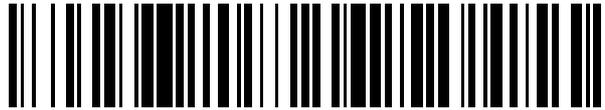


19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 535 314**

51 Int. Cl.:

H04N 19/597 (2014.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

96 Fecha de presentación y número de la solicitud europea: **26.12.2007 E 07860170 (5)**

97 Fecha y número de publicación de la concesión europea: **21.01.2015 EP 2099226**

54 Título: **Método de codificación de vídeo, método de decodificación, dispositivo de los mismos, programa de los mismos, y medio de almacenamiento que contiene el programa**

30 Prioridad:

28.12.2006 JP 2006353628

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

08.05.2015

73 Titular/es:

**NIPPON TELEGRAPH AND TELEPHONE CORPORATION (100.0%)
3-1, OTEMACHI 2-CHOME CHIYODA-KU
TOKYO 100-8116, JP**

72 Inventor/es:

**SHIMIZU, SHINYA;
KITAHARA, MASAKI;
KAMIKURA, KAZUTO y
YASHIMA, YOSHIYUKI**

74 Agente/Representante:

LINAGE GONZÁLEZ, Rafael

ES 2 535 314 T3

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

DESCRIPCIÓN

Método de codificación de vídeo, método de decodificación, dispositivo de los mismos, programa de los mismos, y medio de almacenamiento que contiene el programa.

5 **Campo técnico**

La presente invención se refiere a técnicas de codificación y decodificación para imágenes de vídeo de puntos de vista múltiples.

10 Se reivindica la prioridad en la Solicitud de Patente Japonesa núm. 2006-353628, presentada el 28 de diciembre de 2006, cuyo contenido se incorpora en este documento a modo de referencia.

15 **Técnica anterior**

Las imágenes de vídeo de puntos de vista múltiples son una pluralidad de imágenes de vídeo obtenidas fotografiando el mismo objeto y el fondo del mismo usando una pluralidad de cámaras. A continuación, una imagen de vídeo obtenida por una única cámara es llamada "imagen de vídeo bidimensional", y un conjunto de múltiples imágenes de vídeo bidimensionales obtenidas fotografiando el mismo objeto y el fondo del mismo es llamado "imagen de vídeo de puntos de vista múltiples".

25 Hay una fuerte correlación temporal en la imagen de vídeo bidimensional de cada cámara, que se incluye en una imagen de vídeo de puntos de vista múltiples. Además, cuando las cámaras se sincronizan las unas con las otras, las imágenes (tomadas por las cámaras) al mismo tiempo capturan el objeto y el fondo del mismo en el mismo estado exacto desde diferentes posiciones, de manera que haya una fuerte correlación entre las cámaras. La eficiencia de codificación de la codificación de vídeo se puede mejorar usando esta correlación.

30 En primer lugar, se mostrarán técnicas convencionales que se refieren a la codificación de imágenes de vídeo bidimensionales.

En muchos métodos conocidos de codificar imágenes de vídeo bidimensionales, como H. 264, MPEG-2, MPEG-4 (que son estándares de codificación internacionales), y similares, se lleva a cabo una codificación altamente eficiente por medio de compensación del movimiento, transformación ortogonal, cuantificación, codificación de entropía, o similares. Por ejemplo, en H.264, es posible llevar a cabo la codificación usando la correlación temporal entre el cuadro presente y los cuadros pasados o futuros.

35 El documento no patente 1 da a conocer técnicas detalladas de compensación del movimiento usadas en H. 264. Explicaciones generales de las mismas se describen a continuación.

40 De acuerdo con la compensación del movimiento en H. 264, un cuadro objetivo de codificación se divide en bloques de cualquier tamaño, y cada bloque puede tener un vector de movimiento individual, consiguiéndose de ese modo un alto nivel de eficiencia de codificación incluso para un cambio local en una imagen de vídeo.

45 Además, como candidatos para una imagen de referencia, los cuadros pasados o futuros (con respecto al cuadro presente), que ya se han codificado, se pueden preparar de manera que cada bloque pueda tener un cuadro de referencia individual, implementándose de ese modo un alto nivel de eficiencia de codificación incluso para una imagen de vídeo en la que se produzca una oclusión debido a un cambio temporal.

50 A continuación, se explicará un método de codificación convencional de imágenes de vídeo de puntos de vista múltiples. Ya que la codificación de las imágenes de vídeo de puntos de vista múltiples usa una correlación entre cámaras, las imágenes de vídeo de puntos de vista múltiples son codificadas de forma altamente eficiente en un método conocido que usa la "compensación de la disparidad" en la que se aplica la compensación del movimiento a las imágenes obtenidas por diferentes cámaras al mismo tiempo. Aquí, la disparidad es la diferencia entre las posiciones, a las que es proyectado el mismo punto de un objeto procesado gráficamente, en los planos de imagen de cámaras que se disponen en diferentes posiciones.

55 La figura 7 es una vista esquemática que muestra el concepto de disparidad generado entre tales cámaras. Es decir, la figura 7 muestra un estado en el que un observador mira hacia abajo a los planos de imagen de las cámaras A y B, cuyos ejes ópticos son paralelos el uno al otro, desde el lado superior de los mismos. Generalmente, tales puntos, a los que es proyectado el mismo punto en un objeto procesado gráficamente, en los planos de imagen de diferentes cámaras, son llamados "puntos correspondientes". En la codificación basada en la compensación de la disparidad, basada en la relación correspondiente anterior, cada valor de píxel de un cuadro objetivo de codificación se predice usando un cuadro de referencia, y se codifica el residuo de predicción pertinente y la información de la disparidad que indica la relación correspondiente.

65 Para cada cuadro en una imagen de vídeo de puntos de vista múltiples, la redundancia temporal y la redundancia

entre cámaras están presentes al mismo tiempo. El documento no patente 2 y el documento patente 1 (que da a conocer un aparato de codificación de imágenes de puntos de vista múltiples) dan a conocer cada uno un método para eliminar ambas redundancias simultáneamente.

5 En los métodos pertinentes, se lleva a cabo la predicción temporal de una imagen diferencial entre una imagen original y una imagen con la disparidad compensada, y se codifica un residuo de la compensación del movimiento en la imagen diferencial.

10 De acuerdo con tales métodos, la redundancia temporal, que no se puede eliminar mediante una compensación de la disparidad para eliminar la redundancia entre cámaras, se puede eliminar usando la compensación del movimiento. Por lo tanto, se reduce un residuo de predicción, que se codifica finalmente, de manera que se consiga un alto nivel de eficiencia de codificación.

15 Documento no patente 1: ITU-T Rec.H.264/ISO/IEC 11496-10, "Editor's Proposed Draft Text Modifications for Joint Video Specification (ITU-T Rec. H. 264 / ISO/IEC 14496-10 AVC), Draft 7", Borrador Final del Comité, Documento JVT-E022, págs. 10-13, y 62-68, septiembre de 2002.

20 Documento no patente 2: Shinya SHIMIZU, Masaki KITAHARA, Kazuto KAMIKURA y Yoshiyuki YASHIMA, "Multi-view Video Coding based on 3-D Warping with Depth Map", En Las Actas del Simposio de Codificación de Imágenes 2006, SS3-6, abril de 2006.

Documento patente 1: solicitud de patente pendiente de examen japonesa, primera publicación núm. H10-191393.

25 **Divulgación de la invención**

Problema que resolver por la invención

30 De acuerdo con el método de codificación de vídeo de puntos de vista múltiples convencional que aplica la compensación del movimiento a una imagen diferencial, se puede reducir el residuo de predicción en una parte que tenga las redundancias tanto temporal como entre cámaras, de manera que la codificación se pueda llevar a cabo de forma eficiente.

35 Sin embargo, si se produce una oclusión entre objetos procesados gráficamente como se muestra en la figura 8, entonces la presencia/ausencia de la redundancia entre cámaras para un único objeto puede cambiar dependiendo del tiempo.

La figura 8 muestra un ejemplo de la oclusión entre objetos en imágenes que son obtenidas por las cámaras A y B en cada uno de los tiempos a y b.

40 La figura 9 muestra un ejemplo de la imagen residual de la compensación de la disparidad cuando hay una oclusión.

45 De acuerdo con el método dado a conocer en el documento no patente 2, para cada píxel de un cuadro de referencia usado en la compensación de la disparidad, se proporciona información de la disparidad que indica un punto correspondiente en otro cuadro. Por lo tanto, si la imagen de la cámara B tiene la disparidad compensada a partir de la imagen de la cámara A (véase la figura 8), se obtiene una imagen diferencial como se muestra en la figura 9.

50 En la figura 9, la profundidad de color indica el valor de la señal diferencial, de manera que cuanto más próximo se encuentre del blanco, menor será la diferencia pertinente.

55 Aquí, se usa la información de la disparidad, que indica a qué parte en la imagen de la cámara B corresponde cada píxel en el cuadro de referencia (es decir, la imagen de la cámara A). Por lo tanto, no se lleva a cabo una compensación de la disparidad en cada parte que esté presente en la imagen de la cámara B pero no esté presente en la imagen de la cámara A. Según esto, el color de cada parte (en la imagen pertinente) para la que se puede obtener un punto correspondiente pasa a ser próximo al blanco. En contraste, en la parte R (en el interior de un círculo punteado en la figura 9) para la que no se puede obtener un punto correspondiente, la imagen original permanece sin cambios como la señal diferencial.

60 Sin embargo, debido a una diferencia en la sensibilidad entre las cámaras o una influencia de reflexión, la señal diferencial no pasa a ser completamente cero ni en cada parte para la que se puede obtener un punto correspondiente. En un área como R en la figura 9, es obvio que el residuo de predicción no se puede reducir ni mediante la compensación del movimiento llevada a cabo a partir de una imagen diferencial en otro tiempo como el tiempo a en la figura 9.

65 En un método simple para resolver el problema anterior, cada bloque se remite a un cuadro diferente. Es decir, como el cuadro de referencia, se puede usar no sólo la imagen decodificada de la imagen diferencial, sino también la

imagen decodificada de una imagen de cámara final obtenida añadiendo una imagen con la disparidad compensada a la imagen diferencial, y se puede remitir de forma conmutable a una de las dos imágenes decodificadas en cada bloque.

5 Según esto, incluso en el área R en la figura 9 donde permanece la imagen de cámara original, el residuo de predicción se puede reducir llevando a cabo la compensación del movimiento a partir de la imagen decodificada de la imagen de cámara.

10 Sin embargo, en tal método, es necesario codificar la información para indicar el cuadro de referencia de cada bloque, que causa un aumento de la cantidad de código.

Además, el método anterior es ineficaz si una parte que tiene una oclusión y una parte que no tiene una oclusión están presentes simultáneamente en un bloque.

