

19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 823 480**

51 Int. Cl.:

H04N 19/513 (2014.01)

H04N 19/61 (2014.01)

H04N 19/96 (2014.01)

H04N 19/70 (2014.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

96 Fecha de presentación y número de la solicitud europea: **04.05.2011** **E 18192011 (7)**

97 Fecha y número de publicación de la concesión europea: **15.07.2020** **EP 3429205**

54 Título: **Método y aparato para procesar una señal de vídeo**

30 Prioridad:

04.05.2010 US 33090210 P

11.05.2010 US 33327310 P

12.11.2010 US 41280110 P

17.11.2010 US 41443610 P

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

07.05.2021

73 Titular/es:

LG ELECTRONICS INC. (100.0%)
20, Yeouido-dong, Yeongdeungpo-gu
Seoul 150-721, KR

72 Inventor/es:

PARK, SEUNGWOOK;
PARK, JOONYOUNG;
KIM, JUNGSUN;
CHOI, YOUNGHEE;
JEON, YONGJOON;
SUNG, JAEWON;
JEON, BYEONGMOON y
LIM, JAEHYUN

74 Agente/Representante:

MIR PLAJA, Mireia

ES 2 823 480 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

Método y aparato para procesar una señal de vídeo

Campo técnico

La presente invención se refiere a un método y a un aparato para procesar una señal de vídeo, y, más particularmente, a un método y a un aparato para codificar o decodificar una señal de vídeo.

Antecedentes de la técnica

Codificación por compresión se refiere a una tecnología de procesamiento de señales para transmitir información digitalizada a través de una línea de comunicaciones o almacenar dicha información en un formato que resulte adecuado para un soporte de almacenamiento. Se pueden codificar por compresión la voz, imágenes, caracteres, etcétera, y, en particular, a una tecnología para llevar a cabo la codificación de imágenes por compresión se le denomina compresión de imágenes de vídeo. La codificación por compresión de una señal de vídeo se puede llevar a cabo mediante la eliminación de información sobrante teniendo en cuenta la correlación espacial, la correlación temporal, la correlación probabilística, etcétera. No obstante, con el desarrollo reciente de diversos soportes y medios de transmisión de datos, existe una necesidad de un método y un aparato de procesamiento de señales de vídeo de alta eficiencia.

En la presentación "Video coding technology proposal by France Telecom, NTT, NTT DoCoMo, Panasonic and Technicolor" correspondiente a la 1ª Reunión de JCT-VC del 15 de abril, 2010 al 23 de abril, 2010, en Dresde, del JOINT COLLABORATIVE TEAM ON VIDEO CODING OF ISO/IEC JTC1/SC29/WG11 AND ITU-TSG.16, publicada como documento JCTVC-A114, Amonou et al. dan a conocer un esquema de predicción temporal en el que se forma un conjunto de candidatos a vector de movimiento con tres predictores de vector de movimiento (vecinos izquierdo y superior, y un vector ubicado conjuntamente). A partir de este conjunto, se infiere un subconjunto de dos predictores de entre los cuales se selecciona el predictor óptimo sobre la base de la comparación de los dos primeros candidatos del conjunto completo. Basándose en su igualdad, el tercer predictor del conjunto completo sustituye o no el segundo del subconjunto. Para la predicción entre bloques, se puede fijar un tamaño de partición, por el cual una señal en una de las particiones se predice usando parámetros de movimiento en un bloque de encima o a la izquierda. La señal de la otra partición se predice usando parámetros de movimiento a decodificar para el bloque en curso.

En la presentación "Test Model under Consideration (TMuC)" correspondiente a la 1ª Reunión del JCT-VC del 15 de abril, 2010 al 23 de abril, 2010, en Dresde, del JOINT COLLABORATIVE TEAM ON VIDEO CODING OF ISO/IEC JTC1/SC29/WG11 AND ITU-TSG.16, publicada como documento JCTVC-A205, se da a conocer un esquema de predicción que usa un proceso de inferencia para candidatos a predictor de vector de movimiento a partir de particiones de unidades de predicción vecinas. La predicción propuesta de vectores de movimiento para particiones rectangulares tiene en cuenta la distancia temporal de los cuadros de referencia. para reducir adicionalmente la velocidad de bits requerida para transmitir los vectores de movimiento, se intercala la predicción y la codificación de los componentes de un vector de movimiento.

Exposición

Problema técnico

Uno de los objetivos de la presente invención concebida para resolver el problema reside en la reducción de la información de predicción transmitida, restaurando una unidad en curso a través de un modo de fusión que usa información de predicción de otra unidad ya restaurada en la predicción inter de la unidad en curso.

Otro de los objetivos de la presente invención concebida para resolver el problema reside en implementar eficientemente un modo de predicción y en predecir de manera más precisa información de predicción de un bloque en curso.

Todavía otro objetivo de la presente invención concebida para resolver el problema reside en seleccionar unidades candidatas a la fusión adecuadas y en determinar eficientemente una unidad a fusionar teniendo en cuenta características de la unidad en curso y áreas vecinas fusionadas.

Todavía otro objetivo de la presente invención concebida para resolver el problema reside en proporcionar un método para mejorar la eficiencia en un método de señalización destinado a implementar un modo de predicción.

Solución técnica

La presente invención se ha diseñado para resolver los problemas anteriores, y el método para procesar una señal de vídeo de acuerdo con la presente invención usa una estructura para segmentar (*partitioning*) de manera recursiva una unidad de codificación en varias unidades de codificación, y un método para ello. Además, dicha unidad de codificación se divide en varias formas de unidades de predicción, y, de este modo, puede mejorarse la precisión de la compensación en la predicción de movimiento.

La presente invención puede usar un modo de fusión para incrementar la eficiencia de codificación. En este caso, se presenta un método de selección de unidades candidatas a la fusión en varias posiciones.

La presente invención presenta un método de señalización eficiente para especificar una unidad a fusionar entre unidades candidatas a la fusión. Además, la presente invención presenta un método para deducir una unidad a fusionar sin transmitir la información. Con este fin, se puede usar un método de determinación adaptativa de una unidad a fusionar teniendo en cuenta varias condiciones, tales como la posición de una unidad en curso y unidades vecinas, el tamaño de la unidad, información de movimiento, etcétera.

Efectos ventajosos

Según un método para procesar una señal de vídeo de acuerdo con la presente invención, se puede reducir la complejidad, que es necesaria para adquirir información de movimiento de una unidad en curso, mediante la fusión entre unidades al llevar a cabo una predicción inter, y se puede mejorar la eficiencia de codificación al no transmitir información de predicción de la unidad en curso.

Además, características de imágenes u objetos dentro de las imágenes pueden verse bien reflejadas por la predicción y la fusión en diversos tamaños de unidad y unidades de partición, y es posible una predicción más precisa.

Además, la flexibilidad de la fusión se puede ampliar seleccionando unidades vecinas de varias posiciones como unidades a fusionar, y se puede adquirir información de predicción más precisa.

Además, se pueden determinar de manera eficiente y adaptativa unidades candidatas a la fusión y/o unidades a fusionar teniendo en cuenta varias condiciones, tales como las posiciones de una unidad en curso y de unidades vecinas, el tamaño de la unidad, información de movimiento, etcétera.

Además, se establece la transmisión de información necesaria para un modo de fusión únicamente cuando es necesario, y se eliminan los casos innecesariamente redundantes, mejorándose así la eficiencia de codificación.

Descripción de los dibujos

La FIG. 1 es un diagrama de bloques esquemático de un dispositivo codificador de señales de vídeo de acuerdo con una realización ejemplificativa de la presente invención.

La FIG. 2 es un diagrama de bloques esquemático de un dispositivo descodificador de señales de vídeo de acuerdo con una realización ejemplificativa de la presente invención.

La FIG. 3 ilustra un ejemplo de segmentación de una unidad de codificación de acuerdo con una realización ejemplificativa de la presente invención.

La FIG. 4 ilustra varios métodos de segmentación de una unidad de predicción y denominaciones de los tipos de partición.

La FIG. 5 ilustra un caso de segmentación asimétrica de una unidad de predicción.

La FIG. 6 ilustra un caso de segmentación geométrica de una unidad de predicción.

La FIG. 7 ilustra candidatos a la fusión en varias posiciones para una unidad en curso.

La FIG. 8 ilustra un método para seleccionar candidatos a la fusión teniendo en cuenta el tamaño de una unidad.

La FIG. 9 es un diagrama de flujo que ilustra un método para seleccionar una unidad a fusionar cuando hay dos candidatos a la fusión.

La FIG. 10 es un diagrama de flujo que ilustra un proceso de obtención de información que es necesario para un modo de fusión usando el número de candidatos a la fusión disponibles.

La FIG. 11 ilustra un método para determinar una unidad que se puede fusionar usando información de movimiento de una partición vecina cuando el modo de partición es del tipo NxN.

5 Las FIGs. 12 a 13 ilustran un método para deducir una unidad a fusionar.

Las FIGs. 14 a 17 ilustran un método para deducir una unidad a fusionar en una partición específica.

Modo óptimo

10

La invención se expone en las reivindicaciones.

Modo para llevar a cabo la invención

15

A continuación se hará referencia detalladamente a las realizaciones preferidas de la presente invención, ilustrándose ejemplos de ellas en los dibujos adjuntos. En primer lugar, las terminologías o vocablos usados en esta memoria descriptiva y las reivindicaciones no se consideran como limitados a los significados generales o del diccionario, y deben considerarse de manera que presentan significados y conceptos que se corresponden con la idea técnica de la presente invención sobre la base del principio de que un inventor puede definir adecuadamente los conceptos de las terminologías para describir una invención de la mejor manera posible. La realización dada a conocer en esta exposición y las configuraciones mostradas en los dibujos adjuntos son simplemente una realización preferida, y no representan todas las ideas técnicas de la presente invención. Por lo tanto, se entiende que la presente invención abarca las modificaciones y variaciones de esta invención siempre que se sitúen dentro del alcance de las reivindicaciones adjuntas y sus equivalentes en el momento de presentar esta solicitud.

25

Los siguientes términos en la presente invención se pueden interpretar basándose en los siguientes criterios, e incluso términos que no se manifiestan pueden interpretarse de acuerdo con la intención que se describe a continuación. "Codificación" puede interpretarse como codificación o descodificación en función de la situación, e "información" incluye valores, parámetros, coeficientes, elementos, etcétera, y puede interpretarse como un concepto adecuado en función de la situación. "Pantalla" o "representación visual" se refiere en general a una unidad para indicar una imagen en una zona de tiempo específica, y franja (*slice*), cuadro, etcétera, es una unidad destinada a formar parte de una representación visual en la codificación de una señal de vídeo concreta, pero puede ser intercambiable. "Píxel" o "Pel" se refiere a una unidad mínima para constituir una imagen. "Muestra" se puede usar como término para indicar el valor de un píxel específico. "Muestra" se puede dividir en elementos de luma y de croma, pero, en general, se usa como un término que incluye ambos elementos. En lo anterior, "croma" indica una diferencia entre colores determinados, y comprende, en general, Cb y Cr. "Unidad" se ha usado para significar la unidad básica de un procesamiento de imágenes o una posición específica de una imagen, y se puede usar junto con términos tales como "bloque" o "área".

30

35

40

La presente invención se refiere a un método y a un aparato para codificar o decidir una señal de vídeo. La FIG. 1 es un diagrama de bloques esquemático de un dispositivo de codificación de señales de vídeo de acuerdo con una realización ejemplificativa de la presente invención. En referencia a la FIG. 1, el dispositivo 100 de codificación de la presente invención incluye una unidad 110 de transformada, una unidad 115 de cuantificación, una unidad 120 de cuantificación inversa, una unidad 125 de transformada inversa, una unidad 130 de filtrado, una unidad 150 de predicción y/o una unidad 160 de codificación entrópica.

45

La unidad 110 de transformada obtiene un coeficiente de transformada transformando un valor de píxel correspondiente a una señal de vídeo de entrada. Algunos ejemplos de un método de transformada de este tipo son la transformada discreta de coseno (DCT), la transformada discreta de seno (DST), la transformada por ondículas, etcétera. La unidad 110 de transformada lleva a cabo una transformada segmentando una señal de imagen de entrada en unidades de un cierto tamaño. A la unidad básica en la transformada se le denomina unidad de transformada. La eficiencia de la codificación se puede cambiar de acuerdo con la distribución y los atributos de valores dentro del área de la transformada.

50

55

La unidad 115 de cuantificación cuantifica un valor de un coeficiente de transformada obtenido a la salida de la unidad 110 de transformada. La unidad 120 de cuantificación inversa cuantifica inversamente el valor de coeficiente de transformada, y la unidad 125 de transformada inversa restaura un valor de píxel usando el valor de coeficiente de transformada cuantificado inversamente.

60

La unidad 130 de filtrado lleva a cabo una operación de filtrado con vistas a una mejora objetiva o subjetiva de la calidad de una imagen. Algunos ejemplos de los filtros usados en la unidad de filtrado son un filtro antibloques y/o un filtro de bucle adaptativo, etcétera. La unidad 156 de almacenamiento almacena una imagen a la cual se ha aplicado el filtrado, para dar salida a la imagen o usar la imagen como imagen de referencia.

En una codificación de vídeo general, una señal de imagen no se codifica como sí misma para mejorar la eficiencia de codificación, sino que se utiliza un método de predicción de una imagen usando un área ya codificada, y obteniendo una imagen restaurada mediante la adición de un valor residual entre la imagen original y la imagen predicha a la imagen predicha. Algunos ejemplos del método de predicción de una imagen son la predicción intra (predicción dentro de una pantalla), la predicción inter (predicción entre pantallas), etcétera, y, por lo tanto, la unidad de predicción puede incluir una unidad 152 de predicción intra y una unidad 154 de predicción inter. La unidad 152 de predicción intra lleva a cabo una predicción intra a partir de áreas restauradas dentro de la imagen en curso, transmitiendo así la información de codificación dentro de la pantalla a la unidad 160 de codificación entrópica. La unidad 154 de predicción inter obtiene un valor de vector de movimiento del área en curso, y lleva a cabo una compensación de movimiento entre pantallas en referencia a un área específica, es decir, área de referencia, de otra imagen restaurada usando el valor de vector de movimiento obtenido. Además, la unidad 154 de predicción inter transmite información de ubicación del área de referencia (información del sistema de referencia, vector de movimiento, etcétera) a la unidad 160 de codificación entrópica, de manera que la información se puede incluir en el flujo continuo de bits.

