

19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 642 839**

51 Int. Cl.:

H04N 7/24 (2011.01)
H04N 21/2387 (2011.01)
H04N 21/426 (2011.01)
H04N 21/6587 (2011.01)
H04N 21/845 (2011.01)
H04N 19/503 (2014.01)
H04N 19/70 (2014.01)
H04N 19/46 (2014.01)
H04N 19/107 (2014.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

- 86 Fecha de presentación y número de la solicitud internacional: **02.07.2012** **PCT/KR2012/005256**
87 Fecha y número de publicación internacional: **10.01.2013** **WO13005969**
96 Fecha de presentación y número de la solicitud europea: **02.07.2012** **E 12806944 (0)**
97 Fecha y número de publicación de la concesión europea: **04.10.2017** **EP 2728861**

54 Título: **Procedimiento y aparato de multiplexión y demultiplexión de datos de vídeo para identificar el estado de reproducción de datos de vídeo**

30 Prioridad:

02.07.2011 US 201161504178 P
28.10.2011 US 201161552549 P

45 Fecha de publicación y mención en BOPI de la traducción de la patente:
20.11.2017

73 Titular/es:

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

72 Inventor/es:

PARK, YOUNG-O;
CHOI, KWANG-PYO;
KIM, CHAN-YUL y
YANG, HEE-CHUL

74 Agente/Representante:

CARPINTERO LÓPEZ, Mario

ES 2 642 839 T3

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

DESCRIPCIÓN

Procedimiento y aparato de multiplexión y desmultiplexión de datos de vídeo para identificar el estado de reproducción de datos de vídeo

Campo técnico

5 La presente invención se refiere a un procedimiento de decodificación de vídeo.

Antecedentes de la técnica

Un códec de vídeo que incluye las normas ITU-T H.261, ISO / IEC MPEG-1 Visual, ITU-T H.262 (ISO / IEC MPEG-2 Visual), ITU-T H.264, ISO / IEC MPEG-4 Visual e ITU-T H.264 (ISO / IEC MPEG-4 AVC) realiza una codificación de proyección sobre un macrobloque por medio de una predicción inter o una predicción intra, genera una secuencia de bits que contiene datos de imagen codificados de acuerdo con un formato previamente determinado definido por cada códec de vídeo, y emite como salida la secuencia de bits. El documento de normalización a nombre de YING CHEN ET AL, "Comments on Clean Decoding Refresh Pictures", 5. JCT-VC MEETING; 96. MPEG MEETING; 16-3-2011 - 23-3-2011; Ginebra; URL: [HTTP://WFTP3.ITU.INT/AV-ARCH/JCTVC-SITE](http://WFTP3.ITU.INT/AV-ARCH/JCTVC-SITE), n.º JCTVC-E400, describe antecedentes adicionales de la técnica.

Descripción detallada de la invención

Problema técnico

La presente invención define un nuevo tipo de una imagen intra para la reproducción de acceso aleatorio, y proporciona un procedimiento y aparato para identificar un estado de reproducción normal y un estado de reproducción de acceso aleatorio mediante soporte físico o soporte lógico relacionado con un aparato de decodificación.

Solución técnica

De acuerdo con algunas realizaciones de la presente invención, un estado de reproducción de una imagen intra puede ser identificado por medio de la sintaxis incluida en unos datos de transmisión previamente determinados.

Efectos ventajosos

25 De acuerdo con algunas realizaciones de la presente invención, un estado de reproducción normal y un estado de reproducción de acceso aleatorio pueden ser identificados por un soporte físico o soporte lógico relacionado con un aparato de decodificación, de tal modo que es posible reservar un recurso del sistema que es usado para decodificar imágenes que no requieren ser decodificadas.

Descripción de los dibujos

30 La figura 1 es un diagrama de bloques de un aparato de codificación de un vídeo.
La figura 2 es un diagrama de bloques de un aparato de decodificación de un vídeo.
La figura 3 es un diagrama para describir un concepto de unas unidades de codificación.
La figura 4 es un diagrama de bloques de un codificador de imágenes sobre la base de una codificación.
35 La figura 5 es un diagrama de bloques de un decodificador de imágenes sobre la base de unas unidades de decodificación.
La figura 6 es un diagrama que ilustra unas unidades de codificación más profundas de acuerdo con unas profundidades y unas particiones.
La figura 7 es un diagrama para describir una relación entre una unidad de codificación y unas unidades de transformación.
40 La figura 8 es un diagrama para describir una información de codificación de unas unidades de codificación que se corresponden con una profundidad codificada.
La figura 9 es un diagrama de unas unidades de codificación más profundas de acuerdo con las profundidades.
Las figuras 10 a 12 son unos diagramas para describir una relación entre las unidades de codificación, las unidades de predicción y las unidades de transformación.
45 La figura 13 es un diagrama para describir una relación entre una unidad de codificación, una unidad de predicción o una partición, y una unidad de transformación, de acuerdo con la información del modo de codificación de la tabla 1.
Las figuras 14A y 14B ilustran un proceso de decodificación o una imagen de acceso aleatorio limpia (CRA, clean random access) en una reproducción normal y un acceso aleatorio, de acuerdo con una realización de la presente invención.
50 La figura 15 ilustra una estructura de un aparato de multiplexión de datos de vídeo.
La figura 16 ilustra una estructura de una unidad de capa de abstracción de red (NAL, network abstraction layer) para su uso en una realización de la presente invención.
Las figuras 17A y 17B ilustran un proceso de decodificación de una imagen de CRA en una reproducción

normal y un acceso aleatorio.

La figura 18 es un diagrama de flujo de un procedimiento de multiplexión de datos de vídeo.

La figura 19 ilustra una estructura de un aparato de multiplexión inversa de datos de vídeo 1900, operable para realizar etapas en unos procedimientos de acuerdo con algunas realizaciones a modo de ejemplo de la presente invención.

La figura 20 es un diagrama de flujo de un procedimiento de multiplexión inversa de datos de vídeo.

Mejor modo

De acuerdo con un aspecto de la presente invención, se proporciona un procedimiento tal como se expone en la reivindicación 1 adjunta.

Modo de la invención

En lo sucesivo en el presente documento, la presente invención se describirá con detalle mediante la explicación de algunas realizaciones a modo de ejemplo de la invención con referencia a los dibujos adjuntos. A lo largo de la totalidad de la memoria descriptiva, una imagen puede incluir una imagen fija y una imagen en movimiento y se puede hacer referencia a la misma como vídeo. Asimismo, a lo largo de la totalidad de la memoria descriptiva, se puede hacer referencia a una trama de imagen como imagen.

La figura 1 es un diagrama de bloques de un aparato de codificación de vídeo 100.

El aparato de codificación de vídeo 100 incluye un divisor de unidades de codificación máxima 110, un determinador de unidades de codificación 120 y una unidad de salida 130.

El divisor de unidades de codificación máxima 110 puede dividir una imagen actual sobre la base de una unidad de codificación máxima para la imagen actual de una imagen. Si la imagen actual es más grande que la unidad de codificación máxima, los datos de imagen de la imagen actual se pueden dividir en al menos una unidad de codificación máxima. La unidad de codificación máxima puede ser una unidad de datos que tiene un tamaño de 32 x 32, 64 x 64, 128 x 128, 256 x 256, etc., en la que una forma de la unidad de datos es un cuadrado que tiene una anchura y una longitud que son, cada una, un múltiplo de 2 y mayores que 8. Los datos de imagen pueden ser enviados hacia el determinador de unidades de codificación 120 de acuerdo al menos con una unidad de codificación máxima.

Una unidad de codificación puede ser caracterizada por un tamaño máximo y una profundidad. La profundidad denota un número de veces que la unidad de codificación es espacialmente dividida a partir de la unidad de codificación máxima, y a medida que la profundidad se hace mayor, unidades de codificación más profundas de acuerdo con las profundidades pueden ser divididas a partir de la unidad de codificación máxima a una unidad de codificación mínima. Una profundidad de la unidad de codificación máxima es una profundidad más alta y una profundidad de la unidad de codificación mínima es una profundidad más baja. Debido a que un tamaño de una unidad de codificación que se corresponde con cada profundidad disminuye a medida que se hace mayor la profundidad de la unidad de codificación máxima, una unidad de codificación que se corresponde con una profundidad superior puede incluir una pluralidad de unidades de codificación que se corresponden con las profundidades más bajas.

Tal como se ha descrito en lo que antecede, los datos de imagen de la imagen actual son divididos 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 son divididas de acuerdo con las profundidades. Debido a que la unidad de codificación máxima es dividida de acuerdo con las profundidades, los datos de imagen de un dominio espacial incluido en la unidad de codificación máxima pueden ser jerárquicamente clasificados de acuerdo con las profundidades.

Una profundidad máxima y un tamaño máximo de una unidad de codificación, que limitan el número total de veces que una altura y una anchura de la unidad de codificación máxima son jerárquicamente divididas, pueden ser determinados.

El determinador de unidades de codificación 120 codifica al menos una región de división obtenida al dividir una región de la unidad de codificación máxima de acuerdo con las profundidades, y determina una profundidad para emitir como salida unos datos de imagen finalmente codificados, de acuerdo al menos con una región dividida. Dicho de otra forma, el determinador de unidades de codificación 120 determina una profundidad codificada por la codificación de 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 imagen actual, y seleccionando una profundidad que tiene el menor error de codificación. De este modo, los datos de imagen codificados de la unidad de codificación que se corresponde con la profundidad codificada determinada, son finalmente emitidos como salida. Asimismo, las unidades de codificación que se corresponden con la profundidad codificada pueden ser consideradas como unidades de codificación codificadas.

La profundidad codificada determinada y los datos de imagen codificados de acuerdo con la profundidad codificada determinada, son enviados a la unidad de salida 130.

Los datos de imagen en la unidad de codificación máxima son codificados sobre la base de las unidades de codificación más profundas que se corresponden al menos con una profundidad igual a o por debajo de la profundidad máxima, y los resultados de la codificación de los datos de imagen son comparados sobre la base de cada una de las unidades de codificación más profundas. Una profundidad que tiene el menor error de codificación puede ser seleccionada después de comparar los errores de codificación de las unidades de codificación más profundas. Al menos una profundidad codificada puede ser seleccionada para cada unidad de codificación máxima.

El tamaño de la unidad de codificación máxima es dividido debido a que una unidad de codificación es jerárquicamente dividida de acuerdo con las profundidades, y a medida que se incrementa el número de unidades de codificación. Asimismo, incluso si las unidades de codificación se corresponden con la misma profundidad en una unidad de codificación máxima, se determina si se dividen o no cada una de las unidades de codificación que se corresponden con la misma profundidad a una profundidad inferior, por la medición de un error de codificación de los datos de imagen de cada unidad de codificación, de forma separada. En consecuencia, incluso cuando los datos de imagen son incluidos en una unidad de codificación máxima, los datos de imagen son divididos en regiones de acuerdo con las profundidades y los errores de codificación pueden diferir de acuerdo con las regiones en una unidad de codificación máxima y, por lo tanto, las profundidades codificadas pueden diferir de acuerdo con las regiones en los datos de imagen. De este modo, una o más profundidades codificadas pueden ser determinadas en una unidad de codificación máxima, y los datos de imagen de la unidad de codificación máxima se pueden dividir de acuerdo con las unidades de codificación de al menos una profundidad codificada.

En consecuencia, el determinador de unidades de codificación 120 puede determinar las unidades de codificación que tienen una estructura de árbol incluida en la unidad de codificación máxima. Las "unidades de codificación que tienen una estructura de árbol" incluyen las unidades de codificación que se corresponden con una profundidad determinada por ser 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 ser jerárquicamente determinada de acuerdo con las profundidades en la misma región de la unidad de codificación máxima, y puede ser independientemente determinada en diferentes regiones. Similarmente, una profundidad codificada en una región actual puede ser independientemente determinada a partir de una profundidad codificada en otra región.

Una profundidad máxima es un índice relacionado con el número de tiempos de división de una unidad de codificación máxima a una unidad de codificación mínima. Una primera profundidad máxima puede denotar el número total de tiempos de división de la unidad de codificación máxima a la unidad de codificación mínima. Una segunda profundidad máxima puede denotar el número total de niveles de profundidad de 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 cual la unidad de codificación máxima es dividida una vez, puede ser ajustada a 1, y una profundidad de una unidad de codificación, en la cual la unidad de codificación máxima es dividida dos veces, puede ser ajustada a 2. En el presente caso, si la unidad de codificación mínima es una unidad de codificación en la cual la unidad de codificación máxima es dividida cuatro veces, existen 5 niveles de profundidad de las profundidades 1, 2, 3 y 4 y, por lo tanto, la primera profundidad máxima puede ser ajustada a 4, y la segunda profundidad máxima puede ser ajustada a 5.

La codificación de predicción y la transformación pueden ser realizadas de acuerdo con la unidad de codificación máxima. La codificación de predicción y la transformación son también realizadas sobre la base de las unidades de codificación más profundas de acuerdo con una profundidad igual a o profundidades menores que la profundidad máxima, de acuerdo con la unidad de codificación máxima. La transformación puede ser realizada de acuerdo con el procedimiento de transformación ortogonal o transformación de números enteros.

Debido a que el número de unidades de codificación más profundas se incrementa siempre que la unidad de codificación máxima es dividida de acuerdo con las profundidades, la codificación que incluye la codificación de predicción y la transformación, es realizada sobre todas las unidades de codificación más profundas generadas a medida que la profundidad se hace mayor. Para conveniencia de la descripción, la codificación de predicción y la transformación se describirá a continuación sobre la base de una unidad de codificación de una profundidad actual, en una unidad de codificación máxima.

El aparato de codificación de vídeo 100 puede seleccionar variadamente un tamaño o forma de una unidad de datos para codificar los datos de imagen. Con el fin de codificar los datos de imagen, las operaciones, tales como la codificación de predicción, la transformación y la codificación de entropía, son realizadas y esta vez la misma unidad de datos puede ser usada para todas las operaciones o diferentes unidades de datos pueden ser usadas para cada operación.

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

Con el fin de realizar la codificación de predicción en la unidad de codificación máxima, la codificación de predicción puede ser realizada sobre la base de una unidad de codificación que se corresponde con una profundidad codificada, es decir, sobre la base de una unidad de codificación que ya no es dividida a las unidades de codificación que se corresponden con una profundidad inferior. En lo sucesivo en el presente documento, se hará referencia a

5 continuación a la unidad de codificación que ya no es dividida y que se vuelve una unidad de base para la codificación de predicción como una “unidad de predicción”. Una partición obtenida por la división de la unidad de predicción puede incluir una unidad de predicción o una unidad de datos obtenida por la división de al menos una de una altura y una anchura de la unidad de predicción.

10 Por ejemplo, cuando una unidad de codificación $2N \times 2N$ (en la que N es un número entero positivo) ya no es dividida y se vuelve una unidad de predicción $2N \times 2N$, y un tamaño de una partición puede ser $2N \times 2N$, $2N \times N$, $N \times 2N$, o $N \times N$. Los ejemplos de un tipo de partición incluyen las particiones simétricas que son obtenidas al dividir simétricamente una altura o anchura de la unidad de predicción, las particiones obtenidas al dividir asimétricamente la altura o la anchura de la unidad de predicción, tal como $1 : n$ o $n : 1$, las particiones que son obtenidas al dividir geométricamente la unidad de predicción, y las particiones que tienen formas arbitrarias.

15 El modo de predicción de la unidad de predicción puede ser al menos uno de un modo intra, un modo inter o un modo de salto. Por ejemplo, el modo intra o el modo inter pueden ser realizados sobre la partición de $2N \times 2N$, $2N \times N$, $N \times 2N$, o $N \times N$. Asimismo, el modo de salto puede ser realizado solo sobre la partición de $2N \times 2N$. La codificación es independientemente realizada sobre una unidad de predicción en una unidad de codificación, con lo cual se selecciona un modo de predicción que tiene el menor error de codificación.

20 El aparato de codificación de vídeo 100 puede también realizar la transformación sobre los datos de imagen en una unidad de codificación no solo sobre la base de la unidad de codificación para codificar los datos de imagen, sino también sobre la base de una unidad de datos que es diferente de la unidad de codificación.

Con el fin de realizar la transformación en la unidad de codificación, la transformación puede ser realizada sobre la base de una unidad de datos que tiene un tamaño más pequeño que o igual a la unidad de codificación. Por

25 ejemplo, la unidad de datos para la transformación puede incluir una unidad de datos para un modo intra y una unidad de datos para un modo inter.

Se hará referencia a continuación a una unidad de datos usada como una base de la transformación como una “unidad de transformación”. Una profundidad de transformación que indica el número de tiempos de división para alcanzar la unidad de transformación por la división de la altura y la anchura de la unidad de codificación, puede

30 también ser ajustada 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 uno cuando cada una de la altura y la anchura de la unidad de codificación actual es dividida en dos partes iguales, totalmente dividida en 4^1 unidades de transformación, y el tamaño de la unidad de transformación es de este modo $N \times N$, y puede ser 2 cuando cada una de la altura y la anchura de la unidad de

35 codificación actual es dividida en cuatro partes iguales, dividida totalmente en 4^2 unidades de transformación y el tamaño de la unidad de transformación es de este modo $N/2 \times N/2$. Por ejemplo, la unidad de transformación puede ser ajustada de acuerdo con una estructura de árbol jerárquica, en la cual una unidad de transformación de una profundidad de transformación superior es dividida en cuatro unidades de transformación de una profundidad de transformación más baja de acuerdo con las características jerárquicas de una profundidad de transformación.

40 De manera similar a la unidad de codificación, la unidad de transformación en la unidad de codificación puede ser recursivamente dividida en regiones de tamaño más pequeño, de tal modo que la unidad de transformación puede ser determinada independientemente en unidades de regiones. De este modo, los datos residuales en la unidad de codificación se pueden dividir de acuerdo con la transformación que tiene la estructura de árbol de acuerdo con las profundidades de transformación.

45 La información de codificación de acuerdo con las unidades de codificación que se corresponden con una profundidad codificada, requiere no solo información con respecto a la profundidad codificada, sino también con respecto a la información relacionada con la codificación de predicción y a la transformación. En consecuencia, el determinador de unidades de codificación 120 no solo determina una profundidad codificada que tiene el menor error de codificación, sino también determina un tipo de partición en una unidad de predicción, un modo de predicción de

50 acuerdo con las unidades de predicción, y un tamaño de una unidad de transformación para la transformación.

Las unidades de codificación de acuerdo con una estructura de árbol en una unidad de codificación máxima y un procedimiento de determinación de una partición se describirán con detalle posteriormente con referencia a las figuras 3 a 12.

55 El determinador de unidades de codificación 120 puede medir un error de codificación de unidades de codificación más profundas de acuerdo con las profundidades mediante el uso de la Optimización de Distorsión de Tasa sobre la base de los multiplicadores Lagrangianos.

La unidad de salida 130 envía los datos de imagen de la unidad de codificación máxima, que se codifican sobre la base de al menos una profundidad codificada determinada por el determinador de unidades de codificación 120, y la

información con respecto al modo de codificación de acuerdo con la profundidad codificada, en secuencias de bits.

Los datos de imagen codificados pueden ser obtenidos mediante la codificación de los datos residuales de una imagen.

- 5 La información con respecto al modo de codificación de acuerdo con la profundidad codificada puede incluir la información con respecto a la profundidad codificada, con respecto al tipo de partición en la unidad de predicción, el modo de predicción y el tamaño de la unidad de transformación.

