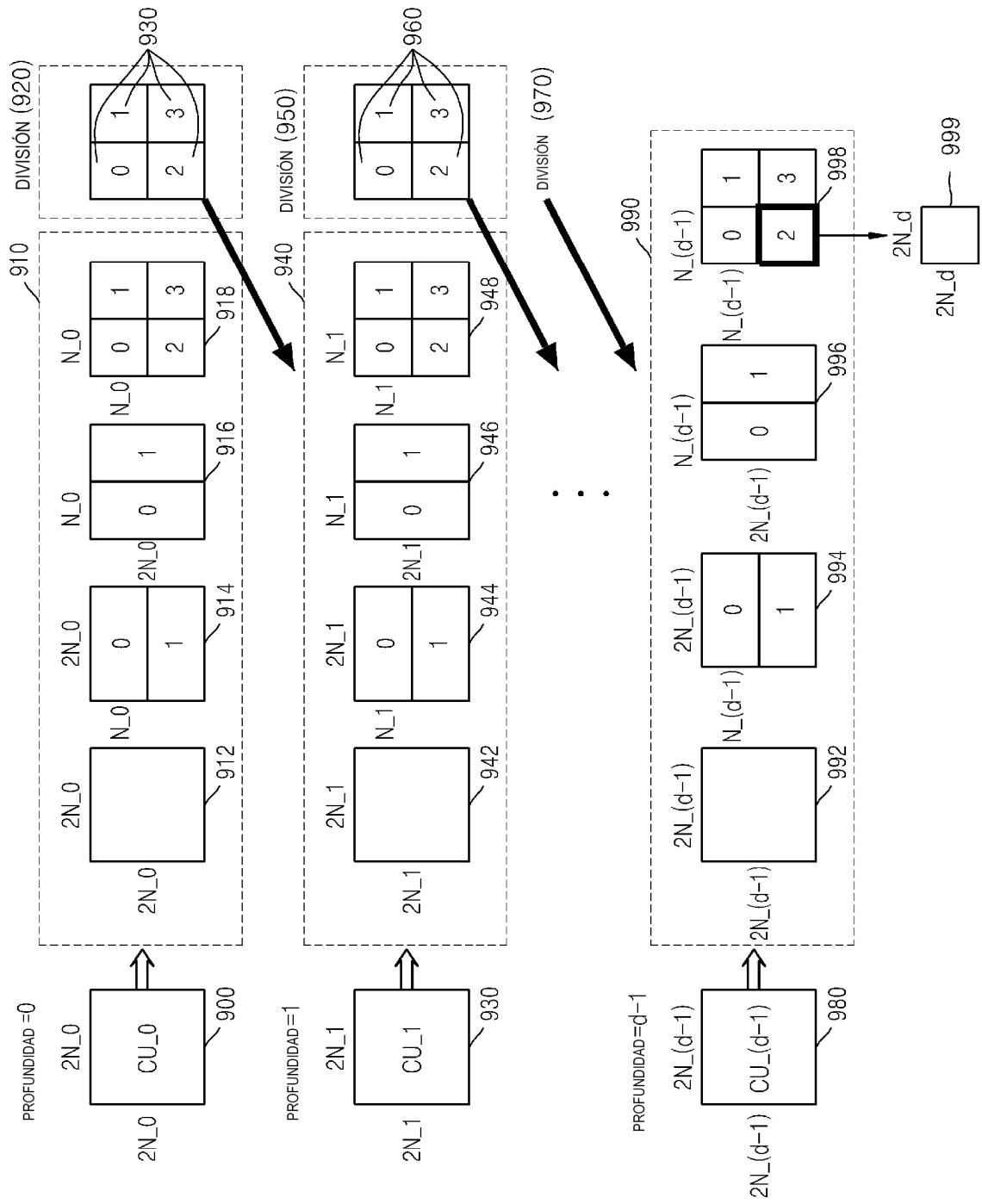
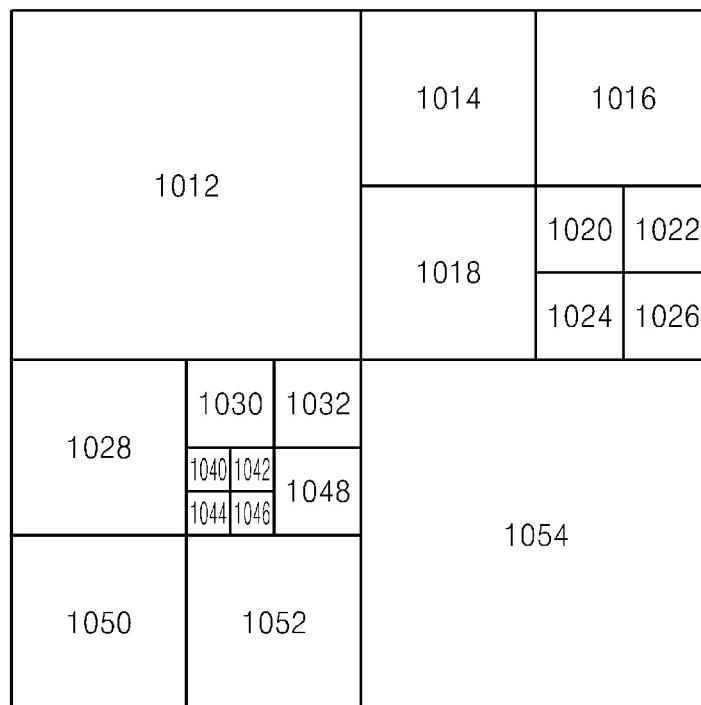