La unidad 160 de codificación entrópica genera un flujo continuo de bits de señal de vídeo codificando entrópicamente un coeficiente de transformada cuantificado, información de codificación inter, información de codificación intra, información de área de referencia, etcétera. La unidad 160 de codificación entrópica puede usar una codificación de longitud variable (VLC), una codificación aritmética, etcétera. El método de codificación de longitud variable transforma símbolos de entrada en palabras de código consecutivas, y la longitud de la palabra de código puede ser variable. Por ejemplo, los símbolos generados de manera frecuente se expresan como palabras de código cortas, y los símbolos generados de manera no frecuente se expresan como palabras de código largas. En calidad de método de codificación de longitud variable se puede usar la Codificación de Longitud Variable Adaptativa basada en el Contexto (CAVLC). La codificación aritmética transforma símbolos de datos consecutivos en un número primo, y puede obtener un bit de número primo óptimo para expresar cada símbolo. Como codificación aritmética, se puede usar la Codificación Aritmética Binaria Adaptativa basada en el Contexto (CABAC).

La FIG. 2 es un diagrama de bloques esquemático de un dispositivo 200 de descodificación de señales de vídeo de acuerdo con una realización ejemplificativa de la presente invención. El funcionamiento del dispositivo 200 de descodificación se corresponde con el funcionamiento de cada parte del codificador, y lleva a cabo un proceso inverso de la operación correspondiente. En referencia a la FIG. 2, el dispositivo 200 de descodificación de la presente invención puede incluir una unidad 210 de descodificación entrópica, una unidad 220 de cuantificación inversa, una unidad 225 de transformada inversa, una unidad 230 de filtrado y/o una unidad 250 de predicción, etcétera.

La unidad 210 de descodificación entrópica descodifica entrópicamente un flujo continuo de bits de señal de vídeo, y, de este modo, extrae el coeficiente de la transformada e información de predicción, etcétera, para cada área. La unidad 220 de cuantificación inversa cuantifica inversamente el coeficiente de transformada descodificado entrópicamente, y la unidad 220 de cuantificación inversa cuantifica inversamente el coeficiente de transformada descodificado entrópicamente y restaura el valor de píxel original usando el coeficiente de transformada cuantificado inversamente.

Además, la unidad 230 de filtrado mejora la calidad de la imagen filtrando la misma. En este caso, se pueden incluir un filtro antibloques para reducir el fenómeno de distorsión de bloques y/o un filtro de bucle adaptativo para eliminar la distorsión de la imagen completa, etcétera. La imagen, que ha sido filtrada, se proporciona en la salida o se almacena en la unidad 256 de almacenamiento para usarse como imagen de referencia.

Además, la unidad 250 de predicción de la presente invención incluye una unidad 252 de predicción intra y una unidad 254 de predicción inter, y restaura una imagen de predicción utilizando información, tal como un tipo de codificación descodificado a través de la unidad 210 de descodificación entrópica antes mencionada, un coeficiente de transformada correspondiente a cada área, un vector de movimiento, etcétera.

En este caso, la unidad 252 de predicción intra lleva a cabo una predicción intra a partir de la muestra descodificada dentro de la imagen en curso, generando, así, una imagen de predicción. La unidad 254 de predicción inter genera una imagen de predicción usando un vector de movimiento y una imagen de referencia almacenada en la unidad 256 de almacenamiento. Cuando se obtiene un vector de movimiento, se puede usar un método tal como la predicción de vectores de movimiento y la competición de vectores de movimiento, etcétera.

El cuadro de vídeo restaurado se genera sumando el valor residual correspondiente a cada píxel restaurado a partir de la unidad de transformada inversa a la imagen de predicción, a la cual se da salida desde la unidad 252 de predicción intra o la unidad 254 de predicción inter. Posteriormente se describirá el funcionamiento de la unidad 254 de predicción inter, en particular varios métodos de procesamiento en el modo de fusión.

En lo sucesivo en la presente, se describirá, en referencia a las FIGs. 3 a 5, un método de segmentación de una unidad de codificación y una unidad de predicción, etcétera, en el funcionamiento del dispositivo 100 de codificación y el dispositivo 200 de descodificación. Unidad de codificación se refiere a una unidad básica para procesar una imagen en un proceso tal como el proceso antes descrito de procesamiento de una señal de vídeo, por ejemplo, predicción intra o inter,

transformada, cuantificación, codificación entrópica, etcétera. El tamaño de la unidad de codificación usada en la codificación de una imagen puede que no sea homogéneo. La unidad de codificación puede tener una forma cuadrilátera, y una unidad de codificación se puede segmentar en varias unidades de codificación.

5 La FIG. 3 ilustra un ejemplo de segmentación de una unidad de codificación de acuerdo con una realización ejemplificativa de la presente invención. Por ejemplo, una unidad de codificación que tiene un tamaño de $2N \times 2N$ se puede segmentar en cuatro unidades de codificación que tienen un tamaño de $N \times N$. Una segmentación de la unidad de codificación de este tipo se puede llevar a cabo recursivamente, y no es necesario que todas las unidades de codificación se segmenten de la misma manera. No obstante, por comodidad en el proceso de codificación y procesamiento,
10 puede haber una restricción sobre el tamaño máximo 310 y el tamaño mínimo 320. Si se determina el tamaño máximo de la unidad de codificación, al tamaño máximo se le denomina tamaño máximo de unidad de codificación, si se determina el tamaño mínimo, al tamaño mínimo se le denomina tamaño mínimo de unidad de codificación.

15 Para una unidad de codificación se puede indicar información sobre si la unidad de codificación se segmenta. Por ejemplo, si el valor de la bandera que indica si la unidad de codificación se va a segmentar es 1, el bloque correspondiente al nodo se segmenta de nuevo en cuatro bloques, y si el valor de la bandera es 0, el bloque ya no se segmenta, y puede llevarse a cabo el proceso de procesamiento de la unidad de codificación.

20 El bloque no se tiene que segmentar necesariamente en cuatro áreas directas. En tal caso, códigos sobre métodos de segmentación predeterminados se pueden mapear con información de segmentación. Por ejemplo, si el valor de la información es 1, el bloque se puede dividir en dos subbloques rectangulares horizontales, si el valor es 2, el bloque se puede dividir en dos subbloques rectangulares verticales, y si el valor es 3, el bloque se puede dividir en cuatro subbloques cuadrados. Un método de este tipo representa algunos ejemplos del método de partición, pero la presente invención no se limita a ellos.

25 La estructura de la unidad de codificación antes descrita se puede indicar usando una estructura de árbol recursiva. Es decir, la unidad de codificación, que se segmenta en otras unidades de codificación usando la imagen individual o la unidad de codificación de tamaño máximo como raíz, acaba teniendo nodos hijo correspondientes al número de unidades de codificación segmentadas. Por tanto, la unidad de codificación, que ya no se segmenta más, se convierte en un nodo hoja. Suponiendo que, para una unidad de codificación, solamente es posible una segmentación de tipo cuadrado, una unidad de codificación se puede segmentar en un máximo de otras 4 unidades de codificación, y por lo tanto el árbol, que indica la unidad de codificación, puede tener un formato de árbol cuaternario. En lo sucesivo en la presente, por comodidad en cuanto a la explicación, a la unidad de codificación que tiene el tamaño de unidad de codificación máximo se le denomina unidad de codificación más grande (LCU), y a la unidad de codificación que tiene el tamaño de unidad de codificación mínimo se le denomina unidad de codificación más pequeña (SCU).
35

En el codificador, el tamaño óptimo de la unidad de codificación se puede seleccionar de acuerdo con características de una imagen de vídeo (por ejemplo, la resolución) o teniendo en cuenta la eficiencia de la codificación, y, en el flujo continuo de bits, se puede incluir información de la misma e información para dibujarla. Por ejemplo, se pueden definir el tamaño de la unidad de codificación máxima y la profundidad máxima del árbol. En el caso de una segmentación de tipo cuadrado, la altura y la anchura de la unidad de codificación son la mitad de la altura y la anchura de la unidad de codificación del nodo padre, y, por lo tanto, el tamaño mínimo de la unidad de codificación se puede calcular usando la información anterior. Además, y a la inversa, el tamaño mínimo de la unidad de codificación y la profundidad máxima del árbol se pueden definir de antemano, y el tamaño de la unidad de codificación máxima se puede deducir usando la información definida. En la segmentación de tipo cuadrado, el tamaño de la unidad se cambia a un formato de un múltiplo de 2, y, por lo tanto, el tamaño de la unidad de codificación concreta se indica como un valor logarítmico que tiene 2 como base, mejorándose así la eficiencia de transmisión.

40 En el decodificador se puede obtener información que indica si la unidad de codificación en curso antes descrita se ha segmentado. La eficiencia se puede mejorar si dicha información se obtiene (o transmite) únicamente bajo una condición específica. Por ejemplo, si la unidad de codificación en curso tiene el tamaño de unidad de codificación mínimo, la unidad no se segmenta en unidades de codificación más pequeñas, y, por lo tanto, en tal caso, no es necesario obtener información sobre si la unidad se ha segmentado.

55 Si la información indica que la unidad de codificación se ha segmentado, el tamaño de la unidad de codificación a segmentar se convierte en la mitad del tamaño de la unidad de codificación en curso, y la unidad de codificación se segmenta en cuatro unidades de codificación de tipo cuadrado sobre la base de la ubicación del procesado en curso. El procesamiento anterior se puede repetir para cada unidad de codificación segmentada.

60 La predicción de la imagen para la codificación se lleva a cabo para la unidad de codificación que ya no se segmenta más (es decir, el nodo hoja del árbol de la unidad de codificación). La unidad de codificación se divide en una o más unidades de predicción (PU), bloques o particiones de predicción, y a una división de este tipo se le denomina también partición. La forma en la que una unidad se ha segmentado se puede indicar mediante la información de tipo de unidad de predicción o la información de tipo de segmentación, etcétera. La FIG. 4 ilustra varios métodos de segmentación de

la unidad de predicción y un ejemplo de la denominación de tipo de segmentación. En referencia a la FIG. 4, una unidad de predicción que tiene un tamaño de $2N \times 2N$ puede no segmentarse (tipo $2N \times 2N$), se puede segmentar en dos particiones rectangulares que tengan un tamaño de $2N \times N$ (tipo $2N \times N$), se puede segmentar en dos particiones rectangulares que tengan un tamaño de $N \times 2N$ (tipo $N \times 2N$), o se puede segmentar en cuatro particiones cuadriláteras que tengan un tamaño de $N \times N$ (tipo $N \times N$). La unidad de predicción es una unidad básica para llevar a cabo la predicción, y la forma de segmentación de una unidad de predicción posible se puede definir de manera diferente en la unidad de codificación intra y la unidad de codificación inter. Por ejemplo, en la unidad de codificación intra, solamente se puede permitir una segmentación de tipo $2N \times 2N$ ó $N \times N$, y en la unidad de codificación inter, se puede permitir una segmentación de tipo $2N \times 2N$, $2N \times N$, $N \times 2N$, ó $N \times N$.

Si el tamaño de la unidad de codificación en curso es mayor que el tamaño mínimo predeterminado de la unidad de codificación, la segmentación de tipo $N \times N$ no se puede permitir debido a que, en este caso, puede obtenerse el mismo resultado que en el caso en el que la unidad de codificación se segmenta nuevamente.

La segmentación no se lleva a cabo necesariamente de manera simétrica. La FIG. 5 ilustra un caso de segmentación asimétrica de una unidad de predicción.

Además, la segmentación de la unidad de predicción también es posible geométricamente. Tal como se muestra en la FIG. 6, pueden generarse particiones con formas diversas diferentes a una forma cuadrilátera. En este caso, la unidad de predicción se puede segmentar en dos particiones 610 y 620 a través de una línea 600 de segmentación arbitraria. La información de segmentación se puede expresar en forma de una distancia (ρ) desde el centro, y un ángulo (θ) entre una línea perpendicular desde el centro (O) a la línea de segmentación y el eje básico (por ejemplo, eje x), etcétera. Dicha diversidad de formas de la segmentación de la unidad de predicción es ventajosa por cuanto es posible una predicción más precisa para una diversidad de formas de objetos incluidos en la imagen.

En la presente memoria descriptiva, a la unidad a codificar se le denomina unidad en curso, y a la imagen que incluye la unidad en curso se le denomina imagen en curso. Para restaurar la unidad en curso, se puede utilizar la información descifrada dentro de la imagen en curso o se puede utilizar la parte descifrada de otras imágenes. A la imagen (franja (*slice*)), que utiliza únicamente la imagen en curso para la restauración, es decir, lleva a cabo únicamente una predicción intra (predicción dentro de la pantalla), se le denomina representación visual intra o representación visual (franja) I, a la imagen (franja), que usa un máximo de un vector de movimiento o índice de referencia para predecir cada unidad, se le denomina representación visual predictiva o representación visual (franja) P, y a la imagen (franja), que usa un máximo de dos vectores de movimiento e índices de referencia, se le denomina representación visual de bipredicción o representación visual (franja) B.

En la unidad de predicción intra, se lleva a cabo la predicción intra, que predice el valor de píxel de la unidad en cuestión a partir de las áreas restauradas dentro de la imagen en curso. Por ejemplo, el valor de píxel de la unidad en curso se puede predecir a partir de los píxeles codificados de las unidades, que están ubicadas en los lados superior, izquierdo, superior izquierdo y/o superior derecho, sobre la base de la unidad en curso.

