

19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 959 874**

51 Int. Cl.:

H04N 19/463 (2014.01)
H04N 19/65 (2014.01)
H04N 19/70 (2014.01)
H04N 19/105 (2014.01)
H04N 19/176 (2014.01)
H04N 19/61 (2014.01)
H04N 19/513 (2014.01)
H04N 19/00 (2014.01)
H04N 19/52 (2014.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

96 Fecha de presentación y número de la solicitud europea: **21.05.2012** **E 20189030 (8)**

97 Fecha y número de publicación de la concesión europea: **10.05.2023** **EP 3751854**

54 Título: **Procedimiento de codificación de imágenes, aparato de codificación de imágenes, procedimiento de decodificación de imágenes, aparato de decodificación de imágenes**

30 Prioridad:

24.05.2011 US 201161489416 P

45 Fecha de publicación y mención en BOPI de la traducción de la patente:
28.02.2024

73 Titular/es:

**SUN PATENT TRUST (100.0%)
437 Madison Avenue, 35th Floor
New York, NY 10022, US**

72 Inventor/es:

**SUGIO, TOSHIYASU;
NISHI, TAKAHIRO;
SHIBAHARA, YOUJI;
TANIKAWA, KYOKO;
SASAI, HISAO y
MATSUNOBU, TORU**

74 Agente/Representante:

MILTENYI, Peter

ES 2 959 874 T3

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

DESCRIPCIÓN

Procedimiento de codificación de imágenes, aparato de codificación de imágenes, procedimiento de decodificación de imágenes, aparato de decodificación de imágenes

Campo técnico

La presente invención se define en las reivindicaciones adjuntas.

Técnica anterior

- Generalmente, en el procesamiento de codificación de una imagen en movimiento, la cantidad de información se reduce mediante compresión para la cual se utiliza la redundancia de una imagen en movimiento en la dirección espacial y la dirección temporal. Generalmente, la conversión a un dominio de frecuencia se realiza como un procedimiento en el que se utiliza la redundancia en la dirección espacial, y la codificación que usa predicción entre imágenes (la predicción se denomina inter predicción en lo sucesivo) se realiza como un procedimiento de compresión en el que se utiliza la redundancia en la dirección temporal. En la codificación de inter predicción, una imagen actual es codificada utilizando, como una imagen de referencia, una imagen codificada que precede o sucede a la imagen actual en el orden de tiempo de visualización. Posteriormente, se deriva un vector de movimiento realizando una estimación de movimiento en la imagen actual con referencia a la imagen de referencia. Entonces, se elimina la redundancia en la dirección temporal utilizando una diferencia calculada entre los datos de imagen de la imagen actual y los datos de imagen de predicción que se obtienen mediante la compensación de movimiento en función del vector de movimiento derivado (véase NPL 1, por ejemplo). En este punto, en la estimación de movimiento, se calculan valores de diferencia entre los bloques actuales en la imagen actual y los bloques en la imagen de referencia, y se determina un bloque que tiene el menor valor de diferencia en la imagen de referencia como un bloque de referencia. A continuación, se estima un vector de movimiento a partir del bloque actual y el bloque de referencia.
- La referencia NPL 3 se refiere a la predicción temporal de los vectores de movimiento. Para evitar un error de análisis sintáctico cuando se elimina un vector de movimiento redundante de una lista de vectores de movimiento candidatos, o bien se fuerza la disponibilidad de un predictor temporal o bien se sustituye el candidato redundante por un candidato con valores diferentes. A continuación, el índice del candidato seleccionado se codifica en base al número variable de candidatos dentro de la lista de candidatos

Lista de referencias

Literatura no relacionada con patentes:

NPL 1: ITU-T Recommendation H.264 "Advanced video coding for generic audiovisual services", marzo de 2010

NPL 2: JCT-VC, "WD3: Working Draft 3 of High-Efficiency Video Coding", JCTVC-E603, marzo de 2011

- NPL 3: Guillaume Laroche *et al.*, "Robust solution for the AMVP parsing issue", JCTVC-E219, 96. MPEG meeting Geneva; (Motion Picture Expert Group or ISO/IEC JTC1/SC29/WG11), marzo de 2011

Resumen de la invención

Problema técnico:

Todavía es deseable mejorar la resistencia a errores en la codificación y decodificación de imágenes en las que se usa la inter predicción, más allá de la técnica convencional descrita anteriormente.

- En vista de esto, el objeto de la presente invención es proporcionar un procedimiento de codificación y decodificación de imágenes y un aparato de codificación y decodificación de imágenes con los que se mejore la resistencia a errores en la codificación de imágenes y la decodificación de imágenes usando la inter predicción.

Esto se consigue mediante las características de las reivindicaciones independientes. Las realizaciones preferidas son objeto de las reivindicaciones dependientes.

- Se debe observar que estos aspectos generales o específicos pueden implementarse como un sistema, un procedimiento, un circuito integrado, un programa informático, un medio de grabación legible por ordenador tal como una memoria de solo lectura de disco compacto (CD-ROM), o como cualquier combinación de un sistema, un procedimiento, un circuito integrado, un programa informático y un medio de grabación legible por ordenador.

Efectos ventajosos de la invención:

- De acuerdo con un aspecto de la presente invención, se puede mejorar la resistencia a errores en la codificación y decodificación de imágenes usando la inter predicción.

Breve descripción de los dibujos

La figura 1A es un diagrama para ilustrar un ejemplo de lista de imágenes de referencia para una imagen B.

La figura 1B es un diagrama para ilustrar un ejemplo de lista de imágenes de referencia de una dirección de predicción 0 para una imagen B.

- 5 La figura 1C es un diagrama para ilustrar un ejemplo de lista de imágenes de referencia de una dirección de predicción 1 para una imagen B.

La figura 2 es un diagrama para ilustrar vectores de movimiento para usar en el modo de predicción de vector de movimiento temporal.

La figura 3 muestra un ejemplo de vector de movimiento de un bloque vecino para usar en el modo de fusión.

La figura 4 es un diagrama para ilustrar un ejemplo de lista de candidatos a bloques de fusión.

- 10 La figura 5 muestra una relación entre el tamaño de una lista de candidatos a bloques de fusión y secuencias de bits asignadas a índices de candidatos a bloques de fusión.

La figura 6 es un diagrama de flujo que muestra un ejemplo de un proceso para codificación cuando se usa el modo de fusión.

- 15 La figura 7 muestra un ejemplo de configuración de un aparato de codificación de imágenes que codifica imágenes usando el modo de fusión.

La figura 8 es un diagrama de flujo que muestra un ejemplo de un proceso para decodificar usando el modo de fusión.

La figura 9 muestra un ejemplo de configuración de un aparato de decodificación de imágenes que decodifica imágenes codificadas usando el modo de fusión.

- 20 La figura 10 muestra una sintaxis para la asociación de índices de candidatos a bloques de fusión a una secuencia de bits (*bitstream*).

La figura 11 es un diagrama de bloques que muestra una configuración de un aparato de codificación de imágenes de acuerdo con una Realización 1.

- 25 La figura 12 es un diagrama de flujo que muestra operaciones de procesamiento del aparato de codificación de imágenes de acuerdo con la Realización 1.

La figura 13 muestra un ejemplo de lista de candidatos a bloques de fusión de acuerdo con la Realización 1.

La figura 14 es un diagrama de flujo que ilustra un proceso para calcular candidatos a bloques de fusión y el tamaño de una lista de candidatos a bloques de fusión de acuerdo con la Realización 1.

- 30 La figura 15 es un diagrama de flujo que ilustra un proceso para actualizar un número total de candidatos utilizables para la fusión de acuerdo con la Realización 1.

La figura 16 es un diagrama de flujo que ilustra un proceso para agregar un nuevo candidato de acuerdo con la Realización 1.

La figura 17 es un diagrama de flujo que ilustra un proceso para seleccionar un candidato a bloque de fusión de acuerdo con la Realización 1.

- 35 La figura 18 es un diagrama de bloques que muestra una configuración de un aparato de codificación de imágenes de acuerdo con una Realización 2.

La figura 19 es un diagrama de flujo que muestra operaciones de procesamiento del aparato de codificación de imágenes de acuerdo con la Realización 2.

- 40 La figura 20 es un diagrama de bloques que muestra una configuración de un aparato de decodificación de imágenes de acuerdo con una Realización 3.

La figura 21 es un diagrama de flujo que muestra operaciones de procesamiento del aparato de decodificación de imágenes de acuerdo con la Realización 3.

La figura 22 es un diagrama de flujo que ilustra un proceso para establecer el tamaño de una lista de candidatos a bloques de fusión de acuerdo con la Realización 3.

- 45 La figura 23 es un diagrama de flujo que ilustra un proceso para calcular un candidato a bloque de fusión según la Realización 3.

La figura 24 muestra una sintaxis para la asociación de índices de candidatos a bloques de fusión a una secuencia de bits.

La figura 25 muestra una sintaxis de ejemplo en el caso en el que el tamaño de una lista de candidatos a bloques de fusión se establece en el valor máximo del número total de candidatos a bloques de fusión.

- 5 La figura 26 es un diagrama de bloques que muestra una configuración de un aparato de decodificación de imágenes de acuerdo con una Realización 4.

La figura 27 es un diagrama de flujo que muestra operaciones de procesamiento del aparato de decodificación de imágenes de acuerdo con la Realización 4.

- 10 La figura 28 muestra una configuración general de un sistema proveedor de contenidos para implementar servicios de distribución de contenidos.

La figura 29 muestra una configuración general de un sistema de difusión digital.

La figura 30 muestra un diagrama de bloques que ilustra un ejemplo de una configuración de un televisor.

- 15 La figura 31 es un diagrama de bloques que ilustra un ejemplo de una configuración de una unidad de reproducción/grabación de información que lee y escribe información de y en un medio de grabación que es un disco óptico.

La figura 32 muestra un ejemplo de una configuración de un medio de grabación que es un disco óptico.

La figura 33A muestra un ejemplo de un teléfono móvil.

La figura 33B es un diagrama de bloques que muestra un ejemplo de una configuración de un teléfono móvil.

La figura 34 ilustra una estructura de datos multiplexados.

- 20 La figura 35 muestra esquemáticamente cómo cada flujo es multiplexado en datos multiplexados.

La figura 36 muestra cómo se almacena un flujo de video en un flujo de paquetes PES con más detalle.

La figura 37 muestra una estructura de paquetes TS y paquetes fuente en los datos multiplexados.

La figura 38 muestra una estructura de datos de una PMT (tabla de mapas de programas).

La figura 39 muestra una estructura interna de información de datos multiplexados.

- 25 La figura 40 muestra una estructura interna de información de atributos de flujo.

La figura 41 muestra unas etapas para identificar datos de video.

La figura 42 es un diagrama de bloques que muestra un ejemplo de una configuración de un circuito integrado para implementar el procedimiento de codificación de imágenes en movimiento y el procedimiento de decodificación de imágenes en movimiento de acuerdo con cada una de las realizaciones.

- 30 La figura 43 muestra una configuración para conmutar entre frecuencias de funcionamiento (*driving frequencies*).

La figura 44 muestra unas etapas para identificar datos de video y cambiar entre frecuencias de funcionamiento.

La figura 45 muestra un ejemplo de una tabla de búsqueda en la que los estándares de datos de video están asociados con frecuencias de funcionamiento.

- 35 La figura 46A es un diagrama que muestra un ejemplo de una configuración para compartir un módulo de una unidad de procesamiento de señal.

La figura 46B es un diagrama que muestra otro ejemplo de una configuración para compartir un módulo de la unidad de procesamiento de señal.

Descripción de realizaciones

(Base subyacente de formación de conocimiento de la presente invención)

- 40 En un esquema de codificación de imágenes en movimiento ya estandarizado, que se denomina H.264, se usan tres tipos de imágenes: imagen I, imagen P e imagen B, para reducir la cantidad de información por compresión.

La imagen I no se codifica por codificación de inter predicción. Específicamente, la imagen I se codifica por predicción dentro de la imagen (predicción que se denomina intra predicción en lo sucesivo). La imagen P se

codifica por codificación de inter predicción con referencia a una imagen codificada que precede o sucede a la imagen actual en el orden de tiempo de visualización. La imagen B se codifica por codificación de inter predicción con referencia a dos imágenes codificadas que preceden y suceden a la imagen actual en el orden de tiempo de visualización.

- 5 En la codificación de inter predicción, se genera una lista de imágenes de referencia para identificar una imagen de referencia. En una lista de imágenes de referencia, se asignan unos índices de imágenes de referencia a unas imágenes de referencia codificadas para que puedan ser referenciadas en la inter predicción. Por ejemplo, se generan dos listas de imágenes de referencia (L0, L1) para una imagen B porque puede ser codificada con referencia a dos imágenes.
- 10 La figura 1A es un diagrama para ilustrar un ejemplo de lista de imágenes de referencia para una imagen B. La figura 1B muestra un ejemplo de lista de imágenes de referencia 0 (L0) para una dirección de predicción 0 en una predicción bidireccional. En la lista de imágenes de referencia 0, se asigna el índice de imagen de referencia 0 que tiene un valor 0 a una imagen de referencia 0 con un número de orden de visualización 2. Se asigna el índice de imagen de referencia 0 que tiene un valor 1 a una imagen de referencia 1 con un número de orden de visualización 1. Se asigna el índice de imagen de referencia 0 que tiene un valor 2 a una imagen de referencia 2 con un número de orden de visualización 0. En otras palabras, cuanto más corta es la distancia temporal de una imagen de referencia con respecto a la imagen actual, más pequeño es el índice de imagen de referencia asignado a la imagen de referencia.
- 15 Por otro lado, la figura 1C muestra un ejemplo de lista de imágenes de referencia 1 (L1) para una dirección de predicción 1 en una predicción bidireccional. En la lista de imágenes de referencia 1, se asigna el índice de imagen de referencia 1 que tiene un valor 0 a una imagen de referencia 1 con un número de orden de visualización 1. Se asigna el índice de imagen de referencia 1 que tiene un valor 1 a una imagen de referencia 0 con un número de orden de visualización 2. Se asigna el índice de imagen de referencia 2 que tiene un valor 2 a una imagen de referencia 2 con un número de orden de visualización 0.
- 20 De esta manera, es posible asignar índices de imagen de referencia que tienen valores diferentes entre direcciones de predicción a una imagen de referencia (imágenes de referencia 0 y 1 en la figura 1A) o asignar el índice de imagen de referencia que tiene el mismo valor para ambas direcciones a una imagen de referencia (imagen de referencia 2 en la figura 1A).
- 25 En un procedimiento de codificación de imágenes en movimiento denominado H.264 (véase NPL 1), está disponible un modo de estimación de vectores de movimiento como un modo de codificación para la inter predicción de cada bloque actual en una imagen B. En el modo de estimación de vector de movimiento, se codifica un valor de diferencia entre los datos de imagen de un bloque actual y los datos de imagen de predicción y un vector de movimiento utilizado para generar los datos de imagen de predicción. Además, en el modo de estimación de vector de movimiento, se pueden realizar selectivamente la predicción bidireccional y la predicción unidireccional. En la predicción bidireccional, se genera una imagen de predicción con referencia a dos imágenes codificadas, una de las cuales precede a una imagen actual a codificar y la otra sucede a la imagen actual. En la predicción unidireccional, se genera una imagen de predicción con referencia a una imagen codificada que precede o sucede a una imagen actual a codificar.
- 30 Además, en el procedimiento de codificación de imágenes en movimiento denominado H.264, se puede seleccionar un modo de codificación denominado modo de predicción de vector de movimiento temporal para la derivación de un vector de movimiento en la codificación de una imagen B. A continuación se describirá utilizando la figura 2, el procedimiento de codificación de inter predicción realizado en el modo de predicción del vector de movimiento temporal.
- 35 La figura 2 es un diagrama para ilustrar vectores de movimiento para uso en el modo de predicción de vector de movimiento temporal. Específicamente, la figura 2 muestra un caso en el que se codifica un bloque a en una imagen B2 en el modo de predicción del vector de movimiento temporal.
- 40 En la codificación, se usa un vector de movimiento v_b que se ha usado para la codificación de un bloque b ubicado en la misma posición en una imagen P3, que es una imagen de referencia que sucede a la imagen B2, que la posición del bloque a en la imagen B2 (en este punto, el bloque b se denomina bloque co-ubicado del bloque a en lo sucesivo). El vector de movimiento v_b es un vector de movimiento utilizado para codificar el bloque b con referencia a la imagen P1.
- 45 Se obtienen dos bloques de referencia para el bloque a, a partir de una imagen de referencia siguiente (*forward reference picture*) y una imagen de referencia anterior (*backward reference picture*), es decir, una imagen P1 y una imagen P3 usando vectores de movimiento paralelos al vector de movimiento v_b . A continuación, se codifica el bloque a mediante predicción bidireccional en función de los dos bloques de referencia obtenidos. Específicamente, en la codificación del bloque a, se usa un vector de movimiento va_1 para referenciar la imagen P1, y se usa un vector de movimiento va_2 para referenciar la imagen P3.
- 50 Se obtienen dos bloques de referencia para el bloque a, a partir de una imagen de referencia siguiente (*forward reference picture*) y una imagen de referencia anterior (*backward reference picture*), es decir, una imagen P1 y una imagen P3 usando vectores de movimiento paralelos al vector de movimiento v_b . A continuación, se codifica el bloque a mediante predicción bidireccional en función de los dos bloques de referencia obtenidos. Específicamente, en la codificación del bloque a, se usa un vector de movimiento va_1 para referenciar la imagen P1, y se usa un vector de movimiento va_2 para referenciar la imagen P3.
- 55 Además, se discute un modo de fusión como un modo de inter predicción para la codificación de cada bloque actual

en una imagen B o una imagen P (ver NPL 2). En el modo de fusión, se codifica un bloque actual utilizando una dirección de predicción, un vector de movimiento y un índice de imagen de referencia que son duplicaciones de los utilizados para la codificación de un bloque vecino del bloque actual. En este momento, las duplicaciones del índice y otras del bloque vecino se asocian a una secuencia de bits de modo que la dirección de movimiento, el vector de movimiento y el índice de imagen de referencia que se han usado para la codificación pueden ser seleccionados en la decodificación. A continuación se proporciona un ejemplo concreto para esto con referencia a La figura 3.

La figura 3 muestra un ejemplo de vector de movimiento de un bloque vecino para usarlo en el modo de fusión. En la figura 3, un bloque vecino A es un bloque codificado ubicado a la izquierda inmediata de un bloque actual. Un bloque vecino B es un bloque codificado ubicado inmediatamente encima del bloque actual. Un bloque vecino C es un bloque codificado ubicado inmediatamente encima a la derecha del bloque actual. Un bloque vecino D es un bloque codificado ubicado inmediatamente debajo a la izquierda del bloque actual.

El bloque vecino A es un bloque codificado por predicción unidireccional en la dirección de predicción 0. El bloque vecino A tiene un vector de movimiento $MvL0_A$ que tiene dirección de predicción 0 como un vector de movimiento con respecto a una imagen de referencia indicada por un índice de imagen de referencia $RefL0_A$ de la dirección de predicción 0. En este punto, $MvL0$ indica un vector de movimiento que hace referencia a una imagen de referencia especificada en una lista de imágenes de referencia 0 ($L0$). $MvL1$ indica un vector de movimiento que hace referencia a una imagen de referencia especificada en una lista de imágenes de referencia 1 ($L1$).

El bloque vecino B es un bloque codificado por predicción unidireccional en la dirección de predicción 1. El bloque vecino B tiene un vector de movimiento $MvL1_B$ que tiene dirección de predicción 1 como un vector de movimiento con respecto a una imagen de referencia indicada por un índice de imagen de referencia $RefL1_B$ de la dirección de predicción 1.

El bloque vecino C es un bloque codificado por intra predicción.

El bloque vecino D es un bloque codificado por predicción unidireccional en la dirección de predicción 0. El bloque vecino D tiene un vector de movimiento $MvL0_D$ que tiene dirección de predicción 0 como un vector de movimiento con respecto a una imagen de referencia indicada por un índice de imagen de referencia $RefL0_D$ de la dirección de predicción 0.

En este caso, por ejemplo, se selecciona una combinación de una dirección de predicción, un vector de movimiento y un índice de imagen de referencia con la que se puede codificar el bloque actual con la máxima eficacia de codificación como una dirección de predicción, un vector de movimiento, y un índice de imagen de referencia del bloque actual a partir de las direcciones de predicción, vectores de movimiento e índices de imagen de referencia de los bloques vecinos A al D, y una dirección de predicción, un vector de movimiento y un índice de imagen de referencia que se calculan usando un bloque co-ubicado en el modo de predicción del vector de movimiento temporal. Entonces, se asocia un índice de candidato a bloque de fusión que indica el bloque seleccionado que tiene la dirección de predicción, el vector de movimiento y el índice de imagen de referencia a una secuencia de bits.

Por ejemplo, cuando se selecciona el bloque vecino A, el bloque actual se codifica utilizando el vector de movimiento $MvL0_A$ que tiene la dirección de predicción 0 y el índice de imagen de referencia $RefL0_A$. Entonces, solamente el índice de candidato a bloque de fusión que tiene un valor 0 que indica el uso del bloque vecino A como se muestra en la figura 4 es asociado a la secuencia de bits. De este modo, se reduce la cantidad de información en una dirección de predicción, un vector de movimiento y un índice de imagen de referencia.

Además, en el modo de fusión, un candidato que no puede usarse para la codificación (denominado candidato inutilizable para la fusión en lo sucesivo), y un candidato que tiene una combinación de una dirección de predicción, un vector de movimiento y un índice de imagen de referencia idéntica a una combinación de una dirección de predicción, un vector de movimiento y un índice de imagen de referencia de cualquier otro bloque de fusión (denominado candidato idéntico en lo sucesivo) son eliminados de entre los candidatos a bloque de fusión como se muestra en la figura 4.

De esta manera, se reduce el número total de candidatos a bloques de fusión, de modo que se puede reducir la cantidad de código asignado a los índices de candidatos a bloques de fusión. En este punto, "inutilizable para la fusión" significa (1) que el candidato a bloque de fusión ha sido codificado por intra predicción, (2) que el candidato a bloque de fusión está fuera del límite de una porción (*slice*) que incluye el bloque actual o del límite de una imagen que incluye el bloque actual, o (3) que todavía no se ha codificado el candidato a bloque de fusión.

En el ejemplo mostrado en la figura 4, el bloque vecino C es un bloque codificado por intra predicción. El candidato a bloque de fusión que tiene el índice de candidato a bloque de fusión 3 es por lo tanto un candidato inutilizable para la fusión y es eliminado de la lista de candidatos a bloque de fusión. El bloque vecino D es idéntico en dirección de predicción, vector de movimiento e índice de imagen de referencia al bloque vecino A. Por lo tanto, el candidato a bloque de fusión que tiene el índice de candidato a bloque de fusión 4 es eliminado de la lista de candidatos a bloque de fusión. Como resultado, el número total de candidatos a bloques de fusión es finalmente tres, y el tamaño de la lista de candidatos a bloques de fusión acaba siendo tres.

Los índices de candidatos a bloques de fusión son codificados mediante codificación de longitud variable asignando secuencias de bits de acuerdo con el tamaño de cada lista de candidatos a bloques de fusión como se muestra en la figura 5. Por lo tanto, en el modo de fusión, se reduce la cantidad de código cambiando secuencias de bits asignadas a los índices del modo de fusión de acuerdo con el tamaño de cada lista de candidatos a bloques de fusión.

La figura 6 es un diagrama de flujo que muestra un ejemplo de un proceso para codificación cuando se usa el modo de fusión. En la etapa S1001, se obtienen vectores de movimiento, índices de imágenes de referencia y direcciones de predicción de candidatos a bloques de fusión a partir de bloques vecinos y un bloque co-ubicado. En la etapa S1002, los candidatos idénticos y los candidatos inutilizables para la fusión son eliminados de entre los candidatos a bloques de fusión. En la etapa S1003, se establece el número total de candidatos a bloques de fusión después de la eliminación como el tamaño de la lista de candidatos a bloques de fusión. En la etapa S1004, se determina el índice de candidato a bloque de fusión a utilizar para la codificación del bloque actual. En la etapa S1005, se codifica el índice de candidato a bloque de fusión determinado realizando una codificación de longitud variable en secuencia de bits de acuerdo con el tamaño de la lista de candidatos a bloques de fusión.

La figura 7 muestra un ejemplo de configuración de un aparato de codificación de imágenes 1000 que codifica imágenes usando el modo de fusión. El aparato de codificación de imágenes 1000 incluye un sustractor 1001, una unidad de transformación ortogonal 1002, una unidad de cuantificación 1003, una unidad de cuantificación inversa 1004, una unidad de transformación ortogonal inversa 1005, un agregador 1006, una memoria de bloques 1007, una memoria de fotogramas (*frames*) 1008, una unidad de intra predicción 1009, una unidad de inter predicción 1010, una unidad de control de inter predicción 1011, una unidad de determinación del tipo de imagen 1012, un conmutador 1013, una unidad de cálculo de candidatos a bloque de fusión 1014, una memoria *colPic* 1015 y una unidad de codificación de longitud variable 1016.

En la figura 7, la unidad de cálculo de candidatos a bloque de fusión 1014 calcula candidatos a bloques de fusión. Entonces, la unidad de cálculo de candidatos a bloque de fusión 1014 transmite el número total de candidatos a bloque de fusión derivados a la unidad de codificación de longitud variable 1016. La unidad de codificación de longitud variable 1016 establece el número total de candidatos a bloque de fusión como el tamaño de la lista de candidatos a bloques de fusión que es un parámetro de codificación. Entonces, la unidad de codificación de longitud variable 1016 realiza la codificación de longitud variable en una secuencia de bits asignando una secuencia de bits según el tamaño de la lista de candidatos a bloques de fusión a un índice de candidato a bloque de fusión a utilizar para la codificación.

La figura 8 es un diagrama de flujo que muestra un ejemplo de un proceso para decodificar usando el modo de fusión. En la etapa S2001, se obtienen vectores de movimiento, índices de imágenes de referencia y direcciones de predicción de candidatos a bloques de fusión a partir de bloques vecinos y un bloque co-ubicado. En la etapa S2002, los candidatos idénticos y los candidatos inutilizables para la fusión son eliminados de entre los candidatos a bloques de fusión. En la etapa S2003, el número total de candidatos a bloques de fusión después de la eliminación se establece como el tamaño de la lista de candidatos a bloques de fusión. En la etapa S2004, se decodifica el índice de candidato a bloque de fusión a utilizar para la decodificación de un bloque actual a partir de una secuencia de bits usando el tamaño de la lista de candidatos a bloques de fusión. En la etapa S2005, se realiza la decodificación de un bloque actual generando una imagen de predicción utilizando el candidato a bloque de fusión indicado por el índice de candidato a bloque de fusión decodificado.

La figura 9 muestra un ejemplo de configuración de un aparato de decodificación de imágenes 2000 que decodifica imágenes codificadas usando el modo de fusión. El aparato de decodificación de imágenes 2000 incluye una unidad de decodificación de longitud variable 2001, una unidad de cuantificación inversa 2002, una unidad de transformación ortogonal inversa 2003, un agregador 2004, una memoria de bloques 2005, una memoria de fotogramas 2006, una unidad de intra predicción 2007, una unidad de inter predicción 2008, una unidad de control de inter predicción 2009, un conmutador de 2010, una unidad de cálculo de candidatos a bloque de fusión 2011 y una memoria *colPic* 2012.

En la figura 9, la unidad de cálculo de candidatos a bloque de fusión 2011 calcula los candidatos a bloque de fusión. A continuación, la unidad de cálculo de candidatos a bloques de fusión 2011 transmite el número total calculado de candidatos a bloques de fusión a la unidad de decodificación de longitud variable 2001. La unidad de decodificación de longitud variable 2001 establece el número total de candidatos a bloques de fusión como el tamaño de la lista de candidatos a bloques de fusión que es un parámetro de decodificación. Entonces, la unidad de decodificación de longitud variable 2001 decodifica un índice de candidato a bloque de fusión a partir de la secuencia de bits usando el tamaño de la lista de candidatos a bloques de fusión.

La figura 10 muestra una sintaxis para la asociación de índices de candidatos a bloques de fusión a una secuencia de bits. En la figura 10, *merge_idx* representa un índice de candidato a bloque de fusión, y *merge_flag* representa un indicador de fusión (*merge flag*). *NumMergeCand* representa el tamaño de una lista de candidatos a bloques de fusión. *NumMergeCand* se establece en el número total de candidatos a bloques de fusión después de eliminar los candidatos inutilizables para la fusión y los candidatos idénticos de entre los candidatos a bloques de fusión.

La codificación o decodificación de una imagen se realiza usando el modo de fusión de la manera descrita anteriormente.

Sin embargo, en el modo de fusión, el número total de candidatos a bloques de fusión se establece como el tamaño de una lista de candidatos a bloques de fusión para uso en la codificación o decodificación de un índice de candidato a bloque de fusión. El número total de candidatos a bloques de fusión se determina después de eliminar los candidatos inutilizables para la fusión o candidatos idénticos en base a información sobre imágenes de referencia que incluyen un bloque co-ubicado.

Por lo tanto, se produce una discrepancia en la secuencia de bits asignada a un índice de candidato a bloque de fusión entre un aparato de codificación de imágenes y un aparato de decodificación de imágenes en el caso de que exista una diferencia en el número total de candidatos a bloques de fusión entre el aparato de codificación de imágenes y el aparato de decodificación de imágenes. Como resultado, el aparato de decodificación de imágenes no puede decodificar una secuencia de bits correctamente.

Por ejemplo, cuando se pierde información sobre una imagen de referencia referenciada como un bloque co-ubicado debido a la pérdida de paquetes en una ruta de transmisión, el vector de movimiento o el índice de imagen de referencia del bloque co-ubicado resultan desconocidos. En consecuencia, la información sobre un candidato a bloque de fusión a generar a partir del bloque co-ubicado resulta desconocida. En tal caso, ya no es posible eliminar correctamente candidatos inutilizables para la fusión o candidatos idénticos de entre los candidatos a bloques de fusión en la decodificación. Como resultado, el aparato de decodificación de imágenes no puede obtener el tamaño correcto de una lista de candidatos a bloques de fusión, y por lo tanto es imposible decodificar normalmente un índice de candidato a bloque de fusión.

A la vista de esto, un procedimiento de codificación de imágenes según un aspecto de la presente invención es un procedimiento para codificar una imagen bloque a bloque para generar una secuencia de bits, e incluye: determinar un número máximo de un candidato a fusión que es una combinación de una dirección de predicción, un vector de movimiento y un índice de imagen de referencia para usar en la codificación de un bloque actual; derivar un primer candidato a fusión; determinar si un número total del primer candidato a fusión es o no es menor que el número máximo; derivar un segundo candidato a fusión cuando se determina que el número total del primer candidato a fusión es menor que el número máximo; seleccionar un candidato a fusión para usar en la codificación del bloque actual a partir del primer candidato a fusión y el segundo candidato a fusión; y codificar, usando el número máximo determinado, un índice para identificar el candidato a fusión seleccionado, y asociar el índice codificado a la secuencia de bits.

Con este procedimiento, se puede codificar un índice para identificar un candidato a fusión usando el número máximo determinado. En otras palabras, se puede codificar un índice independientemente del número total de candidatos a fusión realmente derivados. Por lo tanto, incluso cuando se pierde la información necesaria para derivar un candidato a fusión (por ejemplo, información sobre un bloque co-ubicado), todavía se puede codificar un índice y de este modo se mejora la resistencia a errores. Además, se puede decodificar un índice independientemente del número total de candidatos a fusión realmente derivados. En otras palabras, se puede decodificar un índice sin esperar a la derivación de los candidatos a fusión. En otras palabras, se puede generar una secuencia de bits para la cual se puede realizar en paralelo la derivación de los candidatos a fusión y la decodificación de los índices.

Además, con este procedimiento, se puede derivar un segundo candidato a fusión cuando se determina que el número total de los primeros candidatos a fusión es menor que el número máximo. En consecuencia, el número total de candidatos a fusión se puede aumentar dentro de un intervalo que no supere el número máximo, de modo que se puede aumentar la eficacia de la codificación.

Por ejemplo, en la derivación de un primer candidato a fusión, se puede derivar un candidato a fusión que es una combinación de una dirección de predicción, un vector de movimiento y un índice de imagen de referencia como el primer candidato a fusión, siendo la combinación diferente de una combinación de una dirección de predicción, un vector de movimiento y un índice de imagen de referencia de cualquier primer candidato a fusión derivado anteriormente.

Con esto, un candidato a fusión que es una combinación de una dirección de predicción, un vector de movimiento y un índice de imagen de referencia idéntica a una combinación de una dirección de predicción, un vector de movimiento y un índice de imagen de referencia de cualquier primer candidato a fusión derivado anteriormente puede ser eliminado de entre los primeros candidatos a fusión. Se puede aumentar el número total de los segundos candidatos a fusión consecuentemente, de modo que se puede aumentar la variedad de combinaciones de una dirección de predicción, un vector de movimiento y un índice de imagen de referencia de un candidato a fusión seleccionable. Por lo tanto, es posible aumentar aún más la eficacia de la codificación.

Por ejemplo, en la derivación de un primer candidato a fusión, se puede derivar el primer candidato a fusión en base a una dirección de predicción, un vector de movimiento y un índice de imagen de referencia utilizados en la codificación de un bloque espacialmente o temporalmente vecino del bloque actual.

Con esto, en la derivación de un primer candidato a fusión, se puede derivar un primer candidato a fusión en base a una dirección de predicción, un vector de movimiento, y un índice de imagen de referencia usados para la codificación de un bloque espacialmente o temporalmente vecino del bloque actual.

5 Por ejemplo, en la derivación de un primer candidato a fusión, se puede derivar una combinación de una dirección de predicción, un vector de movimiento y un índice de imagen de referencia como el primer candidato a fusión, habiendo sido usada la combinación de la dirección de predicción, vector de movimiento, e índice de imagen de referencia en la codificación de un bloque de entre bloques espacialmente vecinos del bloque actual excepto un bloque codificado por intra predicción, un bloque fuera de un límite de una porción que incluye el bloque actual o un límite de una imagen que incluye el bloque actual, y un bloque aún por codificar.

10 Con esto, se puede derivar un primer candidato a fusión a partir de bloques apropiados para la obtención de un candidato a fusión.

Por ejemplo, en la derivación de un segundo candidato a fusión, se puede derivar como el segundo candidato a fusión un candidato a fusión que es diferente en al menos uno de entre la dirección de predicción, el vector de movimiento y el índice de imagen de referencia del primer candidato a fusión.

15 Con esto, se puede derivar como un segundo candidato a fusión un candidato a fusión que es diferente en al menos uno de entre la dirección de predicción, el vector de movimiento y el índice de imagen de referencia de un primer candidato a fusión. Por lo tanto, es posible aumentar el número total de candidatos a fusión, teniendo cada uno una combinación diferente de una dirección de predicción, un vector de movimiento y un índice de imagen de referencia, de modo que se puede aumentar adicionalmente la eficacia de la codificación.

20 Por ejemplo, en la codificación, se puede asociar además a la secuencia de bits información que indica el número máximo determinado.

Con esto, se puede asociar a una secuencia de bits información que indica el número máximo determinado. Por lo tanto, es posible cambiar los números máximos por la unidad adecuada para aumentar la eficacia de la codificación.

25 Por ejemplo, el procedimiento de codificación de imágenes puede incluir además: conmutar un proceso de codificación entre un primer proceso de codificación conforme a un primer estándar y un segundo proceso de codificación conforme a un segundo estándar; y asociar, a la secuencia de bits, información de identificación que indica el primer estándar o el segundo estándar conforme al cual se ajusta el proceso de codificación después de la conmutación, en el que cuando el proceso de codificación conmuta al primer proceso de codificación, la determinación de un número máximo de un candidato a fusión, la derivación de un primer candidato a fusión, la determinación de si el número total del primer candidato a fusión es o no es menor que el número máximo, la derivación de un segundo candidato a fusión, la selección y la codificación se realizan como el primer proceso de codificación.

Con esto, es posible realizar de forma conmutable el primer proceso de codificación conforme al primer estándar y el segundo proceso de codificación conforme al segundo estándar.

35 Además, un procedimiento de decodificación de imágenes según un aspecto de la presente invención es un procedimiento para decodificar, bloque a bloque, una imagen codificada incluida en una secuencia de bits, e incluye: determinar un número máximo de un candidato a fusión que es una combinación de una dirección de predicción, un vector de movimiento y un índice de imagen de referencia para usar en la decodificación de un bloque actual; derivar un primer candidato a fusión; determinar si el número total del primer candidato a fusión es o no es mayor que el número máximo; derivar un segundo candidato a fusión cuando se determina que el número total del primer candidato a fusión es menor que el número máximo; decodificar un índice codificado y asociado a la secuencia de bits, usando el número máximo determinado, siendo el índice un índice para identificar un candidato a fusión; y seleccionar, en base al índice decodificado, un candidato a fusión para usarlo en la decodificación de un bloque actual, seleccionándose el candidato a fusión seleccionado de entre el primer candidato a fusión y el segundo candidato a fusión.

40 Con esto, se puede decodificar un índice para identificar un candidato a fusión usando un número máximo determinado. En otras palabras, se puede decodificar un índice independientemente del número total de candidatos a fusión realmente derivados. Por lo tanto, incluso cuando se pierde información necesaria para derivar un candidato a fusión (por ejemplo, información sobre un bloque co-ubicado), todavía se puede decodificar un índice y de este modo se mejora la resistencia a errores. Además, se puede decodificar un índice sin esperar a la derivación de los candidatos a fusión, de modo que se pueden realizar en paralelo la derivación de candidatos a fusión y la decodificación de índices.

45 Además, con este procedimiento, se puede derivar un segundo candidato a fusión cuando se determina que el número total de los primeros candidatos a fusión es menor que el número máximo. En consecuencia, se puede aumentar el número total de candidatos a fusión dentro de un intervalo que no supera el número máximo de manera que se puede decodificar de forma apropiada una secuencia de bits codificada con una mayor eficacia de la codificación.

Por ejemplo, en la derivación de un primer candidato a fusión, se puede derivar como el primer candidato a fusión un candidato a fusión que es una combinación de una dirección de predicción, un vector de movimiento y un índice de imagen de referencia, siendo la combinación diferente de una combinación de una dirección de predicción, un vector de movimiento y un índice de imagen de referencia de cualquier primer candidato a fusión derivado anteriormente.

- 5 Con esto, se puede eliminar de entre los primeros candidatos a fusión un candidato a fusión que es una combinación de una dirección de predicción, un vector de movimiento y un índice de imagen de referencia idéntica a una combinación de una dirección de predicción, un vector de movimiento y un índice de imagen de referencia de cualquier primer candidato a fusión derivado anteriormente. El número total de los segundos candidatos a fusión se puede aumentar consecuentemente, de modo que se puede aumentar la variedad de combinaciones de una
- 10 dirección de predicción, un vector de movimiento y un índice de imagen de referencia de un candidato a fusión seleccionable. Por lo tanto, es posible decodificar una secuencia de bits codificada con una mayor eficacia de la codificación.

- Por ejemplo, en la derivación de un primer candidato a fusión, se puede derivar el primer candidato a fusión en base a una dirección de predicción, un vector de movimiento y un índice de imagen de referencia utilizados en la
- 15 decodificación de un bloque espacialmente o temporalmente vecino del bloque actual.

Con esto, se puede derivar un primer candidato a fusión en base a una dirección de predicción, un vector de movimiento y un índice de imagen de referencia usados para decodificar un bloque actual espacialmente o temporalmente vecino del bloque actual.

- 20 Por ejemplo, en la derivación de un primer candidato a fusión, se puede derivar como el primer candidato a fusión una combinación de una dirección de predicción, un vector de movimiento y un índice de imagen de referencia, habiéndose usado la combinación de la dirección de predicción, vector de movimiento e índice de imagen de referencia en la decodificación de un bloque de entre bloques espacialmente vecinos del bloque actual excepto un bloque decodificado por intra predicción, un bloque fuera de un límite de una porción que incluye el bloque actual o un límite de una imagen que incluye el bloque actual, y un bloque aún por decodificar.

- 25 Con esto, se puede derivar un primer candidato a fusión a partir de bloques apropiados para la obtención de un candidato a fusión.

Por ejemplo, en la derivación de un segundo candidato a fusión, se puede derivar como el segundo candidato a fusión un candidato a fusión que es diferente del primer candidato a fusión en al menos uno de entre la dirección de predicción, vector de movimiento e índice de imagen de referencia.

- 30 Con esto, se puede derivar como un segundo candidato a fusión un candidato a fusión que es diferente de un primer candidato a fusión en al menos uno de entre la dirección de predicción, vector de movimiento e índice de imagen de referencia. Por lo tanto, es posible aumentar el número total de candidatos a fusión teniendo cada uno una combinación diferente de una dirección de predicción, un vector de movimiento y un índice de imagen de referencia, de modo que se puede decodificar de forma apropiada una secuencia de bits codificada con una eficacia de
- 35 codificación aún mayor.

Por ejemplo, en la determinación de un número máximo de un candidato a fusión, el número máximo se puede determinar en base a la información asociada a la secuencia de bits y que indica el número máximo.

Con esto, se puede determinar un número máximo en base a información asociada a una secuencia de bits. Por lo tanto, es posible decodificar una imagen codificada usando números máximos modificados por la unidad apropiada.