FIG. 16



UNIDAD DE CODIFICACIÓN (1010)

FIG. 17

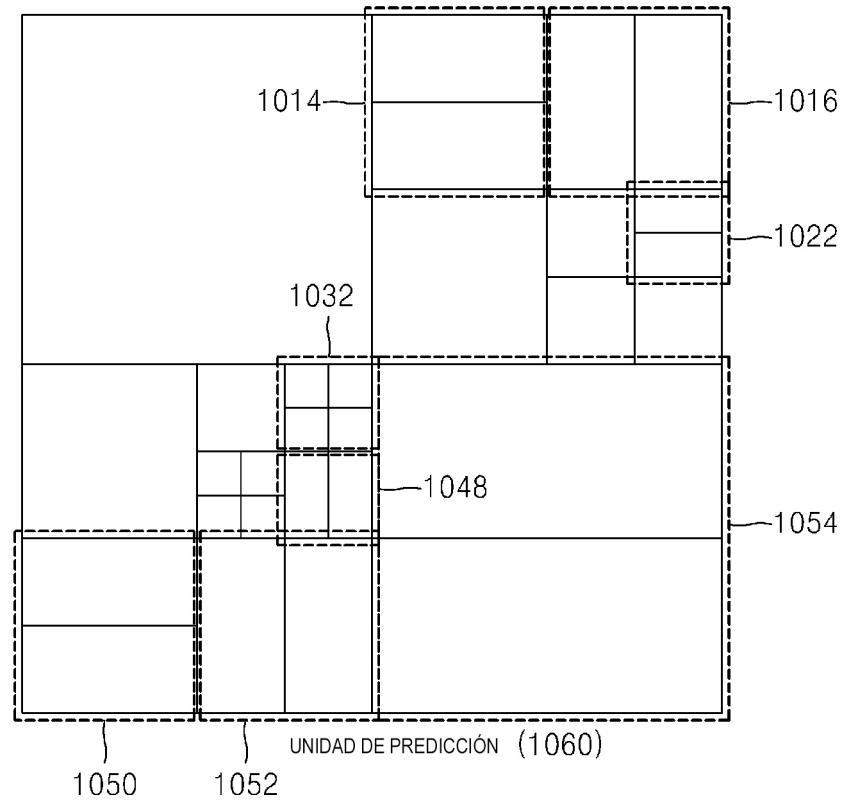


FIG. 18

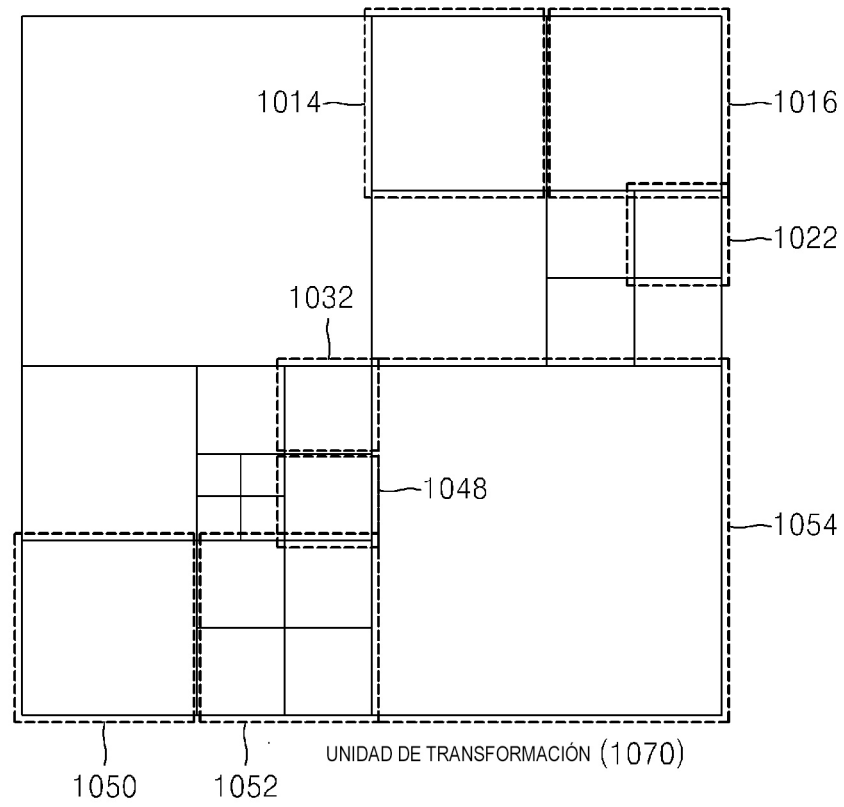


FIG. 19