10 La información con respecto a la profundidad codificada puede ser definida mediante el uso de la información de división de acuerdo con las profundidades, que indica si la codificación es realizada o no sobre unidades de codificación de una profundidad inferior en vez de una profundidad actual. Si la profundidad actual de la unidad de codificación actual es la profundidad codificada, los datos de imagen de la unidad de codificación actual son codificados y enviados y, por lo tanto, la información de división puede ser definida no para dividir la unidad de codificación actual a una profundidad más baja. Como alternativa, si la profundidad actual de la unidad de codificación actual no es la profundidad codificada, la codificación es realizada sobre la unidad de codificación de la profundidad inferior y, por lo tanto, la información de división puede ser definida para dividir la unidad de codificación actual para obtener las unidades de codificación de la profundidad inferior.

15 Si la profundidad actual no es la profundidad codificada, la codificación es realizada sobre la unidad de codificación que es dividida en la unidad de codificación de la profundidad inferior. Debido a que al menos una unidad de codificación de la profundidad inferior existe en una unidad de codificación de la profundidad actual, la codificación es repetidamente realizada sobre cada unidad de codificación de la profundidad inferior y, por lo tanto, la codificación puede ser recursivamente realizada para las unidades de codificación que tienen la misma profundidad.

20 Debido a que las unidades de codificación que tienen una estructura de árbol son determinadas para una unidad de codificación máxima, y la información con respecto al menos a un modo de codificación es determinada para una unidad de codificación de una profundidad codificada, la información con respecto al menos a un modo de codificación puede ser determinada para una unidad de codificación máxima. Asimismo, una profundidad codificada de los datos de imagen de la unidad de codificación máxima puede ser diferente de acuerdo con las posiciones debido a que los datos de imagen son jerárquicamente divididos de acuerdo con las profundidades y, por lo tanto, la información con respecto a la profundidad codificada y el modo de codificación puede ser ajustada para los datos de imagen.

25 En consecuencia, la unidad de salida 130 puede asignar la información de codificación con respecto a una profundidad codificada correspondiente y a un modo de codificación a al menos una unidad de codificación, la unidad de predicción, y una unidad mínima incluida en la unidad de codificación máxima.

30 La unidad mínima puede ser una unidad de datos rectangular obtenida al dividir la unidad de codificación mínima que constituye la profundidad más baja entre 4, y puede ser una unidad de datos rectangular máximo que puede ser incluida en todas las unidades de codificación, las unidades de predicción, las unidades de partición y las unidades de transformación incluidas en la unidad de codificación máxima.

35 Por ejemplo, la salida de la información de codificación a través de la unidad de salida 130 puede ser clasificada en información de codificación de acuerdo con las unidades de codificación, e información de codificación de acuerdo con las unidades de predicción. La información de codificación de acuerdo con las unidades de codificación puede incluir la información con respecto al modo de predicción y con respecto al tamaño de las particiones. La información de codificación de acuerdo con las unidades de predicción puede incluir la información con respecto a una dirección estimada de un modo inter, con respecto a un índice de imagen de referencia del modo inter, con respecto a un vector de movimiento, con respecto a un componente de croma de un modo intra, y con respecto a un procedimiento de interpolación del modo intra. Asimismo, la información con respecto a un tamaño máximo de la unidad de codificación definido de acuerdo con las imágenes, sectores o GOP, y la información con respecto a una profundidad máxima puede ser insertada en un encabezamiento de una secuencia de bits.

40 El divisor de unidades de codificación máxima 110 y el determinador de unidades de codificación 120 se corresponden con las capas de codificación de vídeo que determinan una trama de referencia de cada una de las tramas de imagen que forman una secuencia de imágenes mediante la realización de la estimación de movimiento y la compensación de movimiento sobre cada trama de imagen de la secuencia de imágenes de acuerdo con las unidades de codificación, y codifican cada trama de imagen mediante el uso de la trama de referencia determinada.

45 Asimismo, tal como se describirá posteriormente, la unidad de salida 130 mapea la sintaxis (almacenamiento en memoria intermedia de max_dec_frame) por una unidad de una capa de abstracción de red (NAL, *network abstraction layer*) y, por lo tanto, genera una secuencia de bits, en la que la sintaxis indica un tamaño máximo de una memoria intermedia requerida para que un descodificador descodifique las tramas de imagen, la sintaxis (num_reorder_frames) que indican el número de tramas de imagen requeridas para ser reordenadas, y la sintaxis (max_latency_increase) que indica la formación de latencia de una trama de imagen que tiene la mayor diferencia entre un orden de codificación y un orden de visualización, y que es proveniente de las tramas de imagen que forman la secuencia de imágenes.

En el aparato de codificación de vídeo 100, la unidad de codificación más profunda puede ser una unidad de codificación obtenida al dividir una altura o anchura de una unidad de codificación de una profundidad superior, que es una capa por arriba, entre dos. Dicho de otra forma, 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 $N \times N$. Asimismo, la unidad de codificación de la profundidad actual que tiene el tamaño de $2N \times 2N$ puede incluir máximo 4 de las unidades de codificación de la profundidad inferior.

En consecuencia, el aparato de codificación de vídeo 100 puede formar las unidades de codificación que tienen la estructura de árbol al determinar las unidades de codificación que tienen una forma óptima y un tamaño óptimo para cada unidad de codificación máxima, sobre la base del tamaño de la unidad de codificación máxima y la profundidad máxima determinada considerando las características de la imagen actual. Asimismo, debido a que la codificación puede ser realizada sobre cada unidad de codificación máxima mediante el uso de cualquiera de los diversos modos de predicción y las transformaciones, un modo de codificación óptimo puede ser determinado considerando las características de la unidad de codificación de diversos tamaños de imagen.

De este modo, si una imagen que tiene alta resolución o una gran cantidad de datos es codificada en un macrobloque convencional, un número de macrobloques por imagen se incrementa de forma excesiva. En consecuencia, un número de piezas de información comprimida generada para cada macrobloque se incrementa y, por lo tanto, es difícil transmitir la información comprimida y la eficiencia de compresión de datos disminuye. No obstante, mediante el uso del aparato de codificación de vídeo 100, la eficiencia de compresión de imagen puede ser incrementada debido a que una unidad de codificación es ajustada mientras se consideran las características de una imagen al tiempo que se incrementa un tamaño máximo de una unidad de codificación, al tiempo que se considera un tamaño de la imagen.

La figura 2 es un diagrama de bloques de un aparato de descodificación de vídeo 200.

El aparato de descodificación de vídeo 200 incluye un receptor 210, un extractor de datos de imagen y de información de codificación 220, y un descodificador de datos de imagen 230. Las definiciones de los diversos términos, tal como una unidad de codificación, una profundidad, una unidad de predicción, una unidad de transformación, y la información con respecto a diversos modos de codificación, para diversas operaciones del aparato de descodificación de vídeo 200, son idénticas a aquellas descritas con referencia a la figura 1, y el aparato de codificación de vídeo 100.

El receptor 210 recibe y analiza sintácticamente una secuencia de bits de un vídeo codificado. El extractor de datos de imagen y de información de codificación 220 extrae los datos de imagen codificados para cada unidad de codificación a partir de la secuencia de bits analizada sintácticamente, en la que las unidades de codificación tienen una estructura de árbol de acuerdo con cada unidad de codificación máxima, y envía los datos de imagen extraídos al descodificador de datos de imagen 230. El extractor de datos de imagen y de información de codificación 220 puede extraer la información con respecto a un tamaño máximo de una unidad de codificación de una imagen actual, a partir de un encabezamiento con respecto a la imagen actual o SPS.

Asimismo, el extractor de datos de imagen y de información de codificación 220 extrae la información con respecto a una profundidad codificada y a un modo de codificación para las unidades de codificación que tienen una estructura de árbol de acuerdo con cada unidad de codificación máxima, a partir de la secuencia de bits analizada sintácticamente. La información extraída con respecto a la profundidad codificada y al modo de codificación, es enviada al descodificador de datos de imagen 230. Dicho de otra forma, los datos de imagen en una secuencia de bits son divididos en la unidad de codificación máxima, de tal modo que el descodificador de datos de imagen 230 descodifica los datos de imagen para cada unidad de codificación máxima.

La información con respecto a la profundidad codificada y al modo de codificación de acuerdo con la unidad de codificación máxima puede ser ajustada para la información con respecto al menos a una unidad de codificación que se corresponde con la profundidad codificada, y la información con respecto a un modo de codificación puede incluir la información con respecto a un tipo de partición de una unidad de codificación correspondiente, que se corresponde con la profundidad codificada con respecto a un modo de predicción, y un tamaño de una unidad de transformación. Asimismo, la información de división de acuerdo con las profundidades puede ser extraída como la información con respecto a la profundidad codificada.

La información con respecto a la profundidad codificada y el modo de codificación de acuerdo con cada unidad de codificación máxima extraída por el extractor de datos de imagen y de información de codificación 220, es información con respecto a una profundidad codificada y un modo de codificación determinado para generar un error de codificación mínimo cuando un codificador, tal como el aparato de codificación de vídeo 100 realiza repetidamente la codificación para cada unidad de codificación más profunda de acuerdo con las profundidades de acuerdo con cada unidad de codificación máxima. En consecuencia, el aparato de descodificación de vídeo 200 puede restablecer una imagen al descodificar los datos de imagen de acuerdo con una profundidad codificada y un modo de codificación que genera el error de codificación mínimo.

Debido a que la información de codificación con respecto a la profundidad codificada y el modo de codificación puede ser asignado a una unidad de datos previamente determinada de entre una unidad de codificación correspondiente, una unidad de predicción, y una unidad mínima, el extractor de datos de imagen y de información de codificación 220 puede extraer la información con respecto a la profundidad codificada y al modo de codificación de acuerdo con las unidades de datos previamente determinadas. Las unidades de datos previamente determinadas a las cuales es asignada la misma información con respecto a la profundidad codificada y al modo de codificación, pueden ser inferidas para ser las unidades incluidas en la misma unidad de codificación máxima.

El descodificador de datos de imagen 230 restablece la imagen actual al descodificar los datos de imagen en cada unidad de codificación máxima sobre la base de la información con respecto a la profundidad codificada y al modo de codificación de acuerdo con las unidades de codificación máxima. Dicho de otra forma, el descodificador de datos de imagen 230 puede descodificar los datos de imagen codificados sobre la base de la información extraída con respecto al tipo de partición, al modo de predicción, y la unidad de transformación para cada unidad de codificación de entre las unidades de codificación que tienen la estructura de árbol incluida en cada unidad de codificación máxima. Un proceso de descodificación puede incluir una predicción que incluye la predicción intra y la compensación de movimiento, y una transformación inversa. La transformación inversa puede ser realizada de acuerdo con el procedimiento de transformación ortogonal inversa o transformación de número interno inverso.

El descodificador de datos de imagen 230 puede realizar la predicción intra o la compensación de movimiento de acuerdo con una partición y a un modo de predicción de cada unidad de codificación, sobre la base de la información con respecto al tipo de partición y al modo de predicción de la unidad de predicción de la unidad de codificación de acuerdo con las profundidades codificadas.

Asimismo, el descodificador de datos de imagen 230 puede realizar la transformación inversa de acuerdo con cada unidad de transformación en la unidad de codificación, sobre la base de la información con respecto al tamaño de la unidad de transformación de la unidad de codificación de acuerdo con las profundidades codificadas, para realizar así la transformación inversa de acuerdo con las unidades de codificación máxima.

El descodificador de datos de imagen 230 puede determinar al menos una profundidad codificada de una unidad de codificación máxima actual mediante el uso de la información de división de acuerdo con las profundidades. Si la información de división indica que los datos de imagen ya no son divididos en la profundidad actual, la profundidad actual es una profundidad codificada. En consecuencia el descodificador de datos de imagen 230 puede descodificar los datos codificados de al menos una unidad de codificación que se corresponde con cada profundidad codificada en la unidad de codificación máxima actual, mediante el uso de la información con respecto al tipo de partición de la unidad de predicción, el modo de predicción, y el tamaño de la unidad de transformación para cada unidad de codificación que se corresponde con la profundidad codificada, y la salida de los datos de imagen de la unidad de codificación máxima actual.

Es decir, las unidades de datos que contienen la información de codificación que incluye la misma información de división puede ser obtenida por la observación de la información de codificación establecida, asignada para la unidad de datos previamente determinada de entre la unidad de codificación, la unidad de predicción y la unidad mínima, y las unidades de datos obtenidas pueden ser consideradas por ser una unidad de datos para ser codificada por el descodificador de datos de imagen 230 en el mismo modo de codificación.

Asimismo, el receptor 210 y el extractor de datos de imagen y de información de codificación 220 realizan un proceso de descodificación NAL en el cual la sintaxis (almacenamiento en memoria intermedia de max_dec_frame) que indica un tamaño máximo de una memoria intermedia requerida para que un descodificador descodifique las tramas de imagen, la sintaxis (num_reorder_frames) que indica el número de tramas de imagen requeridas para ser reordenadas, y sintaxis (max_latency_increase) que indica la información de latencia de una trama de imagen que tiene la mayor diferencia entre un orden de codificación y un orden de visualización, y que es proveniente de las tramas de imagen que forman una secuencia de imágenes, son obtenidas de una secuencia de bits y son enviadas al descodificador de datos de imagen 230.

El aparato de descodificación de vídeo 200 puede obtener la información con respecto al menos a una unidad de codificación que genera el error de codificación mínima cuando la codificación es recursivamente realizada para cada unidad de codificación máxima, y puede usar la información para descodificar la imagen actual. Dicho de otra forma, las unidades de codificación que tienen la estructura de árbol determinada para hacer las unidades de codificación óptimas en cada unidad de codificación máxima, pueden ser descodificadas. Asimismo, el tamaño máximo de la unidad de codificación es determinado considerando la resolución y una cantidad de datos de imagen.

En consecuencia, incluso si los datos de imagen tienen una alta resolución y una gran cantidad de datos, los datos de imagen pueden ser eficientemente descodificados y restablecidos mediante el uso de un tamaño de una unidad de codificación y un modo de codificación, que son determinados de forma adaptativa de acuerdo con las características de los datos de imagen, mediante el uso de información con respecto a un modo de codificación óptimo recibido de un codificador.

Un procedimiento de determinar las unidades de codificación que tienen una estructura de árbol, una unidad de predicción, y una unidad de transformación se describirá a continuación con referencia a las figuras 3 a 13.

La figura 3 es un diagrama para describir un concepto de unas unidades de codificación.

5 Un tamaño de una unidad de codificación puede ser expresado en anchura por altura, y puede ser 64 x 64, 32 x 32, 16 x 16, y 8 x 8. Una unidad de codificación de 64 x 64 puede ser dividida en particiones de 64 x 64, 64 x 32, 32 x 64, o 32 x 32, y una unidad de codificación 32 x 32 puede ser dividida en particiones de 32 x 32, 32 x 16, 16 x 32, o 16 x 16, una unidad de codificación de 16 x 16 puede ser dividida en particiones de 16 x 16, 16 x 8, 8 x 16, o 8 x 8, y una unidad de codificación de 8 x 8 puede ser dividida en particiones de 8 x 8, 8 x 4, 4 x 8, o 4 x 4.

10 En los datos de vídeo 310, una resolución es 1920 x 1080, un tamaño máximo de una unidad de codificación es 64, y una profundidad máxima es 2. En los datos de vídeo 320, una resolución es 1920 x 1080, un tamaño máximo de una unidad de codificación es 64, y una profundidad máxima es 3. En los datos de vídeo 330, una resolución es 352 x 288, un tamaño máximo de una unidad de codificación 16, y una profundidad máxima es 1. La profundidad máxima mostrada en la figura 3 denota un número total de divisiones a partir de una unidad de codificación máxima a una unidad de descodificación mínima.

15 Si una resolución es alta o una cantidad de datos es grande, un tamaño máximo de una unidad de codificación puede ser grande no solo para incrementar la eficiencia de codificación, sino también para reflejar de manera precisa las características de una imagen. En consecuencia, el tamaño máximo de la unidad de codificación de los datos de vídeo 310 y 320 que tienen una resolución más alta que los datos de vídeo 330, puede ser 64.

20 Debido a que la profundidad máxima de los datos de vídeo 310 es 2, las unidades de codificación 315 de los datos de vídeo 310 pueden incluir una unidad de codificación máxima que tiene un tamaño de eje largo de 64, y unidades de codificación que tienen tamaños de eje largo de 32 y 16 debido a que las profundidades son aumentadas a dos capas por la división de la unidad de codificación máxima dos veces. Mientras tanto, debido a que la profundidad máxima de los datos de vídeo 330 es 1, las unidades de codificación 335 de los datos de vídeo 330 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 debido a que las profundidades son aumentadas a una capa por la división de la unidad de codificación máxima una vez.

30 Debido a que la profundidad máxima de los datos de vídeo 320 es 3, las unidades de codificación 325 de los datos de vídeo 320 pueden incluir una unidad de codificación máxima que tiene un tamaño de eje largo de 64, y las unidades de codificación que tienen tamaños de eje largo de 32, 16 y 8 debido a que las profundidades son aumentadas a 3 capas por la división de la unidad de codificación máxima tres veces. A medida que una profundidad aumenta, la información detallada puede ser más precisamente expresada.

La figura 4 es un diagrama de bloques de un codificador de imágenes 400 sobre la base de unas unidades de codificación.

35 El codificador de imágenes 400 realiza las operaciones del determinador de unidades de codificación 120 del aparato de codificación de vídeo 100 para codificar los datos de imagen. Dicho de otra forma, un predictor intra 410 realiza la predicción intra sobre las unidades de codificación en un modo intra, de entre una trama actual 405 y un estimador de movimiento 420 y un compensador de movimiento 425 realiza la estimación inter y la compensación de movimiento sobre las unidades de codificación en un modo inter de entre la trama actual 405 mediante el uso de la trama actual 405 y una trama de referencia 495.

40 Los datos enviados desde el predictor intra 410, el estimador de movimiento 420 y el compensador de movimiento 425 son enviados como un coeficiente de transformación cuantificado a través de un transformador 430 y un cuantificador 440. El coeficiente de transformación cuantificado es restablecido como datos en un dominio espacial a través de un cuantificador inverso 460 y un transformador inverso 470, y los datos restablecidos en el dominio espacial son enviados como la trama de referencia 495 después de ser post-procesados a través de una unidad de desbloqueo 480 y una unidad de filtración de bucle 490. El coeficiente de transformación cuantificado puede ser enviado como una secuencia de bits 455 a través de un codificador de entropía 450. En particular, el codificador de entropía 450 puede mapear las sintaxis de almacenamiento en memoria intermedia de trama descodificada máxima (almacenamiento en memoria intermedia de max_dec_frame) por una unidad de un NAL y, por lo tanto, puede generar una secuencia de bits, en la que la sintaxis de almacenamiento en memoria intermedia de trama descodificada máxima (almacenamiento en memoria intermedia de max_dec_frame) indica un tamaño máximo de una memoria intermedia requerida para que un descodificador descodifique las tramas de imagen, la sintaxis del número de tramas de reordenación (num_reorder_frames) indicando el número de las tramas de imagen requeridas para ser reordenadas, las sintaxis de trama de latencia máxima (MaxLatencyFrame) indicando un valor máximo de un valor de diferencia entre un orden de codificación y un orden de codificación y un orden de visualización de las tramas de imagen que forman una secuencia de imágenes, o las sintaxis de incremento de latencia máxima (max_latency_increase) para determinar la sintaxis de trama de latencia máxima (MaxLatencyFrame). En particular, el codificador de entropía 450 de acuerdo con la presente realización incluye la sintaxis de memoria intermedia de trama descodificada máxima (almacenamiento en memoria intermedia de max_dec_frame) que indica un tamaño

máximo de una memoria intermedia requerida para que un decodificador descodifique las tramas de imagen, la sintaxis del número de tramas de reordenación (`num_reorder_frames`) que indica el número de las tramas de imagen requeridas para ser reordenadas, y la sintaxis de incremento de latencia máxima (`max_latency_increase`) para determinar la sintaxis de trama de latencia máxima (`MaxLatencyFrame`), como elementos esenciales, en un conjunto de parámetros de secuencia (SPS, *sequence parameter set*) que es la información de encabezamiento que incluye la información relacionada con la codificación de la secuencia de imagen completa.