El modo intra se puede clasificar, en términos generales, en un modo vertical, un modo horizontal, un modo DC, un modo angular, etcétera, en función de la dirección del área de referencia en la que están ubicados los píxeles de referencia, que se usan en la predicción de los valores de píxel, y del método de predicción. El modo vertical usa el valor de un área vecina en vertical de la unidad en cuestión como valor de predicción de la unidad en curso, y el modo horizontal usa, como área de referencia, el área vecina en horizontal. En el modo DC, como valor de predicción se usa el valor medio de las áreas de referencia. Además, el modo angular es un caso en el que el área de referencia está ubicada en una dirección arbitraria, y la dirección se puede indicar por el ángulo entre el píxel en curso y el píxel de referencia. Por comodidad, se pueden usar el ángulo predeterminado y el número de modo de predicción, y el número de ángulos usados se puede cambiar en función del tamaño de la unidad en cuestión.

En relación con dicha diversidad de métodos de predicción se pueden definir y utilizar varios modos específicos. El modo de predicción se puede transmitir con el propio valor que indica el modo, aunque, para mejorar la eficiencia de transmisión, puede utilizarse el método de predicción del valor del modo de predicción de la unidad en curso. En este caso, el modo de predicción de la unidad en curso se puede obtener basándose en la información sobre si el valor predicho en el modo de predicción se usa como él mismo en el descodificador y la diferencia con el valor real.

En la unidad de predicción inter, se lleva a cabo la predicción inter la cual predice el valor de píxel de la unidad en cuestión usando información de imágenes ya restauradas diferentes a la imagen en curso. A la imagen usada en la predicción se le denomina representación visual de referencia. El área de referencia que se utiliza en la predicción de la unidad en curso en el proceso de predicción inter se puede indicar usando el índice que indica la imagen de referencia que incluye el área de referencia (a la que, en lo sucesivo en la presente, se denominará como "índice de referencia") e información de vector de movimiento, etcétera.

Algunos ejemplos de la predicción inter (predicción entre pantallas) son la predicción en sentido directo, la predicción en sentido retrógrado y la bipredicción. La predicción en sentido directo es una predicción que utiliza una representación visual de referencia visualizada (o a la que se ha dado salida) temporalmente antes de la representación visual en curso, y la predicción retrógrada es una predicción que usa una representación visual de referencia visualizada (o a la que se ha dado salida) temporalmente después de la representación visual en curso. Con este fin, puede que sea necesario un conjunto de información de movimiento (por ejemplo, un vector de movimiento y un índice de representación visual de referencia). En la bipredicción, puede utilizarse un máximo de dos áreas de referencia, y estas dos áreas de referencia pueden existir en la misma representación visual de referencia, y pueden existir en representaciones visuales diferentes. Las representaciones visuales de referencia se pueden visualizar (o se les puede dar salida) temporalmente tanto antes como después de la representación visual en curso. En el método de bipredicción, puede utilizarse un máximo de dos conjuntos de información de movimiento (por ejemplo, un vector de movimiento y un índice de representación visual de referencia).

La información de movimiento de la unidad de predicción en curso puede incluir la información de vector de movimiento y el índice de representación visual de referencia. Información de vector de movimiento puede significar el vector de movimiento, el valor de predicción del vector de movimiento, o el vector de movimiento diferencial, o también puede significar información de índice que especifique un valor de predicción del vector de movimiento. Vector de movimiento diferencial significa una diferencia entre el vector de movimiento y un valor de predicción del vector de movimiento.

El bloque de referencia de la unidad de predicción en curso se puede obtener usando el vector de movimiento y el índice de representación visual de referencia. El bloque de referencia existe dentro de la representación visual de referencia que tiene el índice de representación visual de referencia. Además, el valor de píxel del bloque especificado por el vector de movimiento se puede utilizar como predictor de la unidad de predicción en curso. Es decir, se usa la compensación por movimiento, la cual predice la imagen de la unidad de predicción en curso estimando el movimiento a partir de la representación visual descodificada previamente.

Además de la imagen en curso, la lista de imágenes de referencia puede estar constituida por imágenes usadas para la predicción inter. Una franja B necesita dos listas de imágenes de referencia, y las listas se denominan, respectivamente, lista de referencia 0 y lista de referencia 1. Entre las franjas B, a la franja en la que la lista de referencia 0 es igual a la lista de referencia 1 se le denomina, en particular, franja GPB.

Para reducir el volumen de transmisión en relación con el vector de movimiento, el valor de predicción del vector de movimiento se puede obtener usando la información de movimiento de las unidades codificadas, y se puede transmitir solamente la diferencia del vector de movimiento. En el descodificador, el valor de predicción del vector de movimiento de la unidad en curso se puede obtener usando información de movimiento de otras unidades descodificadas, y el valor del vector de movimiento en la unidad en curso se puede obtener usando la diferencia transmitida. Cuando se obtiene el valor de predicción del vector de movimiento, puede usarse un método de competición de vectores de movimiento, método que obtiene diversos valores de candidatos de vector de movimiento usando la información de movimiento de las unidades ya codificadas y obtiene un valor de predicción de vector de movimiento entre los valores candidatos.

La información de predicción (por ejemplo, el índice de referencia, el vector de movimiento, el sentido de la predicción, etcétera), que es necesaria en la predicción inter de la unidad en curso, no se incluye directamente en el flujo continuo de bits cuando se transmite, y se puede deducir usando la unidad vecina. Usando un método de este tipo, se puede mejorar la tasa de compresión utilizando el número de bits asignados a la información de predicción. Específicamente, como información de predicción de la unidad en curso se puede utilizar la información de predicción de la unidad vecina codificada usando la predicción inter. Cuando se usa un método de este tipo, se expresa que la unidad en curso se ha fusionado con la unidad vecina que entregó información de predicción, y a dicho método de predicción se le denomina modo de fusión.

Para el modo de fusión, el flujo continuo de bits puede incluir información que indique si la unidad en curso se ha fusionado (por ejemplo, una bandera como bandera_fusión (merge_flag)), información de fusión que indique qué unidad vecina se ha fusionado con la unidad en curso (por ejemplo, una bandera que indique si la unidad en curso se ha fusionado con un cierto vecino, o información de índice que indique un cierto vecino, etcétera), etcétera. La información que indica qué unidad vecina se ha fusionado con la unidad en curso se puede establecer de manera que se obtenga únicamente cuando se indique que la unidad en curso se ha fusionado (en el ejemplo anterior, cuando bandera_fusión sea verdadero ó 1).

Además, cuando se usa el modo de predicción inter, se puede aplicar el modo de omisión en unidades de la unidad de codificación. El modo de omisión es un método de predicción que transmite únicamente cierta información (por ejemplo, información que indica cuál de los diversos candidatos de predicción por vectores de movimiento se usará) entre información para la restauración por predicción de las unidades en curso. En este caso, se puede utilizar en sí misma información de otras unidades ya codificadas. Cuando se usa el modo de omisión, se puede reducir el volumen de información transmitida y, por lo tanto, en primer lugar, se determina si la unidad de codificación está en el modo de

omisión y, en caso contrario, se puede usar otro modo (por ejemplo, un modo de fusión, un modo de predicción directa o un modo de predicción inter general).

El modo de fusión se puede aplicar en unidades de la unidad de codificación, o se puede aplicar en unidades de la unidad de predicción. En caso de que el modo de fusión se aplique en unidades de la unidad de codificación, se transmite una bandera que indica si aplicar fusión a cada unidad de codificación mínima que usa un modo de predicción inter. Tal como se ha descrito anteriormente, el modo de omisión se puede aplicar a la unidad de codificación y, por lo tanto, después de comprobar, en primer lugar, si el modo de omisión se va a aplicar en unidades de la unidad de codificación mínima (por ejemplo, usando un método tal como análisis sintáctico y comprobación de la bandera que indica si aplicar el modo de omisión), únicamente cuando no se aplica el modo de omisión, se puede obtener la bandera de aplicación del modo de fusión.

Tal como se ha descrito anteriormente, la unidad de codificación se puede segmentar de diversas maneras en unidades de la unidad de predicción. En caso de aplicar el modo de fusión en unidades de la unidad de predicción, la bandera de fusión se obtiene, respectivamente, para todas las particiones de modo inter en las cuales no se aplica el modo de omisión (o el modo de predicción directa).

En lo sucesivo en la presente se describe el funcionamiento específico del modo de fusión. Cuando se trabaja en el modo de fusión, la unidad que se tomará como objetivo puede incluir tanto las unidades de la unidad de codificación como la unidad de predicción (segmentación).

A las unidades que se pueden fusionar con la unidad en curso se les denomina candidatos a la fusión. Los candidatos a la fusión se pueden seleccionar como unidades adyacentes a la unidad en curso entre unidades ya restauradas. A la unidad que se fusiona con la unidad en curso, es decir, la unidad que lleva la información de predicción, se le denomina unidad a fusionar. Con el fin de determinar la unidad a fusionar, se puede utilizar información, obtenida a partir del flujo continuo de bits, que indique qué unidad vecina se ha fusionado, y la unidad a fusionar se puede deducir usando una regla específica.

En lo sucesivo en la presente, se describirán, en referencia a las FIGs. 7 y 8, los tipos de unidades que se pueden convertir en candidatos a la fusión basándose en la unidad en curso, se describirá, en referencia a las FIGs. 9 y 10, el método para transmitir información necesaria para la fusión, y se describirán, en referencia a las FIGs. 11 a 17, diversas realizaciones para deducir la unidad a fusionar entre unidades candidatas a la fusión.

La FIG. 7 ilustra candidatos a la fusión en diversas posiciones para una unidad en curso. Los candidatos a la fusión pueden existir dentro de la imagen 700 en curso en la que está ubicada la unidad en curso, o pueden existir dentro de otra imagen ya restaurada 750. A las unidades dentro de la imagen en curso se les puede denominar candidatos a la fusión espaciales, y a las unidades dentro de otra imagen se les puede denominar candidatos a la fusión temporales. Algunos ejemplos de candidatos a la fusión espaciales son las unidades 710 del área izquierda adyacente a la unidad en curso, las unidades 720 del área superior, y/o las unidades del área de las esquinas (C, C-1, C-2). Los candidatos a la fusión temporales pueden llegar a ser unidades (D) ubicadas en una posición específica de otra imagen, por ejemplo, una posición correspondiente a la unidad en curso, etcétera. Aún cuando la FIG. 7 ilustra unidades de un tamaño específico, el tamaño de cada unidad puede variar en función del nivel segmentado según se ha descrito anteriormente.

Con vistas a la eficiencia en la codificación, los candidatos a la fusión se pueden seleccionar como un cierto número de unidades teniendo en cuenta la eficiencia en la codificación y la complejidad del cálculo entre las posiciones.

Algunos ejemplos de los candidatos a la fusión son la unidad superior (A) entre las unidades del área izquierda, una unidad (A-1) seleccionada entre las unidades del área izquierda excepto A, la unidad de más a la izquierda (B) entre las unidades del área superior, una unidad (B-1) seleccionada entre las unidades del área superior excepto B, la unidad (C) en la esquina superior derecha, la unidad (C-1) en la esquina inferior izquierda, la unidad (C-2) en la esquina superior izquierda, y una unidad (D) correspondiente a otra imagen. Es evidente que, cuando hay solamente una unidad adyacente a la unidad en curso en el área izquierda o superior, no hay ninguna unidad seleccionada adicionalmente.

De manera similar, es posible constituir candidatos a la fusión con la unidad seleccionada del área izquierda, la unidad seleccionada del área superior, las unidades del área de las esquinas (C, C-1, C-2) y la unidad (D) correspondiente a otra imagen, etcétera. Aquí, el número de unidades que se seleccionan en el área izquierda o superior puede ser preferentemente uno (en este caso, el número máximo de candidatos a la fusión sería 6), o un número específico predeterminado. Las unidades del área de las esquinas también se pueden usar seleccionando solamente algunas de las unidades en función de las necesidades.

Para simplificar, como candidatos a la fusión se puede determinar el número menor de unidades. Por ejemplo, la unidad superior (A) entre las unidades del área izquierda, la unidad de más a la izquierda (B) entre las unidades del área superior, la unidad (C) en la esquina superior derecha, la unidad (C-1) en la esquina inferior izquierda, y una unidad (D)

correspondiente a otra imagen, etcétera, pueden convertirse en candidatos a la fusión. En este caso, el número máximo de candidatos a la fusión sería 5.

En su expresión más sencilla, únicamente dos unidades, por ejemplo, la unidad superior (A; a la que en lo sucesivo en la presente se hará referencia como unidad vecina superior) entre las unidades del área izquierda, y la unidad de más a la izquierda (B; a la que, en lo sucesivo en la presente, se hará referencia como unidad vecina izquierda) entre las unidades del área superior pueden llegar a ser los candidatos a la fusión.

Tal como se ha mencionado anteriormente, el tamaño de la unidad puede variar en función del nivel de segmentación o del modo de segmentación. Por tanto, los candidatos a la fusión se pueden determinar teniendo en cuenta el tamaño de la unidad así como la posición (más a la izquierda o superior, etcétera). Por ejemplo, la selección se puede realizar basándose en la longitud del borde adyacente a la unidad en curso. Es decir, se selecciona la unidad que es tangente al contorno más largo entre las unidades de las áreas izquierda o superior adyacentes a la unidad en curso. En referencia a la FIG. 8, por ejemplo, las unidades B, C, y D están posicionadas en el área superior adyacente a la unidad en curso X. Tal como se ha descrito anteriormente, las unidades B y C pueden ser unidades de codificación que se han segmentado una vez más en comparación con la unidad D, o se pueden haber segmentado en un modo (N x N) diferente al modo de la unidad D (2N x 2N). Entre las unidades B, C y D, la unidad D es tangente a la unidad en curso X por el contorno más largo, y, por lo tanto, la unidad D se selecciona como candidata a la fusión entre las unidades vecinas superiores. También es posible seleccionar una o más unidades en el orden anterior entre las unidades vecinas superiores B, C, y D. Los candidatos a la fusión también se pueden seleccionar teniendo en cuenta el área de la unidad además de la longitud del contorno adyacente.