- 40 Por ejemplo, el procedimiento de decodificación de imágenes puede incluir además: conmutar un proceso de decodificación entre un primer proceso de decodificación conforme a un primer estándar y un segundo proceso de decodificación conforme a un segundo estándar, de acuerdo con la información de identificación que está asociada a la secuencia de bits e indica el primer estándar o el segundo estándar, en el que cuando se conmuta el proceso de decodificación al primer proceso de decodificación, la determinación de un número máximo de un candidato a fusión,
- 45 la derivación de un primer candidato a fusión, la determinación de si el número total del primer candidato a fusión es menor que el número máximo, la derivación de un segundo candidato a fusión, la decodificación y la selección se realizan como el primer proceso de decodificación.

Con esto, es posible realizar de forma conmutable el primer proceso de decodificación conforme al primer estándar y el segundo proceso de decodificación conforme al segundo estándar.

- 50 Se debe observar que estos aspectos generales o específicos se pueden implementar como un sistema, un procedimiento, un circuito integrado, un programa informático, un medio de grabación legible por ordenador tal como una memoria de solo lectura de disco compacto (CD-ROM), o como cualquier combinación de un sistema, un procedimiento, un circuito integrado, un programa informático y un medio de grabación legible por ordenador.

- 55 A continuación se describirán específicamente con referencia a los dibujos, un aparato de codificación de imágenes y un aparato de decodificación de imágenes según un aspecto de la presente invención.

Cada una de las realizaciones a modo de ejemplo descritas a continuación muestra un ejemplo específico de la presente invención.

Los valores numéricos, formas, materiales, elementos constituyentes, la disposición y conexión de los elementos constituyentes, las etapas, el orden de procesamiento de las etapas, etc. que se muestran en las siguientes realizaciones a modo de ejemplo son meros ejemplos, y por lo tanto no limitan el concepto inventivo de la presente invención. Por lo tanto, entre los elementos constituyentes de las siguientes realizaciones a modo de ejemplo, los elementos constituyentes que no se citan en cualquiera de las reivindicaciones independientes que definen la parte más genérica del concepto inventivo no son necesariamente requeridos para superar las desventajas.

[Realización 1]

10 La figura 11 es un diagrama de bloques que muestra una configuración de un aparato de codificación de imágenes 100 de acuerdo con la Realización 1. El aparato de codificación de imágenes 100 codifica una imagen bloque a bloque para generar una secuencia de bits (*bitstream*).

15 Como se muestra en la figura 11, el aparato de codificación de imágenes 100 incluye un sustractor 101, una unidad de transformación ortogonal 102, una unidad de cuantificación 103, una unidad de cuantificación inversa 104, una unidad de transformación ortogonal inversa 105, un agregador 106, una memoria de bloques 107, una memoria de fotogramas (*frames*) 108, una unidad de intra predicción 109, una unidad de inter predicción 110, una unidad de control de inter predicción 111, una unidad de determinación del tipo de imagen 112, un conmutador 113, una unidad de cálculo de candidatos a bloque de fusión 114, una memoria *colPic* 115, y una unidad de codificación de longitud variable 116.

20 El sustractor 101 sustrae, bloque a bloque, datos de imagen de predicción a partir de datos de imagen de entrada incluidos en una secuencia de imágenes de entrada para generar datos de error de predicción.

La unidad de transformación ortogonal 102 transforma los datos de error de predicción generados a partir de un dominio de imagen en un dominio de frecuencia.

25 La unidad de cuantificación 103 cuantifica los datos de error de predicción transformados en un dominio de frecuencia.

La unidad de cuantificación inversa 104 cuantifica inversamente los datos de error de predicción cuantificados por la unidad de cuantificación 103.

La unidad de transformación ortogonal inversa 105 transforma los datos de error de predicción cuantificados inversamente de un dominio de frecuencia en un dominio de imagen.

30 El agregador 106 agrega, bloque a bloque, datos de imagen de predicción y los datos de error de predicción cuantificados inversamente por la unidad de transformación ortogonal inversa 105 para generar datos de imagen reconstruida.

La memoria de bloques 107 almacena los datos de imagen reconstruida en unidades de un bloque.

La memoria de fotogramas 108 almacena los datos de imagen reconstruida en unidades de un fotograma (*frame*).

35 La unidad de determinación del tipo de imagen 112 determina en qué tipos de imagen de entre imagen I, imagen B e imagen P se deben codificar los datos de imagen de entrada. Entonces, la unidad de determinación del tipo de imagen 112 genera información del tipo de imagen que indica el tipo de imagen determinado.

40 La unidad de intra predicción 109 genera datos de imagen de intra predicción de un bloque actual realizando una intra predicción que usa datos de imagen reconstruida almacenados en la memoria de bloques 107 en unidades de un bloque.

La unidad de inter predicción 110 genera datos de imagen de inter predicción de un bloque actual realizando una inter predicción que usa datos de imagen reconstruida almacenados en la memoria de fotogramas 108 en unidades de un fotograma y un vector de movimiento derivado por un proceso que incluye estimación de movimiento.

45 Cuando un bloque actual se codifica por codificación de intra predicción, el conmutador 113 genera datos de imagen de intra predicción generados por la unidad de intra predicción 109 como datos de imagen de predicción del bloque actual para el sustractor 101 y el agregador 106. Por otro lado, cuando un bloque actual se codifica por codificación de inter predicción, el conmutador 113 genera datos de imagen de inter predicción generados por la unidad de inter predicción 110 como datos de imagen de predicción del bloque actual para el sustractor 101 y el agregador 106.

50 La unidad de cálculo de candidatos a bloque de fusión 114 deriva candidatos a bloque de fusión para el modo de fusión usando vectores de movimiento y otros de bloques vecinos del bloque actual y un vector de movimiento y otros del bloque co-ubicado (información *colPic*) almacenados en la memoria *colPic* 115. Entonces, la unidad de cálculo de candidatos a bloque de fusión 114 calcula el número total de candidatos utilizables para la fusión usando

un procedimiento descrito más adelante.

Además, la unidad de cálculo de candidatos a bloque de fusión 114 asigna índices de candidatos a bloque de fusión que tienen cada uno un valor diferente a los candidatos a bloque de fusión derivados. A continuación, la unidad de cálculo de candidatos a bloque de fusión 114 transmite los candidatos a bloque de fusión y los índices de candidatos a bloque de fusión a la unidad de control de inter predicción 111. Además, la unidad de cálculo de candidatos a bloque de fusión 114 transmite el número calculado de candidatos utilizables para la fusión a la unidad de codificación de longitud variable 116.

La unidad de control de inter predicción 111 selecciona un modo de predicción utilizando cuál es el error de predicción que es el menor de entre un modo de predicción que usa un vector de movimiento derivado por estimación de movimiento (modo de estimación de movimiento) y un modo de predicción que usa un vector de movimiento derivado a partir de un candidato a bloque de fusión (modo de fusión). La unidad de control de inter predicción 111 también transmite un indicador de fusión que indica si el modo de predicción seleccionado es o no es el modo de fusión a la unidad de codificación de longitud variable 116. Además, la unidad de control de inter predicción 111 transmite un índice de candidato a bloque de fusión correspondiente a los candidatos a bloque de fusión determinados a la unidad de codificación de longitud variable 116 cuando el modo de predicción seleccionado es el modo de fusión. Además, la unidad de control de inter predicción 111 transfiere la información *colPic* que incluye el vector de movimiento y otros del bloque actual a la memoria *colPic* 115.

La unidad de codificación de longitud variable 116 genera una secuencia de bits (*bitstream*) realizando una codificación de longitud variable en los datos de error de predicción cuantificados, el indicador de fusión y la información del tipo de imagen. La unidad de codificación de longitud variable 116 también establece el número total de candidatos utilizables para la fusión como el tamaño de la lista de candidatos a bloques de fusión. Además, la unidad de codificación de longitud variable 116 realiza la codificación de longitud variable en un índice de candidato a bloque de fusión para usarlo en la codificación, asignando, de acuerdo con el tamaño de la lista de candidatos a bloques de fusión, una secuencia de bits al índice de candidato a bloque de fusión.

La figura 12 es un diagrama de flujo que muestra operaciones de procesamiento del aparato de codificación de imágenes 100 de acuerdo con la Realización 1.

En la etapa S101, la unidad de cálculo de candidatos a bloque de fusión 114 deriva candidatos a bloques de fusión a partir de bloques vecinos y un bloque co-ubicado de un bloque actual. Además, la unidad de cálculo de candidatos a bloque de fusión 114 calcula el tamaño de una lista de candidatos a bloque de fusión usando un procedimiento descrito más adelante.

Por ejemplo, en el caso mostrado en la figura 3, la unidad de cálculo de candidatos a bloque de fusión 114 selecciona los bloques vecinos del A al D como candidatos a bloque de fusión. Además, la unidad de cálculo de candidatos a bloque de fusión 114 calcula, como candidato a bloque de fusión, un bloque de fusión co-ubicado que tiene un vector de movimiento, un índice de imagen de referencia y una dirección de predicción que se calculan a partir del vector de movimiento de un bloque co-ubicado usando el modo de predicción de tiempo.

La unidad de cálculo de candidatos a bloque de fusión 114 asigna índices de candidatos a bloque de fusión a los respectivos candidatos a bloque de fusión como se muestra en (a) en la figura 13. A continuación, la unidad de cálculo de candidatos a bloque de fusión 114 calcula una lista de candidatos a bloque de fusión como se muestra en (b) en la figura 13 y el tamaño de la lista de candidatos a bloques de fusión eliminando candidatos inutilizables para la fusión y candidatos idénticos y agregando nuevos candidatos usando un procedimiento descrito más adelante.

Se asignan códigos más cortos a los índices de candidatos a bloques de fusión con valores menores. En otras palabras, cuanto menor sea el valor de un índice de candidato a bloque de fusión, menor será la cantidad de información necesaria para indicar el índice de candidato a bloque de fusión.

Por otra parte, cuanto mayor sea el valor de un índice de candidato a bloque de fusión, mayor será la cantidad de información necesaria para el índice de candidato a bloque de fusión. Por lo tanto, la eficacia de la codificación aumentará cuando se asignen índices de candidatos a bloques de fusión con valores menores a candidatos a bloques de fusión que tienen más probabilidades de tener vectores de movimiento de mayor precisión e índices de imagen de referencia de mayor precisión.

Por lo tanto, un caso posible es que la unidad de cálculo de candidatos a bloque de fusión 114 cuente el número total de veces de selección de cada candidato a bloque de fusión como un bloque de fusión, y asigne índices de candidatos a bloques de fusión con valores menores a los bloques con un mayor número total de veces. Específicamente, esto se puede conseguir cuando se codifica un bloque actual, mediante la especificación de un bloque de fusión seleccionado de entre bloques vecinos y la asignación de un índice de candidato a bloque de fusión de menor valor al bloque de fusión especificado.

Cuando un candidato a bloque de fusión no tiene información tal como un vector de movimiento (por ejemplo, cuando el bloque de fusión ha sido un bloque codificado por intra predicción, se encuentra fuera del límite de una imagen o del límite de una porción, o aún no ha sido codificado), el candidato a bloque de fusión no se puede usar

para la codificación.

En la Realización 1, un candidato a bloque de fusión inutilizable para la codificación se denomina candidato inutilizable para la fusión, y un candidato a bloque de fusión utilizable para la codificación se denomina candidato utilizable para la fusión. Además, entre una pluralidad de candidatos a bloques de fusión, un candidato a bloque de fusión idéntico en vector de movimiento, índice de imagen de referencia y dirección de predicción a cualquier otro bloque de fusión se denomina candidato idéntico.

En el caso mostrado en la figura 3, el bloque vecino C es un candidato inutilizable para la fusión porque es un bloque codificado por intra predicción. El bloque vecino D es un candidato idéntico porque es idéntico al bloque vecino A en el vector de movimiento, el índice de imagen de referencia y la dirección de predicción.

En la etapa S102, la unidad de control de inter predicción 111 selecciona un modo de predicción en base a una comparación, usando un procedimiento descrito más adelante, entre el error de predicción de una imagen de predicción que se ha generado usando un vector de movimiento derivado por estimación de movimiento y el error de predicción de una imagen de predicción que se ha generado usando un vector de movimiento obtenido a partir de un candidato a bloque de fusión. Cuando el modo de predicción seleccionado es el modo de fusión, la unidad de control de inter predicción 111 pone el indicador de fusión a 1, y cuando no, la unidad de control de inter predicción 111 pone el indicador de fusión a 0.

En la etapa S103, se determina si el indicador de fusión es o no es 1 (es decir, si el modo de predicción seleccionado es o no es el modo de fusión).

Cuando el resultado de la determinación en la etapa S103 es verdadero (Sí, S103), la unidad de codificación de longitud variable 116 asocia el indicador de fusión a una secuencia de bits en la etapa S104. Posteriormente, en la etapa S105, la unidad de codificación de longitud variable 116 asigna secuencias de bits de acuerdo con el tamaño de la lista de candidatos a bloques de fusión como se muestra en la figura 5 a los índices de candidatos a bloque de fusión de los candidatos a bloque de fusión a utilizar en la codificación. Entonces, la unidad de codificación de longitud variable 116 realiza la codificación de longitud variable en la secuencia de bits asignada.

Por otro lado, cuando el resultado de la determinación en la etapa S103 es falso (S103, No), la unidad de codificación de longitud variable 116 asocia información sobre un indicador de fusión y un modo de vector de estimación de movimiento a una secuencia de bits en la etapa S106.

En la Realización 1, se asigna un índice de candidato a bloque de fusión que tiene un valor "0" al bloque vecino A como se muestra en (a) en la figura 13. Se asigna un índice de candidato a bloque de fusión que tiene un valor "1" al bloque vecino B. Se asigna un índice de candidato a bloque de fusión que tiene un valor "2" al bloque de fusión co-ubicado. Se asigna un índice de candidato a bloque de fusión que tiene un valor "3" al bloque vecino C. Se asigna un índice de candidato a bloque de fusión que tiene un valor "4" al bloque vecino D.

Debe observarse que los índices de candidatos a bloques de fusión que tienen dicho valor pueden asignarse de otro modo. Por ejemplo, cuando se agrega un nuevo candidato usando un procedimiento descrito más adelante, la unidad de codificación de longitud variable 116 puede asignar valores menores a los candidatos a bloques de fusión preexistentes y un valor mayor al nuevo candidato. En otras palabras, la unidad de codificación de longitud variable 116 puede asignar un índice de candidato a bloque de fusión de un valor menor a un candidato a bloque de fusión preexistente dando prioridad a un nuevo candidato.

Además, los candidatos a bloques de fusión no se limitan a los bloques en las posiciones de los bloques vecinos A, B, C y D. Por ejemplo, un bloque vecino ubicado encima del bloque vecino inferior izquierdo D se puede usar como un candidato a bloque de fusión. Además, no es necesario usar todos los bloques vecinos como candidatos a bloques de fusión. Por ejemplo, también es posible usar solo los bloques vecinos A y B como candidatos a bloques de fusión.

Además, aunque la unidad de codificación de longitud variable 116 asocia un índice de candidato a bloque de fusión a una secuencia de bits en la etapa S105 de la figura 12 en la Realización 1, la asociación de dicho índice de candidato a bloque de fusión a una secuencia de bits no siempre es necesario. Por ejemplo, la unidad de codificación de longitud variable 116 no necesita asociar un índice de candidato a bloque de fusión a una secuencia de bits cuando el tamaño de la lista de candidatos a bloques de fusión es "1". De este modo, se reduce la cantidad de información sobre el índice de candidato a bloque de fusión.

La figura 14 es un diagrama de flujo que muestra detalles del proceso en la etapa S101 de la figura 12. Específicamente, la figura 14 ilustra un procedimiento de cálculo de candidatos a bloques de fusión y el tamaño de una lista de candidatos a bloques de fusión. La figura 14 se describirá a continuación.

En la etapa S111, la unidad de cálculo de candidatos a bloque de fusión 114 determina si un candidato a bloque de fusión [N] es o no es un candidato utilizable para la fusión utilizando un procedimiento descrito más adelante. Entonces, la unidad de cálculo de candidatos a bloque de fusión 114 actualiza el número total de candidatos utilizables para la fusión de acuerdo con el resultado de la determinación.

En este punto, N denota un valor de índice para identificar cada candidato a bloque de fusión. En la Realización 1, N toma valores del 0 al 4. Específicamente, el bloque vecino A de la figura 3 es asignado a un candidato a bloque de fusión [0]. El bloque vecino B de la figura 3 es asignado a un candidato a bloque de fusión [1]. El bloque de fusión co-ubicado es asignado a un candidato a bloque de fusión [2]. El bloque vecino C de la figura 3 es asignado a un candidato a bloque de fusión [3]. El bloque vecino D de la figura 3 es asignado a un candidato a bloque de fusión [4].

En la etapa S112, la unidad de cálculo de candidatos a bloque de fusión 114 obtiene el vector de movimiento, el índice de imagen de referencia y la dirección de predicción del candidato a bloque de fusión [N], y los agrega a una lista de candidatos a bloque de fusión.

En la etapa S113, la unidad de cálculo de candidatos a bloque de fusión 114 busca en la lista de candidatos a bloque de fusión un candidato inutilizable para la fusión y un candidato idéntico, y elimina el candidato inutilizable para la fusión y el candidato idéntico de la lista de candidatos a bloque de fusión como se muestra en la figura 13.

En la etapa S114, la unidad de cálculo de candidatos a bloque de fusión 114 agrega un nuevo candidato a la lista de candidatos a bloque de fusión usando un procedimiento descrito más adelante. En este punto, cuando se agrega un nuevo candidato, la unidad de cálculo de candidatos a bloque de fusión 114 puede reasignar índices de candidatos a bloque de fusión de modo que los índices de candidatos a bloque de fusión de valores menores son asignados a candidatos a bloques de fusión preexistentes dando prioridad al nuevo candidato. En otras palabras, la unidad de cálculo de candidatos a bloque de fusión 114 puede reasignar los índices de candidatos a bloque de fusión de modo que se asigna un índice de candidato a bloque de fusión de mayor valor al candidato nuevo. De este modo, se reduce la cantidad de código de índices de candidatos a bloques de fusión.

En la etapa S115, la unidad de cálculo de candidatos a bloque de fusión 114 establece el número total de candidatos utilizables para la fusión calculado en la etapa S111 como el tamaño de la lista de candidatos a bloques de fusión. En el ejemplo mostrado en la figura 13, el número calculado de candidatos utilizables para la fusión es "4", y el tamaño de la lista de candidatos a bloque de fusión se pone a 4.

El nuevo candidato en la etapa S114 es un candidato agregado nuevamente a los candidatos a bloques de fusión usando un procedimiento descrito más adelante cuando el número total de candidatos a bloques de fusión es menor que el número total de candidatos utilizables para la fusión. Ejemplos de dicho nuevo candidato incluyen un bloque vecino situado encima del bloque vecino inferior izquierdo D de la figura 3, un bloque que corresponde a cualquiera de los bloques vecinos A, B, C y D de un bloque co-ubicado. Además, ejemplos de dicho nuevo candidato incluyen además un bloque que tiene un vector de movimiento, un índice de imagen de referencia, una dirección de predicción, y similares, que se obtienen estadísticamente para toda la región o una cierta región de una imagen de referencia.

Por lo tanto, cuando el número total de candidatos a bloques de fusión es menor que el número total de candidatos utilizables para la fusión, la unidad de cálculo de candidatos a bloques de fusión 114 agrega un nuevo candidato que tiene un nuevo vector de movimiento, un nuevo índice de imagen de referencia y una nueva dirección de predicción de modo que se incrementa la eficacia de la codificación.

La figura 15 es un diagrama de flujo que muestra detalles del proceso en la etapa S111 de la figura 14. Específicamente, la figura 15 ilustra un procedimiento para determinar si un candidato a bloque de fusión [N] es o no es un candidato utilizable para la fusión y actualiza el número total de candidatos utilizables para la fusión. A continuación se describirá la figura 15.

En la etapa S121, la unidad de cálculo de candidatos a bloque de fusión 114 determina si es verdadero o falso que (1) un candidato a bloque de fusión [N] ha sido codificado por intra predicción, (2) el candidato a bloque de fusión [N] es un bloque fuera del límite de una porción que incluye el bloque actual o del límite de una imagen que incluye el bloque actual, o (3) el candidato a bloque de fusión [N] aún está por codificar.

Cuando el resultado de la determinación en la etapa S121 es verdadero (S121, Si), la unidad de cálculo de candidatos a bloque de fusión 114 establece el candidato a bloque de fusión [N] como candidato inutilizable para la fusión en la etapa S122. Por otra parte, cuando el resultado de la determinación en la etapa S121 es falso (S121, No), la unidad de cálculo de candidatos a bloque de fusión 114 establece el candidato a bloque de fusión [N] como candidato utilizable para la fusión en la etapa S123.

En la etapa S124, la unidad de cálculo de candidatos a bloque de fusión 114 determina si es verdadero o falso que el candidato a bloque de fusión [N] es un candidato utilizable para la fusión o un candidato a bloque de fusión co-ubicado. En este punto, cuando el resultado de la determinación en la etapa S124 es verdadero (S124, Si), la unidad de cálculo de candidatos a bloque de fusión 114 actualiza el número total de candidatos a bloque de fusión incrementándolo en uno en la etapa S125. Por otra parte, cuando el resultado de la determinación en la etapa S124 es falso (S124, No), la unidad de cálculo de candidatos a bloque de fusión 114 no actualiza el número total de candidatos utilizables para la fusión.

Así, cuando un candidato a bloque de fusión es un bloque de fusión co-ubicado, la unidad de cálculo de candidatos a bloque de fusión 114 incrementa en uno el número total de candidatos utilizables para la fusión, con

independencia de si el bloque co-ubicado es un candidato utilizable para la fusión o un candidato inutilizable para la fusión. Esto evita que haya discrepancia en los números de candidatos utilizables para la fusión entre el aparato de codificación y el aparato de decodificación de imágenes, incluso cuando se pierde información en un bloque de fusión co-ubicado debido a un incidente tal como la pérdida de paquete(s).

- 5 El número total de candidatos utilizables para la fusión se establece como el tamaño de la lista de candidatos a bloques de fusión en la etapa S115 mostrada en la figura 14. Además, se usa el tamaño de la lista de candidatos a bloques de fusión en la codificación de longitud variable de los índices de candidatos a bloques de fusión en la etapa S105 mostrada en la figura 12. Esto hace posible que el aparato de codificación de imágenes 100 genere una secuencia de bits que se puede decodificar normalmente de modo que pueden obtenerse índices de candidatos a bloques de fusión incluso cuando se pierde información sobre la imagen de referencia que incluye un bloque co-ubicado.

La figura 16 es un diagrama de flujo que muestra detalles del proceso en la etapa S114 de la figura 14. Específicamente, la figura 16 ilustra un procedimiento para agregar un nuevo candidato. La figura 16 se describirá a continuación.

- 15 En la etapa S131, la unidad de cálculo de candidatos a bloque de fusión 114 determina si el número total de candidatos a bloques de fusión es o no es menor que el número total de candidatos utilizables para la fusión. En otras palabras, la unidad de cálculo de candidatos a bloque de fusión 114 determina si el número total de candidatos a bloque de fusión está o no está todavía por debajo del número total de candidatos utilizables para la fusión.

- 20 En este punto, cuando el resultado de la determinación en la etapa S131 es verdadero (S131, Si), en la etapa S132, la unidad de cálculo de candidatos a bloque de fusión 114 determina si hay o no hay un nuevo candidato que puede agregarse como un candidato a bloque de fusión a la lista de candidatos a bloque de fusión. En este punto, cuando el resultado de la determinación en la etapa S132 es verdadero (S132, Si), la unidad de cálculo de candidatos a bloque de fusión 114 asigna un índice de candidato a bloque de fusión que tiene un valor al nuevo candidato y agrega el nuevo candidato a la lista de candidatos a bloque de fusión en la etapa S133. Además, en la etapa S134, la unidad de cálculo de candidatos a bloque de fusión 114 incrementa en uno el número total de candidatos a bloque de fusión.

- 25 Por otra parte, cuando el resultado de la determinación en la etapa S131 o en la etapa S132 es falso (S131 o S132, No), el proceso para agregar un nuevo candidato finaliza. En otras palabras, el proceso para agregar un nuevo candidato finaliza cuando el número total de candidatos a bloques de fusión alcanza el número total de candidatos utilizables para la fusión o cuando no hay un candidato nuevo.

La figura 17 es un diagrama de flujo que muestra detalles del proceso en la etapa S102 de la figura 12. Específicamente, la figura 17 ilustra un proceso para seleccionar un candidato a bloque de fusión. La figura 17 se describirá a continuación.

- 35 En la etapa S141, la unidad de control de inter predicción 111 pone a 0 un índice de candidato a bloque de fusión, establece el error de predicción mínimo en el error de predicción (coste) del modo de estimación del vector de movimiento, y pone a 0 un indicador de fusión. En este punto, el coste se calcula usando la siguiente fórmula para un modelo de optimización de tasa-distorsión (*RD optimization model*), por ejemplo.

(Ecuación 1)

$$\text{Coste} = D + \lambda R$$

- 40 En la ecuación 1, D denota la distorsión de la codificación. Por ejemplo, D es la suma de las diferencias absolutas entre los valores de píxel originales de un bloque actual a codificar y los valores de píxel obtenidos mediante la codificación y decodificación del bloque actual usando una imagen de predicción generada usando un vector de movimiento. R denota la cantidad de códigos generados. Por ejemplo, R es la cantidad de código necesaria para codificar un vector de movimiento utilizado para la generación de una imagen de predicción. λ denota un multiplicador de Lagrange indeterminado.

En la etapa S142, la unidad de control de inter predicción 111 determina si el valor de un índice de candidato a bloque de fusión es o no es menor que el número total de candidatos a bloque de fusión de un bloque actual. En otras palabras, la unidad de control de inter predicción 111 determina si todavía hay o no hay un candidato a bloque de fusión en el que no se ha realizado todavía el proceso de la etapa S143 a la etapa S145.

- 50 Cuando el resultado de la determinación en la etapa S142 es verdadero (S142, Si), en la etapa S143, la unidad de control de inter predicción 111 calcula el coste para un candidato a bloque de fusión al cual se le asigna un índice de candidato a bloque de fusión. A continuación, en la etapa S144, la unidad de control de inter predicción 111 determina si el coste calculado para un candidato a bloque de fusión es o no es menor que el error de predicción mínimo.

- 55 En este punto, cuando el resultado de la determinación en la etapa S144 es verdadero, (S144, Si), la unidad de

control de inter predicción 111 actualiza el error de predicción mínimo, el índice de candidato a bloque de fusión y el valor del indicador de fusión en la etapa S145. Por otro lado, cuando el resultado de la determinación en la etapa S144 es falso (S144, No), la unidad de control de inter predicción 111 no actualiza el error de predicción mínimo, el índice de candidato a bloque de fusión o el valor del indicador de fusión.

- 5 En la etapa S146, la unidad de control de inter predicción 111 incrementa en uno el índice de candidato a bloque de fusión, y repite desde la etapa S142 hasta la etapa S146.

Por otro lado, cuando el resultado de la etapa S142 es falso (S142, No), es decir, no hay más candidatos a bloque de fusión sin procesar, la unidad de control de inter predicción 111 fija los valores finales del indicador de fusión y del índice de candidato a bloque de fusión en la etapa S147.

- 10 De este modo, el aparato de codificación de imágenes 100 según la realización 1 es capaz de calcular el tamaño de una lista de candidatos a bloques de fusión para su uso en la codificación o decodificación de un índice de candidato a bloque de fusión, utilizando un procedimiento independiente de la información sobre imágenes de referencia que incluyen un bloque co-ubicado. El aparato de codificación de imágenes 100 consigue de este modo una resistencia a errores mejorada.

- 15 Más específicamente, con independencia de si un bloque de fusión co-ubicado es un candidato utilizable para la fusión, el aparato de codificación de imágenes 100 de acuerdo con la Realización 1 incrementa en uno el número total de candidatos utilizables para la fusión cada vez que se determina un candidato a bloque de fusión como bloque de fusión co-ubicado. Entonces, el aparato de codificación de imágenes 100 determina una secuencia de bits para asignarla a un índice de candidato a bloque de fusión, usando el número total de candidatos utilizables para la fusión calculado de la manera indicada. El aparato de codificación de imágenes 100 es por lo tanto capaz de generar una secuencia de bits a partir de la cual se puede decodificar normalmente el índice de candidato a bloque de fusión incluso cuando se pierde información sobre imágenes de referencia que incluyen un bloque co-ubicado.

- 20 Además, cuando el número total de candidatos a bloques de fusión es menor que el número total de candidatos utilizables para la fusión, el aparato de codificación de imágenes 100 de acuerdo con la Realización 1 agrega, como candidato a bloque de fusión, un candidato nuevo que tiene un nuevo vector de movimiento, un nuevo índice de imagen de referencia y una nueva dirección de predicción de modo que se incrementa la eficacia de la codificación.

- 25 Debe observarse que el ejemplo descrito en la Realización 1 en el que el indicador de fusión es siempre asociado a una secuencia de bits en el modo de fusión no es limitante. Por ejemplo, el modo de fusión se puede seleccionar a la fuerza dependiendo de una forma de bloque para uso en la inter predicción de un bloque actual. En este caso, es posible reducir la cantidad de información no asociando ningún indicador de fusión a una secuencia de bits.

- 30 Debe observarse que el ejemplo descrito en la Realización 1 en el cual se usa el modo de fusión en el que se codifica un bloque actual usando una dirección de predicción, un vector de movimiento y un índice de imagen de referencia copiados de un bloque vecino del bloque actual no es limitante. Por ejemplo, se puede usar un modo de fusión con omisión (*skip merging mode*). En el modo de fusión con omisión, se codifica un bloque actual de la misma manera que en el modo de fusión, utilizando una dirección de predicción, un vector de movimiento y un índice de imagen de referencia copiados de un bloque vecino del bloque actual con referencia a una lista de candidatos a bloque de fusión creada como se muestra en (b) en la figura 13. Cuando todos los errores de predicción resultantes son cero para el bloque actual, se pone a 1 un indicador de omisión y el indicador de omisión y un índice de candidato a bloque de fusión son asociados a una secuencia de bits. Cuando cualquiera de los errores de predicción resultantes es distinto de cero, se pone a 0 un indicador de omisión y el indicador de omisión, un indicador de fusión, un índice de candidato a bloque de fusión y los errores de predicción son asociados a una secuencia de bits.

- 35 Debe observarse que el ejemplo descrito en la Realización 1 en el cual se usa el modo de fusión en el que se codifica un bloque actual utilizando una dirección de predicción, un vector de movimiento y un índice de imagen de referencia copiados de un bloque vecino del bloque actual no es limitante. Por ejemplo, se puede codificar un vector de movimiento en el modo de estimación de vector de movimiento usando una lista de candidatos a bloques de fusión creada como se muestra en (b) en la figura 13. Específicamente, se calcula una diferencia sustrayendo un vector de movimiento de un candidato a bloque de fusión indicado por un índice de candidato a bloque de fusión de un vector de movimiento en el modo de estimación de vector de movimiento. Entonces, la diferencia calculada y el índice de candidato a bloque de fusión que indica el índice de candidato a bloque de fusión pueden ser asociados a una secuencia de bits.

- 40 Debe observarse que el ejemplo descrito en la Realización 1 en el cual se usa el modo de fusión en el que se codifica un bloque actual utilizando una dirección de predicción, un vector de movimiento y un índice de imagen de referencia copiados de un bloque vecino del bloque actual no es limitante. Por ejemplo, se puede codificar un vector de movimiento en el modo de estimación de vector de movimiento usando una lista de candidatos a bloques de fusión creada como se muestra en (b) en la figura 13. Específicamente, se calcula una diferencia sustrayendo un vector de movimiento de un candidato a bloque de fusión indicado por un índice de candidato a bloque de fusión de un vector de movimiento en el modo de estimación de vector de movimiento. Entonces, la diferencia calculada y el índice de candidato a bloque de fusión que indica el índice de candidato a bloque de fusión pueden ser asociados a una secuencia de bits. El escalado se realiza usando la siguiente fórmula.

(Ecuación 2)

$$\text{scaledMV_Merge} = \text{MV_Merge} \times (\text{POC}(\text{Refldx_ME}) - \text{curPOC}) / (\text{POC}(\text{Refldx_Merge}) - \text{curPOC})$$

En este punto, POC (Refldx_ME) indica el orden de visualización de una imagen de referencia indicada por un índice de imagen de referencia Refldx_ME; POC (Refldx_Merge) denota el orden de visualización de una imagen de referencia indicada por un índice de imagen de referencia Refldx_Merge; y curPOC denota el orden de visualización de una imagen actual a codificar.

[Realización 2]

En la Realización 1, el aparato de codificación de imágenes determina una secuencia de bits para su asignación a un índice de candidato a bloque de fusión usando el número total de candidatos utilizables para la fusión incrementándolo en uno cada vez que se determina un candidato a bloque de fusión como bloque de fusión co-ubicado, con independencia de si un bloque de fusión co-ubicado es o no es un candidato utilizable para la fusión. Opcionalmente, por ejemplo, el aparato de codificación de imágenes puede determinar una secuencia de bits para su asignación a un índice de candidato a bloque de fusión utilizando el número total de candidatos utilizables para la fusión calculado incrementándolo en uno para cada candidato a bloque de fusión, con independencia de si el candidato a bloque de fusión es o no es un bloque de fusión co-ubicado en la etapa S124 de la figura 15. En otras palabras, el aparato de codificación de imágenes puede asignar una secuencia de bits a un índice de candidato a bloque de fusión usando el tamaño de una lista de candidatos a bloques de fusión fijado en un número máximo N del número total de candidatos a bloques de fusión. En otras palabras, el aparato de codificación de imágenes puede codificar índices de candidatos a bloque de fusión usando el tamaño de una lista de candidatos a bloques de fusión fijado en un valor máximo N del número total de candidatos a bloques de fusión en el supuesto de que todos los candidatos a bloques de fusión son candidatos utilizables para la fusión.

Por ejemplo, en el caso mostrado en la Realización 1, cuando el valor máximo N del número total de candidatos a bloques de fusión es cinco (el bloque vecino A, el bloque vecino B, el bloque de fusión co-ubicado, el bloque vecino C, y el bloque vecino D), el aparato de codificación de imágenes puede codificar los índices de candidatos a bloques de fusión usando el tamaño de la lista de candidatos a bloques de fusión establecido en cinco. Además, por ejemplo, cuando el valor máximo N del número total de candidatos a bloques de fusión es cuatro (el bloque vecino A, el bloque vecino B, el bloque vecino C y el bloque vecino D), el aparato de codificación de imágenes puede codificar los índices de candidatos a bloque de fusión usando el tamaño de la lista de candidatos a bloques de fusión fijado en cuatro permanentemente.

De esta manera, el aparato de codificación de imágenes puede determinar el tamaño de una lista de candidatos a bloques de fusión en base al valor máximo del número total de candidatos a bloques de fusión. Por lo tanto, es posible generar una secuencia de bits a partir de la cual una unidad de decodificación de longitud variable de un aparato de decodificación de imágenes pueda decodificar un índice de candidato a bloque de fusión sin referencia a información sobre un bloque vecino o un bloque co-ubicado, de modo que se puede reducir la complejidad computacional para la unidad de decodificación de longitud variable.

Dicha modificación del aparato de codificación de imágenes según la Realización 1 se describirá específicamente a continuación como un aparato de codificación de imágenes de acuerdo con la Realización 2.

La figura 18 es un diagrama de bloques que muestra una configuración de un aparato de codificación de imágenes 200 de acuerdo con la Realización 2. El aparato de codificación de imágenes 200 codifica una imagen bloque a bloque para generar una secuencia de bits. El aparato de codificación de imágenes 200 incluye una unidad de derivación de candidatos a fusión 210, una unidad de control de predicción 220 y una unidad de codificación 230.

La unidad de derivación de candidatos a fusión 210 corresponde a la unidad de cálculo de candidatos a bloque de fusión 114 de la Realización 1. La unidad de derivación de candidatos a fusión 210 deriva candidatos a fusión. La unidad de derivación de candidatos a fusión 210 genera una lista de candidatos a fusión en la que, por ejemplo, unos índices que identifican a cada candidato a fusión derivado diferente (en lo sucesivo denominados índices de candidatos a fusión) se asocian con los respectivos candidatos a fusión derivados.

Los candidatos a fusión son candidatos en los que cada uno de ellos es una combinación de una dirección de predicción, un vector de movimiento y un índice de imagen de referencia para uso en la codificación de un bloque actual. Específicamente, cada uno de los candidatos a fusión es una combinación que incluye al menos un conjunto de una dirección de predicción, un vector de movimiento y un índice de imagen de referencia.

Los candidatos a fusión corresponden a los candidatos a bloques de fusión de la Realización 1. La lista de candidatos a fusión corresponde a la lista de candidatos a bloques de fusión.

Como se muestra en la figura 18, la unidad de derivación de candidatos a fusión 210 incluye una primera unidad de determinación 211, una primera unidad de derivación 212, una unidad de especificación 213, una segunda unidad de determinación 214 y una segunda unidad de derivación 215.

La primera unidad de determinación 211 determina un número máximo de candidatos a fusión. En otras palabras, la primera unidad de determinación 211 determina un valor máximo N del número total de candidatos a bloques de

fusión.

Por ejemplo, la primera unidad de determinación 211 determina un número máximo de candidatos a fusión en base a características de la secuencia de imagen de entrada (tal como una secuencia, una imagen, una porción (*slice*) o un bloque). Opcionalmente, por ejemplo, la primera unidad de determinación 211 puede determinar un número

La primera unidad de derivación 212 deriva primeros candidatos a fusión. Específicamente, la primera unidad de derivación 212 deriva primeros candidatos a fusión dentro de un intervalo en el que el número total de los primeros candidatos a fusión no supera el número máximo. Más específicamente, la primera unidad de derivación 212 deriva primeros candidatos a fusión en base a, por ejemplo, una dirección de predicción, un vector de movimiento y un índice de imagen de referencia utilizados para la codificación de un bloque espacialmente o temporalmente vecino del bloque actual. Entonces, por ejemplo, la primera unidad de derivación 212 registra los primeros candidatos a fusión derivados de este modo en la lista de candidatos a fusión en asociación con los respectivos índices de candidatos a fusión.

El bloque vecino espacialmente es un bloque que está dentro de una imagen que incluye el bloque actual y es vecino del bloque actual. Específicamente, los bloques vecinos A al D mostrados en la figura 3 son ejemplos del bloque espacialmente vecino.

El bloque temporalmente vecino es un bloque que está dentro de una imagen diferente de una imagen que incluye el bloque actual y corresponde al bloque actual. Específicamente, un bloque co-ubicado es un ejemplo de bloque temporalmente vecino.

Debe observarse que el bloque temporalmente vecino no necesita ser un bloque ubicado en la misma posición que el bloque actual (bloque co-ubicado). Por ejemplo, el bloque temporalmente vecino puede ser un bloque vecino del bloque co-ubicado.

Debe observarse que la primera unidad de derivación 212 puede derivar, como primer candidato a fusión, una combinación de una dirección de predicción, un vector de movimiento y un índice de imagen de referencia utilizados para la codificación de bloques que son espacialmente vecinos del bloque actual a excepción de los bloques inutilizables para la fusión. Un bloque inutilizable para la fusión es un bloque codificado por intra predicción, un bloque fuera del límite de una porción que incluye el bloque actual o del límite de una imagen que incluye el bloque actual, o un bloque aún por codificar. Con esta configuración, la primera unidad de derivación 212 puede derivar primeros candidatos a fusión a partir de bloques apropiados para obtener candidatos a fusión.

Cuando se ha derivado una pluralidad de primeros candidatos a fusión, la unidad de especificación 213 especifica un candidato idéntico, es decir, un primer candidato a fusión que es una combinación de una dirección de predicción, un vector de movimiento y un índice de imagen de referencia idéntica a una combinación de una dirección de predicción, un vector de movimiento y un índice de imagen de referencia de cualquier otro de los primeros candidatos a fusión. Entonces, la unidad de especificación 213 elimina el candidato idéntico especificado de la lista de candidatos a fusión.

La segunda unidad de determinación 214 determina si el número total de los primeros candidatos a fusión es o no es menor que el número máximo determinado. En este punto, la segunda unidad de determinación 214 determina si el número total de los primeros candidatos a fusión, excepto el primer candidato a fusión idéntico especificado, es o no es menor que el número máximo determinado.

Cuando se determina que el número total de los primeros candidatos a fusión es menor que el número máximo determinado, la segunda unidad de derivación 215 deriva segundos candidatos a fusión. Específicamente, la segunda unidad de derivación 215 deriva segundos candidatos a fusión dentro de un intervalo en el que la suma del número total de primeros candidatos a fusión y el número total de los segundos candidatos a fusión no supera el número máximo. En este punto, la segunda unidad de derivación 215 deriva los segundos candidatos a fusión dentro de un intervalo en el que la suma del número total de primeros candidatos a fusión excepto el candidato idéntico y el número total de los segundos candidatos a fusión no supera el número máximo.

El segundo candidato a fusión corresponde al nuevo candidato en la Realización 1. Por consiguiente, la segunda unidad de derivación 215 puede derivar un segundo candidato a fusión en base a, por ejemplo, una combinación de una dirección de predicción, un vector de movimiento y un índice de imagen de referencia utilizada para la codificación de un bloque vecino diferente de los bloques vecinos utilizados para el primer candidato a fusión.