Con el fin de que el codificador de imágenes 400 sea aplicado en el aparato de codificación de vídeo 100, todos los elementos del codificación de imágenes 400, es decir, el predictor intra 410, el estimador de movimiento 420, el compensador de movimiento 425, el transformador 430, el cuantificador 440, el codificador de entropía 450, el cuantificador inverso 460, el transformador inverso 470, la unidad de desbloqueo 480 y la unidad de filtración de bucle 490 realizan operaciones sobre la base de 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 predictor intra 410, el estimador de movimiento 420 y el compensador de movimiento 425 determina las 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 la unidad de transformación en cada unidad de codificación de entre las unidades de codificación que tienen una estructura de árbol.

La figura 5 es un diagrama de bloques de un decodificador de imágenes 500 sobre la base de unas unidades de codificación.

Un analizador sintáctico 510 analiza sintácticamente los datos de imagen codificados que van a ser descodificados, y la información con respecto a la codificación requerida para la descodificación a partir de una secuencia de bits 505. En particular, el analizador sintáctico 510 obtiene las sintaxis de almacenamiento en memoria intermedia de trama descodificada máxima (almacenamiento en memoria intermedia de `max_dec_frame`) que indica un tamaño máximo de una memoria intermedia requerida para descodificar las tramas de imagen incluidas como un elemento esencial en un SPS, la sintaxis número-de-tramas-de-reordenación (`num_reorder_frames`) que indica el número de las tramas de imagen requeridas para ser reordenadas, y las sintaxis de incremento de latencia máxima (`max_latency_increase`) para determinar una sintaxis de trama de latencia máxima (`MaxLatencyFrame`) a partir de una secuencia de bits, y las envía a un decodificador de entropía 520. En la figura 5, el analizador sintáctico 510 y el decodificador de entropía 520 son elementos separados. No obstante, la obtención de los datos de imagen y la obtención de cada elemento de información de sintaxis relacionada con los datos de imagen codificados que son realizados por el analizador sintáctico 510, pueden ser implementadas para ser realizadas por el decodificador de entropía 520.

Los datos de imagen codificados son enviados como datos cuantificados de forma inversa a través del decodificador de entropía 520 y un cuantificador inverso 530, y los datos cuantificados de forma inversa son restablecidos a los datos de imagen en un dominio espacial a través de un transformador inverso 540.

Un predictor intra 550 realiza la predicción intra sobre las unidades de codificación en un modo intra con respecto a los datos de imagen en el dominio espacial, y un compensador de movimiento 560 realiza la compensación de movimiento sobre las unidades de codificación en un modo inter mediante el uso de una trama de referencia 585.

Las tramas de imagen que son restablecidas mientras pasan a través del predictor intra 550 y el compensador de movimiento 560 pueden ser post-procesadas a través de la unidad de desbloqueo 570 y pueden ser enviadas a una memoria intermedia de imágenes descodificadas (DPB, *decoded picture buffer*) 580. El DPB 580 almacena una trama de referencia, cambia un orden de visualización de las tramas de imagen, y almacena las tramas de imagen restablecidas, para enviar las tramas de imagen. El DPB 580 almacena las tramas de imagen restablecidas, y ajusta un tamaño máximo de una memoria intermedia requerida para descodificar normalmente la secuencia de imágenes, mediante el uso de la sintaxis de almacenamiento en memoria intermedia de trama descodificada máxima (almacenamiento en memoria intermedia de `max_dec_frame`) que indica el tamaño máximo de la memoria intermedia requerida para descodificar las tramas de imagen, que es emitida como salida desde el analizador sintáctico 510 o el decodificador de entropía 520.

Asimismo, el DPB 580 puede determinar si envía o no una trama de imagen de referencia que es previamente descodificada y almacenada, mediante el uso de la sintaxis número-de-tramas-de-reordenación (`num_reorder_frames`) que indica el número de las tramas requeridas para ser reordenadas, y la sintaxis de incremento de latencia máxima (`max_latency_increase`) para determinar la sintaxis de trama de latencia máxima (`MaxLatencyFrame`). Un proceso de envío de la trama de imagen de referencia almacenada en el DPB 580 se describirá con detalle posteriormente.

Con el fin de descodificar los datos de imagen en el decodificador de datos de imagen 230 del aparato de descodificación de vídeo 200, el decodificador de imágenes 500 pueda realizar las operaciones que son realizadas

después del analizador sintáctico 510.

Con el fin de que el descodificador de imágenes 500 sea aplicado en el aparato de descodificación de vídeo 200, todos los elementos del descodificador de imágenes 500, es decir, el analizador sintáctico 510, el descodificador de entropía 520, el cuantificador inverso 530, el transformador inverso 540, el predictor intra 550, el compensador de movimiento 560, la unidad de desbloqueo 570 y la unidad de filtración de bucle 580 pueden realizar las operaciones de descodificación sobre la base de las unidades de codificación que tienen una estructura de árbol para cada unidad de codificación máxima. Específicamente, la predicción intra 550 y el compensador de movimiento 560 pueden determinar las particiones y un modo de predicción para cada una de las unidades de codificación que tienen una estructura de árbol, y el transformador inverso 540 puede determinar un tamaño de una unidad de transformación para cada unidad de codificación.

La figura 6 es un diagrama que ilustra unas unidades de codificación más profundas de acuerdo con unas profundidades y unas particiones.

En aparato de codificación de vídeo 100 y el aparato de descodificación de vídeo 200 usan unidades de codificación jerárquica para considerar las características de una imagen. Una altura máxima, una anchura máxima y una profundidad máxima de las unidades de codificación pueden ser determinadas de forma adaptativa de acuerdo con las características de la imagen, o pueden ser ajustadas de manera diferente por un usuario. Los tamaños de las unidades de codificación más profundas de acuerdo con las profundidades pueden ser determinadas de acuerdo con el tamaño máximo previamente determinado de la unidad de codificación.

En una estructura jerárquica 600 de las unidades de codificación, la altura máxima y la anchura máxima de las unidades de codificación son, cada una, de 64, y la profundidad máxima es de 4. Debido a que una profundidad aumenta a lo largo de un eje vertical de la estructura jerárquica 600, una altura y una anchura de la unidad de codificación más profunda son, cada una, divididas. Asimismo, una unidad de predicción y las particiones, que son bases para la codificación de predicción de cada unidad de codificación más profunda, son mostradas a lo largo de un eje horizontal de la estructura jerárquica 600.

Dicho de otra forma, una unidad de codificación 610 es una unidad de codificación máxima en la estructura jerárquica 600 en la que una profundidad es 0 y un tamaño, es decir, una altura por anchura, es 64 x 64. La profundidad aumenta a lo largo del eje vertical, y existen una unidad de codificación 620 que tiene un tamaño de 32 x 32 y una profundidad de 1, una unidad de codificación 630 que tiene un tamaño de 16 x 16 y una profundidad de 2, una unidad de codificación 640 que tiene un tamaño de 8 x 8 y una profundidad de 3, y una unidad de codificación 650 que tiene un tamaño de 4 x 4 y una profundidad de 4. La unidad de codificación 650 que tiene el tamaño de 4 x 4 y la profundidad de 4 es una unidad de codificación mínima.

La unidad de predicción y las particiones de una unidad de codificación son acomodadas a lo largo del eje horizontal de acuerdo con cada profundidad. Dicho de otra forma, si la unidad de codificación 610 que tiene el tamaño de 64 x 64 y la profundidad de 0 es una unidad de predicción, la unidad de predicción puede ser dividida en particiones incluidas en la unidad de codificación 610, es decir, una partición 610 que tiene un tamaño de 64 x 64, las particiones 612 que tienen el tamaño de 64 x 32, las particiones 614 que tienen el tamaño de 32 x 64, o las particiones 616 que tienen el tamaño 32 x 32.

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

De manera similar, una unidad de predicción de la unidad de codificación 630 que tiene el tamaño de 16 x 16 y la profundidad de 2 puede ser dividida en particiones incluidas en la unidad de codificación 630, es decir una partición que tiene un tamaño de 16 x 16 e incluida en la unidad de codificación 630, las particiones 632 que tienen un tamaño de 16 x 8, las particiones 634 que tienen un tamaño de 8 x 16, y las particiones 636 que tienen un tamaño de 8 x 8.

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

La unidad de codificación 650 que tiene el tamaño de 4 x 4 y la profundidad de 4 es la unidad de codificación mínima y una unidad de codificación de la profundidad más baja. Una unidad de predicción de la unidad de codificación 650 solo es asignada a una partición que tiene un tamaño de 4 x 4.

Con el fin de determinar al menos una profundidad codificada de las unidades de codificación que constituyen la unidad de codificación máxima 610, el determinador de unidades de codificación 120 del aparato de codificación de vídeo 100 realiza la codificación para las unidades de codificación que se corresponden con cada profundidad incluida en la unidad de codificación máxima 610.

Un número de unidades de codificación más profundas de acuerdo con las profundidades que incluyen datos en el mismo intervalo y el mismo tamaño se incrementa a medida que se incrementa la profundidad. Por ejemplo, se requiere que cuatro unidades de codificación que se corresponden con una profundidad de 2 cubran los datos que son incluidos en una unidad de codificación que se corresponde con una profundidad de 1. En consecuencia, con el fin de comparar los resultados de la codificación de los mismos datos de acuerdo con las profundidades, la unidad de codificación que se corresponde con la profundidad de 1 y cuatro unidades de codificación que se corresponden con la profundidad de 2 son, cada una, codificadas.

Con el fin de realizar la codificación para una profundidad actual, de entre las profundidades, al menor error de codificación puede ser seleccionado para la profundidad actual mediante la realización de la codificación para cada unidad de predicción en las unidades de codificación que se corresponden con la profundidad actual, a lo largo del eje horizontal de la estructura jerárquica 600. Como alternativa, el error de codificación mínimo puede ser buscado mediante una comparación de los menores errores de codificación de acuerdo con las profundidades, mediante la realización de la codificación para cada profundidad a medida que la profundidad aumenta a lo largo del eje vertical de la estructura jerárquica 600. Una profundidad y una partición que tiene el error de codificación mínimo en la unidad de codificación 610 pueden ser seleccionadas como la profundidad codificada y un tipo de partición de la unidad de codificación 610.

La figura 7 es un diagrama para describir una relación entre una unidad de codificación 710 y las unidades de transformación 720.

El aparato de codificación de vídeo 100 o 200 codifica o descodifica una imagen de acuerdo con las unidades de codificación que tienen tamaños más pequeños que o iguales a una unidad de codificación máxima para cada unidad de codificación máxima. Los tamaños de las unidades de transformación para la transformación durante la codificación pueden ser seleccionados sobre la base de las unidades de datos que no son más grandes que una unidad de codificación correspondiente.

Por ejemplo, en el aparato de codificación de vídeo 100 o 200, si un tamaño de la unidad de codificación 710 es de 64 x 64, la transformación puede ser realizada mediante el uso de las unidades de transformación 720 que tienen un tamaño de 32 x 32.

Asimismo, los datos de la unidad de codificación 710 que tiene el tamaño de 64 x 64 pueden ser codificados por la realización de la transformación sobre cada una de las unidades de transformación que tienen el tamaño de 32 x 32, 16 x 16, 8 x 8 y 4 x 4, que son más pequeñas de 64 x 64, y luego una unidad de transformación que tiene el menor error de codificación, puede ser seleccionada.

La figura 8 es un diagrama para describir una información de codificación de unas unidades de codificación que se corresponden con una profundidad codificada.

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

La información 800 indica la información con respecto a una forma de una partición obtenida al dividir una unidad de predicción de una unidad de codificación actual, en la que la partición es una unidad de datos para la codificación de predicción de la unidad de codificación actual. Por ejemplo, una unidad de codificación actual CU_0 que tiene un tamaño de 2N x 2N puede ser dividida en cualquiera de una partición 802 que tenga un tamaño de 2N x 2N, una partición 804 que tenga un tamaño de 2N x N, una partición 802 que tenga un tamaño de N x 2N, y una partición 808 que tenga un tamaño de N x N. En el presente caso, la información 800 con respecto a un tipo de partición es ajustada para indicar una de la partición 804 que tiene un tamaño de 2N x N, la partición 806 que tiene un tamaño de N x 2N, y la partición 808 que tiene un tamaño de N x N.

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

La información 820 indica una unidad de transformación que va a estar basada en cuándo la información es realizada sobre una unidad de codificación actual. Por ejemplo, la unidad de transformación puede ser una primera unidad de transformación intra 822, una segunda unidad de transformación intra 824, una primera unidad de transformación inter 826, o una segunda unidad de transformación intra 828.

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

La figura 9 es un diagrama de unas unidades de codificación más profundas de acuerdo con las profundidades.

La información de división puede ser usada para indicar un cambio de una profundidad. La información de división indica si una unidad de codificación de una profundidad actual es dividida en unidades de codificación de una profundidad inferior.

Una unidad de predicción 910 para la codificación de predicción de una unidad de codificación 900 que tiene una profundidad de 0 y un tamaño de $2N_0 \times 2N_0$ puede incluir las particiones de un tipo de partición 912 que tiene un tamaño de $2N_0 \times 2N_0$, un tipo de partición 914 que tiene un tamaño de $2N_0 \times N_0$, un tipo de partición 916 que tiene un tamaño de $N_0 \times 2N_0$, y un tipo de partición 918 que tiene un tamaño de $N_0 \times N_0$. La figura 9 solo ilustra los tipos de partición 912 al 918 que son obtenidos al dividir simétricamente la unidad de predicción 910, pero un tipo de partición no está limitado a estos, y las particiones de la unidad de predicción 910 pueden incluir particiones asimétricas, particiones que tienen una forma previamente determinada, y particiones que tienen una forma geométrica.

La codificación de predicción es repetidamente realizada sobre 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 $N_0 \times 2N_0$, y cuatro particiones que tienen un tamaño de $N_0 \times N_0$, de acuerdo con cada tipo de partición. La codificación de predicción en un modo de intra y un modo inter puede ser realizada sobre las particiones que tienen 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 de predicción en un modo de salto es realizado solo sobre la partición que tiene el tamaño de $2N_0 \times 2N_0$.

Los errores de la codificación que incluyen la codificación de predicción en los tipos de partición 912 al 918 son comparados, y el menor error de codificación es determinado de acuerdo con los tipos de partición. Si un error de codificación es más pequeño en uno de los tipos de partición 912 al 916, la unidad de predicción 910 puede no ser dividida en una profundidad más baja.

Si el error de codificación es el más pequeño en el tipo de partición 918, una profundidad es cambiada de 0 a 1 para dividir el tipo de partición 918 en la operación 920, y la codificación es repetidamente realizada sobre las unidades de codificación 930 que tienen una profundidad de 2 y un tamaño de $N_0 \times N_0$ para buscar un error de codificación mínimo.

Una unidad de predicción 940 para la codificación de predicción de la unidad de codificación 930 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 de partición 942 que tiene un tamaño de $2N_1 \times 2N_1$, un tipo de partición 944 que tiene un tamaño de $2N_1 \times N_1$, un tipo de partición 946 que tiene un tamaño de $N_1 \times 2N_1$, y un tipo de partición 948 que tiene un tamaño de $N_1 \times N_1$.

Si un error de codificación es el más pequeño en el tipo de partición 948, una profundidad es cambiada de 1 a 2 para dividir el tipo de partición 948 en la operación 950, y la codificación es repetidamente realizada sobre las unidades de codificación 960, que tiene 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 operación y división de acuerdo con cada profundidad puede ser realizada hasta cuando una profundidad se vuelve $d - 1$, y la información de división puede ser codificada como hasta cuando una profundidad es de 1 de 0 a $d - 2$. Dicho de otra forma, cuando la codificación es realizada hasta cuando la profundidad es $d - 1$ después de que una unidad de codificación que se corresponde con una profundidad de $d - 2$ es dividida en la operación 970, una unidad de predicción 990 para la codificación de predicción de una unidad de codificación 980 que tiene una profundidad de $d - 1$ y un tamaño de $2N_{(d-1)} \times 2N_{(d-1)}$ puede incluir particiones de un tipo de partición 992 que tiene un tamaño de $2N_{(d-1)} \times 2N_{(d-1)}$, un tipo de partición 994 que tiene un tamaño de $2N_{(d-1)} \times N_{(d-1)}$, un tipo de partición 996 que tiene un tamaño de $N_{(d-1)} \times 2N_{(d-1)}$, y un tipo de partición 998 que tiene un tamaño de $N_{(d-1)} \times N_{(d-1)}$.

La codificación de predicción puede ser repetidamente realizada sobre 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 de partición 992 al 998 para buscar un tipo de partición que tiene un error de codificación mínimo.

Incluso cuando el tipo de partición 998 tiene el error de codificación mínimo, debido a que una profundidad máxima es d , una unidad de codificación $CU_{(d-1)}$ que tiene una profundidad de $d - 1$ ya no es dividida a una profundidad inferior, y una profundidad codificada para las unidades de codificación que constituyen una unidad de codificación máxima actual 900 es determinada para ser $d - 1$ y un tipo de partición de la unidad de codificación máxima actual 900 puede ser determinada para ser $N_{(d-1)} \times N_{(d-1)}$. Asimismo, debido a que la profundidad máxima es d y una unidad de codificación mínima 980 que tiene una profundidad más baja de $d - 1$ ya no es dividida a una profundidad inferior, la información de división para la unidad de codificación mínima 980 no es ajustada.

Una unidad de datos 999 puede ser una "unidad mínima" para la unidad de codificación máxima actual. Una unidad mínima puede ser una unidad de datos rectangular obtenida al dividir una unidad de codificación mínima 980 entre 4. Mediante la realización de la codificación repetidamente, el aparato 100 de codificación de vídeo puede seleccionar una profundidad que tiene el menor error de codificación al comparar los errores de codificación de acuerdo con las profundidades de la unidad de codificación 900, para determinar una profundidad codificada, y

ajustar un tipo de partición correspondiente y un modo de predicción como un modo de codificación de la profundidad codificada.

5 Como tal, los errores de codificación mínimos de acuerdo con las profundidades son comparados en todas las profundidades de 1 a d, y una profundidad que tiene el menor error de codificación puede ser determinada como una profundidad codificada. La profundidad codificada, el tipo de partición de la unidad de predicción, y el modo de predicción pueden ser codificados y transmitidos como información con respecto a un modo de codificación. Asimismo, debido a que una unidad de codificación es dividida de una profundidad de 0 a una profundidad codificada, solo la información dividida de la profundidad codificada es ajustada a 0, y la información dividida de las profundidades excluyendo la profundidad codificada es ajustada a 1.