En caso de que el número máximo de los candidatos a la fusión sea 2, la información sobre cuál de los dos candidatos se fusionará puede venir indicada por una bandera de un bit. La FIG. 9 es un diagrama de flujo que ilustra un proceso de selección de una unidad a fusionar en el ejemplo anterior, en el cual la unidad vecina izquierda A ó la unidad vecina derecha B se convierte en el candidato a la fusión. En primer lugar, se obtiene (S901) una bandera de fusión A (bandera_fusión) que indica si la unidad en curso está codificada en un modo de fusión. Si la unidad en curso está codificada en un modo de fusión sobre la base de la información de la bandera de fusión (es decir, bandera_fusión = VERDADERO ó 1), se obtiene (S930) la bandera de sentido de fusión (bandera_fusión_izquierda (merge_left_flag)) que indica con qué unidad se ha fundido. La bandera de sentido de fusión presenta el valor VERDADERO (ó 1) cuando se fusiona con la unidad vecina izquierda, y presenta el valor FALSO (ó 0) cuando se fusiona con la unidad vecina superior. Por tanto, si el valor de la bandera de sentido de fusión es 1 (S940), como información de predicción de la unidad en curso X se usa la información de predicción de la unidad vecina izquierda A (S950), y si el valor de la bandera de sentido de fusión es 0, como información de predicción de la unidad en curso X se usa la información de predicción de la unidad vecina superior B (S970). En caso de que la bandera de fusión indique un caso diferente al modo de fusión, la información de predicción de la unidad en curso se obtiene usando un modo de predicción general (S960). En lo anterior, la bandera de sentido de fusión es meramente una cierta realización, y, por lo tanto, dicha bandera no indica necesariamente la fusión con la unidad vecina izquierda y se puede usar en un formato que indique si se va a producir una fusión con una unidad específica entre candidatos a la fusión.

En caso de que el número máximo de candidatos a la fusión sea dos o mayor, la unidad a fusionar no se puede especificar solamente con la información de la bandera, y, por lo tanto, es necesario utilizar otros métodos. Por ejemplo, se puede usar información de índice (índice de fusión) sobre qué unidad entre los candidatos a la fusión se fusionará, es decir, información que especifica la unidad a fusionar. El índice se puede fijar en un cierto orden, o se puede cambiar de manera adaptativa de acuerdo con la situación.

Además, según una realización ejemplificativa de la presente invención, la información anterior se puede transmitir de manera más eficiente en función del caso. Es decir, en una cierta condición en la que no es necesario transmitir la información anterior, puede omitirse la transmisión de esta información.

La FIG. 10 es un diagrama de flujo que ilustra un proceso de obtención de información que es necesaria para un modo de fusión usando el número de candidatos a la fusión disponibles. En primer lugar, se determina el número de candidatos a la fusión (NumCandidatos (NumCandidate)) que puede ser usado por la unidad en curso, comprobando las posiciones de los candidatos a la fusión antes descritos (S1010). Estar disponible significa aquí que la unidad en curso puede obtener información de predicción a partir de la unidad candidata a la fusión. Es decir, este es el caso en el que la unidad existe en la posición, la unidad está codificada en el modo inter y tiene información de predicción, y ya ha sido descodificada. Si el número de candidatos a la fusión disponibles es 0, no se puede aplicar el modo de fusión, y, por lo tanto, no es necesario transmitir la bandera de fusión (bandera_fusión), que indica si se va a producir fusión. En este caso, la información de predicción de la unidad en curso se obtiene usando un modo de predicción inter general (S1070). Por lo tanto, comprobando la magnitud del número de candidatos a la fusión disponibles (NumCandidatos) (S1020), la bandera de fusión (bandera_fusión) se obtiene únicamente cuando el número de los candidatos a la fusión es mayor que 0 (S1030). En caso de que el número de los candidatos a la fusión sea 1, como unidad a fusionar se puede determinar un candidato a la fusión. Por tanto, no es necesario transmitir por separado información (bandera o índice) sobre la unidad con la que se fusiona la unidad en curso, y, de este modo, solamente se obtiene dicha

información de fusión (S1050) en caso de que la bandera que indica si se va a producir fusión sea 1 y el número de candidatos a la fusión sea mayor que 1 (S1040). La unidad a fusionar se determina basándose en la información de fusión obtenida, y la información de predicción de la unidad a fusionar se usa como información de predicción de la unidad en curso (S1060).

En el caso en el que el tipo de segmentación de la unidad de codificación de un tamaño $2N \times 2N$ sea $2N \times 2N$, es decir, en el caso en el que la unidad de codificación esté constituida por una partición, puede usarse el modo de predicción directa. Por tanto, en este caso, el número de candidatos a la fusión (NumCandidatos (NumCandidates)) se fija a cero con independencia de la información de movimiento de la unidad vecina, y, de este modo, se puede establecer que el modo de fusión no se aplique a la unidad de codificación y también se puede establecer que no se transmita la información relacionada.

No obstante, en este caso, puede que resulte imposible que la información de predicción de la unidad en curso (sentido de predicción, índice de referencia, vector de movimiento, etcétera) se exprese con el modo de predicción de sentido $2N \times 2N$. También puede que resulte necesario usar un modo mejorado de predicción de sentido para compensar una pérdida generada en tal caso.

De acuerdo con una realización ejemplificativa de la presente invención, puede ejecutarse una señalización variable en función del número de candidatos a la fusión. Por ejemplo, en caso de que el número de candidatos a la fusión, que resulten posibles para la unidad en curso, sea 3, se transmite información para distinguir de 0 a 2, y en caso de que el número de los posibles candidatos a la fusión sea 2, se transmite únicamente información para distinguir 0 y 1. En este caso, se produce la ventaja de que el número de bits asignados al índice se puede reducir cuando el número de candidatos a la fusión es pequeño.

Es posible usar un método de cambio de las unidades que pertenecen a los candidatos a la fusión después de fijar el número máximo de candidatos. Por ejemplo, se comprueban en orden posiciones predeterminadas, y solamente las unidades correspondientes, en cuanto a número, al número máximo de candidatos se convierten en los candidatos a la fusión, y la unidad a fusionar se selecciona solamente entre los candidatos a la fusión. Por ejemplo, se supone que el número máximo de candidatos es 4, y los candidatos a la fusión se comprueban en el orden de A, B, C, C-1 y D. Si está disponible la totalidad de las unidades anteriores, las cuatro unidades A, B, C y C-1 se convierten en los candidatos a la fusión de acuerdo con dicho orden. Si la unidad C no está disponible en el ejemplo anterior, como candidatos a la fusión se determinarán cuatro unidades A, B, C-1 y D. Si el número de candidatos a la fusión disponibles es menor que el número máximo de candidatos, es posible usar el método de señalización variable antes explicado, en función del número de candidatos a la fusión.

Cuando el modo de segmentación es del tipo $N \times N$, es decir, la unidad de codificación está constituida por cuatro particiones cuadriláteras, la condición para determinar candidatos a la fusión disponibles se puede añadir en la cuarta partición. Una condición de este tipo se puede usar en el proceso de contar el número de candidatos a la fusión disponibles o deducir las unidades a fusionar. En referencia a la FIG. 11, por ejemplo, en el caso de la partición n.º 3, la partición con la cual la partición en curso se puede fusionar de manera adaptativa se determina de acuerdo con la información de movimiento de otras tres particiones (partición n.º 0 a 2).

Por ejemplo, cuando la partición n.º 0 y la partición n.º 1 tienen la misma información de movimiento, y la partición n.º 2 tiene una información de movimiento diferente, la partición n.º 3 no se puede fusionar con la partición n.º 2 ya que, en tal caso, la segmentación sería redundante con la segmentación de un tipo $2N \times N$. Por lo tanto, en este caso, la partición n.º 2 queda exenta de los candidatos a la fusión.

Cuando la partición n.º 0 y la partición n.º 2 tienen la misma información de movimiento, y la partición n.º 1 tiene información de movimiento diferente, la partición n.º 3 no se puede fusionar con la partición n.º 1 debido a que, en tal caso, la segmentación sería redundante con la segmentación de un tipo $N \times 2N$. Por lo tanto, en este caso, la partición n.º 3 queda exenta de los candidatos a la fusión.

Cuando la partición n.º 0, la partición n.º 1 y la partición n.º 2 tienen, todas ellas, la misma información de movimiento, la partición n.º 3 no se puede fusionar con ninguna de la partición n.º 1 y la partición n.º 2 ya que, en tal caso, la fragmentación sería redundante con la fragmentación del tipo $2N \times 2N$. Por lo tanto, en este caso, la partición n.º 1 y la partición n.º 2 quedan exentas de los candidatos a la fusión.

Además, con el fin de reducir el número de bits para la transmisión de información con vistas a determinar una unidad a fusionar, en lugar de transmitir la información, puede usarse un método de deducción de la unidad a fusionar de acuerdo con una cierta regla. En lo sucesivo en la presente, en referencia a la FIG. 12, se describirá un método de determinación de la unidad a fusionar entre unidades candidatas a la fusión disponibles. En la FIG. 12, se supone que solamente las unidades A y B son unidades candidatas a la fusión para facilitar la explicación, aunque las unidades de las diversas posiciones antes descritas pueden llegar a ser candidatos a la fusión.

De acuerdo con una realización ejemplificativa de la presente invención, la unidad a fusionar se puede seleccionar sobre la base de la longitud del contorno según el cual la unidad en curso es tangente a la unidad vecina. En referencia a la FIG. 12(a), por ejemplo, el contorno entre la unidad en curso X y la unidad vecina superior A es mayor que el contorno entre la unidad en curso A y la unidad vecina izquierda B, y, por lo tanto, la unidad a fusionar se selecciona como unidad vecina superior A. Es decir, la unidad en curso X se fusiona con la unidad vecina superior A, y, como información de predicción de la unidad en curso X, se usa la información de predicción de la unidad vecina superior A.

De acuerdo con una realización ejemplificativa de la presente invención, como unidad a fusionar de la unidad en curso se puede seleccionar la unidad que tiene un vector de movimiento más similar a la unidad vecina de una cierta posición (por ejemplo, superior izquierda) entre las unidades vecinas de la unidad en curso. En referencia a la FIG. 12(b), por ejemplo, la unidad vecina superior A tiene un vector de movimiento que es más similar a la unidad vecina superior izquierda C que la unidad vecina izquierda B, y, por lo tanto, como unidad a fusionar, se selecciona la unidad vecina superior A.

De acuerdo con una realización ejemplificativa de la presente invención, como unidad a fusionar se puede seleccionar la unidad con un área más amplia entre los candidatos a la fusión. En referencia a la FIG. 12(c), por ejemplo, el área de la unidad vecina izquierda B es mayor que la unidad superior A, y, por tanto, como unidad a fusionar se selecciona la unidad vecina izquierda B.

Además, las realizaciones anteriores se pueden combinar con un cierto orden de prioridad.

De acuerdo con una realización ejemplificativa de la presente invención, cuando se usa un modo de fusión, la unidad en curso se segmenta geoméricamente de manera automática, y cada área de segmentación se fusiona con la unidad vecina adyacente. En referencia a la FIG. 13, por ejemplo, la unidad en curso se segmenta en forma de X1 y X2 según la línea de segmentación 1300. La línea de segmentación 1300 se determina usando la distancia euclídea de las unidades vecinas A y B que son candidatos a la fusión. Consecuentemente, X1 se fusiona con A, y X2 se fusiona con B.

Además, en la unidad segmentada, cuando no hay ningún área adyacente o hay solamente un área adyacente, la unidad a fusionar se puede deducir sin transmitir información específica.

De acuerdo con una realización ejemplificativa de la presente invención, en el caso de una segmentación geométrica, puede que no haya ningún área restaurada adyacente. La FIG. 14 ilustra un caso en el que la unidad en curso se ha segmentado geoméricamente. En este caso, la unidad en curso se segmenta en dos particiones, es decir, X1 y X2, y X2 no es adyacente a la región reconstruida. Aquí, en el caso de la partición X2, no se puede permitir la fusión con X1. Esto es debido a que, cuando X1 y X2, que pertenecen a la misma unidad de codificación, tienen la misma información de predicción debido a la fusión, no hay necesidad de segmentación. Por lo tanto, en una situación de este tipo, el modo de fusión no se puede aplicar a la partición X2. En este caso, como en el caso en el que el número de candidatos a la fusión disponibles es 0, no se pueden obtener ni la bandera que indica si se va a producir fusión ni la información que indica la unidad a fusionar.

La FIG. 15 ilustra un caso en el que la unidad en curso se segmenta en dos particiones rectangulares. En referencia a la FIG. 15(a), en el caso de la partición X2, el área superior es adyacente al área restaurada, pero el área izquierda es adyacente a la partición X1 que pertenece a la misma unidad de codificación. De manera similar a la descripción anterior, la situación en la que la partición X2 se fusiona con la partición X1, que pertenece a la misma unidad de codificación, y que acaba teniendo la misma información de predicción, es redundante con el caso sin segmentación, y, por lo tanto, es necesario que dicho caso quede exento. Por lo tanto, en el caso de X2, puede que resulte necesario aplicar el modo de fusión, pero las unidades del área izquierda no pueden ser candidatos a la fusión. Cuando se cuenta el número de candidatos a la fusión disponibles, pueden tenerse en cuenta dichos puntos. Tal como se ha descrito antes, si solamente pueden llegar a ser candidatos a la fusión dos tipos, es decir, la unidad izquierda y la unidad superior (es decir, si se transmite información de tipo bandera para indicar la unidad a fusionar), no es necesaria la transmisión de la bandera para indicar la unidad a fusionar con respecto a X2. El número de candidatos a la fusión es grande, y cuando se usa información de índice adaptativa que se ha descrito anteriormente, esto resulta ventajoso por cuanto se reduce el volumen de la información a usar en el índice.