Además, por ejemplo, la segunda unidad de derivación 215 puede derivar, como segundo candidato a fusión, un candidato a fusión que es diferente en al menos uno de entre una dirección de predicción, un vector de movimiento y un índice de imagen de referencia de los que se puede derivar cualquier otro de los primeros candidatos a fusión. Esto hace posible aumentar el número total de candidatos a fusión, siendo cada uno de los cuales una combinación diferente de una dirección de predicción, un vector de movimiento y un índice de imagen de referencia, de modo que puede incrementarse aún más la eficacia de la codificación.

Debe observarse que la segunda unidad de derivación 215 no necesita derivar, como segundo candidato a fusión, un candidato a fusión diferente de los primeros candidatos a fusión. En otras palabras, la segunda unidad de derivación 215 puede derivar un segundo candidato a fusión que, como resultado, es idéntico a un primer candidato a fusión.

5 A continuación, por ejemplo, la segunda unidad de derivación 215 registra los segundos candidatos a fusión derivados de este modo en la lista de candidatos a fusión, cada uno en asociación con un índice de candidato a fusión diferente. En este momento, la segunda unidad de derivación 215 puede registrar los segundos candidatos a fusión en la lista de candidatos a fusión de modo que los índices de candidatos a fusión asignados a los primeros candidatos a fusión son más pequeños que los índices de candidatos a fusión asignados a los segundos candidatos a fusión como en la Realización 1. Con esto, el aparato de codificación de imágenes 200 puede reducir la cantidad de código cuando es más probable que se seleccione un primer candidato a fusión como candidato a fusión para uso en la codificación que un segundo candidato a fusión, de modo que se puede aumentar la eficacia de la codificación.

10 Se debe observar que la segunda unidad de derivación 215 no necesita derivar un segundo candidato a fusión para que la suma del número total de los primeros candidatos a fusión y el número total de segundos candidatos a fusión sea igual a un número máximo determinado. Cuando la suma del número total de los primeros candidatos a fusión y el número total de los segundos candidatos a fusión es menor que el número máximo determinado, por ejemplo, puede haber un índice de candidato a fusión con el que ningún candidato a fusión está asociado.

15 La unidad de control de predicción 220 selecciona un candidato a fusión para su uso en la codificación de un bloque actual de entre los primeros candidatos a fusión y los segundos candidatos a fusión. En otras palabras, la unidad de control de predicción 220 selecciona un candidato a fusión para usarlo en la codificación de un bloque actual de entre la lista de candidatos a fusión.

20 La unidad de codificación 230 codifica el índice para identificar el candidato a fusión seleccionado (índice de candidato a fusión) usando el número máximo determinado. Específicamente, la unidad de codificación 230 realiza una codificación de longitud variable en una secuencia de bits asignada al valor del índice del candidato a fusión seleccionado como se muestra en la figura 5. Además, la unidad de codificación 230 asocia el índice codificado a una secuencia de bits.

25 En este punto, la unidad de codificación 230 puede asociar además información que indica el número máximo determinado por la primera unidad de determinación 211 a la secuencia de bits. Específicamente, por ejemplo, la unidad de codificación 230 puede escribir la información que indica el número máximo en una cabecera de porción (*slice header*). Esto hace posible cambiar los números máximos por la unidad apropiada para que se pueda aumentar la eficacia de la codificación.

30 La unidad de codificación 230 no necesita asociar un número máximo a una secuencia de bits. Por ejemplo, cuando se especifica el número máximo en un estándar, o cuando el número máximo es el mismo que un valor por defecto, la unidad de codificación 230 no necesita asociar información que indique el número máximo a una secuencia de bits.

A continuación, se describirán operaciones del aparato de codificación de imágenes 200 en la configuración descrita anteriormente.

35 La figura 19 es un diagrama de flujo que muestra operaciones de procesamiento del aparato de codificación de imágenes 200 de acuerdo con la Realización 2.

40 En primer lugar, la primera unidad de determinación 211 determina un número máximo de candidatos a fusión (S201). La primera unidad de derivación 212 deriva un primer candidato a fusión (S202). Cuando se ha derivado una pluralidad de primeros candidatos a fusión, la unidad de especificación 213 especifica un primer candidato a fusión que es una combinación de una dirección de predicción, un vector de movimiento y un índice de imagen de referencia idéntica a una combinación de una dirección de predicción, un vector de movimiento, y un índice de imagen de referencia de cualquier otro de los primeros candidatos a fusión (S203).

45 La segunda unidad de determinación 214 determina si el número total de los primeros candidatos a fusión, excepto el candidato idéntico, es o no es menor que el número máximo determinado (S204). En este punto, cuando se determina que el número total de los primeros candidatos a fusión excepto el candidato idéntico es menor que el número máximo determinado (S204, Sí), la segunda unidad de derivación 215 deriva segundos candidatos a fusión (S205). Por otra parte, cuando se determina que el número total de los primeros candidatos a fusión excepto el candidato idéntico no es menor que el número máximo determinado (S204, No), la segunda unidad de derivación 215 no deriva ningún segundo candidato a fusión. Estas etapas S204 y S205 corresponden a la etapa S114 de la Realización 1.

50 La unidad de control de predicción 220 selecciona un candidato a fusión a utilizar para la codificación de un bloque actual de entre los primeros candidatos a fusión y los segundos candidatos a fusión (S206). Por ejemplo, la unidad de control de predicción 220 selecciona un candidato a fusión para el cual el coste representado por la ecuación 1 es

un mínimo de la lista de candidatos a fusión como en la Realización 1.

La unidad de codificación 230 codifica un índice para identificar el candidato a fusión seleccionado, usando el número máximo determinado (S207). Además, la unidad de codificación 230 asocia el índice codificado a una secuencia de bits.

- 5 De esta manera, el aparato de codificación de imágenes 200 de acuerdo con la Realización 2 puede codificar un índice para identificar un candidato a fusión usando un número máximo determinado. En otras palabras, se puede codificar un índice independientemente del número total de candidatos a fusión realmente derivados. Por lo tanto, incluso cuando se pierde información necesaria para derivar un candidato a fusión (por ejemplo, información sobre un bloque co-ubicado), aún se puede decodificar un índice y de este modo se mejora la resistencia a errores.
- 10 Además, se puede decodificar un índice independientemente del número total de candidatos a fusión realmente derivados. En otras palabras, se puede decodificar un índice sin esperar a la derivación de los candidatos a fusión. En otras palabras, se puede generar una secuencia de bits para la que se pueden realizar en paralelo la derivación de los candidatos a fusión y la decodificación de los índices.

- 15 Además, con el aparato de codificación de imágenes 200 de acuerdo con la realización 2, se puede derivar un segundo candidato a fusión cuando se determina que el número total de los primeros candidatos a fusión es menor que el número máximo. En consecuencia, se puede aumentar el número total de candidatos a fusión dentro de un intervalo que no supere el número máximo, de modo que se puede aumentar la eficacia de la codificación.

- 20 Además, con el aparato de codificación de imágenes 200 de acuerdo con la Realización 2, se puede derivar un segundo candidato a fusión en base al número total de primeros candidatos a fusión, excepto los primeros candidatos a fusión idénticos. Como resultado, se puede aumentar el número total de los segundos candidatos a fusión, de modo que se puede aumentar la variedad de combinaciones de una dirección de predicción, un vector de movimiento y un índice de imagen de referencia para un candidato a fusión seleccionable. Por lo tanto, es posible aumentar aún más la eficacia de la codificación.

- 25 En la Realización 2, la unidad de especificación 213 incluida en el aparato de codificación de imágenes 200 no siempre es necesaria para el aparato de codificación de imágenes 200. En otras palabras, la etapa S203 del diagrama de flujo mostrado en la figura 19 no siempre es necesaria. Incluso en tal caso, el aparato de codificación de imágenes 200 puede codificar un índice para identificar un candidato a fusión usando un número máximo determinado de modo que se puede mejorar la resistencia al error.

- 30 Además, en la Realización 2, aunque la unidad de especificación 213 especifica un candidato idéntico después de que la primera unidad de derivación 212 derive primeros candidatos a fusión como se muestra en la figura 19, el proceso no necesita realizarse en este orden. Por ejemplo, la primera unidad de derivación 212 puede identificar un candidato idéntico en el proceso de derivar los primeros candidatos a fusión, y deriva los primeros candidatos a fusión de modo que el candidato idéntico especificado es excluido de entre los primeros candidatos a fusión. En otras palabras, la primera unidad de derivación 212 puede derivar, como primer candidato a fusión, un candidato a fusión que es una combinación de una dirección de predicción, un vector de movimiento y un índice de imagen de referencia diferente de una combinación de una dirección de predicción, un movimiento vector, y un índice de imagen de referencia de cualquier primer candidato a fusión derivado anteriormente. Más específicamente, por ejemplo, en el caso en el que ya se ha seleccionado un candidato a fusión en base a un bloque vecino izquierdo como primer candidato a fusión, la primera unidad de derivación 212 puede derivar un candidato a fusión basado en un bloque vecino superior como primer candidato a fusión cuando el candidato a fusión basado en el bloque vecino superior es diferente del candidato a fusión que se basa en el bloque vecino izquierdo.
- 40

[Realización 3]

- 45 La figura 20 es un diagrama de bloques que muestra una configuración de un aparato de decodificación de imágenes 300 de acuerdo con la Realización 3. El aparato de decodificación de imágenes 300 es un dispositivo correspondiente al aparato de codificación de imágenes 100 según la Realización 1. Por ejemplo, el aparato de decodificación de imágenes 300 decodifica, bloque a bloque, imágenes codificadas incluidas en una secuencia de bits generada por el aparato de codificación de imágenes 100 de acuerdo con la Realización 1.

- 50 Como se muestra en la figura 20, el aparato de decodificación de imágenes 300 incluye una unidad de decodificación de longitud variable 301, una unidad de cuantificación inversa 302, una unidad de transformación ortogonal inversa 303, un agregador 304, una memoria de bloques 305, una memoria de fotogramas (*frames*) 306, una unidad de intra predicción 307, una unidad de inter predicción 308, una unidad de control de inter predicción 309, un conmutador 310, una unidad de cálculo de candidatos a bloque de fusión 311, y una memoria *colPic* 312.

- 55 La unidad de decodificación de longitud variable 301 genera información del tipo de imagen, un indicador de fusión y un coeficiente cuantificado realizando una decodificación de longitud variable en una secuencia de bits de entrada. Además, la unidad de decodificación de longitud variable 301 obtiene un índice de candidato a bloque de fusión realizando la decodificación de longitud variable usando el número total de candidatos utilizables para la fusión que se describe a continuación.

La unidad de cuantificación inversa 302 cuantifica de forma inversa el coeficiente cuantificado obtenido por la decodificación de longitud variable.

La unidad de transformación ortogonal inversa 303 genera datos de error de predicción transformando un coeficiente de transformación ortogonal obtenido por la cuantificación inversa desde un dominio de frecuencia a un dominio de imagen.

La memoria de bloques 305 almacena, en unidades de un bloque, datos de imagen decodificados generados agregando datos de error de predicción y datos de imagen de predicción.

La memoria de fotogramas 306 almacena datos de imagen decodificados en unidades de un fotograma (*frame*).

La unidad de intra predicción 307 genera datos de imagen de predicción de un bloque actual a decodificar, realizando una intra predicción usando los datos de imagen decodificados almacenados en la memoria de bloques 305 en unidades de un bloque.

La unidad de inter predicción 308 genera datos de imagen de predicción de un bloque actual a decodificar, realizando una inter predicción usando los datos de imagen decodificados almacenados en la memoria de fotogramas 306 en unidades de un fotograma.

Cuando se decodifica un bloque actual mediante decodificación de intra predicción, el conmutador 310 genera datos de imagen de intra predicción generados por la unidad de intra predicción 307 como datos de imagen de predicción del bloque actual para el agregador 304. Por otro lado, cuando se decodifica un bloque actual mediante decodificación de inter predicción, el conmutador 310 genera datos de imagen de inter predicción generados por la unidad de inter predicción 308 como datos de imagen de predicción del bloque actual para el agregador 304.

La unidad de cálculo de candidatos a bloque de fusión 311 deriva candidatos a bloque de fusión para el modo de fusión a partir de vectores de movimiento y otros de bloques vecinos con el bloque actual y un vector de movimiento y otros de un bloque co-ubicado (información *colPic*) almacenados en la memoria *colPic* 312, usando un procedimiento descrito más adelante. Además, la unidad de cálculo de candidatos a bloque de fusión 311 asigna índices de candidatos a bloque de fusión que tienen cada uno un valor diferente a los candidatos a bloque de fusión derivados. Entonces, la unidad de cálculo de candidatos a bloque de fusión 311 transmite los candidatos a bloque de fusión y los índices de candidatos a bloque de fusión a la unidad de control de inter predicción 309.

La unidad de control de inter predicción 309 hace que la unidad de inter predicción 308 genere una imagen de inter predicción utilizando información sobre el modo de estimación de vector de movimiento cuando el indicador de fusión decodificado es "0". Por otro lado, cuando el indicador de fusión es "1", la unidad de control de inter predicción 309 determina, en base a un índice de candidato a bloque de fusión decodificado, un vector de movimiento, un índice de imagen de referencia y una dirección de predicción para uso en la inter predicción a partir de una pluralidad de candidatos a bloques de fusión. Entonces, la unidad de control de inter predicción 309 hace que la unidad de inter predicción 308 genere una imagen de inter predicción utilizando el vector de movimiento, el índice de imagen de referencia y la dirección de predicción determinados. Además, la unidad de control de inter predicción 309 transfiere información *colPic* que incluye el vector de movimiento del bloque actual a la memoria *colPic* 312.

Finalmente, el agregador 304 genera datos de imagen decodificados agregando los datos de imagen de predicción y los datos de error de predicción.

La figura 21 es un diagrama de flujo que muestra operaciones de procesamiento del aparato de decodificación de imágenes 300 de acuerdo con la Realización 3.

En la etapa S301, la unidad de decodificación de longitud variable 301 decodifica un indicador de fusión.

En la etapa S302, cuando el indicador de fusión es "1" (S302, Sí), en la etapa S303, la unidad de cálculo de candidatos a bloque de fusión 311 calcula el número total de candidatos utilizables para la fusión utilizando un procedimiento descrito más adelante. Entonces, la unidad de cálculo de candidatos a bloque de fusión 311 determina el número calculado de candidatos utilizables para la fusión como el tamaño de una lista de candidatos a bloques de fusión.

En la etapa S304, la unidad de decodificación de longitud variable 301 realiza la decodificación de longitud variable en un índice de candidato a bloque de fusión a partir de una secuencia de bits usando el tamaño de la lista de candidatos a bloques de fusión.

En la etapa S305, la unidad de cálculo de candidatos a bloque de fusión 311 genera candidatos a bloques de fusión a partir de bloques vecinos y un bloque co-ubicado de un bloque actual a decodificar usando un procedimiento descrito más adelante.

En la etapa S306, la unidad de control de inter predicción 309 hace que la unidad de inter predicción 308 genere una imagen de inter predicción usando el vector de movimiento, el índice de imagen de referencia y la dirección de predicción del candidato a bloque de fusión indicado por el índice de candidato a bloque de fusión decodificado.

Cuando el indicador de fusión es "0" en la etapa S302 (S302, No), en la etapa S307, la unidad de inter predicción 308 genera una imagen de inter predicción usando información sobre el modo de estimación de vector de movimiento decodificado por la unidad de decodificación de longitud variable 301.

5 Opcionalmente, cuando el tamaño de una lista de candidatos a bloques de fusión que se ha calculado en la etapa S303 es "1", se puede estimar un índice de candidato a bloque de fusión igual a "0" sin ser decodificado.

La figura 22 es un diagrama de flujo que muestra detalles del proceso en la etapa S303 mostrada en la figura 21. Específicamente, la figura 22 ilustra un procedimiento para determinar si un candidato a bloque de fusión [N] es o no es un candidato utilizable para la fusión y calcular el número total de candidatos utilizables para la fusión. La figura 22 se describirá a continuación.

10 En la etapa S311, la unidad de cálculo de candidatos a bloque de fusión 311 determina si es verdadero o falso que (1) un candidato a bloque de fusión [N] ha sido decodificado por intra predicción, (2) el candidato a bloque de fusión [N] es un bloque fuera del límite de una porción que incluye el bloque actual o del límite de una imagen que incluye el bloque actual, o (3) el candidato a bloque de fusión [N] aún está por decodificar.

15 Cuando el resultado de la determinación en la etapa S311 es verdadero (S311, Si), la unidad de cálculo de candidatos a bloque de fusión 311 establece el candidato a bloque de fusión [N] como un candidato inutilizable para la fusión en la etapa S312. Por otra parte, cuando el resultado de la determinación en la etapa S311 es falso (S311, No), la unidad de cálculo de candidatos a bloque de fusión 311 establece el candidato a bloque de fusión [N] como un candidato utilizable para la fusión en la etapa S313.

20 En la etapa S314, la unidad de cálculo de candidatos a bloque de fusión 311 determina si es verdadero o falso que el candidato a bloque de fusión [N] es un candidato utilizable para la fusión o un candidato a bloque de fusión co-ubicado. En este punto, cuando el resultado de la determinación en la etapa S314 es verdadero (S314, Si), la unidad de cálculo de candidatos a bloque de fusión 311 actualiza el número total de candidatos a bloques de fusión incrementándolo en uno en la etapa S315. Por otro lado, cuando el resultado de la determinación en la etapa S314 es falso (S314, No), la unidad de cálculo de candidatos a bloque de fusión 311 no actualiza el número total de
25 candidatos utilizables para la fusión.

Por lo tanto, cuando un candidato a bloque de fusión es un bloque de fusión co-ubicado, la unidad de cálculo de candidatos a bloques de fusión 311 incrementa en uno el número total de candidatos utilizables para la fusión con independencia de si el bloque co-ubicado es un candidato utilizable para la fusión o un candidato inutilizable para la fusión. Esto evita la discrepancia de los números de candidatos utilizables para la fusión entre el aparato de
30 codificación de imágenes y el aparato de decodificación de imágenes, incluso cuando se pierde información sobre un bloque de fusión co-ubicado debido a un incidente tal como la pérdida de paquete(s).

El número total de candidatos utilizables para la fusión se establece como el tamaño de una lista de candidatos a bloques de fusión en la etapa S303 mostrada en la figura 21. Además, el tamaño de la lista de candidatos a bloques de fusión se usa en la codificación de longitud variable de índices de candidatos a bloques de fusión en la etapa
35 S304 mostrada en la figura 21. Esto hace posible que el aparato de decodificación de imágenes 300 decodifique los índices de candidatos a bloques de fusión normalmente incluso cuando se pierde información sobre la imagen de referencia que incluye un bloque co-ubicado.

La figura 23 es un diagrama de flujo que muestra detalles del proceso en la etapa S305 mostrada en la figura 21. Específicamente, la figura 23 ilustra un procedimiento para calcular un candidato a bloque de fusión. La figura 23 se describirá a continuación.
40

En la etapa S321, la unidad de cálculo de candidatos a bloque de fusión 311 obtiene el vector de movimiento, el índice de imagen de referencia y la dirección de predicción de un candidato a bloque de fusión [N], y los agrega a una lista de candidatos a bloque de fusión.

45 En la etapa S322, la unidad de cálculo de candidatos a bloque de fusión 311 busca en la lista de candidatos a bloque de fusión un candidato inutilizable para la fusión y un candidato idéntico, y elimina el candidato inutilizable para la fusión y el candidato idéntico de la lista de candidatos a bloque de fusión como se muestra en la figura 13.

En la etapa S323, la unidad de cálculo de candidatos a bloque de fusión 311 agrega un nuevo candidato a la lista de candidatos a bloque de fusión usando el procedimiento según se ilustra en la figura 16.

50 La figura 24 muestra una sintaxis de ejemplo para la asociación de índices de candidatos a bloques de fusión a una secuencia de bits. En la figura 24, merge_idx representa un índice de candidato a bloque de fusión, y merge_flag representa un indicador de fusión. NumMergeCand representa el tamaño de una lista de candidatos a bloques de fusión. En la Realización 3, NumMergeCand se establece igual al número total de candidatos utilizables para la fusión calculado en el flujo del proceso que se muestra en la figura 22.

55 Así, el aparato de decodificación de imágenes 300 según la Realización 3 es capaz de calcular el tamaño de una lista de candidatos a bloques de fusión para usarlo en la codificación o decodificación de un índice de candidato a

bloque de fusión, usando un procedimiento independiente de la información sobre imágenes de referencia que incluyen un bloque co-ubicado. El aparato de decodificación de imágenes 300 por lo tanto puede decodificar de forma apropiada una secuencia de bits con una resistencia a errores mejorada.

Más específicamente, con independencia de si un bloque de fusión co-ubicado es o no es un candidato utilizable para la fusión, el aparato de decodificación de imágenes 300 de acuerdo con la Realización 3 incrementa en uno el número total de candidatos utilizables para la fusión cada vez que se determina un candidato a bloque de fusión como un bloque de fusión co-ubicado. Entonces, el aparato de decodificación de imágenes 300 determina una secuencia de bits asignada a un índice de candidato a bloque de fusión usando el número total de candidatos utilizables para la fusión calculado de la manera descrita. Esto hace posible que el aparato de decodificación de imágenes 300 decodifique índices de candidatos a bloques de fusión normalmente incluso cuando se pierde información sobre la imagen de referencia que incluye un bloque co-ubicado.

Además, cuando el número total de candidatos a bloques de fusión es menor que el número total de candidatos utilizables para la fusión, es posible que el aparato de decodificación de imágenes 300 según la Realización 3 decodifique de forma apropiada una secuencia de bits codificada con una eficacia de codificación aumentada agregando un nuevo candidato que tiene un nuevo vector de movimiento, un nuevo índice de imagen de referencia y una nueva dirección de predicción.

[Realización 4]

La Realización 4 es una realización de la invención. En la Realización 3, el aparato de decodificación de imágenes determina una secuencia de bits para su asignación a un índice de candidato a bloque de fusión usando el número total de candidatos utilizables para la fusión incrementado en uno cada vez que se determina un candidato a bloque de fusión como bloque de fusión co-ubicado, con independencia de si un bloque de fusión co-ubicado es o no es un candidato utilizable para la fusión. Opcionalmente, por ejemplo, el aparato de decodificación de imágenes puede determinar una secuencia de bits para su asignación a un índice de candidato a bloque de fusión usando el número total de candidatos utilizables para la fusión calculado incrementándolo en uno para cada candidato a bloque de fusión con independencia de si el candidato a bloque de fusión es o no es un bloque de fusión co-ubicado en la etapa S314 de la figura 22. En otras palabras, el aparato de decodificación de imágenes puede asignar una secuencia de bits a un índice de candidato a bloque de fusión usando el tamaño de una lista de candidatos a bloques de fusión establecido en un número máximo N del número total de candidatos a bloques de fusión. En otras palabras, el aparato de decodificación de imágenes puede decodificar índices de candidatos a bloques de fusión usando el tamaño de una lista de candidatos a bloques de fusión establecido en un valor máximo N del número total de candidatos a bloques de fusión asumiendo que todos los candidatos a bloques de fusión son candidatos utilizables para la fusión.

Por ejemplo, en el caso mostrado en la Realización 3, cuando el valor máximo N del número total de candidatos a bloques de fusión es cinco (el bloque vecino A, el bloque vecino B, el bloque de fusión co-ubicado, el bloque vecino C y el bloque vecino D), el aparato de decodificación de imágenes puede decodificar los índices de candidatos a bloques de fusión usando el tamaño de la lista de candidatos a bloques de fusión establecido en cinco. Por lo tanto, es posible que la unidad de decodificación de longitud variable del aparato de decodificación de imágenes decodifique un índice de candidato a bloque de fusión a partir de una secuencia de bits sin hacer referencia a información sobre un bloque vecino o un bloque co-ubicado. Como resultado, por ejemplo, la etapa S314 y la etapa S315 mostradas en la figura 22 se pueden omitir para que se pueda reducir la complejidad computacional de la unidad de decodificación de longitud variable.

La figura 25 muestra una sintaxis de ejemplo en el caso en el que el tamaño de una lista de candidatos a bloques de fusión se ha fijado en el valor máximo del número total de candidatos a bloques de fusión. Como se puede ver en la figura 25, el NumMergeCand puede ser omitido de la sintaxis cuando el tamaño de una lista de candidatos a bloques de fusión se ha fijado en el valor máximo del número total de candidatos a bloques de fusión.

Dicha modificación del aparato de decodificación de imágenes según la Realización 3 se describirá específicamente a continuación como un aparato de decodificación de imágenes de acuerdo con la Realización 4.

La figura 26 es un diagrama de bloques que muestra una configuración de un aparato de decodificación de imágenes 400 de acuerdo con la Realización 4. El aparato de decodificación de imágenes 400 es un dispositivo que corresponde al aparato de codificación de imágenes 200 según la Realización 2. Específicamente, por ejemplo, el aparato de decodificación de imágenes 400 decodifica, bloque a bloque, imágenes codificadas incluidas en una secuencia de bits generada por el aparato de codificación de imágenes 200 de acuerdo con la realización 2. El aparato de decodificación de imágenes 400 incluye una unidad de derivación de candidatos a fusión 410, una unidad de decodificación 420 y una unidad de control de predicción 430.

La unidad de derivación de candidatos a fusión 410 corresponde a la unidad de cálculo de candidatos a bloque de fusión 311 de la Realización 3. La unidad de derivación de candidatos a fusión 410 deriva candidatos a fusión. La unidad de derivación de candidatos a fusión 410 genera una lista de candidatos a fusión en la que, por ejemplo, los índices que identifican a los diferentes candidatos a fusión derivados (índices de candidatos a fusión) son asociados

con los respectivos candidatos a fusión derivados.

Como se muestra en la figura 26, la unidad de derivación de candidatos a fusión 410 incluye una primera unidad de determinación 411, una primera unidad de derivación 412, una unidad de especificación 413, una segunda unidad de determinación 414 y una segunda unidad de derivación 415.

- 5 La primera unidad de determinación 411 determina un número máximo de candidatos a fusión. En otras palabras, la primera unidad de determinación 211 determina un valor máximo N del número total de candidatos a bloques de fusión.

Por ejemplo, la primera unidad de determinación 411 determina un número máximo de candidatos a fusión usando el mismo procedimiento usado por la primera unidad de determinación 211 de la realización 2.

- 10 La primera unidad de determinación 411 determina un número máximo en base a la información asociada a una secuencia de bits y que indica un número máximo.

En este punto, aunque la primera unidad de determinación 411 está incluida en la unidad de derivación de candidatos a fusión 410, la segunda unidad de determinación 411 puede incluirse en la unidad de decodificación 420.

- 15 La primera unidad de derivación 412 deriva primeros candidatos a fusión. Específicamente, la primera unidad de derivación 412 deriva los primeros candidatos a fusión de la misma manera que la primera unidad de derivación 212 de la Realización 2. Por ejemplo, la primera unidad de derivación 412 deriva los primeros candidatos a fusión dentro de un intervalo en el que el número total de los primeros candidatos a fusión no supera el número máximo. Más específicamente, la primera unidad de derivación 412 deriva los primeros candidatos a fusión en base a, por
20 ejemplo, una dirección de predicción, un vector de movimiento y un índice de imagen de referencia utilizados para la decodificación de un bloque espacialmente o temporalmente vecino a un bloque actual a decodificar. Entonces, por ejemplo, la primera unidad de derivación 412 registra los primeros candidatos a fusión derivados de este modo en la lista de candidatos a fusión en asociación con los respectivos índices de candidatos a fusión.

- 25 Debe observarse que la primera unidad de derivación 412 puede derivar, como un primer candidato a fusión, una combinación de una dirección de predicción, un vector de movimiento y un índice de imagen de referencia usada para decodificar bloques que son espacialmente vecinos del bloque actual excepto los bloques inutilizables para la fusión. Con esta configuración, la primera unidad de derivación 412 puede derivar primeros candidatos a bloques de fusión apropiados para obtener candidatos a fusión.

- 30 Cuando se ha derivado una pluralidad de primeros candidatos a fusión, la unidad de especificación 413 especifica un candidato idéntico, es decir, un primer candidato a fusión que es una combinación de una dirección de predicción, un vector de movimiento y un índice de imagen de referencia idéntica a una combinación de una dirección de predicción, un vector de movimiento y un índice de imagen de referencia de cualquier otro de los primeros candidatos a fusión. A continuación, la unidad de especificación 413 elimina el candidato idéntico especificado de la lista de candidatos a fusión.

- 35 La segunda unidad de determinación 414 determina si el número total de los primeros candidatos a fusión es o no es menor que un número máximo determinado. En este punto, la segunda unidad de determinación 414 determina si el número total de los primeros candidatos a fusión, excepto el primer candidato a fusión idéntico especificado, es o no es mayor que el número máximo determinado.

- 40 Cuando se determina que el número total de los primeros candidatos a fusión es menor que el número máximo determinado, la segunda unidad de derivación 415 deriva segundos candidatos a fusión. Específicamente, la segunda unidad de derivación 415 deriva los segundos candidatos a fusión de la misma manera que la segunda unidad de derivación 215 de la Realización 2.

- 45 Por ejemplo, la segunda unidad de derivación 415 puede derivar, como un segundo candidato a fusión, un candidato a fusión que es diferente en al menos uno de entre la dirección de predicción, el vector de movimiento y el índice de imagen de referencia de cada uno de los primeros candidatos a fusión. Esto hace posible aumentar el número total de candidatos a fusión, cada uno de los cuales es una combinación diferente de una dirección de predicción, un vector de movimiento y un índice de imagen de referencia, de modo que se puede decodificar de forma apropiada una secuencia de bits codificada con mayor eficacia de la codificación.

- 50 Entonces, por ejemplo, la segunda unidad de derivación 415 registra los segundos candidatos a fusión derivados de este modo en la lista de candidatos a fusión en asociación con los respectivos índices de candidatos a fusión de la misma manera que la segunda unidad de derivación 215 de la Realización 2.

La unidad de decodificación 420 decodifica un índice codificado y asociado a una secuencia de bits, que es un índice para identificar un candidato a fusión, usando el número máximo determinado.

La unidad de control de predicción 430 selecciona, en base al índice decodificado, un candidato a fusión para usarlo

en la decodificación de un bloque actual de entre los primeros candidatos a fusión y los segundos candidatos a fusión. En otras palabras, la unidad de control de predicción 430 selecciona un candidato a fusión para usarlo en la decodificación de un bloque actual de la lista de candidatos a fusión.

5 A continuación, las operaciones del aparato de decodificación de imágenes 400 en la configuración descrita anteriormente se explicarán a continuación.

La figura 27 es un diagrama de flujo que muestra operaciones de procesamiento del aparato de decodificación de imágenes 400 de acuerdo con la Realización 4.

10 Primero, la primera unidad de determinación 411 determina un número máximo de candidatos a fusión (S401). La primera unidad de derivación 412 deriva un primer candidato a fusión (S402). Cuando se ha derivado una pluralidad de primeros candidatos a fusión, la unidad de especificación 413 especifica un primer candidato a fusión que es una combinación de una dirección de predicción, un vector de movimiento y un índice de imagen de referencia idéntica a una combinación de una dirección de predicción, un vector de movimiento, y un índice de imagen de referencia de cualquier otro de los primeros candidatos a fusión (S403).

15 La segunda unidad de determinación 414 determina si el número total de los primeros candidatos a fusión, excepto el candidato idéntico, es o no es menor que el número máximo determinado (S404). En este punto, cuando se determina que el número total de los primeros candidatos a fusión, excepto el candidato idéntico, es menor que el número máximo determinado (S404, Si), la segunda unidad de derivación 415 deriva los segundos candidatos a fusión (S405). Por otro lado, cuando se determina que el número total de los primeros candidatos a fusión excepto el candidato idéntico no es menor que el número máximo determinado (S404, No), la segunda unidad de derivación 20 415 no deriva ningún segundo candidato a fusión.

La unidad de decodificación 420 decodifica un índice codificado y asociado a una secuencia de bits, que es un índice para identificar un candidato a fusión, usando el número máximo determinado (S406).

25 La unidad de control de predicción 430 selecciona, en base al índice decodificado, un candidato a fusión para usarlo en la decodificación de un bloque actual de entre los primeros candidatos a fusión y los segundos candidatos a fusión (S407). Por ejemplo, la unidad de control de predicción 430 selecciona un candidato a fusión para el cual el coste representado por la ecuación 1 es un mínimo de la lista de candidatos a fusión como en la realización 1.

30 Aunque el proceso se realiza de manera que la decodificación de un índice (S406) se realiza después de derivar un candidato a fusión, no es necesario realizar el proceso en este orden. Por ejemplo, se puede derivar un candidato a fusión (S402 a S405) después de decodificar un índice (S406). Opcionalmente, la decodificación de un índice (S406) y la derivación de un candidato a fusión (S402 a S405) pueden realizarse en paralelo. Esto aumenta la velocidad de procesamiento para la decodificación.

35 De esta manera, el aparato de decodificación de imágenes 400 de acuerdo con la Realización 4 puede decodificar un índice para identificar un candidato a fusión, usando un número máximo determinado. En otras palabras, se puede decodificar un índice independientemente del número total de candidatos a fusión realmente derivados. Por lo tanto, incluso cuando se pierde información necesaria para derivar un candidato a fusión (por ejemplo, información sobre un bloque co-ubicado), aún se puede decodificar un índice y de este modo se mejora la resistencia a errores. Además, se puede decodificar un índice sin esperar a la derivación de los candidatos a fusión, de modo que la derivación de los candidatos a fusión y la decodificación de los índices se pueden realizar en paralelo.

40 Además, el aparato de decodificación de imágenes 400 de acuerdo con la Realización 4 puede derivar un segundo candidato a fusión cuando se determina que el número total de los primeros candidatos a fusión es menor que un número máximo. En consecuencia, se puede aumentar el número total de candidatos a fusión dentro de un intervalo que no supere el número máximo de modo que se puede decodificar de forma apropiada una secuencia de bits codificada con una mayor eficacia de la codificación.

45 Además, el aparato de decodificación de imágenes 400 de acuerdo con la Realización 4 puede derivar un segundo candidato a fusión en base al número total de primeros candidatos a fusión, excepto los primeros candidatos a fusión idénticos. Como resultado, se puede aumentar el número total de los segundos candidatos a fusión, de modo que se puede aumentar la variedad de combinaciones de una dirección de predicción, un vector de movimiento y un índice de imagen de referencia para un candidato a fusión seleccionable. Por lo tanto, es posible decodificar de forma apropiada una secuencia de bits codificada con una mayor eficacia de la codificación.

50 Como en la Realización 2, la unidad de especificación 413 incluida en el aparato de decodificación de imágenes 400 no siempre es necesaria para el aparato de decodificación de imágenes 400 en la Realización 4. En otras palabras, la etapa S403 del diagrama de flujo mostrado en la figura 27 no siempre es necesaria. Incluso en tal caso, el aparato de decodificación de imágenes 400 puede decodificar un índice para identificar un candidato a fusión usando un número máximo determinado de modo que se puede mejorar la resistencia a errores.

55 Además, en la Realización 4, aunque la unidad de especificación 413 especifica un candidato idéntico después de que la primera unidad de derivación 412 derive los primeros candidatos a fusión como se muestra en la figura 27, el

proceso no necesita realizarse en este orden. Por ejemplo, la primera unidad de derivación 412 puede derivar, como primer candidato a fusión, un candidato a fusión que es una combinación de una dirección de predicción, un vector de movimiento y un índice de imagen de referencia diferente de una combinación de una dirección de predicción, un vector de movimiento, y un índice de imagen de referencia de cualquier primer candidato a fusión derivado anteriormente.

Aunque el aparato de codificación de imágenes y el aparato de decodificación de imágenes según uno o más aspectos de la presente invención se han descrito en base a las realizaciones, la presente invención no se limita a estas realizaciones. Los expertos en la técnica apreciarán fácilmente que son posibles muchas modificaciones de las realizaciones de ejemplo o realizaciones en las que se combinan los elementos constituyentes de las realizaciones de ejemplo sin apartarse materialmente de las nuevas enseñanzas y ventajas descritas en la presente invención. Todas estas modificaciones y realizaciones se encuentran también dentro del alcance de uno o más aspectos de la presente invención.

En las realizaciones de ejemplo, cada uno de los elementos constituyentes puede implementarse como una pieza de hardware dedicado o puede implementarse ejecutando un programa de software apropiado para el elemento constituyente. Los elementos constituyentes pueden implementarse mediante una unidad de ejecución de programas tal como una CPU o un procesador que lee y ejecuta un programa de software grabado en un medio de grabación tal como un disco duro o una memoria semiconductora. En este punto, ejemplos del programa de software que implementa el aparato de codificación de imágenes o el aparato de decodificación de imágenes en las realizaciones incluyen un programa como se describe a continuación.

Específicamente, el programa hace que un ordenador ejecute un procedimiento que es un procedimiento de codificación de imágenes para codificar una imagen bloque a bloque para generar una secuencia de bits, y el procedimiento incluye: determinar un número máximo de un candidato a fusión que es una combinación de una dirección de predicción, un vector de movimiento y un índice de imagen de referencia para su uso en la codificación de un bloque actual; derivar un primer candidato a fusión; determinar si un número total del primer candidato a fusión es o no es menor que el número máximo; derivar un segundo candidato a fusión cuando se determina que el número total del primer candidato a fusión es menor que el número máximo; seleccionar un candidato a fusión para su uso en la codificación del bloque actual de entre el primer candidato a fusión y el segundo candidato a fusión; y codificar, usando el número máximo determinado, un índice para identificar el candidato a fusión seleccionado, y asociar el índice codificado a la secuencia de bits.

Además, el programa hace que un ordenador ejecute un procedimiento de decodificación de imágenes para decodificar, bloque a bloque, una imagen codificada incluida en una secuencia de bits, y el procedimiento incluye: determinar un número máximo de un candidato a fusión que es una combinación de una dirección de predicción, un vector de movimiento y un índice de imagen de referencia para uso en la decodificación de un bloque actual; derivar un primer candidato a fusión; determinar si el número total del primer candidato a fusión es o no es menor que el número máximo; derivar un segundo candidato a fusión cuando se determina que el número total del primer candidato a fusión es menor que el número máximo; decodificar un índice codificado y asociado a la secuencia de bits, usando el número máximo determinado, siendo el índice un índice para identificar un candidato a fusión; y seleccionar, en base al índice decodificado, un candidato a fusión para uso en la decodificación de un bloque actual, seleccionándose el candidato a fusión seleccionado de entre el primer candidato a fusión y el segundo candidato a fusión.

[Realización 5]

El procesamiento descrito en cada una de las realizaciones puede implementarse simplemente en un sistema informático independiente, grabando, en un medio de grabación, un programa para implementar las configuraciones del procedimiento de codificación de imágenes en movimiento (procedimiento de codificación de imágenes) y del procedimiento de decodificación de imágenes en movimiento (procedimiento de decodificación de imágenes) descritas en cada una de las realizaciones. Los medios de grabación pueden ser cualquier medio de grabación, siempre que el programa se pueda grabar, como un disco magnético, un disco óptico, un disco óptico magnético, una tarjeta IC y una memoria semiconductora.

En lo sucesivo, se describirán las aplicaciones al procedimiento de codificación de imágenes en movimiento (procedimiento de codificación de imágenes) y al procedimiento de decodificación de imágenes en movimiento (procedimiento de decodificación de imágenes) descritos en cada una de las realizaciones y sistemas que las usan. El sistema tiene la característica de tener un aparato de codificación y decodificación de imágenes que incluye un aparato de codificación de imágenes que usa el procedimiento de codificación de imágenes y un aparato de decodificación de imágenes que usa el procedimiento de decodificación de imágenes. Se pueden modificar otras configuraciones en el sistema según corresponda en función de los casos.

La figura 28 ilustra una configuración general de un sistema proveedor de contenidos ex100 para implementar servicios de distribución de contenidos. El área para proporcionar servicios de comunicación se divide en celdas de un tamaño deseado, y en cada una de las celdas se ubican unas estaciones base ex106, ex107, ex108, ex109 y ex110 que son estaciones inalámbricas fijas.

El sistema proveedor de contenidos ex100 se conecta a dispositivos, como un ordenador ex111, un asistente digital personal (PDA) ex112, una cámara ex113, un teléfono móvil ex114 y una máquina de juegos ex115, a través de Internet ex101, un proveedor de servicios de Internet ex102, una red telefónica ex104, así como las estaciones base de la ex106 a la ex110, respectivamente.

5 Sin embargo, la configuración del sistema proveedor de contenidos ex100 no se limita a la configuración que se muestra en la figura 28, y es aceptable una combinación en la que se conectan cualquiera de los elementos. Además, cada dispositivo puede conectarse directamente a la red telefónica ex104, en lugar de a través de las estaciones base de la ex106 a la ex110 que son las estaciones inalámbricas fijas. Además, los dispositivos pueden interconectarse entre sí a través de una comunicación inalámbrica de corta distancia y otras.

10 La cámara ex113, tal como una cámara de video digital, es capaz de capturar video. Una cámara ex116, tal como una cámara digital, es capaz de capturar imágenes fijas y videos. Además, el teléfono móvil ex114 puede ser el que cumple con cualquiera de los estándares tales como Sistema Global para Comunicaciones Móviles (GSM) (marca registrada), Acceso Múltiple por división de Código (CDMA), Acceso Múltiple por división de Código de Banda Ancha (W-CDMA), Evolución a Largo Plazo (LTE) y Acceso a Paquetes de Alta Velocidad (HSPA). Alternativamente, el
15 teléfono móvil ex114 puede ser un sistema de teléfono móvil personal (PHS).

