

19



OFICINA ESPAÑOLA DE  
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 524 973**

51 Int. Cl.:

**H04N 19/17** (2014.01)

**H04N 19/18** (2014.01)

**H04N 19/597** (2014.01)

**H04N 19/105** (2014.01)

**H04N 19/136** (2014.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

96 Fecha de presentación y número de la solicitud europea: **23.02.2010 E 10743600 (8)**

97 Fecha y número de publicación de la concesión europea: **27.08.2014 EP 2400759**

54 Título: **Codificación y decodificación de imágenes multivista usando corrección de iluminación y de color localizada**

30 Prioridad:

**23.02.2009 JP 2009038786**

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

**16.12.2014**

73 Titular/es:

**NIPPON TELEGRAPH AND TELEPHONE CORPORATION (50.0%)**  
**3-1 Otemachi 2-chome**  
**Chiyoda-ku, Tokyo 100-8116, JP y**  
**NATIONAL UNIVERSITY CORPORATION**  
**NAGOYA UNIVERSITY (50.0%)**

72 Inventor/es:

**SHIMIZU, SHINYA;**  
**KIMATA, HIDEAKI y**  
**TANIMOTO, MASAYUKI**

74 Agente/Representante:

**LINAGE GONZÁLEZ, Rafael**

**ES 2 524 973 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**

Codificación y decodificación de imágenes multivista usando corrección de iluminación y de color localizada

**5 Campo técnico**

La presente invención se refiere a un método de codificación de imágenes multivista y un dispositivo para codificar imágenes fotografiadas por una pluralidad de cámaras que están fotografiando un objeto en particular, y también a un método de decodificación de imágenes multivista y un dispositivo para decodificación de datos codificados que han sido codificados usando este método de codificación de imágenes multivista, y también a un programa de codificación de imágenes multivista que se usa para implementar este método de codificación de imágenes multivista, y a un programa de decodificación de imágenes multivista que se usa para implementar este método de decodificación de imágenes multivista.

**15 Técnica anterior**

El término 'imágenes multivista' se refiere a una pluralidad de imágenes obtenidas fotografiando el mismo objeto y el fondo usando una pluralidad de cámaras, mientras que el término imágenes en movimiento multivista (es decir, 'vídeo multivista') se refiere a imágenes en movimiento obtenidas de esta manera.

La predicción con compensación de movimiento y la predicción con compensación de disparidad se han propuesto como tecnologías para su uso en la codificación de imágenes en movimiento general y la codificación de imágenes en movimiento multivista.

La predicción con compensación de movimiento es un método que se emplea también en las normas internacionales para los formatos de codificación de imágenes en movimiento de los últimos años tipificados por H.264. En este método, el movimiento de un objeto se compensa entre una trama destinada a codificación y una trama de referencia que ya ha sido codificada de manera que se obtiene una diferencia inter-trama para la señal de imagen, y sólo se codifica esta señal de diferencia (véase el documento no patente 1).

En cambio, en la predicción con compensación de disparidad, al compensar las disparidades en un objeto usando una trama fotografiada por una cámara diferente a la de la trama de referencia, la codificación puede realizarse como las diferencias inter-trama entre las señales de imagen que se están obteniendo (véase el documento no patente 2).

El término 'disparidad' que se usa en el presente documento se refiere a diferencias en las posiciones en los planos de imagen de cámaras colocadas en diferentes posiciones en las que se proyecta la misma posición en un objeto. En la predicción con compensación de disparidad, esto se representa mediante vectores bidimensionales y después se codifica. Tal como se muestra en la figura 9, dado que las disparidades son información cuya creación depende de la posición de la cámara y de la distancia desde la cámara (es decir, la profundidad), existe un método conocido como predicción de síntesis de vistas (predicción de interpolación de vistas) que usa este principio.

En la predicción de síntesis de vistas (predicción de interpolación de vistas), existe un método en el que la profundidad de un objeto se estima usando la información de la posición de la cámara y la teoría de la triangulación para los vídeos multivista obtenidos en el lado de codificación o el lado de decodificación, y las tramas destinadas a su codificación se sintetizan (es decir, se interpolan) usando esta información de profundidad estimada de manera que se crea una imagen de predicción (véanse el documento no patente 1 y el documento no patente 3). Debe observarse que si la profundidad se estima en el lado de codificación, es necesario codificar la profundidad que se usa.

En la predicción que usa imágenes fotografiadas usando estas cámaras separadas, si existen diferencias individuales entre las respuestas de los elementos de obtención de imágenes de las cámaras, o si se aplica control de ganancia o corrección gamma en cada cámara, o si los ajustes de la profundidad de campo o la apertura o similares son diferentes en cada cámara, o si existe un efecto de iluminación dependiente de la dirección en la escena, o similares, entonces la eficacia de codificación se deteriora. Esto se debe a que la predicción se realiza con la suposición de que la iluminación y el color del objeto son los mismos tanto en la trama destinada a codificación como en la trama de referencia.

Se están investigando métodos tales como compensación de iluminación y corrección de color como formas de abordar los cambios en la iluminación y el color de un objeto. En estos métodos, usando una trama de referencia cuya iluminación y cuyo color hayan sido corregidos como trama que se usa para realizar una predicción, es posible limitar a un mínimo la magnitud de la predicción residual que se codifica.

En H.264, se adopta la predicción ponderada en la que se usa una función lineal como modelo de corrección (véase el documento no patente 1), mientras que en el documento no patente 3 se propone un método en el que las correcciones se realizan usando una tabla de colores.

[DOCUMENTOS DE LA TÉCNICA ANTERIOR]

5 [Documento de patente 1] Solicitud de patente japonesa abierta a consulta por el público (JP-A) nº 2007-036.800 "Método de codificación de vídeo, método de decodificación de vídeo, programa de codificación de vídeo, programa de decodificación de vídeo y medio de grabación legible por ordenador en el que se graban estos programas"

[Documento no patente 1]

10 ITU-T Rec. H.264/ISO/IEC 11496-10, "Codificación de vídeo avanzada para los servicios audiovisuales genéricos", Borrador de Comité Final, Documento JVT-E022d7, septiembre de 2002. (pág. 10-13, pág. 62-73)

[Documento no patente 2]

15 Hideaki Kimata y Masaki Kitahara, "Preliminary results on multiple view video coding (3DAV)", documento M10976 MPEG Redmond Meeting, julio de 2004.

[Documento no patente 3]

20 K. Yamamoto, M. Kitahara, H. Kimata, T. Yendo, T. Fujii, M. Tanimoto, S. Shimizu, K. Kamikura y Y. Yashima, "Multi-view Video Coding Using View Interpolation and Color Correction", IEEE Transactions on Circuits and System for Video Technology, Vol. 17, nº 11, pág. 1436-1449, noviembre de 2007.

**Descripción de la invención**

25 Problema que debe ser resuelto por la invención

Cuando se realiza una codificación usando la compensación de iluminación y corrección de color mencionada anteriormente se producen los dos problemas siguientes.

30 El primer problema es el aumento en la cantidad de código que procede de la adición de parámetros para compensación de iluminación y corrección de color y similares. En predicción con compensación de disparidad normal y predicción de síntesis (interpolación) de vistas, dado que se hace necesario codificar parámetros para compensación de iluminación y corrección de color y similares que anteriormente no habían necesitado codificación, existe un deterioro en la eficacia de codificación.

40 El segundo problema es la exactitud de la corrección. En caso de fundido y parpadeo en la codificación normal de imágenes en movimiento, dado que toda la pantalla cambia de la misma forma, es posible realizar una compensación de iluminación y una corrección de color y similares satisfactorias usando un único parámetro de corrección. Sin embargo, las ausencias de correspondencia (es decir, las discrepancias en iluminación y color) que son causadas porque el objeto no es un reflector difuso completo, o porque la profundidad de campo y el enfoque no guardan una total correspondencia en cada cámara no dependen de la escena, sino del objeto. En consecuencia, en la corrección que se basa en un único parámetro de corrección, se producen casos en los que, dependiendo del objeto, se incrementa la ausencia de correspondencia.

45 Para contrarrestar este problema, puede considerarse un método en el que se usa una pluralidad de parámetros de corrección con el fin de abordar la ausencia de correspondencia en cada objeto individual. Sin embargo, si se usa este método, entonces además de la cantidad de código requerida para codificar una pluralidad de parámetros de corrección, es necesario también codificar información que muestre el parámetro de corrección que deberá usarse en cada zona de imagen. En consecuencia, la cantidad de código aumenta todavía más, y no es posible resolver el primer problema.

50 La presente invención se concibió a la luz de las circunstancias descritas anteriormente, y un objeto de la misma es proporcionar una nueva tecnología de codificación y decodificación de imágenes multivista que consiga una codificación altamente eficaz incluso en imágenes multivista (es decir, imágenes estáticas e imágenes en movimiento multivista) en las que se genere una ausencia de correspondencia localizada de iluminación y color entre las cámaras, y que consiga también una reducción en la cantidad de código requerida cada vez que se emplee esta nueva codificación.

60 Medios para resolver el problema

[1] Idea tecnológica básica que subyace a la presente invención

65 Con el fin de resolver los problemas descritos anteriormente, en la presente invención, se han ideado los medios siguientes para los casos en que una trama destinada a codificación-decodificación se divida primero en zonas, y a continuación se realice codificación y decodificación de imágenes multivista en cada zona.

En primer lugar, se determina la información de profundidad para un objeto que se va a fotografiar en una zona destinada a tratamiento. A continuación, en una zona adyacente a la zona destinada a tratamiento en la que ya se ha completado la codificación (decodificación), a partir de la información de profundidad se determina un grupo de 5 píxeles en el que se ha fotografiado el mismo objeto que en la zona destinada a tratamiento y se establece como un grupo de píxeles de muestra. A continuación, se crea una imagen de síntesis de vistas para los píxeles contenidos en este grupo de píxeles de muestra, y se crea también una imagen de síntesis de vistas para los píxeles contenidos en la zona destinada a tratamiento. A continuación, se estiman los parámetros de corrección para corregir la ausencia de correspondencia en iluminación y color a partir de la imagen de síntesis de vistas y de la imagen 10 decodificada en el grupo de píxeles de muestra. A continuación, mediante corrección de la imagen de síntesis de vistas creada para la zona destinada a tratamiento usando los parámetros de corrección estimados, se crea una imagen de predicción que se usará para codificar (decodificar) señales de imagen en la zona destinada a tratamiento.

15 En el caso del método convencional en el que se calculan los parámetros de corrección comparando la trama destinada a codificación con una trama de referencia, dado que la trama destinada a codificación no puede ser adquirida por el lado de decodificación, es necesario codificar los parámetros de corrección.

20 En cambio, en la presente invención, los parámetros de corrección se calculan comparando señales de imagen de zonas ya codificadas/decodificadas de una trama destinada a codificación/decodificación con información de una imagen de síntesis de vistas creada usando una trama de referencia. Dado que estas señales pueden ser adquiridas por el lado de decodificación, no es necesario codificar los parámetros de corrección. Es decir, es posible, por medio de la presente invención, resolver el problema de un aumento en la cantidad de código.

25 Por otra parte, dado que la codificación es un proceso en el que las señales de entrada son convertidas de la forma más correcta posible, puede considerarse que las señales de imagen que han sido ya codificadas/decodificadas son sustancialmente las mismas que las señales de imagen destinadas a codificación. Es decir, los parámetros de corrección calculados por medio de la presente invención pueden llevar la imagen sintetizada extraordinariamente 30 significativamente cerca de la imagen destinada a codificación, y la predicción residual que debe codificarse puede reducirse significativamente.

35 Por otra parte, en la presente invención, los parámetros de corrección se estiman usando la información de zonas adyacentes en las que ha sido fotografiado el mismo objeto que el objeto fotografiado en la zona destinada a tratamiento. Al actuar de esta manera, es posible corregir la iluminación y el color que dependen del objeto. Debe observarse que dado que en esta determinación del objeto se usa la información de profundidad que se requiere cuando se va a crear una imagen de síntesis de vistas, no hay necesidad de codificar y transmitir información adicional.

40 En la codificación/decodificación multivista descrita anteriormente, comparando la variación de información de profundidad en una zona destinada a tratamiento con un valor umbral predefinido, es posible determinar si una pluralidad de objetos ha sido fotografiada o no dentro de una zona destinada a tratamiento. Si una pluralidad de objetos ha sido fotografiada, entonces se establece la información de profundidad y el grupo de píxeles de muestras para cada objeto, y se estiman los parámetros de corrección. Debe observarse que al tratar los objetos que tienen menos de un número fijado de píxeles en una zona destinada a tratamiento con otros objetos, es posible evitar todo 45 incremento en la cantidad de cálculos.

