

19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 780 126**

51 Int. Cl.:

H04N 19/176 (2014.01)
H04N 19/96 (2014.01)
H04N 19/593 (2014.01)
H04N 19/11 (2014.01)
H04N 19/117 (2014.01)
H04N 19/182 (2014.01)
H04N 19/82 (2014.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

96 Fecha de presentación y número de la solicitud europea: **17.08.2010 E 18199130 (8)**

97 Fecha y número de publicación de la concesión europea: **11.03.2020 EP 3448030**

54 Título: **Procedimiento de decodificación de una imagen, aparato de codificación de una imagen y medio legible por ordenador para almacenar datos asociados con un vídeo**

30 Prioridad:

17.08.2009 KR 20090075855

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

24.08.2020

73 Titular/es:

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

72 Inventor/es:

**MIN, JUNG-HYE;
HAN, WOO-JIN y
KIM, IL-KOO**

74 Agente/Representante:

GONZÁLEZ PECES, Gustavo Adolfo

ES 2 780 126 T3

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

DESCRIPCIÓN

Procedimiento de decodificación de una imagen, aparato de codificación de una imagen y medio legible por ordenador para almacenar datos asociados con un vídeo

Campo técnico

- 5 La presente invención se refiere a un procedimiento para decodificar una imagen, un aparato para codificar una imagen y un medio legible a ordenador para almacenar datos asociados con un vídeo, capaz de mejorar la eficiencia de compresión de vídeo al realizar intra predicción usando píxeles vecinos filtrados.

Antecedentes

- 10 En procedimientos de compresión de vídeo tales como MPEG-1, MPEG-2, MPEG-4 y Codificación de Vídeo Avanzada H.264/MPEG-4 (AVC), una imagen se divide en macrobloques para codificar un vídeo. Después de eso, cada macrobloque se codifica de acuerdo con todos los modos de codificación disponibles en inter predicción e intra predicción, y luego un modo de codificación se selecciona de acuerdo con una velocidad de bits requerida para codificar el macrobloque y la distorsión entre el macrobloque original y un macrobloque decodificado, para de esta manera codificar el macrobloque.
- 15 Al desarrollarse y proporcionarse hardware para reproducir y almacenar contenido de vídeo de alta resolución o alta calidad, aumenta la necesidad de un códec de vídeo para codificar y decodificar de manera efectiva el contenido de vídeo de alta resolución o alta calidad. En un códec de vídeo convencional, un vídeo se codifica de acuerdo con un modo de predicción limitado en base a un macrobloque que tiene un tamaño predeterminado.
- 20 El documento WIEN M, "Variable block-size transforms for H.264/AVC", IEEE TRANSACTIONS ON CIRCUITS AND SYSTEMS FOR VIDEO TECHNOLOGY, IEEE SERVICE CENTER, PISCATAWAY, NJ, USA, vol. 13, n.º 7, 1 de julio de 2003 (páginas 604-613) describe un concepto para codificación de transformada de tamaño de bloque variable. Se llaman transformadas de tamaño de bloque adaptativo (ABT) y se propuso para codificación de alta resolución y vídeo entrelazado en el estándar de codificación de vídeo emergente H.264/AVC. La idea básica de ABT inter es alinear el tamaño de bloque usado para la codificación de transformada del error de predicción al tamaño del bloque usado para la compensación de movimiento. La ABT intra emplea predicción de tamaño de bloque variable y transformadas para codificación. Con ABT, se explota la máxima longitud de señal factible para la codificación de transformada. Los resultados de la simulación revelan un aumento del rendimiento de hasta el 12 % de ahorro de velocidad total y 0,9 dB en la relación de señal y ruido pico.
- 25
- 30 La contribución KIM J Y COL.: "Enlarging MB size for high fidelity video coding beyond HD", 36. VCEG MEETING; 8-10-2008 - 10-10-2008; SAN DIEGO, USA; (VIDEO CODINGEXPERTS GROUP OF ITU-T SG.16), n.º VCEG-AJ21, 5 de octubre de 2008 extiende la arquitectura de H.264 con tamaños MB ampliados. Por tanto, las divisiones de modo de macrobloque (MB) constituyen una estructura jerárquica de manera vertical y el tipo mb en la sintaxis MB del H.264 actual se extiende a los modos de bloque grande de forma que las cantidades de bit totales para representar el tipo mb se vuelven eficaces.
- 35

Divulgación

Problema técnico

En un códec de vídeo convencional, un vídeo se codifica de acuerdo con un modo de predicción limitada en base a un macrobloque que tiene un tamaño predeterminado.

Solución técnica

La presente invención proporciona un procedimiento para decodificar una imagen, un aparato para codificar una imagen y un medio legible a ordenador para almacenar datos asociados con un vídeo según las reivindicaciones adjuntas, capaz de mejorar la eficiencia de compresión de vídeo al filtrar píxeles vecinos de un bloque actual y realizar intra predicción en el bloque actual usando los píxeles vecinos filtrados.

Efectos ventajosos

De acuerdo con la presente invención, se puede mejorar la eficiencia de codificación.

Descripción de los dibujos

- La figura 1 es un diagrama de bloques de un aparato para codificar un vídeo, según una realización ejemplar;
 La figura 2 es un diagrama de bloques de un aparato para decodificar un vídeo;
 La figura 3 es un diagrama para describir un concepto de unidades de codificación;
 La figura 4 es un diagrama de bloques de un codificador de imagen en base a unidades de codificación según una realización ejemplar;
 La figura 5 es un diagrama de bloques de un decodificador de imagen en base a unidades de codificación;
- 50

- La figura 6 es un diagrama que ilustra unidades de codificación más profundas de acuerdo con profundidades, y divisiones;
- La figura 7 es un diagrama para describir una relación entre una unidad de codificación y unidades de transformación;
- 5 La figura 8 es un diagrama para describir información de codificación de unidades de codificación que corresponden a una profundidad codificada;
- La figura 9 es un diagrama de unidades de codificación más profundas de acuerdo con profundidades;
- Las figuras 10 a 12 son diagramas para describir una relación entre unidades de codificación, unidades de predicción y unidades de transformación;
- 10 La figura 13 es un diagrama para describir una relación entre una unidad de codificación, una unidad de predicción o una división, y una unidad de transformación, de acuerdo con la información de modo de codificación de la tabla 1;
- La figura 14 es un diagrama de bloques de un aparato de intra predicción;
- 15 La figura 15 es una tabla que muestra los números de modos de intra predicción de acuerdo con los tamaños de unidades de codificación;
- Las figuras 16A a 16C ilustran modos de intra predicción aplicados a una unidad de codificación que tiene un tamaño predeterminado;
- La figura 17 ilustra modos de intra predicción aplicados a una unidad de codificación que tiene un tamaño predeterminado;
- 20 Las figuras 18A a 18C son diagramas para describir modos de intra predicción que tienen varias directividades;
- La figura 19 es un diagrama que ilustra una unidad de codificación actual y píxeles vecinos que serán filtrados;
- La figura 20 es un diagrama para describir un proceso de filtrar píxeles vecinos;
- La figura 21 es un diagrama de flujo que ilustra un procedimiento de codificación de vídeo; y
- 25 La figura 22 es un diagrama de flujo que muestra un procedimiento de decodificación de vídeo según una realización ejemplar.

Mejor modo

- De acuerdo con un aspecto de una realización de ejemplo, se proporciona un procedimiento de decodificación de una imagen tal como se indica en la reivindicación 1 adjunta.
- 30 De acuerdo con otro aspecto de una realización de ejemplo, se proporciona un aparato de codificación de una imagen como se indica en la reivindicación 2 adjunta.
- De acuerdo con otro aspecto de una realización de ejemplo, se proporciona un medio legible a ordenador para almacenar datos asociados con un vídeo como se indica en la reivindicación 3 adjunta.

Modo para la invención

- 35 A continuación, se describirán ejemplos útiles para entender la invención más completamente con referencia a las figuras adjuntas. En la presente memoria descriptiva, una unidad puede o no referirse a una unidad de tamaño, dependiendo del contexto. En la presente memoria descriptiva, una imagen puede indicar una imagen fija para un vídeo o una imagen en movimiento, es decir, el propio vídeo.
- A continuación, una "unidad de codificación" es una unidad de datos de codificación en la cual los datos de imagen son codificados en un lado de codificador y una unidad de datos codificados en la cual los datos de imagen codificados son decodificados en un lado de decodificador, de acuerdo con ejemplos de realización. Asimismo, una profundidad codificada significa una profundidad donde una unidad de codificación es codificada.
- 40 En primer lugar, se describirá con referencia a las figuras 1 a 13 un procedimiento y un aparato para codificar vídeo y un procedimiento y un aparato para decodificar vídeo.
- La figura 1 es un diagrama de bloques de un aparato 100 de codificación de vídeo según una realización ejemplar.
- 45 El aparato 100 de codificación de vídeo incluye un divisor 110 de unidades de codificación máximas, un determinador 120 de unidades de codificación y una unidad 130 de salida.
- El divisor 110 de unidades de codificación máximas puede dividir una imagen actual en base a 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, datos de imagen de la imagen actual pueden ser divididos en la por lo menos una unidad de
- 50 codificación máxima. La unidad de codificación máxima de acuerdo con un ejemplo de realización puede ser una unidad de datos que tenga 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 sea un cuadrado que tenga una anchura y una altura en cuadrados de 2. Los datos de imagen pueden ser enviados al determinador 120 de unidades de codificación de acuerdo con la por lo menos una unidad de codificación máxima.
- 55 Una unidad de codificación según una realización ejemplar puede caracterizarse por un tamaño máximo y una profundidad. La profundidad indica un número de veces en que la unidad de codificación es dividida espacialmente a

partir de la unidad de codificación máxima, y al hacerse más profunda o incrementarse la profundidad, unidades de codificación más profundas de acuerdo con profundidades pueden dividirse 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. Ya que un tamaño de una unidad de codificación que corresponde a cada profundidad se reduce al hacerse más profunda la profundidad de la unidad de codificación máxima, una unidad de codificación que corresponda a una profundidad superior puede incluir una pluralidad de unidades de codificación que correspondan a profundidades más bajas.

Como se describió anteriormente, los datos de imagen de la imagen actual se dividen en las unidades de codificación máximas de acuerdo con un tamaño máximo de la unidad de codificación y cada una de las unidades de codificación máximas puede incluir unidades de codificación más profundas que se dividan de acuerdo con profundidades. Ya que la unidad de codificación máxima de acuerdo con un ejemplo de realización se divide de acuerdo con profundidades, los datos de imagen de un dominio espacial incluido en la unidad de codificación máxima pueden clasificarse jerárquicamente de acuerdo con profundidades.

Una profundidad máxima y un tamaño máximo de una unidad de codificación, las cuales limitan el número total de veces en que una altura y una anchura de la unidad de codificación máxima se dividen jerárquicamente puede predefinirse.

El determinador 120 de unidades de codificación codifica por lo menos una región dividida obtenida al dividir una región de la unidad de codificación máxima de acuerdo con profundidades, y determina una profundidad para enviar datos de imagen codificados finalmente de acuerdo con la por lo menos una región dividida. En otras palabras, el determinador 120 de unidades de codificación determina una profundidad codificada al codificar los datos de imagen en las unidades de codificación más profundas de acuerdo con profundidades, de acuerdo con la unidad de codificación máxima de la imagen actual, y selecciona una profundidad que tenga el mínimo error de codificación. Así, los datos de imagen codificados de la unidad de codificación que corresponden a la profundidad codificada determinada son finalmente enviados. También, las unidades de codificación que correspondan 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 130 de salida.

Los datos de imagen en la unidad de codificación máxima se codifican en base a las unidades de codificación más profundas que corresponden a por lo menos una profundidad igual a o debajo de la profundidad máxima, y los resultados de codificar los datos de imagen se comparan en base a cada una de las unidades de codificación más profundas. Una profundidad que tenga el mínimo error de codificación puede seleccionarse después de comparar errores de codificación de las unidades de codificación más profundas. Al menos, una profundidad codificada puede seleccionarse para cada unidad de codificación máxima.

El tamaño de la unidad de codificación máxima se divide al ser dividida jerárquicamente en una unidad de codificación de acuerdo con profundidades, y al incrementarse el número de unidades de codificación. Asimismo, incluso si las unidades de codificación corresponden a la misma profundidad en una unidad de codificación máxima, se determina si se divide cada una de las unidades de codificación que corresponda a la misma profundidad hasta una profundidad más baja al medir un error de codificación de los datos de imagen de cada unidad de codificación, por separado. En consecuencia, incluso cuando datos de imagen se incluyen 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 regiones en una de las unidades de codificación máximas, y de esta manera las profundidades codificadas pueden diferir de acuerdo con regiones en los datos de imagen. Así, una o más profundidades codificadas pueden determinarse en una unidad de codificación máxima, y los datos de imagen de la unidad de codificación máxima pueden dividirse de acuerdo con unidades de codificación de al menos una profundidad codificada.

En consecuencia, el determinador 120 de unidades de codificación puede determinar unidades de codificación que tengan una estructura de árbol incluida en la unidad de codificación máxima. Las unidades de codificación que tienen una estructura de árbol de acuerdo con un ejemplo de realización incluyen unidades de codificación que corresponden a una profundidad determinada como la profundidad codificada, de entre todas las unidades de codificación más profundas incluidas en la unidad de codificación máxima. Una unidad de codificación de una profundidad codificada puede determinarse jerárquicamente de acuerdo con profundidades en la misma región de la unidad de codificación máxima, y puede determinarse independientemente en regiones diferentes. De forma similar, una profundidad codificada en una región actual puede determinarse independientemente a partir de una profundidad codificada en otra región.

Una profundidad máxima de acuerdo con un ejemplo de realización es un índice relacionado con el número de veces de división a partir de una unidad de codificación máxima hasta una unidad de codificación mínima. Una primera profundidad máxima de acuerdo con una realización ejemplar puede indicar el número total de veces de división a partir de la unidad de codificación máxima hasta la unidad de codificación mínima. Una segunda profundidad máxima de acuerdo con un ejemplo de realización puede indicar el número total de niveles de profundidad a partir de la unidad de codificación máxima hasta 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 se divide una vez, puede establecerse en 1, y una profundidad de una unidad de codificación, en la cual la unidad de codificación máxima se divide dos veces, puede establecerse en 2. Aquí, 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 profundidades con 0, 1, 2, 3 y 4, y de esta manera la primera profundidad máxima puede establecerse en 4, y la segunda profundidad máxima puede establecerse en 5.

La codificación por predicción y transformación pueden realizar de acuerdo con la unidad de codificación máxima. La codificación por predicción y la transformación se realizan también en base a 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 realizar de acuerdo con el procedimiento de transformación ortogonal o transformación de números enteros.

Ya que el número de unidades de codificación más profundas se incrementa siempre que la unidad de codificación máxima se divide de acuerdo con profundidades, la codificación que incluye la codificación de predicción y la transformación se realiza en todas las unidades de codificación más profundas generadas al hacerse más profunda la profundidad. Para facilidad de descripción, la codificación de predicción y la transformación se describirán ahora en base a una unidad de codificación de una profundidad actual, en una unidad de codificación máxima.

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

Por ejemplo, el aparato 100 de codificación de vídeo 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 de esta manera realizar la codificación por predicción en los datos de imagen en la unidad de codificación.

Para realizar codificación por predicción en la unidad de codificación máxima, la codificación por predicción puede realizarse en base a una unidad de codificación que corresponda a una profundidad codificada, es decir, en base a una unidad de codificación que ya no se divide en unidades de codificación que correspondan a una profundidad más baja. A continuación, la unidad de codificación ya no se divide más y se vuelve una unidad básica para la codificación por predicción será llamada ahora una unidad de predicción. Una división obtenida al dividir la unidad de predicción puede incluir una unidad de predicción o una unidad de datos obtenida al dividir al menos una de una altura y una anchura de la unidad de predicción.

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

Un modo de predicción de la unidad de predicción puede ser al menos uno de un modo intra, un modo inter y un modo de salto. Por ejemplo, el modo intra o el modo inter pueden realizarse en la división de $2N \times 2N$, $2N \times N$, $N \times 2N$ o $N \times N$. Asimismo, el modo de salto puede realizarse solo en la división de $2N \times 2N$. La codificación se realiza independientemente en una unidad de predicción en una unidad de codificación, seleccionando así un modo de predicción que tenga al menos un error de codificación.

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

Para realizar la transformación en la unidad de codificación, la transformación puede realizarse en base a una unidad de datos que tenga un tamaño más pequeño que o igual a la unidad de codificación. Por 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.

Una unidad de datos usada como una base de la transformación será llamada ahora una unidad de transformación. Una profundidad de transformación que indique el número de veces de división para alcanzar la unidad de transformación al dividir la altura y la anchura de la unidad de codificación también puede establecerse 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 sea también $2N \times 2N$, puede ser 1 cuando cada una de la altura y la anchura de la unidad de codificación actual se divida en dos partes iguales, dividida totalmente en 4^1 unidades de transformación, y el tamaño de la unidad de transformación es entonces $N \times N$, y puede ser 2 cuando cada una de la altura y la anchura de la unidad de codificación actual se divida en cuatro partes iguales, divididas totalmente en 4^2 unidades de transformación y el tamaño de la unidad de transformación sea entonces $N/2 \times N/2$. Por ejemplo, la unidad de transformación puede establecerse de acuerdo con una estructura de árbol jerárquica, en

la cual una unidad de transformación de una profundidad de transformación superior se divide en cuatro unidades de transformación de una profundidad de transformación inferior de acuerdo con las características jerárquicas de una profundidad de transformación.

5 De manera similar a la unidad de codificación, la unidad de transformación en la unidad de codificación puede dividirse recursivamente en regiones de tamaño más pequeño, de tal forma que la unidad de transformación puede determinarse independientemente en unidades de regiones. Así, datos residuales en la unidad de codificación pueden dividirse de acuerdo con la transformación que tenga la estructura de árbol de acuerdo con profundidades de transformación.

10 La información de codificación de acuerdo con unidades de codificación que corresponden a una profundidad codificada requiere no solo información acerca de la profundidad codificada, sino también acerca de la información relacionada con codificación por predicción y transformación. En consecuencia, el determinador 120 de unidades de codificación no solo determina una profundidad codificada que tiene un mínimo error de codificación, sino también determina un tipo de división en una unidad de predicción, un modo de predicción de acuerdo con unidades de predicción y un tamaño de una unidad de transformación para transformación.

15 Las unidades de codificación de acuerdo con una estructura de árbol en una unidad de codificación máxima y un procedimiento para determinar una división, de acuerdo con realizaciones ejemplares, se describirán en detalle más adelante con referencia a las figuras 3 a 12.