En el sistema proveedor de contenidos ex100, un servidor de transmisión (*streaming*) ex103 se conecta a la cámara ex113 y otros a través de la red telefónica ex104 y la estación base ex109, lo que permite la distribución de imágenes de un espectáculo en vivo y otros. En dicha distribución, un contenido (por ejemplo, video de un espectáculo de música en vivo) capturado por el usuario usando la cámara ex113 se codifica como se describió
20 anteriormente en cada una de las realizaciones (es decir, la cámara funciona como el aparato de codificación de imágenes según un aspecto de la presente invención), y el contenido codificado se transmite al servidor de transmisión ex103. Por otro lado, el servidor de transmisión ex103 realiza la distribución en secuencia (*stream*) de los datos de contenido transmitido a los clientes a petición de los mismos. Los clientes incluyen el ordenador ex111, la PDA ex112, la cámara ex113, el teléfono móvil ex114 y la máquina de juegos ex115 que son capaces de
25 decodificar los datos codificados mencionados anteriormente. Cada uno de los dispositivos que ha recibido los datos distribuidos, decodifica y reproduce los datos codificados (es decir, funciona como el aparato de decodificación de imágenes de acuerdo con un aspecto de la presente invención).

Los datos capturados pueden ser codificados por la cámara ex113 o el servidor de transmisión ex103 que transmite los datos, o los procesos de codificación pueden ser compartidos entre la cámara ex113 y el servidor de transmisión ex103. De forma similar, los datos distribuidos pueden ser decodificados por los clientes o el servidor de transmisión ex103, o los procesos de decodificación pueden ser compartidos entre los clientes y el servidor de transmisión ex103. Además, los datos de las imágenes fijas y los videos capturados no solo por la cámara ex113 sino también por la cámara ex116 pueden ser transmitidos al servidor de transmisión ex103 a través del ordenador ex111. Los procesos de codificación pueden ser realizados por la cámara ex116, el ordenador ex111, o el servidor de
30 transmisión ex103, o compartidos entre ellos.

Además, los procesos de codificación y decodificación pueden ser realizados por un LSI ex500 generalmente incluido en cada uno de los ordenadores ex111 y los dispositivos. El LSI ex500 puede configurarse en un solo chip o una pluralidad de chips. Software para codificar y decodificar video puede integrarse en algún tipo de medio de grabación (como un CD-ROM, un disco flexible y un disco duro) que sea legible por el ordenador ex111 y otros, y los
40 procesos de codificación y decodificación pueden realizarse utilizando el software. Además, cuando el teléfono móvil ex114 está equipado con una cámara, pueden transmitirse los datos de video obtenidos por la cámara. Los datos de video son datos codificados por el LSI ex500 incluido en el teléfono móvil ex114.

Además, el servidor de transmisión ex103 puede estar compuesto por servidores y ordenadores, y puede descentralizar datos y procesar los datos descentralizados, grabar o distribuir datos.

45 Como se describió anteriormente, los clientes pueden recibir y reproducir los datos codificados en el sistema proveedor de contenidos ex100. En otras palabras, los clientes pueden recibir y decodificar información transmitida por el usuario, y reproducir los datos decodificados en tiempo real en el sistema proveedor de contenidos ex100, de modo que el usuario que no tiene ningún derecho y equipo particular puede realizar la difusión personal.

Aparte del ejemplo del sistema proveedor de contenido ex100, al menos uno de entre el aparato de codificación de imágenes en movimiento (aparato de codificación de imágenes) y el aparato de decodificación de imágenes en movimiento (aparato de decodificación de imágenes) descritos en cada una de las realizaciones puede implementarse en un sistema de difusión digital ex200 ilustrado en la figura 29. Más específicamente, una estación de difusión ex201 comunica o transmite, a través de ondas de radio, a un satélite de difusión ex202, datos multiplexados obtenidos multiplexando datos de audio y otros sobre datos de video. Los datos de video son datos
50 codificados por el procedimiento de codificación de imágenes en movimiento descrito en cada una de las realizaciones (es decir, datos codificados por el aparato de codificación de imágenes de acuerdo con un aspecto de la presente invención). Al recibir los datos multiplexados, el satélite de difusión ex202 transmite ondas de radio para hacer la difusión. A continuación, una antena de uso doméstico ex204 con una función de recepción de difusión satelital recibe las ondas de radio. A continuación, un dispositivo tal como un televisor (receptor) ex300 y un
55

decodificador *set top box* (STB) ex217 decodifica los datos multiplexados recibidos, y reproduce los datos decodificados (es decir, funciona como el aparato de decodificación de imágenes según un aspecto de la presente invención).

Además, un lector/grabador ex218 (i) lee y decodifica los datos multiplexados grabados en un medio de grabación ex215, tal como un DVD y un BD, o (i) codifica señales de video en el medio de grabación ex215, y en algunos casos, escribe los datos obtenidos multiplexando una señal de audio en los datos codificados. El lector/grabador ex218 puede incluir el aparato de decodificación de imágenes en movimiento o el aparato de codificación de imágenes en movimiento como se muestra en cada una de las realizaciones. En este caso, las señales de video reproducidas se muestran en el monitor ex219, y pueden ser reproducidas por otro dispositivo o sistema utilizando el medio de grabación ex215 en el que se graban los datos multiplexados. También es posible implementar el aparato de decodificación de imágenes en movimiento en el decodificador *set top box* ex217 conectado al cable ex203 para un televisor por cable o a la antena ex204 para difusión por satélite y/o terrestre, para visualizar las señales de video en el monitor ex219 del televisor ex300. El aparato de decodificación de imágenes en movimiento puede no implementarse en el decodificador *set top box*, sino en el televisor ex300.

La figura 30 ilustra la televisión (receptor) ex300 que usa el procedimiento de codificación de imágenes en movimiento y el procedimiento de decodificación de imágenes en movimiento descrito en cada una de las realizaciones. El televisor ex300 incluye: un sintonizador ex301 que obtiene o proporciona datos multiplexados obtenidos multiplexando datos de audio sobre datos de video, a través de la antena ex204 o el cable ex203, etc. que recibe una difusión; una unidad de modulación/demodulación ex302 que demodula los datos multiplexados recibidos o modula los datos en datos multiplexados para su suministro al exterior; y una unidad de multiplexación/demultiplexación ex303 que demultiplexa los datos multiplexados modulados en datos de video y datos de audio, o multiplexa datos de video y datos de audio codificados por una unidad de procesamiento de señales ex306 en datos.

El televisor ex300 incluye además: una unidad de procesamiento de señales ex306 que incluye una unidad de procesamiento de señales de audio ex304 y una unidad de procesamiento de señales de video ex305 que decodifican datos de audio y datos de video y codifican datos de audio y datos de video, respectivamente (funcionando como el aparato de codificación de imágenes y el aparato de decodificación de imágenes según los aspectos de la presente invención); y una unidad de salida ex309 que incluye un altavoz ex307 que proporciona la señal de audio decodificada, y una unidad de visualización ex308 que visualiza la señal de video decodificada, tal como una pantalla. Además, el televisor ex300 incluye una unidad de interfaz ex317 que incluye una unidad de entrada de operación ex312 que recibe una entrada de una operación de usuario. Además, la televisión ex300 incluye una unidad de control ex310 que controla en general cada elemento constituyente de la televisión ex300, y una unidad de circuito de suministro de energía ex311 que suministra energía a cada uno de los elementos. Además de la unidad de entrada de operación ex312, la unidad de interfaz ex317 puede incluir: un puente (*bridge*) ex313 que está conectado a un dispositivo externo, tal como el lector/grabador ex218; una unidad de ranura (*slot unit*) ex314 para permitir el acoplamiento del medio de grabación ex216, tal como una tarjeta SD; un controlador ex315 conectable a un medio de grabación externo, como un disco duro; y un módem ex316 conectable a una red telefónica. En este punto, el medio de grabación ex216 puede grabar información eléctricamente usando un elemento de memoria semiconductor no volátil/volátil para almacenamiento. Los elementos constituyentes de la televisión ex300 están conectados entre sí a través de un bus síncrono.

En primer lugar, se describirá la configuración en la que el televisor ex300 decodifica datos multiplexados obtenidos del exterior a través de la antena ex204 y otros y reproduce los datos decodificados. En el televisor ex300, tras una operación de usuario a través de un controlador remoto ex220 y otros, la unidad de multiplexación/demultiplexación ex303 demultiplexa los datos multiplexados demodulados por la unidad de modulación/demodulación ex302, bajo el control de la unidad de control ex310 que incluye una CPU. Además, la unidad de procesamiento de señales de audio ex304 decodifica los datos de audio demultiplexados, y la unidad de procesamiento de señales de video ex305 decodifica los datos de video demultiplexados, usando el procedimiento de decodificación descrito en cada una de las realizaciones, en el televisor ex300. La unidad de salida ex309 proporciona la señal de video y la señal de audio decodificadas al exterior, respectivamente. Cuando la unidad de salida ex309 proporciona la señal de video y la señal de audio, las señales pueden almacenarse temporalmente en búferes ex318 y ex319, y otras de manera que las señales se reproducen en sincronización entre sí. Además, el televisor ex300 puede leer datos multiplexados no a través de una difusión y otros, sino de los medios de grabación ex215 y ex216, tales como un disco magnético, un disco óptico y una tarjeta SD. A continuación, se describirá una configuración en la que el televisor ex300 codifica una señal de audio y una señal de video, y transmite los datos al exterior o escribe los datos en un medio de grabación. En el televisor ex300, tras una operación de usuario a través del controlador remoto ex220 y otros, la unidad de procesamiento de señales de audio ex304 codifica una señal de audio, y la unidad de procesamiento de señales de video ex305 codifica una señal de video, bajo control de la unidad de control ex310 usando el procedimiento de codificación descrito en cada una de las realizaciones. La unidad de multiplexación/demultiplexación ex303 multiplexa la señal de video y la señal de audio codificadas, y proporciona la señal resultante al exterior. Cuando la unidad de multiplexación/demultiplexación ex303 multiplexa la señal de video y la señal de audio, las señales se pueden almacenar temporalmente en los búferes ex320 y ex321, y otras de manera que las señales se reproducen en sincronización entre sí. En este punto, los búferes ex318, ex319, ex320 y ex321 pueden ser diversos según se ilustra, o al menos un búfer puede ser compartido en la televisión ex300.

Además, los datos pueden almacenarse en un búfer, de modo que se puede evitar el sobre-desbordamiento (*overflow*) y sub-desbordamiento (*underflow*) del sistema entre la unidad de modulación/demodulación ex302 y la unidad de multiplexación/demultiplexación ex303, por ejemplo.

5 Además, el televisor ex300 puede incluir una configuración para recibir una entrada AV de un micrófono o una cámara distinta de la configuración para obtener datos de audio y video de una difusión o un medio de grabación, y puede codificar los datos obtenidos. Aunque el televisor ex300 puede codificar, multiplexar y proporcionar datos al exterior en la descripción, puede ser capaz sólo de recibir, decodificar y proporcionar datos al exterior, pero no la codificación, multiplexación y suministro de datos al exterior.

10 Además, cuando el lector/grabador ex218 lee o escribe datos multiplexados de o en un medio de grabación, uno de entre el televisor ex300 y el lector/grabador ex218 pueden decodificar o codificar los datos multiplexados, y el televisor ex300 y el lector/grabador ex218 pueden compartir la decodificación o codificación.

15 Como un ejemplo, la figura 31 ilustra una configuración de una unidad de reproducción/grabación de información ex400 cuando los datos son leídos o escritos de o en un disco óptico. La unidad de reproducción/grabación de información ex400 incluye los elementos constituyentes ex401, ex402, ex403, ex404, ex405, ex406 y ex407 que se describirán a continuación. El cabezal óptico ex401 irradia un punto láser en una superficie de grabación del medio de grabación ex215 que es un disco óptico para escribir información, y detecta la luz reflejada en la superficie de grabación del medio de grabación ex215 para leer la información. La unidad de grabación de modulación ex402 controla eléctricamente un láser semiconductor incluido en el cabezal óptico ex401, y modula la luz láser de acuerdo con los datos grabados. La unidad de demodulación de reproducción ex403 amplifica una señal de reproducción obtenida detectando eléctricamente la luz reflejada en la superficie de grabación usando un foto detector incluido en el cabezal óptico ex401, y demodula la señal de reproducción separando un componente de señal grabado en el medio de grabación ex215 para reproducir la información necesaria. El búfer ex404 contiene temporalmente la información a grabar en el medio de grabación ex215 y la información reproducida del medio de grabación ex215. El motor de disco ex405 hace girar el medio de grabación ex215. La unidad de servo control ex406 mueve el cabezal óptico ex401 a una pista de información predeterminada mientras controla la rotación del motor de disco ex405 para seguir el punto láser. La unidad de control del sistema ex407 controla en general la unidad de reproducción/grabación de información ex400. Los procesos de lectura y escritura pueden ser implementados por la unidad de control del sistema ex407 utilizando información almacenada en el búfer ex404 y generando y agregando nueva información según sea necesario, y por la unidad de grabación de modulación ex402, la unidad de demodulación de reproducción ex403 y la unidad de servo control ex406 que graba y reproduce información a través del cabezal óptico ex401 durante una operación coordinada. La unidad de control del sistema ex407 incluye, por ejemplo, un microprocesador, y ejecuta el procesamiento haciendo que un ordenador ejecute un programa para lectura y escritura.

25 Aunque el cabezal óptico ex401 irradia un punto láser en la descripción, puede realizar una grabación de alta densidad utilizando luz de campo cercano.

30 La figura 32 ilustra el medio de grabación ex215 que es el disco óptico. En la superficie de grabación del medio de grabación ex215, hay unas ranuras de guía en forma de espiral, y una pista de información ex230 graba, por adelantado, información de dirección que indica una posición absoluta en el disco según el cambio en la forma de las ranuras de guía. La información de dirección incluye información para determinar posiciones de bloques de grabación ex231 que son una unidad para grabación de datos. La reproducción de la pista de información ex230 y la lectura de la información de dirección en un aparato que graba y reproduce datos puede causar la determinación de las posiciones de los bloques de grabación. Además, el medio de grabación ex215 incluye un área de grabación de datos ex233, un área de circunferencia interna ex232 y un área de circunferencia externa ex234. El área de grabación de datos ex233 es un área para uso en la grabación de los datos del usuario. El área de circunferencia interna ex232 y el área de circunferencia externa ex234 que están dentro y fuera del área de grabación de datos ex233, respectivamente, son para uso específico excepto para la grabación de los datos de usuario. La unidad de reproducción/grabación de información 400 lee y escribe audio codificado, datos de video codificados, o datos multiplexados obtenidos multiplexando los datos codificados de audio y video, de y en el área de grabación de datos ex233 del medio de grabación ex215.

40 Aunque un disco óptico que tiene una capa, tal como un DVD y un BD se describe como un ejemplo en la descripción, el disco óptico no se limita a esto, y puede ser un disco óptico que tiene una estructura multicapa y puede ser grabado en una parte que no sea la superficie. Además, el disco óptico puede tener una estructura para grabación/reproducción multidimensional, tal como la grabación de información usando luz de colores con diferentes longitudes de onda en la misma porción del disco óptico y para la grabación de información que tiene capas diferentes desde varios ángulos.

55 Además, un automóvil ex210 que tiene una antena ex205 puede recibir datos del satélite ex202 y otros, y reproducir video en un dispositivo de visualización tal como un sistema de navegación para automóvil ex211 instalado en el automóvil ex210, en el sistema de difusión digital ex200. En este punto, una configuración del sistema de navegación de automóvil ex211 será una configuración que, por ejemplo, incluye una unidad de recepción de GPS de la configuración ilustrada en la figura 30. Lo mismo ocurrirá con la configuración del ordenador ex111, el teléfono

60

móvil ex114 y otros.

La figura 33A ilustra el teléfono móvil ex114 que usa el procedimiento de codificación de imágenes en movimiento y el procedimiento de decodificación de imágenes en movimiento descrito en las realizaciones. El teléfono móvil ex114 incluye: una antena ex350 para transmitir y recibir ondas de radio a través de la estación base ex110; una unidad de cámara ex365 capaz de capturar imágenes en movimiento e imágenes fijas; y una unidad de visualización ex358 tal como una pantalla de cristal líquido para visualizar los datos tales como de video decodificado capturado por la unidad de cámara ex365 o recibido por la antena ex350. El teléfono móvil ex114 incluye además: una unidad de cuerpo principal que incluye una unidad de teclado de operación ex366; una unidad de salida de audio ex357 como un altavoz para salida de audio; una unidad de entrada de audio ex356 tal como un micrófono para la entrada de audio; una unidad de memoria ex367 para almacenar video o imágenes fijas capturadas, audio grabado, datos codificados o decodificados del video recibido, las imágenes fijas, correos electrónicos u otros; y una unidad de ranura (*slot unit*) ex364 que es una unidad de interfaz para un medio de grabación que almacena datos de la misma manera que la unidad de memoria ex367.

A continuación, se describirá un ejemplo de una configuración del teléfono móvil ex114 con referencia a la figura 33B. En el teléfono móvil ex114, una unidad de control principal ex360 diseñada para controlar en general cada unidad del cuerpo principal que incluye la unidad de visualización ex358 así como la unidad de teclado de operación ex366 se conecta entre sí, a través de un bus síncrono ex370, a una unidad de circuito de suministro de energía ex361, una unidad de control de entrada de operación ex362, una unidad de procesamiento de señales de video ex355, una unidad de interfaz de cámara ex363, una unidad de control de pantalla de cristal líquido (LCD) ex359, una unidad de modulación/demodulación ex352, una unidad de multiplexación/demultiplexación ex353, una unidad de procesamiento de señales de audio ex354, la unidad de ranura ex364 y la unidad de memoria ex367.

Cuando se activa una tecla de finalización de llamada o una tecla de encendido mediante operación de usuario, la unidad de circuito de suministro de energía ex361 suministra energía de un paquete de baterías a las respectivas unidades para activar el teléfono móvil ex114.

En el teléfono móvil ex114, la unidad de procesamiento de señales de audio ex354 convierte las señales de audio recogidas por la unidad de entrada de audio ex356 en modo de conversación de voz en señales de audio digitales bajo el control de la unidad de control principal ex360 que incluye una CPU, ROM y RAM. A continuación, la unidad de modulación/demodulación ex352 realiza el procesamiento de espectro expandido en las señales de audio digital, y la unidad de transmisión y recepción ex351 realiza la conversión de digital a analógico y la conversión de frecuencia en los datos, para transmitir los datos resultantes a través de la antena ex350. Además, en el teléfono móvil ex114, la unidad de transmisión y recepción ex351 amplifica los datos recibidos por la antena ex350 en el modo de conversación de voz y realiza la conversión de frecuencia y la conversión de analógico a digital en los datos. A continuación, la unidad de modulación/demodulación ex352 realiza el procesamiento de espectro expandido inverso en los datos, y la unidad de procesamiento de señales de audio ex354 realiza su conversión en señales de audio analógicas, para emitirlos a través de la unidad de salida de audio ex357.

Además, cuando se transmite un correo electrónico en modo de comunicación de datos, los datos de texto del correo electrónico introducidos operando la unidad de teclado de operación ex366 y otros del cuerpo principal se envían a la unidad de control principal ex360 a través de la unidad de control de entrada de operación ex362. La unidad de control principal ex360 hace que la unidad de modulación/demodulación ex352 realice el procesamiento de espectro expandido en los datos de texto, y la unidad de transmisión y recepción ex351 realiza la conversión digital a analógica y la conversión de frecuencia en los datos resultantes para transmitir los datos a la estación base ex110 a través de la antena ex350. Cuando se recibe un correo electrónico, se realiza en los datos recibidos el procesamiento que es aproximadamente el inverso al procesamiento para transmitir un correo electrónico, y los datos resultantes se proporcionan a la unidad de visualización ex358.

Cuando se transmite video, imágenes fijas o video y audio en el modo de comunicación de datos, la unidad de procesamiento de señales de video ex355 comprime y codifica las señales de video suministradas por la unidad de cámara ex365 usando el procedimiento de codificación de imágenes en movimiento mostrado en cada una de las realizaciones (es decir, funciona como el aparato de codificación de imágenes de acuerdo con el aspecto de la presente invención), y transmite los datos de video codificados a la unidad de multiplexación/demultiplexación ex353. Por el contrario, cuando la unidad de cámara ex365 captura video, imágenes fijas y otras, la unidad de procesamiento de señales de audio ex354 codifica señales de audio recogidas por la unidad de entrada de audio ex356, y transmite los datos de audio codificados a la unidad de multiplexación/demultiplexación ex353.

La unidad de multiplexación/demultiplexación ex353 multiplexa los datos de video codificados suministrados por la unidad de procesamiento de señales de video ex355 y los datos de audio codificados suministrados por la unidad de procesamiento de señales de audio ex354, usando un procedimiento predeterminado. A continuación, la unidad de modulación/demodulación (unidad de circuito de modulación/demodulación) ex352 realiza el procesamiento de espectro expandido en los datos multiplexados, y la unidad de transmisión y recepción ex351 realiza conversión de digital a analógico y la conversión de frecuencia en los datos para transmitir los datos resultantes a través de la antena ex350.

Cuando se reciben datos de un archivo de video que está vinculado a una página web y otros en el modo de comunicación de datos o cuando se recibe un correo electrónico con video y/o audio adjunto, para decodificar los datos multiplexados recibidos a través de la antena ex350, la unidad de multiplexación/demultiplexación ex353 demultiplexa los datos multiplexados en una secuencia de bits (*bitstream*) de datos de video y una secuencia de bits (*bitstream*) de datos de audio, y suministra los datos de video codificados a la unidad de procesamiento de señales de video ex355 y los datos de audio codificados a la unidad de procesamiento de señales de audio ex354, a través del bus síncrono ex370. La unidad de procesamiento de señales de video ex355 decodifica la señal de video usando un procedimiento de decodificación de imágenes en movimiento correspondiente al procedimiento de codificación de imágenes en movimiento mostrado en cada una de las realizaciones (es decir, funciona como el aparato de decodificación de imágenes según el aspecto de la presente invención) y, luego, la unidad de visualización ex358 visualiza, por ejemplo, el video y las imágenes fijas incluidas en el archivo de video vinculado a la página web a través de la unidad de control de LCD ex359. Además, la unidad de procesamiento de señales de audio ex354 decodifica la señal de audio, y la unidad de salida de audio ex357 proporciona el audio.

Además, de forma similar al televisor ex300, un terminal tal como el teléfono móvil ex114 probablemente tiene 3 tipos de configuraciones de implementación que incluyen no solo (i) un terminal de transmisión y recepción que incluye tanto un aparato de codificación como un aparato de decodificación, sino también (ii) un terminal de transmisión que incluye solo un aparato de codificación y (iii) un terminal de recepción que incluye solo un aparato de decodificación. Aunque el sistema de emisión de difusión digital ex200 recibe y transmite los datos multiplexados obtenidos multiplexando datos de audio sobre datos de video en la descripción, los datos multiplexados pueden ser datos obtenidos multiplexando datos que no son de audio sino datos de caracteres relacionados con video sobre datos de video, y pueden no ser datos multiplexados sino datos de video en sí.

Como tal, el procedimiento de codificación de imágenes en movimiento y el procedimiento de decodificación de imágenes en movimiento en cada una de las realizaciones se pueden usar en cualquiera de los dispositivos y sistemas descritos. Por lo tanto, se pueden obtener las ventajas descritas en cada una de las realizaciones.

Además, la presente invención no se limita a las realizaciones, y son posibles diversas modificaciones y revisiones sin apartarse del alcance de la presente invención.

[Realización 6]

Se pueden generar datos de video conmutando, según sea necesario, entre (i) el procedimiento de codificación de imágenes en movimiento o el aparato de codificación de imágenes en movimiento mostrado en cada una de las realizaciones y (ii) un procedimiento de codificación de imágenes en movimiento o un aparato de codificación de imágenes en movimiento conforme a un estándar diferente, como MPEG-2, MPEG-4 AVC y VC-1.

En este punto, cuando se genera una pluralidad de datos de video que se ajustan a los diferentes estándares y luego se decodifican, deben seleccionarse los procedimientos de decodificación para cumplir con los diferentes estándares. Sin embargo, dado que no se puede detectar a qué estándar corresponde cada uno de la pluralidad de los datos de video a decodificar, existe el problema de que no se puede seleccionar un procedimiento de decodificación apropiado.

Para resolver el problema, los datos multiplexados obtenidos multiplexando datos de audio y otros sobre datos de video tienen una estructura que incluye información de identificación que indica a qué estándar corresponden los datos de video. A continuación se describirá la estructura específica de los datos multiplexados que incluye los datos de video generados en el procedimiento de codificación de imágenes en movimiento y por el aparato de codificación de imágenes en movimiento mostrado en cada una de las realizaciones. Los datos multiplexados son una secuencia digital en formato de secuencia de transporte MPEG-2.

La figura 34 ilustra una estructura de los datos multiplexados. Como se ilustra en la figura 34, los datos multiplexados pueden obtenerse multiplexando al menos uno de entre un flujo de video (*video stream*), un flujo de audio (*audio stream*), un flujo de gráficos de presentación (PG) y un flujo de gráficos interactivos. El flujo de video representa el video principal y video secundario de una película, el flujo de audio (IG) representa una parte de audio principal y una parte de audio secundaria para mezclarla con la parte de audio principal, y el flujo de gráficos de presentación representa los subtítulos de la película. En este punto, el video principal es el video normal que se mostrará en una pantalla, y el video secundario es el video que se mostrará en una ventana más pequeña en el video principal. Además, el flujo de gráficos interactivos representa una pantalla interactiva a generar organizando componentes GUI en una pantalla. El flujo de video se codifica en el procedimiento de codificación de imágenes en movimiento o en el aparato de codificación de imágenes en movimiento mostrado en cada una de las realizaciones, o en un procedimiento de codificación de imágenes en movimiento o por un aparato de codificación de imágenes en movimiento conforme a un estándar convencional, como MPEG-2, MPEG-4 AVC y VC-1. El flujo de audio se codifica de acuerdo con un estándar, como Dolby-AC-3, Dolby Digital Plus, MLP, DTS, DTS-HD y PCM lineal.

Cada flujo incluido en los datos multiplexados es identificado por un PID. Por ejemplo, se asigna 0x1011 al flujo de video a usar para el video de una película, se asignan del 0x1100 al 0x111F a los flujos de audio, se asignan del 0x1200 al 0x121F a los flujos de gráficos de presentación, se asignan del 0x1400 al 0x141F a los flujos de gráficos

interactivos, se asignan del 0x1B00 al 0x1B1F a los flujos de video a utilizar para el video secundario de la película, y se asignan del 0x1A00 al 0x1A1F a los flujos de audio a utilizar para el audio secundario a mezclar con el audio principal.

La figura 35 ilustra esquemáticamente cómo se multiplexan los datos. Primero, un flujo de video ex235 compuesto por fotogramas de video y un flujo de audio ex238 compuesto por tramas de audio se transforman en un flujo de paquetes PES ex236 y un flujo de paquetes PES ex239, y además en paquetes TS ex237 y paquetes TS ex240, respectivamente. De manera similar, los datos de un flujo de gráficos de presentación ex241 y datos de un flujo de gráficos interactivos ex244 se transforman en un flujo de paquetes PES ex242 y un flujo de paquetes PES ex245, y además en paquetes TS ex243 y paquetes TS ex246, respectivamente. Estos paquetes TS se multiplexan en un flujo para obtener datos multiplexados ex247.

La figura 36 ilustra cómo se almacena un flujo de video en un flujo de paquetes PES con más detalle. La primera barra de la figura 36 muestra un flujo de fotogramas de video en un flujo de video. La segunda barra muestra el flujo de paquetes PES. Como denotan las flechas indicadas como yy1, yy2, yy3 e yy4 en la figura 36, el flujo de video se divide en imágenes como imágenes I, imágenes B e imágenes P siendo cada una de ellas una unidad de presentación de video, y las imágenes se almacenan en una carga útil de cada uno de los paquetes PES. Cada uno de los paquetes PES tiene una cabecera PES, y la cabecera PES almacena una marca de tiempo de presentación (PTS) que indica un momento/hora de visualización de la imagen, y una marca de tiempo de decodificación (DTS) que indica un momento/hora de decodificación de la imagen.

La figura 37 ilustra un formato de paquetes TS para ser finalmente escritos en los datos multiplexados. Cada uno de los paquetes TS es un paquete de longitud fija de 188 bytes que incluye una cabecera TS de 4 bytes que tiene información, tal como un PID para identificar un flujo y una carga útil TS de 184 bytes para almacenar datos. Los paquetes PES se dividen y almacenan en las cargas útiles TS, respectivamente. Cuando se utiliza un BD-ROM, se le atribuye a cada uno de los paquetes TS una cabecera TP_Extra_Header de 4 bytes, dando como resultado unos paquetes fuente de 192 bytes. Los paquetes fuente son escritos en los datos multiplexados. La TP_Extra_Header almacena información tal como una marca de tiempo de llegada Arrival_Time_Stamp (ATS). La ATS muestra un momento/hora de inicio de transferencia en el que cada uno de los paquetes TS debe transferirse a un filtro PID. Los paquetes fuente se disponen en los datos multiplexados como se muestra en la parte inferior de la figura 37. Los números que se incrementan a partir de la cabecera de los datos multiplexados se denominan números de paquete fuente (SPN).

Cada uno de los paquetes TS incluidos en los datos multiplexados incluye no solo flujos de audio, video, subtítulos y otros, sino también una tabla de asociación de programas (PAT), una tabla de mapas de programas (PMT) y una referencia de reloj de programa (PCR). La PAT muestra lo que indica un PID en una PMT que se usa en los datos multiplexados, y se registra igual a cero un PID de la propia PAT. La PMT almacena los PID de los flujos de video, audio, subtítulos y otros incluidos en los datos multiplexados, y la información de atributos de los flujos correspondientes a los PID. La PMT también tiene varios descriptores relacionados con los datos multiplexados. Los descriptores tienen información tal como información de control de copia que muestra si se permite o no la copia de los datos multiplexados. La PCR almacena información de tiempo STC correspondiente a una ATS que muestra cuando se transfiere el paquete PCR a un decodificador, para conseguir la sincronización entre un reloj de tiempo/momento de llegada (ATC) que es un eje de tiempo/momento de ATS y un reloj de tiempo/momento del sistema (STC) que es un eje de tiempo de PTS y DTS.

La figura 38 ilustra la estructura de datos de la PMT en detalle. En la parte superior de la PMT se dispone una cabecera PMT. La cabecera PMT describe la longitud de los datos incluidos en la PMT y otros. Después de la cabecera PMT se dispone una pluralidad de descriptores relacionados con los datos multiplexados. Información tal como la información de control de copia se describe en los descriptores. Después de los descriptores, se dispone una pluralidad de fragmentos de información de flujo relacionada con los flujos incluidos en los datos multiplexados. Cada fragmento de información de flujo incluye descriptores que describen cada uno información, tal como un tipo de flujo para identificar un códec de compresión de un flujo, un PID de flujo e información de atributos de flujo (tal como una velocidad de fotogramas o una relación de aspecto). Los descriptores de flujo son iguales en número a la cantidad de flujos en los datos multiplexados.

Cuando los datos multiplexados son grabados en un medio de grabación y otros, se graban junto con archivos de información de los datos multiplexados.

Cada uno de los archivos de información de datos multiplexados es información de gestión de los datos multiplexados como se muestra en la figura 39. Los archivos de información de datos multiplexados están en correspondencia uno a uno con los datos multiplexados, y cada uno de los archivos incluye información de datos multiplexados, información de atributos de flujo y un mapa de entrada.

Como se ilustra en la figura 39, la información de datos multiplexados incluye una velocidad del sistema, un momento de inicio de la reproducción y un momento de finalización de la reproducción. La velocidad del sistema indica la velocidad de transferencia máxima a la que se transfieren los datos multiplexados a un filtro PID por parte de un decodificador destino del sistema que se describirá más adelante. Los intervalos de las ATS incluidos en los

datos multiplexados se establecen en una velocidad no superior a la del sistema. El momento de inicio de la reproducción indica una PTS en un fotograma de video en la cabecera de los datos multiplexados. Se agrega un intervalo de un fotograma a una PTS en un fotograma de video al final de los datos multiplexados, y la PTS se establece en el momento/hora de finalización de la reproducción.

5 Como se muestra en la figura 40, se registra un fragmento de la información de atributos en la información de atributos del flujo, para cada PID de cada flujo incluido en los datos multiplexados. Cada fragmento de la información de atributos tiene información diferente dependiendo de si el flujo correspondiente es un flujo de video, un flujo de audio, un flujo de gráficos de presentación o un flujo de gráficos interactivos. Cada fragmento de información de atributos de flujo de video transporta información que incluye qué tipo de códec de compresión se usa para comprimir el flujo de video, y la resolución, la relación de aspecto y la velocidad de fotogramas de los fragmentos de datos de imagen que se incluyen en el flujo de video. Cada fragmento de la información de atributos de flujo de audio transporta información que incluye qué tipo de códec de compresión se utiliza para comprimir el flujo de audio, cuántos canales se incluyen en el flujo de audio, qué idioma admite el flujo de audio y cuán alta es la frecuencia de muestreo. La información de atributos de flujo de video y la información de atributos de flujo de audio se usan para la inicialización de un decodificador antes de que el reproductor reproduzca la información.

En la presente realización, los datos multiplexados a usar son de un tipo de flujo incluido en la PMT. Además, cuando los datos multiplexados se graban en un medio de grabación, se usa la información de atributos de flujo de video incluida en la información de los datos multiplexados. Más específicamente, el procedimiento de codificación de imágenes en movimiento o el aparato de codificación de imágenes en movimiento descritos en cada una de las realizaciones incluye una etapa o una unidad para asignar información única que indica los datos de video generados por el procedimiento de codificación de imágenes en movimiento o el aparato de codificación de imágenes en movimiento, para el tipo de flujo incluido en la PMT o la información de atributos del flujo de video. Con la configuración, los datos de video generados por el procedimiento de codificación de imágenes en movimiento o el aparato de codificación de imágenes en movimiento descritos en cada una de las realizaciones se pueden distinguir de los datos de video correspondientes a otro estándar.

Además, la figura 41 ilustra las etapas del procedimiento de decodificación de imágenes en movimiento de acuerdo con la presente realización. En la etapa exS100, el tipo de flujo incluido en la PMT o la información de atributos de flujo de video incluida en la información de datos multiplexados se obtiene de los datos multiplexados. A continuación, en la etapa exS101, se determina si el tipo de flujo o la información de atributos del flujo de video indican que los datos multiplexados están generados por el procedimiento de codificación de imágenes en movimiento o el aparato de codificación de imágenes en movimiento en cada una de las realizaciones. Cuando se determina que el tipo de flujo o la información de atributos del flujo de video indica que los datos multiplexados están generados por el procedimiento de codificación de imágenes en movimiento o el aparato de codificación de imágenes en movimiento en cada una de las realizaciones, en la etapa exS102, la decodificación se realiza mediante el procedimiento de decodificación de imágenes en movimiento en cada una de las realizaciones. Además, cuando el tipo de flujo o la información de atributos del flujo de video indican correspondencia con los estándares convencionales, tales como MPEG-2, MPEG-4 AVC y VC-1, en la etapa exS103, la decodificación se realiza mediante un procedimiento de decodificación de imágenes en movimiento de acuerdo con los estándares convencionales.

Como tal, asignar un nuevo valor único al tipo de flujo o a la información de atributos del flujo de video permite determinar si el procedimiento de decodificación de imágenes en movimiento o el aparato de decodificación de imágenes en movimiento que se describe en cada una de las realizaciones puede o no puede realizar la decodificación. Incluso en el caso de datos multiplexados correspondientes a un estándar diferente, se puede seleccionar un procedimiento o aparato de decodificación apropiado. Por lo tanto, es posible decodificar información sin errores. Además, el procedimiento o aparato de codificación de imágenes en movimiento, o el procedimiento o aparato de decodificación de imágenes en movimiento de la presente realización se puede usar en los dispositivos y sistemas descritos anteriormente.

[Realización 7]

Cada uno de entre el procedimiento de codificación de imágenes en movimiento, el aparato de codificación de imágenes en movimiento, el procedimiento de decodificación de imágenes en movimiento y el aparato de decodificación de imágenes en movimiento de cada una de las realizaciones normalmente se consigue en forma de un circuito integrado o de un circuito integrado a gran escala (LSI). Como un ejemplo de LSI, la figura 42 ilustra una configuración de LSI ex500 que está fabricado en un chip. El LSI ex500 incluye los elementos ex501, ex502, ex503, ex504, ex505, ex506, ex507, ex508 y ex509 que se describirán a continuación, y los elementos están conectados entre sí a través de un bus ex510. La unidad de circuito de suministro de energía ex505 se activa suministrando energía a cada uno de los elementos cuando se enciende la unidad de circuito de suministro de energía ex505.

Por ejemplo, cuando se realiza la codificación, el LSI ex500 recibe una señal AV de un micrófono ex117, una cámara ex113, y otros a través de una E/S de AV ex509 bajo el control de una unidad de control ex501 que incluye una CPU ex502, un controlador de memoria ex503, un controlador de flujos ex504 y una unidad de control de la frecuencia de funcionamiento ex512. La señal AV recibida se almacena temporalmente en una memoria externa ex511, tal como

una SDRAM. Bajo el control de la unidad de control ex501, los datos almacenados se segmentan en porciones de datos de acuerdo con la cantidad de procesamiento y la velocidad a transmitir a una unidad de procesamiento de señales ex507. Entonces, la unidad de procesamiento de señales ex507 codifica una señal de audio y/o una señal de video. En este punto, la codificación de la señal de video es la codificación descrita en cada una de las realizaciones. Además, la unidad de procesamiento de señales ex507 a veces multiplexa los datos de audio codificados y los datos de video codificados, y una E/S de flujo ex506 proporciona los datos multiplexados al exterior. Los datos multiplexados proporcionados se transmiten a la estación base ex107, o se escriben en el medio de grabación ex215. Cuando los conjuntos de datos se multiplexan, los datos deben almacenarse temporalmente en el búfer ex508 para que los conjuntos de datos se sincronicen entre sí.

Aunque la memoria ex511 es un elemento externo al LSI ex500, puede estar incluida en el LSI ex500. El búfer ex508 no se limita a un búfer, sino que puede estar compuesto de búferes. Además, el LSI ex500 puede estar fabricado en un chip o una pluralidad de chips.

Además, aunque la unidad de control ex501 incluye la CPU ex502, el controlador de memoria ex503, el controlador de flujos ex504, la unidad de control de la frecuencia de funcionamiento ex512, la configuración de la unidad de control ex501 no se limita a esto. Por ejemplo, la unidad de procesamiento de señales ex507 puede incluir además una CPU. La inclusión de otra CPU en la unidad de procesamiento de señales ex507 puede mejorar la velocidad de procesamiento. Además, como otro ejemplo, la CPU ex502 puede servir o ser parte de la unidad de procesamiento de señales ex507, y, por ejemplo, puede incluir una unidad de procesamiento de señales de audio. En tal caso, la unidad de control ex501 incluye la unidad de procesamiento de señales ex507 o la CPU ex502 que incluye una parte de la unidad de procesamiento de señales ex507.

El nombre utilizado aquí es LSI, pero también se puede llamar IC, sistema LSI, súper LSI o ultra LSI según el grado de integración.

Además, las formas de conseguir la integración no se limitan al LSI, y un circuito especial o un procesador de propósito general y demás también pueden conseguir la integración. Para el mismo propósito, puede usarse una matriz de puerta programable de campo (FPGA) que se puede programar después de la fabricación de los LSI o un procesador reconfigurable que permite la reconfiguración de la conexión o la configuración de un LSI.

En el futuro, con el avance en la tecnología de semiconductores, una nueva tecnología puede reemplazar al LSI. Los bloques funcionales se pueden integrar usando dicha tecnología. La posibilidad es que la presente invención se aplique a la biotecnología.

[Realización 8]

Cuando se decodifican datos de video generados en el procedimiento de codificación de imágenes en movimiento o por el aparato de codificación de imágenes en movimiento descritos en cada una de las realizaciones, en comparación con la decodificación de datos de video que satisfacen un estándar convencional, como MPEG-2, MPEG-4 AVC, y VC-1, la cantidad de procesamiento probablemente aumente. Por lo tanto, se requiere que el LSI ex500 tenga una frecuencia de funcionamiento mayor que la de la CPU ex502 a utilizar para decodificar los datos de video que satisfacen el estándar convencional. Sin embargo, cuando se incrementa la frecuencia de funcionamiento, existe el problema de que aumente el consumo de energía.

Para resolver el problema, el aparato de decodificación de imágenes en movimiento, tal como el televisor ex300 y el LSI ex500 está configurado para determinar a qué estándar corresponden los datos de video, y conmutar entre las frecuencias de funcionamiento según el estándar determinado. La figura 43 ilustra una configuración ex800 en la presente realización. Una unidad de conmutación de la frecuencia de funcionamiento ex803 establece una frecuencia de funcionamiento a una frecuencia de funcionamiento mayor cuando se generan los datos de video por parte del procedimiento de codificación de imágenes en movimiento o del aparato de codificación de imágenes en movimiento descrito en cada una de las realizaciones. Entonces, la unidad de conmutación de la frecuencia de funcionamiento ex803 da instrucciones a una unidad de procesamiento de decodificación ex801 que ejecuta el procedimiento de decodificación de imágenes en movimiento descrito en cada una de las realizaciones para decodificar los datos de video. Cuando la información de video satisface el estándar convencional, la unidad de conmutación de la frecuencia de funcionamiento ex803 establece una frecuencia de funcionamiento a una frecuencia de funcionamiento inferior a la de los datos de video generados por el procedimiento de codificación de imágenes en movimiento o el aparato de codificación de imágenes en movimiento descrito en cada una de las realizaciones. Entonces, la unidad de conmutación de la frecuencia de funcionamiento ex803 da instrucciones a la unidad de procesamiento de decodificación ex802 que se ajuste al estándar convencional para decodificar los datos de video.