Además, en la codificación/decodificación de imágenes multivista descrita anteriormente, los modelos de corrección de los cuales existe una pluralidad (es decir, el número de parámetros de corrección) se modifican basándose en el número de píxeles presentes en un grupo de píxeles de muestra.

50 [2] Estructura de la presente invención

A continuación, se describirá la estructura del dispositivo de codificación de imágenes multivista y el dispositivo de decodificación de imágenes multivista de la presente invención.

55 [2-1] Estructura del dispositivo de codificación de imágenes multivista de la presente invención

El dispositivo de codificación de imágenes multivista de la presente invención es un dispositivo que codifica 60 imágenes multivista (es decir, imágenes estáticas e imágenes en movimiento fotografiadas desde múltiples puntos de vista) dividiendo una imagen de entrada de un objeto que es fotografiado por una primera cámara en una pluralidad de zonas objeto de codificación y, usando una imagen de síntesis de vistas que se sintetiza a partir de la información de profundidad para la imagen de entrada y a partir de una imagen ya codificada del mismo objeto fotografiado por una segunda cámara que está situada en una posición diferente a la de la primera cámara, realizando codificación predictiva para cada una de las zonas objeto de codificación, y que incluye: (1) una unidad de 65 determinación de profundidad representativa que establece la información de profundidad representativa para un objeto fotografiado en la zona objeto de codificación; (2) una unidad de determinación de grupo de píxeles de

muestra que, basándose en la información de profundidad para una zona ya codificada que es adyacente a la zona objeto de codificación y en la información de profundidad representativa, determina un grupo de píxeles en el que ha sido fotografiado el mismo objeto que en la zona objeto de codificación y establece el grupo de píxeles como un grupo de píxeles de muestra; (3) una unidad de estimación de parámetros de corrección que, basándose en la imagen de síntesis de vistas para el grupo de píxeles de muestra y en una imagen decodificada que ha sido decodificada para el grupo de píxeles de muestra, estima los parámetros de corrección para corregir las ausencias de correspondencia de iluminación o de color; (4) una unidad de corrección de imágenes de síntesis de vistas que, usando los parámetros de corrección, corrige la imagen de síntesis de vistas para la zona objeto de codificación de manera que se crea una imagen de síntesis de vistas corregida; (5) una unidad de codificación de imagen que, usando la imagen de síntesis de vistas corregida, codifica una señal de imagen de la imagen objeto de codificación de manera que se crean datos codificados; y (6) una unidad de decodificación de imagen que decodifica los datos codificados de manera que se crea una imagen decodificada para la zona objeto de codificación.

También es posible que al dispositivo de codificación de imágenes multivista según una realización de la presente invención se le proporcione además (7) una unidad de determinación de objetos que divide los píxeles en la zona objeto de codificación en uno o varios grupos usando la información de profundidad para los píxeles pertinentes como referencia. En este caso, también es posible que la unidad de determinación de profundidad representativa establezca la información de profundidad representativa para cada grupo determinado por la unidad de determinación de objetos, que la unidad de determinación de grupo de píxeles de muestra establezca un grupo de píxeles de muestra para cada grupo determinado por la unidad de determinación de objetos, que la unidad de estimación de parámetros de corrección estime los parámetros de corrección para cada grupo determinado por la unidad de determinación de objetos y que la unidad de corrección de imágenes de síntesis de vistas corrija una imagen de síntesis de vistas para cada grupo determinado por la unidad de determinación de objetos.

Por otra parte, también es posible que al dispositivo de codificación de imágenes multivista según una realización de la presente invención se le proporcione además (8) una unidad de selección de modelo de corrección que seleccione un modelo de corrección para corregir la imagen de síntesis de vistas para la zona objeto de codificación de acuerdo con el número de píxeles en el grupo de píxeles de muestra. En este caso, también es posible que la unidad de estimación de parámetros de corrección estime los parámetros de corrección para el modelo de corrección seleccionado por la unidad de selección de modelo de corrección, y que la unidad de corrección de imágenes de síntesis de vistas corrija la imagen de síntesis de vistas usando el modelo de corrección seleccionado por la unidad de selección de modelo de corrección.

El método de codificación de imágenes multivista de la presente invención que se implementa como resultado de cada uno de los dispositivos de tratamiento mencionados anteriormente que realizan sus operaciones respectivas puede aplicarse también por medio de un programa informático. Este programa informático es suministrado grabado en un medio de grabación adecuado legible por el ordenador, o es suministrado a través de una red. Cuando vaya a aplicarse la presente invención, el programa informático se instala en un ordenador y la presente invención se consigue cuando el programa informático se ejecuta en una unidad de control tal como una CPU.

[2-2] Estructura del dispositivo de decodificación de imágenes multivista de la presente invención

El dispositivo de decodificación de imágenes multivista de la presente invención es un dispositivo que decodifica datos codificados para imágenes multivista (es decir, imágenes estáticas e imágenes en movimiento fotografiadas desde múltiples puntos de vista) dividiendo una imagen objeto de decodificación de un objeto que es fotografiado por una primera cámara en una pluralidad de zonas objeto de decodificación y, usando una imagen de síntesis de vistas que se sintetiza a partir de la información de profundidad para la imagen objeto de decodificación y a partir de una imagen decodificada del mismo objeto fotografiado por una segunda cámara que está situada en una posición diferente a la de la primera cámara, realizando decodificación predictiva para cada una de las zonas objeto de decodificación, y que incluye: (1) una unidad de determinación de profundidad representativa que establece la información de profundidad representativa para un objeto fotografiado en la zona objeto de decodificación; (2) una unidad de determinación de grupo de píxeles de muestra que, basándose en la información de profundidad para una zona ya decodificada que es adyacente a la zona objeto de decodificación y en la información de profundidad representativa, determina un grupo de píxeles en el que el mismo objeto que en la zona objeto de decodificación ha sido fotografiado y establece el grupo de píxeles como un grupo de píxeles de muestra; (3) una unidad de estimación de parámetros de corrección que, basándose en la imagen de síntesis de vistas para el grupo de píxeles de muestra y en una imagen decodificada que ha sido decodificada para el grupo de píxeles de muestra, estima los parámetros de corrección para corregir las ausencias de correspondencia de iluminación o de color; (4) una unidad de corrección de imágenes de síntesis de vistas que, usando los parámetros de corrección, corrige la imagen de síntesis de vistas para la zona objeto de decodificación de manera que se crea una imagen de síntesis de vistas corregida; y (5) una unidad de decodificación de imagen que, usando la imagen de síntesis de vistas corregida, decodifica una señal de imagen de la imagen objeto de decodificación.

También es posible que al dispositivo de decodificación de imágenes multivista según una realización de la presente invención se le proporcione además (6) una unidad de determinación de objetos que divide los píxeles en la zona objeto de decodificación en uno o varios grupos usando la información de profundidad para los píxeles pertinentes

5 como referencia. En este caso, también es posible que la unidad de determinación de profundidad representativa establezca la información de profundidad representativa para cada grupo determinado por la unidad de determinación de objetos, que la unidad de determinación de grupo de píxeles de muestra establezca un grupo de píxeles de muestra para cada grupo determinado por la unidad de determinación de objetos, que la unidad de estimación de parámetros de corrección estime los parámetros de corrección para cada grupo determinado por la unidad de determinación de objetos y que la unidad de corrección de imágenes de síntesis de vistas corrija una imagen de síntesis de vistas para cada grupo determinado por la unidad de determinación de objetos.

10 También es posible que al dispositivo de decodificación de imágenes multivista según una realización de la presente invención se le proporcione además (7) una unidad de selección de modelo de corrección que selecciona un modelo de corrección para corregir la imagen de síntesis de vistas para la zona objeto de decodificación de acuerdo con el número de píxeles en el grupo de píxeles de muestra. En este caso, la unidad de estimación de parámetros de corrección estima los parámetros de corrección para el modelo de corrección seleccionado por la unidad de selección de modelo de corrección, y la unidad de corrección de imágenes de síntesis de vistas corrige la imagen de síntesis de vistas usando el modelo de corrección seleccionado por la unidad de selección de modelo de corrección.

15 El método de decodificación de imágenes multivista de la presente invención que se implementa como resultado de cada uno de los dispositivos de tratamiento descritos anteriormente que realizan sus operaciones respectivas también puede conseguirse por medio de un programa informático. Este programa informático se suministra grabado en un medio de grabación adecuado legible por el ordenador, o se suministra a través de una red. Cuando vaya a aplicarse la presente invención, el programa informático se instala en un ordenador y la presente invención se consigue cuando el programa informático se ejecuta en una unidad de control tal como una CPU.

20 Efecto de la invención

25 Según la presente invención, incluso en casos en los que tiene lugar una ausencia de correspondencia de iluminación y de color entre cámaras localmente, es posible reducir la predicción residual dado que los parámetros de corrección se determinan para cada objeto por separado y de forma local. En consecuencia, es posible conseguir una codificación y una decodificación altamente eficaces de imágenes multivista e imágenes en movimiento multivista.

30 Por otra parte, según la presente invención, dado que los parámetros de corrección se determinan de una forma que no necesita codificación/decodificación adicional, es posible reducir considerablemente la cantidad de código necesaria cuando se realiza esta codificación y decodificación de imágenes multivista e imágenes en movimiento multivista.

35 **Breve descripción de los dibujos**

40 La figura 1 es un diagrama de bloques que muestra un dispositivo de codificación de vídeo multivista según una primera realización de la presente invención.

La figura 2 es un organigrama que muestra el tratamiento ejecutado por el dispositivo de codificación de vídeo multivista según la primera realización de la presente invención.

45 La figura 3 es un organigrama que muestra detalles del tratamiento ejecutado por el dispositivo de codificación de vídeo multivista según la primera realización de la presente invención.

50 La figura 4 es un organigrama que muestra el tratamiento ejecutado por el dispositivo de codificación de vídeo multivista según la primera realización de la presente invención.

La figura 5 es un diagrama de bloques que muestra un dispositivo de decodificación de vídeo multivista según una segunda realización de la presente invención.

55 La figura 6 es un organigrama que muestra el tratamiento ejecutado por el dispositivo de decodificación de vídeo multivista según la segunda realización de la presente invención.

La figura 7 es un organigrama que muestra detalles del tratamiento ejecutado por el dispositivo de decodificación de vídeo multivista según la segunda realización de la presente invención.

60 La figura 8 es un diagrama de bloques que muestra una unidad de creación de parámetros de corrección en las realizaciones primera y segunda de la presente invención.

La figura 9 es una vista que muestra un modo de predicción con compensación de disparidad.

65 **Realizaciones para llevar a cabo la invención**

A continuación se describirá la presente invención en detalle con referencia a los dibujos que ilustran las realizaciones de la presente invención.

5 Debe observarse que en la siguiente descripción, al asociar la información de posición (en concreto, valores de coordenadas o índices que pueden asociarse con valores de coordenadas) rodeada por el símbolo [ ] para vídeo (es decir, tramas) y la información de profundidad, se muestran las señales de imagen y la información de profundidad (definida para cada píxel) de los objetos fotografiados en los píxeles en esa posición.

10 [1] Dispositivo de codificación de vídeo multivista según una primera realización de la presente invención

En la figura 1 se muestra la estructura de un dispositivo de codificación de vídeo multivista 100 según la primera realización de la presente invención.

15 Tal como se muestra en la figura 1, en el dispositivo de codificación de vídeo multivista 100 de la presente realización, una unidad de entrada de imágenes objeto de codificación 101 recibe entradas de imágenes (es decir, tramas) de un objeto o escena fotografiados por una primera cámara como un objeto de codificación. La memoria de imágenes objeto de codificación 102 almacena las tramas objeto de codificación introducidas desde la unidad de entrada de imágenes objeto de codificación 101. Las tramas objeto de codificación almacenadas son suministradas a una unidad de codificación de imagen 109.

20 Una unidad de entrada de imágenes de síntesis de vistas 103 recibe entradas de imágenes de síntesis de vistas para codificar imágenes objeto. Las imágenes de síntesis de vistas han sido generadas usando las imágenes ya codificadas en las que se fotografió el mismo objeto o escena por las segundas cámaras colocadas en posiciones diferentes a la de la primera cámara. La memoria de imágenes de síntesis de vistas 104 almacena las imágenes de síntesis de vistas introducidas desde la unidad de entrada de imágenes de síntesis de vistas 103. Las imágenes de síntesis de vistas almacenadas son suministradas a una unidad de creación de parámetros de corrección 107 y una unidad de corrección de imágenes de síntesis de vistas 108.