Además, la unidad que tiene una información de predicción que es igual a la información de predicción de otra partición X1, que pertenece a la misma unidad de codificación que la unidad en curso, no puede ser un candidato a la fusión disponible. Cuando se segmenta la misma unidad de codificación en dos particiones, es decir, X1 y X2, el caso en el que X2 se fusiona con una cierta unidad y llega a tener el mismo valor que la información de predicción de X2 conduce a un resultado igual que en la fusión con X1. Por tanto, en este caso, resulta que la segmentación de una unidad de codificación en X1 y X2 no tiene sentido. Por esta razón, en la presente invención, de los candidatos a unidad de fusión disponibles se eliminan las unidades que buscan la misma información de predicción que X1 entre unidades candidatas a la fusión de diversas posiciones que pueden llegar a convertirse en candidatos de predicción en lo anterior. Es decir, el número de candidatos a unidad de fusión que pueden ser usados por la partición en curso, se reduce por el número de candidatos a fusión que tienen una información de predicción igual a la de otra partición X1 que pertenece a la

misma unidad de codificación. De esta manera, el número de unidades candidatas a la fusión disponibles se convierte en 0, y puede que no sea necesario transmitir la bandera de fusión que indica si aplicar el modo de fusión. Por ejemplo, en referencia a la FIG. 15(a), en el caso en el que la información de predicción de la unidad vecina superior A es igual a la de X1, X2 no se puede fusionar con ninguna de entre A y X1. Por tanto, en este caso, puesto que no hay ninguna unidad candidata a la fusión disponible, no es necesario transmitir ni la bandera que indica si aplicar el modo de fusión ni información que indica la unidad a fusionar. Es decir, esto es igual al caso en el que el número de candidatos a la fusión disponibles es 0.

Por contraposición, en el caso de la FIG. 15(b), el lado izquierdo de la partición X2 es adyacente a la región restaurada, pero el lado superior no es adyacente a la región restaurada y resulta adyacente a otra partición X1 que pertenece a la misma unidad de codificación. Por tanto, en este caso, X2 no se fusiona con X1, y únicamente el área izquierda se convierte en un candidato a la fusión disponible. En el caso del ejemplo de la FIG. 15(b), el candidato a la fusión que queda es solamente la unidad A, y, por lo tanto, no es necesario transmitir la información de tipo bandera para designar la unidad a fusionar. Asimismo, las unidades que tienen la misma información de predicción que X1 no se pueden fusionar con X2, y, por lo tanto, en el caso en el que el área adyacente izquierda A de X2 tiene la misma información de predicción que la correspondiente de X1, no queda ninguna unidad candidata a la fusión disponible y no es necesario transmitir la información de bandera de fusión que indica si aplicar el modo de fusión.

La FIG. 16 ilustra varios ejemplos en el caso en el que una unidad se segmenta de manera asimétrica en dos particiones. En este caso, se puede establecer que una partición, que ocupa un área más pequeña, se fusione con solamente una unidad adyacente al lado largo. En el caso de la FIG. 16(a), el lado largo de X1 es el área izquierda, y en el caso en el que se seleccione una de entre el área izquierda o el área superior, la unidad a fusionar se determina como unidad A sin transmitir la bandera (bandera_fusión_izquierda) que designa la unidad. En el caso de la FIG. 16(b), el lado largo de la X1 es el área derecha, y la unidad a fusionar se determina como unidad A en la misma situación que la anterior.

La FIG. 17 ilustra varios ejemplos en el caso en el que una unidad se segmenta geométricamente en dos particiones. En este caso, la unidad a fusionar se puede deducir de acuerdo con el tipo de segmentación. Aquí, el tipo de segmentación se puede considerar usando la longitud de la línea de borde entre la partición y la región reconstruida. En referencia a la FIG. 17, con respecto a la primera partición geométrica X1, se supone que la longitud de la línea de borde entre X1 y el área restaurada izquierda es a, la longitud de la línea de borde entre X1 y el área restaurada superior es b, y c1 y c2 son ciertos umbrales predeterminados. Las FIGs. 17(a) y 17(b) indican los tipos de segmentación de varias segmentaciones geométricas, y esto se puede especificar como un valor específico usando a y b, por ejemplo, un valor como a/b. Además, la unidad a fusionar se determina usando la relación entre este valor y el umbral.

Por ejemplo, tal como se muestra en la FIG. 17(a), en el caso en el que a/b es mayor que $c1$ ($a/b > c1$), la unidad a fusionar de X1 se determina como A. Por contraposición, tal como se muestra en la FIG. 17(b), en el caso en el que a/b es menor que $c2$ ($a/b < c2$), la unidad a fusionar de X1 se determina como B. En estos dos casos, no es necesario transmitir información que indique cuál de A y B se fusiona. Tal como se muestra en la FIG. 17(c), en el caso en el que a/b está entre $c1$ y $c2$ ($c2 \leq a/b \leq c1$), se transmite información aparte que indica cuál de A y B es la unidad a fusionar.

Las realizaciones ejemplificativas antes descritas son combinaciones de elementos y características de la presente invención. Los elementos o características se pueden considerar selectivos a no ser que se mencione lo contrario. Cada elemento o característica se puede llevar a la práctica sin combinarse con otros elementos o características. Además, una realización de la presente invención se puede constituir combinando partes de los elementos y/o características. Los órdenes de operaciones que se han descrito en realizaciones de la presente invención se pueden reorganizar. Algunas construcciones de una realización cualquiera se pueden incluir en otra realización, y se pueden sustituir con construcciones correspondientes de otra realización.

Un método de decodificación/codificación, en el que se aplica la presente invención, se configura con un programa para su ejecución en ordenador y, a continuación, se almacena en un soporte de grabación legible por ordenador. Además, se pueden almacenar, en soporte de grabación legible por ordenador, datos multimedia que presenten una estructura de datos de la presente invención. Los soportes de grabación legibles por ordenador incluyen todos los tipos de dispositivos de almacenamiento para almacenar datos que pueden ser leídos por un sistema de ordenador. Los soportes de grabación legibles por ordenador incluyen ROM, RAM, CD-ROM, cintas magnéticas, discos flexibles, dispositivos de almacenamiento óptico de datos, etcétera, y también incluye un dispositivo implementado con ondas portadoras (por ejemplo, transmisión por internet). Además, un flujo continuo de bits generado por el método de codificación se almacena en un soporte de grabación legible por ordenador o se transmite a través de una red de comunicaciones por cable/inalámbrica.

Diversas realizaciones descritas en la presente se pueden implementar en un soporte legible por ordenador usando, por ejemplo, *software* de ordenador, *hardware*, o alguna combinación de los mismos. Para una implementación en *hardware*, las realizaciones descritas en la presente se pueden implementar en uno o más circuitos integrados de

aplicación específica (ASICs), procesadores de señal digital (DSPs), dispositivos de procesamiento de señales digitales (DSPDs), dispositivos de lógica programable (PLDs), matrices de puertas programables in situ (FPGAs), procesadores, controladores, microcontroladores, microprocesadores, otras unidades electrónicas diseñadas para llevar a cabo las funciones descritas en la presente, o una combinación selectiva de los mismos. En algunos casos, dichas realizaciones se implementan mediante controlador.

Para una implementación en *software*, las realizaciones descritas en la presente se pueden implementar con módulos de *software* independientes, tales como procedimientos y funciones, cada uno de los cuales lleva a cabo una o más de las funciones y operaciones descritas en este documento. Los códigos de *software* se pueden implementar con una aplicación de *software* escrita en cualquier lenguaje de programación adecuado y se pueden almacenar en memoria, y pueden ser ejecutados por un controlador.

Se pondrá de manifiesto para aquellos versados en la materia que, en la presente invención, pueden aplicarse diversas modificaciones y variaciones sin desviarse con respecto al alcance de la invención. De este modo, se pretende que la presente invención abarque las modificaciones y variaciones de la misma siempre que estas se sitúen dentro del alcance de las reivindicaciones adjuntas y sus equivalentes.

Aplicabilidad industrial

La presente invención es aplicable a la codificación o decodificación de una señal de vídeo.

Resumiendo, el presente documento da a conocer un método y un aparato para codificar o decodificar una señal de vídeo. El método para procesar una señal de vídeo usa un modo de fusión en el que, en lugar de transmitir información de predicción sobre la presente unidad, se usa información de predicción sobre una unidad vecina, con el fin de mejorar la eficiencia de la codificación. En este caso, se determina el número de unidades candidatas disponibles para la fusión entre las unidades en una posición predeterminada, y se obtiene información para el modo de fusión sobre la base del número de las unidades candidatas disponibles para la fusión. La unidad a fusionar se determina usando la información correspondiente al modo de fusión, y se obtiene información de predicción sobre la unidad a fusionar. El valor de predicción para la unidad presente se obtiene usando la información de predicción sobre la unidad a fusionar, y la unidad presente se restaura usando el valor de predicción obtenido.

Para resumir, la materia en cuestión de los siguientes párrafos que se numeran para permitir referencias pertenece a la exposición del presente documento:

[Párrafo 1] Un método para procesar una señal de vídeo, comprendiendo el método: determinar un número de unidades candidatas a la fusión disponibles; obtener una bandera de fusión que indica si una unidad en curso usa un modo de fusión en caso de que el número de las unidades candidatas a la fusión disponibles sea mayor de 0; obtener información de fusión de la unidad en curso si la bandera de fusión indica que el bloque en curso es un modo de fusión y el número de las unidades candidatas a la fusión disponibles es mayor de 1; determinar una unidad a fusionar usando la información de fusión; obtener información de predicción de la unidad a fusionar; obtener un valor de predicción de píxel de la unidad en curso usando la información de predicción de la unidad a fusionar; y restaurar un valor de píxel de la unidad en curso usando el valor de predicción de píxel de la unidad en curso.

[Párrafo 2] El método según el párrafo 1, en el que la determinación determina un número de unidades disponibles entre una pluralidad de unidades candidatas que se seleccionan basándose en una posición de la unidad en curso.

[Párrafo 3] El método según el párrafo 2, en el que las unidades candidatas incluyen por lo menos una de un primer grupo compuesto por unidades adyacentes a un contorno superior de la unidad en curso, un segundo grupo compuesto por unidades adyacentes a un contorno izquierdo de la unidad en curso, un tercer grupo compuesto por unidades adyacentes a una esquina de la unidad en curso, y un cuarto grupo compuesto por unidades ubicadas en otra imagen que no incluye la unidad en curso.

[Párrafo 4] El método según el párrafo 3, que comprende, además: determinar cada una de las unidades entre los grupos como unidades candidatas a la fusión, en donde las unidades consideran una longitud del contorno adyacente a la unidad en curso, áreas de las unidades, o vectores de movimiento de las unidades.

[Párrafo 5] El método según el párrafo 2, en el que la determinación cuenta un número de unidades, que se han codificado con un modo Inter y tienen información de predicción, y se han decodificado por delante de la unidad en curso, entre las unidades candidatas.

[Párrafo 6] El método según el párrafo 1, en el que la información de fusión es una bandera de sentido de fusión que indica si una unidad vecina que se encuentra en un cierto sentido con respecto a la unidad en curso es la unidad a fusionar en caso de que el número de los candidatos a la fusión sea 2, y la información de fusión es un índice que indica la unidad a fusionar en caso de que el número de los candidatos a la fusión sea mayor de 2.

[Párrafo 7] El método según el párrafo 1, en el que la información de predicción incluye un índice e información de vector de movimiento de una imagen de referencia.

REIVINDICACIONES

1. Método para descodificar un flujo continuo de bits correspondiente a una señal de vídeo que comprende por lo menos un bloque de codificación, con un aparato de descodificación, comprendiendo el método:
5
obtener, con el aparato de descodificación, una pluralidad de bloques de predicción que se segmentan a partir del bloque de codificación;
10
obtener, con el aparato de descodificación, información de bandera de fusión de un bloque de predicción en curso de la pluralidad de bloques de predicción del flujo continuo de bits, indicando la información de bandera de fusión si el bloque de predicción en curso está codificado en un modo de fusión;
15
obtener, con el aparato de descodificación, información de índice de fusión del bloque de predicción en curso del flujo continuo de bits, especificando la información de índice de fusión un candidato a la fusión a fusionar con el bloque de predicción en curso de entre candidatos a la fusión que incluyen bloques vecinos espaciales y temporales del bloque de predicción en curso;
20
inferir, con el aparato de descodificación, información de predicción del bloque de predicción en curso basándose en los candidatos a la fusión y la información de índice de fusión, de manera que la información de predicción incluye información de vector de movimiento e información de índice de referencia; y
descodificar, con el aparato de descodificación, el bloque de predicción en curso basándose en la información de predicción del bloque de predicción en curso,
25
en donde, cuando el bloque de codificación 2Nx2N se segmenta en un bloque de predicción 2NxN superior y un bloque de predicción 2NxN inferior y el bloque de predicción en curso es el bloque de predicción 2NxN inferior del bloque de codificación, el bloque de predicción 2NxN superior del bloque de codificación se excluye de los candidatos a la fusión.
- 30 2. Método de la reivindicación 1, en el que cada uno de los bloques vecinos espaciales y temporales se codifica en un modo de predicción inter.
3. Método de la reivindicación 1 ó la reivindicación 2, en el que, cuando el bloque de codificación 2Nx2N se segmenta en un bloque de predicción Nx2N izquierdo y un bloque de predicción Nx2N derecho y el bloque de predicción en curso es el bloque de predicción Nx2N derecho del bloque de codificación, el bloque de predicción Nx2N izquierdo del bloque de codificación se excluye de los candidatos a la fusión.
35
4. Método de una cualquiera de las reivindicaciones 1 a 3, en el que los bloques vecinos espaciales incluyen un bloque de predicción vecino izquierdo, un bloque de predicción vecino superior izquierdo, un bloque de predicción vecino superior derecho o un bloque de predicción vecino inferior izquierdo.
40
5. Método de una cualquiera de las reivindicaciones 1 a 4,
45
en el que los bloques vecinos espaciales se corresponden con bloques de predicción vecinos adyacentes al bloque de predicción en curso dentro de una representación visual que incluye el bloque de predicción en curso.
6. Aparato configurado para llevar a cabo el método de una cualquiera de las reivindicaciones 1 a 5.
7. Método para codificar un flujo continuo de bits correspondiente a una señal de vídeo que comprende por lo menos un bloque de codificación, con un aparato de codificación, comprendiendo el método:
50
obtener, con el aparato de codificación, una pluralidad de bloques de predicción que se segmentan a partir del bloque de codificación;
55
codificar, con el aparato de codificación, información de bandera de fusión de un bloque de predicción en curso de la pluralidad de bloques de predicción en el flujo continuo de bits, indicando la información de bandera de fusión si el bloque de predicción en curso está codificado en un modo de fusión;
60
codificar, con el aparato de codificación, información de índice de fusión del bloque de predicción en curso en el flujo continuo de bits, especificando la información de índice de fusión un candidato a la fusión a fusionar con el bloque de predicción en curso de entre candidatos a la fusión que incluyen bloques vecinos espaciales y temporales del bloque de predicción en curso;

inferir, con el aparato de codificación, información de predicción del bloque de predicción en curso basándose en los candidatos a la fusión y la información de índice de fusión, de manera que la información de predicción incluye información de vector de movimiento e información de índice de referencia; y

5 codificar, con el aparato de codificación, el bloque de predicción en curso basándose en la información de predicción del bloque de predicción en curso,

10 en donde, cuando el bloque de codificación $2N \times 2N$ se segmenta en un bloque de predicción $2N \times N$ superior y un bloque de predicción $2N \times N$ inferior y el bloque de predicción en curso es el bloque de predicción $2N \times N$ inferior del bloque de codificación, el bloque de predicción $2N \times N$ superior del bloque de codificación se excluye de los candidatos a la fusión.