20 El determinador 120 de unidades de codificación puede medir un error de codificación de unidades de codificación más profundas de acuerdo con profundidades usando Optimización de Distorsión de Velocidad en base a multiplicadores Lagrangianos.

La unidad de salida 130 envía los datos de imagen de la unidad de codificación máxima, que son codificados en base a la por lo menos una profundidad codificada determinada por el determinador 120 de unidades de codificación, y la información acerca del modo de codificación de acuerdo con la profundidad codificada, en flujos de bits.

Los datos de imagen codificados pueden obtenerse al codificar datos residuales de una imagen.

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

30 La información acerca de la profundidad codificada puede definirse usando información de división de acuerdo con profundidades, que indique si se realiza la codificación en unidades de codificación de una profundidad más baja en lugar de una profundidad actual. Si la profundidad actual de la unidad de codificación actual es la profundidad codificada, datos de imagen en la unidad de codificación actual son codificados y enviados, y de esta manera la información de división puede definirse para no 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 se realiza en la unidad de codificación de la profundidad más baja, y de esta manera la información de división puede definirse para dividir la unidad de codificación actual y obtener las unidades de codificación de la profundidad más baja.

40 Si la profundidad actual no es la profundidad codificada, la codificación se realiza en la unidad de codificación que se divide en la unidad de codificación de la profundidad más baja. Ya que al menos una unidad de codificación de la profundidad más baja existe en una unidad de codificación de la profundidad actual, la codificación se realiza repetidamente en cada unidad de codificación de la profundidad más baja, y de esta manera la codificación puede realizarse de manera recursiva para las unidades de codificación que tengan la misma profundidad.

45 Ya que las unidades de codificación que tienen una estructura de árbol se determinan para una unidad de codificación máxima, e información acerca de al menos un modo de codificación se determina para una unidad de codificación de una profundidad codificada, información acerca de al menos un modo de codificación puede determinarse 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 ubicaciones, toda vez que los datos de imagen son divididos jerárquicamente de acuerdo con profundidades, y de esta manera la información acerca de la profundidad codificada y el modo de codificación puede establecerse para los datos de imagen.

50 En consecuencia, la unidad 130 de salida puede asignar información de codificación acerca de una profundidad codificada correspondiente y un modo de codificación o por lo menos una de la unidad de codificación, la unidad de predicción y una unidad mínima incluida en la unidad de codificación máxima.

55 La unidad mínima de acuerdo con un ejemplo de realización es una unidad de datos rectangular obtenida al dividir la unidad de codificación mínima que constituye la profundidad más baja entre 4. Como alternativa, la unidad mínima puede ser una unidad de datos rectangular máxima que pueda ser incluida en todas las unidades de codificación, unidades de predicción, unidades de división y unidades de transformación incluidas en la unidad de codificación máxima.

Por ejemplo, la información de codificación enviada a través de la unidad de salida 130 puede clasificarse en información de codificación de acuerdo con unidades codificadas, e información de codificación de acuerdo con unidades de predicción. La información de codificación de acuerdo con las unidades de codificación puede incluir la información acerca del modo de predicción y acerca del tamaño de las divisiones. La información de codificación de acuerdo con las unidades de predicción puede incluir información acerca de una dirección estimada de un modo inter, acerca de un índice de imágenes de referencia del modo inter, acerca de un vector de movimiento, acerca de un componente cromático de un modo intra, y acerca de un procedimiento de interpolación del modo intra. Asimismo, información acerca de un tamaño máximo de la unidad de codificación definida de acuerdo con imágenes, segmentos o GOP, e información acerca de una profundidad máxima puede insertarse en SPS (Conjunto de Parámetros de Secuencia) o un encabezado de una secuencia de bits.

En el aparato 100 de codificación de vídeo, la unidad de codificación más profunda puede ser una unidad de codificación obtenida al dividir una altura o una anchura de una unidad de codificación de una profundidad superior, que está una capa por encima, entre dos. En otras palabras, cuando el tamaño de la unidad de codificación de la profundidad actual es $2N \times 2N$, el tamaño de la unidad de codificación de la profundidad más baja 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 la unidad de codificación de la profundidad más baja.

En consecuencia, el aparato 100 de codificación de vídeo puede formar las unidades de codificación que tengan la estructura de árbol al determinar unidades de codificación que tengan una forma óptima y un tamaño óptimo para cada unidad de codificación máxima, en base al tamaño de la unidad de codificación máxima y la profundidad máxima determinada considerando características de la imagen actual. Asimismo, ya que la codificación puede realizarse en cada unidad de codificación máxima usando cualquiera de varios modos de predicción y transformaciones, un modo de codificación óptimo puede determinarse considerando características de la unidad de codificación de varios tamaños de imagen.

Así, si una imagen que tenga alta resolución o gran cantidad de datos se codifica en un macrobloque convencional, un número de macrobloques por imagen se incrementa excesivamente. En consecuencia, un número de piezas de información comprimida generada para cada macrobloque se incrementa, y de esta manera es difícil transmitir la información comprimida y la eficiencia de compresión de datos se reduce. Sin embargo, usando el aparato 100 de codificación de vídeo, la eficiencia de compresión de imágenes puede incrementarse, ya que una unidad de codificación se ajusta mientras se consideran características de una imagen al incrementar un tamaño máximo de una unidad de codificación mientras se considera un tamaño de la imagen.

La figura 2 es un diagrama de bloques de un aparato 200 de decodificación de vídeo, de acuerdo con un ejemplo que no entra dentro del ámbito de la presente invención.

El aparato 200 de decodificación de vídeo incluye un receptor 210, un extractor 220 de datos de imagen e información de codificación, y un decodificador 230 de datos de imagen. Las definiciones de varios términos, tales como una unidad de codificación, una profundidad, una unidad de predicción, una unidad de transformación e información acerca de varios modos de codificación, para varias operaciones del aparato 200 de decodificación de vídeo son idénticas a las descritas con referencia a la figura 1 y al aparato de codificación de vídeo 100.

El receptor 210 recibe y analiza una secuencia de bits de un vídeo codificado. El extractor 220 de datos de imagen e información de codificación extrae datos de imagen codificados para cada unidad de codificación a partir de la secuencia de bits analizada, donde 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 decodificador 230 de datos de imagen. El extractor 220 de datos de imagen e información de codificación puede extraer información acerca de un tamaño máximo de una unidad de codificación de una imagen actual, de un encabezado acerca de la imagen actual o SPS.

Asimismo, el extractor 220 de datos de imagen e información de codificación extrae información acerca de una profundidad codificada y un modo de codificación para las unidades de codificación que tengan una estructura de árbol de acuerdo con cada unidad de codificación máxima, de la secuencia de bits analizada. La información extraída acerca de la profundidad codificada y el modo de codificación es enviada al decodificador 230 de datos de imagen. En otras palabras, los datos de imagen en una secuencia de bits se dividen en la unidad de codificación máxima de tal forma que el decodificador 230 de datos de imagen decodifique los datos de imagen para cada unidad de codificación máxima.

La información acerca de la profundidad codificada y el modo de codificación de acuerdo con la unidad de codificación máxima puede establecerse para información acerca de por lo menos una unidad de codificación, que corresponda a la profundidad codificada, e información acerca de un modo de codificación puede incluir información acerca de un tipo de división de una unidad de codificación correspondiente que corresponda a la profundidad codificada, acerca de un modo de predicción, y un tamaño de una unidad de transformación. Asimismo, la información de división de acuerdo con profundidades puede extraerse como la información acerca de la profundidad codificada.

La información acerca de la profundidad codificada y el modo de codificación de acuerdo con cada unidad de codificación máxima extraída por el extractor 220 de datos de imagen e información de codificación es información

acerca de una profundidad codificada y un modo de codificación determinado para generar un error de codificación mínimo cuando un codificador, tal como el aparato 100 de codificación de vídeo, realice repetidamente codificación para cada unidad de codificación más profunda de acuerdo con profundidades de conformidad con cada unidad de codificación máxima. En consecuencia, el aparato 200 de decodificación de vídeo puede restablecer una imagen al decodificar los datos de imagen de acuerdo con una profundidad codificada y un modo de codificación que genere el error de codificación mínimo.

Ya que información de codificación acerca de la profundidad codificada y el modo de codificación puede asignarse a una unidad de datos predeterminada de entre una unidad de codificación, una unidad de predicción y una unidad mínima correspondientes, el extractor 220 de datos de imagen e información de codificación puede extraer la información acerca de la profundidad codificada y el modo de codificación de acuerdo con las unidades de datos predeterminadas. Las unidades de datos predeterminadas a las cuales se asigna la misma información acerca de la profundidad codificada y el modo de codificación pueden inferirse como siendo las unidades de datos incluidas en la misma unidad de codificación máxima.

El decodificador 230 de datos de imagen almacena la imagen actual al decodificar los datos de imagen en cada unidad de codificación máxima con base en la información acerca de la profundidad codificada y el modo de codificación de acuerdo con las unidades de codificación máximas. En otras palabras, el decodificador 230 de datos de imagen puede decodificar los datos de imagen codificados con base en la información extraída acerca del tipo de división, el modo de predicción y la unidad de transformación para cada unidad de codificación de entre las unidades de codificación que tengan la estructura de árbol incluidas en cada unidad de codificación máxima. Un proceso de decodificación puede incluir una predicción que incluya intra predicción y compensación de movimiento, y una transformación inversa. La transformación inversa se puede realizar de acuerdo con el procedimiento de transformación ortogonal inversa o transformación de números enteros inversa.

El decodificador 230 de datos de imagen puede realizar intra predicción o compensación de movimiento de acuerdo con una división y un modo de predicción de cada unidad de codificación, en base a la información acerca del tipo de división y el modo de predicción de la unidad de predicción de la unidad de codificación de acuerdo con profundidades codificadas.

Asimismo, el decodificador 230 de datos de imagen puede realizar transformación inversa de acuerdo con cada unidad de transformación en la unidad de codificación, en base a la información acerca del tamaño de la unidad de transformación de la unidad de codificación de acuerdo con profundidades codificadas, para de esta manera realizar la transformación inversa de acuerdo con unidades de codificación máximas.

El decodificador 230 de datos de imagen puede determinar por lo menos una profundidad codificada de una unidad de codificación máxima actual usando información de división de acuerdo con 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 decodificador 230 de datos de imagen puede decodificar datos codificados de al menos una unidad de codificación que corresponda a cada profundidad codificada en la unidad de codificación máxima actual usando la información acerca del tipo de división de la unidad de predicción, el modo de predicción y el tamaño de la unidad de transformación para cada unidad de codificación que corresponda a la profundidad codificada, y enviar los datos de imagen de la unidad de codificación máxima actual.

En otras palabras, las unidades de datos que contengan la información de codificación que incluya la misma información de división pueden acumularse al observar el conjunto de información de codificación asignado para la unidad de datos predeterminada de entre la unidad de codificación, la unidad de predicción y la unidad mínima, y las unidades de datos acumuladas pueden considerarse como una unidad de datos que será decodificada por el decodificador 230 de datos de imagen en el mismo modo de codificación.

El aparato 200 de decodificación de vídeo puede obtener información acerca de al menos una unidad de codificación que genere el error de codificación mínimo cuando la codificación se lleve a cabo recursivamente para cada unidad de codificación máxima, y puede usar la información para decodificar la imagen actual. En otras palabras, las unidades de codificación que tengan estructura de árbol determinada como unidades de codificación óptimas en cada unidad de codificación máxima pueden ser decodificadas. También, el tamaño máximo de unidad de codificación se determina considerando la resolución y una cantidad de datos de imagen.

En consecuencia, incluso si datos de imagen tienen alta resolución y una gran cantidad de datos, los datos de imagen pueden ser decodificados y restablecidos eficientemente usando un tamaño de una unidad de codificación y un modo de codificación, que se determinan de manera adaptiva de acuerdo con características de los datos de imagen, al usar información acerca de un modo de codificación óptimo recibido desde un codificador.

Se describirá ahora con referencia a las figuras 3 a 13 un procedimiento para determinar unidades de codificación que tengan una estructura de árbol, una unidad de predicción y una unidad de transformación.

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

Un tamaño de una unidad de codificación puede expresarse en anchura x altura, y puede ser 64 x 64, 32 x 32, 16 x

ES 2 780 126 T3

16 y 8 x 8. Una unidad de codificación de 64 x 64 puede dividirse en divisiones de 64 x 64, 64 x 32, 32 x 64 o 32 x 32, y una unidad de 32 x 32 puede dividirse en divisiones de 32 x 32, 32 x 16, 16 x 32 o 16 x 16, una unidad de codificación de 16 x 16 puede dividirse en divisiones de 16 x 16, 16 x 8, 8 x 16 u 8 x 8, y una unidad de codificación de 8x8 puede dividirse en divisiones de 8 x 8, 8 x 4, 4 x 8 o 4 x 4.

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

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

- 15 Ya que la profundidad máxima de los datos 310 de vídeo es 2, las unidades 315 de codificación de los datos 310 de vídeo pueden incluir una unidad de codificación máxima que tenga un tamaño de eje largo de 64, y unidades de codificación que tengan tamaños de eje largo de 32 y 16, toda vez que las profundidades son profundizadas hasta dos capas al dividir la unidad de codificación máxima dos veces. Mientras tanto, ya que la profundidad máxima de los datos 330 de vídeo es 1, las unidades de codificación 335 de los datos 330 de vídeo pueden incluir una unidad de codificación máxima que tenga un tamaño de eje largo de 16, y unidades de codificación que tengan un tamaño de eje largo de 8, toda vez que las profundidades son profundizadas hasta una capa al dividir la unidad de codificación máxima una vez.

- 20 Ya que la profundidad máxima de los datos 320 de vídeo es 3, las unidades 325 de codificación de los datos 320 de vídeo pueden incluir una unidad de codificación máxima que tenga un tamaño de eje largo de 64, y unidades de codificación que tengan tamaños de eje largo de 32, 16 y 8. toda vez que las profundidades son profundizadas hasta 3 capas al dividir la unidad de codificación máxima tres veces. Al hacerse más profunda una profundidad, la información detallada puede expresarse en forma precisa.

La figura 4 es un diagrama de bloques de un codificador 400 de imágenes en base a unidades de codificación, según una realización ejemplar.

- 30 El codificador de imágenes 400 realiza operaciones del determinador 120 de unidades de codificación del aparato 100 de codificación de vídeo para codificar datos de imagen. En otras palabras, un intra predictor 410 realiza intra predicción en unidades de codificación en un modo intra, de entre una trama 405 actual, y un estimador 420 de movimiento y un compensador 425 de movimiento realizan inter estimación y compensación de movimiento en unidades de codificación en un modo inter de entre la trama 405 actual usando la trama 405 actual, y la trama 495 de referencia.

- 35 Los datos enviados desde el intra predictor 410; el estimador 420 de movimiento y el compensador 425 de movimiento son enviados como un coeficiente de transformación cuantificado a través de un transformador 430 y un cuantificador 440. El coeficiente de transformación cuantificado se restablece como datos en un dominio espacial a través de un cuantificador 460 inverso y un transformador 470 inverso, y los datos restablecidos en el dominio espacial son enviados como tramas 495 de referencia después de haber sido post-procesadas a través de una unidad 480 de desbloqueo y una unidad 490 de filtración recurrente. El coeficiente de transformación cuantificado puede enviarse como un flujo 455 de bits a través de un codificador 450 por entropía.

- 40 Para que el codificador 400 de imágenes sea aplicado en el aparato 100 de codificación de vídeo, todos los elementos del codificador 400 de imágenes, es decir, el intra predictor 410, el estimador 420 de movimiento, el compensador 425 de movimiento, el transformador 430, el cuantificador 440, el codificador 450 por entropía, el cuantificador 460 inverso, el transformador 470 inverso, la unidad 480 de desbloqueo y la unidad 490 de filtración recurrente realizan operaciones en base a cada unidad de codificación de entre unidades de codificación que tienen una estructura de árbol mientras considera la profundidad máxima de cada unidad de codificación máxima.

- 45 Específicamente, el intra predictor 410, el estimador 420 de movimiento y el compensador 425 de movimiento determinan divisiones y un modo de predicción de cada unidad de codificación de entre las unidades de codificación que tengan una estructura de árbol mientras consideran 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 tengan una estructura de árbol.

La figura 5 es un diagrama de bloques de un decodificador 500 de imágenes en base a unidades de codificación.

- 55 Un analizador 510 analiza datos de imagen codificados que serán decodificados e información acerca de la codificación requerida para decodificar a partir de un flujo 505 de bits. Los datos de imagen codificados son enviados como datos cuantificados inversos a través de un decodificador 520 por entropía y un cuantificador 530 inverso, y los

datos cuantificados inversos son restablecidos a datos de imagen en un dominio espacial a través de un transformador 540 inverso.

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

10 Los datos de imagen en el dominio espacial, que pasaron a través del intra predictor 550 y el compensador 560 de movimiento, pueden ser enviados como una trama 595 restablecida después de haber sido post-procesados a través de una unidad 570 de desbloqueo y una unidad 580 de filtración recurrente. Asimismo, los datos de imagen que son post-procesados a través de la unidad 570 de desbloqueo y la unidad 580 de filtración recurrente pueden ser enviados como la trama 585 de referencia.

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

15 Para que el decodificador 500 de imágenes sea aplicado en el aparato 200 de decodificación de vídeo, todos los elementos del decodificador 500 de imágenes, es decir, el analizador 510, el decodificador 520 por entropía, el cuantificador 530 inverso, el transformador 540 inverso, el intra predictor 550, el compensador 560 de movimiento, la unidad 570 de desbloqueo y la unidad 580 de filtración recurrente realizan operaciones en base a unidades de codificación que tienen una estructura de árbol para cada unidad de codificación máxima.

20 Específicamente, el intra predictor 550 y el compensador 560 de movimiento realizan operaciones en base a divisiones y un modo de predicción para cada una de las unidades de codificación que tengan una estructura de árbol, y el transformador 540 inverso realiza operaciones en base a un tamaño de una unidad de transformación para cada unidad de codificación.

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

25 El aparato 100 de codificación de vídeo y el aparato 200 de decodificación de vídeo usan unidades de codificación jerárquicas para considerar así las características de una imagen. Una altura máxima, una anchura máxima y una profundidad máxima de unidades de codificación pueden determinarse en forma adaptiva de acuerdo con las características de la imagen, o pueden establecerse de manera diferente por un usuario. Los tamaños de unidades de codificación más profundas de acuerdo con profundidades pueden determinarse de acuerdo con el tamaño máximo predeterminado de la unidad de codificación.