15 Por otro lado, en el método dado a conocer en el documento no patente 1, cuando se codifica un cuadro objetivo de codificación, en lugar de codificar todo el cuadro sometiendo la imagen diferencial a la compensación del movimiento, es posible seleccionar para cada bloque si (i) la imagen diferencial se somete a la compensación del movimiento, (ii) sólo se lleva a cabo la compensación de la disparidad, o (iii) sólo se lleva a cabo la compensación del movimiento.

20 Según esto, incluso en una parte en la que se produce una oclusión entre las cámaras, el residuo de predicción se puede reducir de forma eficiente si hay redundancia temporal.

25 Sin embargo, en el método anterior, es necesario codificar la información para indicar qué método de predicción se usó en cada bloque, es decir, codificar una gran cantidad de información adicional aunque el residuo de predicción se pueda reducir. Por lo tanto, es imposible conseguir un alto nivel de eficiencia de codificación.

30 Asimismo, ya que sólo se puede seleccionar un método de predicción en cada bloque, el método pertinente es ineficaz si la presencia/ausencia de oclusión no es fija en cada parte de un bloque.

En vista de las circunstancias anteriores, un objeto de la presente invención es proporcionar técnicas de codificación y decodificación de vídeo novedosas para conseguir un alto nivel de eficiencia de codificación implementando una predicción apropiada sin aumentar la cantidad de código requerida para la predicción.

35 Medios para resolver el problema

Con el fin de resolver los problemas anteriores, la presente invención proporciona un método de codificación de vídeo para generar, en base a la información de la disparidad entre una imagen de cámara de referencia ya codificada y una imagen de cámara objetivo de codificación correspondiente a la imagen de cámara de referencia,
40 una imagen con la disparidad compensada por medio de la predicción de imágenes entre cámaras; y codificar una imagen diferencial entre la imagen de cámara objetivo de codificación y la imagen con la disparidad compensada. El método tiene una etapa de seleccionar, para cada unidad de sección predeterminada (por ejemplo, cada píxel) en la imagen diferencial, uno de los siguientes dos grupos de imágenes como un objetivo de referencia, determinando si hay o no una imagen con la disparidad compensada en la posición correspondiente, en otras palabras, si el píxel correspondiente en la imagen con la disparidad compensada tiene o no un valor eficaz:
45

(i) un grupo de imágenes diferenciales decodificadas, que se obtiene decodificando una imagen diferencial entre la imagen de cámara ya codificada y la imagen con la disparidad compensada (es decir, un conjunto de imágenes diferenciales decodificadas que incluyen una imagen diferencial decodificada que se obtuvo en otro tiempo y ya se ha codificado, o una imagen diferencial decodificada de una parte ya codificada de la imagen diferencial para la imagen de cámara objetivo de decodificación); y
50

(ii) un grupo de imágenes de cámara decodificadas, que se obtiene decodificando la imagen de cámara ya codificada (un conjunto de imágenes de cámara decodificadas que se representan cada una por la suma de cada imagen diferencial decodificada (que puede ser una imagen diferencial decodificada en otro tiempo) y una imagen con la disparidad compensada correspondiente).
55

De acuerdo con el método anterior, la codificación predictiva que usa una imagen de referencia apropiada para cada unidad de sección deseada se puede llevar a cabo sin añadir nueva información adicional acerca de la selección para el modo o imagen de referencia. Es decir, es posible reducir el elemento residual que codificar, sin aumentar la cantidad de código requerida para la información adicional, implementándose de ese modo un alto nivel de eficiencia de codificación.
60

En el método anterior, la información de la disparidad se puede proporcionar como un vector de disparidad, o representar como información geométrica como una distancia desde la cámara hasta el objeto procesado gráficamente, es decir, se puede proporcionar en cualquier formato mediante el cual se puedan obtener las
65

relaciones correspondientes entre las imágenes pertinentes.

Cuando se codifica la imagen diferencial usando un objetivo de referencia designado, la codificación se puede llevar a cabo usando un método de predicción de señal en un bloque de procesamiento de codificación.

5 En la codificación predictiva general, diferentes objetivos de referencia probablemente tienen diferentes métodos de predicción óptimos. Por lo tanto, si un bloque de procesamiento de codificación incluye píxeles o áreas que tienen diferentes objetivos de referencia, el método de predicción se selecciona para cada objetivo de referencia, y se codifica información adicional que indica una pluralidad de métodos de predicción.

10 Sin embargo, si la codificación se lleva a cabo usando un único método de predicción en un bloque de procesamiento de codificación, como se describe anteriormente, la información que sólo indica un método de predicción se debería codificar para cada bloque de procesamiento de codificación, reduciéndose de ese modo la cantidad de código requerida para toda la codificación. En este caso, para una degradación anticipada de la eficiencia de predicción, la eficiencia de predicción no se debería degradar de forma considerable en consideración a un caso como se muestra en la figura 9, es decir, cuando se lleva a cabo la predicción de vídeo aplicando un único método (es decir, la compensación del movimiento usando un único vector de movimiento) incluso para una parte limítrofe.

15 Adicionalmente, en consideración al equilibrio entre la eficiencia de predicción y la cantidad de código requerida para la información adicional, es posible seleccionar si se usa un único método de predicción en cada bloque, o se asigna un método de predicción individual a cada objetivo de referencia. Cuando se considera tal equilibrio entre la eficiencia de predicción y la cantidad de código requerida para la información adicional, es posible implementar más codificación flexible que tenga un mayor nivel de eficiencia de codificación.

20 En este caso, el lado de decodificación puede determinar a partir de datos codificados si se ha conmutado o no el método de predicción, de manera que el lado de decodificación pueda decodificar de forma apropiada la imagen pertinente.

25 **Efecto de la invención**

De acuerdo con la presente invención, cuando se somete una imagen diferencial obtenida mediante la compensación de la disparidad a la codificación predictiva, no se usa ninguna información adicional, y se usa una imagen de referencia apropiada mediante una operación de conmutación para cada píxel. Por lo tanto, el residuo para el objetivo de codificación se puede reducir sin aumentar la cantidad de código requerida para la información adicional, implementándose de ese modo una codificación de vídeo eficiente de toda la imagen de vídeo de puntos de vista múltiples.

30 **Breve descripción de los dibujos**

La figura 1 es un diagrama que muestra un aparato de codificación de vídeo como una forma de realización de la presente invención.

La figura 2 es un diagrama de flujo de codificación de vídeo llevado a cabo en la forma de realización.

La figura 3 es un diagrama de flujo detallado para la determinación del modo de predicción en la forma de realización.

La figura 4 es un diagrama que muestra un aparato de decodificación de vídeo como una forma de realización de la presente invención.

La figura 5 es un diagrama de flujo de decodificación de vídeo llevado a cabo en la forma de realización.

La figura 6 es un diagrama de flujo detallado de la decodificación de datos codificados para la imagen diferencial en la forma de realización.

La figura 7 es un diagrama esquemático que muestra el concepto de disparidad generado entre cámaras.

La figura 8 es un diagrama que muestra un ejemplo en el que se produce una oclusión entre objetos procesados gráficamente.

La figura 9 es un diagrama que muestra un ejemplo de la imagen residual de la compensación de la disparidad cuando hay una oclusión.

65 **Símbolos de referencia**

- 100 aparato de codificación de vídeo
- 101 unidad de entrada de imagen
- 5 102 unidad de entrada de imagen de referencia
- 103 unidad de entrada de información de la disparidad
- 104 generador de imagen con la disparidad compensada
- 10 105 codificador predictivo de imagen diferencial
- 106 unidad de establecimiento de imagen de referencia
- 15 107 memoria de imágenes de cámara decodificadas
- 108 memoria de imágenes diferenciales decodificadas
- 109 conmutador de referencia
- 20 110 decodificador de imagen diferencial
- 200 aparato de decodificación de vídeo
- 25 201 unidad de entrada de datos codificados
- 202 unidad de entrada de imagen de referencia
- 203 unidad de entrada de información de la disparidad
- 30 204 generador de imagen con la disparidad compensada
- 205 unidad de establecimiento de imagen de referencia
- 35 206 memoria de imágenes de cámara decodificadas
- 207 memoria de imágenes diferenciales decodificadas
- 208 conmutador de referencia
- 40 209 decodificador de imagen diferencial

Mejor modo para llevar a cabo la invención

- 45 Antes de la explicación de las formas de realización de la presente invención, se describirá un resumen de la presente invención.

En las técnicas convencionales para la codificación de puntos de vista múltiples, es necesario codificar y transmitir información que indique la imagen seleccionada. En contraste, la presente invención está caracterizada porque la información que indica la imagen seleccionada se puede codificar usando otra información. Según esto, cuando la presente invención tiene la misma cantidad de código que las técnicas convencionales, la presente invención tiene una mejor calidad de imagen en comparación a las técnicas convencionales.
- 50
- 55 Con el fin de codificar una imagen de vídeo de puntos de vista múltiples, si se obtiene una diferencia entre una imagen de vídeo obtenida por una cámara y una imagen correspondiente generada por medio de la compensación de la disparidad desde otra cámara, y se lleva a cabo la codificación en consideración a la correlación temporal con respecto a la imagen de vídeo diferencial, entonces en las técnicas convencionales, una parte de oclusión como R en la figura 9 no se puede codificar de forma eficiente.
- 60 En contraste, en la presente invención, una parte de oclusión como R en la figura 9 se puede codificar de forma eficiente sometiendo la parte pertinente a la codificación predictiva la cual usa, no una imagen de vídeo diferencial a partir de una imagen generada mediante la compensación de la disparidad, sino una imagen decodificada de la imagen de vídeo original.
- 65 Adicionalmente, en un método fácilmente previsto para resolver un problema tal que la parte de oclusión no se puede codificar de forma eficiente, se usa una técnica de remisión a cuadros múltiples empleada en H.264 de

manera que la codificación predictiva se lleve a cabo asignando una imagen de referencia individual a cada bloque.

Sin embargo, en tal caso, (i) se codifica información para designar una imagen de referencia para cada bloque, que causa un aumento de la cantidad de código, y (ii) la conmutación de la imagen de referencia sólo se puede llevar a cabo por una unidad de bloque. Por lo tanto, el método pertinente es ineficaz cuando un bloque tiene tanto una parte de oclusión como una parte de no oclusión.

En contraste, en la presente invención, la conmutación de la imagen de referencia se lleva a cabo usando información obtenida cuando se genera una imagen mediante la compensación de la disparidad. Por lo tanto, es innecesario codificar nueva información. Asimismo, de acuerdo con la información usada, se puede determinar la presencia/ausencia de oclusión para cada píxel. Por lo tanto, la conmutación de la imagen de referencia se puede llevar a cabo por una unidad de píxel.