10 El extractor de datos de imagen y de información de codificación 220 del aparato de descodificación de vídeo 200 puede extraer y usar la información con respecto a la profundidad codificada y la unidad de predicción de la unidad de codificación 900 para descodificar la partición 912. El aparato de descodificación de vídeo 200 puede determinar una profundidad, en la cual la información de división es 0, como una profundidad codificada mediante el uso de la información de división de acuerdo con las profundidades, y usar la información con respecto a un modo de codificación de la profundidad correspondiente, para la descodificación.

15 Las figuras 10 a 12 son unos diagramas para describir una relación entre las unidades de codificación 1010, las unidades de predicción 1060 y las unidades de transformación 1070.

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

25 Cuando una profundidad de una unidad de codificación máxima es 0 en las unidades de codificación 1010, las profundidades de las unidades de codificación 1012 y 1054 son 1, las profundidades de las unidades de codificación 1014, 1016, 1018, 028, 1050 y 1052 son 2, las profundidades de las unidades de codificación 1020, 1024, 1026, 1030, 1032 y 1048 son 3, y las profundidades de las unidades de codificación 1040, 1042, 1044 y 1046 son 4.

30 En las unidades de predicción 1060, algunas unidades de codificación 1014, 1016, 1022, 1032, 1048, 1050, 1052 y 1054 son obtenidas al dividir la unidad de codificación en las unidades de codificación 1010. Dicho de otra forma, los tipos de partición en las unidades de codificación 1014, 1022, 1050 y 1054 tienen un tamaño de $2N \times N$, los tipos de partición en las unidades de codificación 1016, 1048 y 1052 tienen un tamaño de $N \times 2N$, y un tipo de partición de la unidad de codificación 1032 tiene un tamaño de $N \times N$. Las unidades de predicción y las particiones de las unidades de codificación 1010 son más pequeñas que o iguales a cada unidad de codificación.

35 La transformación o transformación inversa es realizada sobre los datos de imagen de la unidad de codificación 1052 en las unidades de transformación 1070 en una unidad de datos que es más pequeña que la unidad de codificación 1052. Asimismo, las unidades de codificación 1014, 1016, 1022, 1032, 1048, 1050 y 1052 en las unidades de transformación 1070 son diferentes de aquellas en las unidades de predicción 1060 en términos de tamaños y formas. Dicho de otra forma, los aparatos de codificación y descodificación 100 y 200 pueden realizar la predicción intra, la estimación de movimiento, la compensación de movimiento, la transformación y la transformación inversa individualmente sobre una unidad de datos en la misma unidad de codificación.

40 En consecuencia, la codificación es recursivamente realizada sobre 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, las unidades de codificación que tienen una estructura de árbol recursiva pueden ser obtenidas. La información de codificación puede incluir la información de división con respecto a una unidad de codificación, la información con respecto a un tipo de partición, la información con respecto a un modo de predicción, y la información con respecto a un tamaño de una unidad de transformación. La tabla 1 muestra la información de codificación que puede ser ajustada por los aparatos de codificación y descodificación de vídeo 100 y 200.

Tabla 1

Información de división 0 (codificación sobre la 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 la unidad de transformación		Codificar repetidamente las unidades de codificación que tienen profundidad inferior de $d + 1$
Intra Inter	Tipo de partición simétrica	Tipo de partición asimétrica	Información de división 0 de la unidad de transformación	Información de división 1 de la unidad de transformación	
Salto (solo $2N \times 2N$)	$2N \times 2N$ $2N \times N$ $N \times 2N$ $N \times N$	$2N \times nU$ $2N \times nD$ $nL \times 2N$ $nR \times 2N$	$2N \times 2N$	$N \times N$ (tipo simétrico) $N / 2 \times N / 2$ (tipo asimétrico)	

La unidad de salida 130 del aparato 100 de codificación de vídeo puede enviar la información de codificación con respecto a las unidades de codificación que tienen una estructura de árbol, y el extractor de datos de imagen y de información de codificación 220 del aparato de decodificación de vídeo 200 puede extraer la información de codificación con respecto a las unidades de codificación que tiene una estructura de árbol a partir de una secuencia de bits recibida.

La información de división indica si una unidad de codificación actual es dividida en unidades de codificación de una profundidad inferior. Si la información de división de una profundidad actual d es 0, una profundidad, en la cual una unidad de codificación actual ya no es dividida en una profundidad inferior, es una profundidad codificada y, por lo tanto, la información con respecto a un tipo de partición, el modo de predicción, y un tamaño de una unidad de transformación pueden ser definidos para la profundidad codificada. Si la unidad de codificación actual es adicionalmente dividida de acuerdo con la información de división, la codificación es independientemente realizada sobre las cuatro unidades de codificación divididas de una profundidad inferior.

Un modo de predicción puede ser uno de un modo intra, un modo inter, y un modo de salto. El modo intra y el modo inter pueden ser definidos en todos los tipos de partición, y el modo de salto es definido solo en un tipo de partición que tiene un tamaño de $2N \times 2N$.

La información con respecto al tipo de partición puede indicar los tipos de partición simétrica que tienen tamaños de $2N \times 2N$, $2N \times N$, $N \times 2N$, y $N \times N$, que son obtenidos al dividir simétricamente una altura o una anchura de una unidad de predicción, y los tipos de partición asimétrica que tienen tamaños de $2N \times nU$, $2N \times nD$, $nL \times 2N$, y $nR \times 2N$, que son obtenidos al dividir asimétricamente la altura o la anchura de la unidad de predicción. Los tipos de partición asimétrica que tienen los tamaños de $2N \times nU$ y $2N \times nD$ pueden ser respectivamente obtenidos al dividir 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 ser respectivamente obtenidos al dividir la anchura de la unidad de predicción en $1 : 3$ y $3 : 1$.

El tamaño de la unidad de transformación puede ser ajustado para ser dos tipos en el modo intra y dos tipos en el modo inter. Dicho de otra forma, si la información de división de la unidad de transformación es 0, el tamaño de la unidad de transformación puede ser $2N \times 2N$, el cual 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 ser obtenidas al dividir la unidad de codificación actual. Asimismo, 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 de $N \times N$, y si el tipo de partición de la unidad de codificación actual está en el tipo de partición asimétrica, el tamaño de la unidad de transformación puede ser $N / 2 \times N / 2$.

La información de codificación con respecto a las unidades de codificación que tienen una estructura de árbol puede incluir al menos una de una unidad de codificación que se corresponde con una profundidad codificada, una unidad de predicción, y una unidad mínima. La unidad de codificación que se corresponde con la profundidad codificada puede incluir al menos una de una unidad de predicción y una unidad mínima que contiene la misma información de codificación.

En consecuencia, se determina si las unidades de datos adyacentes son o no incluidas en la misma unidad de codificación que se corresponde con la profundidad codificada al comparar la información de codificación de las unidades de datos adyacentes. Asimismo, una unidad de codificación correspondiente que se corresponde con una profundidad codificada es determinada mediante el uso de la información de codificación de una unidad de datos y, por lo tanto, una distribución de las profundidades codificadas en una unidad de codificación máxima puede ser

determinada.

En consecuencia, si una unidad de codificación actual es predicha sobre la base de la información de codificación de las unidades de datos adyacentes, se puede hacer directamente referencia a, y usar, la información de codificación de las unidades de datos en unidades de codificación más profundas adyacentes a la unidad de codificación actual.

- 5 Como alternativa, si una unidad de codificación actual es predicha sobre la base de la información de codificación de las unidades de datos adyacentes, las unidades de datos adyacentes a la unidad de codificación actual son buscadas usando la información de codificación de las unidades de datos, y se puede hacer referencia a las unidades de codificación adyacentes buscadas para la predicción de la unidad de codificación actual.

- 10 La figura 13 es un diagrama para describir una relación entre una unidad de codificación, una unidad de predicción o una partición, y una unidad de transformación, de acuerdo con la información del modo de codificación de la tabla 1.

- 15 Una unidad de codificación máxima 1300 incluye las unidades de codificación 1302, 1304, 1306, 1312, 1314, 1316 y 1318 de las profundidades codificadas. En el presente caso, debido a que la unidad de codificación 1318 es una unidad de codificación de una profundidad codificada, la información de división puede ser ajustada a 0. La información con respecto a un tipo de partición de la unidad de codificación 1318 que tiene un tamaño de $2N \times 2N$, puede ser ajustada para ser una de un tipo de partición 1322 que tiene un tamaño de $2N \times 2N$, un tipo de partición 1324 que tiene un tamaño de $2N \times 2N$, un tipo de partición 1326 que tiene un tamaño de $2N \times 2N$, un tipo de partición 1328 que tiene un tamaño de $N \times N$, un tipo de partición 1332 que tiene un tamaño de $2N \times nU$, un tipo de partición 1334 que tiene un tamaño de $2N \times nD$, un tipo de partición 1336 que tiene un tamaño de $nL \times 2N$, y un tipo de partición 1338 que tiene un tamaño de $nR \times 2N$.

- 20 Cuando el tipo de partición es ajustado para ser simétrico, es decir, el tipo de partición 1322, 1324, 1326 o 1328, un número de transformación 1342 que tiene un tamaño de $2N \times 2N$ es ajustado si la información de división (bandera de tamaño TU) de una unidad de transformación es 0, y una unidad de transformación 1344 que tiene un tamaño de $N \times N$ es ajustada si una bandera de tamaño TU es 1.

- 25 Cuando el tipo de partición es ajustado para ser asimétrico, es decir, el tipo de partición 1332, 1334, 1336 o 1338 una unidad de transformación 1352 que tiene un tamaño de $2N \times 2N$ es ajustado si una bandera de tamaño TU es 0, y una unidad de transformación 1354 que tiene un tamaño de $N / 2 \times N / 2$ es ajustado si una bandera de tamaño TU es 1.

- 30 Tal como se ha descrito en lo que antecede, el aparato de codificación de vídeo 100 y el aparato de decodificación de vídeo 200 realizan la codificación y la decodificación al dividir una unidad de codificación máxima mediante el uso de una unidad de codificación igual a o menor que la unidad de codificación máxima. Los datos codificados en el aparato de codificación de vídeo 100 son multiplexados mediante el uso de una unidad de datos de transmisión apropiada para un protocolo o un formato de un canal de comunicación, un medio de almacenamiento, un sistema de edición de vídeo, una estructura de medios, o similares, y la unidad de datos de transmisión es transmitida al aparato de decodificación de vídeo 200.

- 35 En un caso de reproducción de datos de vídeo, el aparato de decodificación de vídeo 200 restablece los datos de vídeo de acuerdo con una de una manera de reproducción especial y una manera de reproducción normal, y reproduce los datos de vídeo. La manera de reproducción especial incluye una manera de reproducción normal, una manera de avance rápido o una manera de retroceso rápido, y una manera de acceso aleatorio. De acuerdo con la manera de reproducción normal, todas las imágenes incluidas en los datos de vídeo son procesadas y reproducidas de forma secuencial. De acuerdo con la manera de avance rápido o la manera de retroceso rápido, una imagen I en cada periodo previamente determinado es seleccionada y reproducida de una manera hacia adelante o hacia atrás de acuerdo con una velocidad de reproducción. De acuerdo con la manera de acceso aleatorio, la reproducción es realizada con un salto a una imagen clave, es decir, una imagen I en una posición previamente determinada. De acuerdo con la norma H.264, se usa una imagen de refresco de decodificador instantáneo (IDR, *instantaneous decoder refresh*) como la imagen clave para la manera de acceso aleatorio. La imagen de IDR es una imagen intra para refrescar un aparato de decodificación cuando una imagen correspondiente es descodificada. Con más detalle, cuando la imagen de IDR es descodificada, un DPB marca una imagen diferente de la imagen de IDR, la cual es previamente descodificada, como una imagen de no referencia, y un recuento de orden de imagen (POC, *picture order count*) es también inicializado. Asimismo, una imagen que es descodificada después de la imagen de IDR puede siempre estar detrás de la imagen de IDR en un orden de visualización, y puede ser descodificada mientras no se hace referencia a una imagen antes de la imagen de IDR.

- De acuerdo con la presente realización, además de la imagen de IDR, se usa una imagen de acceso aleatorio limpia (CRA, *clean random access*) como la imagen clave para la manera de acceso aleatorio. Se puede hacer referencia a la imagen de CRA como una imagen de refresco de decodificación limpia (CDR, *clean decoding refresh*) o una imagen de refresco de decodificación diferida (DDR, *deferred decoding refresh*). La imagen de CRA es una imagen intra que tiene unas imágenes que preceden en el orden de visualización a, pero son codificadas (o descodificados) después de, la imagen de CRA en un orden de codificación (o de decodificación). Después de esto, una imagen que está en el mismo grupo de imágenes (GOP, *group of pictures*) que la imagen de CRA y que precede a la imagen

de CRA en el orden de visualización pero es codificado (o descodificado) después de la imagen de CRA en el orden de codificación (o de descodificación) es definido como una imagen de comienzo.

La imagen de IDR y la imagen de CRA son comunes debido a que las mismas son unas imágenes clave para la manera de acceso aleatorio y son unas imágenes intra que van a ser codificadas (o descodificadas) sin hacer referencia a otra imagen. Con respecto a la imagen de IDR, una imagen que sigue a la imagen de IDR en el orden de codificación (o de descodificación) no precede a la imagen de IDR en el orden de visualización. No obstante, tal como se ha descrito en lo que antecede, con respecto a la imagen de CRA, la imagen de comienzo sigue a la imagen de CRA en el orden de codificación (o de descodificación) pero precede a la imagen de CRA en el orden de visualización. El orden de descodificación y el orden de codificación indican los órdenes en un descodificador y un codificador mediante los cuales son procesadas las imágenes, y un orden de codificación de las imágenes es igual a un orden de descodificación de las imágenes. En consecuencia, a lo largo de la totalidad de la memoria descriptiva, el orden de codificación puede querer decir el orden de descodificación, o el orden de descodificación puede querer decir el orden de codificación.

Las figuras 14A y 14B ilustran un proceso de descodificación de una imagen de CRA en una reproducción normal y un acceso aleatorio, de acuerdo con una realización de la presente invención.

En las figuras 14A y 14B, Bi y bi son unas imágenes Bi-predictivas que son predichas mediante el uso de la predicción de L0 y la predicción de L1, en la cual i indica un orden de visualización, es decir un POC. Asimismo, Bi que tiene una letra mayúscula "B" indica una imagen que es usada como una imagen de referencia de otra imagen, y bi que tiene una letra minúscula "b" indica una imagen que no es usada como una imagen de referencia de otra imagen.

Con referencia a la figura 14A, en la reproducción normal, es decir, cuando todas las imágenes son descodificadas y reproducidas, las imágenes de comienzo 1420 que son descodificadas después de una imagen de CRA₂₄ 1410, pueden ser descodificados normalmente por referencia a una imagen de referencia que es previamente descodificada. Por ejemplo, cuando una imagen de B₂₂ 1421 es predicha bi-direccionalmente mediante el uso de la predicción de L0 que se refiere a una imagen de B₁₈ 1401 y la predicción de L1 que se refiere a la imagen de CRA₂₄ 1410, en la reproducción normal, la imagen de B₂₂ 1421 puede ser descodificada normalmente por referencia a la imagen de CRA₂₄ 1410 y a la imagen de B₁₈ 1401 que son previamente descodificadas y almacenadas en un DPB 1430.

Con referencia a la figura 14B, en un caso en el que la imagen de CRA₂₄ 1410 es descodificada por medio del acceso aleatorio después de que una imagen de B₆ 1402 sea descodificada, la predicción de L0 para la predicción de la imagen de B₂₂ 1421 determina una imagen de referencia de acuerdo con un índice de imagen de referencia en una dirección de la predicción de L0. En este caso, la imagen de B₆ 1402 que es previamente descodificada y almacenada en un DPB 1440 puede ser determinada como la imagen de referencia para la predicción de L0 de la imagen de B₂₂ 1421. En este caso, la imagen de referencia para la predicción de L0 de la imagen de B₂₂ 1421 tiene que ser la imagen de B₁₈ 1401 pero, debido al acceso aleatorio, una imagen que se refiere a las imágenes de referencia entre la imagen de B₆ 1402 y la imagen de CRA₂₄ 1410 no puede ser descodificada normalmente.

Asimismo, con referencia nuevamente a la figura 14A, debido a que las imágenes de comienzo 1420 preceden a la imagen de CRA₂₄ 1410 en un orden de visualización, después de que la imagen de CRA₂₄ 1410 sea descodificada y visualizada por medio del acceso aleatorio, las imágenes de comienzo 1420 no se visualizan y, por lo tanto, no se requiere que sean descodificadas.

No obstante, el aparato de descodificación de vídeo 200 descodifica y envía simplemente y de forma secuencial una pluralidad de piezas de datos de vídeo de entrada y, por lo tanto, no puede reconocer si las imágenes de comienzo 1420 son unas imágenes que son descodificadas después de la imagen de CRA₂₄ 1410 en la reproducción normal o por medio del acceso aleatorio. Dicho de otra forma, el aparato de descodificación de vídeo 200 no puede reconocer si la imagen de CRA₂₄ 1410 es una imagen que es descodificada por medio del acceso aleatorio o es descodificada en la reproducción normal.

De este modo, las realizaciones de la presente invención proporcionan procedimientos de multiplexión y desmultiplexión mediante los cuales las sintaxis para la identificación de un estado de reproducción de una imagen de CRA es agregada a los datos en una unidad de transmisión previamente determinada, en la que los datos son obtenidos mediante la multiplexión de la imagen de CRA que es codificada para ser identificada si la imagen de CRA es codificada o de acuerdo con un acceso aleatorio o es codificada de acuerdo con una reproducción normal.

En primer lugar, se describirán un aparato de multiplexión de datos de vídeo y un procedimiento del mismo.

La figura 15 ilustra una estructura de un aparato de multiplexión de datos de vídeo 1500.

Con referencia a la figura 15, el aparato de multiplexión de datos de vídeo 1500 incluye un codificador de vídeo 1510, un multiplexor 1520 y un identificador de estado de reproducción 1530.

El codificador de vídeo 1510 se corresponde con el aparato de codificación de vídeo 100 de la figura 1, y codifica datos de vídeo sobre la base de la unidad de codificación jerárquica en una capa de codificación de vídeo en la cual es procesada la codificación de los datos de vídeo. El multiplexor 1520 multiplexa los datos de vídeo mediante el uso de una unidad de datos de transmisión apropiada para un protocolo o un formato de un canal de comunicación, un medio de almacenamiento, un sistema de edición de vídeo, una estructura de medios, o similares. Tal como se describirá posteriormente, el multiplexor 1520 puede multiplexar los datos de vídeo mediante el uso de una unidad de NAL que es una unidad de transmisión en una NAL.