30 En una estructura 600 jerárquica de unidades de codificación, según una realización ejemplar, la altura máxima y la anchura máxima de las unidades de codificación son cada una 64, y la profundidad máxima es 4. Ya que una profundidad se hace más profunda a lo largo de un eje vertical de la estructura 600 jerárquica, una altura y una anchura de la unidad de codificación más profunda se dividen cada una. Asimismo, una unidad de predicción y divisiones, que son las bases para la predicción de acuerdo con cada unidad de codificación más profunda, se muestran a lo largo de un eje horizontal de la estructura 600 jerárquica.

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

45 La unidad de predicción y las divisiones de una unidad de codificación están dispuestas a lo largo del eje horizontal de acuerdo con cada profundidad. En otras palabras, si la unidad 610 de codificación que tiene el tamaño de 64 x 64 y la profundidad de 0 es una unidad de predicción, la unidad de predicción puede dividirse en divisiones que incluyan la unidad 610 de codificación, es decir, una división 610 que tenga un tamaño de 64 x 64, divisiones 612 que tengan el tamaño de 64 x 32, divisiones 614 que tengan el tamaño de 32 x 64, o divisiones 616 que tengan el tamaño de 32 x 32.

50 De manera similar, una unidad de predicción de la unidad 620 de codificación que tenga el tamaño de 32 x 32 y la profundidad de 1 puede dividirse en divisiones incluidas en la unidad de codificación 620, es decir, una división 620 que tenga un tamaño de 32 x 32, divisiones 622 que tengan un tamaño de 32 x 16, divisiones 624 que tengan un tamaño de 16 x 32, y divisiones 626 que tengan un tamaño de 16 x 16.

55 De manera similar, una unidad de predicción de la unidad 630 de codificación que tenga el tamaño de 16 x 16 y la profundidad de 2 puede dividirse en divisiones incluidas en la unidad 630 de codificación, es decir, una división que tenga un tamaño de 16 x 16 incluida en la unidad 630 de codificación, divisiones 632 que tengan un tamaño de 16 x 8, divisiones 634 que tengan un tamaño de 8 x 16, y divisiones 636 que tengan un tamaño de 8 x 8.

De manera similar, una unidad de predicción de la unidad 640 de codificación que tenga el tamaño de 8 x 8 y la profundidad de 3 puede dividirse en divisiones incluidas en la unidad 640 de codificación, es decir, una división que tenga un tamaño de 8 x 8 incluida en la unidad 640 de codificación, divisiones 642 que tengan un tamaño de 8 x 4, divisiones 644 que tengan un tamaño de 4 x 8, y divisiones 646 que tengan un tamaño de 4 x 4.

- 5 La unidad 650 de codificación 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 650 de codificación solo se asigna a una división que tenga un tamaño de 4 x 4.

- 10 Para poder determinar la por lo menos una profundidad codificada de las unidades de codificación que constituyen la unidad 610 de codificación máxima, el determinador 120 de unidades de codificación del aparato 100 de codificación de vídeo realiza codificación para unidades de codificación que correspondan a cada profundidad incluida en la unidad 610 de codificación máxima.

- 15 Un número de unidades de codificación más profundas de acuerdo con profundidades que incluyen datos en el mismo intervalo y el mismo tamaño se incrementa al hacerse más profunda la profundidad. Por ejemplo, cuatro unidades de codificación que correspondan a una profundidad de 2 se requieren para cubrir datos que estén incluidos en una unidad de codificación que corresponda a una profundidad de 1. En consecuencia, para poder comparar los resultados de codificación de los mismos datos de acuerdo con profundidades, la unidad de codificación que corresponda a la profundidad de 1 y cuatro unidades de codificación que correspondan a la profundidad de 2 son cada una codificadas.

- 20 Para realizar la codificación para una profundidad actual de entre las profundidades, al menos un error de codificación puede seleccionarse para la profundidad actual al realizar codificación para cada unidad de predicción en las unidades de codificación que correspondan a la profundidad actual, a lo largo del eje horizontal de la estructura 600 jerárquica. Como alternativa, el error de codificación mínimo puede ser buscado al comparar los mínimos errores de codificación de acuerdo con profundidades, al realizar codificación para cada profundidad al hacerse más profunda la profundidad a lo largo del eje vertical de la estructura 600 jerárquica. Una profundidad y una división que tengan el error de codificación mínimo en la unidad 610 de codificación pueden seleccionarse como la profundidad codificada y un tipo de división de la unidad 610 de codificación.

- 25 La figura 7 es un diagrama para describir una relación entre una unidad 710 de codificación y unidades 720 de transformación según una realización ejemplar.

- 30 El aparato 100 o 200 de codificación de vídeo codifica o decodifica una imagen de acuerdo con unidades de codificación que tengan tamaños más pequeños que o iguales a una unidad de codificación máxima para cada unidad de codificación máxima. Los tamaños de unidades de transformación para la transformación durante la codificación pueden seleccionarse con base en unidades de datos que no sean más grandes que una unidad de codificación correspondiente.

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

- 35 Asimismo, datos de la unidad 710 de codificación que tengan el tamaño de 64 x 64 pueden codificarse al realizar la transformación en cada una de las unidades de transformación que tengan el tamaño de 32 x 32, 16 x 16, 8 x 8 y 4 x 4, que sean más pequeñas que 64 x 64, y luego una unidad de transformación que tenga el mínimo error de codificación puede seleccionarse.

- 40 La figura 8 es un diagrama para describir información de codificación de unidades de codificación que corresponden a una profundidad codificada según una realización ejemplar.

La unidad 130 de salida del aparato 100 de codificación de vídeo puede codificar y transmitir información 800 acerca de un tipo de división, información 810 acerca de un modo de predicción, e información 820 acerca de un tamaño de una unidad de transformación para cada unidad de codificación que corresponda a una unidad codificada, como información acerca de un modo de codificación.

- 45 La información 800 indica información acerca de una forma de una división obtenida al dividir una unidad de predicción de una unidad de codificación actual, en el que la división es una unidad de datos para codificación por predicción de la unidad de codificación actual. Por ejemplo, una unidad de codificación actual CU_0 que tenga un tamaño de 2N x 2N puede dividirse en cualquiera de una división 802 que tenga un tamaño de 2N x 2N, una división 804 que tenga un tamaño de 2N x 2N, una división 806 que tenga un tamaño de N x 2 y una división 808 que tenga un tamaño de N x N. Aquí, la información 800 acerca de un tipo de división se establece para indicar una de la división 804 que tiene un tamaño de 2N x 2N, la división 806 que tiene un tamaño de N x 2N, y la división 808 que tiene un tamaño de N x N.

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

- 55 La información 820 indica una unidad de transformación a la que se recurrirá cuando la transformación se realice en

una unidad de codificación actual. Por ejemplo, la unidad de transformación puede ser una primera unidad 822 de intra transformación, una segunda unidad 824 de intra transformación, una primera unidad 826 de inter transformación o una segunda unidad 828 de intra transformación.

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

La figura 9 es un diagrama de unidades de codificación más profundas de acuerdo con profundidades, según una realización ejemplar.

10 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 más baja.

15 Una unidad 910 de predicción para codificación por predicción de una unidad 900 de codificación que tenga una profundidad de 0 y un tamaño de $2N_0 \times 2N_0$ puede incluir divisiones de un tipo de división 912 que tenga un tamaño $2N_0 \times 2N_0$, un tipo de división 914 que tenga un tamaño de $2N_0 \times N_0$, un tipo de división 916 que tenga un tamaño de $N_0 \times 2N_0$, y un tipo de división 918 que tenga un tamaño de $N_0 \times N_0$. La figura 9 solo ilustra dos tipos de división 918 que se obtienen al dividir simétricamente la unidad 910 de predicción, pero un tipo de división no está limitado a lo mismo, y las divisiones de la unidad de predicción pueden incluir divisiones asimétricas, divisiones que tengan una forma predeterminada, y divisiones que tengan una forma geométrica.

20 La codificación por predicción se realiza repetidamente en una división que tiene un tamaño de $2N_0 \times 2N_0$, dos divisiones que tienen un tamaño de $2N_0 \times N_0$, dos divisiones que tienen un tamaño de $N_0 \times 2N_0$, y cuatro divisiones que tienen un tamaño de $N_0 \times N_0$, de acuerdo con cada tipo de división. La codificación por predicción en un modo intra y un modo inter se puede realizar en las divisiones que tengan los tamaños de $2N_0 \times 2N_0$, $2N_0 \times N_0$, y $N_0 \times N_0$. La codificación por predicción en un modo de salto se realiza solo en la división que tiene el tamaño de $2N_0 \times 2N_0$.

25 Los errores de codificación que incluyen la codificación por predicción en los tipos de división 912 a 918 son comparados, y el mínimo error de codificación se determina entre los tipos de división. Si un error de codificación es más pequeño en uno de los tipos de división 912 a 916, la unidad 910 de predicción puede no dividirse en una profundidad más baja.

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

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

35 Si un error de codificación es el más pequeño en el tipo de división 948, una profundidad se cambia de 1 a 2 para dividir el tipo de división 948 en la operación 950, y la codificación se realiza repetidamente en las unidades 960 de codificación, las cuales tienen una profundidad de 2 y un tamaño de $N_2 \times N_2$ para buscar un error de codificación mínimo.

40 Cuando una profundidad máxima es d , la operación de división de acuerdo con cada profundidad puede realizarse hasta cuando una profundidad se vuelva $d-1$, y la información de división puede codificarse como hasta cuando una profundidad sea una de 0 a $d-2$. En otras palabras, cuando se realiza codificación hasta cuando la profundidad es $d-1$ después de que una unidad de codificación que corresponde a una profundidad de $d-2$ es dividida en la operación 970, una unidad 990 de predicción para codificación por predicción de una unidad 980 de codificación que tenga una profundidad de $d-1$ y un tamaño de $2N_{(d-1)} \times 2N_{(d-1)}$ puede incluir divisiones de un tipo de división 992 que tenga un tamaño de $2N_{(d-1)} \times 2N_{(d-1)}$, un tipo de división 994 que tenga un tamaño de $2N_{(d-1)} \times N_{(d-1)}$, un tipo de división 996 que tenga un tamaño de $N_{(d-1)} \times 2N_{(d-1)}$, y un tipo de división 998 que tenga un tamaño de $N_{(d-1)} \times N_{(d-1)}$.

50 La codificación por predicción puede realizarse repetidamente en una división que tenga un tamaño de $2N_{(d-1)} \times 2N_{(d-1)}$, dos divisiones que tengan un tamaño de $2N_{(d-1)} \times N_{(d-1)}$, dos divisiones que tengan un tamaño de $N_{(d-1)} \times 2N_{(d-1)}$, cuatro divisiones que tengan un tamaño de $N_{(d-1)} \times N_{(d-1)}$ de entre los tipos de división 992 a 998 para buscar un tipo de división que tenga un error de codificación mínimo.

55 Incluso cuando el tipo de división 998 tiene el error de codificación mínimo, ya que una profundidad máxima es d , una unidad de codificación $CU_{(d-1)}$ que tenga una profundidad de $d-1$ ya no es dividida a una profundidad más baja, y una profundidad codificada para las unidades de codificación que constituyan una unidad 900 de codificación máxima actual se determina como $d-1$ y un tipo de división de la unidad 900 de codificación máxima actual puede determinarse como $N_{(d-1)} \times N_{(d-1)}$. Asimismo, ya que la profundidad máxima es d y una unidad 980 de codificación mínima que

tiene una profundidad más baja de d-1, ya no es dividida a una profundidad más baja, no se establece la información de división para la unidad 980 de codificación mínima.

Una unidad 999 de datos 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 980 de codificación mínima entre 4. Al realizar la codificación repetidamente, el aparato 100 de codificación de vídeo puede seleccionar una profundidad que tenga el mínimo error de codificación al comparar errores de codificación de acuerdo con profundidades de la unidad 900 de codificación para determinar una profundidad codificada, y establecer un tipo de división y un modo de predicción correspondientes como un modo de codificación de la profundidad codificada.

De esta manera, los errores de codificación mínimos de acuerdo con profundidades se comparan en todas las profundidades de 1 a d, y una profundidad que tenga el mínimo error de codificación puede determinarse como una profundidad codificada. La profundidad codificada, el tipo de división de la unidad de predicción y el modo de predicción pueden codificarse y transmitirse como información acerca de un modo de codificación. Asimismo, ya que una unidad de codificación se divide a partir de una profundidad de 0 hasta una profundidad codificada, solo la información de división de la profundidad codificada se establece en 0, y la información de división de profundidades que excluyan la profundidad codificada se establece en 1.

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

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

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

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

En las unidades 1060 de predicción, algunas unidades 1014, 1016, 1022, 1032, 1048, 150, 1052 y 1054 de codificación se obtienen al dividir las unidades de codificación en las unidades 1010 de codificación. En otras palabras, tipos de división en las unidades 1014, 1022, 1050 y 1054 de codificación tienen un tamaño de $2N \times N$, tipos de división en las unidades 1016, 1048 y 1052 de codificación tienen un tamaño de $N \times 2N$, y un tipo de división de la unidad 1032 de codificación tiene un tamaño de $N \times N$. Las unidades de predicción y divisiones de las unidades 1010 de codificación son más pequeñas que o iguales a cada unidad de codificación.

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

En consecuencia, la codificación se realiza recursivamente en cada una de las unidades de codificación que tienen una estructura jerárquica en cada región de una unidad de codificación máxima para determinar una unidad de codificación óptima, y de esta manera pueden obtenerse unidades de codificación que tengan una estructura de árbol recursiva. La información de codificación puede incluir información de división acerca de una unidad de codificación, información acerca de un tipo de división, información acerca de un modo de predicción e información acerca de un tamaño de una unidad de transformación. La tabla 1 muestra la información de codificación que puede establecerse por los aparatos 100 y 200 de codificación y decodificación de vídeo.

Tabla 1

Información de división 0 (Codificación en la unidad de codificación que tiene un 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

Salto Intra Inter (Solamente 2Nx2N)	Tipo de división simétrica	Tipo de división asimétrica	Información de división 0 de la unidad de transformación	Información de división 1 de la unidad de transformación	unidades de codificación que tienen una profundidad inferior de d+1
	2Nx2N2Nx NNx2NNxN	2NxnU2Nxn DnLx2NnRx 2N	2Nx2N	NxN(Tipo simétrico)N/2xN/2 (Tipo asimétrico)	

5 La unidad 130 de salida del aparato 100 de codificación de vídeo puede enviar la información de codificación acerca de las unidades de codificación que tengan una estructura de árbol, y los datos de imagen y el extractor 220 de datos de imagen e información de codificación del aparato 200 de decodificación de vídeo pueden extraer la información de codificación acerca de las unidades de codificación tengan una estructura de árbol a partir de una secuencia de bits recibida.

10 La información de división indica si una unidad de codificación actual es dividida en unidades de codificación de una profundidad más baja. Si información de división de una profundidad actual d es 0, una profundidad, en la cual una unidad de codificación actual ya no sea dividida en una profundidad más baja, es una profundidad codificada, y de esta manera información acerca de un tipo de división, modo de predicción y un tamaño de una unidad de transformación puede ser definida para la profundidad codificada. Si la unidad de codificación actual se divide más de acuerdo con la información de división, la codificación se realiza independientemente en cuatro unidades de codificación divididas de una profundidad más baja.

15 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 definirse en todos los tipos de división, y el modo de salto se define solo en un tipo de división que tiene un tamaño de 2N x 2N.

20 La información acerca del tipo de división puede incluir tipos de división simétricos que tengan tamaños de 2N x 2N, 2N x N, N x 2N y N x N, los cuales se obtienen al dividir simétricamente una altura o un ancho de una unidad de predicción, y tipos de división asimétricos que tienen tamaños de 2N x nU, 2N x nD, nL x 2N y nR x 2N, los cuales se obtienen al dividir asimétricamente la altura o la anchura de la unidad de predicción. Los tipos de división asimétricos que tienen los tamaños de 2N x nU y 2N x nD pueden obtenerse respectivamente al dividir la altura de la unidad de predicción en 1:3 y 3:1, y los tipos de división asimétricos que tienen los tamaños de nL x 2N y nR x 2N pueden obtenerse respectivamente al dividir la anchura de la unidad de predicción en 1:3 y 3:1.

25 El tamaño de la unidad de transformación puede establecerse para que sea dos tipos en el modo intra y dos tipos en el modo inter. En otras palabras, si la información de división de la unidad de transformación es 0, el tamaño de la unidad de transformación puede ser 2N x 2N, que es el tamaño de la unidad de codificación actual. Si la información de división de la unidad de transformación es 1, las unidades de transformación pueden obtenerse al dividir la unidad de codificación actual. Asimismo, si un tipo de división de la unidad de codificación actual que tiene el tamaño de 2N x 2N es un tipo de división simétrico, un tamaño de una unidad de transformación puede ser N x N, y si el tipo de división de la unidad de codificación actual es un tipo de división asimétrico, el tamaño de la unidad de transformación puede ser N/2 x N/2.

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

40 En consecuencia, se determina si unidades de datos adyacentes están incluidas en la misma unidad de codificación que corresponde a la profundidad codificada al comparar información de codificación de las unidades de datos adyacentes. Asimismo, una unidad de codificación correspondiente que corresponda a una unidad codificada se determina usando información de codificación de una unidad de datos, y de esta manera puede determinarse una distribución de profundidades codificadas en una unidad de codificación máxima.

En consecuencia, si una unidad de codificación actual se predice en base a información de codificación de unidades de datos adyacentes, la información de codificación de unidades de datos en unidades de datos en unidades de codificación más profundas adyacentes a la unidad de codificación actual puede ser referida y usada directamente.

45 Como alternativa, si una unidad de codificación actual se predice con base en información de codificación de unidades de datos adyacentes, unidades de datos adyacentes a la unidad de codificación actual se buscan usando información de codificación de las unidades de datos, y las unidades de codificación adyacentes buscadas pueden ser referidas para predecir la unidad de codificación actual.

La figura 13 es un diagrama para describir una relación entre una unidad de codificación, una unidad de predicción o una división, y una unidad de transformación, de acuerdo con información de modo de codificación de la Tabla 1.

50 Una unidad 1300 de codificación máxima incluye unidades 1302, 1304, 1306, 1312, 1314, 1316 y 1318 de codificación de profundidades codificadas. Aquí, ya que la unidad 1318 de codificación es una unidad de codificación de una

profundidad codificada, la información de división puede establecerse en 0. La información acerca de un tipo de división de la unidad 1318 de codificación que tenga un tamaño de $2N \times 2N$ puede establecerse para que sea uno de un tipo de división 1322 que tenga un tamaño de $2N \times 2N$, un tipo de división 1324 que tenga un tamaño de $2N \times N$, un tipo de división 1326 que tenga un tamaño de $N \times 2N$, un tipo de división 1328 que tenga un tamaño de $N \times N$, un tipo de división 1332 que tenga un tamaño de $2N \times nU$, un tipo de división 1334 que tenga un tamaño de $2N \times nD$, un tipo de división 1336 que tenga un tamaño de $nL \times 2N$, y un tipo de división 1338 que tenga un tamaño de $nR \times 2N$.