Con el fin de implementar las funciones anteriores, en la presente invención, se genera la información Ref que indica si una imagen con la disparidad compensada Syn de un objetivo de codificación se puede generar o no usando una imagen obtenida por otra cámara, y la conmutación de la imagen de referencia se lleva a cabo usando la información Ref.

Un resultado de la determinación de que se puede generar una imagen con la disparidad compensada Syn significa que el mismo objeto aparece en una imagen de otra cámara, y de ese modo no se ha producido una oclusión. En contraste, un resultado de la determinación de que no se puede generar una imagen con la disparidad compensada Syn significa que el mismo objeto no aparece en una imagen de otra cámara, y de ese modo se ha producido una oclusión. Es decir, la conmutación de la imagen de referencia se lleva a cabo usando información que indica si se puede generar o no una imagen con la disparidad compensada Syn, de manera que sea posible resolver el problema de que no se puede llevar a cabo una predicción apropiada en una parte de oclusión como R en la figura 9.

La técnica de remisión a cuadros múltiples usada en H.264 también se conoce como un método para seleccionar una imagen de referencia usada para codificar de entre una pluralidad de candidatas a imagen de referencia. En este método conocido, es necesario codificar y transmitir información que indique qué imagen se seleccionó.

En contraste, en la presente invención, la información usada para seleccionar una imagen de referencia se genera usando información que se usa para generar una imagen con la disparidad compensada y que también se debe transmitir cuando se usa el método conocido anterior. Por lo tanto, es innecesario transmitir información adicional usada para conmutar la imagen de referencia, reduciéndose de ese modo la cantidad de código.

En las formas de realización mostradas más adelante, la información que se usa para seleccionar una imagen de referencia y se genera en base a la información usada para generar una imagen con la disparidad compensada se representa como la información Ref, y la misma información se puede usar en los lados de codificación y de decodificación (por ejemplo, la información en el lado de codificación se genera en la etapa A6 en la figura 2, y la información en el lado de decodificación se genera en la etapa C4 en la figura 5).

Adicionalmente, si la imagen de referencia se selecciona para cada píxel en un método conocido como la técnica de remisión a cuadros múltiples en H.264 o similares, se puede reducir la cantidad de código requerida para una señal de error de predicción que codificar. Sin embargo, en tal caso, es necesario codificar la información usada para conmutar la imagen de referencia para cada píxel, lo cual causa un aumento de la cantidad de código. Es decir, en este caso, la imagen de referencia sólo se puede conmutar por una unidad de bloque de modo que se reduzca la cantidad total de código.

En contraste, en la presente invención, la información para conmutar la imagen de referencia no es una que transmitir adicionalmente como se describe anteriormente. Por lo tanto, cuando se genera tal información en correspondencia a cada píxel, la conmutación de la imagen de referencia para cada píxel se puede llevar a cabo sin aumentar la cantidad de código, reduciéndose de ese modo la cantidad de código requerida para la señal de error de predicción.

A continuación, la presente invención se explicará en detalle de acuerdo con las formas de realización.

En las formas de realización, se asume que se codifica una imagen de vídeo de puntos de vista múltiples obtenida usando dos cámaras, donde la imagen de la cámara B se codifica usando la imagen de la cámara A como la imagen de referencia.

La figura 1 muestra un aparato de codificación de vídeo 100 como una forma de realización de la presente invención.

El aparato de codificación de vídeo 100 incluye una unidad de entrada de imagen 101 para introducir una imagen (como una imagen objetivo de codificación) obtenida por la cámara B en el aparato; una unidad de entrada de

imagen de referencia 102 para introducir una imagen decodificada (como una imagen de referencia) de la cámara A en el aparato; una unidad de entrada de información de la disparidad 103 para introducir información de la disparidad, que indica a qué posición en la imagen objetivo de codificación corresponde cada píxel en la imagen de referencia, en el aparato; un generador de imagen con la disparidad compensada 104 para generar una imagen con la disparidad compensada en la posición de la cámara B usando la imagen de referencia y la información de la disparidad; un codificador predictivo de imagen diferencial 105 para someter una imagen diferencial entre la imagen objetivo de codificación y la imagen con la disparidad compensada a la codificación predictiva; una unidad de establecimiento de imagen de referencia 106 para asignar un grupo de imágenes de referencia a cada píxel en base al estado de la imagen con la disparidad compensada; una memoria de imágenes de cámara decodificadas 107 para almacenar imágenes decodificadas (como un grupo de imágenes de referencia) de imágenes obtenidas por la cámara B; una memoria de imágenes diferenciales decodificadas 108 para almacenar imágenes decodificadas (como otro grupo de imágenes de referencia) de las imágenes diferenciales codificadas; un conmutador de referencia 109 para conmutar la memoria de imágenes de referencia en base a la información del establecimiento acerca del grupo de imágenes de referencia; y un decodificador de imagen diferencial 110 para decodificar cada imagen diferencial codificada.

Las figuras 2 y 3 son diagramas de flujo ejecutados por el aparato de codificación de vídeo 100 que tiene la estructura anterior. Con referencia a los diagramas de flujo, se explicará en detalle la operación llevada a cabo por el aparato de codificación de vídeo 100.

Como se muestra en el flujo de operación en la figura 2 que muestra la operación de codificación general, una imagen de la cámara B se introduce por la unidad de entrada de imagen 101 en el aparato de codificación de vídeo 100 (véase la etapa A1).

Aquí, una imagen decodificada de la cámara A, que tiene el mismo tiempo de visualización (es decir, el mismo tiempo de procesamiento gráfico) que la imagen introducida de la cámara B, es introducida por la unidad de entrada de imagen de referencia 102, y la información de la disparidad entre las imágenes en el tiempo pertinente es introducida por la unidad de entrada de información de la disparidad 103.

A continuación, la imagen introducida de la cámara B es llamada "imagen de cámara objetivo de codificación", y la imagen decodificada de la cámara A es llamada "imagen de cámara de referencia".

En base a la imagen de cámara de referencia introducida por la unidad de entrada de imagen de referencia 102 y la información de la disparidad introducida por la unidad de entrada de información de la disparidad 103, una imagen con la disparidad compensada es generada por el generador de imagen con la disparidad compensada 104 (véase la etapa A2).

La generación de la imagen con la disparidad compensada se lleva a cabo de una manera en la que el valor de píxel de cada píxel en la imagen de cámara de referencia se usa como el valor de píxel de un píxel correspondiente en la imagen con la disparidad compensada.

En este proceso, si una pluralidad de puntos en la imagen de cámara de referencia corresponde al mismo punto en la imagen con la disparidad compensada, entonces se determinan las relaciones delante/detrás entre los objetos procesados gráficamente, y se usa el valor de píxel de un punto en un objeto procesado gráficamente más próximo a la cámara. Por ejemplo, si las cámaras tienen ejes ópticos paralelos, es posible determinar que cuanto mayor sea la disparidad, más próximo se encontrará el punto pertinente a la cámara.

Además, ya que la información de la disparidad indica un punto correspondiente en la imagen de cámara objetivo de codificación para cada píxel en la imagen de cámara de referencia, no hay una imagen con la disparidad compensada para una parte que está ausente en la imagen de cámara de referencia.

Usando la imagen con la disparidad compensada obtenida anteriormente, la imagen de cámara objetivo de codificación se codifica para cada bloque de procesamiento de codificación (véanse las etapas A3 a A13).

En el presente flujo, "blk" es un índice para cada bloque, y maxBlk indica el número total de bloques en una imagen. Es decir, después de que blk se inicialice a cero (véase la etapa A3), el siguiente proceso (etapas A4 a A11) se lleva a cabo repetidamente mientras blk se incrementa en 1 (véase la etapa A12), hasta que blk alcance el número total maxBlk de bloques (véase la etapa A13).

En primer lugar, se define para el bloque blk que Org indica la imagen de cámara objetivo de codificación y Syn indica la imagen con la disparidad compensada (véase la etapa A4). La diferencia Org-Syn se computa para cada píxel que pertenece al bloque blk, de manera que se genere una imagen diferencial Sub (véase la etapa A5).

Para cada parte de píxel que no tiene una imagen con la disparidad compensada, el valor de píxel de Syn se establece en 0 cuando se computa la imagen diferencial.

A continuación, la unidad de establecimiento de imagen de referencia 106 establece la información Ref, que indica una referencia usada cuando se somete la imagen diferencial de cada píxel en el bloque blk a la codificación predictiva como se muestra a continuación (véase la etapa A6).

5 $\forall \text{pix} \in \{0, 1, \dots, \text{maxPix}-1\}$

Ref[pix]=0 cuando no se almacena una imagen con la disparidad compensada en Syn[pix]

10 Ref[pix]=1 cuando se almacena una imagen con la disparidad compensada en Syn[pix]

donde pix es un índice que indica cada píxel en el bloque blk, maxPix indica el número total de píxeles en el bloque blk. Además, la información de la unidad de bloque a la que "[pix]" se anexa indica la información del píxel en la posición del índice de píxel pix.

15 Adicionalmente, aunque se puede usar cualquier método para determinar si una imagen con la disparidad compensada se establece o no como Syn de modo que se genere Ref, un ejemplo de lo mismo se mostrará a continuación.

20 Antes de generar la imagen con la disparidad compensada Syn, el valor de píxel de cada píxel se inicializa en un valor absolutamente inalcanzable (por ejemplo, -1). La imagen con la disparidad compensada Syn se genera actualizando el valor de píxel de una manera en la que el valor de píxel de cada píxel en la imagen de cámara de referencia se determina como el valor de píxel en la posición correspondiente en la imagen con la disparidad compensada. Ref se genera entonces de una manera en la que si la Syn asignada a cada píxel es -1, Ref se establece en 0, y en la que si Syn tiene un valor distinto a -1, Ref se establece en 1. Según esto, la Ref objetivo se puede generar en correspondencia unívoca a Syn.

En el método anterior, en la computación de la imagen diferencial Sub en la etapa descrita anteriormente A5, en cada parte en la que Syn es -1, el valor de píxel es considerado como 0 y Sub se establece en el valor de Org.

30 En otro método, antes de computar la imagen diferencial Sub, Ref se genera de acuerdo con el método descrito anteriormente o similares; en cada parte en la que Syn es -1, Syn se reescribe como 0; y entonces la imagen diferencial Sub se computa computando Org-Syn.

35 Aquí, para facilitar las explicaciones, la Syn total, que también incluye cada parte en la que no hay presente una imagen con la disparidad compensada y el valor de píxel es considerado como 0, es llamada imagen con la disparidad compensada.

40 A continuación, se obtiene un modo de predicción, que proporciona el coste de distorsión de velocidad mínimo cuando se somete la imagen diferencial Sub a la codificación predictiva usando la Ref anterior, y se establece como PMODE (véase la etapa A7).