Cuando existe una solicitud de transmisión para los datos de vídeo codificados de parte de un cliente conectado por medio del canal de comunicación, un aparato que maneja el medio de almacenamiento, el sistema de edición de vídeo, y la estructura de medios (a la que se hace referencia de forma colectiva en lo sucesivo en el presente documento como un “aparato de descodificación”), el identificador de estado de reproducción 1530 identifica si la solicitud de transmisión es para la reproducción secuencial de los datos de vídeo de acuerdo con la reproducción normal o es para la transmisión de los datos de vídeo de acuerdo con un acceso aleatorio. El identificador de estado de reproducción 1530 puede comparar un tiempo de visualización de una imagen requerida por el aparato de descodificación con un tiempo de visualización de una imagen actualmente visualizada por un aparato de descodificación actual, y luego puede identificar que la solicitud de transmisión de los datos de vídeo tiene lugar de acuerdo con el acceso aleatorio.

El multiplexor 1520 agrega la sintaxis previamente determinada a la unidad de NAL que incluye información con respecto a una imagen de CRA que es una imagen clave para el acceso aleatorio, sobre la base del resultado de identificación por el identificador de estado de reproducción 1530, en la que la sintaxis previamente determinada indica qué solicitud de entre una solicitud por medio de la reproducción normal y una solicitud por medio del acceso aleatorio está relacionada con la imagen de CRA.

La figura 16 ilustra una estructura de una unidad de NAL 1600 para su uso en una realización de la presente invención.

Con referencia a la figura 16, la unidad de NAL 1600 es formada de un encabezamiento de NAL 1610 y una carga útil de secuencia de bytes bruta (RBSP, *raw byte sequence payload*) 1620. Un bit de relleno de RBSP 1630 es un bit de ajuste de longitud anexado a un extremo del RBSP 1620 para expresar así una longitud del RBSP 1620 como un múltiplo de 8 bits. El bit de relleno de RBSP 1630 tiene un patrón tal como “100...” que comienza con “1” y continúa con “0” secuencial determinado de acuerdo con la longitud de RBSP 1620. A este respecto, mediante la búsqueda de “1” que es un valor de bit inicial, una posición de un último bit de RBSP 1620 que está justo antes del valor del bit inicial, puede ser determinada.

El encabezamiento de NAL 1610 incluye un *forbidden_zero_bit* 1611 que tiene un valor de 0, una bandera *nal_ref_idc* que indica si la unidad de NAL 1600 incluye un sector que va a ser una imagen de referencia, y similares. En particular, el encabezamiento de NAL 1610 de acuerdo con la presente realización incluye la sintaxis de identificación de estado 1612 que indica cual solicitud de entre una solicitud por medio de la reproducción normal y una solicitud por medio del acceso aleatorio está relacionada con una imagen de CRA que se agrega a la unidad de NAL 1600 que incluye la información con respecto a la imagen de CRA.

La sintaxis de identificación de estado 1612 para identificar un estado de reproducción de la imagen de CRA puede ser incluida en un identificador (tipo unidad de nal) que indica un tipo de la unidad de NAL 1600. Es decir, una unidad de NAL que se usa para descodificar una imagen de CRA que se proporciona de acuerdo con una solicitud por medio de la reproducción normal puede tener un identificador (tipo de unidad de nal) que es un tipo diferente de un identificador (tipo de unidad de nal) de una unidad de NAL que se usa para descodificar una imagen de CRA que se proporciona de acuerdo con una solicitud por medio de un acceso aleatorio.

La tabla 2 muestra los tipos de la unidad de NAL 1600 de acuerdo con los valores de un identificador (tipo de unidad de nal).

[Tabla 2]

tipo de unidad de nal	Tipos de unidad de NAL
0	No especificado
1	Sector de imagen diferente de las imágenes de RAP, de TFD, de TLA
2	Sector de imagen de TFD
3	Sector de imagen de TLA diferente de TFD
4, 5	Sector de imagen de CRA
6, 7	Sector de imagen de BLA

(continuación)

tipo de unidad de nal	Tipos de unidad de NAL
8	Sector de imagen de IDR
9 - 24	Reservado para ampliación futura
25	VPS
26	SPS
27	PPS
28	APS
29	Delimitador de unidad de acceso (AU, <i>Access Unit</i>)
30	Datos de relleno
31	Información de potenciación complementaria (SEI, <i>Supplemental Enhancement Information</i>)
32 - 47	Reservado para ampliación futura
48 - 63	No especificado

El multiplexor 1520 asigna diferentes niveles de 4 y 5 en la tabla 2 anterior, como un identificador (tipo de unidad de nal), a la unidad de NAL que se usa para descodificar la imagen de CRA que se proporciona de acuerdo con la solicitud por medio de la reproducción normal, y a la unidad de NAL que se usa para descodificar la imagen de CRA que se proporciona de acuerdo con la solicitud por medio del acceso aleatorio. Al hacer esto, el multiplexor 1520 puede señalar que una unidad de NAL que incluya la información con respecto a una imagen de CRA es cuál de entre una imagen de CRA que se proporciona de acuerdo con la reproducción normal y una imagen de CRA que se proporciona de acuerdo con un acceso aleatorio.

Asimismo, el multiplexor 1520 puede usar una bandera como sintaxis agregada a un encabezamiento de la unidad de NAL, en la que la bandera es ajustada para tener uno de los diferentes valores de 0 y 1, con respecto a la unidad de NAL usada para descodificar la imagen de CRA que se proporciona de acuerdo con la solicitud por medio de la reproducción normal, y la unidad de NAL usada para descodificar la imagen de CRA que se proporciona de acuerdo con la solicitud por medio del acceso aleatorio.

Las figuras 17A y 17B ilustran un proceso de descodificación de una imagen de CRA en una reproducción normal y un acceso aleatorio. Estas figuras ilustran un proceso de descodificación fuera del alcance de la reivindicación 1 adjunta.

El multiplexor 1520 agrega una información de tipo, que indica directamente el estado de reproducción de la imagen de CRA, a la unidad de NAL. En otra realización más, el multiplexor 1520 puede no señalar directamente un tipo de estado de reproducción de la imagen de CRA pero puede identificar si la imagen de CRA es reproducida de acuerdo con la producción normal o bien es reproducido de acuerdo con el acceso aleatorio, mediante el uso de un contador de imágenes clave que se descodifican antes de la imagen de CRA, y mediante el uso de la información con respecto a un POC de una imagen que es usada como una imagen de referencia de las imágenes de comienzo.

Con referencia a las figuras 17A y 17B, una imagen de B₃₈ 1710 y una imagen de B₄₀ 1720 son unas imágenes que se descodifican antes de una imagen de CRA₄₄ y a las que se hace referencia mediante b₄₁, B₄₂ y b₄₃ que son unas imágenes de comienzo. Se hace referencia a la imagen de B₃₈ 1710 mediante la imagen de B₄₂ que es la imagen de comienzo, y se hace referencia a la imagen de B₄₀ 1720 mediante la imagen de B₄₁ que es la imagen de comienzo. La imagen de B₃₈ 1710 y la imagen de B₄₀ 1720 que preceden a una imagen de CRA en un orden de descodificación y son usadas como imágenes de referencia de las imágenes de comienzo son definidas como una imagen hermana. La razón por lo que la imagen hermana es definida es que es posible identificar si la imagen de CRA es reproducida de acuerdo con la reproducción normal o es reproducido de acuerdo con un acceso aleatorio, mediante el uso de un POC de la imagen hermana. La idea de usar la imagen hermana se ha desvelado en el documento de normalización JCTVC-E400 que se ha citado en lo que antecede

Por ejemplo, con referencia a la figura 17A, mientras las imágenes son descodificadas de forma secuencial en un estado de reproducción normal, cuando la imagen de CRA₄₄ es descodificada, la imagen de B₃₈ 1710 y la imagen de B₄₀ 1720 que son previamente descodificadas son almacenadas en un DPB 1740. Si un valor de POC de 38 de la imagen de B₃₈ 1710, y un valor de POC de 40 de la imagen de B₄₀ 1720, que son imágenes hermanas, son agregados a un encabezamiento de sector de la imagen de CRA₄₄, un descodificador puede comparar un POC de imágenes que son previamente descodificadas y almacenadas en el DPB 1740 a un tiempo de la descodificación de la imagen de CRA₄₄ con el POC de las imágenes hermanas incluidas en el encabezamiento del sector de la imagen de CRA₄₄ y, por lo tanto, puede identificar si la imagen de CRA₄₄ es reproducida de acuerdo con el acceso aleatorio o bien es reproducido de acuerdo con la reproducción normal. Debido a que un orden de reproducción, es decir, un orden de descodificación, se salta la imagen de CRA₄₄, en la reproducción de acuerdo con el acceso aleatorio, si el POC de las imágenes que son previamente descodificadas y almacenadas en el DPB 1740 a un tiempo de la

descodificación la imagen de CRA₄₄ no se acopla con el POC de las imágenes hermanas, existe una alta posibilidad de que la imagen de CRA₄₄ sea reproducida de acuerdo con el acceso aleatorio.

No obstante, existe también una posibilidad de que la imagen de CRA₄₄ se pueda identificar mal con respecto a si la imagen de CRA₄₄ es reproducida de acuerdo con el acceso aleatorio o es reproducida de acuerdo con la reproducción normal, mediante el uso solo de información con respecto al POC de las imágenes hermanas.

Por ejemplo, con referencia a la figura 17B en un caso en el que la imagen de CRA₄₄ es descodificada de acuerdo con el acceso aleatorio después de que una imagen de B₄₀ 1745 sea descodificada, cuando la imagen de CRA₄₄ es descodificada, un valor de POC de 40 de la imagen de B₄₀ 1745 es almacenado en un DPB 1780, y un valor de POC de 40 de una imagen hermana almacenada en el encabezamiento de sector de la imagen de CRA₄₄ es igual al valor de POC de una imagen previa almacenada en el DPB 1780, de tal modo que un descodificador puede identificar mal que la imagen de CRA₄₄ es reproducida de acuerdo con la reproducción normal. Tal como se ha descrito en lo que antecede, un estado de reproducción de una imagen de CRA puede no ser correctamente identificado mediante el uso solo de información con respecto a un POC de una imagen hermana. Es decir, debido a que cada vez que una imagen de IDR es descodificada, un POC es restablecido, de tal modo que una imagen de referencia que es diferente de una imagen de referencia real a la que se va a hacer referencia mediante las imágenes hermanas en un acceso aleatorio, puede tener el mismo POC que un POC de las imágenes hermanas.

De este modo, el multiplexor 1520 agrega la información de POC de una imagen hermana a la sintaxis de los datos de unidad de transmisión de una imagen de CRA, obtiene un recuento proveniente de un contador de discontinuidad de POC (PDC, *POC discontinuity counter*) que es un contador que tiene un valor que se incrementa en 1 siempre que un POC es restablecido o la imagen de CRA es codificada mientras las imágenes antes de la imagen de CRA son codificadas, y agrega el recuento desde el PDC a la sintaxis.

Con referencia nuevamente a la figura 17A, tal como se ha descrito en lo que antecede, el multiplexor 1520 incrementa el recuento del PDC en 1 siempre que el POC es restablecido o la imagen de CRA es codificada mientras las imágenes son codificadas. Debido a que el POC es restablecido siempre que la imagen de IDR es codificada, un valor de PDC se incrementa en 1 siempre que la imagen de IDR es codificada, y el valor del recuento del PDC se incrementa en 1 cuando se codifica una imagen de CRA previa excepto por la imagen de CRA 44. El multiplexor 1520 agrega el valor de PDC con el POC de las imágenes hermanas a un encabezamiento 1730 de los datos de unidad de transmisión de la imagen de CRA. El descodificador de la misma manera que un codificador incrementa el recuento del PDC en 1 siempre que el POC es restablecido mientras los datos de unidad de transmisión de entrada son descodificados, es decir, siempre que la imagen de IDR es descodificada o la imagen de CRA es descodificada. Tal como se ilustra en la figura 17A, en un caso de reproducción normal, cuando la imagen de CRA₄₄ es descodificada, un valor de PDC incluido en el encabezamiento 1730 de los datos de unidad de transmisión que se usan para multiplexar la imagen de CRA₄₄, y un valor de PDC 1745 contado mientras las imágenes son descodificadas por el descodificador, son 3 todos ellos.

Con referencia nuevamente a la figura 17B, en un caso de un acceso aleatorio, existen una imagen de IDR y una imagen de CRA que se descodifican antes de la imagen de CRA₄₄, de modo que, cuando la imagen de CRA₄₄ es descodificada de acuerdo con el acceso aleatorio, un valor de PDC 1785 es 2, el cual es diferente de un valor de PDC que tiene 3 e incluido en un encabezamiento 1770 de los datos de unidad de transmisión que se usan para multiplexar la imagen de CRA₄₄. De este modo, sobre la base de la no concordancia entre los valores de PDC, el descodificador puede determinar que una imagen de CRA₄₄ actual se reproduce de acuerdo con el acceso aleatorio.

La figura 18 es un diagrama de flujo de un procedimiento de multiplexión de datos de vídeo.

Con referencia a la figura 18, en la operación 1810, el codificador de vídeo 1510 codifica las imágenes que forman datos de vídeo sobre la base de las unidades de datos que tienen una estructura jerárquica. Tal como se ha descrito en lo que antecede, el codificador de vídeo 1510 determina las unidades de codificación de estructura de árbol que incluyen las unidades de codificación con profundidades de codificación de entre las unidades de codificación más profundas de acuerdo con las profundidades que son jerárquicamente formadas de acuerdo con las profundidades que indican un recuento de división espacial de al menos una unidad de codificación máxima obtenida al dividir una imagen que forma el vídeo mediante el uso de una unidad de codificación de tamaño máximo, determina una partición para la codificación de predicción para cada una de las unidades de codificación con profundidades de codificación, realiza la transformación sobre la base de las unidades de transformación de estructura jerárquica, y determina luego las unidades de transformación de estructura de árbol. En la determinación de una unidad de datos jerárquico, una estructura de una unidad de datos jerárquico óptima puede ser determinada sobre la base de un costo de distorsión de tasa (RD, *rate distortion*).

En la operación 1820, en respuesta a una solicitud de transmisión para los datos codificados provenientes de un aparato de descodificación, el identificador de estado de reproducción 1530 determina si la solicitud de transmisión es de acuerdo con la reproducción normal o es para un acceso aleatorio. Tal como se ha descrito en lo que antecede, el aparato de descodificación puede indicar de forma colectiva los aparatos para el almacenamiento, la reproducción, y la edición de datos de vídeo codificados, y puede incluir un cliente conectado por medio de un canal de comunicación, un aparato que maneja un medio de almacenamiento, un sistema de edición de vídeo, una

estructura de medios y similares.

En la operación 1830, el multiplexor 1520 agrega la sintaxis previamente determinada a los datos de unidad de transmisión que se usan para multiplexar una imagen de CRA que es una imagen intra que tiene imágenes de comienzo, de acuerdo con un resultado de la determinación en la operación 1820, en la que la sintaxis previamente determinada indica qué solicitud de entre una solicitud por medio de la reproducción normal y una solicitud por medio del acceso aleatorio está relacionada con la imagen de CRA. Tal como se ha descrito en lo que antecede, la imagen de comienzo quiere decir una imagen que precede a la imagen de CRA en un orden de visualización, pero es codificada después de la imagen de CRA en un orden de codificación. Asimismo, los datos de unidad de transmisión pueden ser los datos de NAL. Asimismo, el multiplexor 1520 puede agregar un identificador (tipo de unidad de NAL) que indica un tipo de una unidad de NAL a un encabezamiento de la unidad de NAL de tal modo que una unidad de NAL que se usa para multiplexar una imagen de CRA que se proporciona de acuerdo con la solicitud por medio de la reproducción normal, puede tener un identificador (tipo de unidad de nal) que es diferente de un identificador (tipo de unidad de nal) de una unidad de NAL que se usa para multiplexar una imagen de CRA que se proporciona de acuerdo con una solicitud por medio del acceso aleatorio. Asimismo, el multiplexor 1520 puede agregar una bandera al encabezamiento de la unidad de NAL, en la que la bandera tiene un valor de 0 o 1 de acuerdo con la unidad de NAL usada para multiplexar la imagen de CRA que se proporciona de acuerdo con la solicitud por medio de la reproducción normal, y la unidad de NAL usada para multiplexar la imagen de CRA que se proporciona de acuerdo con la solicitud por medio del acceso aleatorio. También el multiplexor 1520 puede obtener un recuento de un PDC que se incrementa en 1 siempre que un POC es restablecido o la imagen de CRA es codificada mientras las imágenes son multiplexadas, y puede agregar información de POC de una imagen hermana de la imagen de CRA, y el recuento del PDC al encabezamiento de la unidad de NAL.

La figura 19 ilustra una estructura de un aparato de multiplexión inversa de datos de vídeo 1900.

Con referencia a la figura 19, el aparato de multiplexión inversa de datos de vídeo 1900 incluye un descodificador de vídeo 1910, un multiplexor inverso 1920, y un identificador de estado de reproducción 1930.

El multiplexor inverso 1920 recibe los datos de unidad de transmisión, es decir, los datos de unidad de NAL que se transmiten desde el aparato de multiplexión de datos de vídeo 1500 de la figura 15, y obtiene las sintaxis que indica si una imagen de CRA se descodifica de acuerdo con un estado de reproducción normal o bien de acuerdo con un estado de acceso aleatorio, a partir de unos datos de unidad de NAL que se usan para multiplexar la imagen de CRA.

El identificador de estado de reproducción 1930 determina si la imagen de CRA es reproducida de acuerdo con la reproducción normal o de acuerdo con un acceso aleatorio, mediante el uso de la sintaxis obtenida.

Tal como se ha descrito en lo que antecede, cuando una unidad de NAL que se usa para multiplexar una imagen de CRA que se proporciona de acuerdo con la solicitud por medio de la reproducción normal es ajustado para tener un identificador (tipo de unidad de nal) diferente de un identificador (tipo de unidad de nal) de una unidad de NAL que se usa para multiplexar una imagen de CRA que se proporciona de acuerdo con una solicitud por medio del acceso aleatorio, el identificador de estado de reproducción 1930 pueda reconocer un estado de descodificación de una imagen de CRA actual por referencia a los valores de los identificadores (tipo de unidad de nal). De acuerdo con el procedimiento de la reivindicación 1 adjunta, la unidad de NAL tiene agregada a la misma una bandera. Por ejemplo, esa bandera tiene un valor de 0 o 1 de acuerdo con la unidad de NAL usada para multiplexar la imagen de CRA que se proporciona de acuerdo con la solicitud por medio de la reproducción normal, y la unidad de NAL usada para multiplexar la imagen de CRA que se proporciona de acuerdo con la solicitud por medio del acceso aleatorio, el identificador de estado de reproducción 1930 puede reconocer el estado de descodificación de la imagen de CRA actual por referencia a la información con respecto a la bandera.

El descodificador de vídeo 1910 se corresponde con el aparato de descodificación de vídeo 200 de la figura 2 o el descodificador de imágenes 500 de la figura 5. El descodificador de vídeo 1910 obtiene los datos de imagen codificados, y la información de división, la información de tipo de partición, la información del modo de predicción, la información del tamaño de la unidad de transformación, y la información de conjunto de parámetros relacionada con un proceso de codificación, que son con respecto a las unidades de codificación usadas para generar los datos de imagen codificados de una unidad de NAL y realiza la descodificación.

La figura 20 es un diagrama de flujo de un procedimiento de multiplexión inversa de datos de vídeo.