Más específicamente, la unidad de conmutación de la frecuencia de funcionamiento ex803 incluye la CPU ex502 y la unidad de control de la frecuencia de funcionamiento ex512 de la figura 42. En este punto, cada una de entre la unidad de procesamiento de decodificación ex801 que ejecuta el procedimiento de decodificación de imágenes en movimiento descrito en cada una de las realizaciones y la unidad de procesamiento de decodificación ex802 que se ajusta al estándar convencional corresponde a la unidad de procesamiento de señales ex507 de la figura 42. La

CPU ex502 determina a qué estándar se ajustan los datos de video. Entonces, la unidad de control de la frecuencia de funcionamiento ex512 determina una frecuencia de funcionamiento en base a una señal de la CPU ex502. Además, la unidad de procesamiento de señales ex507 decodifica los datos de video en base a la señal de la CPU ex502. Por ejemplo, la información de identificación descrita en la Realización 6 probablemente se usa para identificar los datos de video. La información de identificación no se limita a la descrita en la Realización 6, sino que puede ser cualquier información siempre que la información indique a qué estándar se ajustan los datos de video. Por ejemplo, cuando se puede determinar a qué estándar se ajustan los datos de video en base a una señal externa para determinar que los datos de video se usan para un televisor o un disco, etc., la determinación se puede hacer en base a dicha señal externa. Además, la CPU ex502 selecciona una frecuencia de funcionamiento en base a, por ejemplo, una tabla de búsqueda en la que los estándares de datos de video están asociados con las frecuencias de funcionamiento como se muestra en la figura 45. La frecuencia de funcionamiento se puede seleccionar almacenando la tabla de búsqueda en el búfer ex508 y en una memoria interna de un LSI, y con referencia a la tabla de búsqueda por parte de la CPU ex502.

La figura 44 ilustra unas etapas para ejecutar un procedimiento en la presente realización. Primero, en la etapa exS200, la unidad de procesamiento de señales ex507 obtiene información de identificación a partir de los datos multiplexados. A continuación, en la etapa exS201, la CPU ex502 determina si los datos de video están generados o no por el procedimiento de codificación y el aparato de codificación descritos en cada una de las realizaciones, en base a la información de identificación. Cuando los datos de video están generados por el procedimiento de codificación de imágenes en movimiento y el aparato de codificación de imágenes en movimiento descrito en cada una de las realizaciones, en la etapa exS202, la CPU ex502 transmite una señal para ajustar la frecuencia de funcionamiento a una frecuencia de funcionamiento mayor a la unidad de control de la frecuencia de funcionamiento ex512. Entonces, la unidad de control de la frecuencia de funcionamiento ex512 establece la frecuencia de funcionamiento a la frecuencia de funcionamiento mayor. Por otro lado, cuando la información de identificación indica que los datos de video cumplen con el estándar convencional, tal como MPEG-2, MPEG-4 AVC y VC-1, en la etapa exS203, la CPU ex502 transmite una señal para ajustar la frecuencia de funcionamiento a una frecuencia de funcionamiento inferior a la unidad de control de la frecuencia de funcionamiento ex512. Entonces, la unidad de control de la frecuencia de funcionamiento ex512 ajusta la frecuencia de funcionamiento a la frecuencia de funcionamiento menor que la del caso en que se han generado los datos de video por parte del procedimiento de codificación de imágenes en movimiento y del aparato de codificación de imágenes en movimiento descrito en cada una de las realizaciones.

Además, junto con la conmutación de las frecuencias de funcionamiento, se puede mejorar el efecto de conservación de energía cambiando el voltaje a aplicar al LSI ex500 o a un aparato que incluye el LSI ex500. Por ejemplo, cuando se ajusta la frecuencia de funcionamiento a un valor menor, el voltaje a aplicar al LSI ex500 o al aparato que incluye el LSI ex500 probablemente se ajuste a un voltaje menor que en el caso en que la frecuencia de funcionamiento se ajuste a un valor mayor.

Además, cuando la cantidad de procesamiento para la decodificación es mayor, la frecuencia de funcionamiento se puede ajustar a una frecuencia mayor, y cuando la cantidad de procesamiento para la decodificación es menor, la frecuencia de funcionamiento se puede ajustar a una frecuencia menor, como el procedimiento para ajustar la frecuencia de funcionamiento. Por lo tanto, el procedimiento de ajuste no se limita a los descritos anteriormente. Por ejemplo, cuando la cantidad de procesamiento para decodificar datos de video en conformidad con el MPEG-4 AVC es mayor que la cantidad de procesamiento para decodificar los datos de video generados por el procedimiento de codificación de imágenes en movimiento y el aparato de codificación de imágenes en movimiento descrito en cada una de las realizaciones, es probable que la frecuencia de funcionamiento se ajuste en orden inverso al ajuste descrito anteriormente.

Además, el procedimiento para ajustar la frecuencia de funcionamiento no se limita al procedimiento para ajustar la frecuencia de funcionamiento a un valor menor. Por ejemplo, cuando la información de identificación indica que los datos de video están generados por el procedimiento de codificación de imágenes en movimiento y el aparato de codificación de imágenes en movimiento descrito en cada una de las realizaciones, el voltaje a aplicar al LSI ex500 o al aparato que incluye el LSI ex500 probablemente se ajuste a un valor mayor. Cuando la información de identificación indica que los datos de video cumplen con el estándar convencional, tal como MPEG-2, MPEG-4 AVC y VC-1, el voltaje a aplicar al LSI ex500 o al aparato que incluye el LSI ex500 probablemente se ajuste a un valor menor. Como otro ejemplo, cuando la información de identificación indica que los datos de video están generados por el procedimiento de codificación de imágenes en movimiento y el aparato de codificación de imágenes en movimiento descrito en cada una de las realizaciones, probablemente no tiene que suspenderse el funcionamiento de la CPU ex502. Cuando la información de identificación indica que los datos de video se ajustan al estándar convencional, tal como MPEG-2, MPEG-4 AVC y VC-1, probablemente se suspenda el funcionamiento de la CPU ex502 en un momento dado porque la CPU ex502 tiene una capacidad de procesamiento adicional. Incluso cuando la información de identificación indica que los datos de video están generados por el procedimiento de codificación de imágenes en movimiento y el aparato de codificación de imágenes en movimiento descrito en cada una de las realizaciones, en el caso en que la CPU ex502 tiene una capacidad de procesamiento adicional, probablemente se suspenda el funcionamiento de la CPU ex502 en un momento dado. En tal caso, el tiempo de suspensión probablemente sea más corto que en el caso en que la información de identificación indique que los datos de video se ajustan al estándar convencional, tal como MPEG-2, MPEG-4 AVC y VC-1.

En consecuencia, se puede mejorar el efecto de conservación de energía conmutando entre las frecuencias de funcionamiento de acuerdo con el estándar al que se ajustan los datos de video. Además, cuando el LSI ex500 o el aparato que incluye el LSI ex500 funciona con una batería, se puede alargar la duración de la batería con el efecto de conservación de energía.

5 [Realización 9]

Hay casos en los que se proporciona una pluralidad de datos de video que se ajustan a estándares diferentes a los dispositivos y sistemas, tales como un televisor y un teléfono móvil. Para permitir la decodificación de la pluralidad de datos de video que se ajustan a los diferentes estándares, la unidad de procesamiento de señales ex507 del LSI ex500 necesita ajustarse a los diferentes estándares. Sin embargo, con el uso individual de las unidades de procesamiento de señales ex507 que se ajustan a los respectivos estándares, aparecen problemas de aumento en la escala del circuito del LSI ex500 y de aumento en el coste.

Para resolver el problema, se concibe una configuración en la que se comparten parcialmente la unidad de procesamiento de decodificación para implementar el procedimiento de decodificación de imágenes en movimiento descrito en cada una de las realizaciones y la unidad de procesamiento de decodificación que se ajusta al estándar convencional, tal como MPEG-2, MPEG-4 AVC y VC-1. En la figura 46A, ex900 muestra un ejemplo de la configuración. Por ejemplo, el procedimiento de decodificación de imágenes en movimiento descrito en cada una de las realizaciones y el procedimiento de decodificación de imágenes en movimiento que se ajusta al MPEG-4 AVC tienen, parcialmente en común, los detalles del procesamiento, tales como la codificación de entropía, la cuantificación inversa, el filtrado de desbloqueo y la predicción compensada de movimiento. Los detalles del procesamiento a compartir probablemente incluyan el uso de una unidad de procesamiento de decodificación ex902 que se ajusta al MPEG-4 AVC. Por el contrario, es probable que se use una unidad de procesamiento de decodificación dedicada ex901 para otro procesamiento exclusivo de un aspecto de la presente invención. Como el aspecto de la presente invención se caracteriza por una cuantificación inversa particular, por ejemplo, la unidad de procesamiento de decodificación dedicada ex901 se usa para la cuantificación inversa. De lo contrario, probablemente se comparta la unidad de procesamiento de decodificación para una de entre la decodificación de entropía, el filtrado de desbloqueo, y la compensación de movimiento, o todo el procesamiento. Se puede compartir la unidad de procesamiento de decodificación para implementar el procedimiento de decodificación de imágenes en movimiento descrito en cada una de las realizaciones para que el procesamiento sea compartido, y se puede usar una unidad de procesamiento de decodificación dedicada para el procesamiento exclusivo del MPEG-4 AVC.

Además, en la figura 46B, ex1000 muestra otro ejemplo en el que el procesamiento es parcialmente compartido. Este ejemplo usa una configuración que incluye una unidad de procesamiento de decodificación dedicada ex1001 que admite el procesamiento exclusivo de un aspecto de la presente invención, una unidad de procesamiento de decodificación dedicada ex1002 que admite el procesamiento exclusivo de otro estándar convencional, y una unidad de procesamiento de decodificación ex1003 que admite un procesamiento a compartir entre el procedimiento de decodificación de imágenes en movimiento de acuerdo con el aspecto de la presente invención y el procedimiento de decodificación de imágenes en movimiento convencional. En este punto, las unidades de procesamiento de decodificación dedicadas ex1001 y ex1002 no están necesariamente especializadas en el procesamiento de acuerdo con el aspecto de la presente invención y el procesamiento del estándar convencional, respectivamente, y pueden ser capaces de implementar un procesamiento general. Además, la configuración de la presente realización puede implementarse mediante el LSI ex500.

Como tal, es posible reducir la escala del circuito de un LSI y reducir el coste compartiendo la unidad de procesamiento de decodificación para que el procesamiento sea compartido entre el procedimiento de decodificación de imágenes en movimiento de acuerdo con el aspecto de la presente invención y el procedimiento de decodificación de imágenes en movimiento conforme al estándar convencional.

45 [Lista de signos de referencia]

100, 200: aparato de codificación de imágenes

101: sustractor

102: unidad de transformación ortogonal

103: unidad de cuantificación

104, 302: unidad de cuantificación inversa

105, 303: unidad de transformación ortogonal inversa

106, 304: agregador

107, 305: memoria de bloques

108, 306: memoria de fotogramas (*frames*)

- 109, 307: unidad de intra predicción
- 110, 308: unidad de inter predicción
- 111, 309: unidad de control de inter predicción
- 112: unidad de determinación del tipo de imagen
- 5 113, 310: conmutador
- 114, 311: unidad de cálculo de candidatos a bloque de fusión
- 115, 312: memoria *colPic*
- 116: unidad de codificación de longitud variable
- 210, 410: unidad de derivación de candidatos a fusión
- 10 211, 411: unidad de determinación
- 212, 412: primera unidad de derivación
- 213, 413: unidad de especificación
- 214, 414: unidad de determinación
- 215, 415: segunda unidad de derivación
- 15 220, 430: unidad de control de predicción
- 230: unidad de codificación
- 300, 400: aparato de decodificación de imágenes
- 301: unidad de decodificación de longitud variable
- 420: unidad de decodificación
- 20

REIVINDICACIONES

1. Un procedimiento de decodificación de imágenes, que comprende:

decodificar información indicativa de un número fijo de candidatos a utilizar para decodificar una pluralidad de bloques codificados utilizando el modo de fusión como modo de inter predicción, siendo el número fijo de candidatos un número máximo fijo de candidatos utilizables para la fusión; y

para un bloque actual de entre la pluralidad de bloques:

derivar una pluralidad de primeros candidatos a fusión, en el que cada candidato a fusión se deriva en base a una combinación de una dirección de predicción, un vector de movimiento y un índice de imagen de referencia que se ha utilizado para decodificar un primer bloque, y

en el que la derivación de la pluralidad de primeros candidatos a fusión excluye, de la pluralidad de primeros candidatos a fusión, cualquier candidato a fusión que es un candidato idéntico, que ha sido codificado por intra predicción, que está fuera de un límite de una porción o una imagen que incluye el bloque actual, y que aún está por codificarse;

determinar si el número total de primeros candidatos a fusión es o no es menor que el número fijo de candidatos;

cuando el número total de primeros candidatos a fusión es menor que el número fijo de candidatos, derivar al menos un segundo candidato a fusión que sea diferente de cada uno de los primeros candidatos a fusión hasta que el número total de primeros candidatos a fusión y de segundos candidatos a fusión sea igual al número fijo de candidatos,

en el que cada uno de los segundos candidatos a fusión tiene una combinación diferente de una dirección de predicción, un vector de movimiento y un índice de imagen de referencia de cada uno de los primeros candidatos a fusión;

decodificar un índice codificado utilizando el número fijo de candidatos independientemente del número total de primeros candidatos a fusión,

en el que el índice codificado corresponde a uno de la pluralidad de primeros candidatos a fusión o a uno del al menos un segundo candidato a fusión; y

decodificar el bloque actual utilizando el correspondiente de la pluralidad de primeros candidatos a fusión o uno del al menos un segundo candidato a fusión.

2. El método de decodificación de imágenes según la reivindicación 1, en el que la decodificación del índice codificado utilizando el número fijo de candidatos comprende la decodificación del índice codificado utilizando una codificación de longitud variable en base al número fijo de candidatos.

3. Un aparato de decodificación de imágenes que comprende un procesador y una memoria, estando configurado el procesador, trabajando conjuntamente con la memoria, para:

decodificar información indicativa de un número fijo de candidatos a utilizar para decodificar una pluralidad de bloques codificados utilizando el modo de fusión como modo de inter predicción, siendo el número fijo de candidatos un número máximo fijo de candidatos utilizables para la fusión; y

para un bloque actual de entre la pluralidad de bloques:

derivar una pluralidad de primeros candidatos a fusión, en el que cada candidato a fusión se deriva en base a una combinación de una dirección de predicción, un vector de movimiento y un índice de imagen de referencia que se ha utilizado para decodificar un primer bloque, y

en el que la derivación de la pluralidad de primeros candidatos a fusión excluye, de la pluralidad de primeros candidatos a fusión, cualquier candidato a fusión que es un candidato idéntico, que ha sido codificado por intra predicción, que está fuera de un límite de una porción o una imagen que incluye el bloque actual, y que aún está por codificarse;

determinar si el número total de primeros candidatos a fusión es o no es menor que el número fijo de candidatos;

cuando el número total de primeros candidatos a fusión es menor que el número fijo de candidatos, derivar al menos un segundo candidato a fusión que sea diferente de cada uno de los primeros candidatos a fusión hasta que el número total de primeros candidatos a fusión y de segundos candidatos a fusión sea igual al número fijo de candidatos,

- en el que cada uno del uno o más segundos candidatos a fusión tiene una combinación diferente de una dirección de predicción, un vector de movimiento y un índice de imagen de referencia de cada uno de los primeros candidatos a fusión;
- 5 decodificar un índice codificado utilizando el número fijo de candidatos independientemente del número total de primeros candidatos a fusión,
- en el que el índice codificado corresponde a uno de la pluralidad de primeros candidatos a fusión o a uno del al menos un segundo candidato a fusión; y
- decodificar el bloque actual utilizando el correspondiente de la pluralidad de primeros candidatos a fusión o uno del al menos un segundo candidato a fusión.
- 10 4. El aparato de decodificación de imágenes según la reivindicación 3, en el que la decodificación del índice codificado utilizando el número fijo de candidatos comprende la decodificación del índice codificado utilizando una codificación de longitud variable en base al número fijo de candidatos.
5. Un procedimiento de codificación de imágenes, que comprende:
- 15 codificar información indicativa de un número fijo de candidatos a utilizar para codificar una pluralidad de bloques a codificar utilizando el modo de fusión como modo de inter predicción, siendo el número fijo de candidatos un número máximo fijo de candidatos utilizables para la fusión; y
- para un bloque actual de entre la pluralidad de bloques:
- 20 derivar una pluralidad de primeros candidatos a fusión, en el que cada candidato a fusión se deriva en base a una combinación de una dirección de predicción, un vector de movimiento y un índice de imagen de referencia que se ha utilizado para codificar un primer bloque, y
- en el que la derivación de la pluralidad de primeros candidatos a fusión excluye, de la pluralidad de primeros candidatos a fusión, cualquier candidato a fusión que es un candidato idéntico, que ha sido codificado por intra predicción, que está fuera de un límite de una porción o una imagen que incluye el bloque actual, y que aún está por codificarse;
- 25 determinar si el número total de primeros candidatos a fusión es o no es menor que el número fijo de candidatos;
- cuando el número total de primeros candidatos a fusión es menor que el número fijo de candidatos, derivar al menos un segundo candidato a fusión que sea diferente de cada uno de los primeros candidatos a fusión hasta que el número total de primeros candidatos a fusión y de segundos candidatos a fusión sea igual al
- 30 número fijo de candidatos,
- en el que cada uno del uno o más segundos candidatos a fusión tiene una combinación diferente de una dirección de predicción, un vector de movimiento y un índice de imagen de referencia de cada uno de los primeros candidatos a fusión;
- 35 codificar un índice codificado utilizando el número fijo de candidatos independientemente del número total de primeros candidatos a fusión,
- en el que el índice codificado corresponde a uno de la pluralidad de primeros candidatos a fusión o a uno del al menos un segundo candidato a fusión; y
- codificar el bloque actual utilizando el correspondiente de la pluralidad de primeros candidatos a fusión o uno del al menos un segundo candidato a fusión.
- 40 6. El procedimiento de codificación de imágenes según la reivindicación 5, en el que la codificación del índice codificado utilizando el número fijo de candidatos comprende la codificación del índice codificado utilizando una codificación de longitud variable en base al número fijo de candidatos.
7. Un aparato de codificación de imágenes que comprende un procesador y una memoria, estando configurado el procesador, trabajando conjuntamente con la memoria, para:
- 45 codificar información indicativa de un número fijo de candidatos a utilizar para codificar una pluralidad de bloques a codificar utilizando el modo de fusión como modo de inter predicción, siendo el número fijo de candidatos un número máximo fijo de candidatos utilizables para la fusión; y
- para un bloque actual de entre la pluralidad de bloques:
- 50 derivar una pluralidad de primeros candidatos a fusión, en el que cada candidato a fusión se deriva en base a una combinación de una dirección de predicción, un vector de movimiento y un índice de imagen de

- referencia que se ha utilizado para codificar un primer bloque, y
- 5 en el que la derivación de la pluralidad de primeros candidatos a fusión excluye, de la pluralidad de primeros candidatos a fusión, cualquier candidato a fusión que es un candidato idéntico, que ha sido codificado por intra predicción, que está fuera de un límite de una porción o una imagen que incluye el bloque actual, y que aún está por codificarse;
- determinar si el número total de primeros candidatos a fusión es o no es menor que el número fijo de candidatos;
- 10 cuando el número total de primeros candidatos a fusión es menor que el número fijo de candidatos, derivar al menos un segundo candidato a fusión que sea diferente de cada uno de los primeros candidatos a fusión hasta que el número total de primeros candidatos a fusión y de segundos candidatos a fusión sea igual al número fijo de candidatos,
- en el que cada uno del uno o más segundos candidatos a fusión tiene una combinación diferente de una dirección de predicción, un vector de movimiento y un índice de imagen de referencia de cada uno de los primeros candidatos a fusión;
- 15 codificar un índice codificado utilizando el número fijo de candidatos independientemente del número total de primeros candidatos a fusión,
- en el que el índice codificado corresponde a uno de la pluralidad de primeros candidatos a fusión o a uno del al menos un segundo candidato a fusión; y
- 20 codificar el bloque actual utilizando el correspondiente de la pluralidad de primeros candidatos a fusión o uno del al menos un segundo candidato a fusión.
8. El aparato de codificación de imágenes según la reivindicación 7, en el que la codificación del índice codificado utilizando el número fijo de candidatos comprende la codificación del índice codificado utilizando una codificación de longitud variable en base al número fijo de candidatos.

FIG. 1A

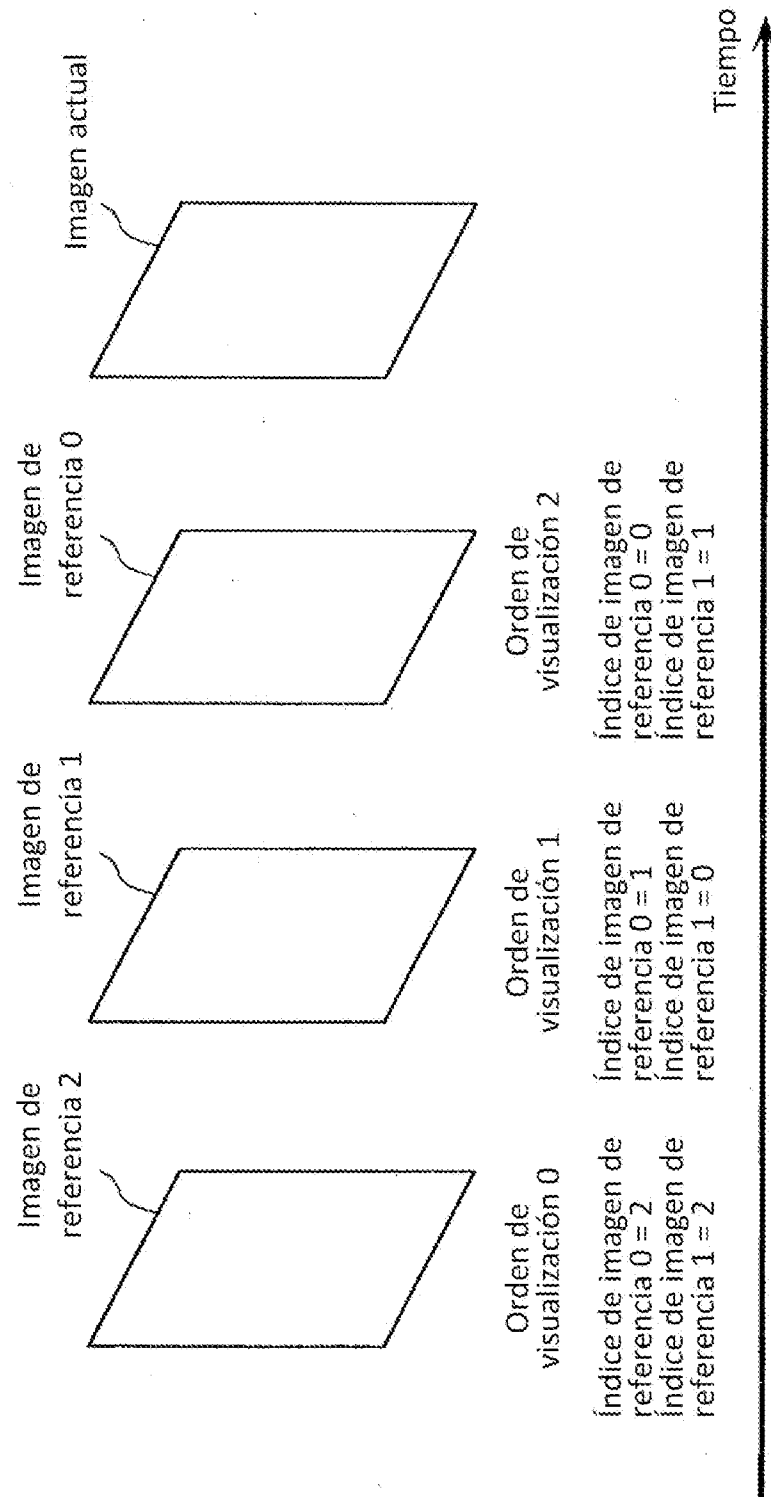


FIG. 1B

Lista de imágenes de referencia 0

Índice de imagen de ref. 0	Orden de visualización
0	2
1	1
2	0

FIG. 1C

Lista de imágenes de referencia 1

Índice de imagen de ref. 1	Orden de visualización
0	1
1	2
2	0

FIG. 2

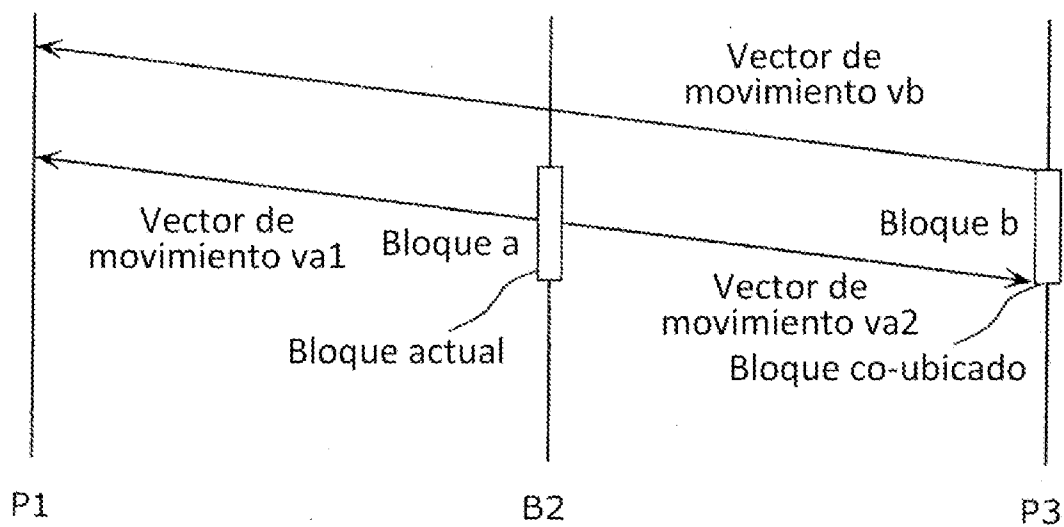


FIG. 3

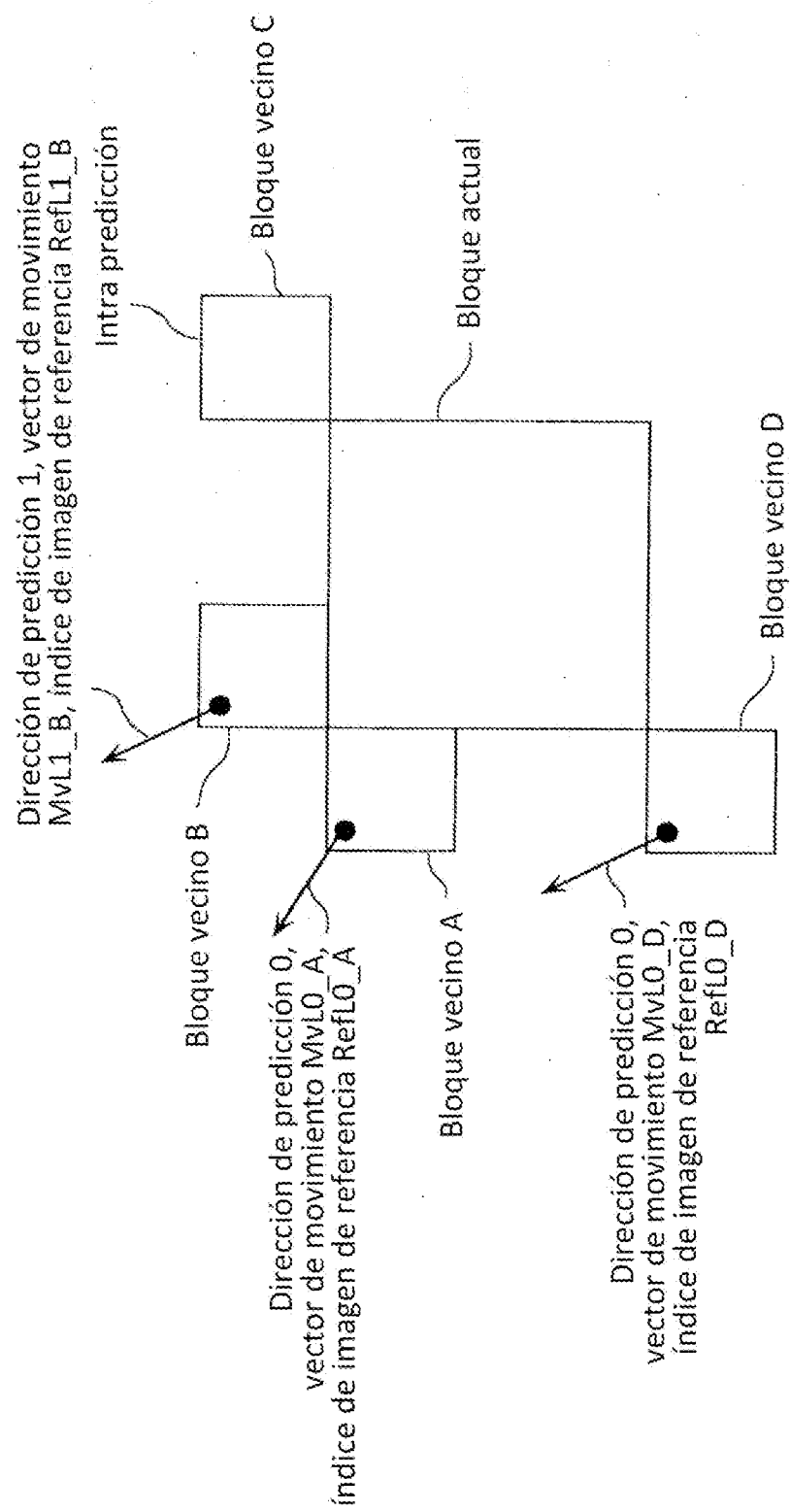


FIG. 4

Índice de candidato a bloque de fusión	Candidato a bloque de fusión
0	Bloque vecino A (dirección de predicción 0, (MvL0_A, RefL0_A))
1	Bloque vecino B (dirección de predicción 1, (MvL1_B, RefL1_B))
2	Bloque de fusión co-ubicado (predicción bi-direccional, (MvL0_Col, RefL0_Col), (MvL1_Col, RefL1_Col))
3	Candidato inutilizable para la fusión (bloque vecino C, intra predicción)
4	Bloque vecino D (dirección de predicción 0, (MvL0_D, RefL0_V))

Cuando MvL0_A = MvL0_D y RefL0_A = RefL0_D, eliminar uno de los bloques vecinos A y D

Eliminar candidato idéntico y candidato inutilizable para la fusión

Índice de candidato a bloque de fusión	Candidato a bloque de fusión
0	Bloque vecino A (dirección de predicción 0, (MvL0_A, RefL0_A))
1	Bloque vecino B (dirección de predicción 1, (MvL1_B, RefL1_B))
2	Bloque de fusión co-ubicado (predicción bi-direccional, (MvL0_Col, RefL0_Col), (MvL1_Col, RefL1_Col))

Número total de candidatos a bloques de fusión = 3

FIG. 5

Tamaño de la lista de candidatos a bloques de fusión = 2

Índice de candidato a bloque de fusión	Secuencia de bits asignada
0	0
1	1

Tamaño de la lista de candidatos a bloques de fusión = 3

Índice de candidato a bloque de fusión	Secuencia de bits asignada
0	0
1	10
2	11

Tamaño de la lista de candidatos a bloques de fusión = 4

Índice de candidato a bloque de fusión	Secuencia de bits asignada
0	0
1	10
2	110
3	111

Tamaño de la lista de candidatos a bloques de fusión = 5

Índice de candidato a bloque de fusión	Secuencia de bits asignada
0	0
1	10
2	110
3	1110
4	1111

FIG. 6

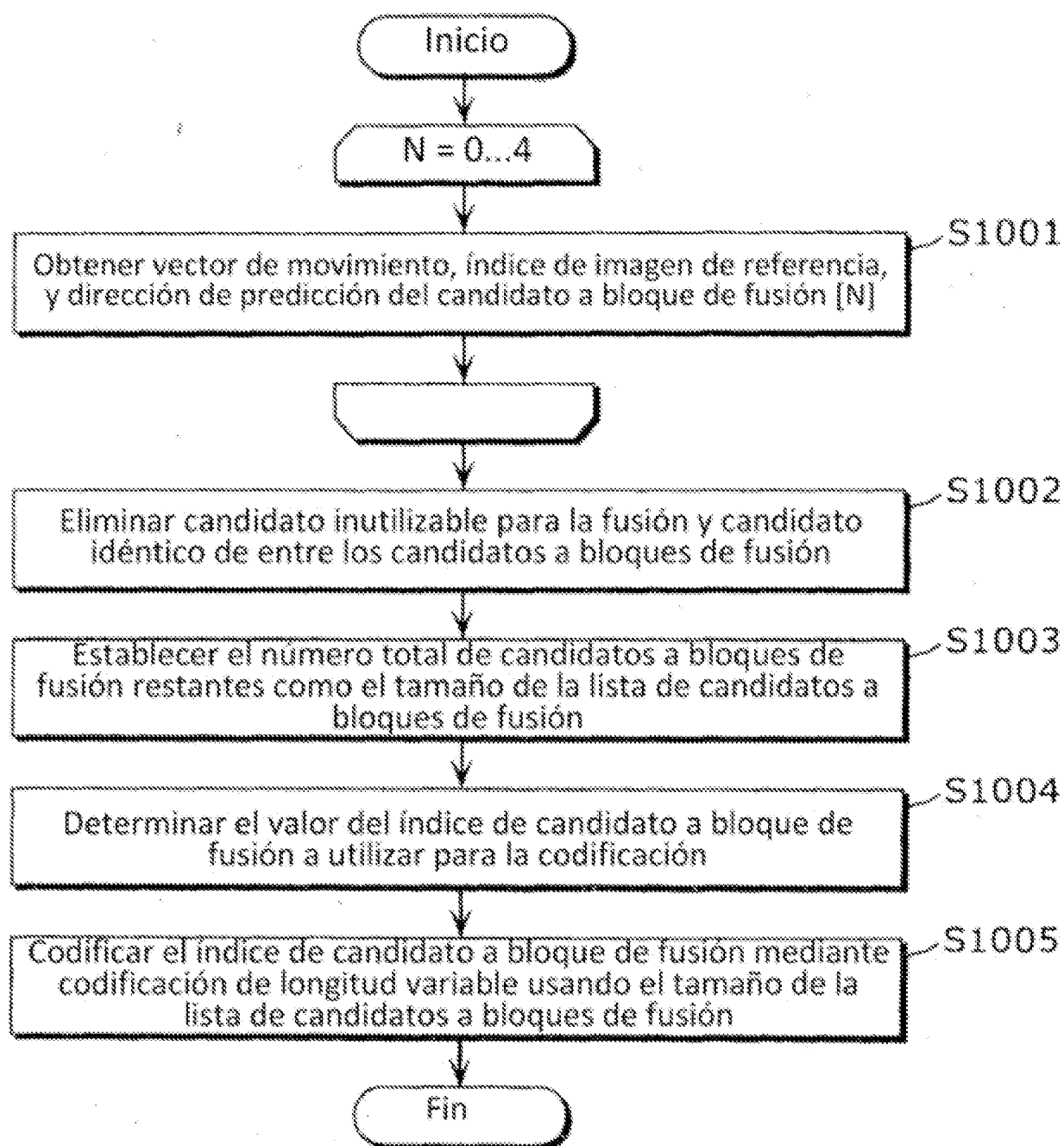


FIG. 7

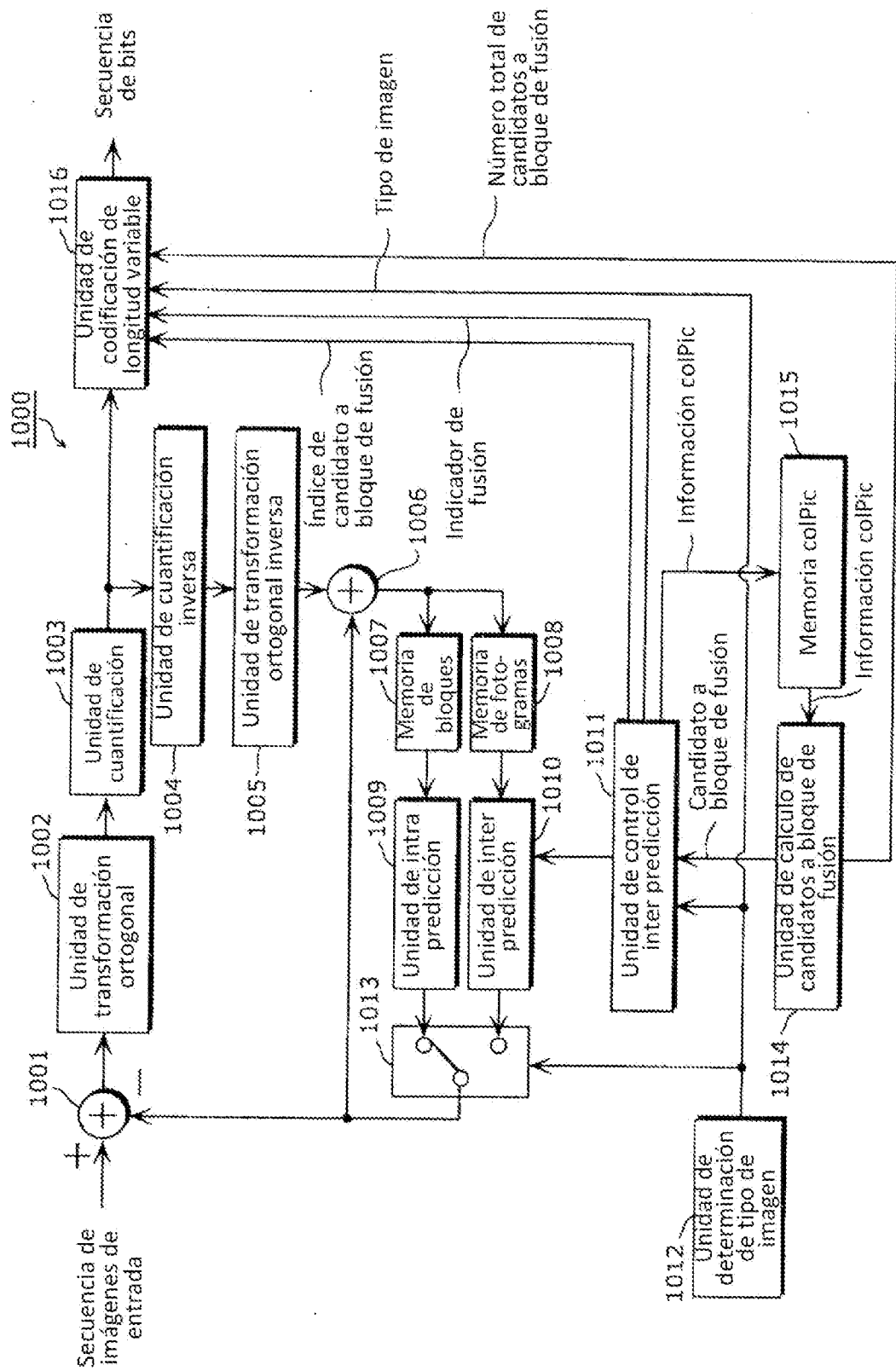


FIG. 8

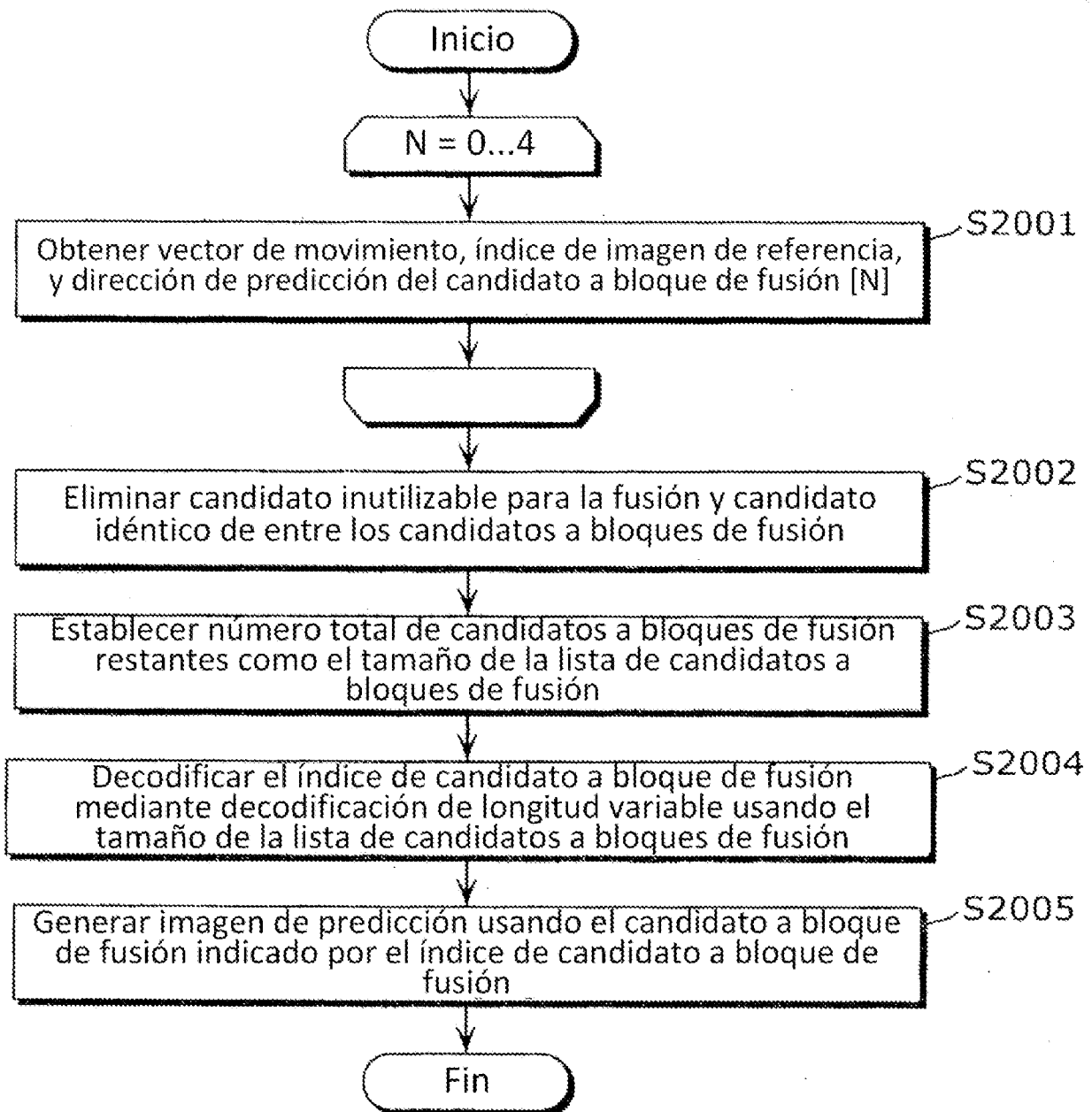


FIG. 9

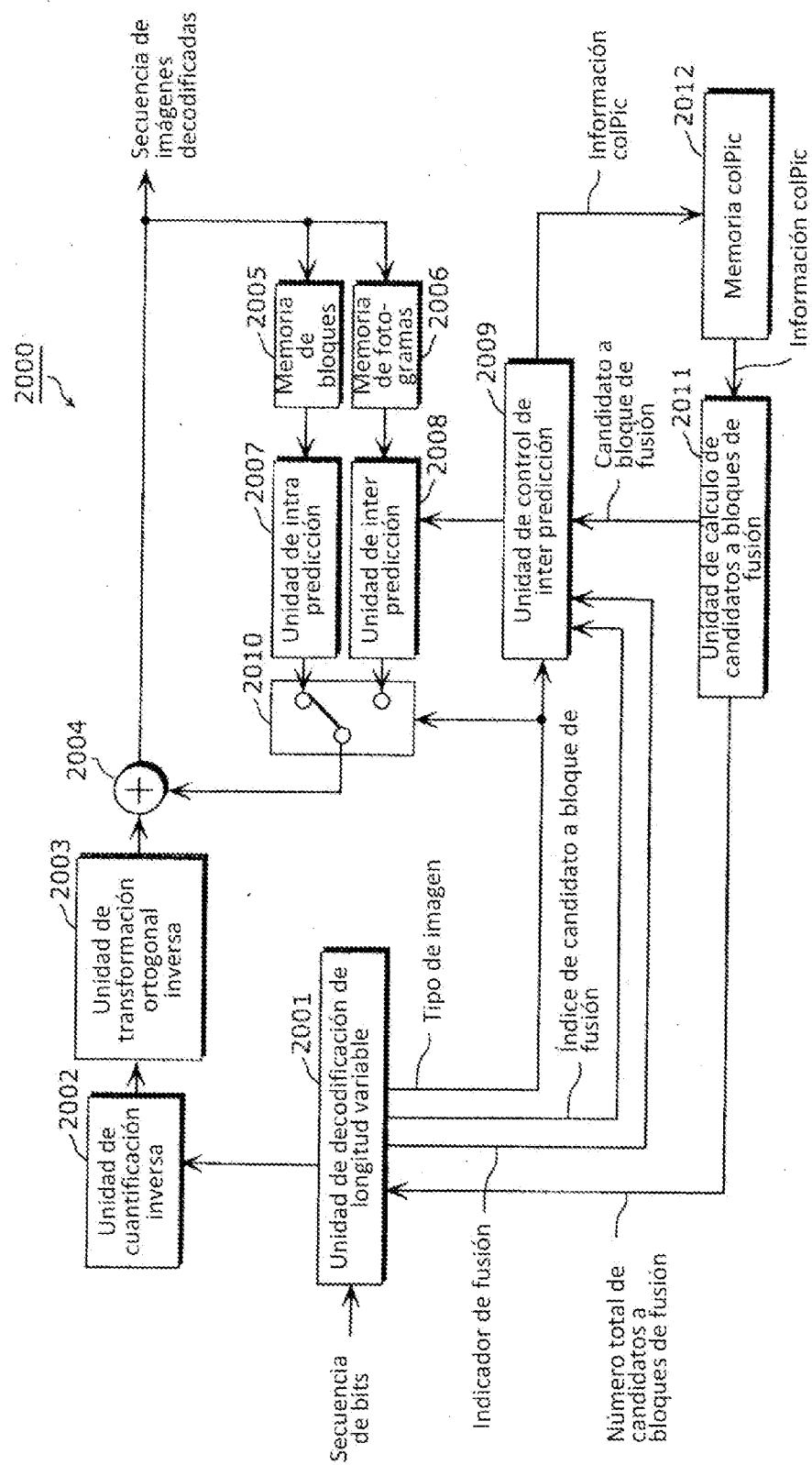


FIG. 10

Tamaño de lista de candidatos a bloques de fusión = número total de candidatos a bloques de fusión