45 Cuando "sad" es la suma de valores absolutos de los errores de predicción en un modo de predicción, y "bin" es un valor predicho de la cantidad de código requerida para codificar información que indica el modo de predicción, el coste de distorsión de velocidad llamado "cost" se computa mediante la siguiente fórmula, donde λ es un multiplicador de Lagrange indefinido y es un valor predeterminado.

$$\text{cost} = \text{sad} + \lambda \cdot \text{bin}$$

50 Usando el PMODE determinado anteriormente, la imagen diferencial Sub del bloque blk en este momento se somete a la codificación predictiva (véase la etapa A8).

La determinación del modo de predicción y la codificación actual se llevan a cabo en el codificador predictivo de imagen diferencial 105. El resultado de la codificación incluye, no sólo datos codificados de la imagen pertinente, sino también datos codificados de la información (por ejemplo, PMODE) requerida para la codificación.

55 El resultado de la codificación se emite desde el aparato de codificación de vídeo 100, y también se decodifica en el decodificador de imagen diferencial 110 (véase la etapa A9).

60 Aquí, Dec indica la imagen decodificada de la imagen diferencial Sub del bloque blk, y se almacena en la memoria de imágenes diferenciales decodificadas 108 de modo que se use Dec en la predicción dentro del cuadro llevada a cabo cuando se codifica otro bloque o la predicción entre cuadros llevada a cabo cuando se codifica un cuadro en otro tiempo (véase la etapa A10).

65 Además, la imagen de cámara decodificada del bloque blk se genera computando la suma de Dec y Syn, y se almacena en la memoria de imágenes de cámara decodificadas 107 (véase la etapa A11).

La operación descrita anteriormente se aplica a cada bloque.

5 La figura 3 muestra un flujo de operación detallado del proceso de determinación del modo de predicción (etapa A7) llevado a cabo en el codificador predictivo de imagen diferencial 105.

En el flujo de operación, un coste de distorsión de velocidad llamado "cost" indicado por la siguiente fórmula se computa para cada modo de predicción, y se determina un modo de predicción que proporciona el valor mínimo.

10 $cost = SAD + \lambda \cdot code(pmode)$

donde SAD es la suma de las diferencias absolutas para cada píxel entre la imagen predicha y la imagen original; pmode indica un índice de modo de predicción; y code(α) es una función que devuelve un valor predicho de la cantidad de código requerida para representar información acerca del modo de predicción proporcionado α .

15 Como se muestra en el flujo de operación, después de que el índice de modo de predicción pmode se inicialice a cero (véase la etapa B1), el siguiente proceso (etapas B2 a B16) se lleva a cabo repetidamente mientras pmode se incrementa en 1 (véase la etapa B17), hasta que pmode alcance el número "maxPmode" de modos de predicción (véase la etapa B18), determinándose de ese modo un modo de predicción que minimiza "cost".

20 En el siguiente proceso, cada modo de predicción se estima usando un coste de distorsión de velocidad, donde maxCost se define como el valor máximo que es inalcanzable como el valor estimado. Asimismo, con el fin de llevar a cabo una estimación repetitiva, los mejores valores estimados en las condiciones individuales explicadas a continuación están indicados respectivamente por minCost, minCost1, y minCost2, y los índices de modo de predicción asignados a los mismos están indicados respectivamente por best_mode, best_mode1, y best_mode2.

30 Las variables minCost y best_mode representan el mejor valor estimado y el índice de modo de predicción correspondiente cuando se predicen todos los píxeles en el bloque pertinente; las variables minCost1 y best_mode1 representan el mejor valor estimado y el índice de modo de predicción correspondiente cuando sólo se predicen píxeles cuya Ref correspondiente tiene un valor de 0; y las variables minCost2 y best_mode2 representan el mejor valor estimado y el índice de modo de predicción correspondiente cuando sólo se predicen píxeles cuya Ref correspondiente tiene un valor de 1.

35 Después de inicializar cada uno de minCost, minCost1, y minCost2 en maxCost (véase la etapa B1), una imagen predicha para el modo de predicción pmode se genera para cada píxel (véanse las etapas B2 a B8).

40 En la generación de la imagen predicha, después de que el índice de píxel "pix" en el bloque pertinente se inicialice a cero (véase la etapa B2), el siguiente proceso (etapas B3 a B6) se lleva a cabo repetidamente mientras pix se incrementa en 1 (véase la etapa B7), hasta que pix alcance el número "maxPix" de píxeles en el bloque (véase la etapa B8), determinándose de ese modo un modo de predicción que minimiza "cost".

En primer lugar, de acuerdo con Ref[pix], se determina si hay o no una imagen con la disparidad compensada en el píxel correspondiente (véase la etapa B3).

45 Si no hay presente una imagen con la disparidad compensada, el conmutador de referencia 109 se opera de manera que el codificador predictivo de imagen diferencial 105 se remita a la memoria de imágenes de cámara decodificadas 107 (véase la etapa B4). En contraste, si hay una imagen con la disparidad compensada, el conmutador de referencia 109 se opera de manera que el codificador predictivo de imagen diferencial 105 se remita a la memoria de imágenes diferenciales decodificadas 108 (véase la etapa B5).

50 Entonces usando la imagen de referencia establecida, un valor predicho Pred[pix] para el píxel pertinente se computa en el modo de predicción pmode (véase la etapa B6).

55 Después de que la generación de la imagen predicha se complete para todos los píxeles en el bloque, tres tipos de la suma de las diferencias absolutas entre la imagen predicha y la imagen original se computan mediante las siguientes fórmulas.

$$SAD1 = \sum |Sub[pix] - Pred[pix]| \cdot (1 - Ref[pix])$$

60 $SAD2 = \sum |Sub[pix] - Pred[pix]| \cdot Ref[pix]$

$$SAD = SAD1 + SAD2$$

65 donde SAD1 indica la suma de las diferencias absolutas para los píxeles que no tienen ninguna imagen con la disparidad compensada; SAD2 indica la suma de las diferencias absolutas para los píxeles que tienen cada uno una

- imagen con la disparidad compensada; y SAD indica la suma de las diferencias absolutas para todos los píxeles en el bloque. Además, E indica la computación de la suma total para $\text{pix}=0$ a $\text{maxPix}-1$. Aunque las fórmulas anteriores emplean operaciones multiplicadoras, la computación de la suma de las diferencias absolutas se puede implementar usando una ramificación condicional simple ya que los píxeles para computar SAD1 y SAD2 son exclusivos los unos para los otros.
- 5 Usando los SAD, SAD1, SAD2, y pmode anteriores, los costes de distorsión de velocidad llamados cost, cost1, y cost2 se computan de acuerdo con la fórmula descrita anteriormente (véase la etapa B10).
- 10 Los resultados computados se comparan respectivamente con los mejores costes de distorsión de velocidad que ya se han computado (véanse las etapas B11, B13, y B15). Para cada caso, si el modo de predicción pmode puede generar un menor coste, la variable que indica el mejor modo de predicción y el mejor valor de coste se actualizan (véanse las etapas B12, B 14, y B 16).
- 15 Después de que se complete la estimación de todos los modos de predicción, se determina si se asignan diferentes modos de predicción a cada píxel que tiene una imagen con la disparidad compensada y cada píxel que no tiene una compensación de la disparidad, o se asigna el mismo modo de predicción a todos los píxeles (véanse las etapas B19 a B21).
- 20 En primer lugar, se determina si best_mode, best_mode1, y best_mode2 son o no todos el mismo (véase la etapa B 19).
- Si todos los elementos anteriores son el mismo, esto indica que se puede usar un único modo en el bloque, y de ese modo best_mode se almacena como el modo de predicción PMODE usado en la codificación predictiva (véase la etapa B22). La operación se termina entonces.
- 25 Si al menos uno de los tres indica un modo diferente, entonces la suma de minCost1, minCost2, y OHCost se compara con minCost (véase la etapa B20).
- 30 Aquí, lo primero (suma) indica el coste requerido cuando se usan diferentes modos en el bloque pertinente, y OHCost representa unos costes generales en consideración a que se codifican dos modos de predicción.
- Si lo primero proporciona un mejor coste, el conjunto de best_mode1 y best_mode2 se almacena como PMODE (véase la etapa B21). La operación se termina entonces.
- 35 Si lo último proporciona un mejor coste, best_mode se almacena como PMODE (véase la etapa B22), y la operación se termina.
- 40 En la presente invención, se selecciona si se usa un modo de predicción o se usan dos modos de predicción en el bloque.
- Sin embargo, siempre se puede usar un modo de predicción. En tal caso, sólo se computa "cost" en la etapa B10; las etapas B13 a B16 y B19 a B22 se omiten; y PMODE se puede usar en lugar de best_mode
- 45 Por otro lado, siempre se pueden usar dos modos de predicción. En tal caso, la computación de SAD en la etapa B9, la computación de Cost en la etapa B 10, y las etapas B 11 a B 12, B 19 a B20, y B22 son innecesarias; y si el resultado de la determinación en la etapa B18 es SÍ, la operación pasa a la etapa B21.
- 50 El modo de predicción en la presente forma de realización indica cómo predecir el valor de píxel.
- Como el método de predicción, hay una predicción dentro del cuadro para predecir el valor de píxel usando un bloque periférico que pertenece al mismo cuadro que el presente bloque y ya se ha codificado y decodificado, y una predicción entre cuadros para predecir el valor de píxel usando otro cuadro que ya se ha codificado y decodificado, y los modos de predicción descritos anteriormente incluyen ambos métodos. Un vector de movimiento requerido en la predicción entre cuadros también funciona como un modo de predicción. Además, cualquier método que pueda predecir el valor de píxel usando datos que ya se han codificado y decodificado puede funcionar como un modo de predicción.
- 55 La figura 4 muestra un aparato de decodificación de vídeo 200 como una forma de realización de la presente invención.
- 60 El aparato de decodificación de vídeo 200 incluye una unidad de entrada de datos codificados 201 para introducir datos codificados en el aparato; una unidad de entrada de imagen de referencia 202 para introducir una imagen decodificada de la cámara A en el aparato; una unidad de entrada de información de la disparidad 203 para introducir información de la disparidad, que indica a qué posición en la imagen como el objetivo de decodificación corresponde cada píxel en la imagen decodificada de la cámara A, en el aparato; un generador de imagen con la
- 65