30 Una unidad de entrada de información de profundidad 105 recibe entradas de información de profundidad para tramas destinadas a codificación. La memoria de información de profundidad 106 almacena la información de profundidad introducida desde la unidad de entrada de información de profundidad 105. La información de profundidad almacenada es suministrada a la unidad de creación de parámetros de corrección 107.

35 La unidad de creación de parámetros de corrección 107 estima los parámetros de corrección usando imágenes de síntesis de vistas, información de profundidad e imágenes decodificadas usadas en zonas periféricas de una zona objeto de codificación, y usando la información de profundidad usada en la zona objeto de codificación. La unidad de corrección de imágenes de síntesis de vistas 108 corrige las imágenes de síntesis de vistas de la zona destinada a codificación usando los parámetros de corrección estimados.

40 La unidad de codificación de imagen 109 codifica señales de imagen de una zona objeto de codificación usando las imágenes corregidas de síntesis de vistas como señales de predicción. La unidad de decodificación de imagen 110 decodifica las señales de imagen codificadas. La memoria de imágenes decodificadas 111 almacena las imágenes decodificadas por la unidad de codificación de imagen 109. Las imágenes decodificadas almacenadas en la memoria de imágenes decodificadas 111 son suministradas a la unidad de creación de parámetros de corrección 107.

50 Tal como se muestra en la figura 8, la unidad de creación de parámetros de corrección 107 tiene una unidad de determinación de objetos 107a a la que se suministra la información de profundidad desde la memoria de información de profundidad 106, y una unidad de determinación de profundidad representativa 107b y una unidad de determinación de grupo de píxeles de muestra 107c que están conectadas en esta secuencia con el lado de flujo descendente de la unidad de determinación de objetos 107a. Una unidad de selección de modelo de corrección 107d y una unidad de estimación de parámetros de corrección 107e están conectadas en esta secuencia con el lado de flujo descendente de la unidad de determinación de grupo de píxeles de muestra 107c. Las imágenes decodificadas de la memoria de imágenes decodificadas 111 y las imágenes de síntesis de vistas de la memoria de imágenes de síntesis de vistas 104 son suministradas a la unidad de estimación de parámetros de corrección 107e, y los parámetros de corrección estimados usando las imágenes decodificadas suministradas y las imágenes de síntesis de vistas son suministrados a la unidad de corrección de imágenes de síntesis de vistas 108.

60 La figura 2 muestra el flujo de tratamiento ejecutado por el dispositivo de codificación de vídeo multivista 100 de la presente realización que se construye de la manera que se ha descrito anteriormente.

A continuación se describirá en detalle el tratamiento ejecutado por el dispositivo de codificación de vídeo multivista 100 de la presente realización de acuerdo con este flujo de tratamiento.

65 En primer lugar, se introduce una trama *Org* que está destinada a codificación por la unidad de entrada de imágenes objeto de codificación 101, y se almacena en la memoria de imágenes objeto de codificación 102. Se introduce una

imagen de síntesis de vistas *Sint* para la trama objeto de codificación *Org* por medio de la unidad de entrada de imágenes de síntesis de vistas 103, y se almacena en la memoria de imágenes de síntesis de vistas 104. Se introduce la información de profundidad *Profundidad* para la trama objeto de codificación *Org* por medio de la unidad de entrada de información de profundidad 105, y se almacena en la memoria de información de profundidad 106 [A1].

La imagen de síntesis de vistas y la información de profundidad que se introducen aquí son las mismas que las obtenidas en el dispositivo de decodificación. El motivo de ello es que, al usar la misma información que la información obtenida en el dispositivo de decodificación, puede suprimirse la generación de ruidos de codificación tales como la deriva. Sin embargo, si la generación de dicho ruido de codificación es admisible, también es posible entonces que se introduzca la información de precodificación original.

Debe observarse que la información de profundidad se proporciona desde el exterior del dispositivo de codificación de vídeo multivista 100, sin embargo, tal como se describe en el Documento no de patente 3, también es posible obtener la información de profundidad estimándola a partir de las tramas ya codificadas en otras cámaras. En consecuencia, no es esencial que la información de profundidad sea transmitida desde el lado de emisión al lado de recepción. La imagen de síntesis de vistas se genera usando tramas ya codificadas en cámaras distintas a la primera cámara y la información de profundidad.

A continuación, la trama objeto de codificación se divide en una pluralidad de zonas objeto de codificación, y la señal de imagen de la trama objeto de codificación se codifica por medio de la unidad de codificación de imagen 109 con corrección de la imagen de síntesis de vistas para cada una de estas zonas [A2-A15].

Es decir, si un índice del bloque de tratamiento de codificación se expresa como *blq*, y si el número total de todos los bloques de tratamiento de codificación se expresa como *numBlqs*, entonces después de que *blq* se haya inicializado a 0 [A2], se repiten las etapas siguientes [A3-A13] añadiéndose 1 a *blq* [A14] hasta que *blq* llegue a *numBlqs* [A15].

En el tratamiento que se repite para cada bloque de tratamiento de codificación, el dispositivo de codificación 100 primero determina un grupo de píxeles  $N_{blq}$  de zonas periféricas ya codificadas al bloque *blq* (es decir, la zona objeto de codificación) [A3].

Para estas zonas periféricas pueden usarse varios tipos de unidad tales como bloques de tratamiento de codificación que son adyacentes al bloque *blq*, o píxeles adyacentes o similares. Aunque puede usarse una diversidad de definiciones para las zonas periféricas, es necesario que se use la misma definición que se usó en el lado de decodificación.

A continuación, la unidad de determinación de objetos 107a de la unidad de creación de parámetros de corrección 107 agrupa píxeles en el bloque *blq* usando la información de profundidad proporcionada para cada píxel como referencia [etapa de determinación de objeto A4].

El índice de cada grupo resultante de lo anterior se expresa como *obj*, el número de grupos se expresa como *numObjs*, y los píxeles pertenecientes al grupo *obj* se expresan como  $C_{obj}$ .

Para esta operación de agrupación puede usarse cualquier método, sin embargo, es necesario que se use el mismo método que el usado en el lado de decodificación. El método de agrupación más sencillo es uno en el que se calcula la variación de la información de profundidad para píxeles en el bloque *blq*, y si el valor de la misma es igual o inferior a un valor umbral entonces todos los píxeles se establecen como un grupo, mientras que si el valor es mayor que el valor umbral, entonces se fija el promedio de los valores de profundidad como un límite y los píxeles se dividen en dos grupos. En este caso, una vez que las variaciones de la información de profundidad de los píxeles en cada grupo se hacen iguales o inferiores a un valor umbral determinado, se pone fin a la agrupación.

Un método de agrupación más complejo es aquél en el que, al inicio, se considera que cada píxel va a formar un grupo. A continuación, cuando se fusionan conjuntamente dos grupos cualesquiera, la fusión se lleva a cabo en secuencia a partir de los grupos que tienen el incremento más reducido en la variación de la información de profundidad de los píxeles dentro del grupo. En este caso, si un grupo que tiene una variación que es superior a un valor umbral determinado termina por crearse sin que importe cuáles son los grupos en particular que se fusionan entre sí, entonces se pone fin a la agrupación.

Debe observarse que dado que la cantidad de cálculo aumenta cuando se incrementa el número de grupos, también es posible definir con antelación un número máximo de grupos. Al poner fin a la operación de agrupación una vez que el número de grupos alcanza este valor máximo, o al definir por adelantado un número mínimo de píxeles que estará contenido dentro de un único grupo de manera que no se creen grupos pequeños, es posible prevenir el crecimiento del número de grupos. Si el tamaño de bloque es 16 x 16 píxeles, que es un tamaño de bloque típico de una unidad para codificación de imagen, entonces en general es raro que se fotografíe un gran número de objetos en el mismo bloque. En consecuencia, puede conseguirse una agrupación altamente precisa incluso si el número máximo de grupos es dos.



Una vez que ha terminado la operación de agrupación, se crea una imagen de predicción *Pred* corrigiendo la imagen de síntesis de vistas para cada píxel de cada grupo [A5-A11].

5 Es decir, después de que el índice de grupo *obj* se ha inicializado a 0 [A5], se llevan a cabo las etapas [A6-A8] en las que, con un incremento de *obj* en 1 [A10], se estiman los parámetros de corrección con el fin de corregir la imagen de síntesis de vistas en la unidad de grupo de creación de parámetros de corrección 107 hasta que *obj* alcanza el valor *numObjs* [A11]. Junto con las etapas, en la unidad de corrección de imágenes de síntesis de vistas 108, se lleva a cabo una etapa [A9] para crear una imagen de predicción corrigiendo las imágenes de síntesis de vistas para píxeles  
10 contenidos en el grupo *obj* mediante el uso de los parámetros de predicción estimados.

En el tratamiento de estimación de parámetros de corrección, en primer lugar, basándose en la información de profundidad, se determina un grupo de píxeles  $N_{blq,obj}$  (es decir, un grupo de píxeles de muestra) en el que se fotografía el mismo objeto que en el grupo *obj* y que se incluye en el grupo de píxeles  $N_{blq}$  en las zonas ya  
15 codificadas periféricas al bloque *blq* (es decir, la zona objeto de codificación) [etapa de determinación de profundidad representativa y etapa de determinación de grupo de píxeles de muestra A6].

Específicamente, en este tratamiento, en primer lugar, en la unidad de determinación de profundidad representativa 107b, se determina un valor de profundidad representativo  $D_{obj}$  para el grupo *obj* (etapa de determinación de  
20 profundidad representativa). Puede usarse cualquier método para determinar el valor de profundidad representativo  $D_{obj}$  siempre que sea posible determinar un único valor de profundidad a partir de la información de profundidad en los píxeles  $C_{obj}$  dentro del grupo *obj*. Sin embargo, es necesario usar el mismo tratamiento que el que se usó en el lado del decodificador. Por ejemplo, puede usarse un valor promedio o un valor medio o similares. En el caso de que se usen valores promedio, el valor de profundidad representativo  $D_{obj}$  puede expresarse mediante la Fórmula (1)  
25 siguiente. Debe observarse que  $\| \cdot \|$  muestra el número de píxeles en el conjunto.

$$D_{obj} = \frac{1}{\|C_{obj}\|} \sum_{p \in C_{obj}} Profundidad[p] \quad (1)$$

Debe observarse que algunos métodos de agrupación, en los que se dividen los píxeles en varios grupos usando la información de profundidad como referencia en la etapa de determinación de objeto A4, determinan los píxeles y/o la  
30 información de profundidad que representa cada grupo como parte de su proceso. Si se usa este tipo de método, puede usarse la información de profundidad para el píxel usado como representativo del grupo en ese momento y también la información de profundidad que muestra el grupo como valores de profundidad representativos. En este caso, el tratamiento de la etapa de determinación de profundidad representativa se lleva a cabo al mismo tiempo en la etapa de determinación de objeto A4. La propagación de afinidad es uno de los métodos más conocidos en el que la decisión del valor representativo del grupo y el tratamiento de agrupación se realizan de forma simultánea. Los  
35 detalles de este método se describen en "Clustering by Passing Messages Between Data Points", B. J. Frey y D. Dueck, Science 2007, Vol. 315(5814): pág. 972-976.

Cuando el valor de profundidad representativo  $D_{obj}$  se ha determinado de esta manera, a continuación, en la unidad de determinación de grupo de píxeles de muestra 107c, se determina un grupo de píxeles  $N_{blq,obj}$  (es decir, un grupo de píxeles de muestra) en el que se fotografía el mismo objeto que en el grupo *obj* y que es un subconjunto del grupo  $N_{blq}$  (etapa de determinación de grupo de píxeles de muestra). Tal como se muestra mediante la fórmula (2) siguiente, el grupo  $N_{blq,obj}$  puede definirse seleccionando los píxeles en los que la diferencia absoluta entre su valor de profundidad y el valor de profundidad representativo es menor que el umbral predefinido *umbMismoObj*.  
45

$$N_{blq,obj} = \{p \mid p \in N_{blq} \cap |D_{obj} - profundidad[p]| < umbMismoObj\} \quad (2)$$

Después de haber determinado el grupo  $N_{blq,obj}$ , se selecciona un modelo de corrección mediante la unidad de selección de modelo de corrección 107d a partir de modelos de corrección predefinidos basándose en el número de  
50 píxeles en este grupo [etapa de selección de modelo de corrección A7].