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

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

Con referencia a la figura 13, el indicador de tamaño de TU es un indicador que tiene un valor de 0 o 1, pero el indicador de tamaño de TU no está limitado a 1 bit, y una unidad de transformación puede dividirse jerárquicamente para que tenga una estructura de árbol mientras el indicador de tamaño de TU se incremente desde 0.

En este caso, el tamaño de una unidad de transformación que se ha usado realmente puede ser expresado usando un indicador de tamaño de TU de una unidad de transformación, según una realización ejemplar, junto con un tamaño máximo y tamaño mínimo de la unidad de transformación. Según una realización ejemplar, el aparato 100 de codificación de vídeo es capaz de codificar información de tamaño de unidad de transformación máxima, información de tamaño de unidad de transformación mínima, y un indicador de tamaño de TU máximo. El resultado de codificar la información del tamaño de unidad de transformación máxima, la información de tamaño de unidad de transformación mínima y el indicador de tamaño de TU máximo puede insertarse en un SPS. De acuerdo con una realización ejemplar, el aparato 200 de decodificación de vídeo puede decodificar vídeo usando la información de tamaño de unidad de transformación máxima, la información de tamaño de unidad de transformación mínima y el indicador de tamaño de TU máximo.

Por ejemplo, si el tamaño de una unidad de codificación actual es 64×64 y un tamaño de unidad de transformación máxima es 32×32 , entonces el tamaño de una unidad de transformación puede ser 32×32 cuando un indicador de tamaño de TU sea 0, puede ser 16×16 cuando el indicador de tamaño de TU sea 1, y puede ser 8×8 cuando el indicador de tamaño de TU sea 2.

Como otro ejemplo, si el tamaño de la unidad de codificación actual es 32×32 y un tamaño de unidad de transformación mínima es 32×32 , entonces el tamaño de la unidad de transformación puede ser 32×32 cuando el indicador de tamaño de TU sea 0. Aquí, el indicador de tamaño de TU no puede establecerse en un valor que no sea 0, toda vez que el tamaño de la unidad de transformación no puede ser inferior a 32×32 .

Como otro ejemplo, si el tamaño de la unidad de codificación actual es 64×64 y un indicador de tamaño de TU máximo es 1, entonces el indicador de tamaño de TU puede ser 0 o 1. Aquí, el indicador de tamaño de TU no puede establecerse en un valor que no sea 0 o 1.

De esta manera, si se define que el indicador de tamaño de TU máximo es "MaxTransformSizeIndex", un tamaño de unidad de transformación mínima es "MinTransformSize", y un tamaño de unidad de transformación es "RootTuSize" cuando el indicador de TU es 0, entonces un tamaño de unidad de transformación mínima actual "CurrMinTuSize" puede determinarse en una unidad de codificación actual, puede definirse por la ecuación (1):

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

En comparación con el tamaño de unidad de transformación mínima actual "CurrMinTuSize" puede determinarse en la unidad de codificación actual, un tamaño de unidad de transformación "RootTuSize" cuando el indicador del tamaño TU sea 0 puede indicar un tamaño de unidad de transformación máxima que se puede seleccionar en el sistema. En la ecuación (1), "RootTuSize/(2^{MaxTransformSizeIndex})" indica un tamaño de unidad de transformación cuando el tamaño de unidad de transformación "RootTuSize", cuando el indicador de tamaño de TU es 0, se divide un número de veces que corresponde al indicador de tamaño de TU máximo; y "MinTransformSize" indica un tamaño de transformación mínimo. Así, un valor más pequeño de entre "RootTuSize/(2^{MaxTransformSizeIndex})" y "MinTransformSize" puede ser el tamaño de unidad de transformación mínima actual "CurrMinTuSize" que se puede determinar en la unidad de codificación actual.

Según una realización ejemplar, el tamaño de unidad de transformación máxima RootTuSize puede variar de acuerdo con el tipo de un modo de predicción.

Por ejemplo, si un modo de predicción actual es un modo inter, entonces RootTuSize puede determinarse usando la ecuación (2) abajo. En la ecuación (2), "MaxTransformSize" indica un tamaño de unidad de transformación máximo, y

“PUSize” indica un tamaño de unidad de predicción actual.

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

5 Es decir, si el modo de predicción actual es el modo inter, el tamaño de unidad de transformación “RootTuSize” cuando el indicador de tamaño de TU es 0, puede ser un valor más pequeño de entre el tamaño de unidad de transformación máxima y el tamaño de unidad de predicción actual.

Si un modo de predicción de una unidad de división actual es un modo intra, “RootTuSize” puede determinarse usando la ecuación (3) abajo. En la ecuación (3), “PartitionSize” indica el tamaño de la unidad de división actual.

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

10 Es decir, si el modo de predicción actual es el modo intra, el tamaño de unidad de transformación “RootTuSize” cuando el indicador de tamaño de TU es 0 puede ser un valor más pequeño de entre el tamaño de unidad de transformación máxima y el tamaño de la unidad de división actual.

Sin embargo, el tamaño de unidad de transformación máxima actual “RootTuSize” que varía de acuerdo con el tipo de un modo de predicción en una unidad de división es solo un ejemplo y no está limitado a esto.

15 La intra predicción realiza por la unidad 410 de intra predicción del apartado 100 de codificación de vídeo ilustrado en la figura 4 y la unidad 550 de intra predicción del aparato 200 de decodificación de vídeo ilustrado en la figura 5 se describirán ahora en detalle. En la siguiente descripción, una unidad de codificación indica un bloque codificado actual en un proceso de codificación de una imagen, y una unidad de decodificación indica un bloque decodificado actual en un proceso de decodificación de una imagen. La unidad de codificación y la unidad de decodificación son diferentes solo en que la unidad de codificación se usa en el proceso de codificación y la unidad de decodificación se usa en la decodificación. Para la consistencia de términos, excepto para un caso particular, la unidad de codificación y la unidad de decodificación son llamadas una unidad de codificación en ambos procesos de codificación y decodificación. Asimismo, la unidad de codificación puede ser la unidad de predicción, división de predicción y bloque. Asimismo, alguien de capacidad ordinaria en la técnica entendería por la presente memoria descriptiva que un procedimiento y un aparato de intra predicción de acuerdo con un ejemplo también se puede aplicar para realizar intra predicción en un códec de vídeo general.

La figura 14 es un diagrama de bloques de un aparato 1200 de intra predicción según una realización ejemplar.

Con referencia a la figura 14, el aparato 1200 de intra predicción incluye una unidad 1210 de filtración de píxeles vecinos, una unidad 1220 de determinación de píxeles de referencia y una unidad 1230 de realización de intra predicción.

30 La unidad 1210 de filtración de píxeles vecinos filtra píxeles vecinos usados para realizar intra predicción en una unidad de codificación actual que será codificada para generar así píxeles vecinos filtrados. La filtración de los píxeles vecinos se describirá ahora con referencia a las figuras 19 y 20.

La figura 19 es un diagrama que ilustra una unidad 1700 de codificación actual y píxeles 1710 y 1720 adyacentes que serán filtrados según una realización ejemplar.

35 Con referencia a la figura 19, la unidad de filtración de píxeles 1210 adyacentes filtra X píxeles 1710 adyacentes en un lado superior de la unidad 1700 de codificación actual e Y píxeles 1720 adyacentes en un lado izquierdo de la unidad 1700 de codificación actual al menos una vez, para de esta manera generar píxeles vecinos filtrados. Aquí, si la unidad 1700 de codificación actual tiene un tamaño de N x N, la unidad 1210 de filtración de píxeles vecinos puede filtrar 4N píxeles vecinos, tal como 2N píxeles 1710 adyacentes en el lado superior de la unidad 1700 de codificación actual y 2N píxeles 1720 adyacentes en el lado izquierdo de la unidad 1700 de codificación actual. Es decir, X = 2N e Y = 2N. El número de los píxeles 1710 y 1720 adyacentes filtrados por la unidad 1210 de filtración de píxeles vecinos no se limita a esto y puede cambiarse teniendo en cuenta la directividad de un modo de intra predicción aplicado a la unidad 1700 de codificación actual.

45 Asimismo, en la figura 19, si X+Y píxeles 1710 y 1720 adyacentes originales en los lados superior e izquierdo de la unidad 1700 de codificación actual son representados por ContextOrg[n] (donde n es un número entero de 0 a X+Y-1), y un píxel vecino más inferior de los Y píxeles 1720 adyacentes tiene un valor de n = 0, es decir, ContextOrg[0], un píxel vecino más a la derecha de los X píxeles 1710 adyacentes tiene un valor de n = X+Y-1, es decir, ContextOrg[X+Y-1].

La figura 20 es un diagrama para describir un proceso de filtración de píxeles vecinos según una realización ejemplar.

50 Con referencia a la figura 20, si 4N píxeles vecinos originales en lados superior e izquierdo de una unidad de codificación actual que tienen un tamaño de N x N son representados por ContextOrg[n] (donde n es un número entero de 0 a 4N-1), la unidad 1210 de filtración de píxeles vecinos filtra los píxeles vecinos originales al calcular valores promedio ponderados entre los píxeles vecinos originales para generar así primeros píxeles vecinos filtrados ContextFiltered[n]. Por ejemplo, la unidad 1210 de filtración de píxeles vecinos genera los primeros píxeles vecinos

filtrados al aplicar un filtro de 3 tomas a los píxeles vecinos originales ContextOrg[n] como se representa por la ecuación (4).

$$\text{ContextFiltered1}[n] = (\text{ContextOrg}[n-1] + 2*\text{ContextOrg}[n] + \text{ContextOrg}[n+1])/4 \quad (4)$$

5 Con referencia a la ecuación (4), la unidad 1210 de filtración de píxeles vecinos calcula un valor promedio ponderado de un píxel vecino ContextOrg[n] que será actualmente filtrado de entre los píxeles vecinos originales y píxeles vecinos ContextOrg[n-1] y ContextOrg[n+1] ubicados en lados izquierdo y derecho del píxel vecino ContextOrg[n] para generar así un primer píxel vecino filtrado. Los píxeles vecinos filtrados más exteriores de entre los primeros píxeles vecinos filtrados tienen valores de los píxeles vecinos originales. Es decir, ContextFiltered1[0] = ContextOrg[0] y ContextFiltered1[4N-1] = ContextOrg[4N-1].

10 De manera similar, la unidad 1210 de filtración de píxeles vecinos puede calcular valores promedio ponderados entre los primeros píxeles vecinos filtrados ContextFiltered1[n] para de esta manera generar segundos píxeles vecinos filtrados ContextFiltered2[n]. Por ejemplo, la unidad 1210 de filtración de píxeles vecinos genera los segundos píxeles vecinos filtrados al aplicar un filtro de 3 tomas a los primeros píxeles vecinos filtrados ContextFiltered1[n] como se representa por la ecuación (5).

$$15 \quad \text{ContextFiltered2}[n] = (\text{ContextFiltered1}[n-1] + 2*\text{ContextFiltered1}[n] + \text{ContextFiltered1}[n+1])/4 \quad (5)$$

Con referencia a la ecuación (5), la unidad 1210 de filtración de píxeles vecinos calcula un valor promedio ponderado de un píxel vecino ContextFiltered1[n] que será actualmente filtrado de entre los primeros píxeles vecinos filtrados y píxeles vecinos ContextFiltered1[n-1] y ContextFiltered1[n+1] ubicados en lados izquierdo y derecho del píxel vecino ContextFiltered1[n] para generar así un segundo píxel vecino filtrado. Los píxeles vecinos filtrados más exteriores de entre los segundos píxeles vecinos filtrados tienen valores de los primeros píxeles vecinos. Es decir, ContextFiltered2[0]=ContextFiltered1[0] y ContextFiltered2[4N-1] = ContextFiltered1[4N-1]. El proceso de filtración de píxeles vecinos descrito anteriormente puede repetirse más de dos veces. Asimismo, el número de tomas de un filtro para filtrar píxeles vecinos no está limitado a tres como se describió anteriormente y puede cambiarse de manera variable. Asimismo, el número de tomas de un filtro y coeficiente del filtro para filtrar píxeles vecinos puede aplicarse de manera adaptiva.

La unidad 1220 de determinación de píxeles de referencia determina los píxeles vecinos filtrados o los píxeles vecinos originales como píxeles de referencia que se usarán para realizar intra predicción en la unidad de codificación actual. En más detalle, la unidad 1220 de determinación de píxeles de referencia selecciona los píxeles vecinos originales, los primeros píxeles vecinos filtrados o los segundos píxeles vecinos filtrados como los píxeles de referencia de acuerdo con el tamaño de la unidad de codificación actual y el tipo de un modo de intra predicción que se realizará actualmente. Por ejemplo, si un índice de referencia de un modo de predicción que usa los píxeles vecinos originales es 0, un índice de referencia de un modo de predicción que usa los primeros píxeles vecinos filtrados es 1, y un índice de referencia de un modo de predicción que usa los segundos píxeles vecinos filtrados es 2, la unidad 1220 de determinación de píxeles de referencia puede determinar el tipo de píxeles vecinos que se usará para realizar intra predicción de acuerdo con el tamaño de la unidad de codificación actual y el tipo de un modo de intra predicción que se realizará actualmente, como se muestra en la Tabla 2.

Tabla 2

Modo de predicción	Tamaño de la unidad de codificación						Modo de predicción	Tamaño de la unidad de codificación					
	4x4	8x8	16x16	32x32	64x64	NxN (N>64)		4x4	8x8	16x16	32x32	64x64	NxN (N>64)
0													
1	0	1	0	0	0	0	17	-	-	2	2	-	-
2	0	1	0	0	0	0	18	-	-	2	2	-	-
3	0	1	0	0	0	0	19	-	-	2	2	-	-
4	0	1	0	0	0	0	20	-	-	2	2	-	-
5	1	2	2	2	2	2	21	-	-	2	2	-	-
6	1	2	2	2	-	-	22	-	-	2	2	-	-
7	1	2	2	2	-	-	23	-	-	2	2	-	-
8	1	2	2	2	-	-	24	-	-	2	2	-	-
9	-	-	2	2	-	-	25	-	-	2	2	-	-
10	-	-	2	2	-	-	26	-	-	2	2	-	-
11	-	-	2	2	-	-	27	-	-	2	2	-	-
12	-	-	2	2	-	-	28	-	-	2	2	-	-
13	-	-	2	2	-	-	29	-	-	2	2	-	-
14	-	-	2	2	-	-	30	-	-	2	2	-	-
15	-	-	2	2	-	-	31	-	-	2	2	-	-
16	-	-	2	2	-	-	32	-	-	2	2	-	-

Con referencia a la tabla 2, por ejemplo, si la unidad de codificación actual tiene un tamaño de 32 x 32 y se realiza intra predicción usando el modo de intra predicción 4, un índice de referencia es 0 y de esta manera la unidad 1220 de determinación de píxeles de referencia determina los píxeles vecinos originales ContextOrg[n] como el píxel de referencia que se usará para realizar intra predicción en la unidad de codificación actual. Los modos de intra predicción en la Tabla 2 representan modos de intra predicción mostrados en la tabla 3. Asimismo, "-" en la Tabla 2 representa que un modo de intra predicción para un tamaño correspondiente de una unidad de codificación no se define. La Tabla 2 se basa en los modos de intra predicción mostrados en la Tabla 3, y se muestra a manera de ejemplo. A diferencia de la Tabla 3, siempre y cuando diferentes modos de intra predicción se establezcan de acuerdo con los tamaños de unidades de codificación, los índices de referencia en la Tabla 2 pueden establecerse de manera diferente.

Con referencia de nuevo a la figura 14, si la unidad 1220 de determinación de píxeles de referencia determina píxeles de referencia que se usarán para realizar intra predicción en la unidad de codificación actual de entre los píxeles vecinos originales y los píxeles vecinos filtrados, la unidad 1230 de realización de intra predicción realiza intra predicción usando los píxeles de referencia determinados de acuerdo con un modo de intra predicción que esté disponible de acuerdo con el tamaño de la unidad de codificación actual, para generar así una unidad de codificación de predicción.

La figura 15 es una tabla que muestra los números de modos de intra predicción de acuerdo con los tamaños de unidades de codificación.

El número de modos de intra predicción que se aplicará a una unidad de codificación (una unidad de decodificación en un proceso de decodificación) puede establecerse de manera variable. Por ejemplo, con referencia a la figura 15, si el tamaño de una unidad de codificación en la cual se realiza intra predicción es $N \times N$, los números de modos de intra predicción actualmente realizados en unidades de codificación con un tamaño de 2×2 , 4×4 , 8×8 , 16×16 , 32×32 , 64×64 y 128×128 pueden establecerse respectivamente como 5, 9, 9, 17, 33, 5 y 5 (en el ejemplo 2). Para otro ejemplo, cuando un tamaño de una unidad de codificación que será intra predicha es $N \times N$, los números de modos de intra predicción que se realizarán actualmente en unidades de codificación que tengan tamaños de 2×2 , 4×4 , 8×8 , 16×16 , 32×32 , 64×64 y 128×128 pueden establecerse para ser 3, 17, 34, 34, 34, 5 y 5. Los números de modos de intra predicción que se realizarán realmente se establecen de manera diferente de acuerdo con los tamaños de las unidades de codificación debido a que las sobrecargas para codificar información de modos de predicción difieren de acuerdo con los tamaños de las unidades de codificación. En otras palabras, una unidad de codificación pequeña ocupa una pequeña porción de datos de imagen completos, pero puede tener una gran sobrecarga para poder transmitir información adicional tal como información del modo de predicción de la unidad de codificación. En consecuencia, si una unidad de codificación pequeña se codifica usando un número excesivamente grande de modos de predicción, el número de bits puede incrementarse y de esta manera se puede predecir la eficiencia de compresión. Asimismo, una gran unidad de codificación, por ejemplo, una unidad de codificación igual a o mayor que 64×64 , generalmente corresponde a una región simple de datos de imagen, y de esta manera la codificación de la unidad de codificación grande usando un número excesivamente grande de modos de predicción también puede reducir la eficiencia de compresión.