- disparidad compensada 204 para generar una imagen con la disparidad compensada en la posición de la cámara B; una unidad de establecimiento de imagen de referencia 205 para asignar un grupo de imágenes de referencia a cada píxel en base al estado de la imagen con la disparidad compensada; una memoria de imágenes de cámara decodificadas 206 para almacenar imágenes decodificadas (como un grupo de imágenes de referencia) de imágenes obtenidas por la cámara B; una memoria de imágenes diferenciales decodificadas 207 para almacenar imágenes decodificadas (como otro grupo de imágenes de referencia) de las imágenes diferenciales; un conmutador de referencia 208 para conmutar la memoria de imágenes de referencia en base a la información del establecimiento acerca del grupo de imágenes de referencia; y un decodificador de imagen diferencial 209 para decodificar los datos codificados introducidos.
- Las figuras 5 y 6 son diagramas de flujo ejecutados por el aparato de decodificación de vídeo 200 que tiene la estructura anterior, y muestran un flujo de operación para decodificar un cuadro de datos codificados de la cámara B. El flujo de operación se explicará en detalle a continuación.
- Aquí, se asume que un cuadro de la cámara A en el mismo tiempo que el cuadro de decodificación objetivo ya se ha decodificado, y la información de la disparidad pertinente ya se ha obtenido también.
- En primer lugar, los datos codificados de la cámara B se introducen en la unidad de entrada de datos codificados 201 (véase la etapa C1). Aquí, una imagen decodificada de la cámara A, que tiene el mismo tiempo de visualización que la imagen introducida de la cámara B, es introducida por la unidad de entrada de imagen de referencia 202.
- A continuación, se genera una imagen con la disparidad compensada en el generador de imagen con la disparidad compensada 204 en base a la imagen decodificada de la cámara A y la información de la disparidad (véase la etapa C2). El proceso en esta etapa es igual al llevado a cabo en la etapa explicada anteriormente A2 en la figura 2.
- Los datos codificados introducidos para cada bloque se decodifican usando la imagen con la disparidad compensada, de manera que se obtenga una imagen decodificada de la cámara B (véanse las etapas C3 a C9).
- En este proceso, "blk" es un índice para cada bloque, y maxBlk indica el número total de bloques en una imagen. Después de que blk se inicialice a cero (véase la etapa C3), el siguiente proceso (etapas C4 a C7) se lleva a cabo repetidamente mientras blk se incrementa en 1 (véase la etapa C8), hasta que blk alcance el número total maxBlk de bloques (véase la etapa C9).
- En primer lugar, la unidad de establecimiento de imagen de referencia 205 genera información Ref para cada píxel en el bloque blk de acuerdo con un método similar al de la etapa descrita anteriormente A6, donde Ref indica una referencia usada cuando la imagen diferencial se sometió a la codificación predictiva (véase la etapa C4).
- Los datos codificados introducidos se decodifican en el decodificador de imagen diferencial 209 usando la información anterior, de manera que se obtenga cada valor decodificado Dec de la imagen diferencial (véase la etapa C5).
- El valor decodificado Dec se usa directamente para decodificar un cuadro en otro tiempo u otro bloque, y de ese modo se almacena en la memoria de imágenes diferenciales decodificadas 207 (véase la etapa C6).
- Además, una imagen decodificada de la imagen obtenida por la cámara B se adquiere computando la suma de Dec y la imagen con la disparidad compensada Syn para cada píxel en el bloque blk.
- La imagen decodificada adquirida se emite desde el aparato de decodificación de vídeo 200, y se almacena simultáneamente en la memoria de imágenes de cámara decodificadas 206 (véase la etapa C7).
- La figura 6 muestra un flujo de operación detallado del proceso (en la etapa C5) de decodificación de la imagen diferencial para cada bloque, que se ejecuta en el decodificador de imagen diferencial 209.
- En primer lugar, la información del modo de predicción se decodifica a partir de los datos codificados, y se establece como pmode[0] (véase la etapa D1). En este proceso, si otro elemento de información del modo de predicción se incluye en los datos codificados (véase la etapa D2), el elemento también se decodifica y se establece como pmode[1] (véase la etapa D3). Si no se incluye ningún otro elemento de información del modo de predicción (véase la etapa D2), pmode[1] se establece en el mismo valor que pmode[0] (véase la etapa D4).
- A continuación, el residuo de predicción para cada píxel, que se incluye en los datos codificados, se decodifica y se establece como Dres (véase la etapa D5).
- La operación después de D5 se lleva a cabo para cada píxel en el bloque. Es decir, después de que el índice de píxel "pix" se inicialice a cero (véase la etapa D6), el siguiente proceso (etapas D7 a D11) se lleva a cabo repetidamente mientras pix se incrementa en 1 (véase la etapa D12), hasta que pix alcance el número maxPix de los píxeles en el bloque (véase la etapa D13).

5 En la operación aplicada a cada píxel, en primer lugar, el conmutador de referencia 208 se opera de acuerdo con el valor de Ref[pix] (véase la etapa D7), de manera que la memoria intermedia de referencia se establezca como la memoria de imágenes de cámara decodificadas 206 (véase la etapa D8) o la memoria de imágenes diferenciales decodificadas 207 (véase la etapa D9).

A continuación, el valor Pred del píxel pix se predice en el modo de predicción pmode[Ref[pix]] (véase la etapa D 10).

10 Después, el valor decodificado Dec[pix] de la imagen diferencial se obtiene por la suma de Dres[pix] y Pred (véase la etapa D11).

15 En la presente forma de realización, la información de la disparidad, que indica a qué posición en la imagen de la cámara B corresponde cada píxel en la imagen de cámara de referencia, se obtiene en el exterior del aparato de codificación de vídeo 100 o el aparato de decodificación de vídeo 200, y la información se codifica, se transmite, y se decodifica.

20 Sin embargo, la información de la disparidad se puede obtener y codificar en el aparato de codificación de vídeo 100, y la información codificada se puede emitir junto con los datos codificados de la imagen diferencial. De forma similar, la información de la disparidad se puede recibir en el aparato de decodificación de vídeo 200, y la información recibida se puede decodificar y usar.

25 Además, la información, que no indica directamente a qué posición en la imagen de la cámara B corresponde cada píxel en la imagen de cámara de referencia, pero proporciona la relación correspondiente descrita anteriormente debido a una conversión aplicada a la información, se puede introducir y convertir a información de la disparidad (que indica la relación correspondiente) en el aparato de codificación de vídeo 100 y el aparato de decodificación de vídeo 200.

30 Como un ejemplo de tal información, hay información que consiste en información tridimensional de un objeto procesado gráficamente y parámetros de cámara. La información tridimensional pueden ser coordenadas tridimensionales de cada parte del objeto procesado gráficamente, o puede indicar una distancia desde el objeto procesado gráficamente a la cámara pertinente.

35 Adicionalmente, en las formas de realización la conmutación del grupo de imágenes de referencia como el objetivo de referencia se puede llevar a cabo para cada píxel en la imagen de cámara objetivo de codificación. Sin embargo, la conmutación se puede llevar a cabo para una unidad de sección específica que incluya una pluralidad de píxeles. Por ejemplo, cuando la conmutación se puede llevar a cabo para cada bloque que tenga un tamaño específico, se puede seleccionar un objetivo de referencia adecuado para un número de píxeles en el bloque, o se puede seleccionar el grupo de imágenes de cámara decodificadas si el bloque incluye al menos un píxel cuyo valor de píxel de la imagen con la disparidad compensada es ineficaz. La selección basada en una unidad de sección apropiada puede reducir el tiempo de procesamiento.

45 Los procesos de codificación y decodificación de vídeo como se describen anteriormente se pueden implementar usando un recurso de hardware o de firmware, y también se pueden implementar usando un ordenador y un programa de software. El programa de ordenador se puede proporcionar almacenándolo en un medio de almacenamiento legible por ordenador, o por medio de una red.

Aplicabilidad industrial

50 De acuerdo con la presente invención, cuando se somete una imagen diferencial obtenida mediante la compensación de la disparidad a la codificación predictiva, no se usa ninguna información nueva adicional, y una imagen de referencia apropiada se usa mediante una operación de conmutación para cada píxel. Por lo tanto, el residuo para el objetivo de codificación se puede reducir sin aumentar la cantidad de código requerida para la información adicional, implementándose de ese modo una codificación de vídeo eficiente de toda la imagen de vídeo de puntos de vista múltiples.

REIVINDICACIONES

1. Un método de codificación de vídeo de codificar una imagen de vídeo de puntos de vista múltiples usando la compensación de la disparidad, comprendiendo el método las etapas de:

5 generar, en base a la información de la disparidad entre una imagen de cámara de referencia, que se obtiene decodificando una imagen de cámara ya codificada, y una imagen de cámara objetivo de codificación correspondiente a la imagen de cámara de referencia, una imagen con la disparidad compensada para la imagen de cámara objetivo de codificación;

10 establecer la información del objetivo de referencia, que indica un objetivo de referencia, para cada una de las secciones de procesamiento de codificación en la imagen de cámara objetivo de decodificación de acuerdo con la presencia o ausencia de la imagen con la disparidad compensada correspondiente a dicha cada una de las secciones de procesamiento de codificación;

15 seleccionar, para cada una de las secciones de procesamiento de codificación en la imagen de cámara objetivo de codificación, una de una imagen diferencial decodificada, que se obtiene decodificando una imagen diferencial entre la imagen de cámara ya codificada y la imagen con la disparidad compensada, y una imagen de cámara decodificada, que se obtiene decodificando la imagen de cámara ya codificada, como el objetivo de referencia en base a la información del objetivo de referencia establecida;

20 codificar de forma predictiva una imagen diferencial entre la imagen de cámara objetivo de codificación y la imagen con la disparidad compensada remitiéndose a la imagen seleccionada como el objetivo de referencia;

25 almacenar una imagen diferencial decodificada, que se obtiene decodificando la imagen diferencial codificada de forma predictiva, y

30 almacenar una imagen de cámara decodificada, que se obtiene decodificando la imagen de cámara ya codificada en base a la imagen diferencial decodificada obtenida decodificando la imagen diferencial codificada de forma predictiva.

2. El método de codificación de vídeo de acuerdo con la reivindicación 1, en el que:

35 la etapa de codificar de forma predictiva la imagen diferencial incluye computar un coste de codificación que se produce cuando cada uno de los métodos de predicción predeterminados se aplica a cada bloque de procesamiento de codificación que pertenece a la imagen diferencial y consiste en una pluralidad de píxeles, donde la predicción de vídeo de un bloque de procesamiento de codificación se lleva a cabo usando un único método de predicción que proporciona el coste de codificación mínimo.

40 3. El método de codificación de vídeo de acuerdo con la reivindicación 1, en el que:

45 la etapa de codificar de forma predictiva la imagen diferencial incluye computar un coste de codificación que se produce cuando cada uno de los métodos de predicción predeterminados se aplica a cada grupo de píxeles en cada bloque de procesamiento de codificación que pertenece a la imagen diferencial y consiste en una pluralidad de píxeles, donde cada grupo de píxeles tiene una imagen de referencia individual, y la codificación se lleva a cabo usando un único método de predicción o una pluralidad de métodos de predicción, de acuerdo con los costes computados.

50 4. El método de codificación de vídeo de acuerdo con la reivindicación 1, en el que cada una de las secciones de procesamiento de codificación es un píxel.