Con referencia a la figura 20, en la operación 2010, el multiplexor inverso 1920 recibe los datos de unidad de transmisión que se usan para multiplexar una secuencia de bits generada por la codificación de las imágenes que forman los datos de vídeo, sobre la base de unidades de datos de estructura jerárquica. Tal como se ha descrito en lo que antecede, los datos de unidad de transmisión pueden ser datos de unidad de NAL.

En la operación 2020, el multiplexor inverso 1920 obtiene la sintaxis que indica si una imagen de CRA se descodifica de acuerdo con un estado de reproducción normal o bien de acuerdo con un estado de acceso aleatorio, a partir de unos datos de unidad de NAL que se usan para multiplexar la imagen de CRA.

En la operación 2030, el identificador de estado de reproducción 1930 identifica si la imagen de CRA es reproducida de acuerdo con la reproducción normal o de acuerdo con un acceso aleatorio, sobre la base de la sintaxis obtenida. Tal como se ha descrito en lo que antecede, cuando un encabezamiento de una unidad de NAL señala un estado de descodificación de la imagen de CRA mediante el uso de un identificador (tipo de unidad de NAL), el identificador de estado de reproducción 1930 puede reconocer un estado de descodificación de una imagen de CRA actual por referencia a un valor del identificador (tipo de unidad de NAL). Si una bandera que tiene un valor de 0 o 1 es agregada al encabezamiento de la unidad de NAL, el identificador de estado de reproducción 1930 puede reconocer el estado de descodificación de la imagen de CRA actual por referencia a la información de la bandera. Asimismo, cuando se incluyen un PDC y una información de POC de una imagen hermana en el encabezamiento de la unidad de NAL, el identificador 1930 puede reconocer el estado de descodificación de la imagen de CRA actual mediante la determinación de una concordancia entre un PDC obtenido en un proceso de descodificación, y el PDC incluido en el encabezamiento de la unidad de NAL, y una concordancia entre un valor POC de imágenes previas almacenadas en un DPB, y un valor POC de la imagen hermana.

Cuando se determina que la imagen de CRA es reproducida de acuerdo con el acceso aleatorio, las imágenes de comienzo de la imagen de CRA no se visualizan y, por lo tanto, no se requiere que sean descodificadas. De acuerdo con las realizaciones de la presente invención, la imagen de CRA que es reproducida de acuerdo con el acceso aleatorio puede ser identificada, de tal modo que las imágenes de comienzo de la imagen de CRA pueden no ser descodificadas y, por lo tanto, se puede reservar un recurso del sistema del aparato de descodificación.

La invención también se puede materializar como códigos legibles por ordenador sobre un medio de grabación legible por ordenador. El medio de grabación legible por ordenador es cualquier dispositivo de almacenamiento de datos que pueda almacenar datos que pueden ser posteriormente leídos por un sistema informático. Los ejemplos del medio de grabación legible por ordenador incluyen memoria de solo lectura (ROM, *read only memory*), memoria de acceso aleatorio (RAM, *random access memory*), CD-ROM, cintas magnéticas, discos flexibles, dispositivos de almacenamiento óptico de datos, etc. El medio de grabación legible por ordenador también puede estar distribuido en sistemas informáticos conectados a la red, de tal modo que el código legible por ordenador se almacene y se ejecute de una manera distribuida.

A pesar de que la presente invención se ha mostrado y descrito en particular con referencia a algunas realizaciones preferidas de la misma, los expertos en la materia podrán comprender que se pueden realizar en la misma diversos cambios en la forma y los detalles sin apartarse del alcance de la invención tal como se define por medio de las reivindicaciones adjuntas. Por lo tanto, el alcance de la invención es definido no por la descripción detallada de la invención, sino por las reivindicaciones adjuntas, y todas las diferencias dentro del alcance se han de considerar como incluidas en la presente invención.

REIVINDICACIONES

1. Un procedimiento realizado por un aparato de descodificación de vídeo, de descodificación de un vídeo, comprendiendo el procedimiento:

- 5 recibir (2010), de una secuencia de bits, una unidad de capa de abstracción de red (NAL) que contiene un encabezamiento;
obtener (2020), del encabezamiento, información de tipo que indica un tipo de una imagen actual y una bandera de acceso aleatorio que indica que la imagen actual se reproduce de acuerdo con una reproducción normal o de acuerdo con un acceso aleatorio;
10 si la información de tipo y la bandera de acceso aleatorio indican que la imagen actual es una imagen de acceso aleatorio limpia usada en la reproducción normal, descodificar imágenes de comienzo que preceden a la imagen actual en un orden de visualización pero siguen a la imagen actual en un orden de descodificación al hacer referencia a una imagen descodificada antes de la imagen actual; y
si la información de tipo y la bandera de acceso aleatorio indican que la imagen actual es una imagen de acceso aleatorio limpia usada en el acceso aleatorio, saltar la descodificación de las imágenes de comienzo,
15 en el que la imagen actual se divide en una pluralidad de unidades de codificación máxima, en el que una de la pluralidad de unidades de codificación máxima se divide en una pluralidad de unidades de codificación.