Así, de acuerdo con un ejemplo, unidades de codificación se clasifican ampliamente en al menos tres tamaños tales como $N_1 \times N_1$ (donde $2 = N_1 = 4$, y N_1 es un número entero), $N_2 \times N_2$ (donde $8 = N_2 = 32$ y N_2 es un número entero), y $N_3 \times N_3$ (donde $64 = N_3$, y N_3 es un número entero). Si el número de modos de intra predicción realizados en las unidades de codificación de $N_1 \times N_1$ es A_1 (donde A_1 es un número entero positivo), el número de modos de intra predicción realizados en las unidades de codificación de $N_2 \times N_2$ es A_2 (donde A_2 es un número entero positivo), y el número de modos de intra predicción realizados en las unidades de codificación de $N_3 \times N_3$ es A_3 (donde A_3 es un número entero positivo), los números de modos de intra predicción realizados de acuerdo con los tamaños de las unidades de codificación pueden establecerse para satisfacer $A_3 = A_1 = A_2$. Es decir, si una imagen actual se divide en unidades de codificación pequeñas, unidades de codificación medias y unidades de codificación grandes, las unidades de codificación medias pueden establecerse para tener el número más grande de modos de predicción y las unidades de codificación pequeñas y las unidades de codificación grandes pueden establecerse para tener un número relativamente pequeño de modos de predicción. Sin embargo, el ejemplo no está limitado a esto y las unidades de codificación pequeñas y grandes también pueden establecerse para tener un gran número de modo de predicción. Los números de modos de predicción de acuerdo con los tamaños de unidades de codificación en la figura 15 se muestran de manera ejemplar y pueden cambiarse.

La figura 16A es una tabla que muestra modos de intra predicción aplicados a una unidad de codificación que tiene un tamaño predeterminado según una realización ejemplar.

Con referencia a las figuras 15 y 16A, por ejemplo, cuando se realiza intra predicción en una unidad de codificación que tiene un tamaño de 4×4 , un modo vertical (modo 0), la unidad de codificación puede tener un modo horizontal (modo 1), un modo de corriente continua (CC) (modo 2), un modo diagonal abajo-izquierda (modo 3), un modo diagonal abajo-derecha (modo 4), un modo vertical-derecha (modo 5), un modo horizontal-abajo (modo 6), un modo vertical-izquierda (modo 7) y un modo horizontal-arriba (modo 8).

La figura 16B ilustra direcciones de los modos de intra predicción mostrados en la figura 16A. En la figura 16B, los números en los extremos de las flechas representan los modos de predicción que corresponden a las direcciones de

predicción indicadas por las flechas. Aquí, el modo 2 es un modo de CC que no tiene directividad y de esta manera no se muestra en la figura 16B.

La figura 16C es un diagrama para describir un procedimiento para realizar intra predicción en una unidad de codificación usando los modos de intra predicción mostrados en la figura 16A.

5 Con referencia a la figura 16C, una unidad de codificación por predicción se genera al realizar un modo de intra predicción disponible determinado de acuerdo con el tamaño de una unidad de codificación actual usando píxeles vecinos A a M de la unidad de codificación actual. Por ejemplo, una operación de realizar codificación por predicción en una unidad de codificación actual que tiene un tamaño de 4 x 4 de acuerdo con el modo 0, es decir, se describirá un modo vertical, mostrado en la figura 16A. Inicialmente, valores de los píxeles vecinos A a D en un lado superior de la unidad de codificación actual se predicen como valores de píxel de la unidad de codificación actual. Es decir, el valor del píxel vecino A se predice como un valor de cuatro píxeles en una primera columna de la unidad de codificación actual, el valor del píxel vecino B se predice como un valor de cuatro píxeles en una segunda columna de la unidad de codificación actual, el valor del píxel vecino C se predice como un valor de cuatro píxeles en una tercera columna de la unidad de codificación actual, y el valor del píxel vecino D se predice como un valor de cuatro píxeles en una cuarta columna de la unidad de codificación actual. Después de eso, los valores de píxel de la unidad de codificación actual predicha usando los píxeles vecinos A a D son restados de los valores de píxel de la unidad de codificación actual original, para de esta manera calcular un valor de error y luego se codifica el valor de error. Mientras tanto, cuando se aplican varios modos de intra predicción, píxeles vecinos usados como píxeles de referencia pueden ser píxeles vecinos originales o píxeles vecinos filtrados como se describe anteriormente.

20 La figura 17 ilustra modos de intra predicción aplicados a una unidad de codificación que tiene un tamaño predeterminado, de acuerdo con otra realización ejemplar.

Con referencia a las figuras 15 y 17, por ejemplo, cuando se realiza intra predicción en una unidad de codificación que tiene un tamaño de 2 x 2, la unidad de codificación puede tener en total cinco modos tales como un modo vertical, un modo horizontal, un modo de CC, un modo de plano y un modo abajo-derecha diagonal.

25 Mientras tanto, si una unidad de codificación, que tiene un tamaño de 32 x 32 tiene 33 modos de intra predicción como se muestra en la figura 15, tienen que establecerse direcciones de los 33 modos de intra predicción. De acuerdo con un ejemplo, para establecer un modo de intra predicción que tenga varias direcciones además de los modos de intra predicción ilustrados en las figuras 16A a 16C, y 17, las direcciones de predicción para seleccionar píxeles vecinos usados como píxeles de referencia de píxeles de la unidad de codificación se establecen usando parámetros (dx, dy). Por ejemplo, si cada uno de los 33 modos de predicción se define como modo N (donde N es un número entero de 0 a 32), el modo 0 puede establecerse como un modo vertical, el modo 1 puede establecerse como un modo horizontal, el modo 2 puede establecerse como un modo de CC, el modo 3 puede establecerse como un modo de plano, y cada uno del modo 4 al modo 31 pueden definirse como un modo de predicción que tenga una directividad de $\tan^{-1}(dy/dx)$ usando (dx, dy) representados como uno de (1, -1), (1, 1), (1, 2), (2, 1), (1, -2), (2, 1), (1, -2), (2, -1), (2, -11), (5, -7), (10, -7), (11, 3), (4, 3), (1, 11), (1, -1), (12, -3), (1, -11), (1, -7), (3, -10), (5, -6), (7, -6), (7, -4), (11, 1), (6, 1), (8, 3), (5, 3), (5, 7), (2, 7), (5, -7) y (4, -3) mostrados en la tabla 3.

Tabla 3

modo #	dx	dy	modo #	dx	dy
modo 4	1	-1	modo 18	1	-11
modo 5	1	1	modo 19	1	-7
modo 6	1	2	modo 20	3	-10
modo 7	2	1	modo 21	5	-6
modo 8	1	-2	modo 22	7	-6
modo 9	2	-1	modo 23	7	-4
modo 10	2	-11	modo 24	11	1
modo 11	5	-7	modo 25	6	1
modo 12	10	-7	modo 26	8	3
modo 13	11	3	modo 27	5	3
modo 14	4	3	modo 28	5	7
modo 15	1	11	modo 29	2	7
modo 16	1	-1	modo 30	5	-7
modo 17	12	-3	modo 31	4	-3
El modo 0 es un modo vertical, el modo 1 es un modo horizontal, el modo 2 es un modo de CC, el modo 3 es un modo plano, y el modo 32 es un modo bilineal.					

40 Las figuras 18A a 18C son diagramas para describir modos de intra predicción que tienen varias directividades.

Como se describió anteriormente en relación con la Tabla 3, modos de intra predicción de acuerdo con un ejemplo pueden tener varias directividades de $\tan^{-1}(dy/dx)$ usando una pluralidad de parámetros (dx, dy).

Con referencia a la figura 18A, píxeles vecinos A y B en una línea de extensión 160 que tiene un ángulo de $\tan^{-1} (dy/dx)$ de acuerdo con valores (dx, dy) en la Tabla 3 con respecto a un píxel actual P en una unidad de codificación actual que será predicha, se pueden usar como un predictor del píxel actual P. En este caso, los píxeles vecinos usados como un predictor pueden codificarse previamente y los píxeles restablecidos de una unidad de codificación previa y lados superior e izquierdo de una unidad de codificación actual. Asimismo, si la línea 160 de extensión pasa entre dos píxeles vecinos colocados en ubicaciones enteras, uno de los píxeles vecinos más cercano a la línea 160 de extensión que el otro puede usarse como el predictor del píxel actual P.

Asimismo, si la línea 160 de extensión pasa entre dos píxeles vecinos colocados en ubicaciones enteras, uno de los píxeles vecinos más cercano al píxel actual P que el otro puede usarse como el predictor del píxel actual P, o un valor promedio ponderado calculado teniendo en cuenta las distancias desde los píxeles vecinos hasta un cruce de la línea 160 de extensión y una línea entre los píxeles vecinos se puede usar como el predictor del píxel actual P.

Las figuras 18B y 18C son diagramas para describir un proceso de generar un predictor cuando la línea 160 de extensión pasa entre dos píxeles vecinos colocados en ubicaciones enteras.

Con referencia a la figura 18B, si la línea 160 de extensión que tiene un ángulo de $\tan^{-1} (dy/dx)$ que se determinará de acuerdo con un valor (dx, dy) de cada modo pasa entre píxeles vecinos A 151 y B 152 colocados en ubicaciones enteras, como se describió anteriormente, uno de los píxeles vecinos A 151 y B 152 más cercano a la línea 160 de extensión o un valor promedio ponderado calculado teniendo en cuenta las distancias desde los píxeles vecinos A 151 y B 152 hasta un cruce de la línea 160 de extensión y una línea entre los píxeles vecinos A 151 y B 152 se puede usar como un predictor del píxel actual P. Por ejemplo, si la distancia entre el cruce y el píxel vecino A 151 es f y la distancia entre el cruce y el píxel vecino B 152 es g, el predictor del píxel actual P puede obtenerse como $(A \cdot g + B \cdot f) / (f + g)$. Aquí, f y g pueden ser distancias reguladas como enteros. En una implementación de software y hardware real, el predictor del píxel actual P puede obtenerse al realizar una operación de desplazamiento tal como $(g \cdot A + f \cdot B + 2) \gg 2$. Como se ilustra en la figura 18B, si la línea 160 de extensión pasa una ubicación de 1/4 entre los píxeles vecinos A 151 y B 152, que está más cerca del píxel vecino A 151, el predictor del píxel actual P puede obtenerse como $(3 \cdot A + B) / 4$. Este valor puede obtenerse al realizar una operación de desplazamiento tal como $(3 \cdot A + B + 2) \gg 2$ teniendo en cuenta el redondeo.

Mientras tanto, si la línea 160 de extensión pasa entre los píxeles vecinos A 151 y B 152, la sección entre los píxeles vecinos A 151 y B 152 puede dividirse en un número predeterminado de secciones, y un valor promedio ponderado puede calcularse teniendo en cuenta las distancias entre el cruce y los píxeles vecinos A 151 y B 152 en cada sección pueden usarse como el predictor. Por ejemplo, con referencia a la figura 18C, la sección entre los píxeles vecinos A 151 y B 152 se divide en cinco secciones P1 a P5, un valor promedio ponderado representativo calculado teniendo en cuenta las distancias entre el cruce y los píxeles vecinos A 151 y B 152 en cada sección puede determinarse y se puede usar como el predictor del píxel actual P. En más detalle, si la línea 160 de extensión pasa la sección P1, un valor del píxel vecino A 151 puede determinarse como el predictor del píxel actual P. Si la línea de extensión pasa a la sección P2, un valor promedio ponderado calculado teniendo en cuenta las distancias entre el centro de la sección P2 y los píxeles vecinos A 151 y B 152, es decir, $(3 \cdot A + 1 \cdot B + 2) \gg 2$, se puede determinar como el predictor del píxel actual P. Si la línea 160 de extensión pasa a la sección P3, un valor promedio ponderado calculado teniendo en cuenta las distancias entre el centro de la sección P3 y los píxeles vecinos A 151 y B 152, es decir, $(2 \cdot A + 2 \cdot B + 2) \gg 2$, puede determinarse como el predictor del píxel actual P. Si la línea 160 de extensión pasa a la sección P4, un valor promedio ponderado calculado teniendo en cuenta las distancias entre el centro de la sección P4 y los píxeles vecinos A 151 y B 152, es decir, $(1 \cdot A + 3 \cdot B + 2) \gg 2$, puede determinarse como el predictor del píxel actual P. Si la línea 160 de extensión pasa a la sección P5, un valor del píxel vecino B 152 puede determinarse como el predictor del píxel actual P.

Asimismo, como se ilustra en la figura 18A, si la línea 160 de extensión encuentra dos píxeles vecinos tales como el píxel vecino A en el lado superior y el píxel vecino B en el lado izquierdo, un valor promedio de los píxeles vecinos A y B se puede usar como el predictor del píxel actual P. Como alternativa, el píxel vecino A puede usarse si un valor de $dx \cdot dy$ es un número positivo, y el píxel vecino B puede usarse si el valor de $dx \cdot dy$ es un número negativo. También, píxeles vecinos usados como píxeles de referencia pueden ser píxeles vecinos originales o píxeles vecinos filtrados como se describió anteriormente.

Los modos de intra predicción que tienen varias directividades en la Tabla 3 pueden establecerse previamente en un lado de codificador y un lado de decodificador, y de esta manera cada unidad de codificación puede transmitir índices que correspondan solo a los modos de intra predicción establecidos.

Como una codificación por predicción se realiza de acuerdo con los modos de intra predicción establecidos de manera variable de acuerdo con el tamaño de una unidad de codificación, la eficiencia de compresión de una imagen puede mejorarse de acuerdo con características de imagen. Asimismo, de acuerdo con un ejemplo de realización, ya que píxeles vecinos y píxeles vecinos filtrados se usan selectivamente para realizar intra predicción, la predicción puede realizarse de manera más variable y de esta manera puede mejorarse la eficiencia de compresión de una imagen.

Según otro ejemplo, en lugar de usar píxeles vecinos determinados previamente de acuerdo con el tamaño de una unidad de codificación actual y el tipo de modo de intra predicción que se realizará actualmente, la unidad de realización de intra predicción 1230 puede realizar predicción en una unidad de codificación actual de acuerdo con

modo de intra predicción disponible al usar por separado píxeles vecinos originales, primeros píxeles vecinos filtrados y segundos píxeles vecinos filtrados como píxeles de referencia, y la unidad de determinación de píxeles de referencia 1220 puede seleccionar píxeles vecinos que tengan costes mínimos como píxeles de referencia que se usarán finalmente para realizar intra predicción en la unidad de codificación actual.

5 La figura 21 es un diagrama de flujo que ilustra un procedimiento de codificación de vídeo.

Con referencia a la figura 21, en la operación 1910, los píxeles vecinos de una unidad de codificación actual que será codificada son filtrados para generar píxeles vecinos filtrados. Como se describió anteriormente, la unidad 1210 de filtración de píxeles vecinos filtra píxeles vecinos en lados superior e izquierdo de la unidad de codificación actual al menos una vez para generar así el píxel vecino filtrado. Aquí, una unidad de codificación puede obtenerse al dividir una imagen actual en base a una unidad de codificación máxima que sea una unidad de codificación que tenga un tamaño máximo, y una profundidad codificada que sea información de división jerárquica de la unidad de codificación máxima.

En la operación 1920, los píxeles vecinos filtrados o los píxeles vecinos originales se seleccionan como píxeles de referencia que se usarán para realizar intra predicción en la unidad de codificación actual. Como se describió anteriormente, la unidad 1220 de determinación de píxeles de referencia puede seleccionar los píxeles de referencia de acuerdo con el tamaño de la unidad de codificación actual y el tipo de un modo de intra predicción que se realizará actualmente, como se muestra en la Tabla 2. De acuerdo con otro ejemplo de realización, la unidad 1220 de determinación de píxeles de referencia puede comparar costes resultantes de la codificación por intra predicción realizar al usar por separado el píxel vecino original y los píxeles vecinos filtrados, y puede determinar el píxel vecino que será filtrado, y puede determinar el píxel vecino que será finalmente usado para realizar intra predicción. Asimismo, la unidad 1220 de determinación de píxeles de referencia puede señalar indicar qué píxeles vecinos se seleccionan entre los píxeles vecinos originales y los píxeles vecinos filtrados para realizar intra predicción en la unidad de codificación actual. En otras palabras, la información de modo de intra predicción puede comprender información de índices de referencia que indique qué píxeles vecinos se seleccionan entre los píxeles vecinos originales y los píxeles vecinos filtrados para realizar intra predicción en la unidad de codificación actual. Cuando los píxeles de referencia que se usarán son preestablecidos en un extremo de codificación y un extremo de decodificación como se muestra en la Tabla 2, la información de índices de referencia no tiene que ser transmitida.

En la operación 1930, se realiza la intra predicción en la unidad de codificación actual usando, los píxeles de referencia seleccionados. Como se describió anteriormente, la unidad 1230 de realización de intra predicción genera una unidad de codificación por predicción al realizar intra predicción en la unidad de codificación actual al aplicar un modo de intra predicción que está disponible en la unidad de codificación actual usando los píxeles de referencia seleccionados, y envía la unidad de codificación por predicción.

La figura 22 es un diagrama de flujo que ilustra un procedimiento de decodificación de vídeo según una realización ejemplar.

35 Con referencia a la figura 22, en la operación 2010, píxeles vecinos de una unidad de decodificación actual que será decodificada se filtran para generar píxeles vecinos filtrados.

En la operación 2020, información acerca de un modo de inter predicción aplicado a la unidad de decodificación actual se extrae de una secuencia de bits. La información acerca del modo de intra predicción puede incluir información acerca de un modo de intra predicción aplicado a la unidad de decodificación actual e información acerca de un índice de referencia que representa si píxeles vecinos originales o píxeles vecinos filtrados se usan como píxeles de referencia. Si, como se muestra en la Tabla 1, el mismo tipo de píxeles de referencia que se usarán de acuerdo con el modo de inter predicción y el tamaño de la unidad de decodificación actual se establece en un lado de codificador y un lado de decodificador, la información acerca del índice de referencia no se transmite necesariamente.

En la operación 2030, los píxeles vecinos filtrados o los píxeles vecinos originales se seleccionan como píxeles de referencia que se usarán para realizar intra predicción en la unidad de decodificación actual. Como se describió anteriormente, si la información acerca del índice de referencia se incluye adicionalmente en la secuencia de bits, los píxeles de referencia se seleccionan de acuerdo con la información extraída acerca del índice de referencia. Si, como se muestra en la Tabla 2, píxeles de referencia pueden determinarse en base al tamaño y al modo de intra predicción de la unidad de decodificación actual, los píxeles vecinos originales o los píxeles vecinos filtrados que se usarán como los píxeles de referencia pueden determinarse en base al tamaño y al modo de intra predicción de la unidad de decodificación actual.