Es posible preparar varios modelos de corrección, sin embargo, es necesario que se prepare también el mismo modelo de corrección en el lado de decodificación, y asimismo es necesario usar los mismos criterios para decidir un modo de corrección que se usará. Por otra parte, si sólo se ha preparado un modelo de corrección, entonces puede omitirse esta etapa.  
55

El tratamiento para crear una imagen de predicción corrigiendo una imagen de síntesis de vistas que se lleva a cabo en la etapa 9A (descrita más adelante) puede expresarse como una función que tiene una imagen de síntesis de vistas como entrada y que produce en salida una imagen de predicción. El término modelo de corrección se refiere a un modelo de la función usado en este momento. Puede tratarse, por ejemplo, de una corrección basada en

*desplazamiento*, una corrección que emplea una función lineal o un tratamiento de filtro lineal bidimensional que tiene una longitud de derivación  $k$ . Cuando el píxel que se va a corregir se denota como  $p$ , lo anterior puede expresarse mediante la fórmula (3) a la fórmula (5) siguientes, respectivamente.

5

$$Pred[p] = Sint[p] + desplazamiento \quad (3)$$

$$Pred[p] = \alpha \cdot Sint[p] + \beta \quad (4)$$

$$Pred[p] = \sum_{i=-k}^{i=k} \sum_{j=-k}^{j=k} (F_{i,j} \cdot Sint[p + (i, j)^r]) + o \quad (5)$$

10 La corrección basada en *desplazamiento* y la corrección que emplea una función lineal son ejemplos típicos de corrección que se basan en tratamiento de filtro lineal. Debe observarse que no es necesario que el tratamiento de corrección sea un tratamiento lineal y también es posible usar un filtro no lineal siempre que sea posible la estimación del parámetro de corrección. Un ejemplo de un método de corrección no lineal es la corrección gamma. La corrección gamma puede expresarse mediante la fórmula (6) siguiente.

$$Pred[p] = (Sint[p] - a)^{\frac{1}{\gamma}} + b \quad (6)$$

15 En los ejemplos de estos modelos de corrección, *desplazamiento*,  $(\alpha, \beta)$ ,  $(\{F_{i,j}\}, o)$  y  $(\gamma, a, b)$  forman respectivamente los parámetros de corrección. Es decir, el número de parámetros de corrección cambia dependiendo del modelo de corrección.

20 Cuando el número de parámetros de corrección aumenta, se hace posible realizar una corrección más precisa, sin embargo, con el fin de decidir estos parámetros de corrección, es necesario que el número de muestras sea igual o superior al número de parámetros de corrección. Tal como se describe más adelante, dado que esta muestra forma una imagen de síntesis de vistas y una imagen decodificada de los píxeles de muestra contenidos en el grupo  $N_{blq,obj}$ , descrito anteriormente decidiendo el modelo de corrección de acuerdo con el número de píxeles del grupo  $N_{blq,obj}$ , se hace posible realizar correcciones precisas. Debe observarse que usando el máximo número de muestras posible en la estimación de los parámetros de corrección se facilita una estimación más robusta de los parámetros de corrección. En consecuencia, es deseable seleccionar el modelo de corrección de manera que asegure una robustez suficiente.

30 Después de haber seleccionado un único modelo de corrección, en la unidad de estimación de parámetros de corrección 107e, se toman como muestras una imagen de síntesis de vistas *Sint* y una imagen decodificada *Dec* para los píxeles de muestra contenidos en el grupo  $N_{blq,obj}$ , y se estiman los parámetros de corrección en el modelo de corrección seleccionado [etapa de estimación de parámetros de corrección A8].

35 En la estimación de los parámetros de corrección realizada en este caso, se supone que el valor de precorrección y el valor ideal corregido son la imagen de síntesis de vistas y la imagen decodificada, respectivamente. A continuación se determinan los parámetros de corrección de manera que se reduzca al mínimo la suma de error, para todos los píxeles en el grupo  $N_{blq,obj}$  entre el valor ideal y el valor que es consecuencia de la corrección usando el modelo de corrección de cada píxel.

40 Por ejemplo, si se usa tratamiento lineal para la corrección, entonces los parámetros de corrección pueden decidirse usando el método de los mínimos cuadrados. Es decir, cuando  $M$  denota la corrección, la estimación puede realizarse mediante la minimización de un valor expresado por la fórmula (7) siguiente:

$$\sum_{p \in N_{blq,obj}} (Dec[p] - M(Sint[p]))^2 \quad (7)$$

45 Es decir, es posible determinar los parámetros de corrección resolviendo ecuaciones simultáneas en las que la derivada parcial de la fórmula (7) con respecto a cada parámetro de corrección es igual a 0.

50 Una vez que se han estimado los parámetros de corrección, en la unidad de corrección de imágenes de síntesis de vistas 108, mediante la corrección de la imagen de síntesis de vistas *Sint* para el grupo *obj* del bloque *blq* usando los parámetros de corrección, se crea una imagen de predicción *Pred* para el grupo *obj* del bloque *blq* [etapa de corrección de imágenes de síntesis de vistas A9].

Específicamente, tal como se muestra en el flujo de tratamiento en la figura 3, el tratamiento para crear esta imagen de predicción *Pred* se realiza para cada píxel. En este caso, en el flujo de tratamiento de la figura 3, *pix* indica información de identificación de píxel, y *numPix<sub>blq,obj</sub>* indica el número de píxeles en el grupo *obj del* bloque *blq*.

- 5 Por ejemplo, en el caso de corrección que se realiza usando valores de desplazamiento, la imagen de predicción *Pred* se crea de acuerdo con la fórmula (3) descrita anteriormente.

Para la creación de esta imagen de predicción *Pred*, se describe un ejemplo en el que la corrección de la fórmula (3) que emplea valores de desplazamiento se realiza como el método de corrección (es decir, el modelo de corrección).  
 10 Tal como se muestra en la figura 4, en la etapa de estimación de parámetros de corrección A8, estimando el desplazamiento cuando los valores de píxel de las imágenes de síntesis de vistas para el mismo objeto que está presente en zonas periféricas que ya han sido codificadas se toman como *Entrada*, y los valores de píxel de imágenes decodificadas de ese objeto se toman como *Salida*, se construye una ecuación de conversión para píxeles como modelo de corrección. A continuación, en la etapa S9, se realiza un tratamiento para generar la imagen de  
 15 predicción para el grupo *obj del* bloque *blq* sustituyendo los valores de píxel de la imagen de síntesis de vistas en el grupo *obj del* bloque *blq* en *Entrada* de la ecuación de conversión construida.

Después de que haya terminado la creación de la imagen de predicción para el bloque *blq*, en la unidad de codificación de imagen 109, se realiza codificación de la trama objeto de codificación *Org* para el bloque *blq* [etapa de codificación de imagen A12] con la imagen de predicción *Pred* creada en la etapa A9 usada para la señal de predicción.  
 20 predicción.

En esta etapa de codificación A12, no existen restricciones en las que pueda usarse el método de codificación, sin embargo, en un método de codificación típico tal como H.264, la codificación se lleva a cabo aplicando codificación DCT-cuantificación-binarización-entropía en la diferencia entre *Org* y *Pred*.  
 25

El tren de bits resultante de la codificación forma la salida del dispositivo de codificación de vídeo multivista 100. Por otra parte, el tren de bits resultante de la codificación es decodificado por la unidad de decodificación de imagen 110 para cada bloque, y una imagen decodificada *Dec* que es el resultado obtenido de la decodificación se almacena en la memoria de imágenes decodificadas 111 que se usará para estimar parámetros de corrección en otros bloques [etapa de decodificación de imagen A13].  
 30

De esta manera, incluso en casos en los que se producen ausencias de correspondencia de iluminación o de color entre cámaras de una forma localizada de acuerdo con el objeto, el dispositivo de codificación de vídeo multivista 100 que se construye de la manera mostrada en la figura 1 determina los parámetros de corrección para compensar estas ausencias de correspondencia de una forma localizada de acuerdo con el objeto con el fin de hacer posible la codificación de vídeo multivista con un alto nivel de eficiencia. Por otra parte, con el fin de evitar cualquier incremento en la cantidad de código, este dispositivo de codificación de vídeo multivista 100 determina estos parámetros de corrección de tal manera que es innecesario codificar/decodificar los parámetros de corrección obtenidos mientras el vídeo multivista se codifica con la corrección local de iluminación y de color.  
 35  
 40

En la presente realización, se ha descrito un caso en el que la señal de imagen del bloque *blq* se codifica con el uso de imágenes de síntesis de vistas que son esenciales, sin embargo, puede usarse también la predicción que usa imágenes de síntesis de vistas descritas en la presente realización como un modo de predicción de entre una pluralidad de modos de predicción.  
 45

#### [2] Dispositivo de decodificación de vídeo multivista según una segunda realización de la presente invención

En la figura 5 se muestra la estructura de un dispositivo de decodificación de vídeo multivista 200 según la segunda realización de la presente invención.  
 50

Tal como se muestra en la figura 5, en el dispositivo de decodificación de vídeo multivista 200 de la presente realización, una unidad de entrada de datos codificados 201 recibe datos codificados de imagen (es decir, tramas) de un objeto o escena fotografiado por una primera cámara como un objeto de decodificación. La memoria de datos codificados 202 almacena datos codificados introducidos desde la unidad de entrada de datos codificados 201. Los datos codificados almacenados son suministrados a una unidad de decodificación de imagen 209.  
 55

Una unidad de entrada de imágenes de síntesis de vistas 203 recibe entradas de imágenes de síntesis de vistas para decodificación de imágenes objeto. Las imágenes de síntesis de vistas han sido generadas usando las imágenes ya decodificadas en las que se fotografió el mismo objeto o escena por segundas cámaras colocadas en diferentes posiciones a la de la primera cámara. La memoria de imágenes de síntesis de vistas 204 almacena imágenes de síntesis de vistas introducidas desde la unidad de entrada de imágenes de síntesis de vistas 203. Las imágenes de síntesis de vistas almacenadas son suministradas a una unidad de creación de parámetros de corrección 207 y una unidad de corrección de imágenes de síntesis de vistas 208.  
 60  
 65

Una unidad de entrada de información de profundidad 205 recibe entradas de información de profundidad para

tramas destinadas a decodificación. La memoria de información de profundidad 206 almacena información de profundidad introducida desde la unidad de entrada de información de profundidad 205. La información de profundidad almacenada es suministrada a la unidad de creación de parámetros de corrección 207.

5 La unidad de creación de parámetros de corrección 207 estima los parámetros de corrección usando imágenes de síntesis de vistas, información de profundidad e imágenes decodificadas usadas en zonas periféricas de una zona objeto de decodificación, y usando la información de profundidad usada en la zona objeto de decodificación. La unidad de corrección de imágenes de síntesis de vistas 208 corrige imágenes de síntesis de vistas de la zona destinada a decodificación usando los parámetros de corrección estimados.

10 La unidad de decodificación de imagen 209 decodifica señales de imagen de la zona objeto de decodificación usando las imágenes de síntesis de vistas corregidas como señales de predicción. (10) La memoria de imágenes decodificadas 210 almacena imágenes decodificadas por la unidad de decodificación de imagen 209.

15 Tal como se muestra en la figura 8, la unidad de creación de parámetros de corrección 207 tiene una unidad de determinación de objetos 207a a la que se suministra información de profundidad desde la memoria de información de profundidad 206, y una unidad de determinación de profundidad representativa 207b y una unidad de determinación de grupo de píxeles de muestra 207c que están conectadas en esta secuencia con el lado de flujo descendente de la unidad de determinación de objetos 207a. Una unidad de selección de modelo de corrección 207d y una unidad de estimación de parámetros de corrección 207e están conectadas en esta secuencia con el lado de flujo descendente de la unidad de determinación de grupo de píxeles de muestra 207c. Las imágenes decodificadas de la memoria de imágenes decodificadas 210 y las imágenes de síntesis de vistas de la memoria de imágenes de síntesis de vistas 204 son suministradas a la unidad de estimación de parámetros de corrección 207e, y los parámetros de corrección estimados usando las imágenes decodificadas y las imágenes de síntesis de vistas 25 suministradas son suministrados a la unidad de corrección de imágenes de síntesis de vistas 208.

La figura 6 muestra el flujo de tratamiento ejecutado por el dispositivo de decodificación de vídeo multivista 200 de la segunda realización que está estructurado de la manera que se ha descrito anteriormente.