5. Un método de decodificación de vídeo de decodificar una imagen de vídeo de puntos de vista múltiples usando la compensación de la disparidad, comprendiendo el método las etapas de:

55 generar, en base a la información de la disparidad entre una imagen de cámara de referencia ya decodificada y una imagen de cámara objetivo de decodificación correspondiente a la imagen de cámara de referencia, una imagen con la disparidad compensada para la imagen de cámara objetivo de decodificación;

60 establecer la información del objetivo de referencia, que indica un objetivo de referencia, para cada una de las secciones de procesamiento de decodificación en la imagen de cámara objetivo de decodificación de acuerdo con la presencia o ausencia de la imagen con la disparidad compensada correspondiente a dicha cada una de las secciones de procesamiento de decodificación;

65 seleccionar, para cada una de las secciones de procesamiento de decodificación en la imagen de cámara objetivo de decodificación, una de una imagen diferencial decodificada que es la imagen diferencial entre la imagen de

cámara ya decodificada y la imagen con la disparidad compensada, y una imagen de cámara decodificada que es la imagen de cámara ya decodificada, como el objetivo de referencia en base a la información del objetivo de referencia establecida;

5 decodificar una imagen diferencial entre la imagen de cámara objetivo de decodificación y la imagen con la disparidad compensada remitiéndose a la imagen seleccionada como el objetivo de referencia, en base a un residuo de predicción de cada píxel, que se obtiene decodificando los datos codificados introducidos;

almacenar la imagen diferencial decodificada obtenida mediante la etapa de decodificación; y

10 añadir la imagen diferencial decodificada a la imagen con la disparidad compensada, emitiendo la suma de ellas como una imagen de cámara decodificada de la imagen de vídeo de puntos de vista múltiples, y almacenar esta imagen de cámara decodificada.

15 6. El método de decodificación de vídeo de acuerdo con la reivindicación 5, en el que:

la etapa de decodificar la imagen diferencial incluye decodificar la imagen diferencial de una manera en la que un único método de predicción designado se aplique a cada bloque de procesamiento de decodificación que pertenece a la imagen de cámara objetivo de decodificación y consiste en una pluralidad de píxeles.

20 7. El método de decodificación de vídeo de acuerdo con la reivindicación 5, en el que:

la etapa de decodificar la imagen diferencial incluye decodificar la imagen diferencial mientras se selecciona para cada bloque de procesamiento de decodificación que pertenece a la imagen de cámara objetivo de decodificación y consiste en una pluralidad de píxeles, si se usa un único método de predicción designado o una pluralidad de métodos de predicción designados.

25 8. El método de decodificación de vídeo de acuerdo con la reivindicación 5, en el que cada una de las secciones de procesamiento de decodificación es un píxel.

30 9. Un aparato de codificación de vídeo (100) para codificar una imagen de vídeo de puntos de vista múltiples usando la compensación de la disparidad, comprendiendo el aparato:

35 un dispositivo (104) para generar, en base a la información de la disparidad entre una imagen de cámara de referencia, que se obtiene decodificando una imagen de cámara ya codificada, y una imagen de cámara objetivo de codificación correspondiente a la imagen de cámara de referencia, una imagen con la disparidad compensada para la imagen de cámara objetivo de codificación;

40 un dispositivo para establecer la información del objetivo de referencia, que indica un objetivo de referencia, para cada una de las secciones de procesamiento de codificación en la imagen de cámara objetivo de codificación de acuerdo con la presencia o ausencia de la imagen con la disparidad compensada correspondiente a dicha cada una de las secciones de procesamiento de codificación;

45 un dispositivo (106) para seleccionar, para cada una de las secciones de procesamiento de codificación en la imagen de cámara objetivo de codificación, una de una imagen diferencial decodificada, que se obtiene decodificando una imagen diferencial entre la imagen de cámara ya codificada y la imagen con la disparidad compensada, y una imagen de cámara decodificada, que se obtiene decodificando la imagen de cámara ya codificada, como el objetivo de referencia en base a la información del objetivo de referencia establecida;

50 un dispositivo (105) para codificar de forma predictiva una imagen diferencial entre la imagen de cámara objetivo de codificación y la imagen con la disparidad compensada remitiéndose a la imagen seleccionada como el objetivo de referencia;

55 un dispositivo (108) para almacenar una imagen diferencial decodificada, que se obtiene decodificando la imagen diferencial codificada de forma predictiva; y

un dispositivo (107) para almacenar una imagen de cámara decodificada, que se obtiene decodificando la imagen de cámara ya codificada en base a la imagen diferencial decodificada obtenida decodificando la imagen diferencial codificada de forma predictiva.

60 10. El aparato de codificación de vídeo de acuerdo con la reivindicación 9, en el que cada una de las secciones de procesamiento de codificación es un píxel.

65 11. Un aparato de decodificación de vídeo (200) para decodificar una imagen de vídeo de puntos de vista múltiples usando la compensación de la disparidad, comprendiendo el aparato:

un dispositivo (204) para generar, en base a la información de la disparidad entre una imagen de cámara de referencia ya decodificada y una imagen de cámara objetivo de decodificación correspondiente a la imagen de cámara de referencia, una imagen con la disparidad compensada para la imagen de cámara objetivo de decodificación;

5 un dispositivo para establecer la información del objetivo de referencia, que indica un objetivo de referencia, para cada una de las secciones de procesamiento de decodificación en la imagen de cámara objetivo de decodificación de acuerdo con la presencia o ausencia de la imagen con la disparidad compensada correspondiente a dicha cada una de las secciones de procesamiento de decodificación;

10 un dispositivo (205) para seleccionar, para cada una de las secciones de procesamiento de decodificación en la imagen de cámara objetivo de decodificación, una de una imagen diferencial decodificada que es la imagen diferencial entre la imagen de cámara ya decodificada y la imagen con la disparidad compensada, y una imagen de cámara decodificada que es la imagen de cámara ya decodificada, como el objetivo de referencia en base a la información del objetivo de referencia establecida;

15 un dispositivo (209) para decodificar una imagen diferencial entre la imagen de cámara objetivo de decodificación y la imagen con la disparidad compensada remitiéndose a la imagen seleccionada como el objetivo de referencia, en base a un residuo de predicción de cada píxel, que se obtiene decodificando los datos codificados introducidos;

20 un dispositivo (207) para almacenar la imagen diferencial decodificada obtenida mediante la etapa de decodificación; y

25 un dispositivo (206) para añadir la imagen diferencial decodificada a la imagen con la disparidad compensada, emitiendo la suma de ellas como una imagen de cámara decodificada de la imagen de vídeo de puntos de vista múltiples, y almacenando esta imagen de cámara decodificada.

30 12. El aparato de decodificación de vídeo de acuerdo con la reivindicación 11, en el que cada una de las secciones de procesamiento de decodificación es un píxel.

13. Un programa de codificación de vídeo mediante el cual un ordenador ejecuta un proceso para implementar el método de codificación de vídeo de acuerdo con la reivindicación 1.

35 14. Un programa de decodificación de vídeo mediante el cual un ordenador ejecuta un proceso para implementar el método de decodificación de vídeo de acuerdo con la reivindicación 5.

15. Un medio de almacenamiento legible por ordenador que almacena un programa de codificación de vídeo mediante el cual un ordenador ejecuta un proceso para implementar el método de codificación de vídeo de acuerdo con la reivindicación 1.

40 16. Un medio de almacenamiento legible por ordenador que almacena un programa de decodificación de vídeo mediante el cual un ordenador ejecuta un proceso para implementar el método de decodificación de vídeo de acuerdo con la reivindicación 5.