En la operación 2040, se realiza intra predicción en la unidad de decodificación actual usando la información extraída acerca del modo de intra predicción y los píxeles de referencia seleccionados.

La presente invención también puede incorporarse como código legible a ordenador en un medio de registro legible a ordenador. El medio de registro legible a ordenador es cualquier dispositivo de almacenamiento de datos que puede almacenar datos que se pueden leer después por un sistema informático. Los ejemplos del medio de registro legible a ordenador incluyen memoria de solo lectura (ROM), memoria de acceso aleatorio (RAM), CD-ROM, cintas magnéticas, disquetes y dispositivos ópticos de almacenamiento de datos. El medio de registro legible a ordenador

también puede distribuirse sobre sistemas informáticos acoplados a la red por lo que el código legible a ordenador se almacena y ejecuta de manera distribuida.

5 Aunque la presente invención ha sido mostrada y descrita particularmente con referencia a ejemplos y a realizaciones de la misma, se entenderá por alguien de capacidad ordinaria en la técnica que se pueden hacer en la misma varios cambios en forma y detalles sin apartarse del alcance de la invención como se define mediante las siguientes reivindicaciones. Los ejemplos y las realizaciones deben considerarse en un sentido descriptivo únicamente y no para efectos de limitación. Por lo tanto, el alcance de la invención se define no por la descripción detallada de la invención, sino por las siguientes reivindicaciones, y todas las diferencias dentro del alcance se considerarán como estando incluidas en la presente invención.

10

REIVINDICACIONES

1. Un procedimiento de decodificación de una imagen, comprendiendo el procedimiento:

5 extraer información que indica un modo de predicción intra aplicado a un bloque actual a decodificar, desde una secuencia de bits;

determinar uno de los píxeles vecinos adyacentes al bloque actual o píxeles vecinos filtrados, filtrados de los píxeles vecinos como píxeles de referencia, en base al modo de predicción intra del bloque actual;

realizar predicción intra en el bloque actual usando la información extraída y los píxeles de referencia determinados, en el que la imagen se divide en una pluralidad de unidades de codificación máximas de forma cuadrada, de acuerdo con la información sobre un tamaño máximo de una unidad de codificación extraída de la secuencia de bits,

10 en el que una unidad de codificación máxima entre la pluralidad de unidades de codificación máxima de forma cuadrada se divide jerárquicamente en una o más unidades de profundidades codificadas de acuerdo con una información de división extraída a partir de la secuencia de bits, las una o más unidades de codificación de profundidades codificadas tienen forma cuadrada,

15 en el que una unidad de codificación de una profundidad codificada actual, $k+1$, donde k es un número entero, es una de unidades de datos de forma cuadrada divididas desde una unidad de codificación de una profundidad codificada menor k ,

en el que, cuando la información de división indica que la unidad de codificación de la profundidad codificada actual, $k+1$ se divide, la unidad de codificación de la profundidad codificada actual, $k+1$, se divide en unidades de codificación de profundidad codificada más profunda, $k+2$,

20 en el que cuando la información de división indica que la unidad de codificación de la profundidad codificada actual, $k+1$, ya no se divide, una o más unidades de predicción para decodificación de predicción se obtienen de la unidad de codificación de la profundidad codificada actual, $k+1$, y una o más unidades de transformación para transformación inversa se obtienen de la unidad de codificación de la profundidad codificada actual, $k+1$,

25 en el que la una o más unidades de predicción se obtienen dividiendo de manera igual uno de una altura o una anchura de la unidad de codificación de la profundidad codificada actual, $k+1$, por 2, o dividiendo de manera igual tanto la altura como la anchura de la unidad de codificación de la profundidad codificada actual, $k+1$, por 2, o determinando la unidad de codificación de la profundidad codificada actual, $k+1$, como la una unidad de predicción según un tipo de división extraído de la secuencia de bits,

30 en el que la una o más unidades de transformación son cuadradas, las unidades de transformación divididas de la unidad de codificación de la profundidad codificada actual tienen el mismo tamaño y la una o más unidades de transformación se obtienen de acuerdo con información sobre el tamaño de la unidad de transformación extraída de la secuencia de bits,

35 en el que el procedimiento de decodificación admite que un tamaño y una forma de la unidad de predicción sean los mismos que un tamaño y una forma de la unidad de transformación o que el tamaño y la forma de la unidad de predicción sean diferentes del tamaño y de la forma de la unidad de transformación, y el bloque actual es uno de la una o más unidades de predicción obtenidas a partir de la unidad de codificación de la profundidad codificada actual, $k+1$.

2. Un aparato de codificación de una imagen, comprendiendo el aparato:

una unidad (1220) de determinación de píxeles de referencia que determina uno de los píxeles vecinos adyacentes a un bloque actual o píxeles vecinos filtrados, filtrados de los píxeles vecinos como píxeles de referencia, en función de un modo de predicción intra del bloque actual;

45 una unidad (1230) de realización de intra predicción que realiza intra predicción en el bloque actual usando los píxeles de referencia determinados; y

una unidad (130) de salida que genera una secuencia de bits incluyendo una información sobre el modo de predicción intra aplicado al bloque actual, información de división que indica si se divide una unidad de codificación de una profundidad codificada, información sobre un tipo de división, información sobre un tamaño de una unidad de transformación e información sobre un tamaño máximo de una unidad de codificación,

50 en el que la imagen se divide en una pluralidad de unidades de codificación máximas con forma cuadrada, y se genera la información sobre el tamaño máximo de la unidad de codificación,

en el que una unidad de codificación máxima entre la pluralidad de unidades de codificación máximas con forma cuadrada se divide jerárquicamente en una o más unidades de codificación de profundidades codificadas, y se genera la información de división para las una o más unidades de codificación de profundidad codificada, las una o más unidades de codificación de profundidades codificadas tienen una forma cuadrada,

55 en el que una unidad de codificación de una profundidad codificada actual, $k+1$, donde k es un número entero, es una de unidades de datos de forma cuadrada divididas de una unidad de codificación de una profundidad codificada menor k ,

en el que, cuando la información de división indica que la unidad de codificación de la profundidad codificada actual, $k+1$, se divide, la unidad de codificación de la profundidad codificada actual, $k+1$, se divide en unidades de codificación de una profundidad codificada más profunda, $k+2$,

60 en el que, cuando la información de división indica que la unidad de codificación de la profundidad codificada actual, $k+1$, ya no se divide, se obtienen una o más unidades de predicción para decodificación de predicción de

la unidad de codificación de la profundidad codificada actual, $k+1$, y se obtienen una o más unidades de transformación para transformación inversa desde la unidad de codificación de la profundidad codificada actual, $k+1$,

5 en el que las una o más unidades de predicción se obtienen dividiendo igualmente una de una altura o un ancho de la unidad de codificación de la profundidad codificada actual, $k+1$, por 2, o dividiendo igualmente tanto la altura como el ancho de la unidad de codificación de la profundidad de codificación actual, $k+1$, por 2, o determinando la unidad de codificación de la profundidad codificada actual, $k+1$ como la una unidad de predicción, y se genera la información sobre el tipo de división,

10 en el que las una o más unidades de transformación son cuadradas, las unidades de transformación divididas desde la unidad de codificación de la profundidad codificada actual tienen el mismo tamaño y las una o más unidades de transformación se obtienen según la información sobre el tamaño de la unidad de transformación,

en el que el procedimiento de codificación soporta que un tamaño y una forma de la unidad de predicción son iguales que un tamaño y una forma de la unidad de transformación o el tamaño y la forma de la unidad de predicción son diferentes del tamaño y la forma de la unidad de transformación, y

15 el bloque actual es uno de la una o más unidades de predicción obtenidas a partir de la unidad de codificación de la profundidad codificada actual, $k+1$.

3. Un medio legible a ordenador para almacenar datos asociados con un vídeo, comprendiendo una secuencia de bits almacenados en el medio legible por ordenador, incluyendo la secuencia de bits información para su uso en un procedimiento de decodificación de una imagen según la reivindicación 1, en el que la información comprende

20 una información sobre un modo de predicción intra aplicado a un bloque actual, información de división que indica si se divide una unidad de codificación de una profundidad codificada, información sobre un tipo de división, información sobre un tamaño de una unidad de transformación e información sobre un tamaño máximo de una unidad de codificación,

25 en el que un valor de predicción del bloque actual se obtiene al determinar uno de píxeles vecinos adyacentes a un bloque actual o píxeles vecinos filtrados, filtrados de los píxeles vecinos como píxeles de referencia, en función de un modo de predicción intra del bloque actual, y realizar intra predicción en el bloque actual usando los píxeles de referencia determinados,

en el que la imagen se divide en una pluralidad de unidades de codificación máximas de forma cuadrada, según la información sobre un tamaño máximo de una unidad de codificación,

30 en el que una unidad de codificación máxima entre la pluralidad de unidades de codificación máximas de forma cuadrada se divide jerárquicamente en una o más unidades de codificación de profundidades codificadas según la información de división, las una o más unidades de codificación de profundidades codificadas tienen una forma cuadrada,

35 en el que una unidad de codificación de una profundidad codificada actual, $k+1$, donde k es un número entero, es una de unidades de datos de forma cuadrada divididas de una unidad de codificación de una profundidad codificada menor k ,

en el que, cuando la información de división indica que se divide la unidad de codificación de la profundidad codificada actual, $k+1$, la unidad de codificación de la profundidad codificada actual, $k+1$, se divide en unidades de codificación de una profundidad codificada más profunda, $k+2$,

40 en el que, cuando la información de división indica que la unidad de codificación de la profundidad codificada actual, $k+1$, ya no se divide, una o más unidades de predicción para decodificación de predicción se obtienen de la unidad de codificación de la profundidad codificada actual, $k+1$, y una o más unidades de transformación para transformación inversa se obtienen de la unidad de codificación de la profundidad codificada actual, $k+1$,

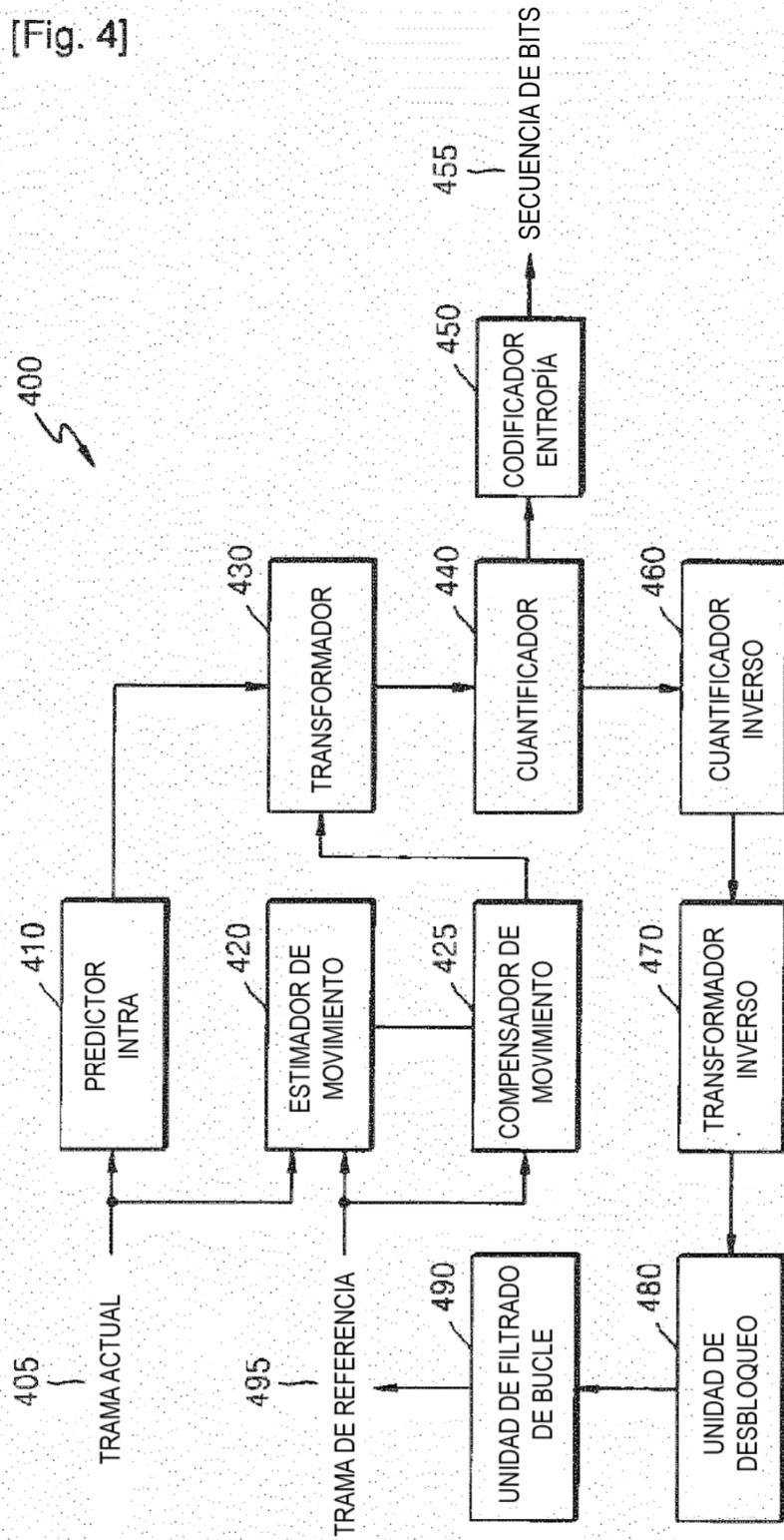
45 en el que las una o más unidades de predicción se obtienen dividiendo igualmente una de una altura o un ancho de la unidad de codificación de la profundidad codificada actual, $k+1$, por 2, o dividiendo igualmente tanto la altura como el ancho de la unidad de codificación de la profundidad codificada actual, $k+1$, por 2, o determinando la unidad de codificación de la profundidad codificada actual, $k+1$ como la una unidad de predicción según la información sobre el tipo de división,

50 en el que las una o más unidades de transformación son cuadradas, las unidades de transformación divididas de la unidad de codificación de la profundidad codificada actual tienen igual tamaño y las una o más unidades de transformación se obtienen según la información sobre el tamaño de la unidad de transformación,

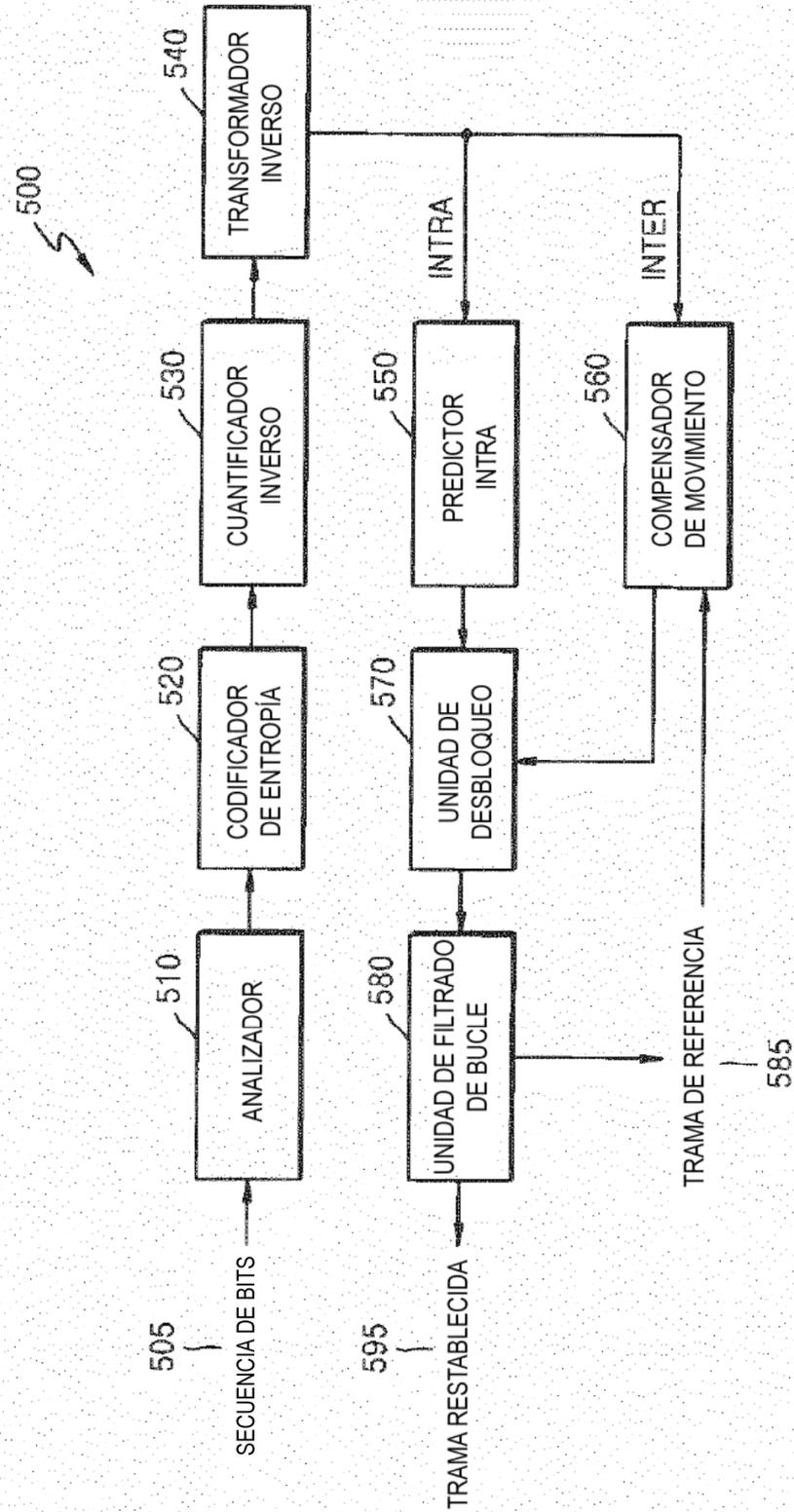
en el que el procedimiento de codificación soporta que un tamaño y una forma de la unidad de predicción son iguales que un tamaño y una forma de la unidad de transformación o el tamaño y la forma de la unidad de predicción son diferentes del tamaño y la forma de la unidad de transformación, y

55 el bloque actual es uno de la una o más unidades de predicción obtenidas a partir de la unidad de codificación de la profundidad codificada actual, $k+1$.