30 A continuación se describirá en detalle el tratamiento ejecutado por el dispositivo de decodificación de vídeo multivista 200 de la presente realización de acuerdo con este flujo de tratamiento.

En primer lugar, los datos codificados para una trama que está destinada a decodificación son introducidos por la unidad de introducción de datos de decodificación 201, y se almacena en la memoria de datos codificados 202. Una imagen de síntesis de vistas *Sint* para el objeto de decodificación trama es introducida por la unidad de entrada de imágenes de síntesis de vistas 203, y se almacena en la memoria de imágenes de síntesis de vistas 204. La información de profundidad *Profundidad* para la trama objeto de decodificación es introducida por la unidad de entrada de información de profundidad 205, y se almacena en la memoria de información de profundidad 206 [B1].

40 La imagen de síntesis de vistas y la información de profundidad que se introducen en este caso son las mismas que las obtenidas por el dispositivo de codificación. El motivo de ello es que, usando la misma información que la información obtenida en el dispositivo de codificación, puede suprimirse la generación de ruidos de codificación como la deriva. Sin embargo, si la generación de dicho ruido de codificación es admisible, entonces también es posible que se introduzca una información diferente de la obtenida por el dispositivo de codificación.

45 Debe observarse que la información de profundidad es proporcionada desde el exterior del dispositivo de decodificación de vídeo multivista 200, sin embargo, tal como se describe en el Documento no de patente 3, también es posible obtener la información de profundidad estimándola a partir de tramas ya decodificadas en otras cámaras. En consecuencia, no es esencial que la información de profundidad sea transmitida desde el lado de emisión al lado de recepción. La imagen de síntesis de vistas se genera usando tramas ya decodificadas en cámaras distintas a la primera cámara e información de profundidad.

50 A continuación, el objeto de decodificación trama se divide en una pluralidad de zonas objeto de decodificación, y la señal de imagen de la trama objeto de decodificación es decodificada por la unidad de decodificación de imagen 209 corrigiendo la imagen de síntesis de vistas para cada una de estas zonas [B2-B14].

55 Es decir, si un índice del bloque de tratamiento de decodificación se expresa como *blq*, y si el número total de todos los bloques de tratamiento de decodificación se expresa como *numBlqs*, entonces después de que *blq* se haya inicializado a 0 [B2], se repiten las etapas siguientes [B3-B12] añadiendo 1 a *blq* [B13] hasta que *blq* llegue a *numBlqs* [B14].

60 En el tratamiento que se repite para cada bloque de tratamiento de decodificación, el dispositivo de decodificación 200 determina primero un grupo de píxeles  $N_{blq}$  de zonas ya decodificadas periféricas al bloque *blq* (es decir, la zona objeto de decodificación) [B3].

65 Pueden usarse varios tipos de unidad para estas zonas periféricas tales como bloques de tratamiento de

decodificación que son adyacentes al bloque *blq*, o píxeles adyacentes o similares. Aunque puede usarse una variedad de definiciones para las zonas periféricas, es necesario que se use la misma definición que se usó en el lado de codificación.

5 A continuación, la unidad de determinación de objetos 207a de la unidad de creación de parámetros de corrección 207 agrupa píxeles en el bloque *blq* usando la información de profundidad proporcionada para cada píxel como referencia [etapa de determinación de objeto B4].

10 El índice de cada grupo resultante de ello se expresa como *obj*, el número de grupos se expresa como *numObjs* y los píxeles pertenecientes al grupo *obj* se expresan como  $C_{obj}$ . El tratamiento realizado en este caso es el mismo que el llevado a cabo en la etapa de determinación de objeto A4 de la primera realización.

15 Una vez que ha terminado la agrupación, se crea una imagen de predicción *Pred* corrigiendo la imagen de síntesis de vistas para cada píxel de cada grupo [B5-B11].

20 Es decir, después de que se ha inicializado el índice del grupo *obj* a 0 [B5], se realizan las etapas [B6-B8] en las que, con incremento de *obj* en 1 [B10], se estiman los parámetros de corrección con el fin de corregir la imagen de síntesis de vistas en la unidad de creación de parámetros de corrección 207 hasta que *obj* llega a *numObjs* [B11]. Junto con las etapas, en la unidad de corrección de imágenes de síntesis de vistas 208, se lleva a cabo una etapa [B9] para crear una imagen de predicción corrigiendo las imágenes de síntesis de vistas para píxeles contenidos en el grupo *obj* usando los parámetros de predicción estimados.

25 El tratamiento en esta etapa B9 es el mismo que en la etapa A9 de la primera realización y, tal como se muestra en el flujo de tratamiento en la figura 7, se realiza para cada píxel. En este caso, en el flujo de tratamiento en la figura 7, *pix* indica información de identificación de píxel, y  $numPix_{blq,obj}$  indica el número de píxeles en el grupo *obj* del bloque *blq*. Las etapas de estimación de parámetros de corrección [B6-B8] que son realizadas por la unidad de determinación de profundidad representativa 207b, la unidad de determinación de grupo de píxeles de muestra 207c, la unidad de selección de modelo de corrección 207d y la unidad de estimación de parámetros de corrección 207e son las mismas que en las etapas A6-A8 de la primera realización.

30 Después de que ha terminado la creación de la imagen de predicción para el bloque *blq*, en la unidad de decodificación de imagen 209, la imagen objeto de decodificación *Dec* se decodifica para el bloque *blq* [B12] con la imagen de predicción *Pred* creada en la etapa B9 usándose para la señal de predicción.

35 Es necesario que el tratamiento de decodificación realizado en este caso se corresponda con el método usado cuando se crearon los datos codificados. Por ejemplo, si la codificación se realizó usando H.264, entonces el tratamiento de decodificación se realiza añadiendo la señal de predicción a la señal residual decodificada aplicando decodificación de entropía, multiplexación de valores, cuantificación inversa y DCT inversa.

40 La imagen decodificada resultante de la decodificación forma la salida del dispositivo de decodificación multivista 200, y se almacena en la memoria de imágenes decodificadas 210 que se usará para estimar los parámetros de corrección en otros bloques.

45 De esta manera, el dispositivo de decodificación de vídeo multivista 200 que está estructurado de la forma descrita anteriormente tal como se muestra en la figura 5 decodifica los datos codificados de la decodificación multivista de vídeo creada por el dispositivo de codificación de vídeo multivista 100 tal como se muestra en la figura 1.

50 En la presente realización, se ha descrito un caso en el que el bloque *blq* se codifica con el uso de imágenes de síntesis de vistas que son esenciales. Incluso en casos en los que los datos codificados se decodifican empleando como uno de una pluralidad de modos de predicción existentes un modo de predicción que hace uso de imágenes de síntesis de vistas, las imágenes sólo se decodifican de acuerdo con el flujo de tratamiento descrito anteriormente cuando este modo de predicción se está usando, aun cuando se estén usando otros modos de predicción, las imágenes pueden decodificarse usando un método de decodificación convencional que corresponde al del modo de predicción (es decir, a uno de estos otros modos de predicción).

55 En las realizaciones primera y segunda descritas anteriormente, se realizan correcciones en todos los bloques, sin embargo, también es posible emplear una estructura en la que se elige si se realiza o no una corrección codificando un bit de información de bandera en cada bloque.

60 Por otra parte, también existe un método en el que se mide la fiabilidad de los parámetros de corrección, y a continuación se elige realizar o no una corrección basándose en el grado de la fiabilidad en lugar de codificar un bit de información de bandera.

65 Específicamente, es posible, después de haber determinado los parámetros de corrección en la etapa A8 y la etapa B8, calcular un valor que expresa la viabilidad y la eficacia de la corrección, por ejemplo, usando las fórmulas (8) a (10) siguientes. Las unidades de corrección de imágenes de síntesis de vistas 108 y 208 corrigen la imagen de

síntesis de vistas en la etapa A9 y la etapa B9 y la envían como imágenes de predicción sólo cuando el valor calculado es mayor que un valor umbral predefinido, mientras que en todos los demás casos, la imagen de síntesis de vistas sin correcciones se envía como imagen de predicción.

$$R = \sum_{p \in N_{blq,obj}} |Dec[p] - Sint[p]| - |Dec[p] - Pred[p]| \quad (8)$$

$$R = \sum_{p \in N_{blq,obj}} |Dec[p] - Sint[p]|^2 - |Dec[p] - Pred[p]|^2 \quad (9)$$

$$R = 1 - \frac{\sum_{p \in N_{blq,obj}} (Dec[p] - Pred[p])^2}{\sum_{p \in N_{blq,obj}} \left( Dec[p] - \frac{1}{\|N_{blq,obj}\|} \sum_{q \in N_{blq,obj}} (Dec[q]) \right)^2} \quad (10)$$

5 El primer término de la fórmula (8) muestra la suma de diferencias absolutas entre la imagen decodificada *Dec* y la imagen de síntesis de vistas *Sint*, mientras que el segundo término muestra la suma de diferencias absolutas entre la imagen decodificada *Dec* y la imagen de predicción *Pred*. Es decir, la fórmula (8) muestra cuántas veces se ha reducido la suma de diferencias absolutas entre el valor verdadero y el valor de predicción mediante la corrección.  
 10 Por otra parte, el primer término de la fórmula (9) muestra la suma de las diferencias al cuadrado entre la imagen decodificada *Dec* y la imagen de síntesis de vistas *Sint*, mientras que el segundo término muestra la suma de las diferencias al cuadrado entre la imagen decodificada *Dec* y la imagen de predicción *Pred*. Es decir, la fórmula (9) muestra la magnitud en que se ha reducido la suma de las diferencias al cuadrado entre el valor verdadero y el valor de predicción mediante la corrección. La fórmula (10) muestra la viabilidad del modelo de corrección en las  
 15 muestras. En este caso, dado que todos estos valores se determinan usando valores no procedentes del bloque objeto de tratamiento, sino de zonas periféricas del mismo, sólo es necesario codificar aquellos valores umbral que se usan universalmente y después proporcionar estos valores al lado de decodificación.

20 Debe observarse que en la presente realización, se ha descrito el tratamiento para codificar o decodificar una trama de una cámara, sin embargo, repitiendo este tratamiento para cada trama, es posible conseguir la codificación o decodificación de un vídeo multivista. Además, repitiendo este tratamiento para cada cámara, es posible conseguir la codificación o decodificación del vídeo multivista de una pluralidad de cámaras.

25 El tratamiento descrito anteriormente puede conseguirse también mediante un ordenador y un programa de software, y dicho programa puede ser suministrado grabándolo en un medio de grabación legible por el ordenador, o puede ser suministrado a través de una red.

30 Por otra parte, en las realizaciones descritas anteriormente, la descripción se centra en un dispositivo de codificación de vídeo multivista y un dispositivo de decodificación de vídeo multivista, sin embargo, el método de codificación de vídeo multivista de la presente invención puede conseguirse por medio de etapas que corresponden a las operaciones de cada parte de este dispositivo de codificación de vídeo multivista. De la misma forma, el método de decodificación de vídeo multivista de la presente invención puede conseguirse por medio de etapas que corresponden a las operaciones de cada parte de este dispositivo de decodificación de vídeo multivista.

35 Si bien anteriormente se han descrito e ilustrado realizaciones preferidas de la invención, debe entenderse que estas realizaciones constituyen ejemplos de la invención y no deben considerarse como limitativas. Pueden realizarse añadidos, omisiones, sustituciones y otras modificaciones sin apartarse del alcance de la presente invención. En consecuencia, no debe considerarse que la invención está limitada por la descripción precedente y está limitada únicamente por el alcance de las reivindicaciones adjuntas.

40 **Aplicabilidad industrial**

La presente invención puede aplicarse a métodos de codificación de imagen multivista. Aplicando la presente invención, es posible conseguir una codificación y decodificación altamente eficaz de imagen multivista y vídeo multivista en las que existen ausencias de correspondencia de iluminación o de color locales y dependientes del objeto entre cámaras, y que es capaz de reducir enormemente la cantidad de código necesaria en los casos en los que se emplea esta nueva invención.