prediction_unit(x0, y0, log2PUWidth, log2PUHeight, PartIdx, InferredMergeFlag){	Descriptor
if(skip_flag[x0][y0]){	
if(NumMergeCand > 1)	
merge_idx[x0][y0]	ue(v) ae(v)
} else if(PredMode == MODE_INTRA){	
...	u(l) ae(v)
} else { /* MODE_INTER */	
if(!InferredMergeFlag)	
merge_flag[x0][y0]	u(l) ae(v)
if(merge_flag[x0][y0] && NumMergeCand > 1){	
merge_idx[x0][y0]	ue(v) ae(v)
} else {	
...	
}	
}	

Índice de candidato a bloque de fusión

Indicador de fusión

Índice de candidato a bloque de fusión

Tamaño de la lista de candidatos a bloques de fusión = número total de candidatos a bloques de fusión

FIG. 11

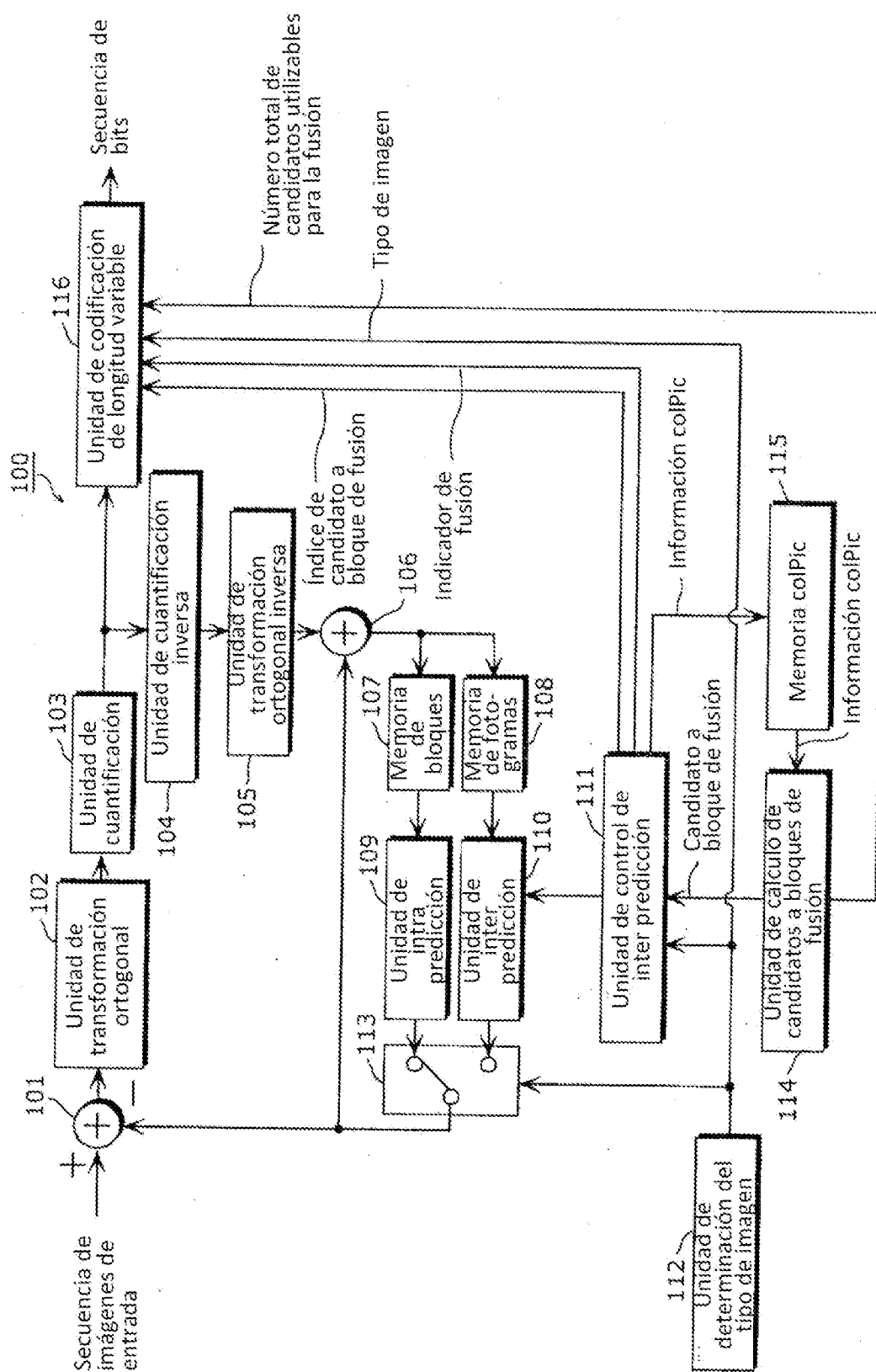


FIG. 12

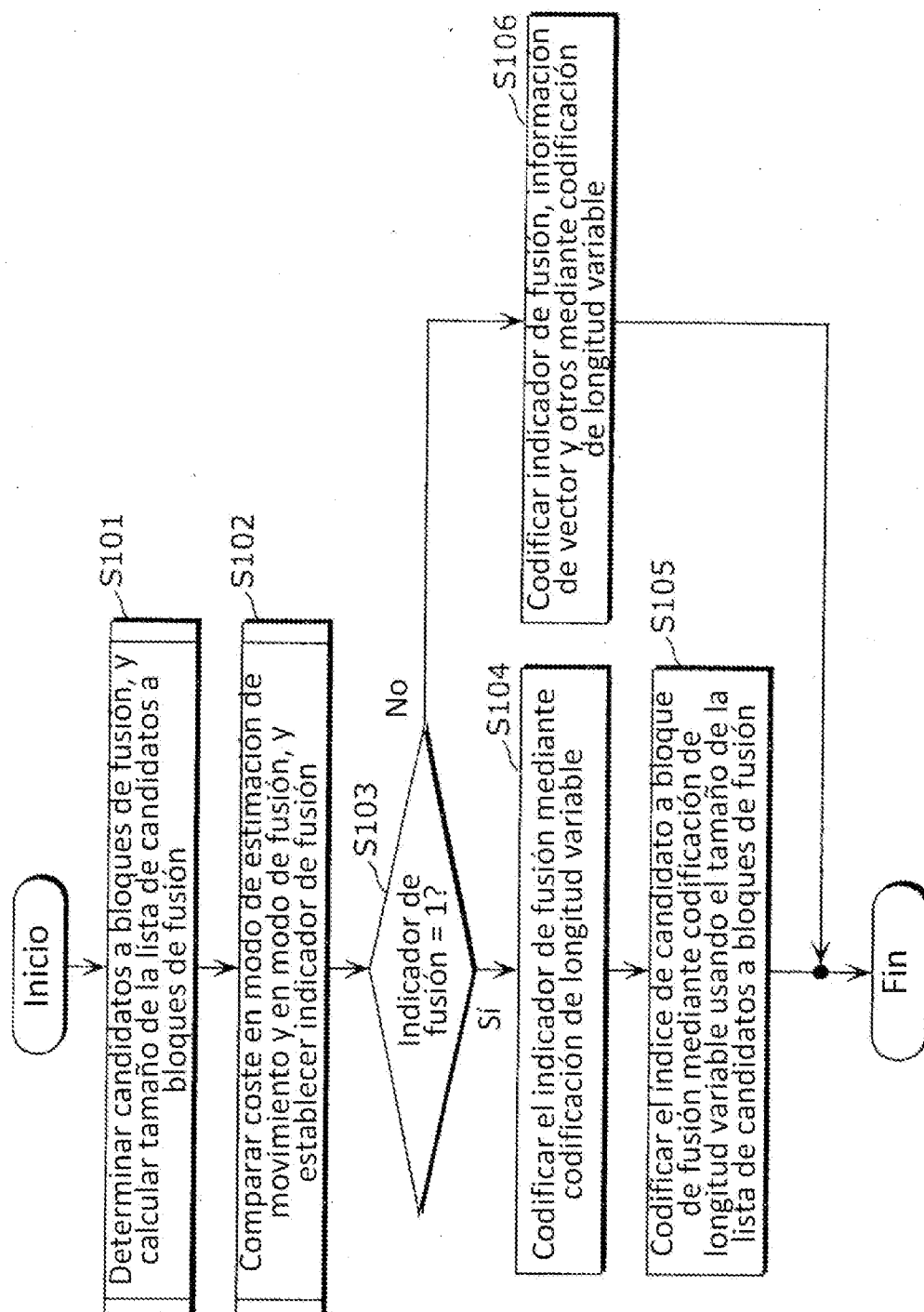


FIG. 13

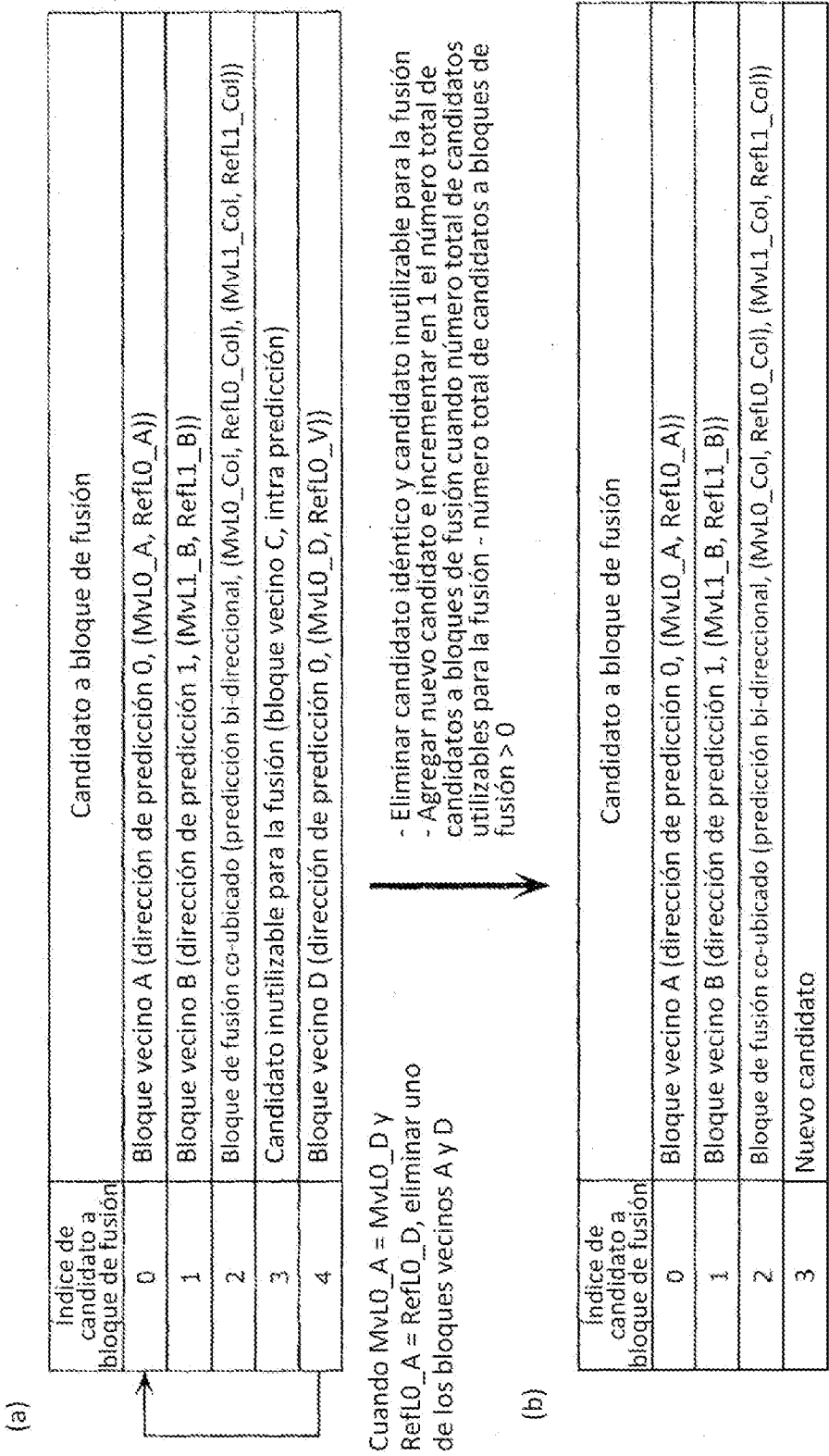


FIG. 14

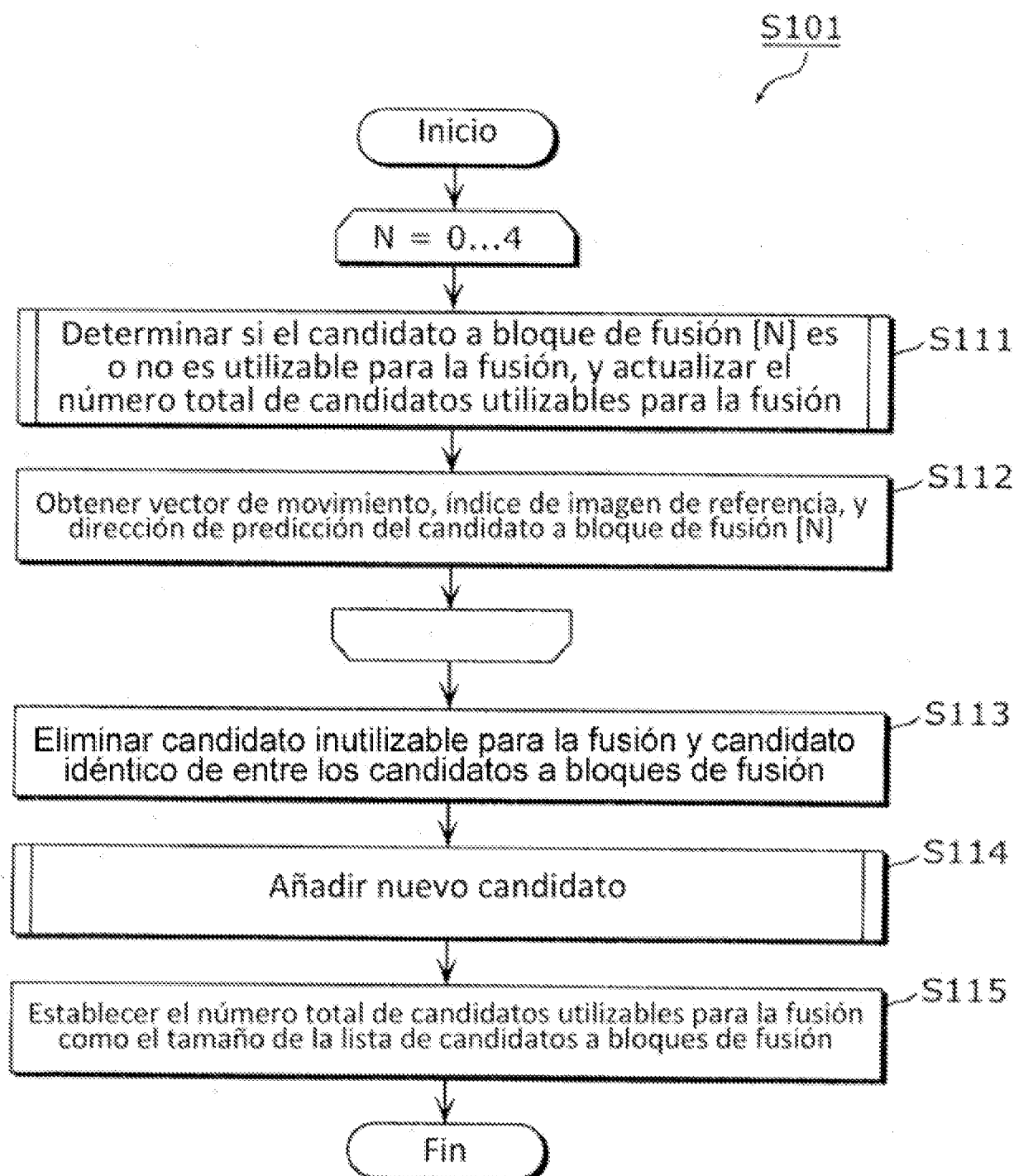


FIG. 15

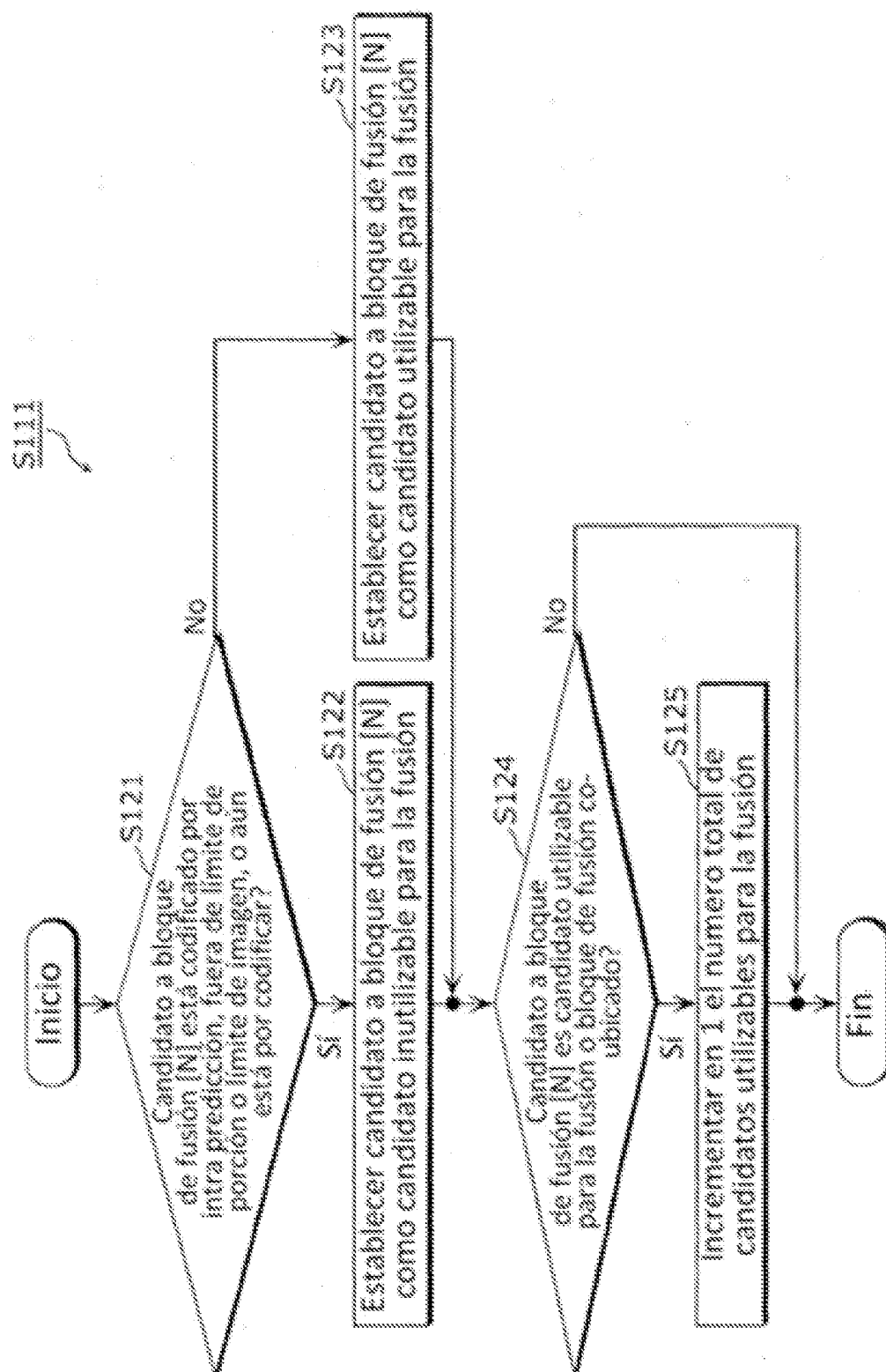


FIG. 16

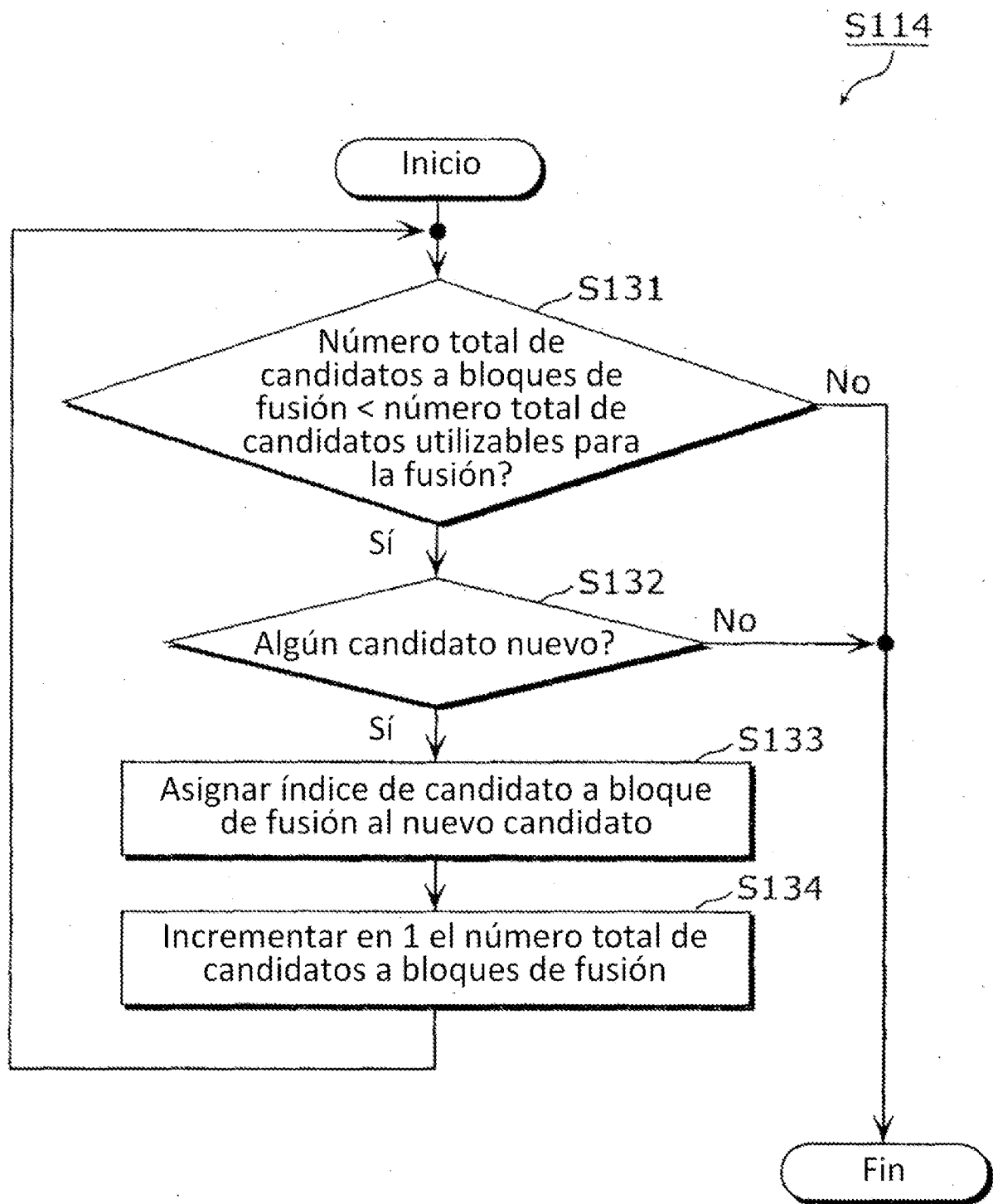


FIG. 17

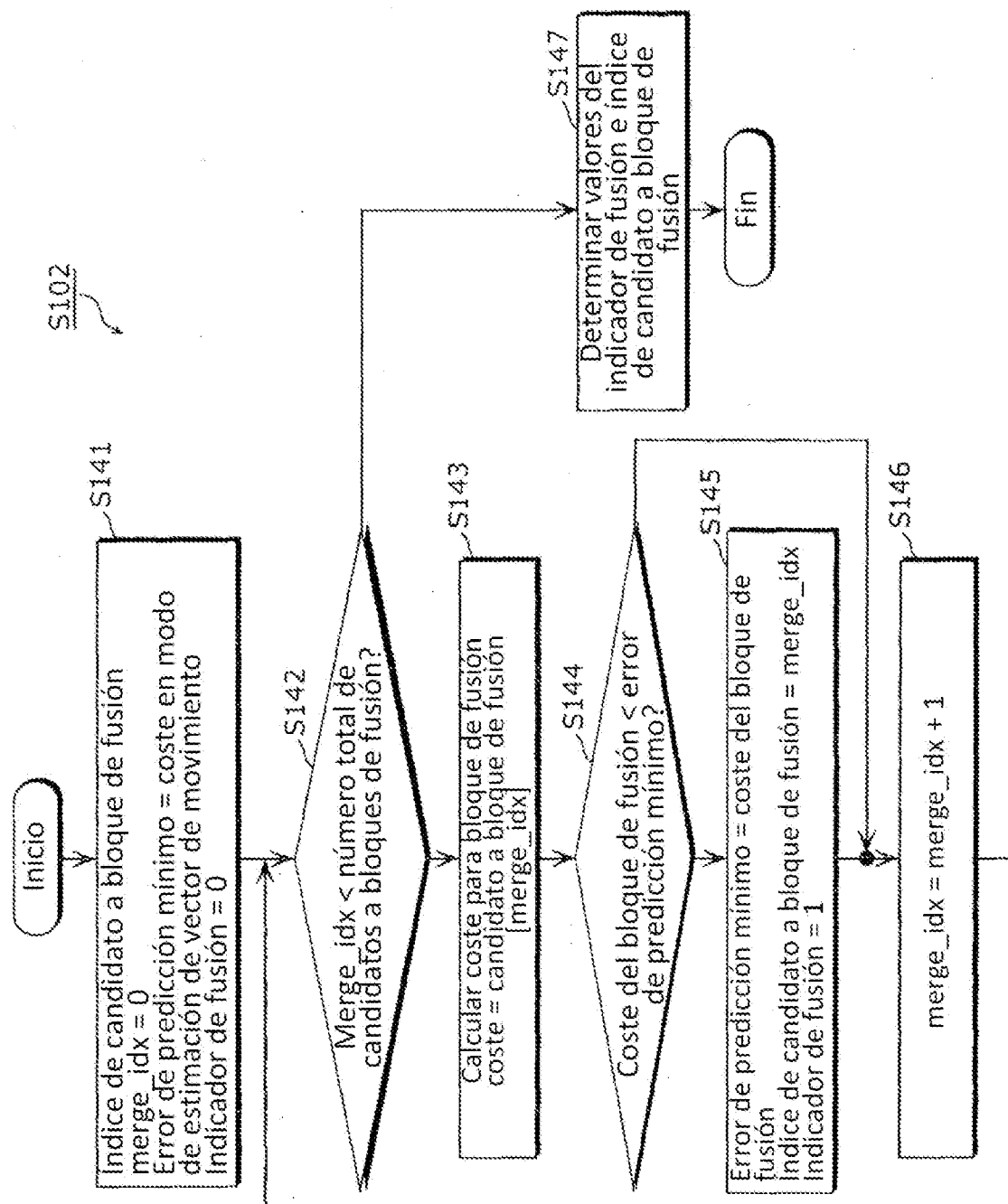


FIG. 18

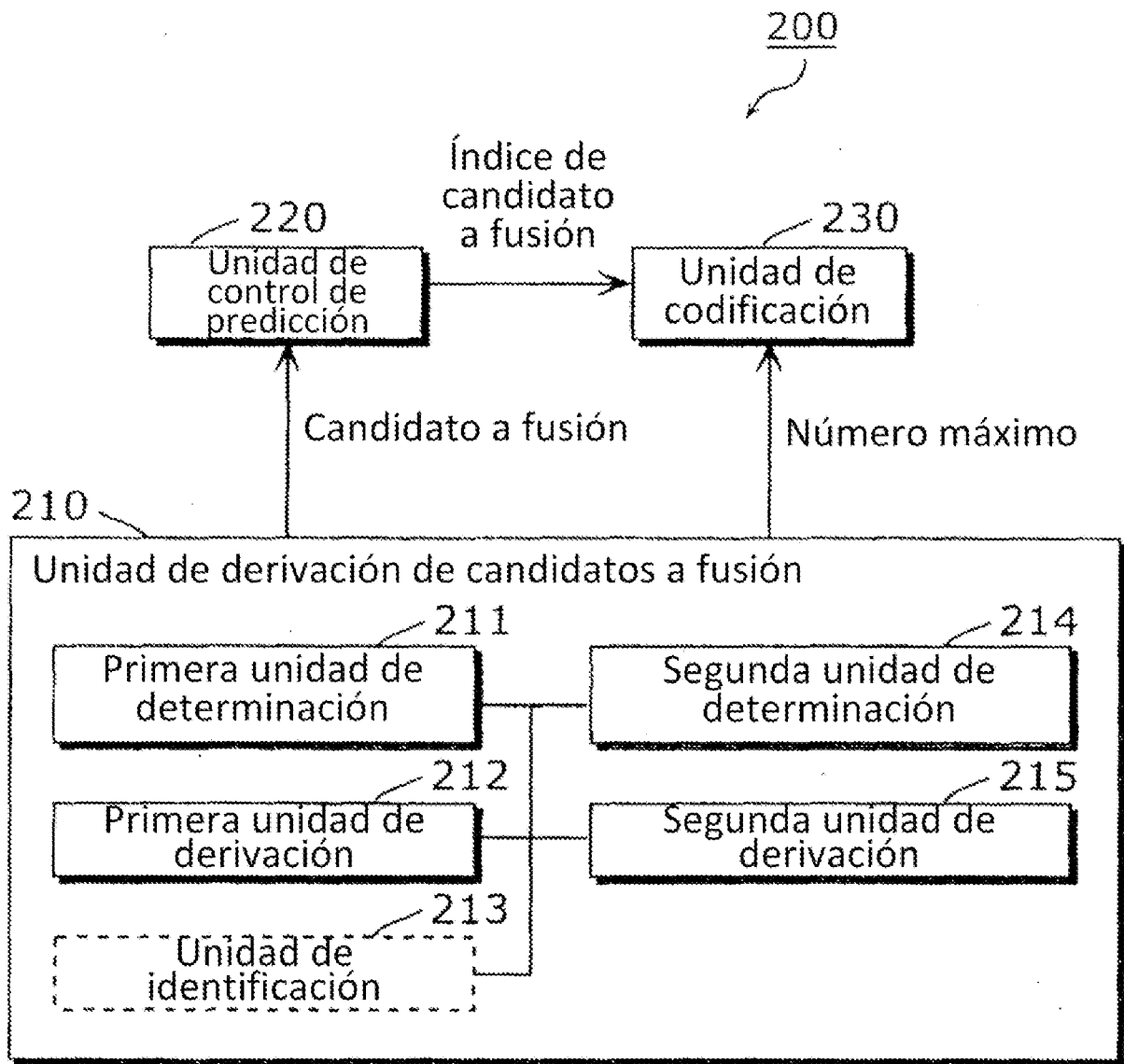


FIG. 19

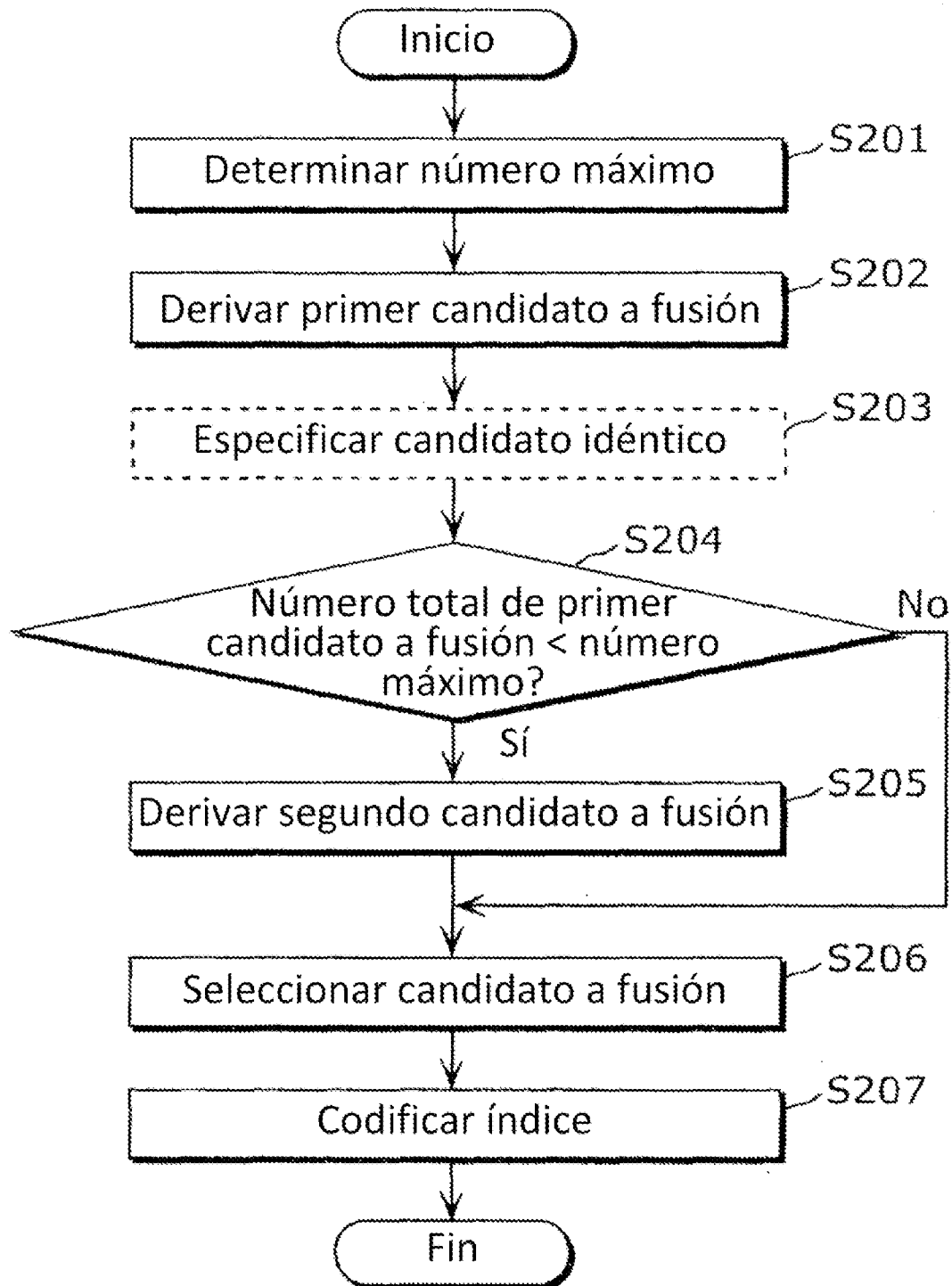


FIG. 20

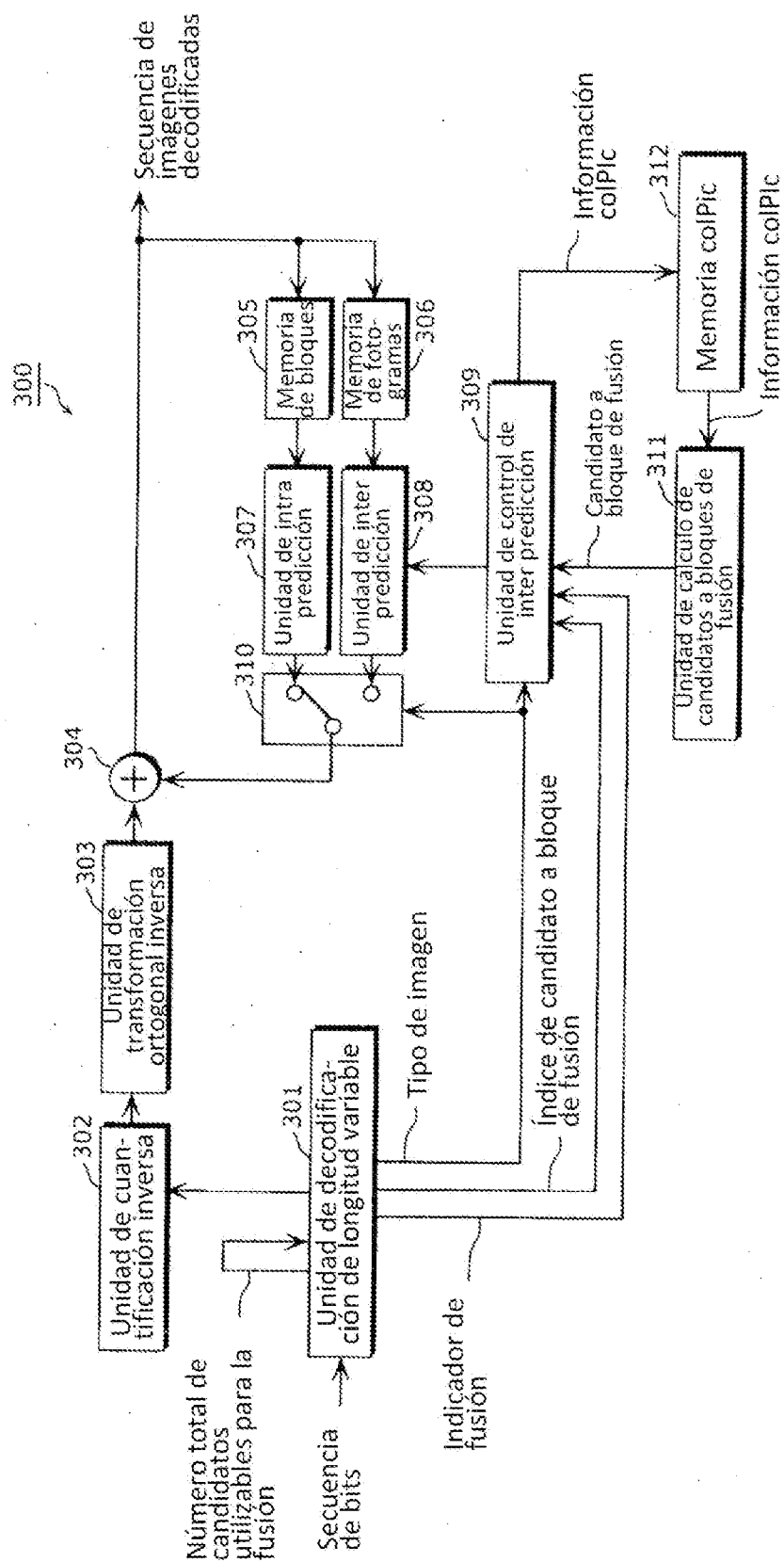


FIG. 21

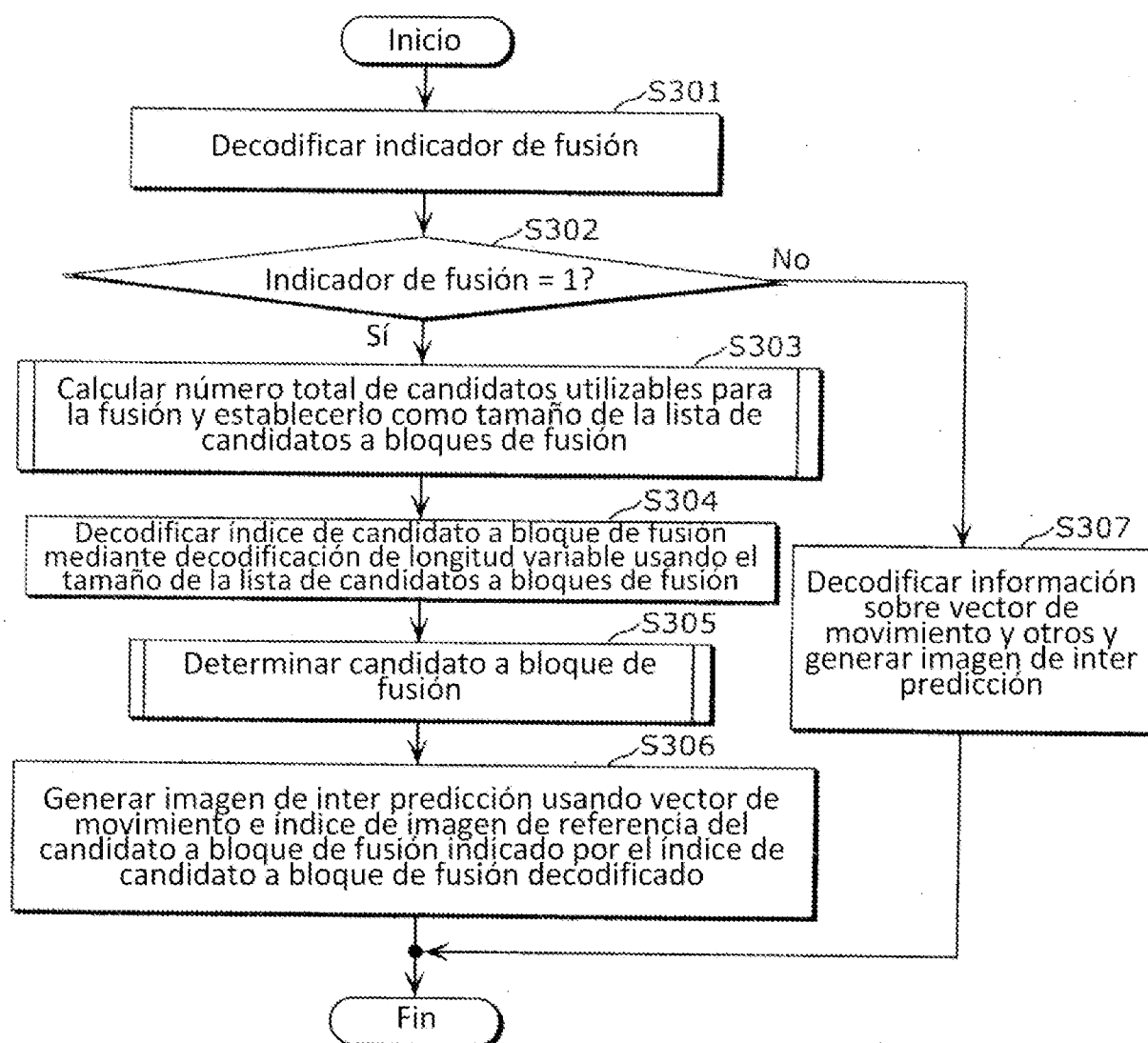


FIG. 22

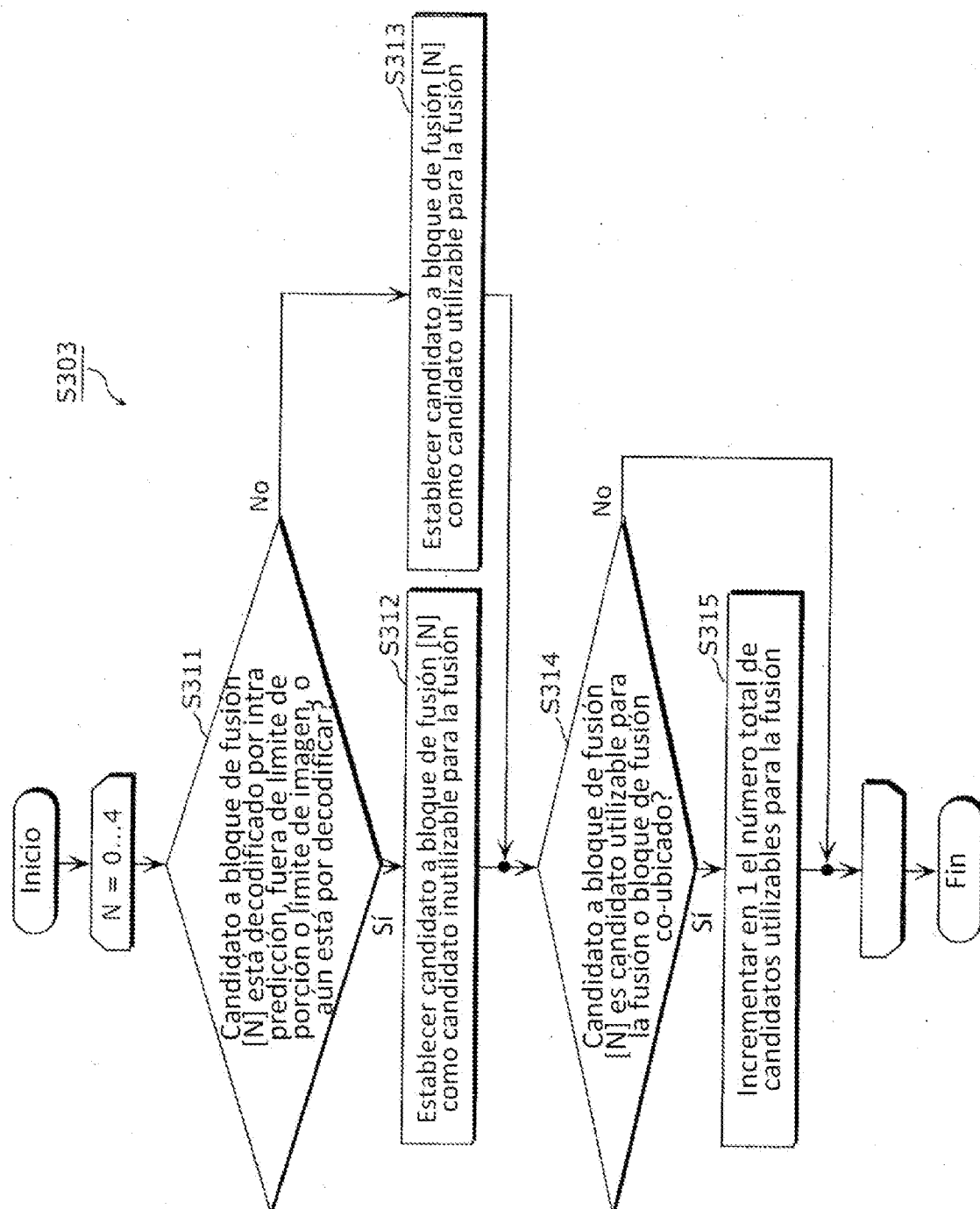


FIG. 23

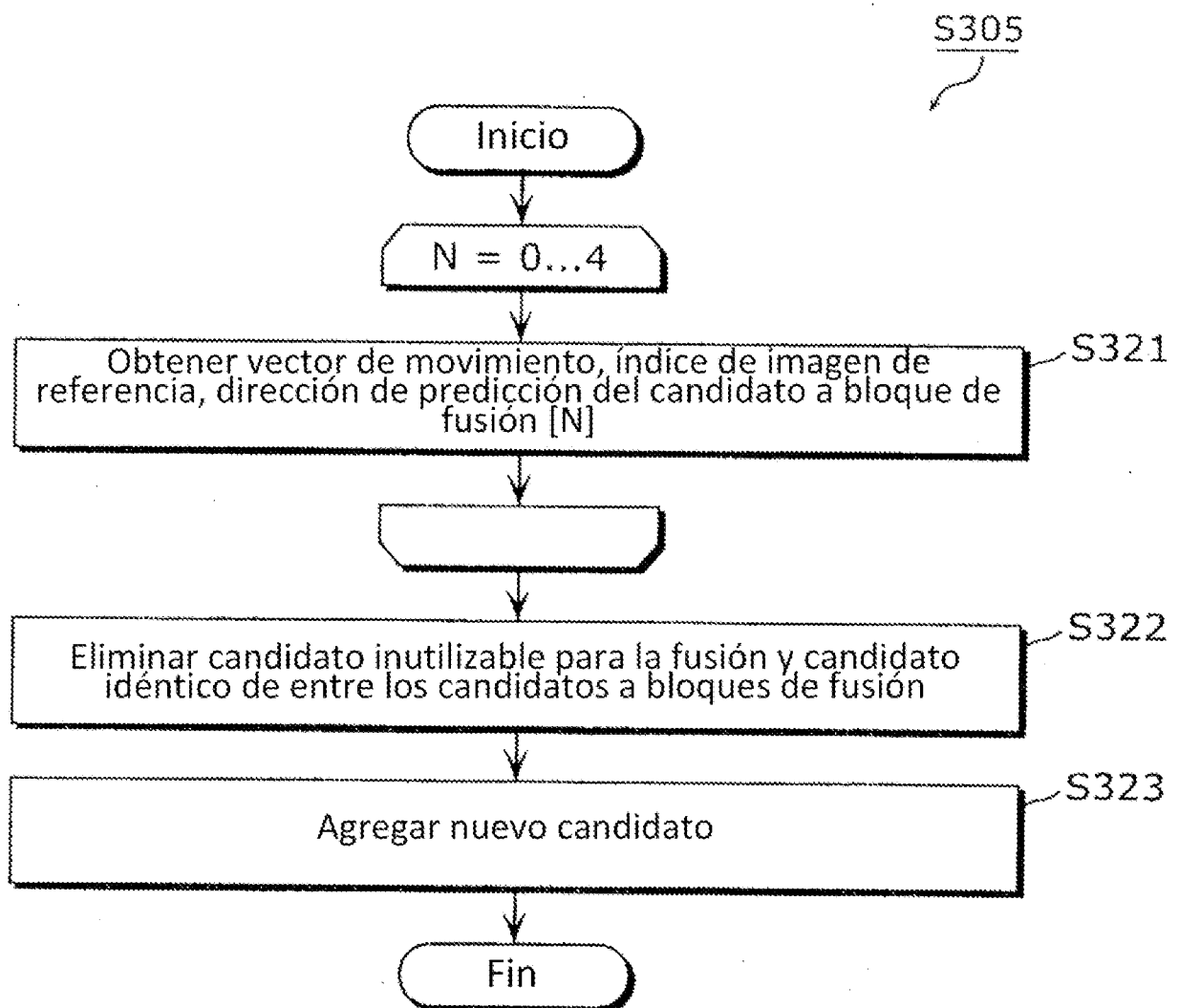


FIG. 24

Tamaño de la lista de candidatos a bloques de fusión = número total de candidatos utilizables para la fusión

prediction_unit(x0, y0, log2PUWidth, log2PUHeight, PartIdx, InferredMergeFlag){			Descriptor
if(skip_flag[x0][y0]){			
if(NumMergeCand > 1)			
merge_idx[x0][y0]			ue(v) ae(v)
} else if(PredMode == MODE_INTRA){			
...			
} else { /* MODE_INTER */			u(l) ae(v)
if(!InferredMergeFlag)			
merge_flag[x0][y0]			u(l) ae(v)
if(merge_flag[x0][y0] && NumMergeCand > 1){			
merge_idx[x0][y0]			ue(v) ae(v)
} else {			
...			
}			
}			

Índice de candidato a bloque de fusión

Indicador de fusión

Índice de candidato a bloque de fusión

Tamaño de la lista de candidatos a bloques de fusión = número total de candidatos utilizables para la fusión

FIG. 25

	prediction_unit(x0, y0, log2PUWidth, log2PUHeight, PartIdx, InferredMergeFlag){	Descriptor
	if(skip_flag[x0][y0]){	
Índice de candidato a bloque de fusión	merge_idx[x0][y0]	ue(v) ae(v)
	} else if(PredMode == MODE_INTRA){	
	...	u(l) ae(v)
	} else { /* MODE_INTER */	
	if(!InferredMergeFlag)	
Indicador de fusión	merge_flag[x0][y0]	u(l) ae(v)
	if(merge_flag[x0][y0]){	
Índice de candidato a bloque de fusión	merge_idx[x0][y0]	ue(v) ae(v)
	} else {	
	...	
	}	
	}	
	}	
	}	