15 8. Soporte de grabación legible por ordenador en el cual está almacenado el flujo continuo de bits generado con el método de la reivindicación 7.

9. Aparato configurado para llevar a cabo el método de la reivindicación 7.

FIG. 1

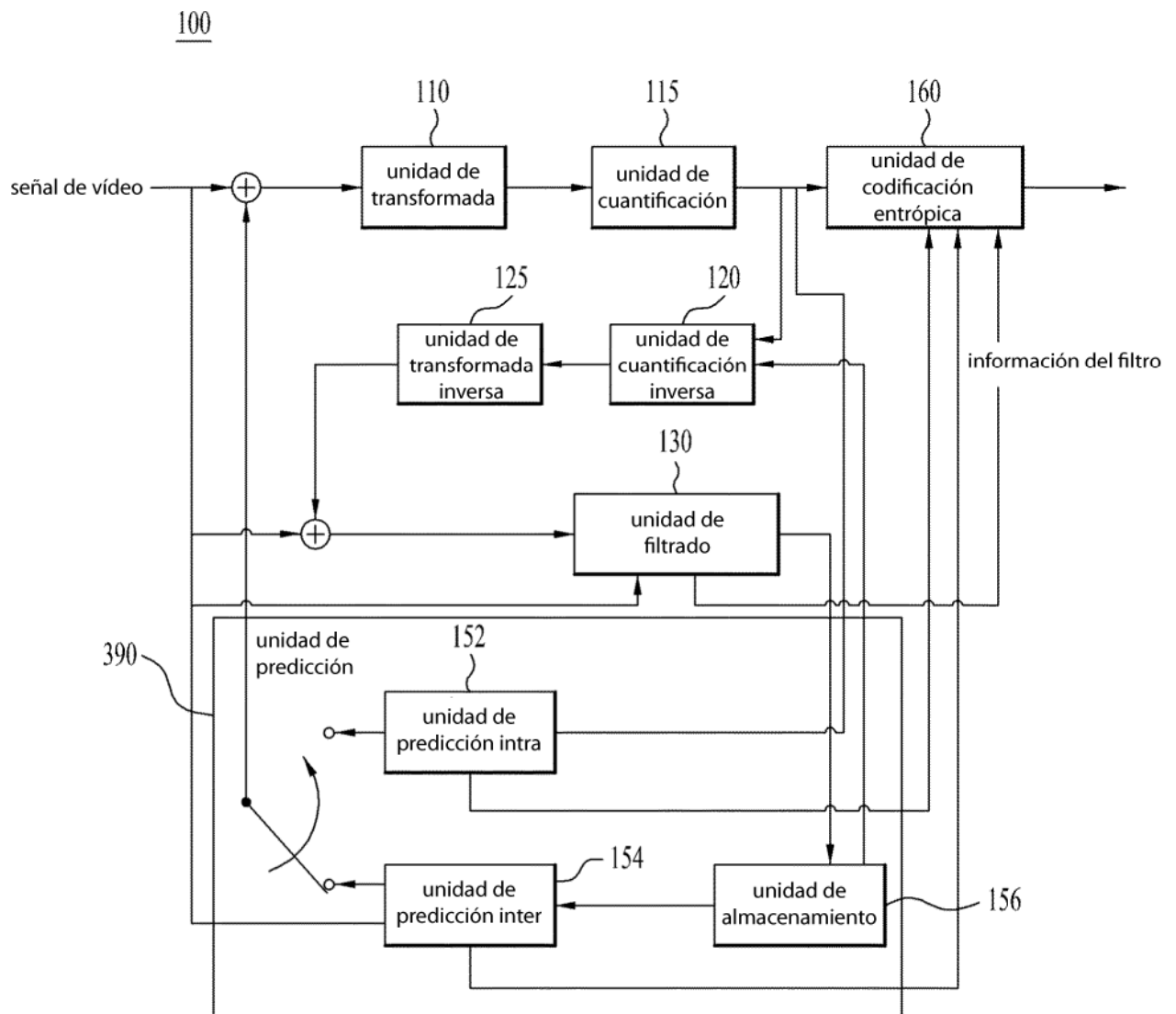


FIG. 2

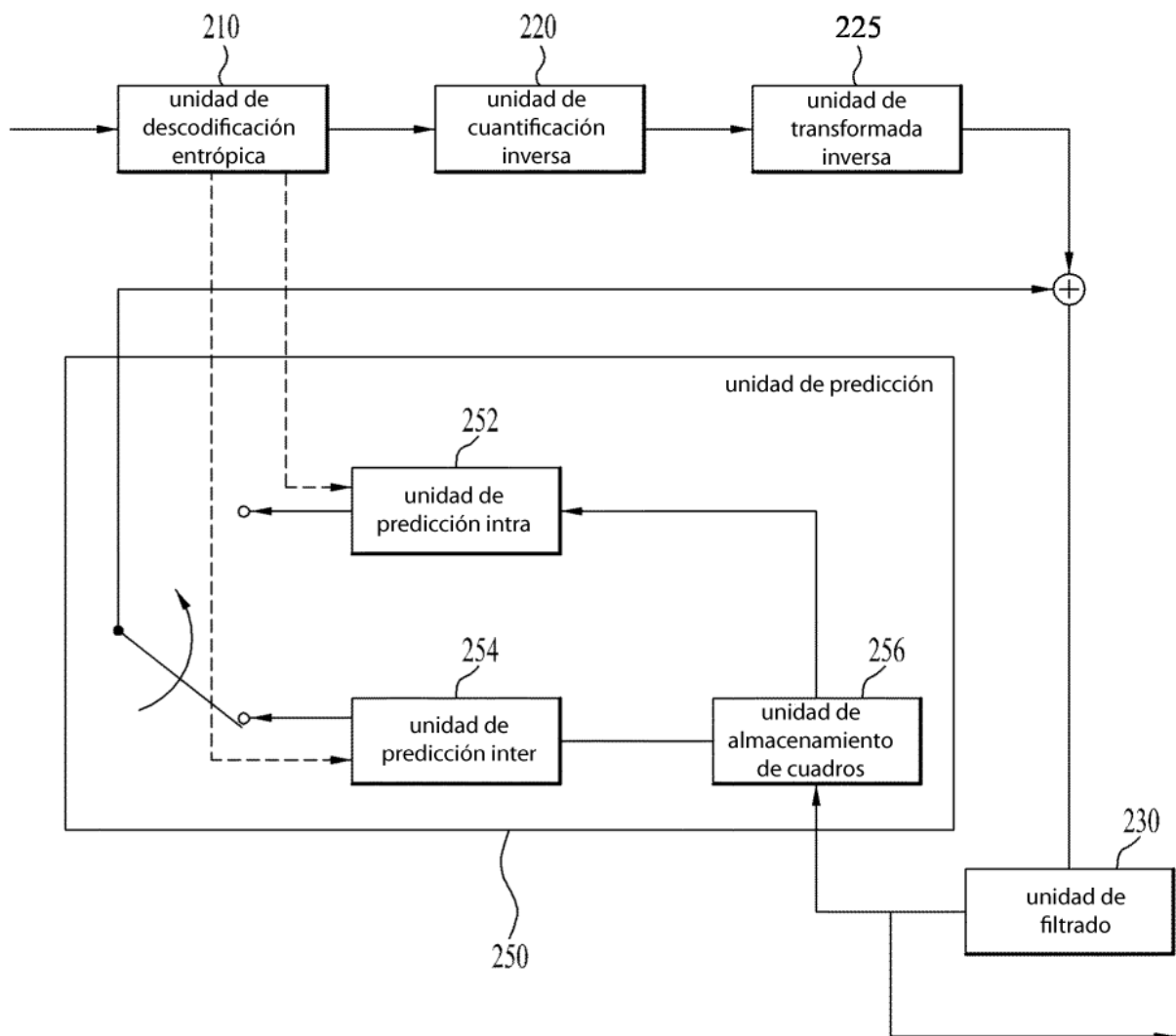


FIG. 3

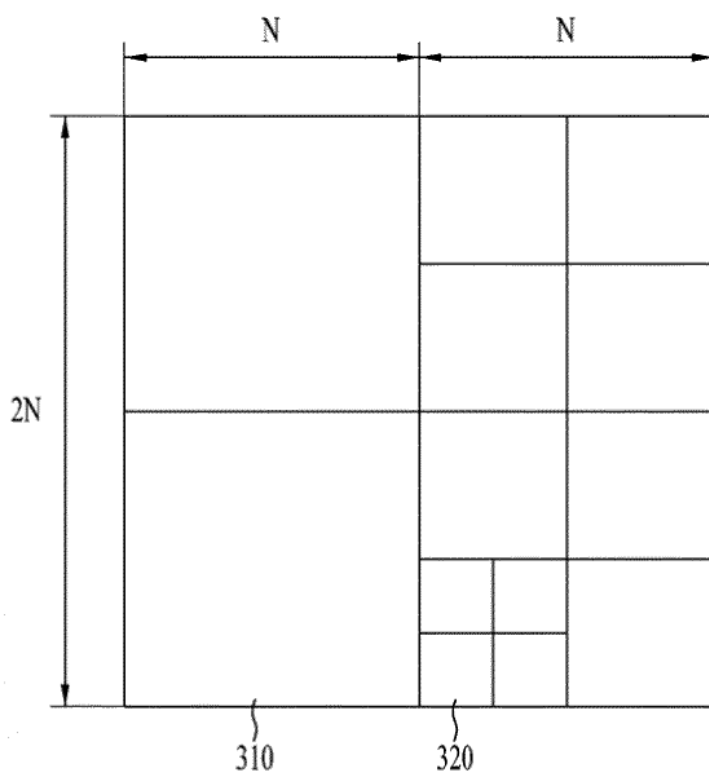


FIG. 4

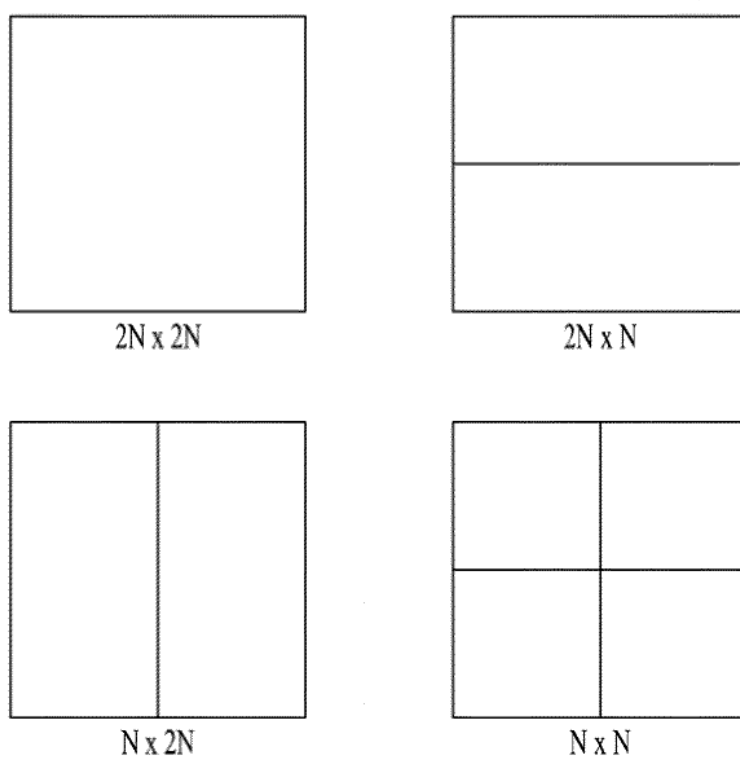


FIG. 5