50 **Símbolos de referencia**

100 Dispositivo de codificación de vídeo multivista

	101	Unidad de entrada de imágenes objeto de codificación
	102	Memoria de imágenes objeto de codificación
	103	Unidad de entrada de imágenes de síntesis de vistas
	104	Memoria de imágenes de síntesis de vistas
5	105	Unidad de entrada de información de profundidad
	106	Memoria de información de profundidad
	107	Unidad de creación de parámetros de corrección
	108	Unidad de corrección de imágenes de síntesis de vistas
	109	Unidad de codificación de imagen
10	110	Unidad de decodificación de imagen
	111	Memoria de imágenes decodificadas
	200	Dispositivo de decodificación de vídeo multivista
	201	Unidad de entrada de datos codificados
	202	Memoria de datos codificados
15	203	Unidad de entrada de imágenes de síntesis de vistas
	204	Memoria de imágenes de síntesis de vistas
	205	Unidad de entrada de información de profundidad
	206	Memoria de información de profundidad
	207	Unidad de creación de parámetros de corrección
20	208	Unidad de corrección de imágenes de síntesis de vistas
	209	Unidad de decodificación de imagen
	210	Memoria de imágenes decodificadas

**REIVINDICACIONES**

1. Un método de codificación de imágenes multivista en el que una imagen de entrada de un objeto que es fotografiado por una primera cámara se divide en una pluralidad de zonas objeto de codificación y, usando una imagen de síntesis de vistas que se sintetiza a partir de la información de profundidad para la imagen de entrada y a partir de una imagen ya codificada del mismo objeto fotografiado por una segunda cámara que está situada en una posición diferente a la de la primera cámara, se realiza codificación predictiva para cada una de las zonas objeto de codificación, que comprende una etapa de codificación de imagen (A12) durante la cual, usando la imagen de síntesis de vistas, se codifica una señal de imagen de la imagen objeto de codificación de manera que se crean datos codificados, siendo los datos codificados decodificados de manera que se crea una imagen decodificada para la zona objeto de codificación en una etapa de decodificación de imagen (A13); estando el método de codificación de imágenes multivista caracterizado porque comprende:
- una etapa de determinación de profundidad representativa (A6) durante la cual se establece la información de profundidad representativa para un objeto fotografiado en la zona objeto de codificación,
  - una etapa de determinación de grupo de píxeles de muestra (A6) durante la cual en la imagen de entrada se determina un grupo de píxeles en el que se ha fotografiado el mismo objeto que en la zona objeto de codificación seleccionando píxeles de una zona ya codificada que es adyacente a la zona objeto de codificación, en la que la diferencia absoluta entre un valor de profundidad de la misma y el valor de profundidad representativo es menor que un umbral predefinido y el grupo de píxeles se establece como un grupo de píxeles de muestra,
  - una etapa de estimación de parámetros de corrección (A8) durante la cual, basándose en la imagen de síntesis de vistas para el grupo de píxeles de muestra y en una imagen decodificada que ha sido decodificada para el grupo de píxeles de muestra, se estiman los parámetros de corrección para corregir al menos una de las ausencias de correspondencia de iluminación o de color de manera que se reduzca al mínimo la suma de valores que representan un error entre los valores de píxel de la imagen decodificada para el grupo de píxeles de muestra y los valores de píxel de la imagen de síntesis de vistas para el grupo de píxeles de muestra que son corregidos usando un modelo de corrección para el grupo de píxeles de muestra, y
  - una etapa de corrección de imágenes de síntesis de vistas (A9) durante la cual la imagen de síntesis de vistas para el objeto es corregida de manera que se crea una imagen de síntesis de vistas corregida aplicando una función de corrección definida por los parámetros de corrección para la imagen de síntesis de vistas;
- en el que la señal de imagen de la imagen objeto de codificación se codifica usando la imagen de síntesis de vistas corregida en la etapa de codificación de imagen (A12).
2. El método de codificación de imágenes multivista según la reivindicación 1, en el que:
- se proporciona una etapa de determinación de objeto (A4) durante la cual los píxeles en la zona objeto de codificación se dividen en uno o varios grupos usando la información de profundidad para los píxeles pertinentes como referencia, y
  - en la etapa de determinación de profundidad representativa, se establece la información de profundidad representativa para cada grupo determinado en la etapa de determinación de objeto,
  - en la etapa de determinación de grupo de píxeles de muestra, se establece un grupo de píxeles de muestra para cada grupo determinado en la etapa de determinación de objeto,
  - en la etapa de estimación de parámetros de corrección, se estiman los parámetros de corrección para cada grupo determinado en la etapa de determinación de objeto, y
  - en la etapa de corrección de imágenes de síntesis de vistas, se corrige una imagen de síntesis de vistas para cada grupo determinado en la etapa de determinación de objeto.
3. El método de codificación de imágenes multivista según la reivindicación 1 ó 2, en el que:
- se proporciona una etapa de selección de modelo de corrección (A7) durante la cual se selecciona un modelo de corrección para corregir la imagen de síntesis de vistas para la zona objeto de codificación de acuerdo con el número de píxeles en el grupo de píxeles de muestra, y
  - en la etapa de estimación de parámetros de corrección, se estiman los parámetros de corrección para el modelo de corrección seleccionado en la etapa de selección de modelo de corrección, y
  - en la etapa de corrección de imágenes de síntesis de vistas, se corrige la imagen de síntesis de vistas usando el modelo de corrección seleccionado en la etapa de selección de modelo de corrección.



4. Un método de decodificación de imágenes multivista en el que una imagen objeto de decodificación de un objeto que es fotografiado por una primera cámara se divide en una pluralidad de zonas objeto de decodificación, y usando una imagen de síntesis de vistas que se sintetiza a partir de la información de profundidad para la imagen objeto de decodificación y a partir de una imagen decodificada del mismo objeto fotografiado por una segunda cámara que está situada en una posición diferente a la de la primera cámara, se realiza decodificación predictiva para cada una de las zonas objeto de decodificación, que comprende una etapa de decodificación de imagen (B12) en la que, usando la imagen de síntesis de vistas, se decodifica una señal de imagen de la imagen objeto de decodificación; estando el método de decodificación de imágenes multivista caracterizado porque comprende:
- 5 - una etapa de determinación de profundidad representativa (B6) durante la cual se establece la información de profundidad representativa para un objeto fotografiado en la zona objeto de decodificación,
  - 10 - una etapa de determinación de grupo de píxeles de muestra (B6) durante la cual se determina un grupo de píxeles en el que el mismo objeto que en la zona objeto de decodificación ha sido fotografiado en la imagen objeto de decodificación seleccionando píxeles de una zona ya codificada que es adyacente a la zona objeto de codificación, en la que la diferencia absoluta entre un valor de profundidad de la misma y el valor de profundidad representativo es menor que un umbral predefinido y el grupo de píxeles se establece como un grupo de píxeles de muestra,
  - 15 - una etapa de estimación de parámetros de corrección (B8) durante la cual, basándose en la imagen de síntesis de vistas para el grupo de píxeles de muestra y en una imagen decodificada que ha sido decodificada para el grupo de píxeles de muestra, se estiman los parámetros de corrección para corregir al menos una de las ausencias de correspondencia de iluminación o de color de manera que se reduzca al mínimo la suma de valores que representan un error entre los valores de píxel de la imagen decodificada para el grupo de píxeles de muestra y los valores de píxel de la imagen de síntesis de vistas para el grupo de píxeles de muestra que son corregidos usando un modelo de corrección para el grupo de píxeles de muestra, y
  - 20 - una etapa de corrección de imágenes de síntesis de vistas (B9) durante la cual se corrige la imagen de síntesis de vistas para el objeto de manera que se crea una imagen de síntesis de vistas corregida aplicando una función de corrección definida por los parámetros de corrección para la imagen de síntesis de vistas;
  - 25 en el que la señal de imagen de la imagen objeto de decodificación se decodifica usando la imagen de síntesis de vistas corregida en la etapa de decodificación de imagen (B12).
  - 30
  - 35 5. El método de decodificación de imágenes multivista según la reivindicación 4, en el que:
    - se proporciona una etapa de determinación de objeto (B4) durante la cual los píxeles en la zona objeto de decodificación se dividen en uno o varios grupos usando la información de profundidad para los píxeles pertinentes como referencia, y
    - 40 - en la etapa de determinación de profundidad representativa, se establece la información de profundidad representativa para cada grupo determinado en la etapa de determinación de objeto,
    - en la etapa de determinación de grupo de píxeles de muestra, se establece un grupo de píxeles de muestra para cada grupo determinado en la etapa de determinación de objeto,
    - 45 - en la etapa de estimación de parámetros de corrección, se estiman los parámetros de corrección para cada grupo determinado en la etapa de determinación de objeto, y
    - 50 - en la etapa de corrección de imágenes de síntesis de vistas, se corrige una imagen de síntesis de vistas para cada grupo determinado en la etapa de determinación de objeto.
  - 6. El método de decodificación de imágenes multivista según la reivindicación 4 ó 5, en el que:
    - 55 - se proporciona una etapa de selección de modelo de corrección (B7) durante la cual se selecciona un modelo de corrección para corregir la imagen de síntesis de vistas para la zona objeto de decodificación de acuerdo con el número de píxeles en el grupo de píxeles de muestra, y
    - en la etapa de estimación de parámetros de corrección, se estiman los parámetros de corrección para el modelo de corrección seleccionado en la etapa de selección de modelo de corrección, y
    - 60 - en la etapa de corrección de imágenes de síntesis de vistas, se corrige la imagen de síntesis de vistas usando el modelo de corrección seleccionado en la etapa de selección de modelo de corrección.
  - 65 7. Un dispositivo de codificación de imágenes multivista que codifica imágenes multivista dividiendo una imagen de entrada de un objeto que es fotografiado por una primera cámara en una pluralidad de zonas objeto de codificación

- y, usando una imagen de síntesis de vistas que se sintetiza a partir de la información de profundidad para la imagen de entrada y a partir de una imagen ya codificada del mismo objeto fotografiado por una segunda cámara que está situada en una posición diferente a la de la primera cámara, realizando codificación predictiva para cada una de las zonas objeto de codificación, que comprende una unidad de codificación de imagen (109) que, usando la imagen de síntesis de vistas, codifica una señal de imagen de la imagen objeto de codificación de manera que se crean datos codificados, siendo los datos codificados decodificados de manera que se crea una imagen decodificada para la zona objeto de codificación mediante una unidad de decodificación de imagen (110), estando el dispositivo de codificación de imágenes multivista caracterizado porque comprende:
- 5
- 10 - una unidad de determinación de profundidad representativa (107b) que establece la información de profundidad representativa para un objeto fotografiado en la zona objeto de codificación,
- una unidad de determinación de grupo de píxeles de muestra (107c) que determina un grupo de píxeles en el que el mismo objeto que en la zona objeto de codificación ha sido fotografiado en la imagen de entrada seleccionando píxeles de una zona ya codificada que es adyacente a la zona objeto de codificación, en la que la diferencia absoluta entre un valor de profundidad de la misma y el valor de profundidad representativo es menor que un umbral predefinido y el grupo de píxeles se establece como un grupo de píxeles de muestra,
- 15
- una unidad de estimación de parámetros de corrección (107e) que, basándose en la imagen de síntesis de vistas para el grupo de píxeles de muestra y en una imagen decodificada que ya ha sido decodificada para el grupo de píxeles de muestra, estima los parámetros de corrección para corregir al menos una de las ausencias de correspondencia de iluminación o de color de manera que se reduzca al mínimo la suma de valores que representan un error entre los valores de píxel de la imagen decodificada para el grupo de píxeles de muestra y los valores de píxel de la imagen de síntesis de vistas para el grupo de píxeles de muestra que son corregidos usando un modelo de corrección para el grupo de píxeles de muestra, y
- 20
- una unidad de corrección de imágenes de síntesis de vistas (108) que corrige la imagen de síntesis de vistas para el objeto de manera que se crea una imagen de síntesis de vistas corregida aplicando una función de corrección definida por los parámetros de corrección para la imagen de síntesis de vistas;
- 30
- en el que la unidad de codificación de imagen (109) codifica la señal de imagen de la imagen objeto de codificación usando la imagen de síntesis de vistas corregida.
8. Un dispositivo de decodificación de imágenes multivista que decodifica datos codificados para imágenes multivista dividiendo una imagen objeto de decodificación de un objeto que es fotografiado por una primera cámara en una pluralidad de zonas objeto de decodificación y, usando una imagen de síntesis de vistas que se sintetiza a partir de la información de profundidad para la imagen objeto de decodificación y a partir de una imagen decodificada del mismo objeto fotografiado por una segunda cámara que está situada en una posición diferente a la de la primera cámara, realizando decodificación predictiva para cada una de las zonas objeto de decodificación, que comprende una unidad de decodificación de imagen (209) que, usando la imagen de síntesis de vistas, decodifica una señal de imagen de la imagen objeto de decodificación; estando el dispositivo de decodificación de imágenes multivista caracterizado porque comprende:
- 35
- 40
- una unidad de determinación de profundidad representativa (207b) que establece la información de profundidad representativa para un objeto fotografiado en la zona objeto de decodificación,
- 45
- una unidad de determinación de grupo de píxeles de muestra (207c) en la que en la imagen objeto de decodificación se determina un grupo de píxeles en el que ha sido fotografiado el mismo objeto que en la zona objeto de decodificación seleccionando píxeles de una zona ya codificada que es adyacente a la zona objeto de codificación, en la que la diferencia absoluta entre un valor de profundidad de la misma y la profundidad representativa es menor que un umbral predefinido y el grupo de píxeles se establece como un grupo de píxeles de muestra,
- 50
- una unidad de estimación de parámetros de corrección (207e) que, basándose en la imagen de síntesis de vistas para el grupo de píxeles de muestra y en una imagen decodificada que ha sido decodificada para el grupo de píxeles de muestra, estima los parámetros de corrección para corregir al menos una de las ausencias de correspondencia de iluminación o de color de manera que se reduzca al mínimo la suma de valores que representan un error entre los valores de píxel de la imagen decodificada para el grupo de píxeles de muestra y los valores de píxel de la imagen de síntesis de vistas para el grupo de píxeles de muestra que son corregidos usando un modelo de corrección para el grupo de píxeles de muestra, y
- 55
- 60
- una unidad de corrección de imágenes de síntesis de vistas (208) que corrige la imagen de síntesis de vistas para el objeto de manera que se crea una imagen de síntesis de vistas corregida aplicando una función de corrección definida por los parámetros de corrección para la imagen de síntesis de vistas;
- 65
- en el que la unidad de decodificación de imagen (209) decodifica la señal de imagen de la imagen objeto de

decodificación usando la imagen de síntesis de vistas corregida.