FIG. 1

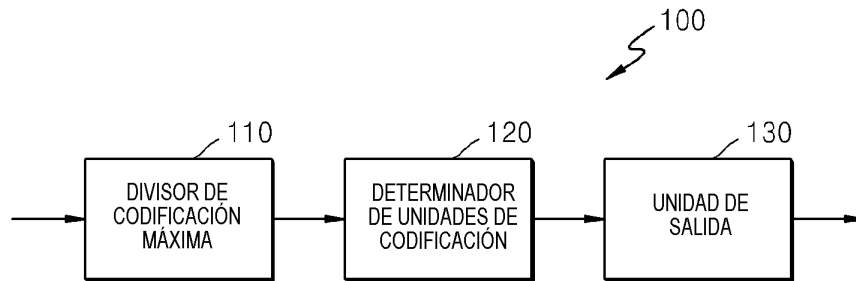


FIG. 2

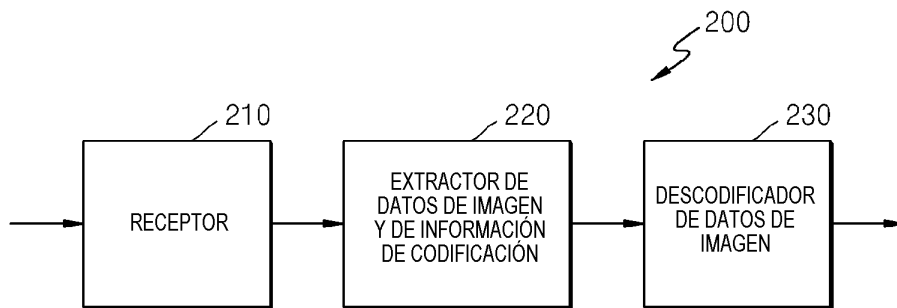


FIG. 3

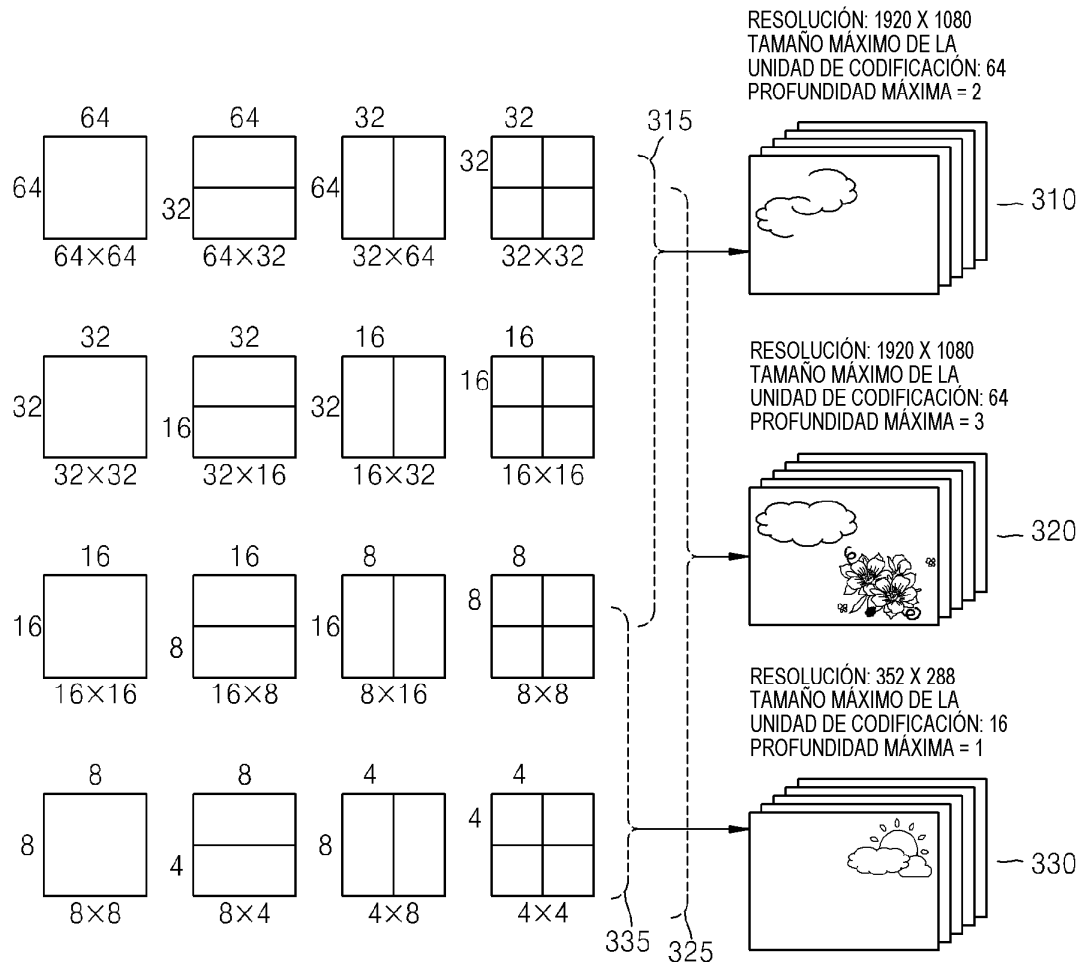


FIG. 4

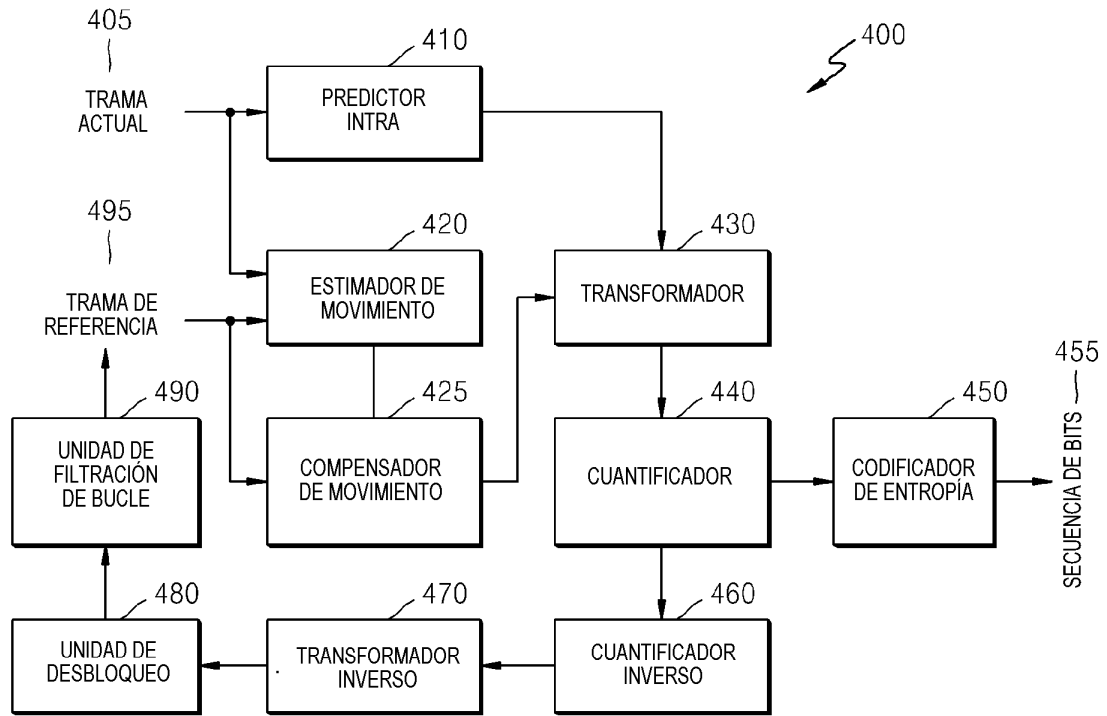


FIG. 5

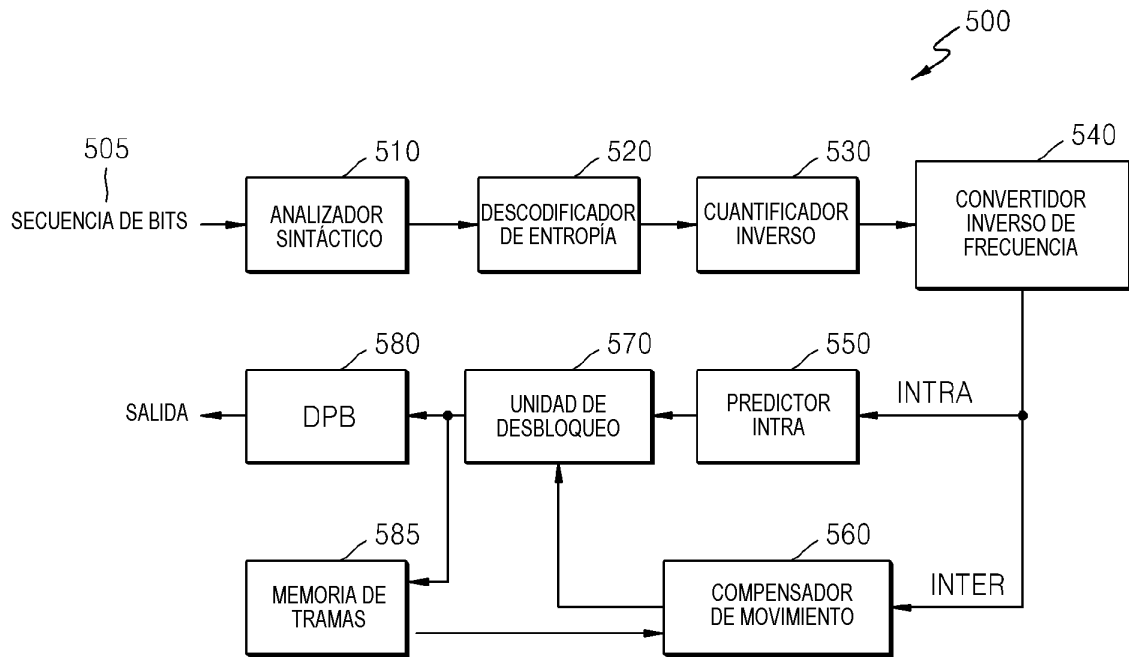


FIG. 6

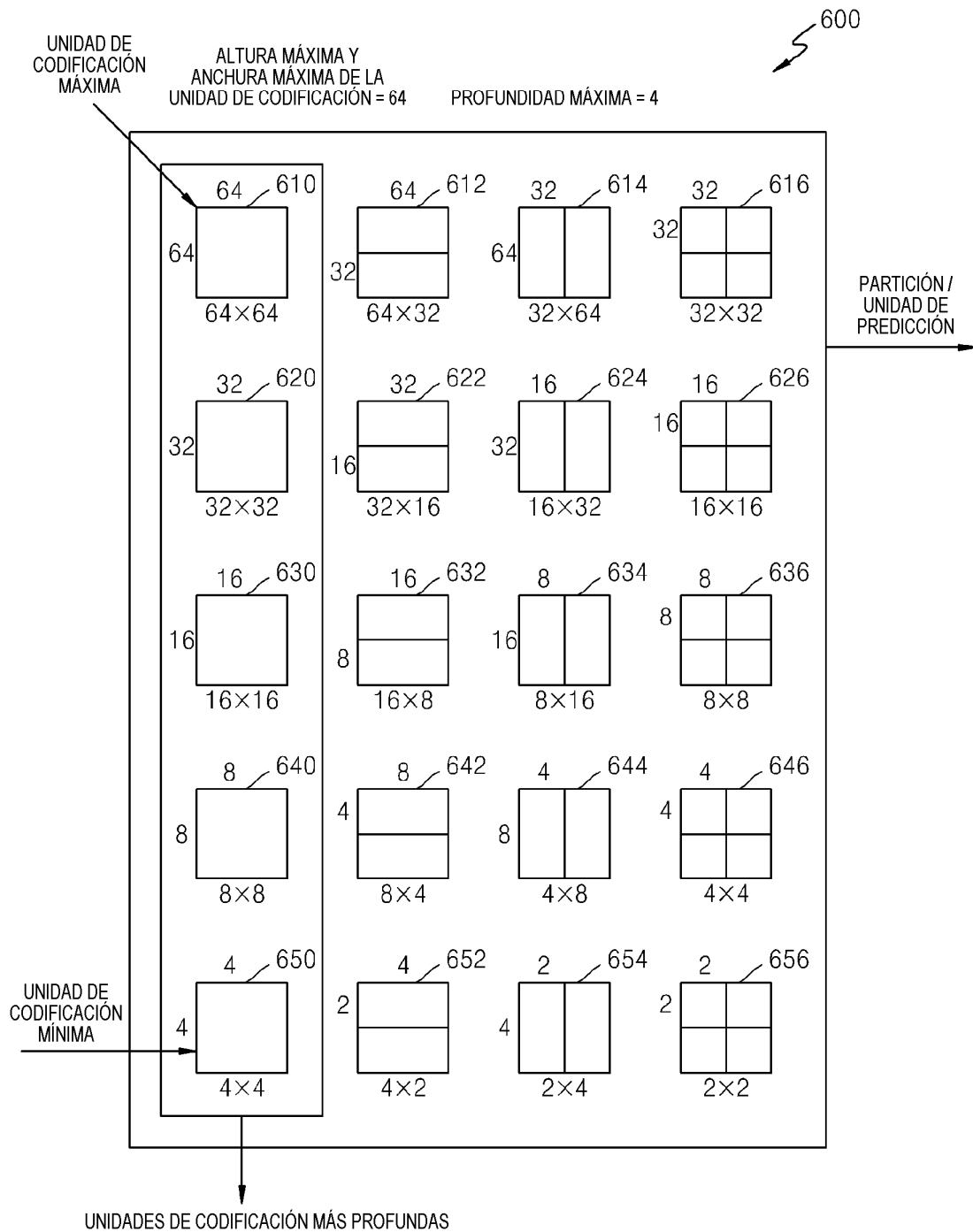


FIG. 7

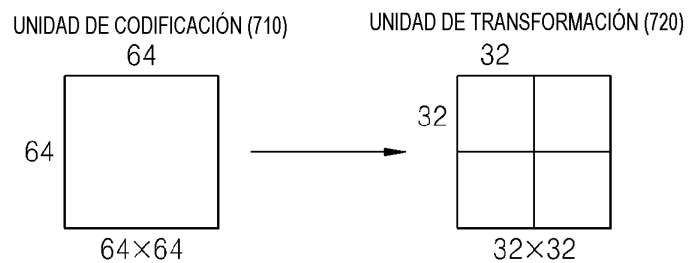


FIG. 8

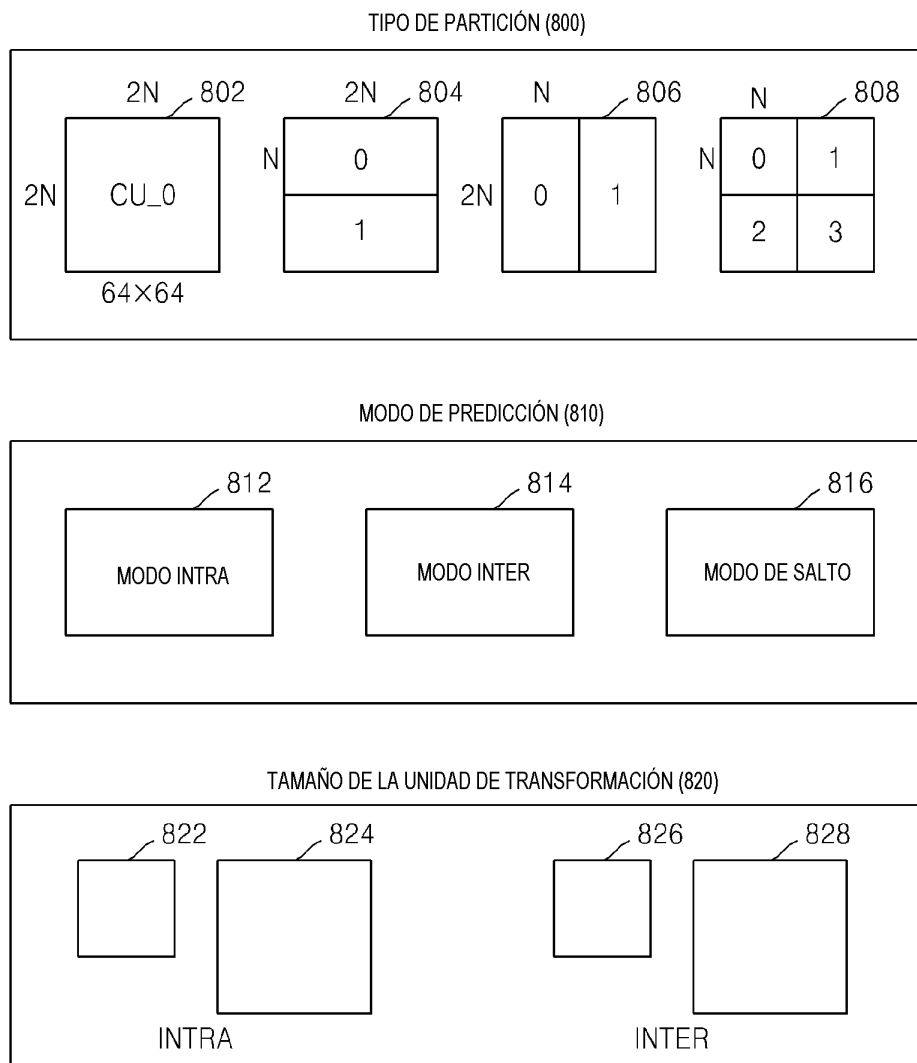


FIG. 9

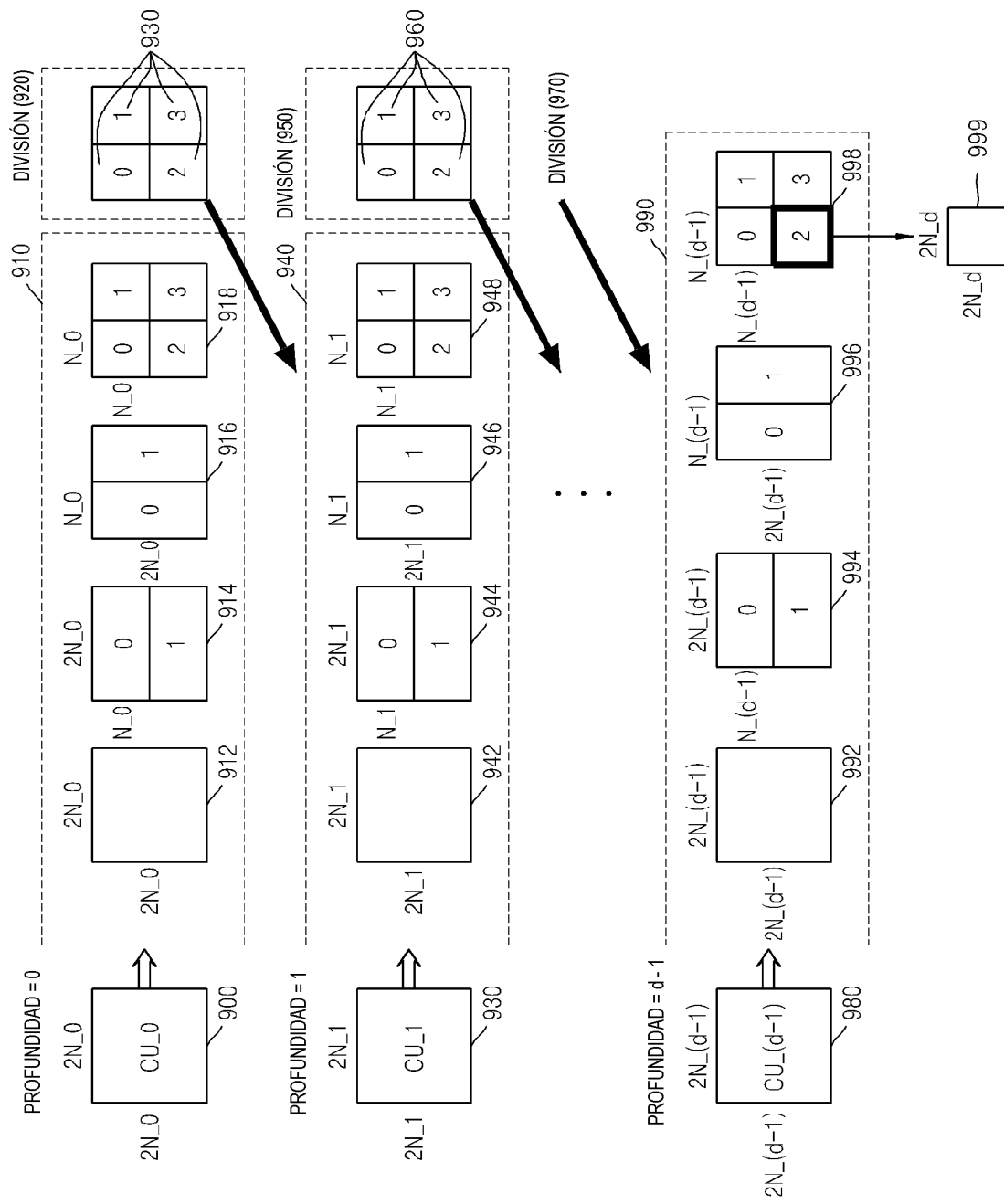


FIG. 10

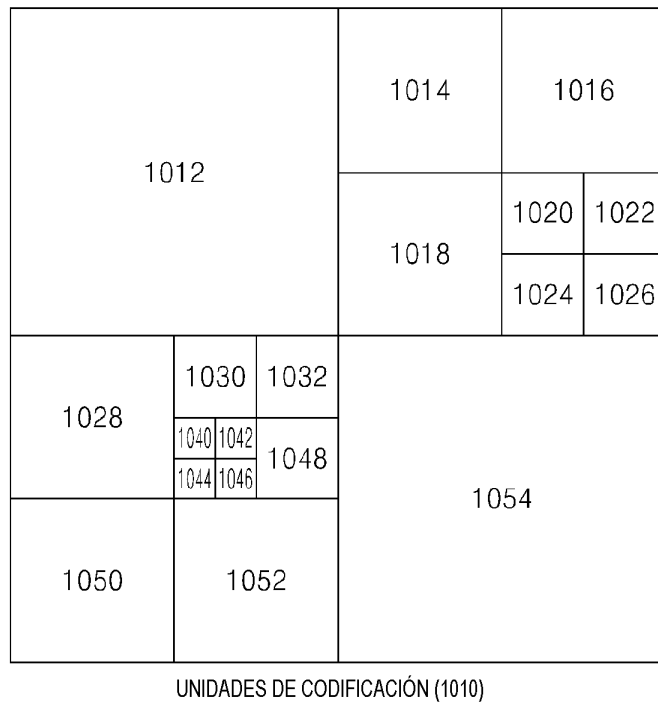


FIG. 11

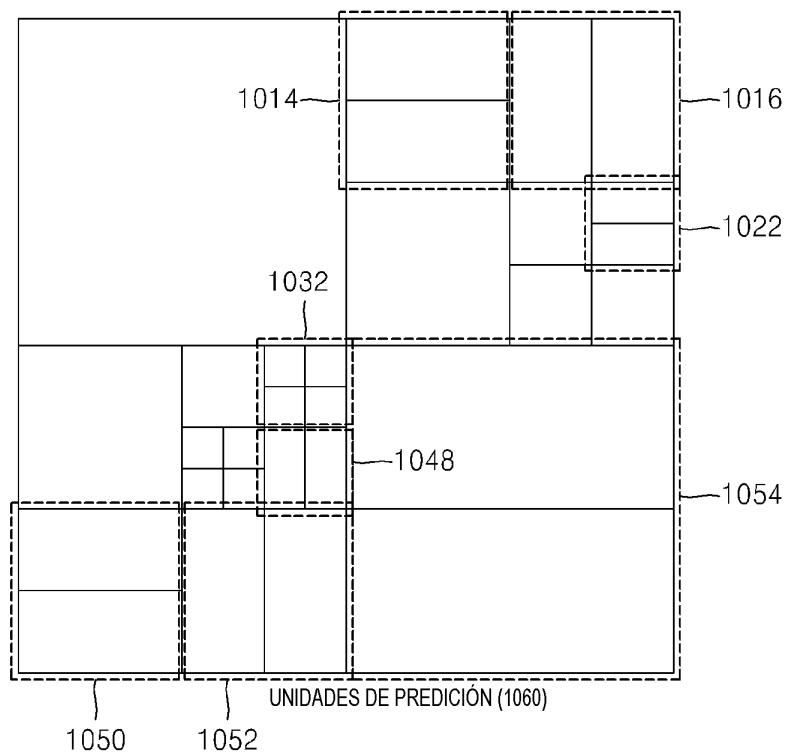