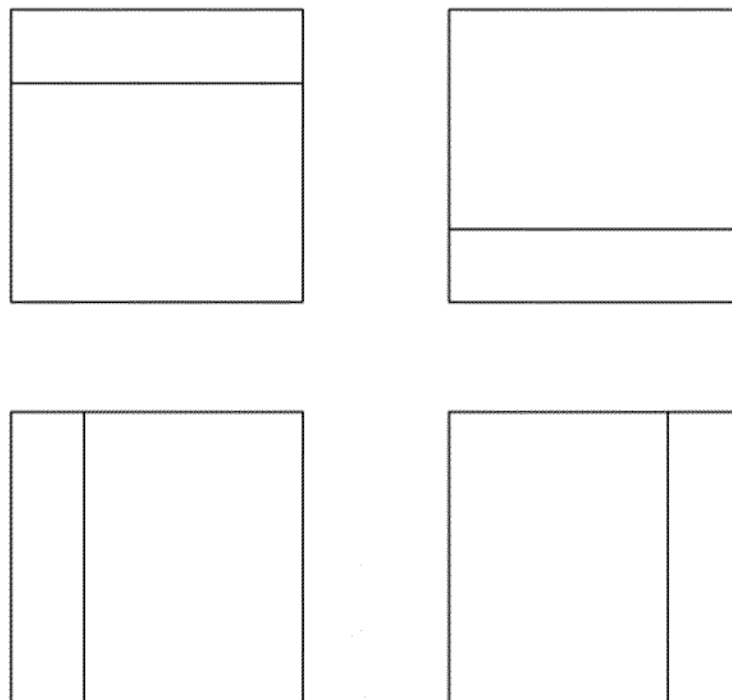


FIG. 6

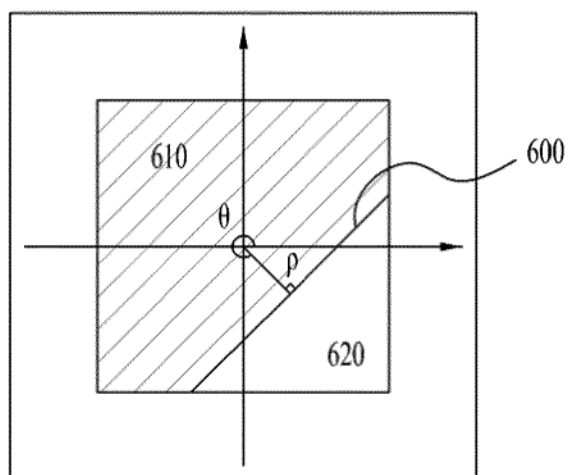


FIG. 7

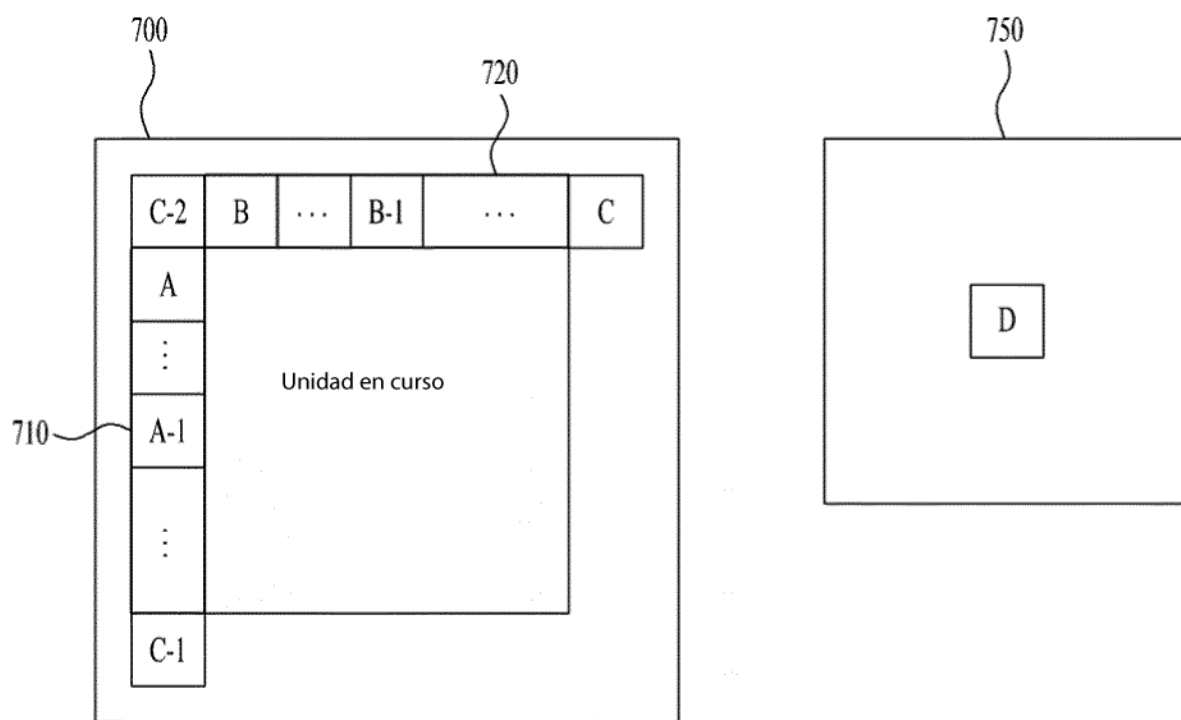


FIG. 8

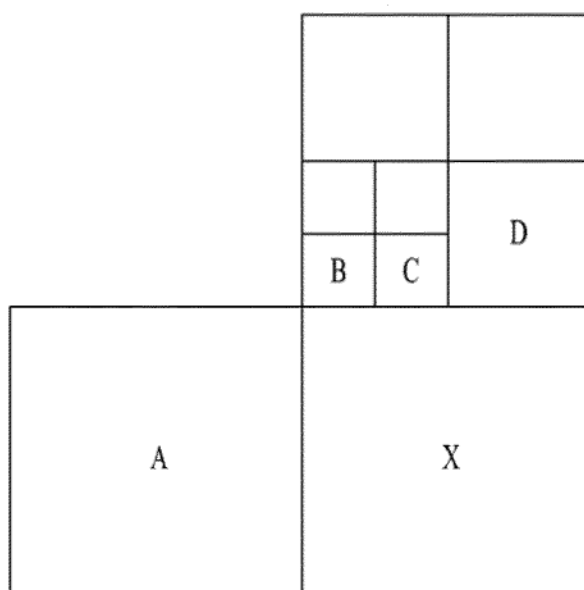


FIG. 9

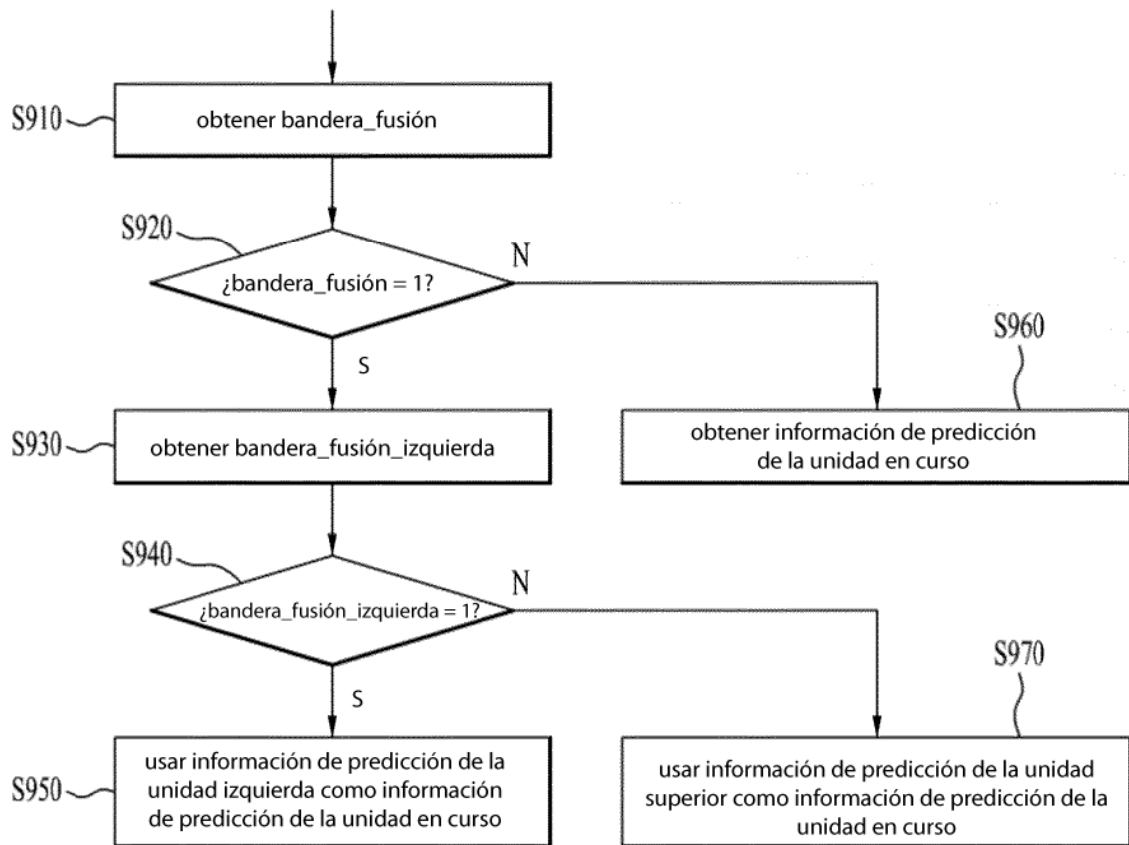


FIG. 10

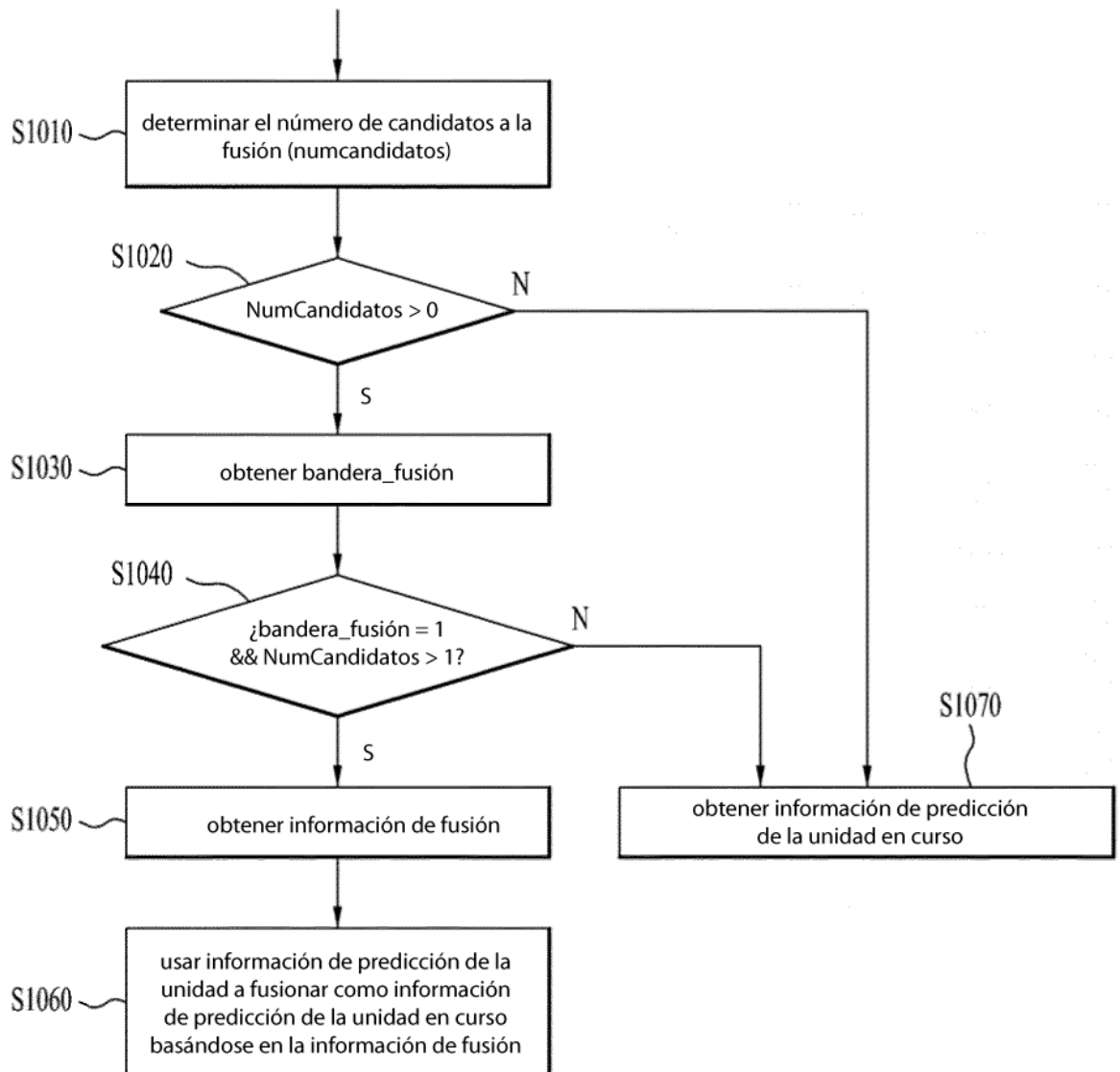


FIG. 11

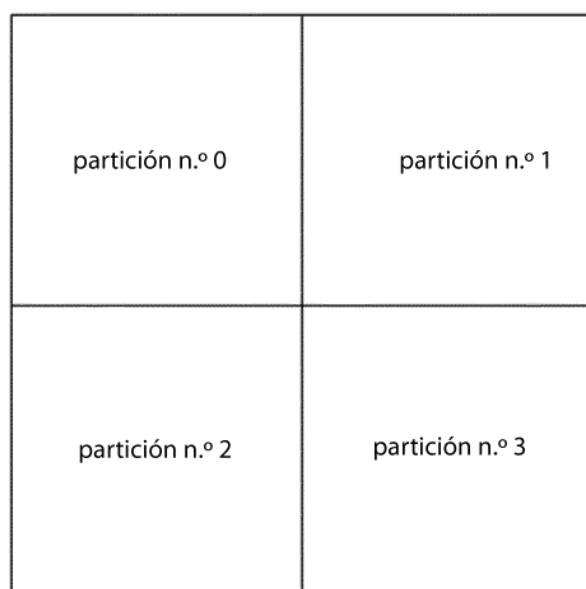


FIG. 12

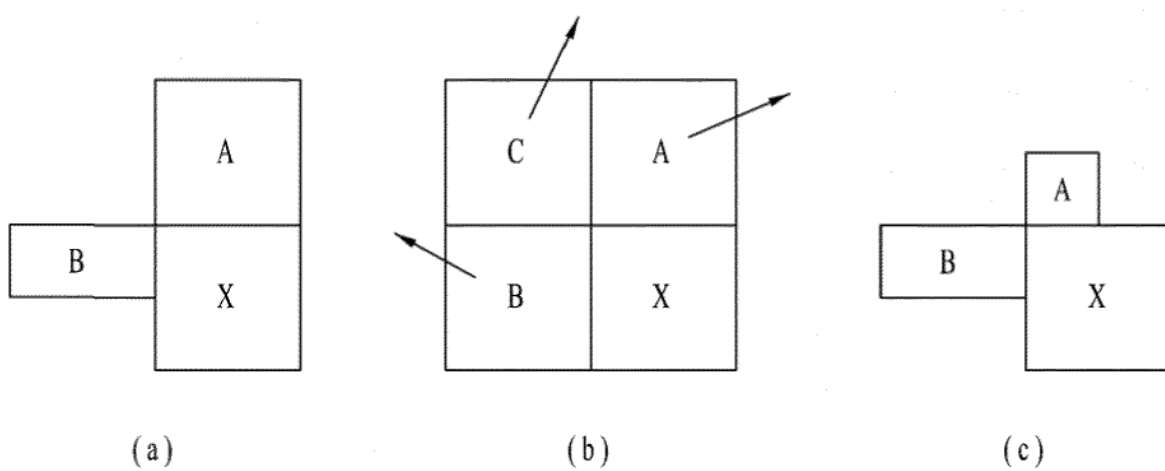


FIG. 13