FIG. 1

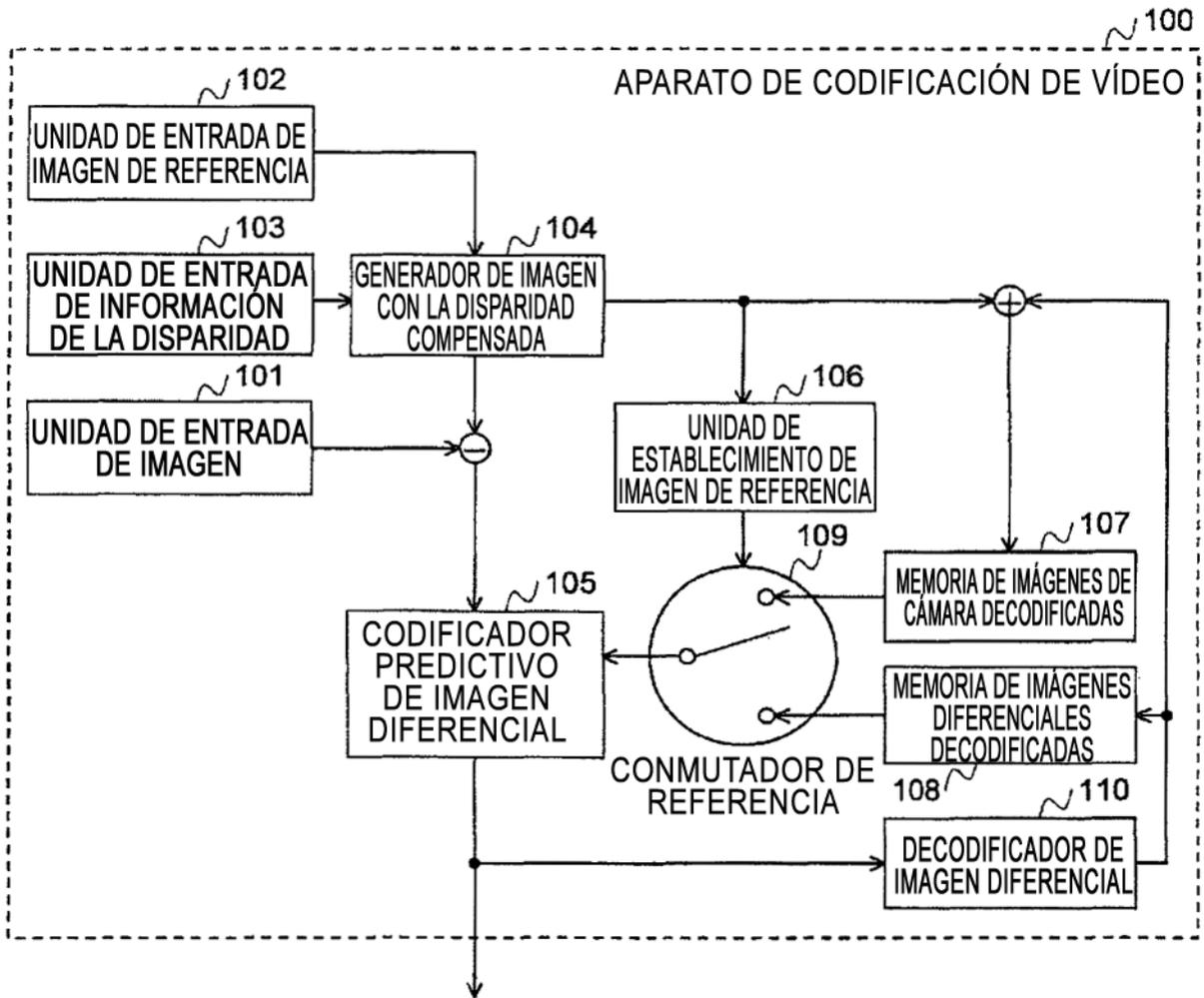


FIG. 2

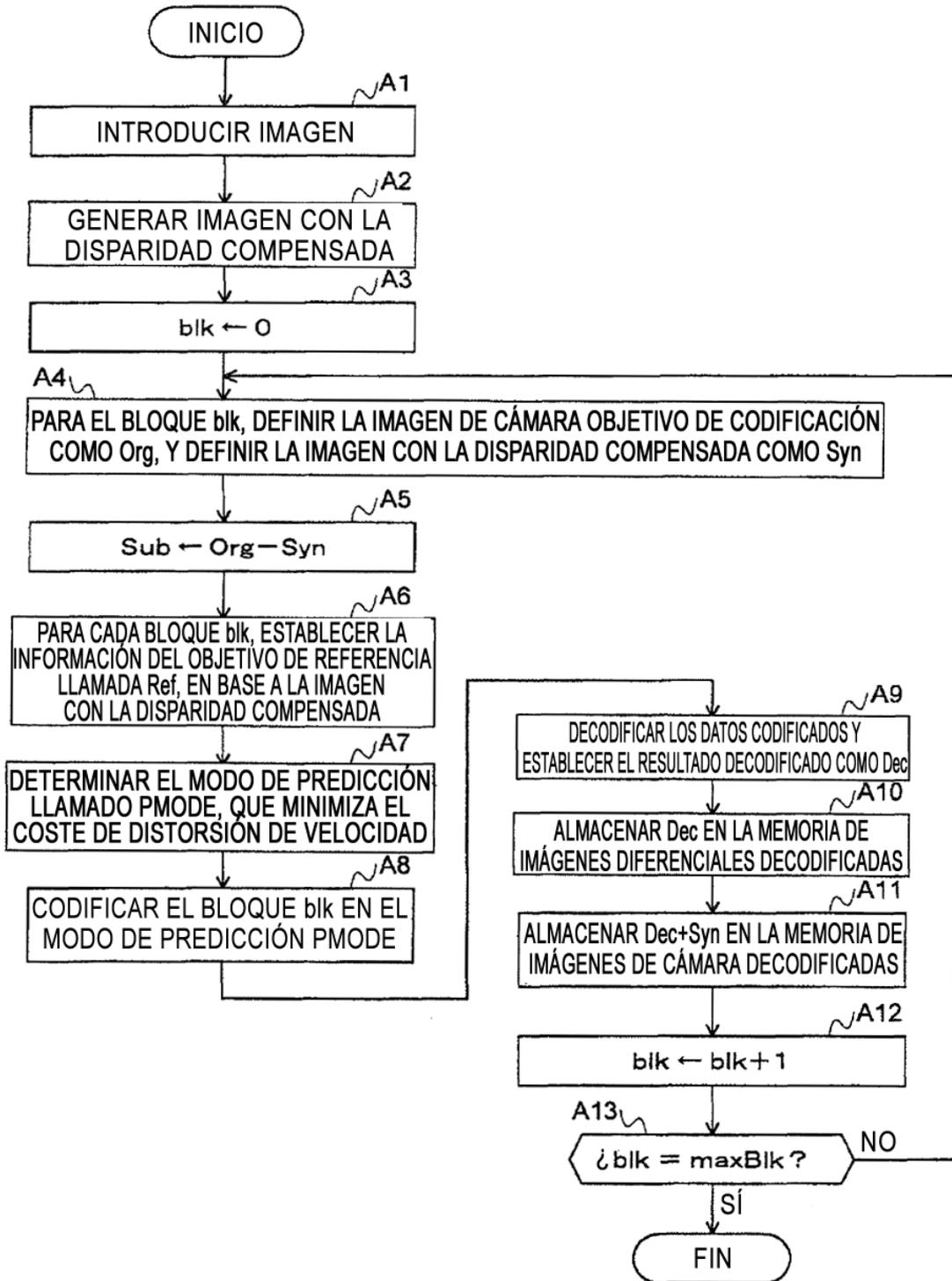


FIG. 3

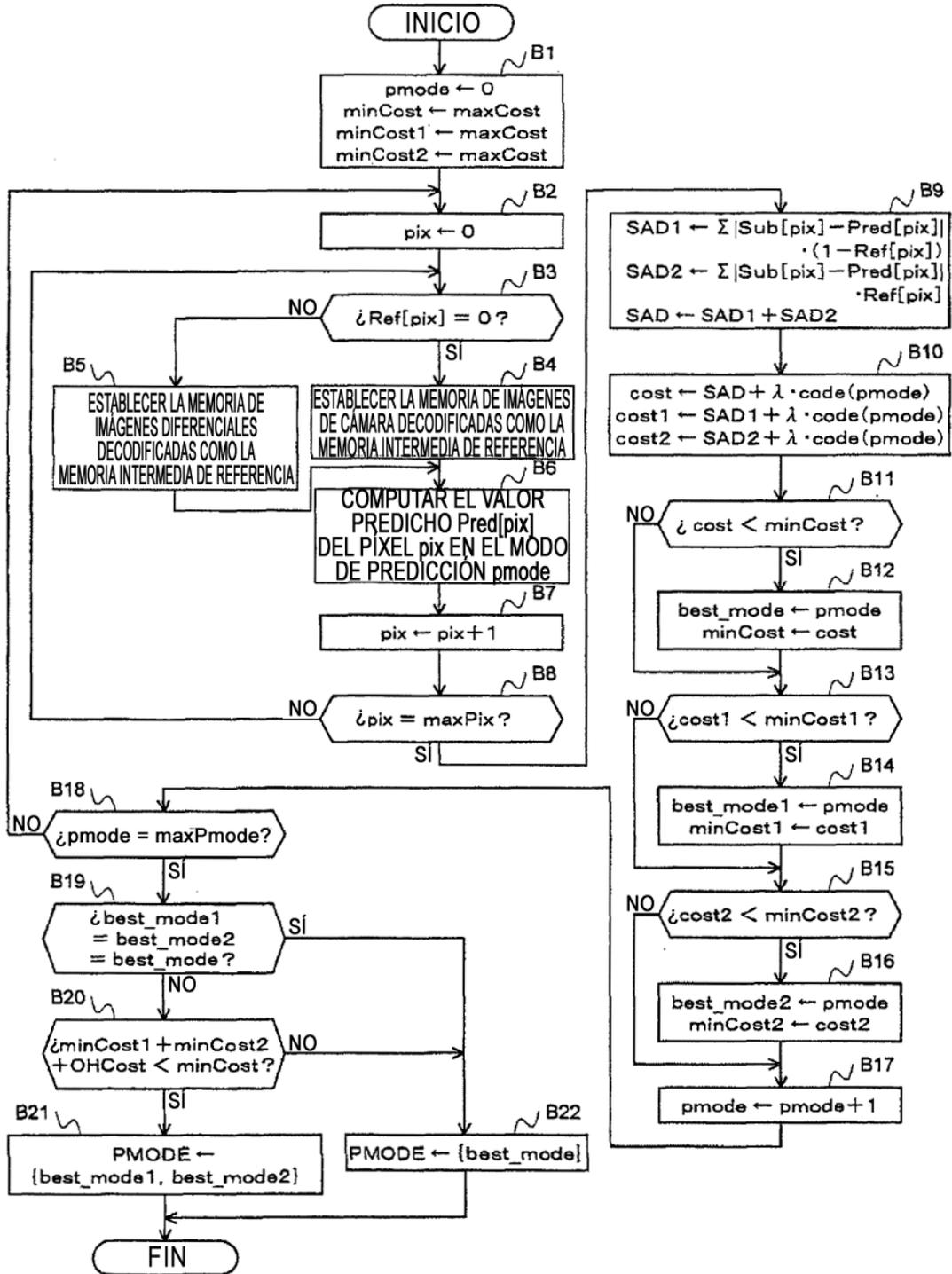


FIG. 4

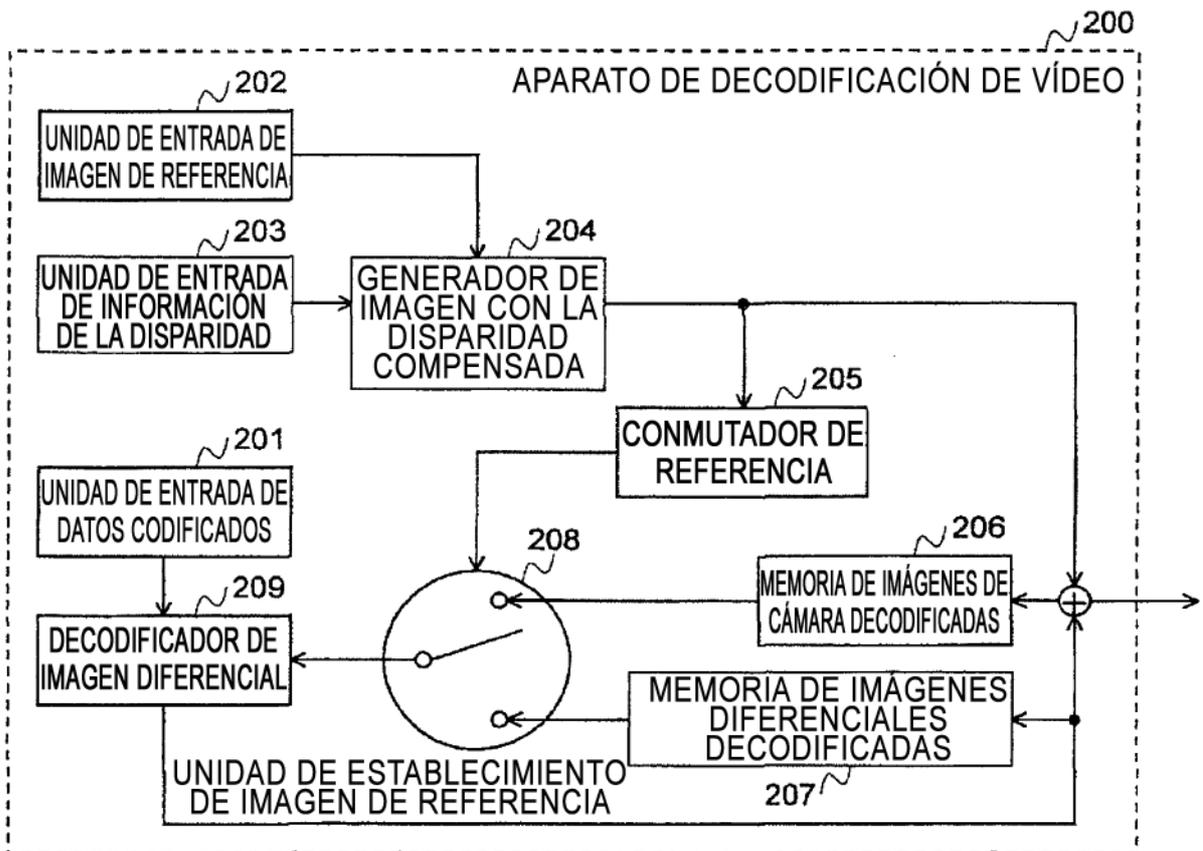