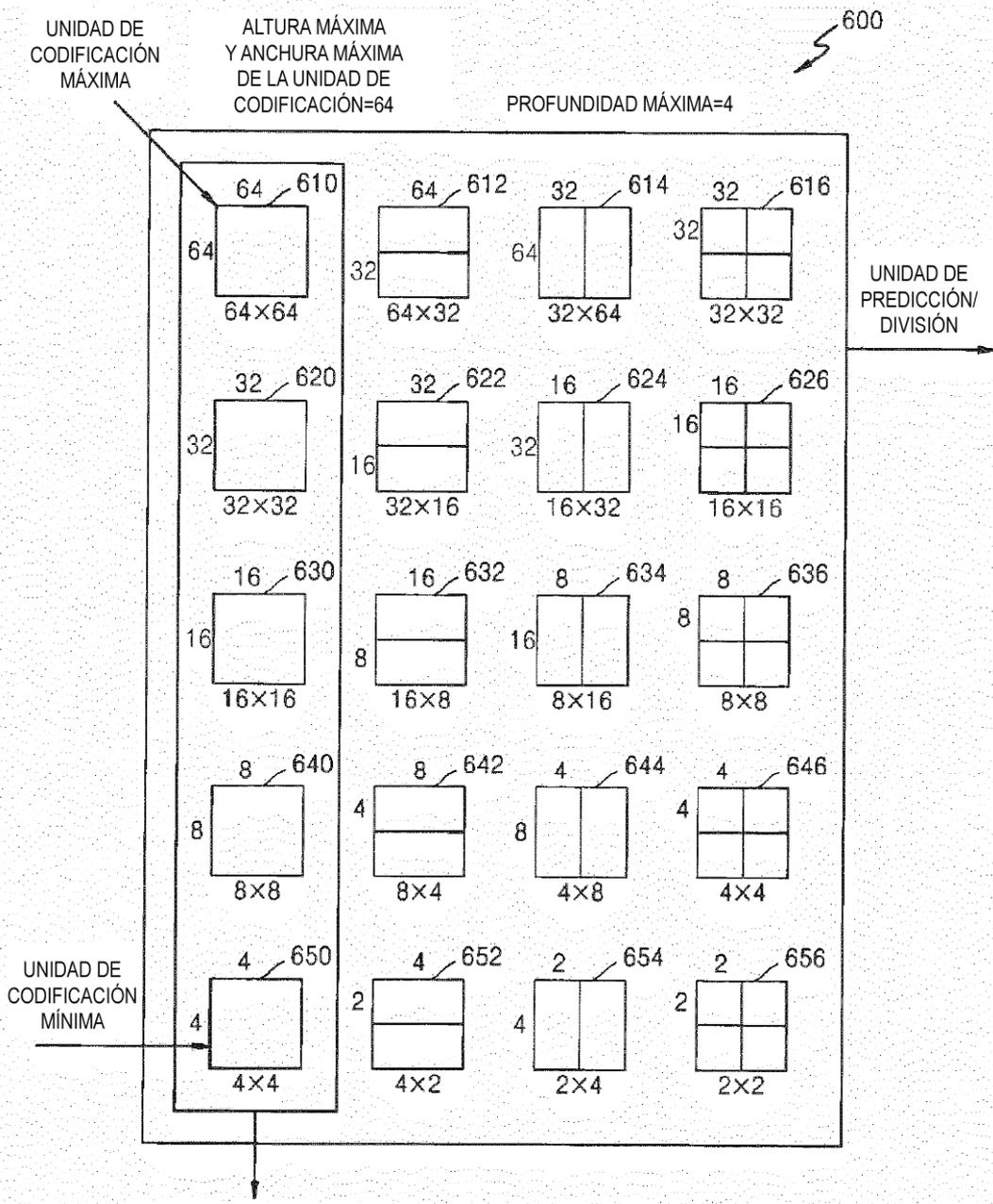
[Fig. 4]



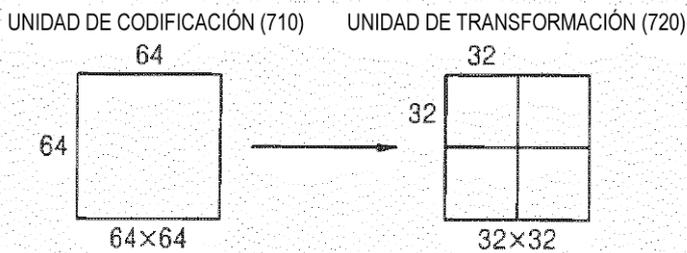
[Fig. 5]



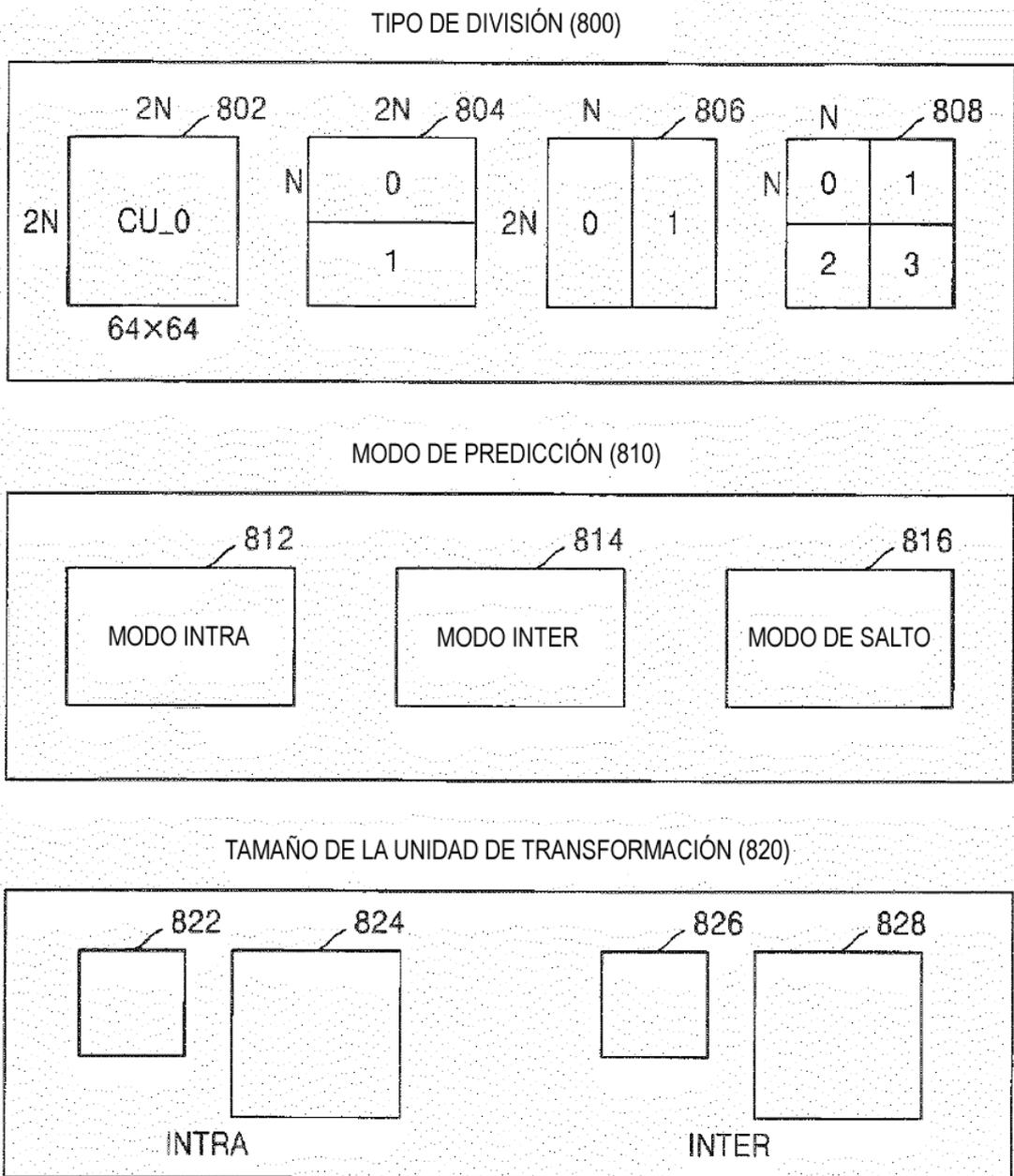
[Fig. 6]



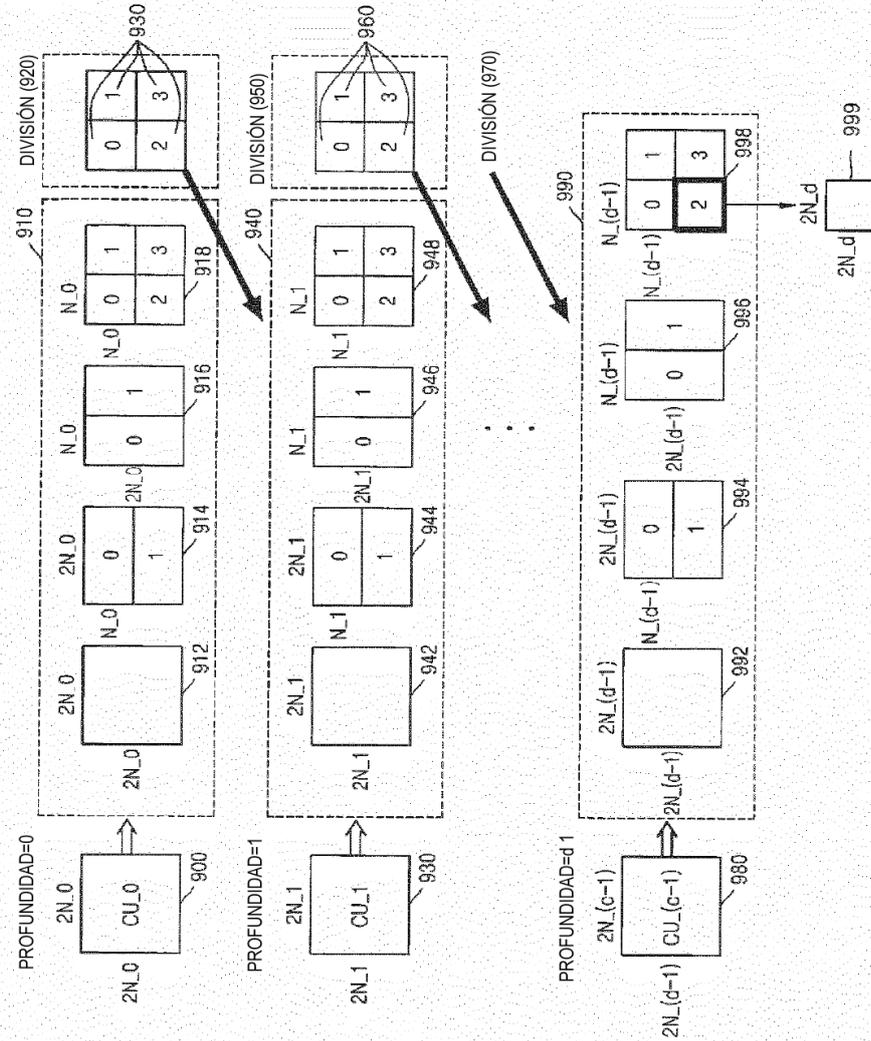
[Fig. 7]



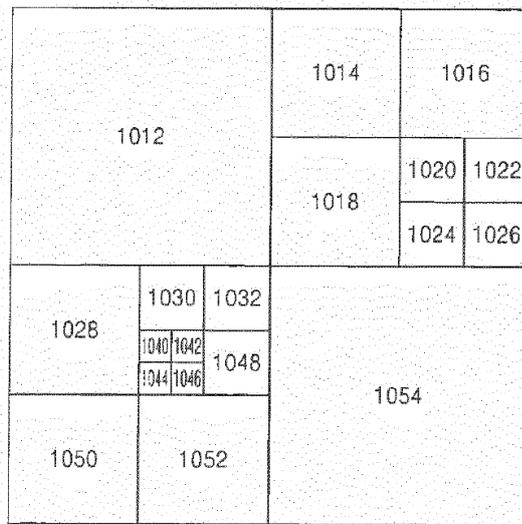
[Fig. 8]



[Fig. 9]

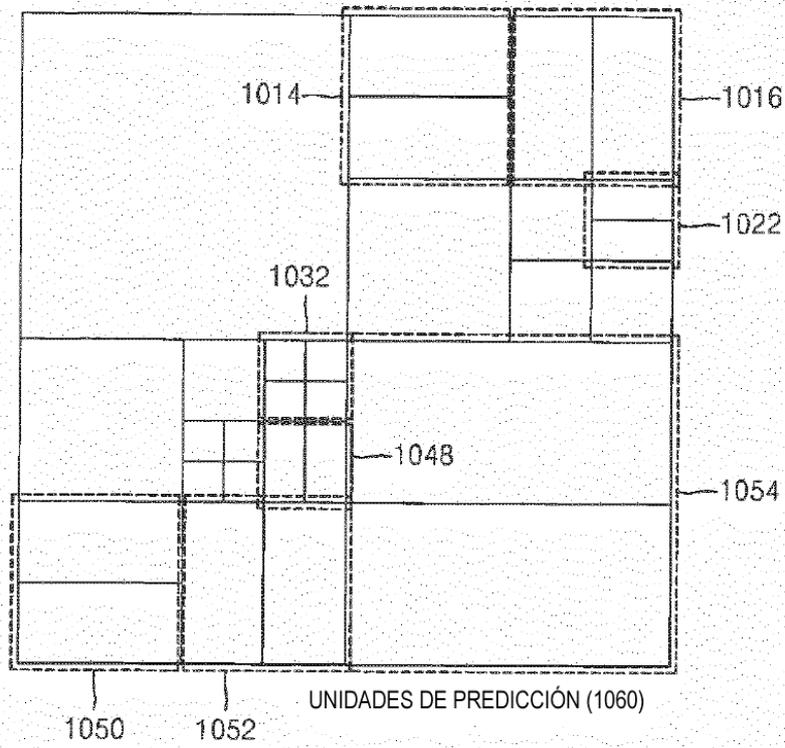


[Fig. 10]

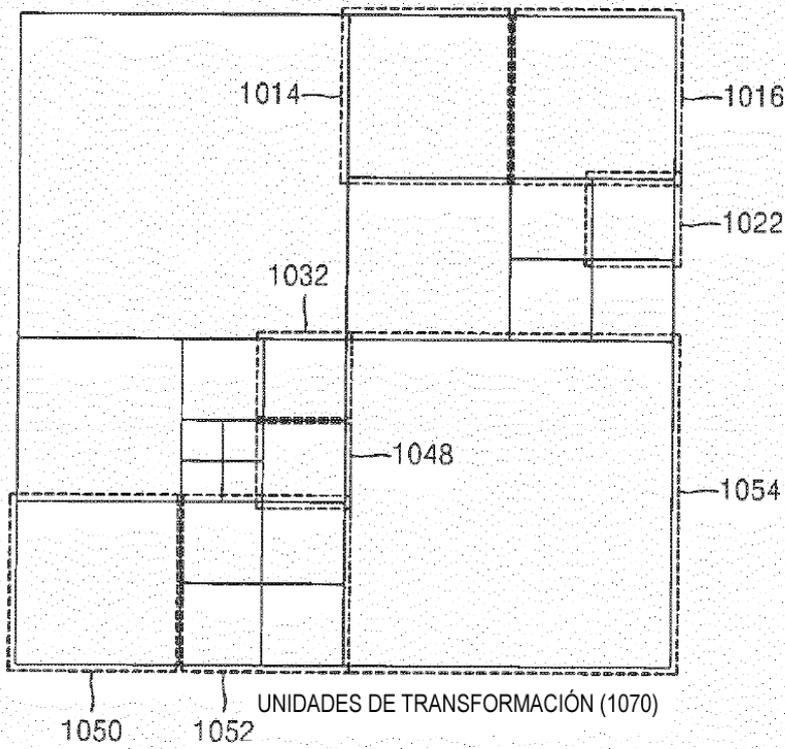


UNIDADES DE CODIFICACIÓN (1010)

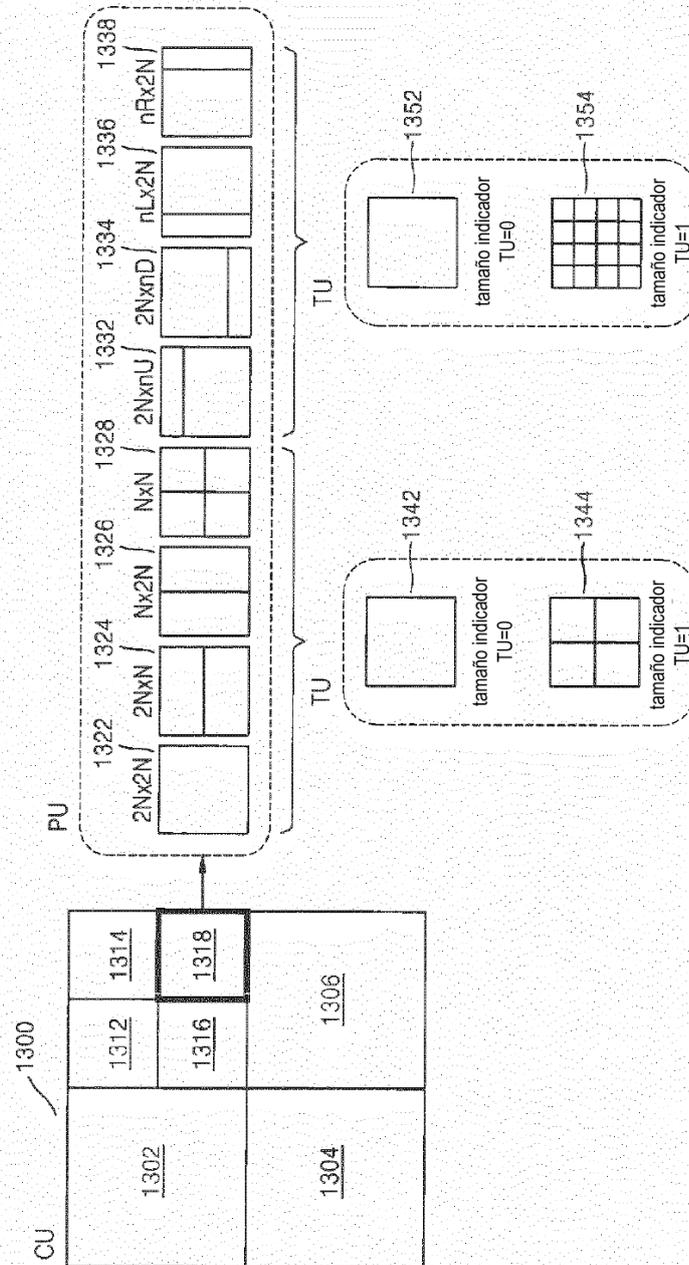
[Fig. 11]