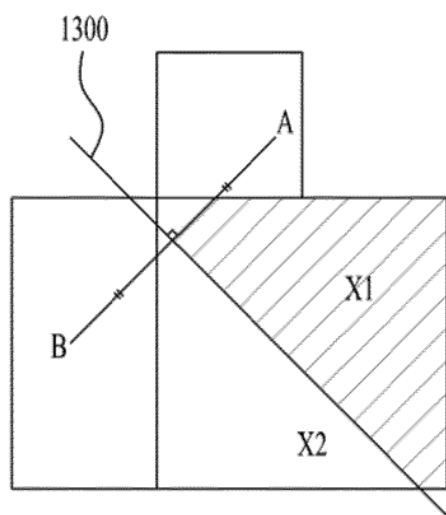


FIG. 14

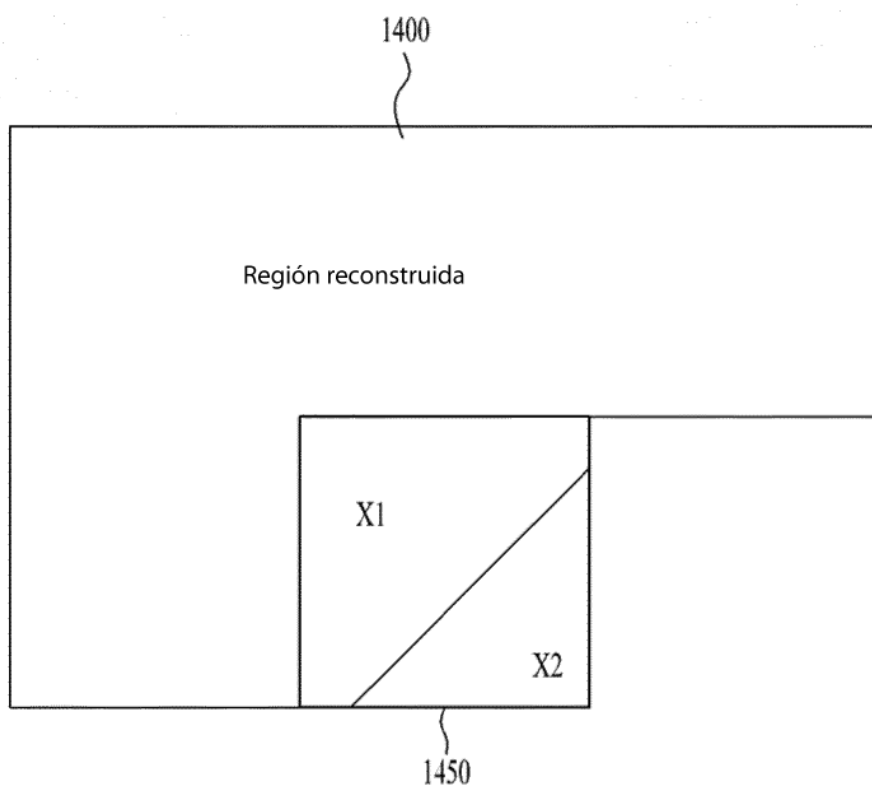
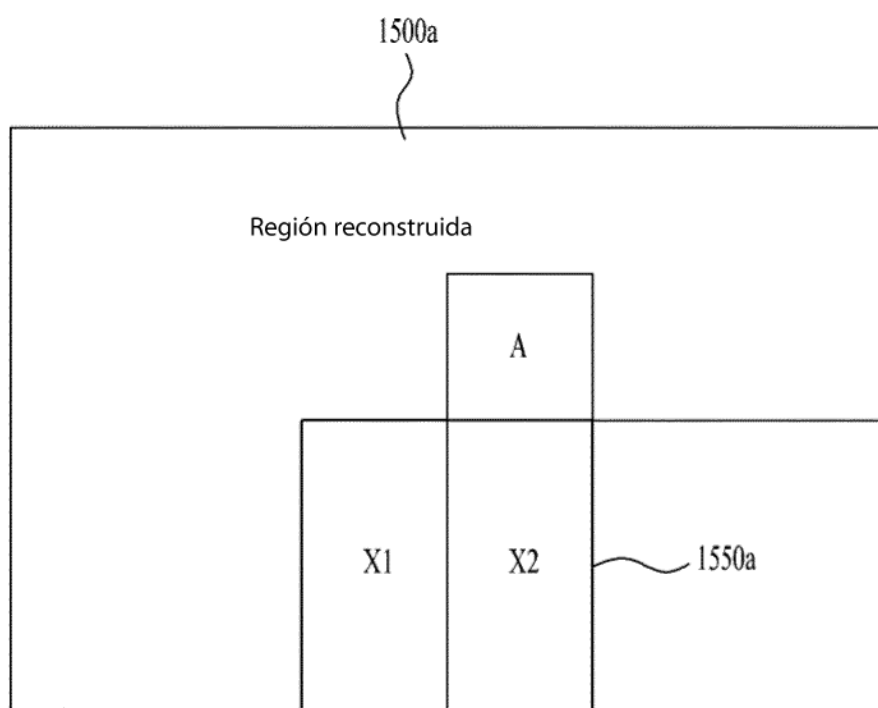
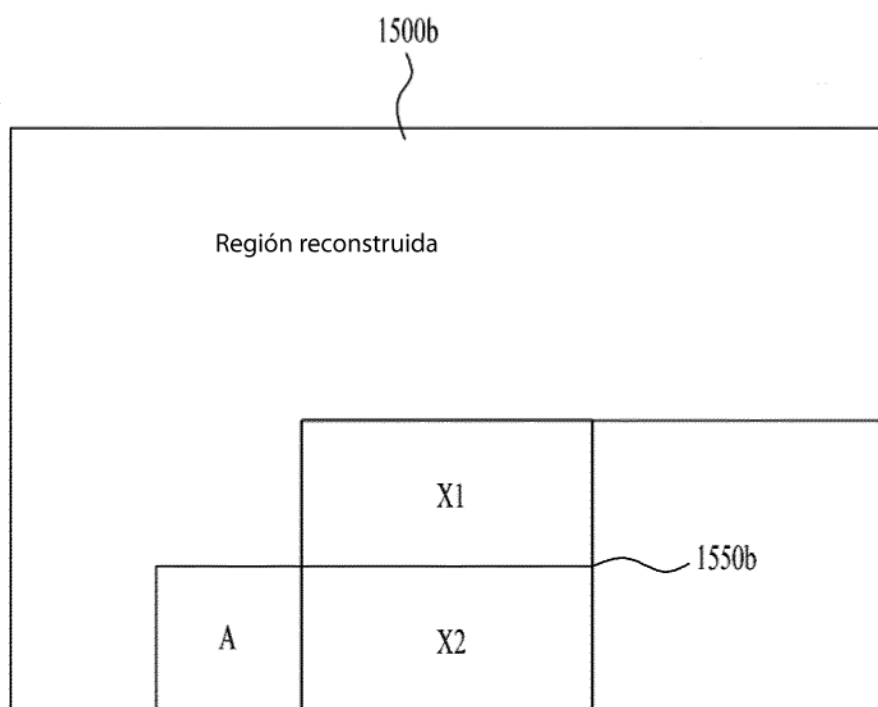


FIG. 15

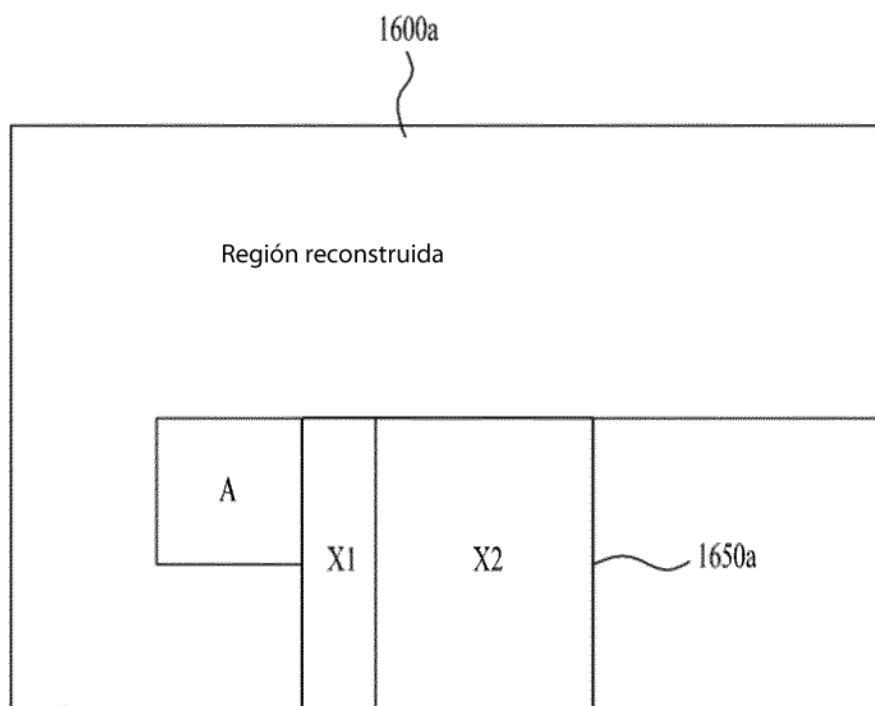


(a)

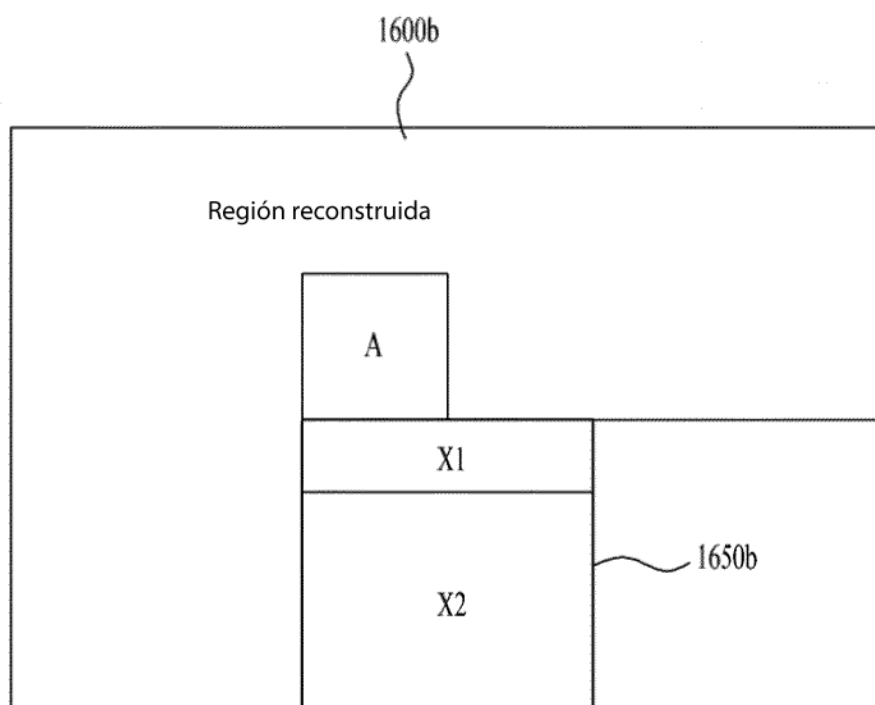


(b)

FIG. 16

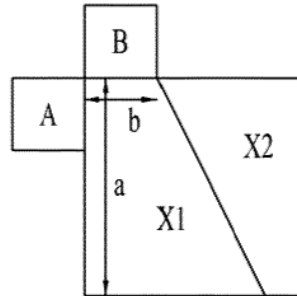


(a)

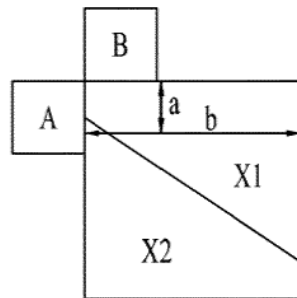


(b)

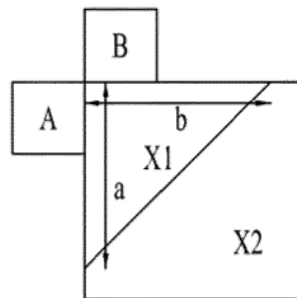
FIG. 17



(a)



(b)



(c)