FIG. 5



FIG. 6

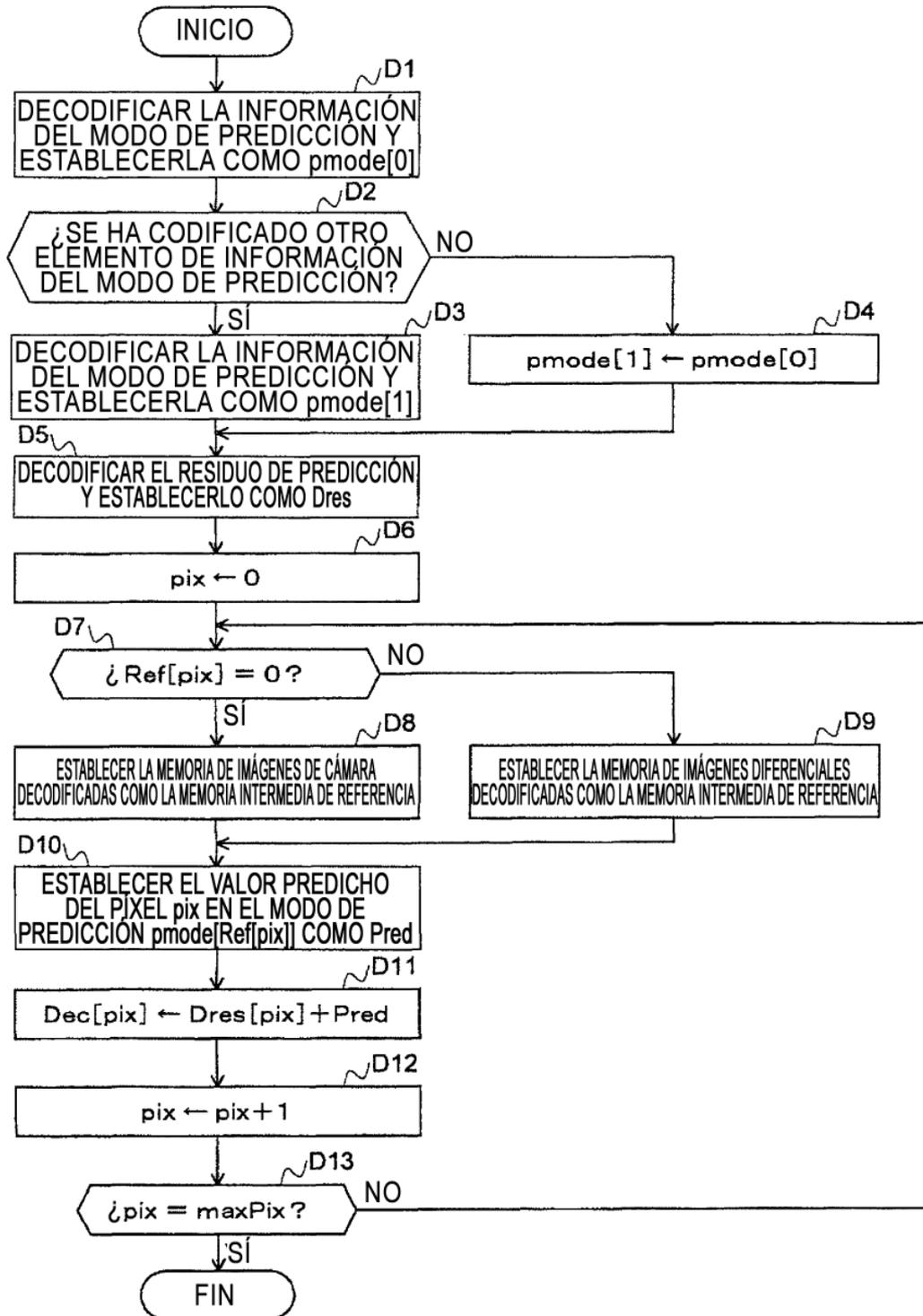


FIG. 7

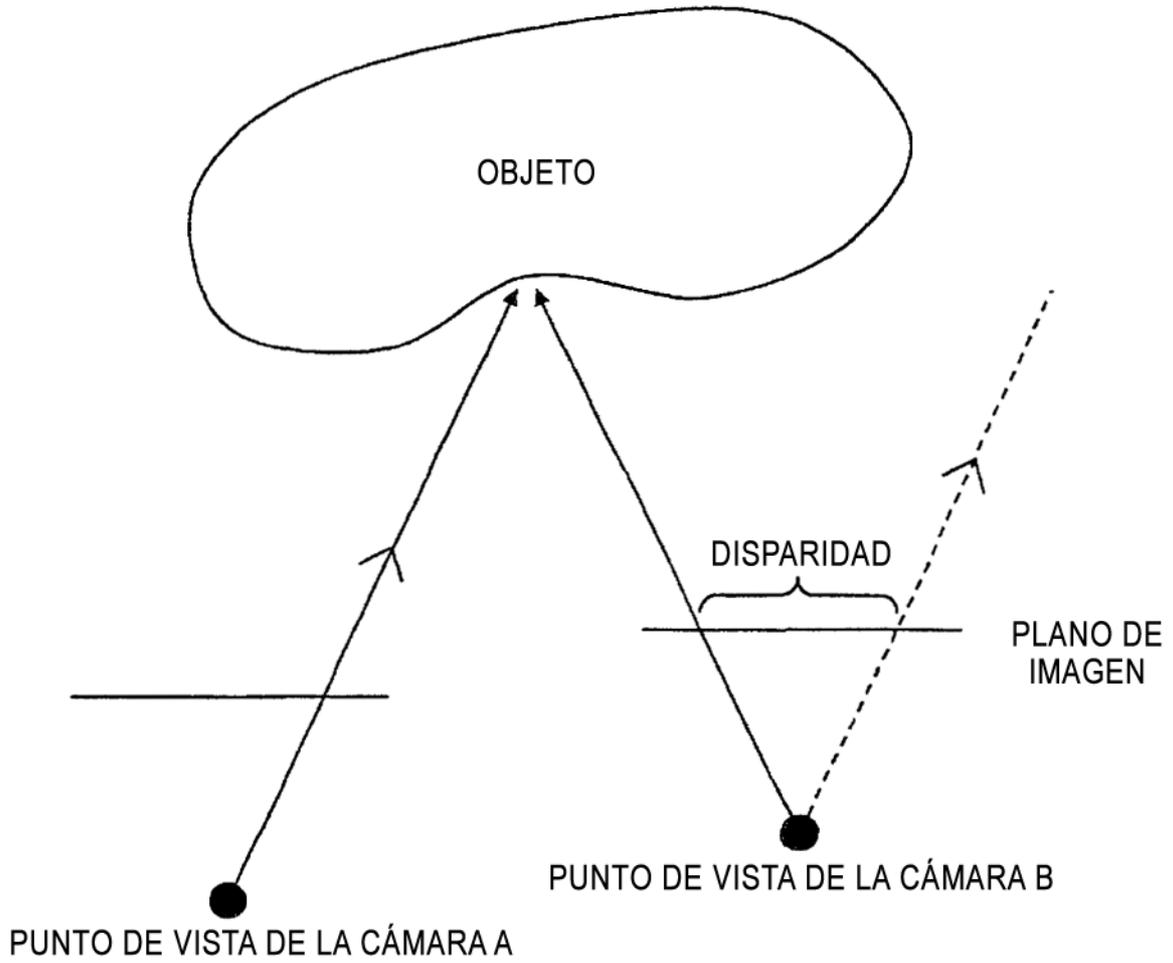


FIG. 8

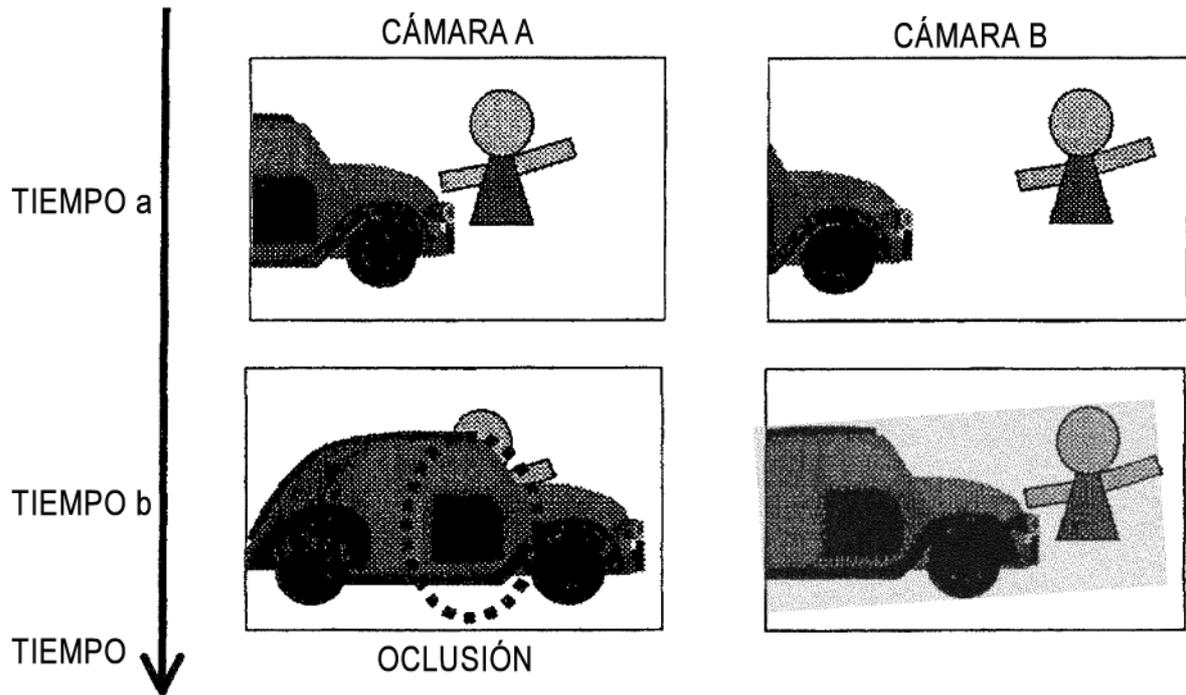


FIG. 9