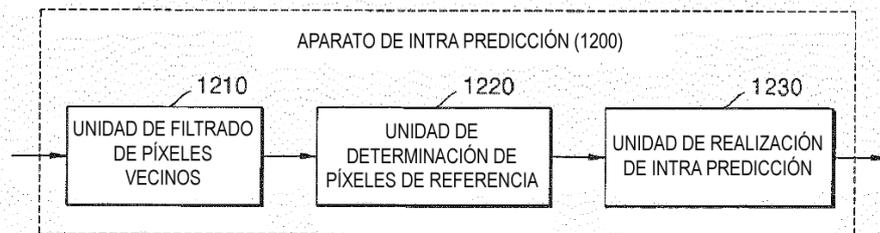
[Fig. 12]



[Fig. 13]



[Fig. 14]



[Fig. 15]

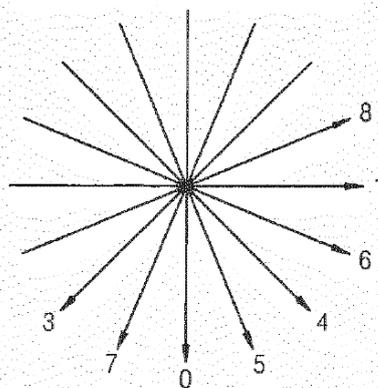
TAMAÑO DE LA UNIDAD DE CODIFICACIÓN	NÚMERO DE MODOS DE PREDICCIÓN		
	EJEMPLO 1	EJEMPLO 2	EJEMPLO 3
2	-	5	5
4	9	9	9
8	9	9	9
16	33	17	11
32	33	33	33
64	5	5	9
128	5	5	5

[Fig. 16a]

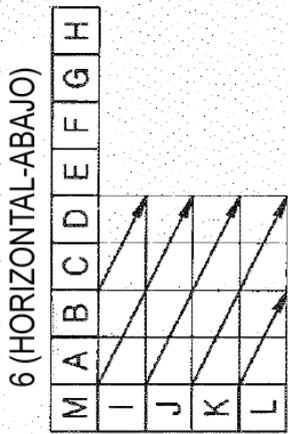
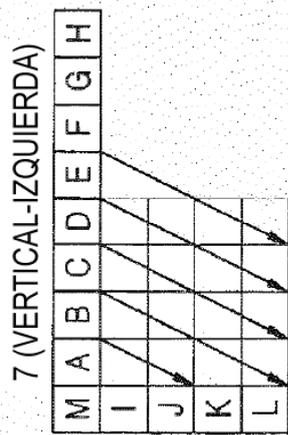
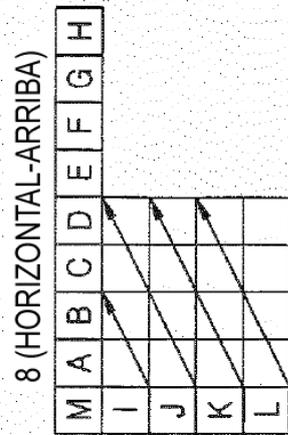
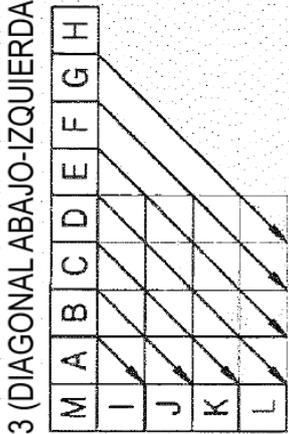
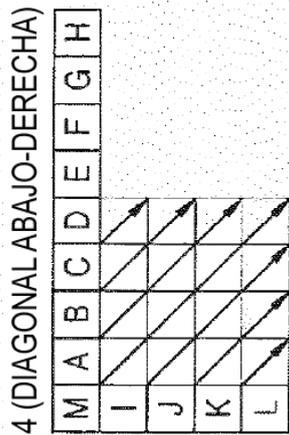
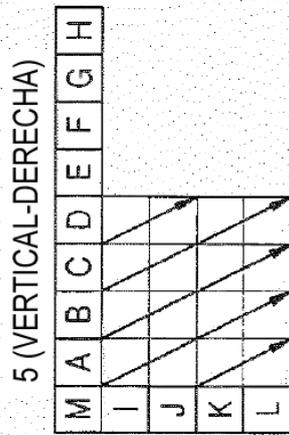
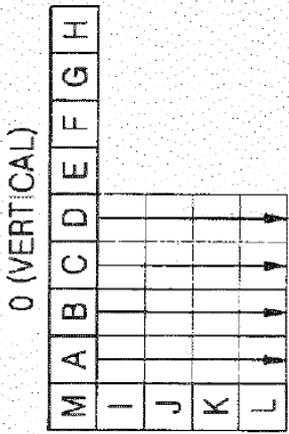
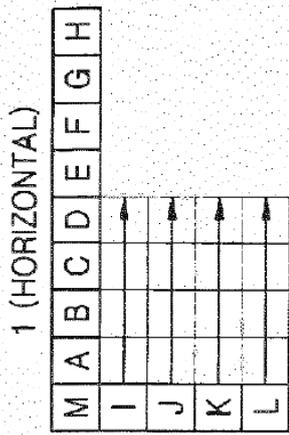
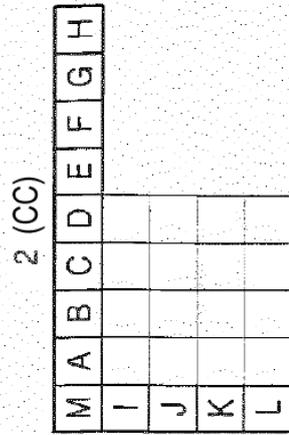
MODO DE PREDICCIÓN	NOMBRE
0	VERTICAL
1	HORIZONTAL
2	CC
3	ABAJO_IZQUIERDO
4	ABAJO_DERECHO
5	VERTICAL_DERECHO
6	HORIZONTAL_ABAJO
7	VERTICAL_IZQUIERDO
8	HORIZONTAL_ARRIBA

[Fig. 16b]

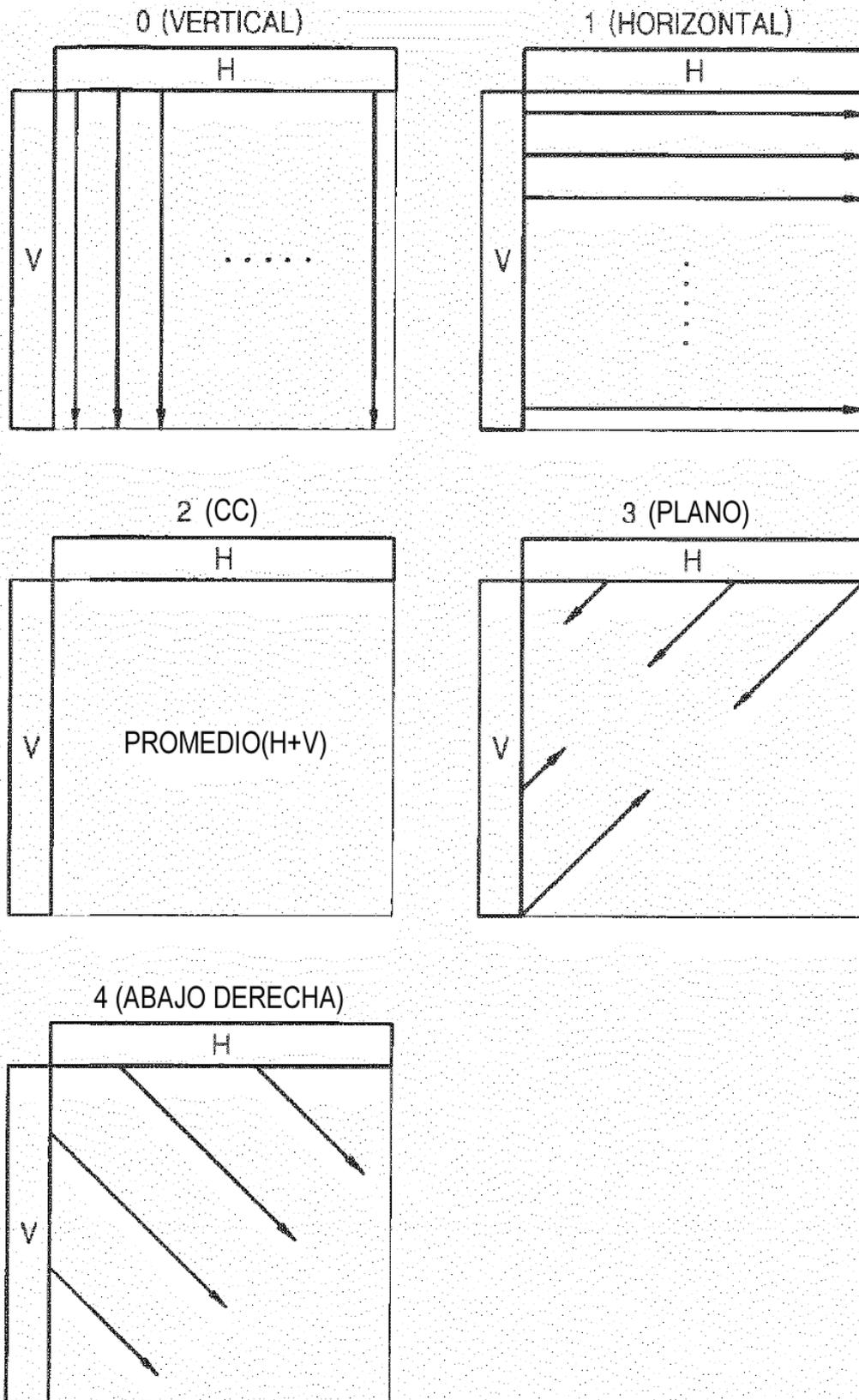
DIRECCIÓN MODO DE PREDICCIÓN



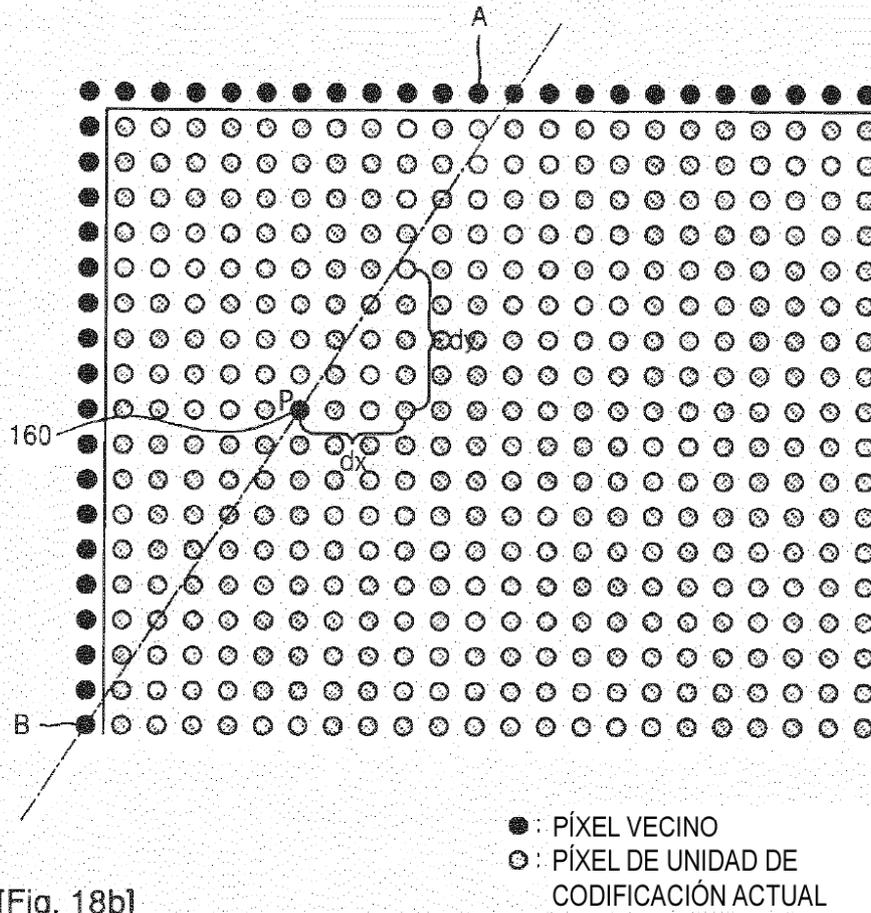
[Fig. 16c]



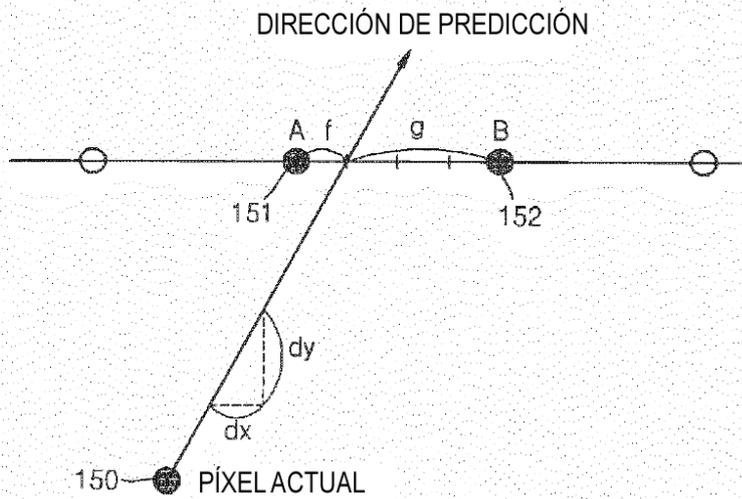
[Fig. 17]



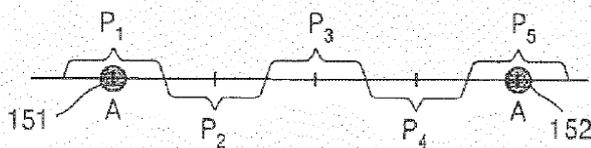
[Fig. 18a]



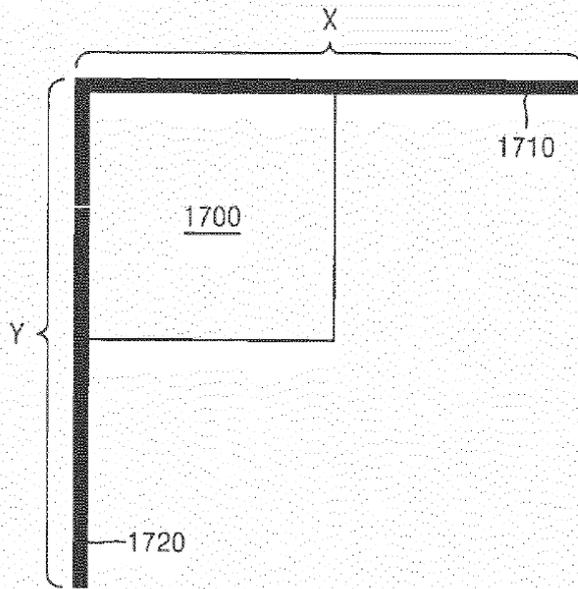
[Fig. 18b]



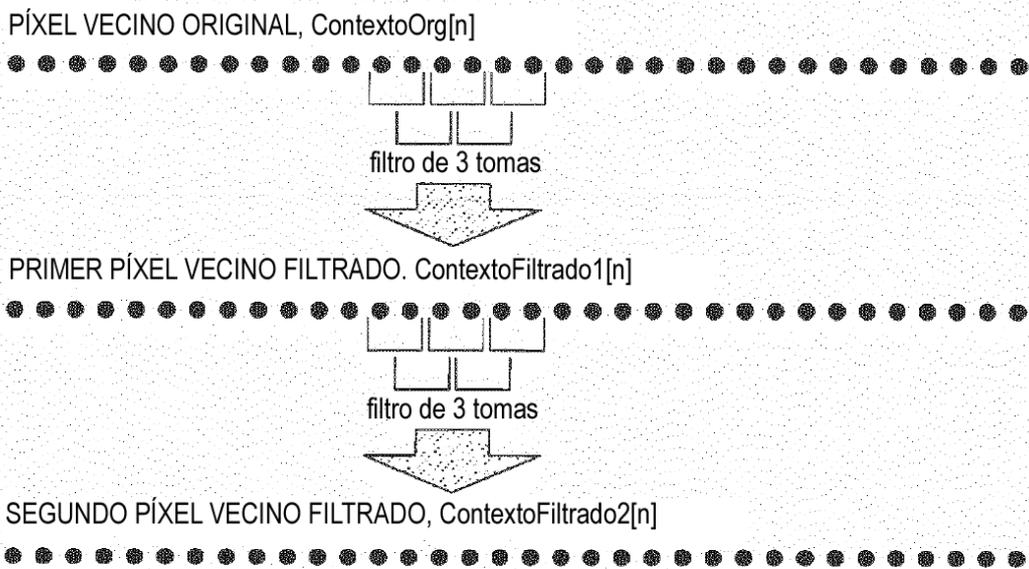
[Fig. 18c]



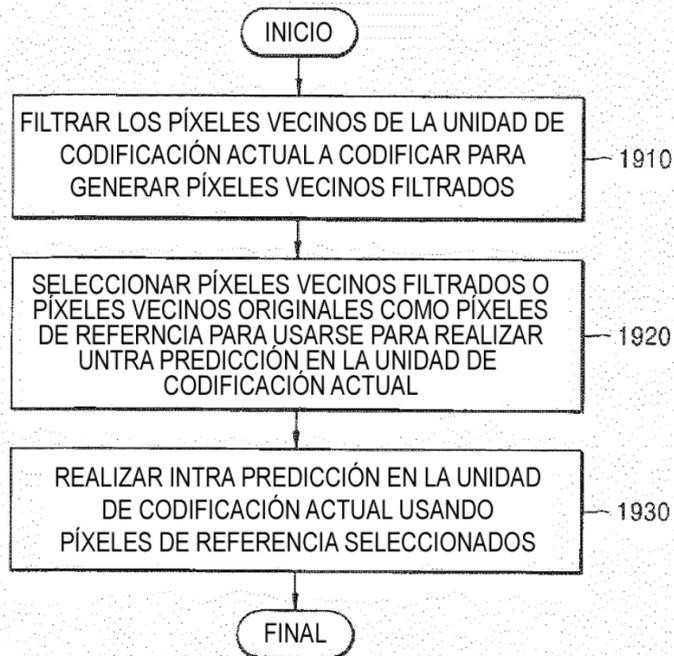
[Fig. 19]



[Fig. 20]



[Fig. 21]



[Fig. 22]