FIG. 26

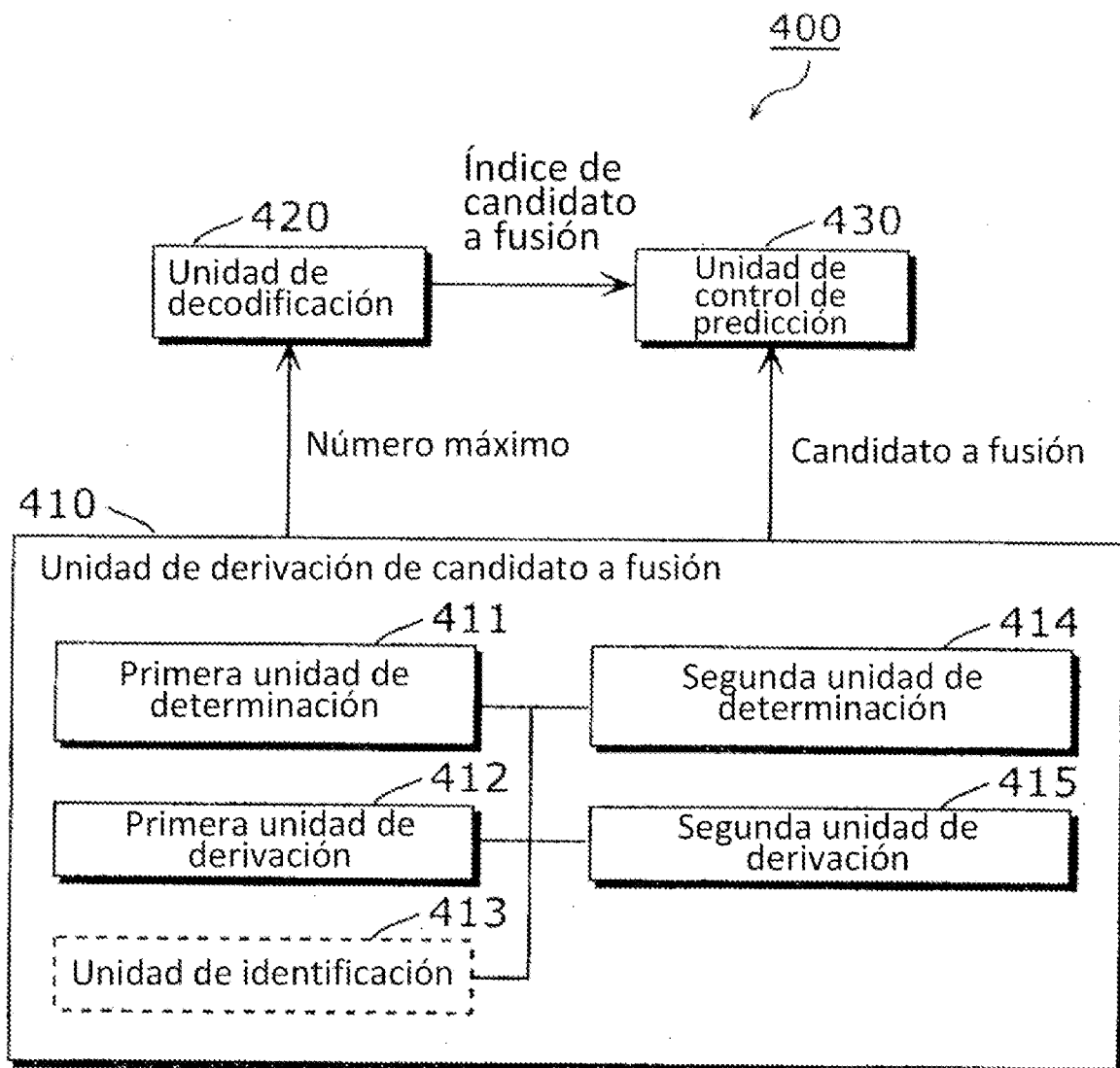


FIG. 27

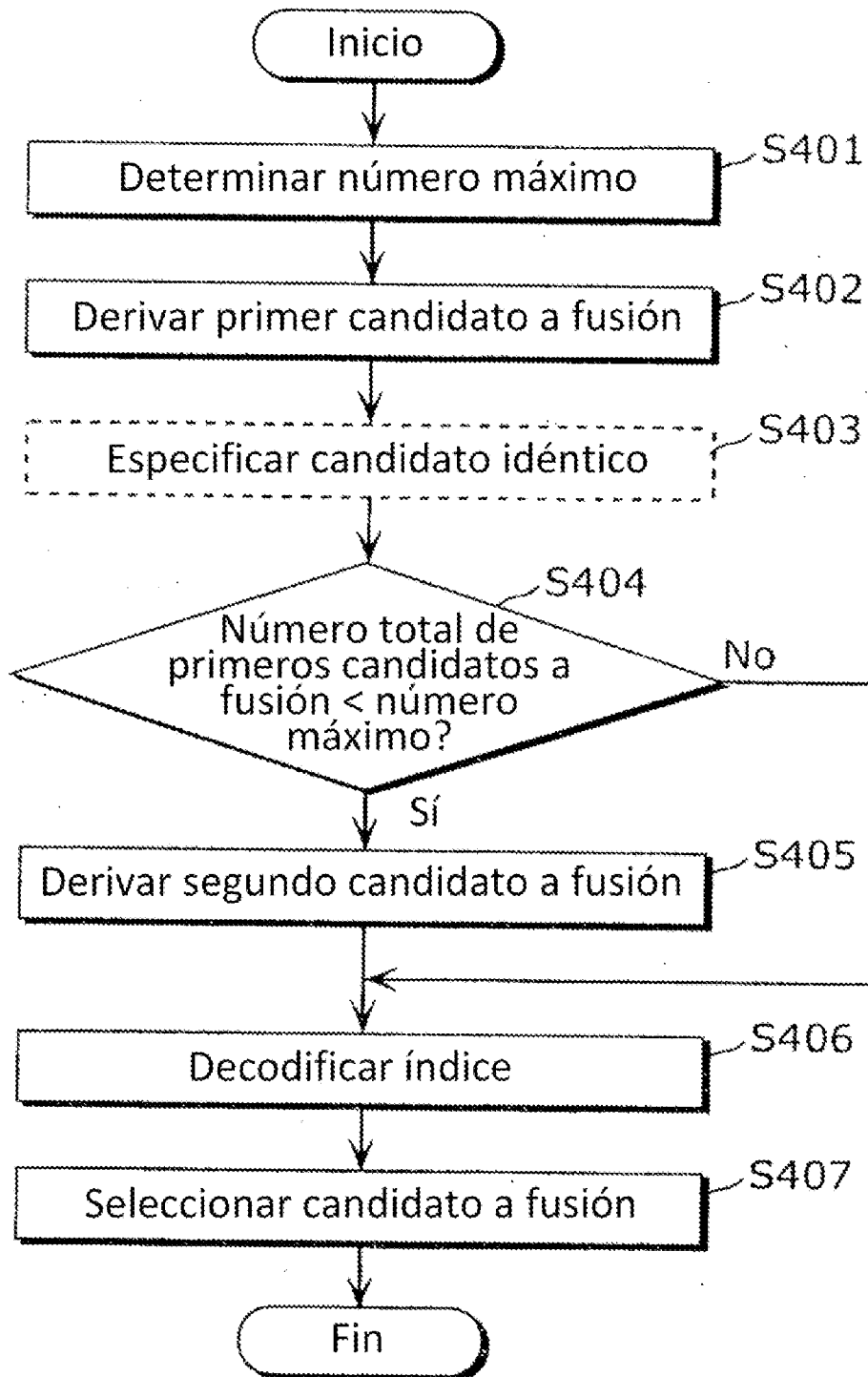


FIG. 28

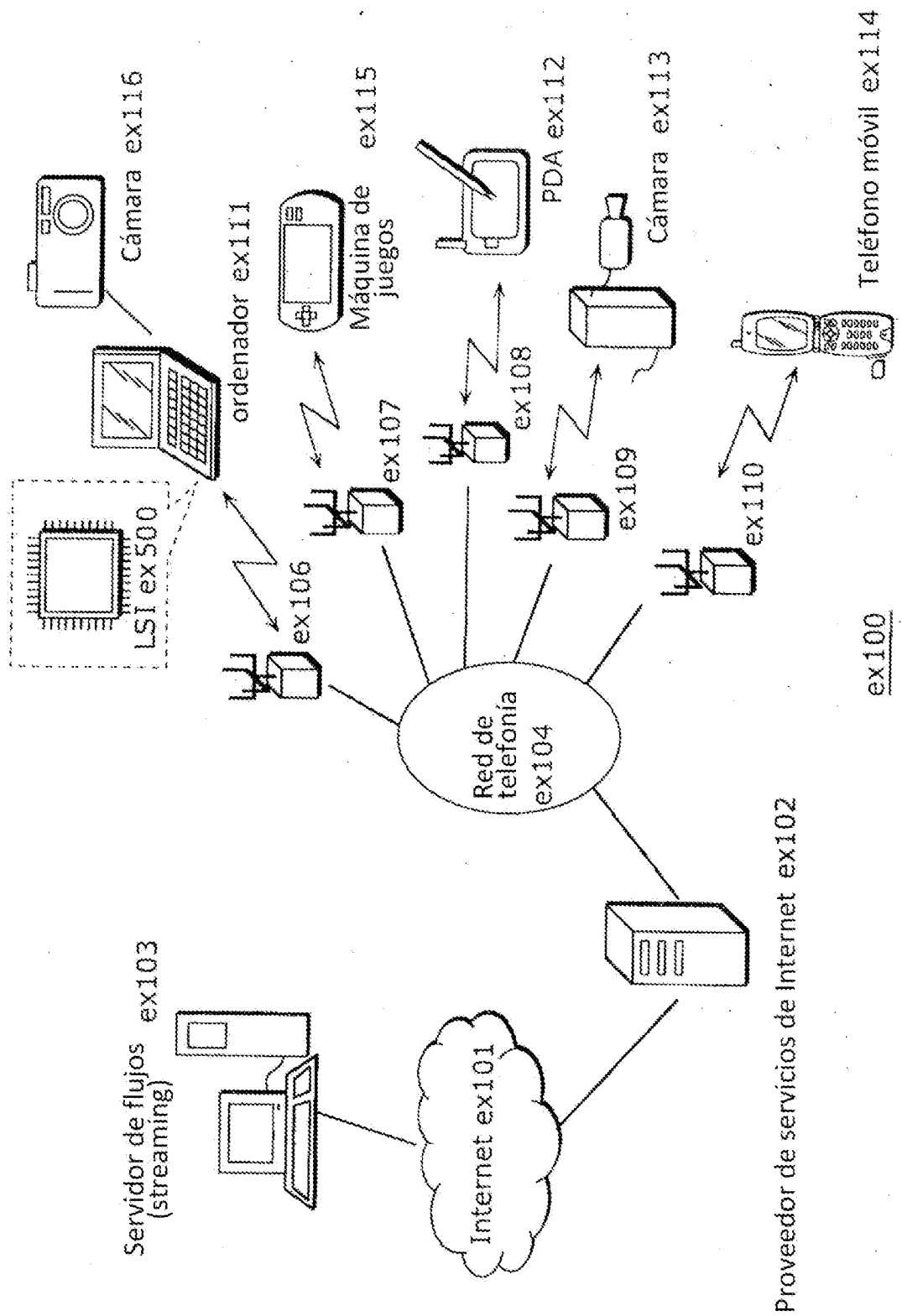


FIG. 29

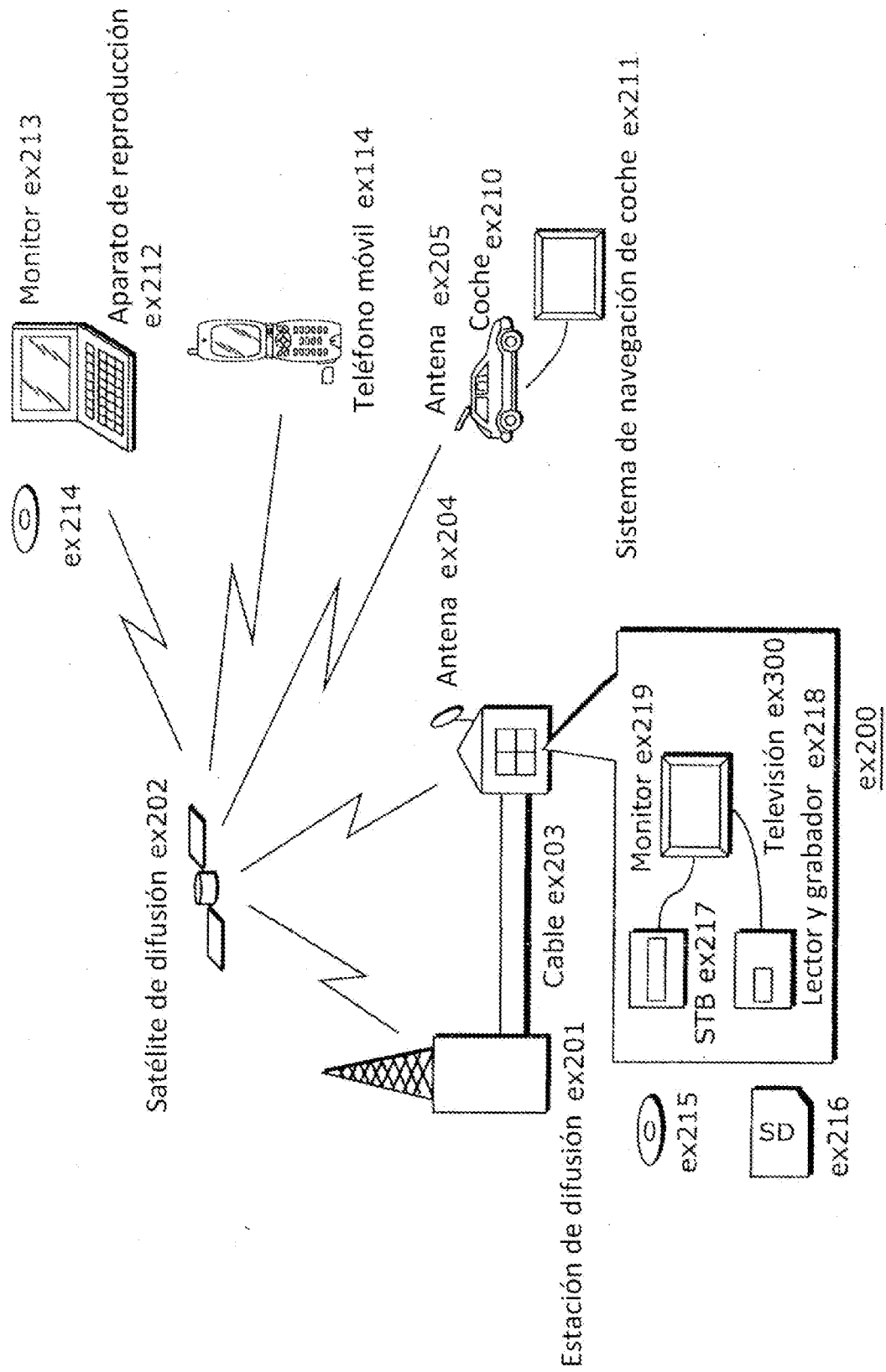


FIG. 30

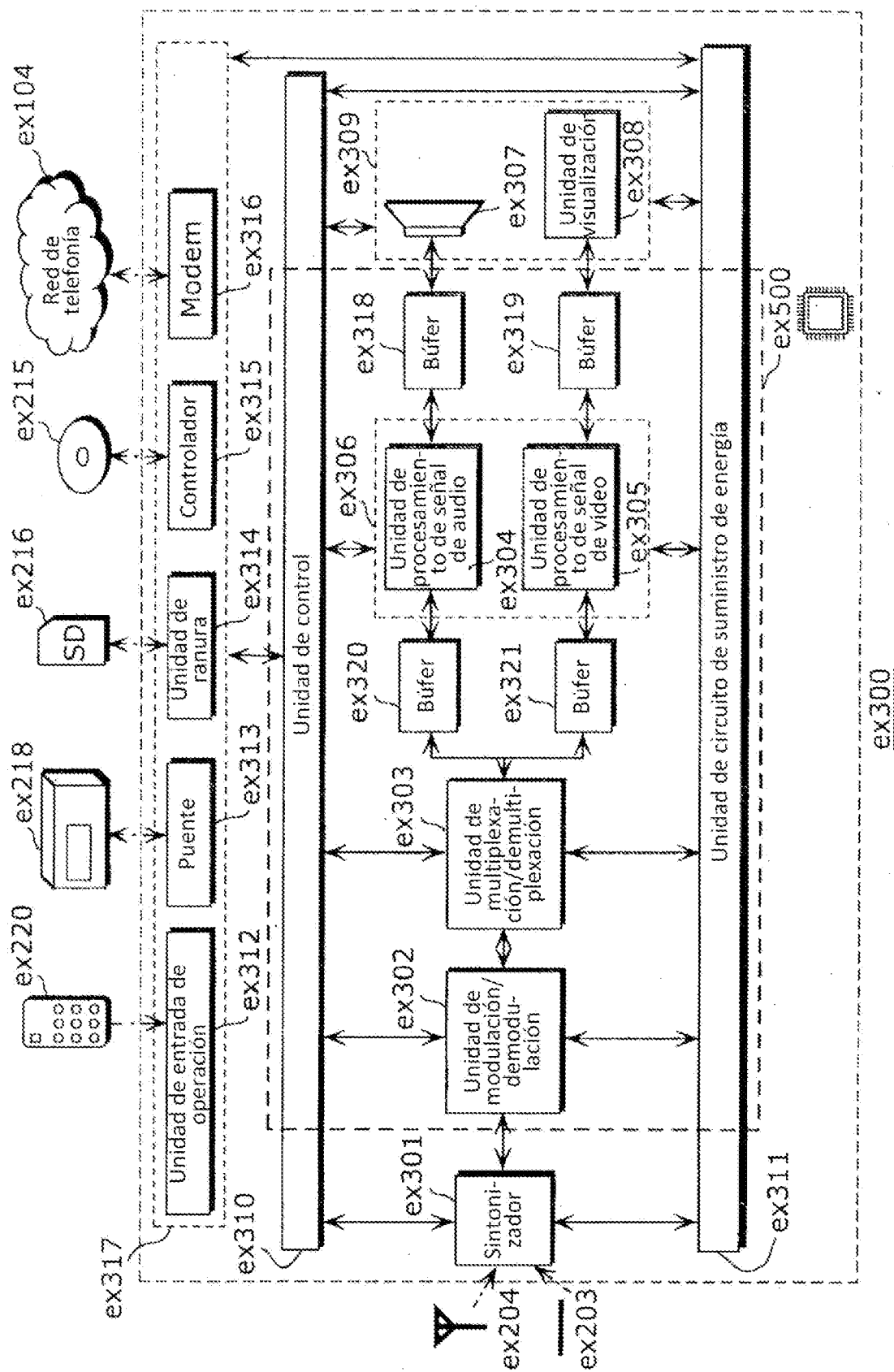


FIG. 31

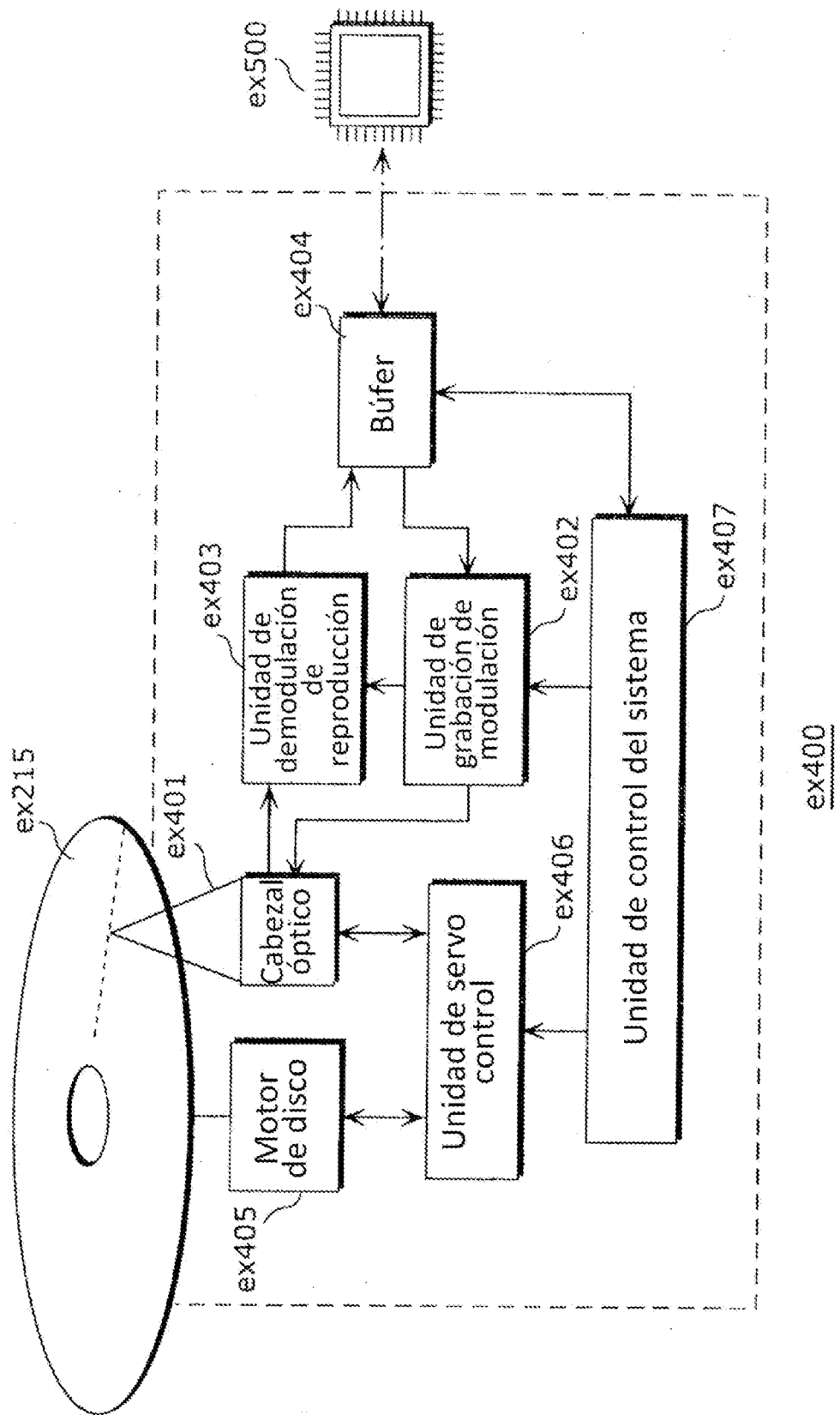


FIG. 32

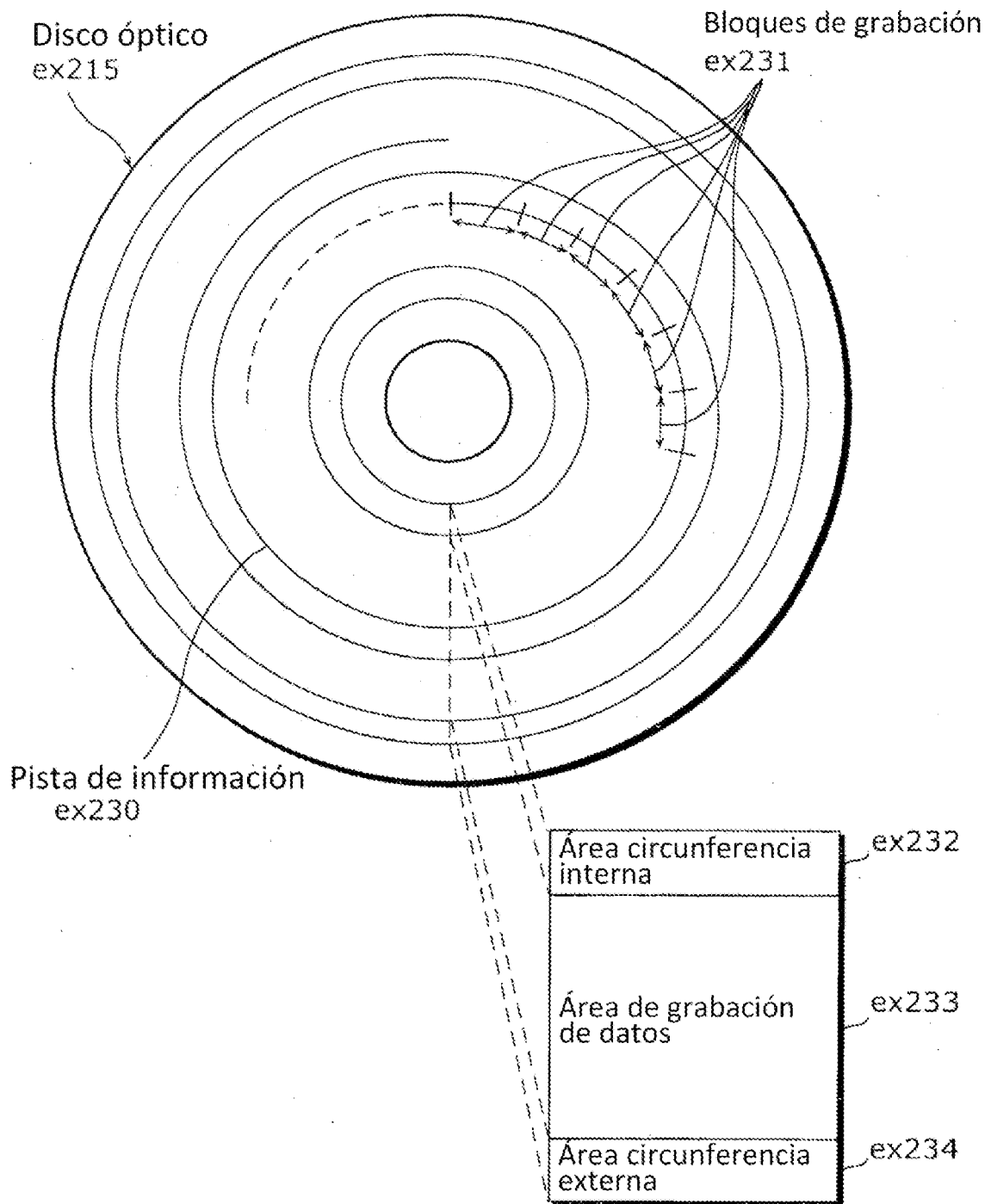


FIG. 33A

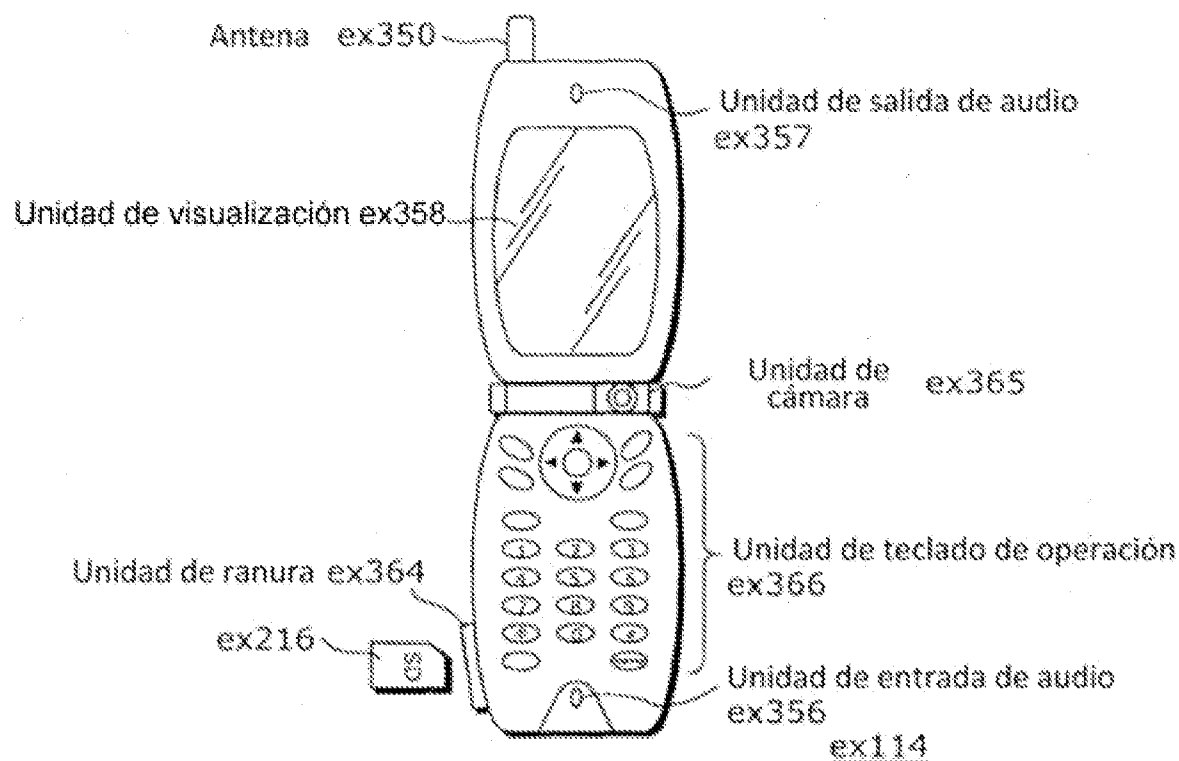


FIG. 33B

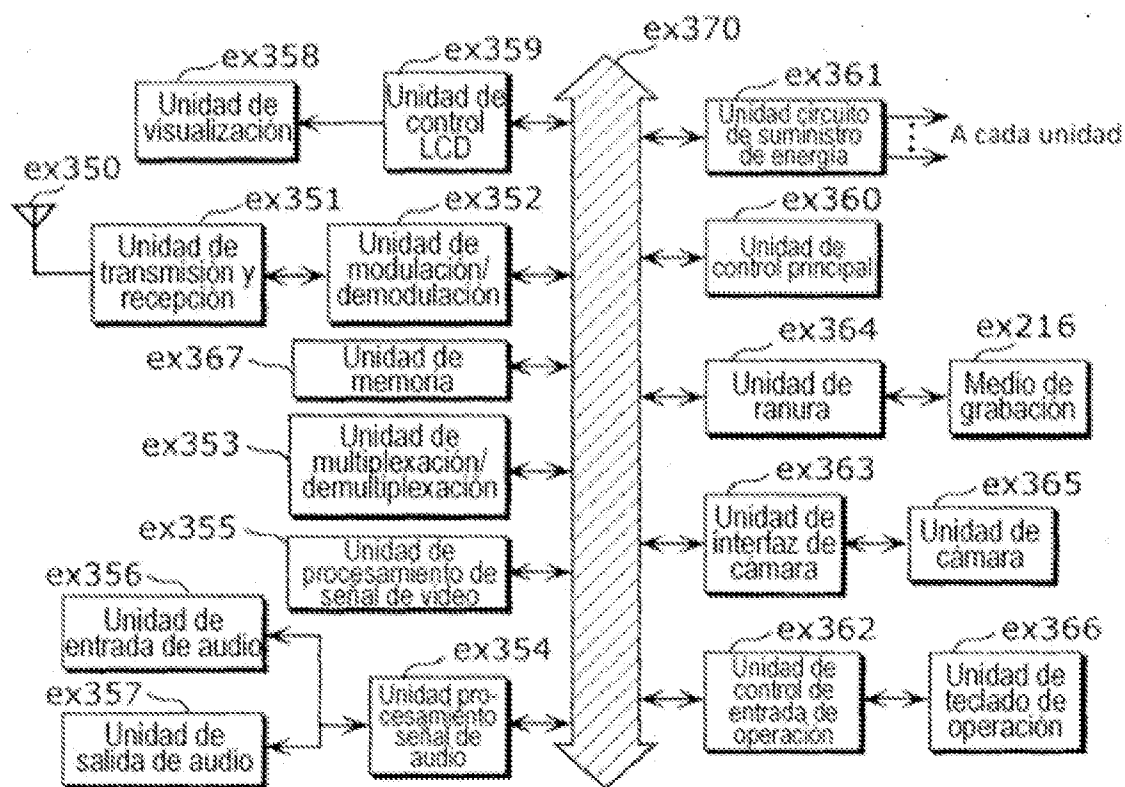


FIG. 34

Flujo de video (PID=0x1011, video principal)
Flujo de audio (PID=0x1100)
Flujo de audio (PID=0x1101)
Flujo de gráficos de presentación (PID=0x1200)
Flujo de gráficos de presentación (PID=0x1201)
Flujo de gráficos interactivos (PID=0x1400)
Flujo de video (PID=0x1B00, video secundario)