FIG. 12

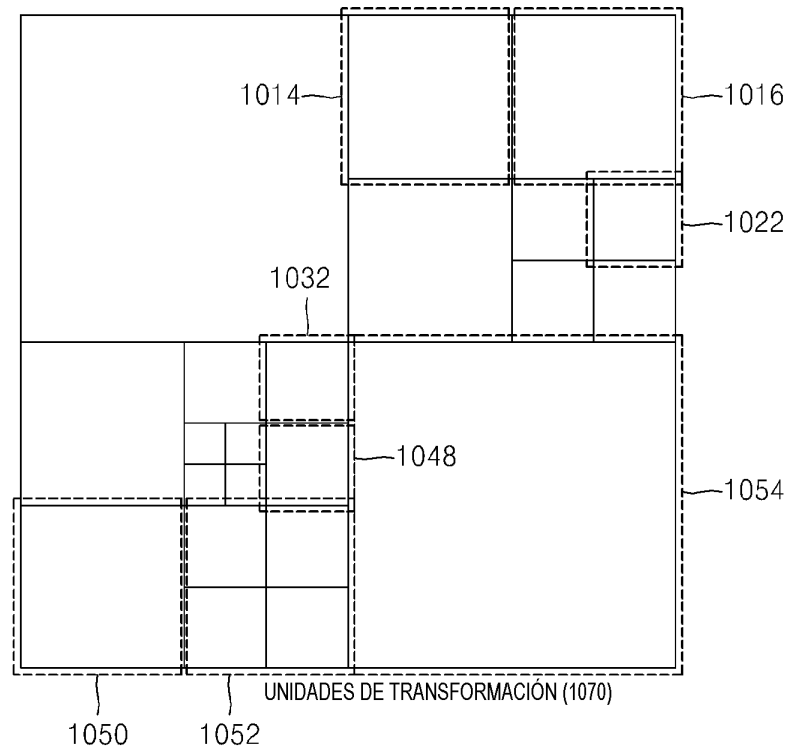


FIG. 13

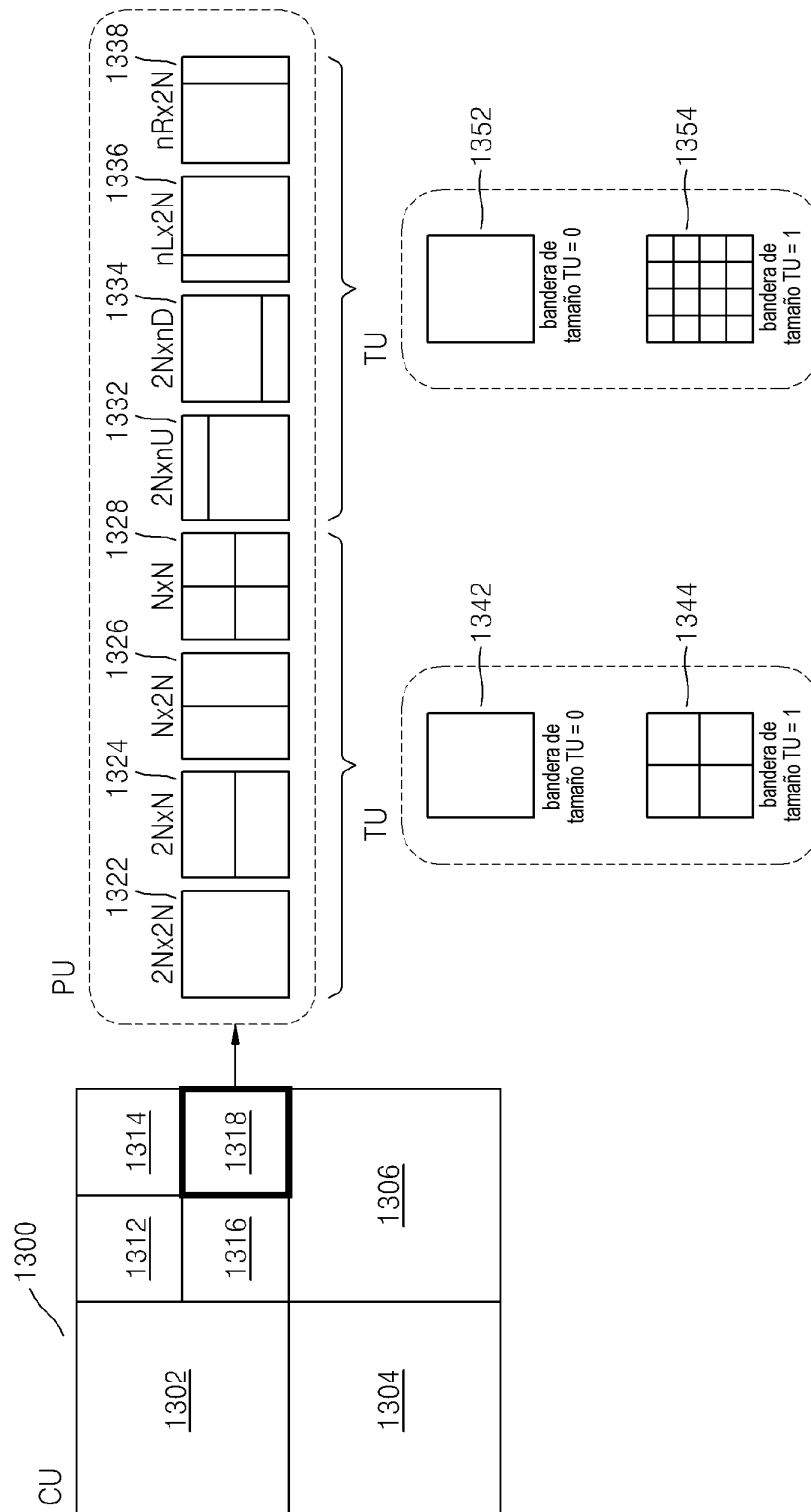


FIG. 14A

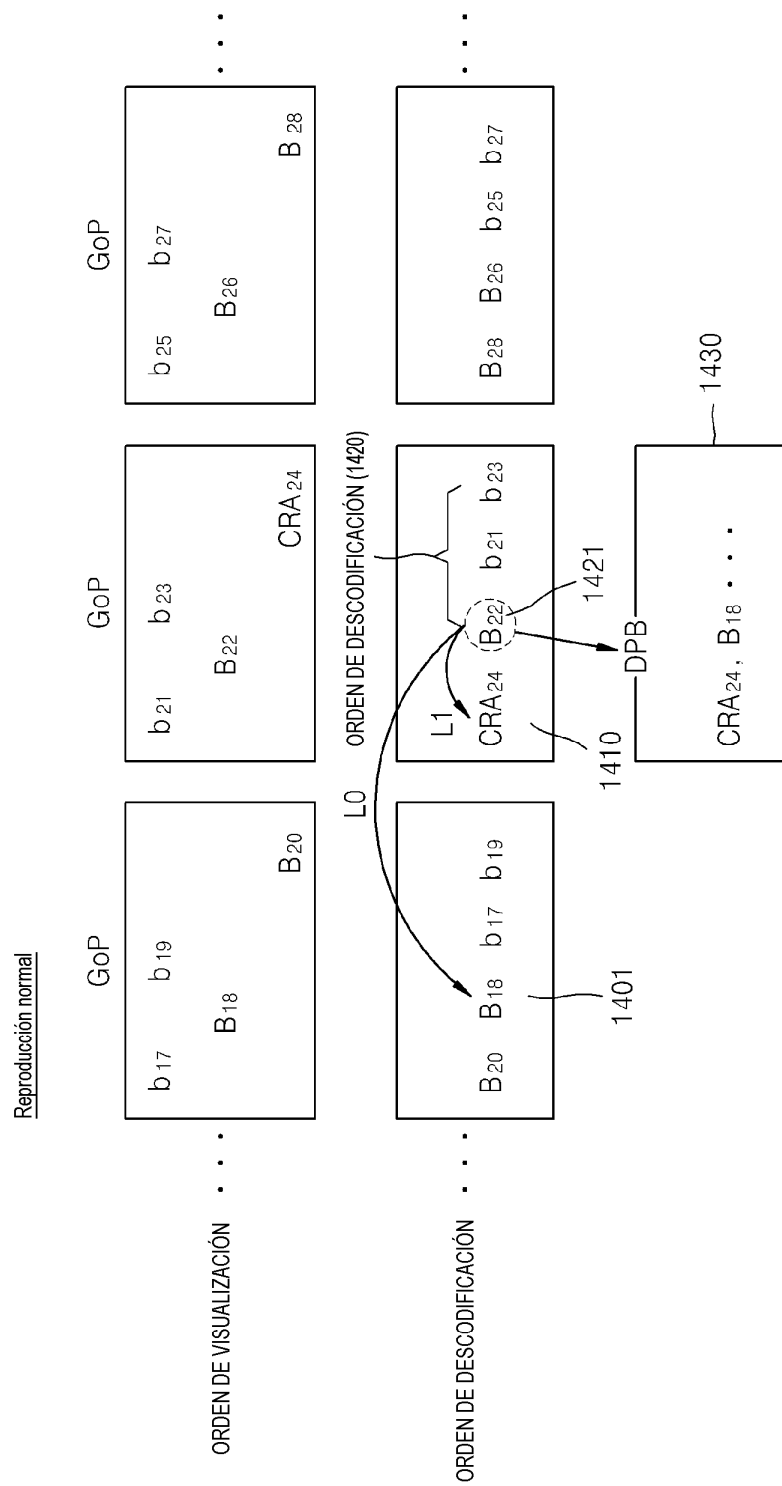


FIG. 14B

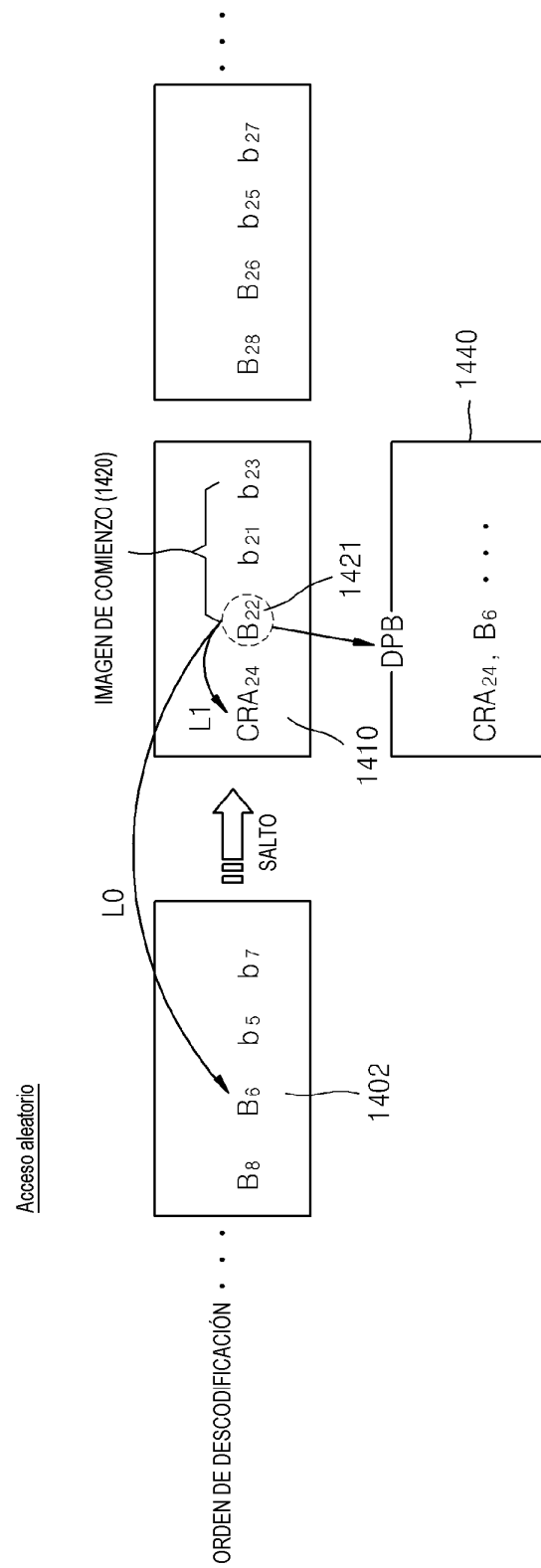


FIG. 15

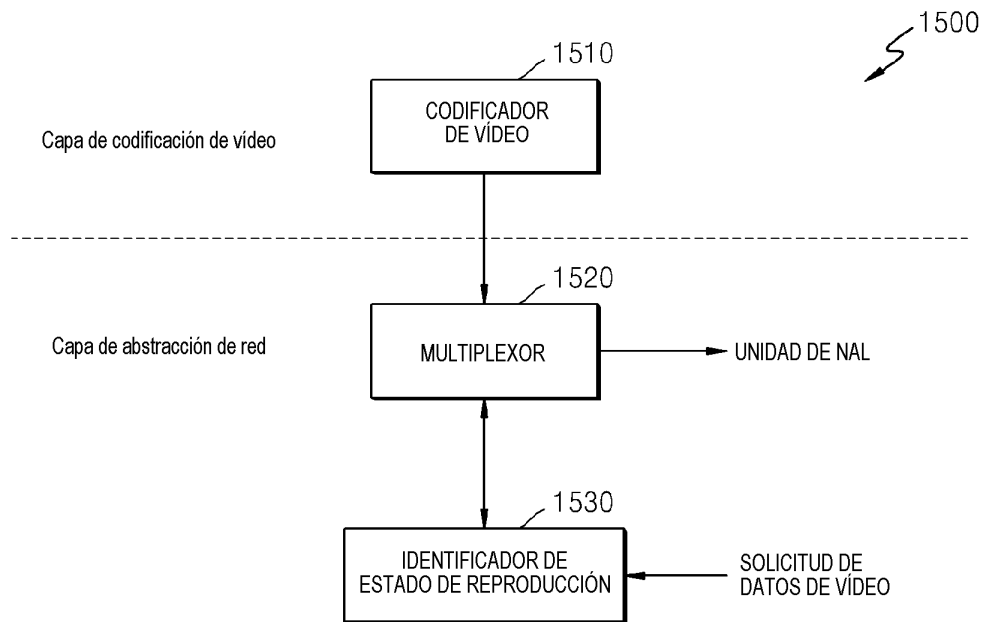


FIG. 16

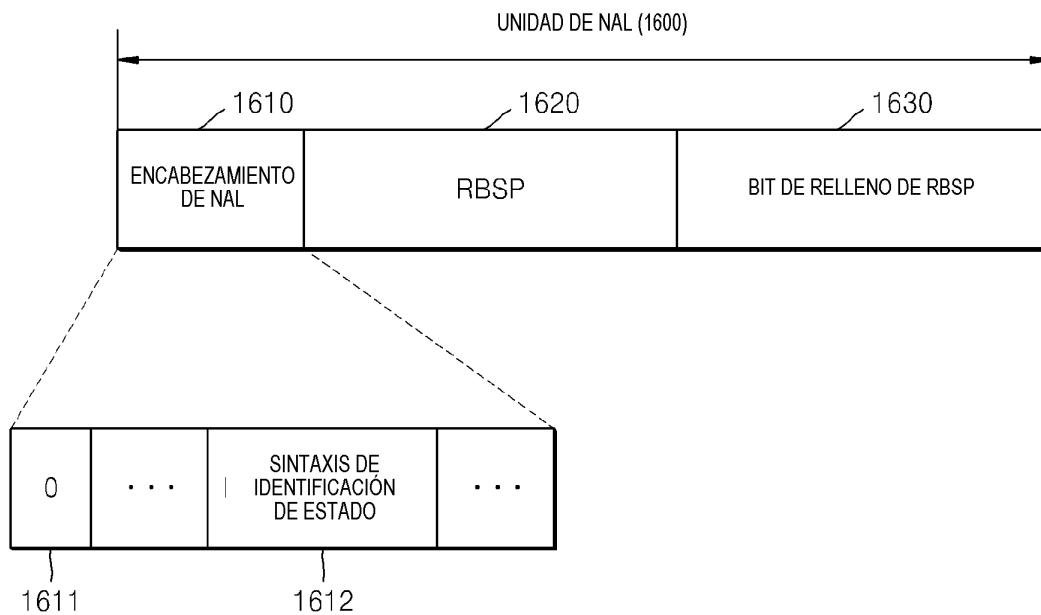


FIG. 17A

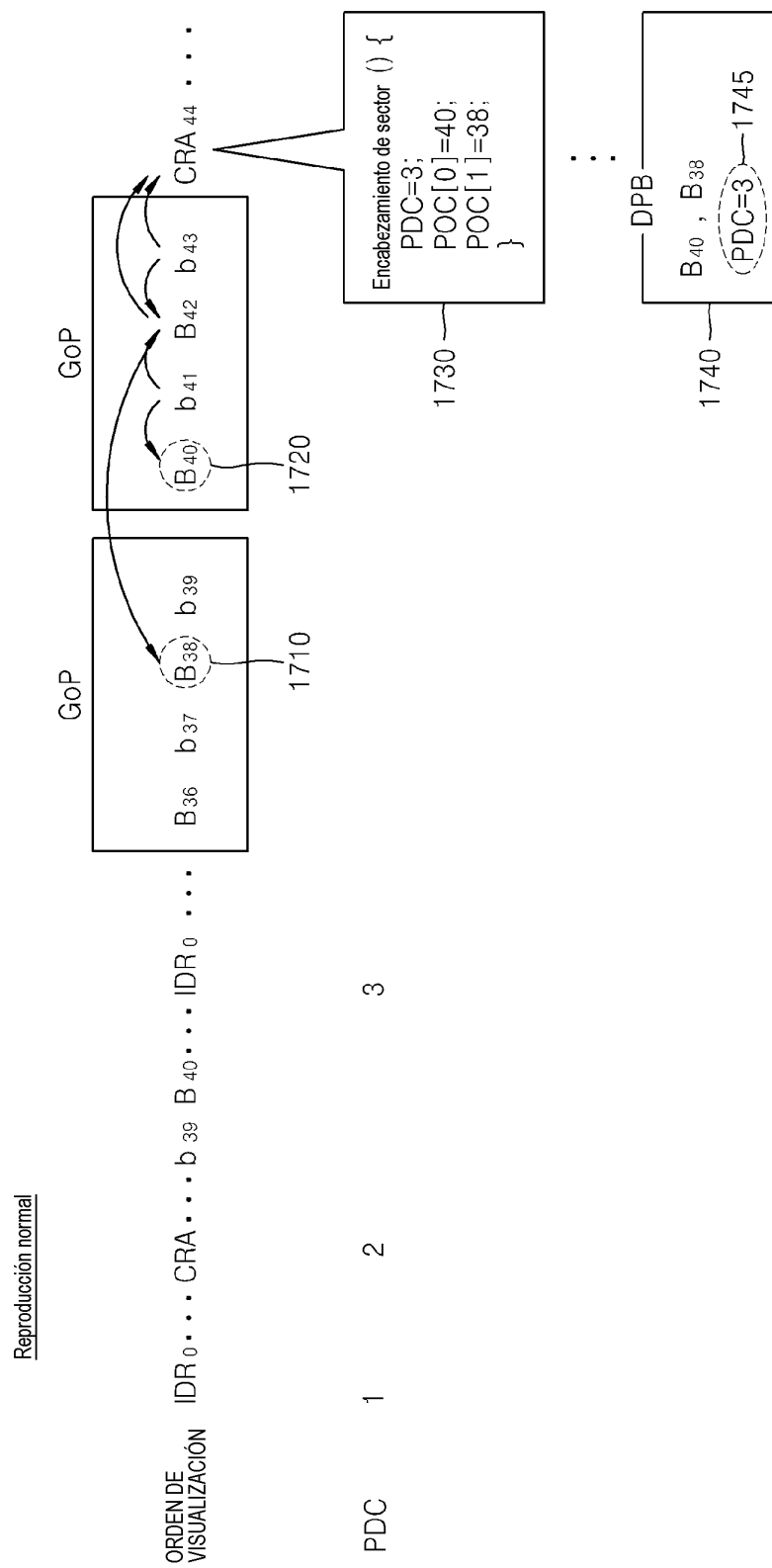


FIG. 17B

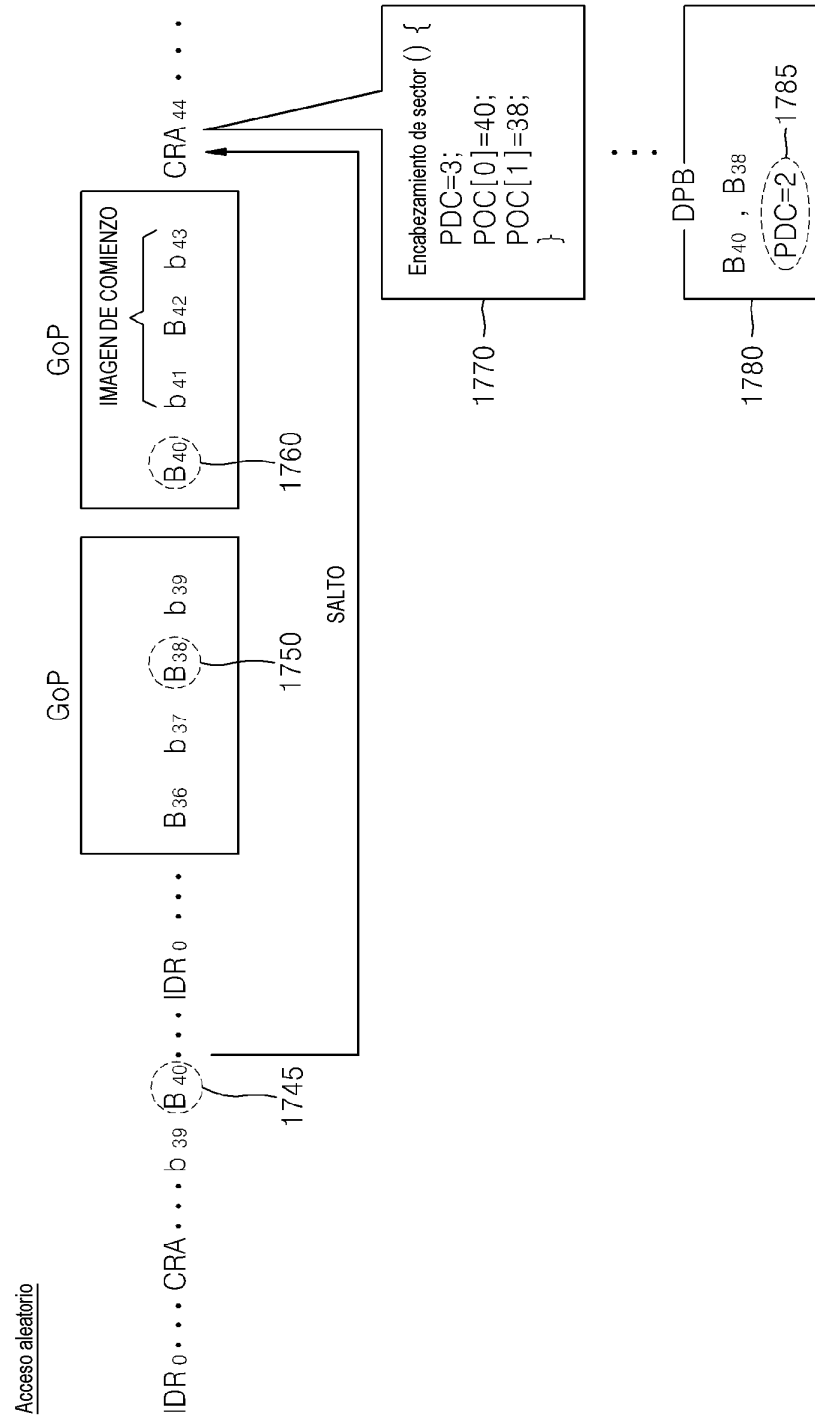


FIG. 18

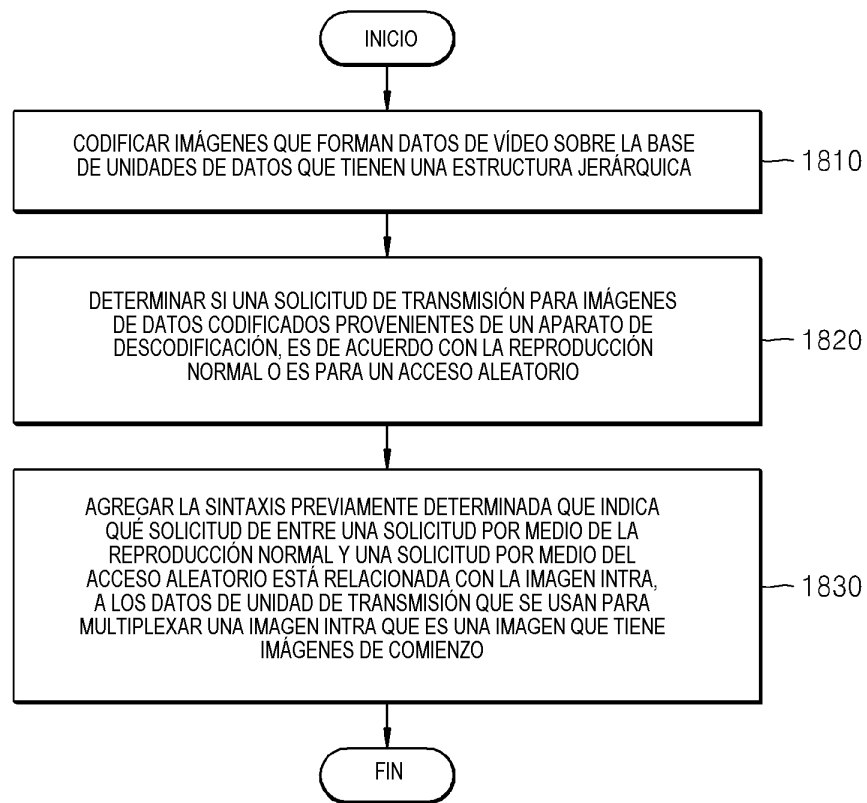


FIG. 19

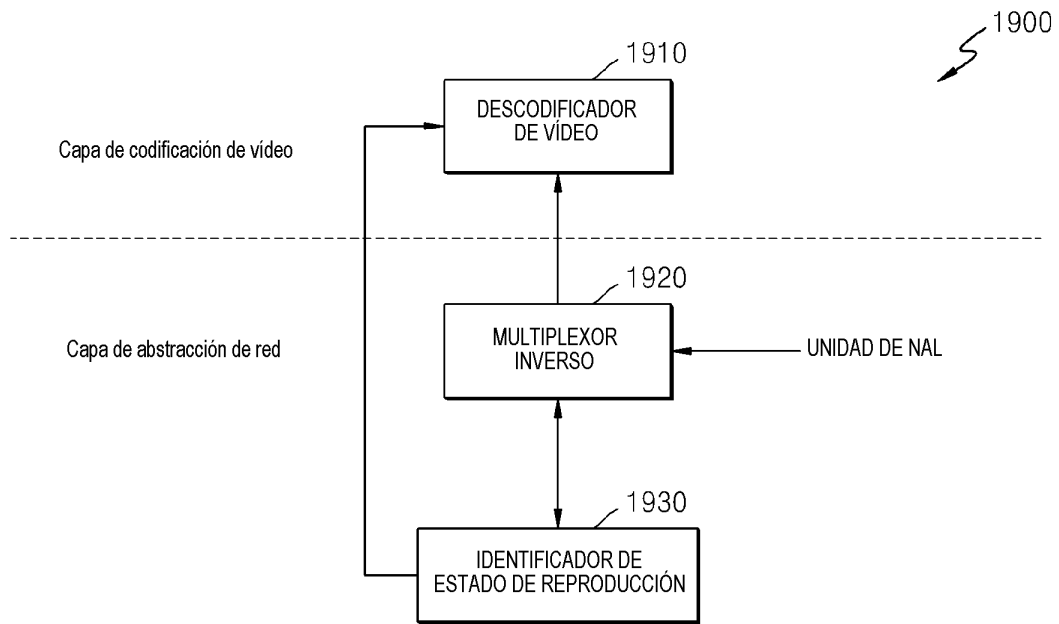


FIG. 20