Flujo de video (PID=0x1B01, video secundario)

FIG. 35

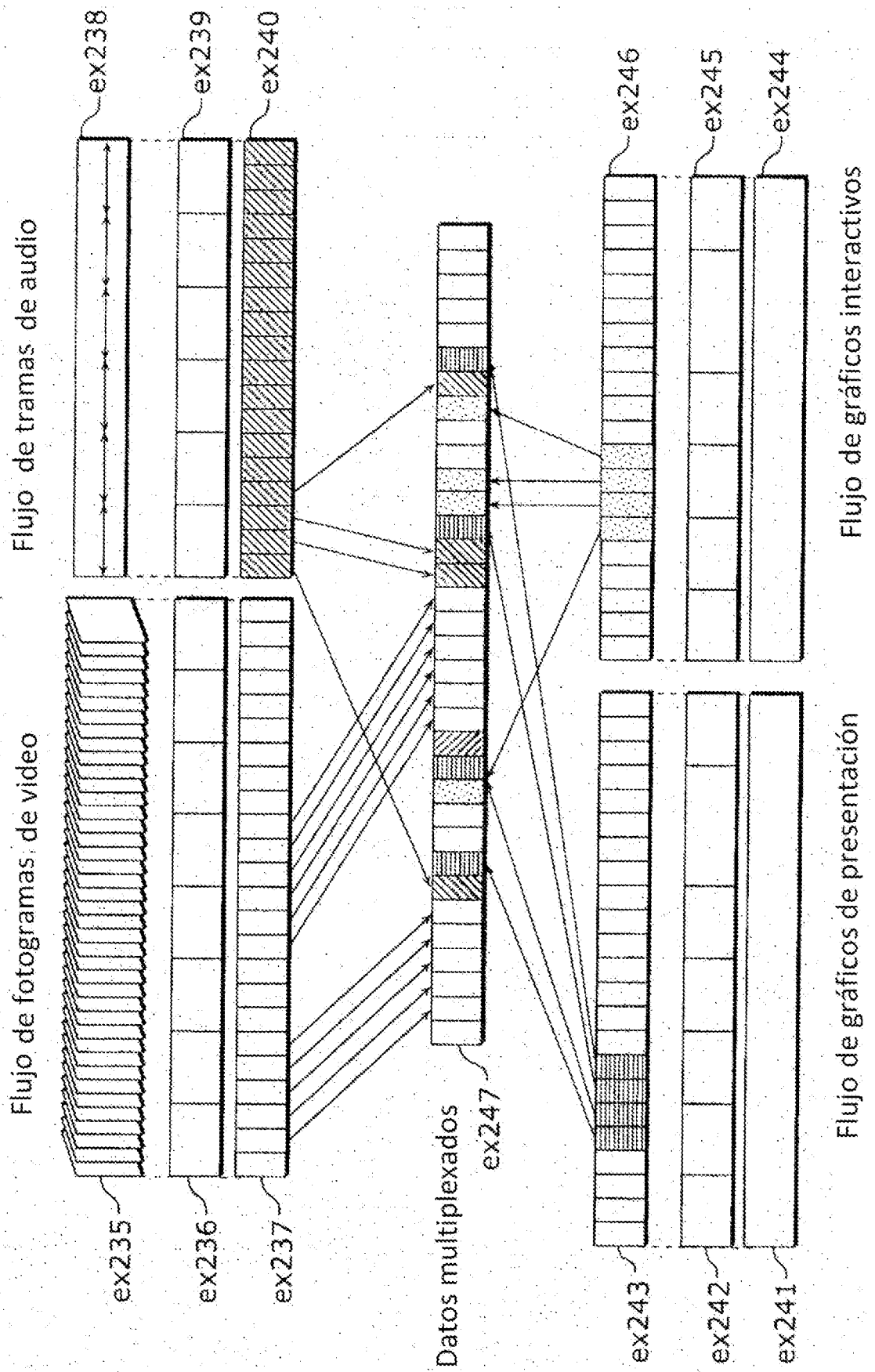


FIG. 36

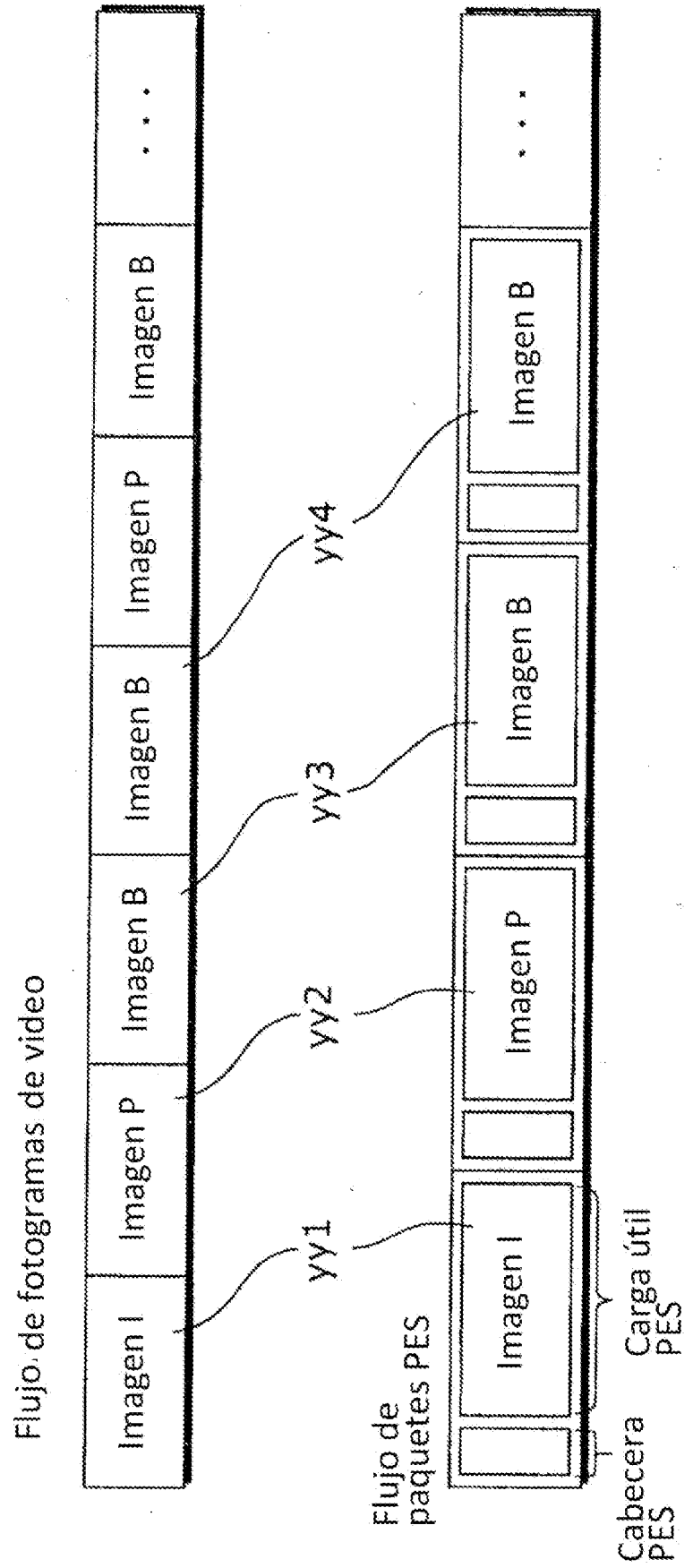


FIG. 37

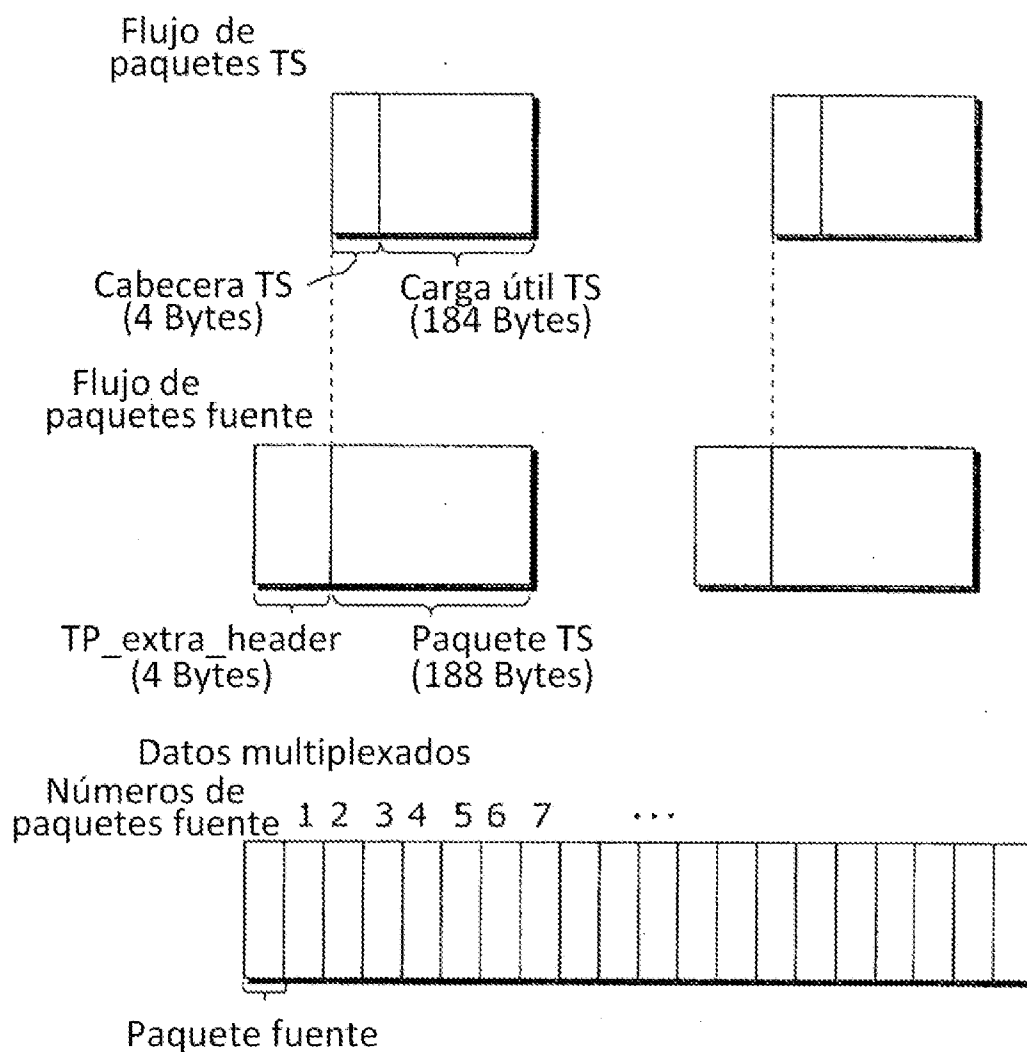


FIG. 38

Estructura de datos de tabla de mapas de programas PMT

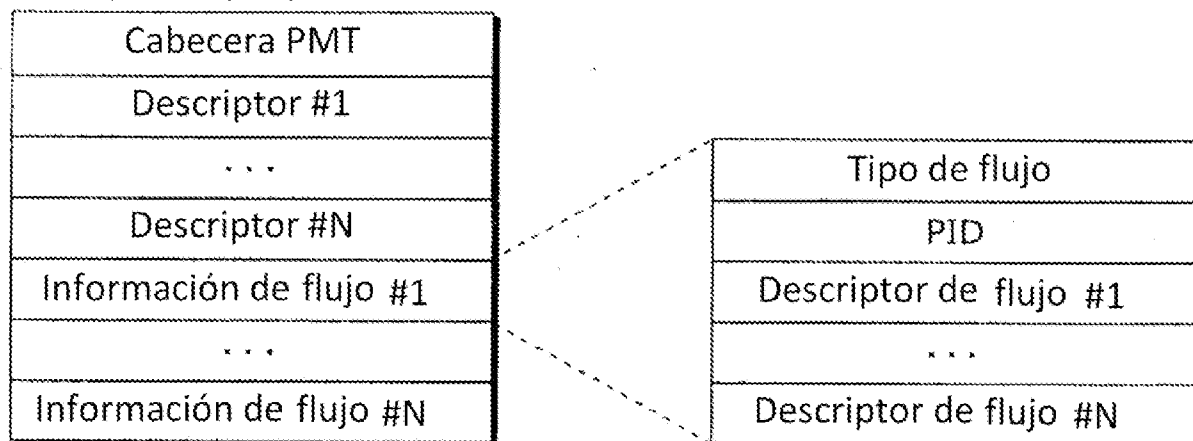


FIG. 39

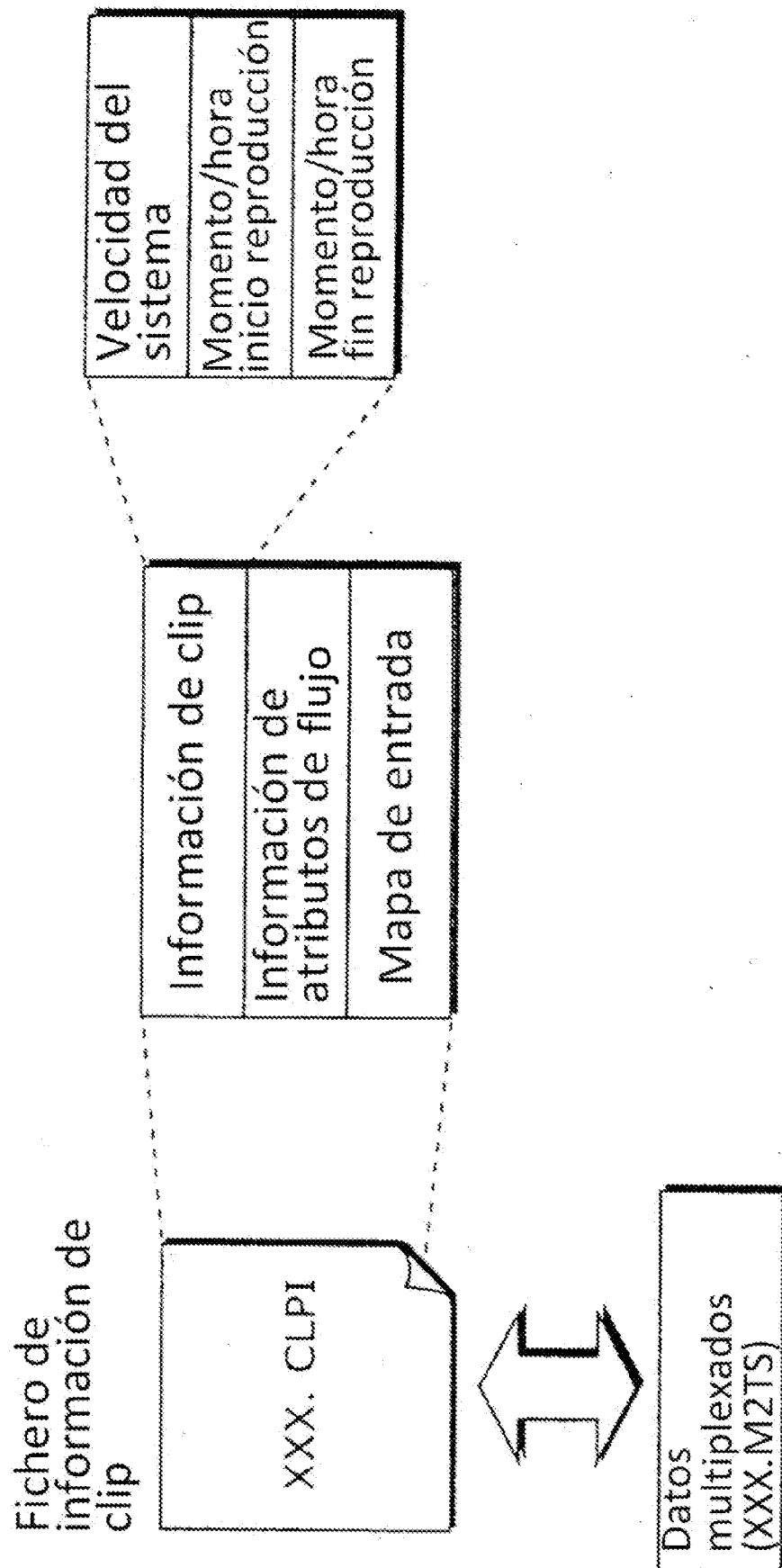


FIG. 40

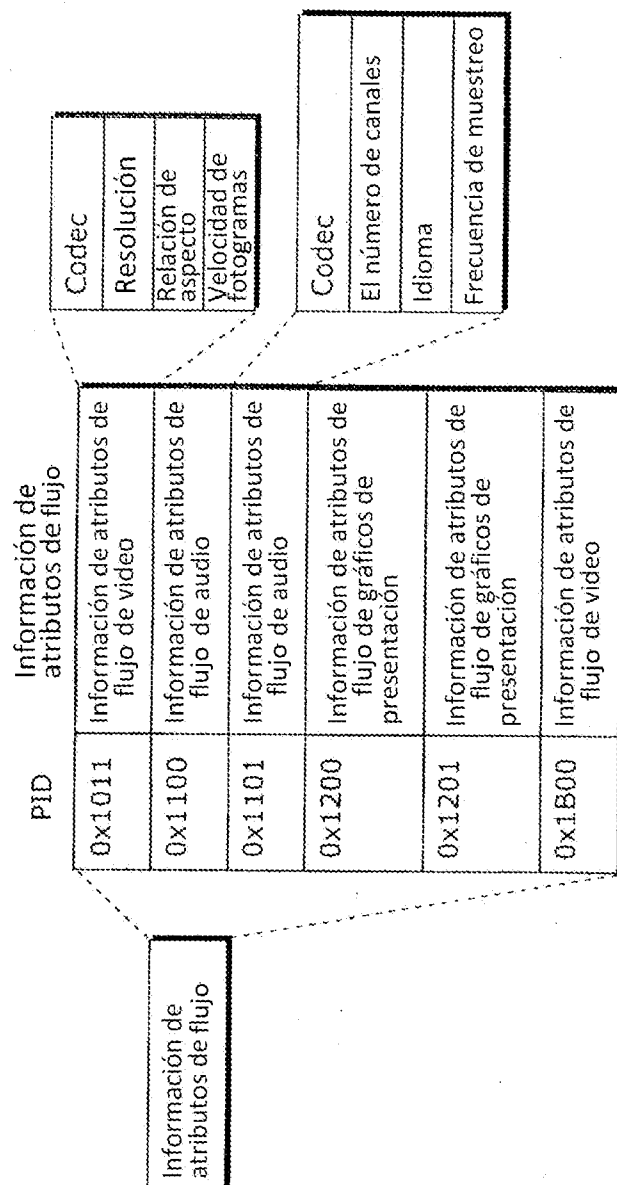


FIG. 41

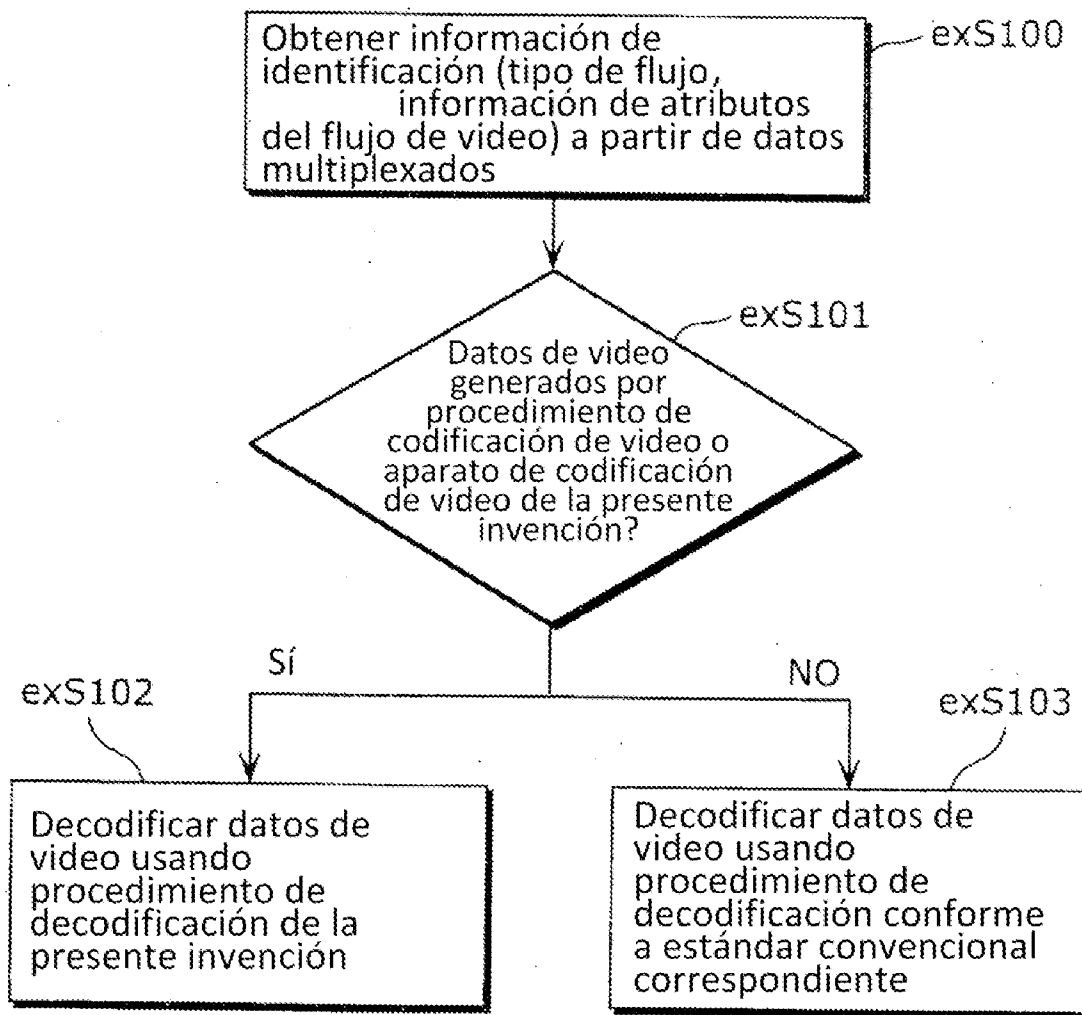


FIG. 42

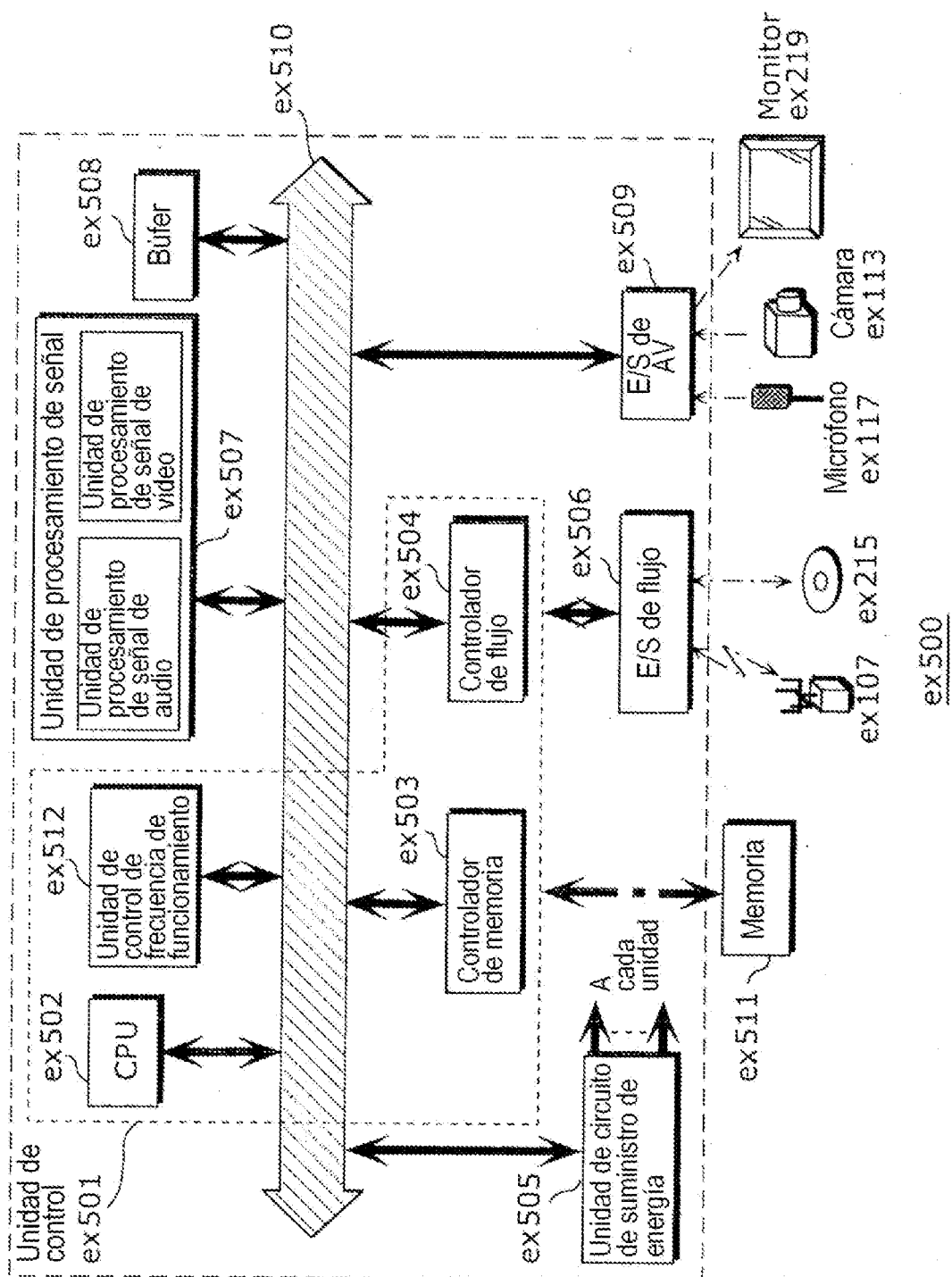


FIG. 43

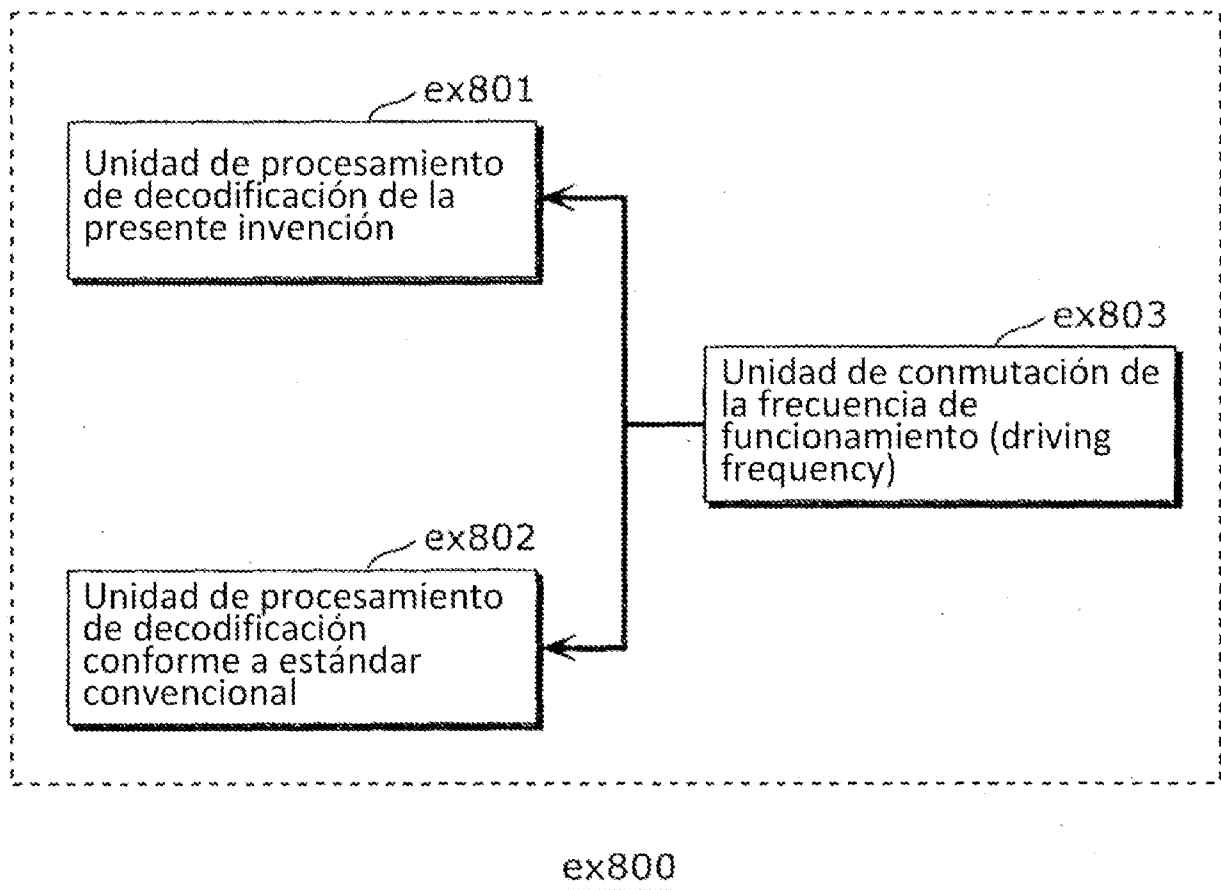


FIG. 44

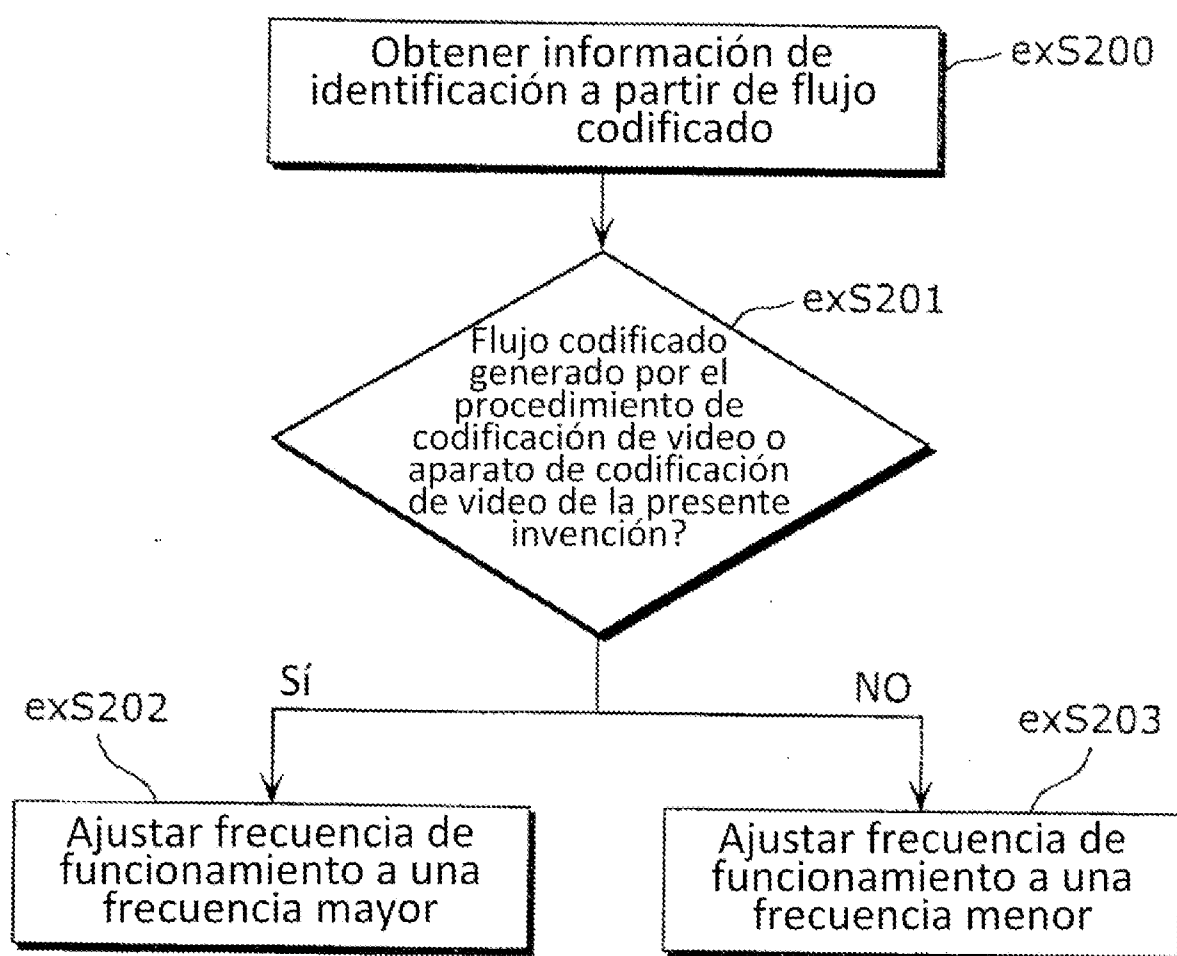


FIG. 45

Estándar correspondiente	Frecuencia de funcionamiento
MPEG-4 AVC	500 MHz
MPEG-2	350 MHz
⋮	⋮

FIG. 46A

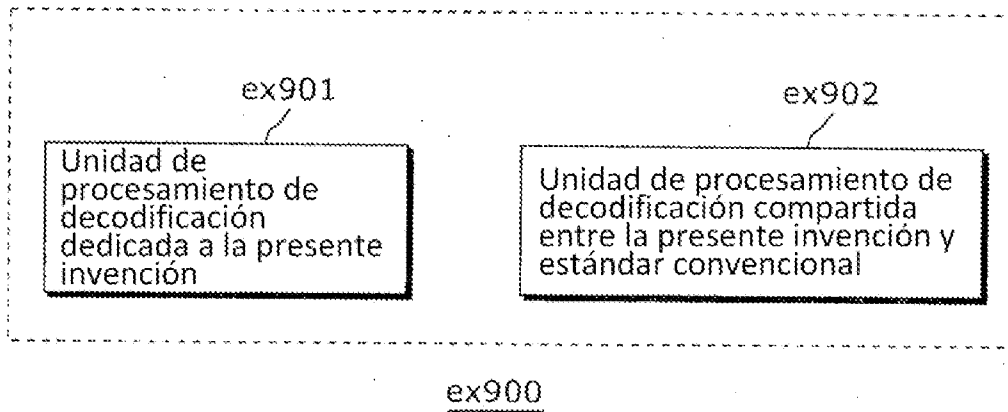


FIG. 46B