9. Un programa de codificación de imágenes multivista que permite la ejecución del método de codificación de imágenes multivista según una cualquiera de las reivindicaciones 1 a 3 en un ordenador.

5

10. Un programa de decodificación de imágenes multivista que permite la ejecución del método de decodificación de imágenes multivista según una cualquiera de las reivindicaciones 4 a 6 en un ordenador.

FIG. 1

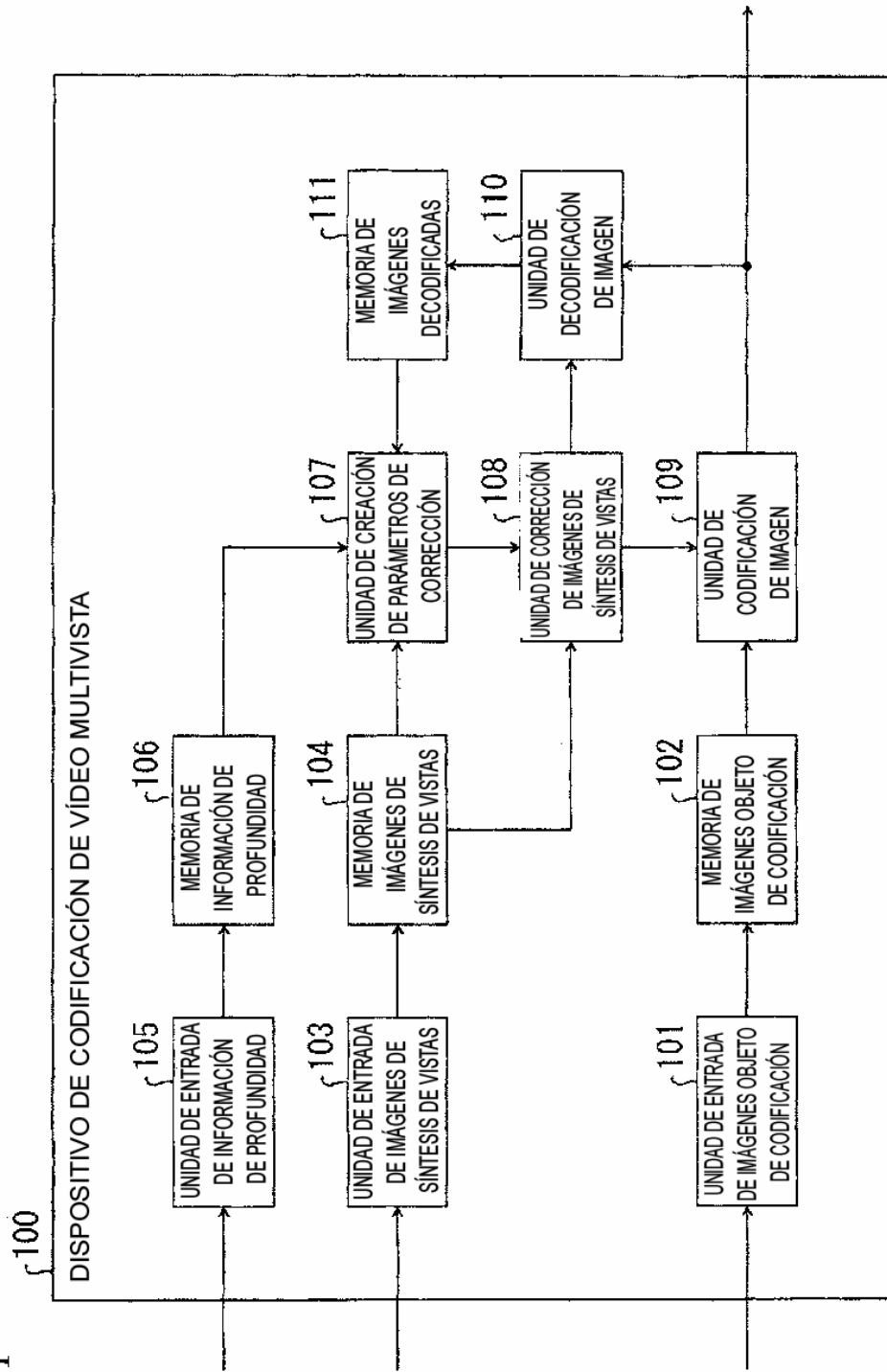


FIG. 2

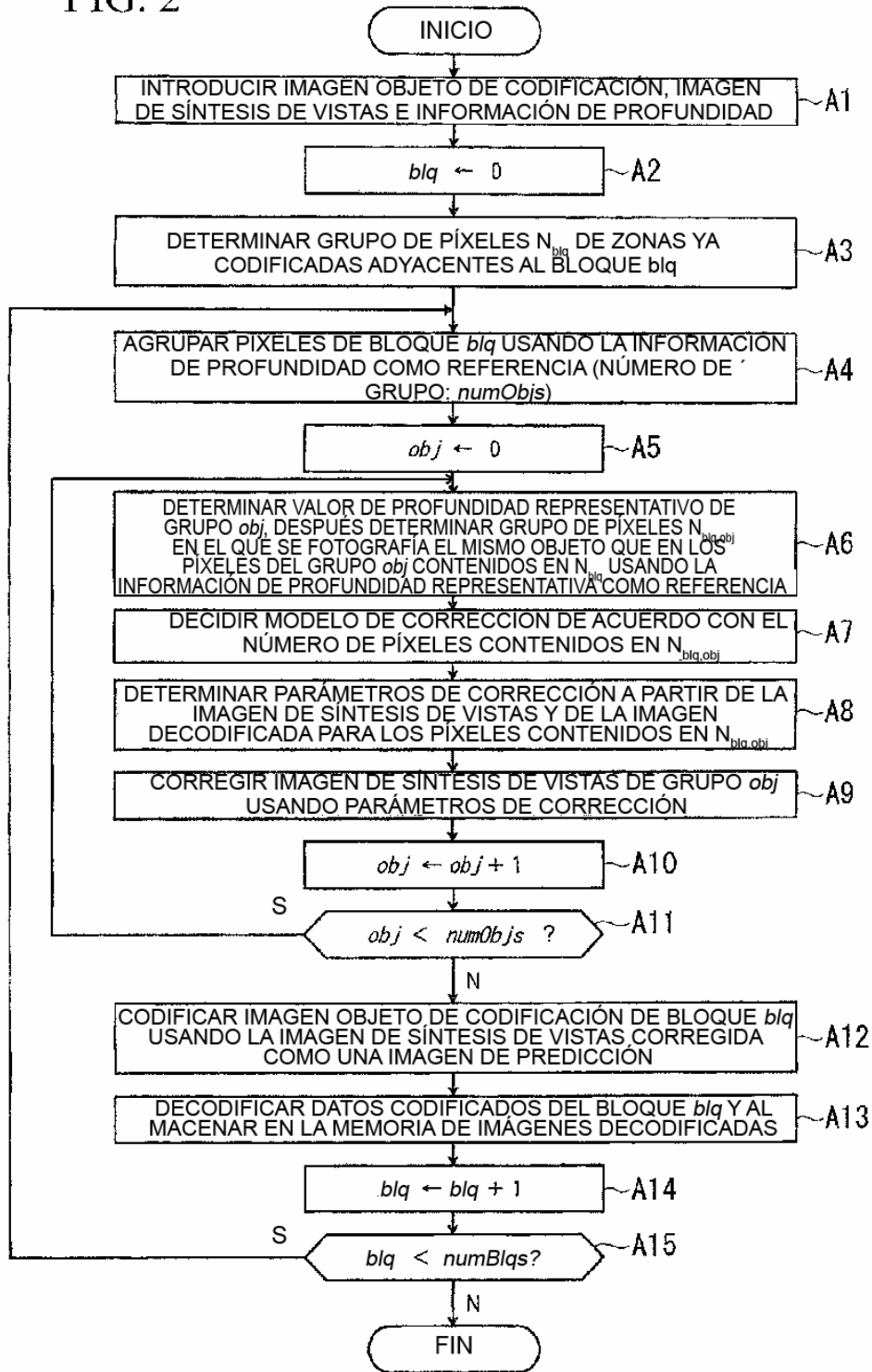


FIG. 3

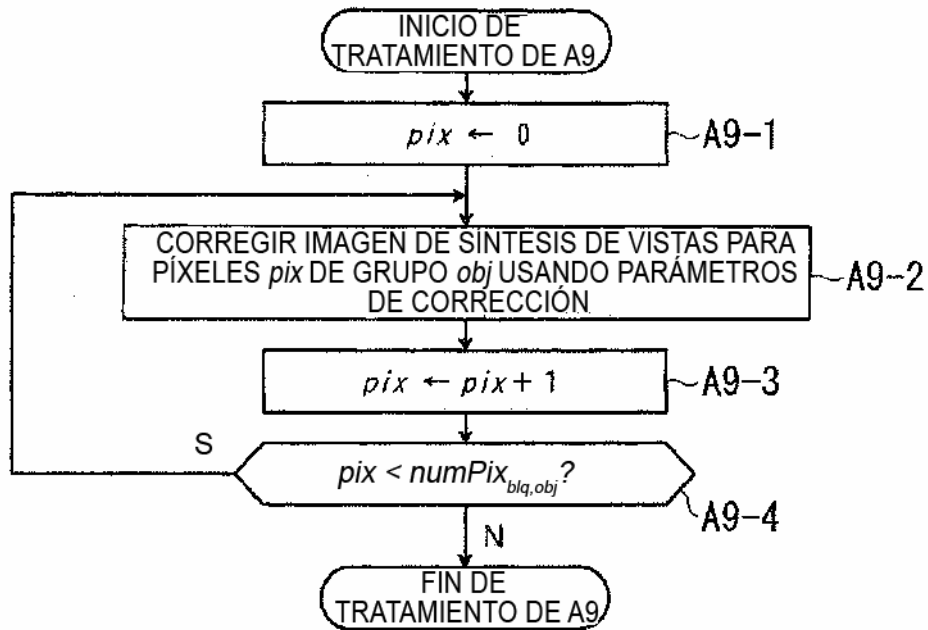


FIG. 4

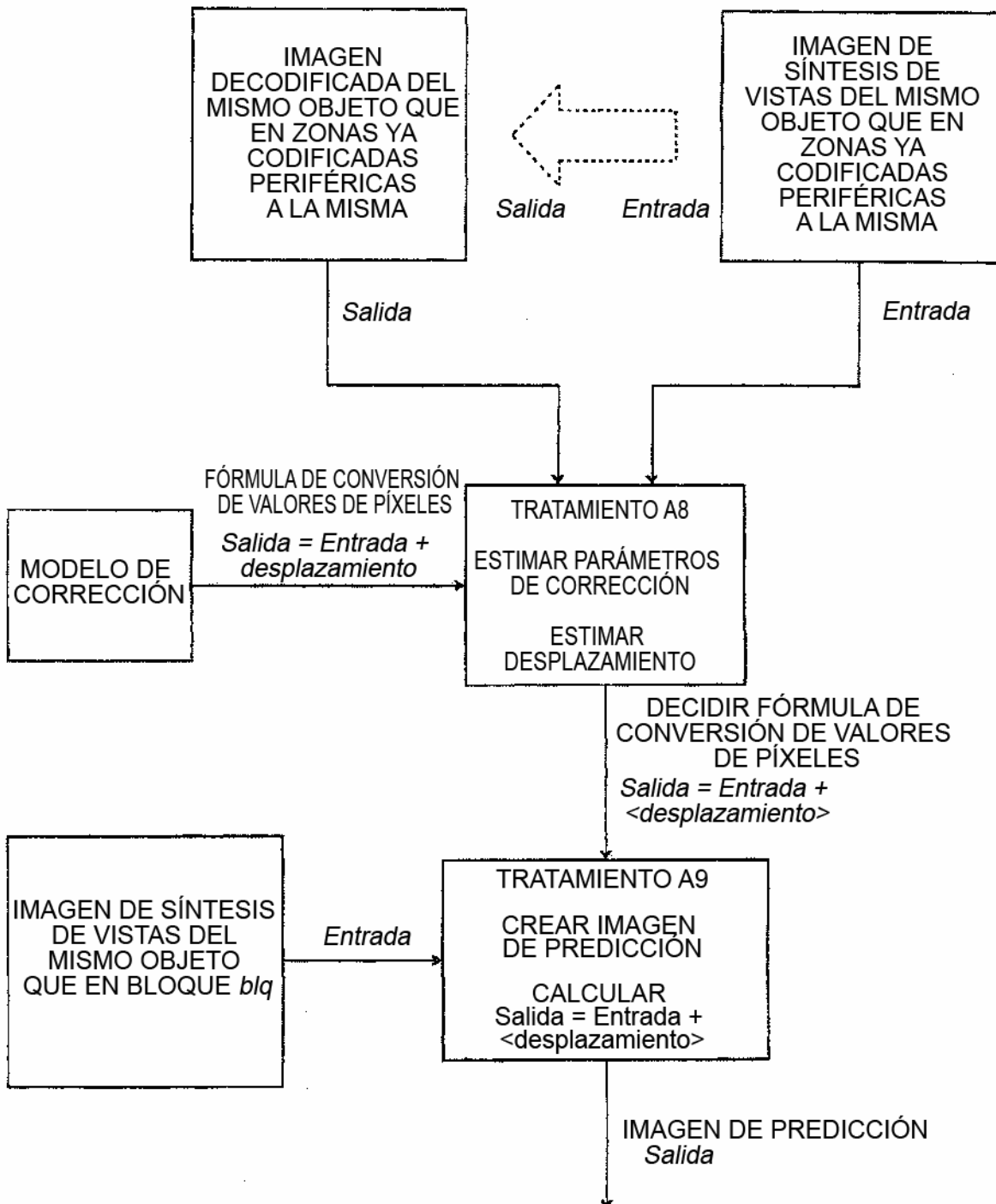


FIG. 5

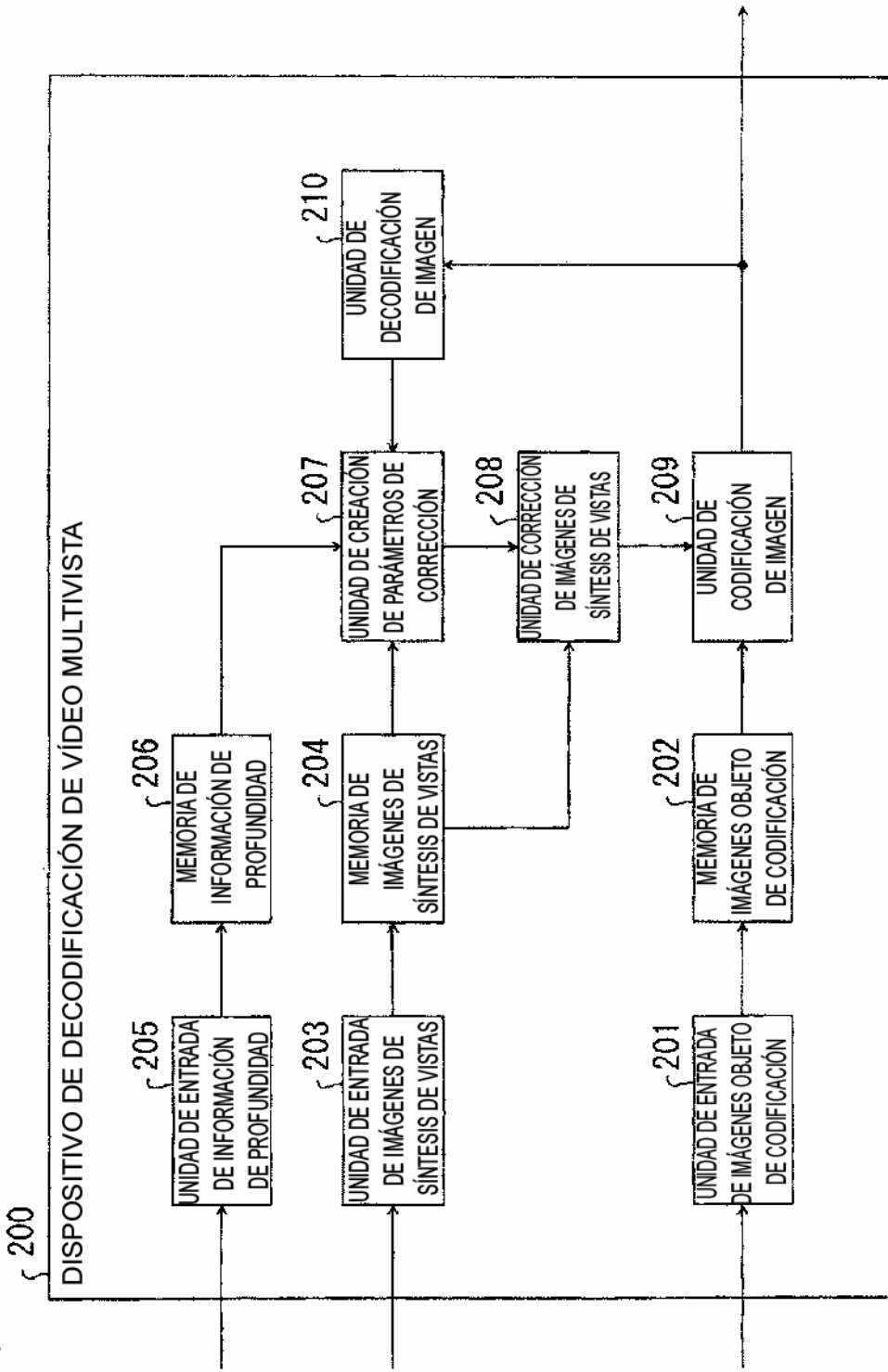




FIG. 6

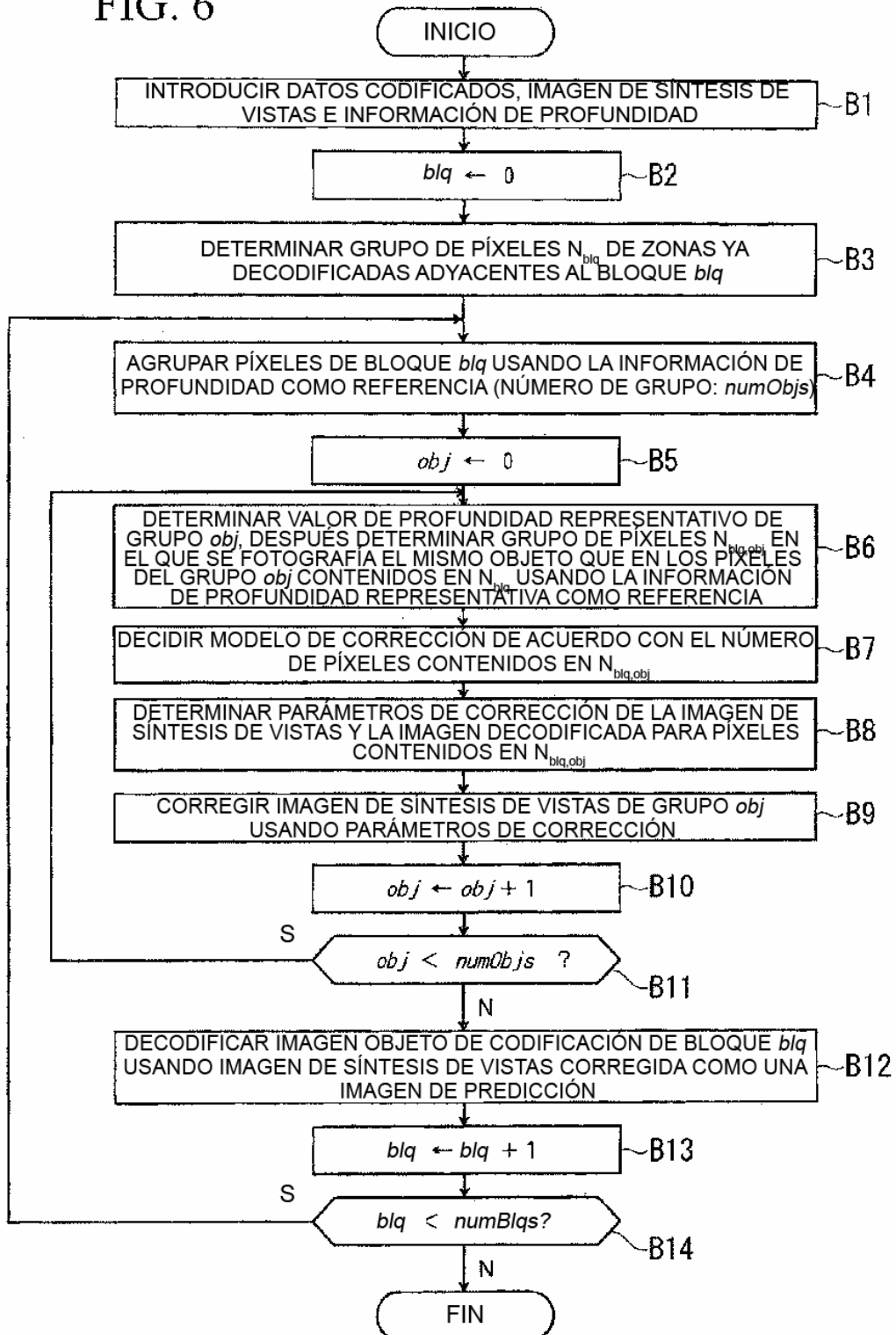


FIG. 7

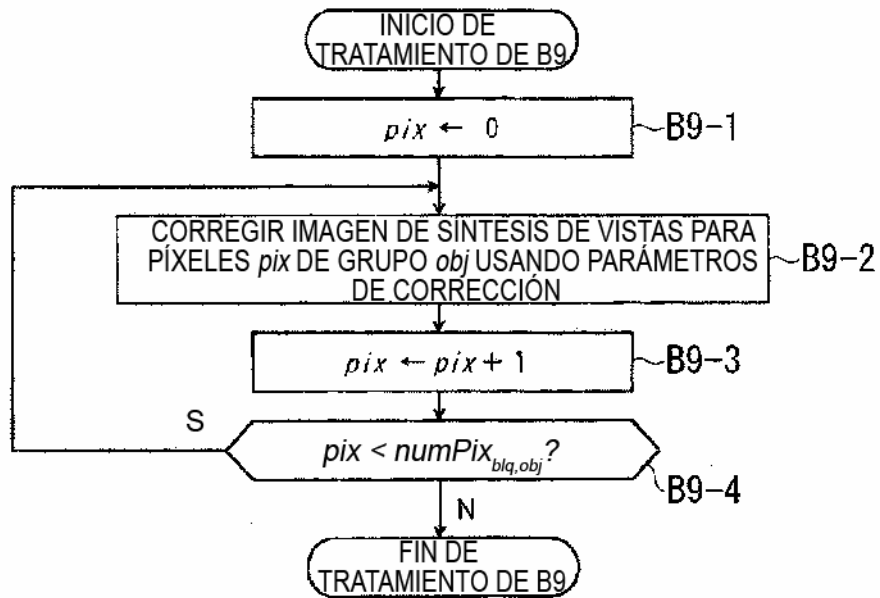


FIG. 8

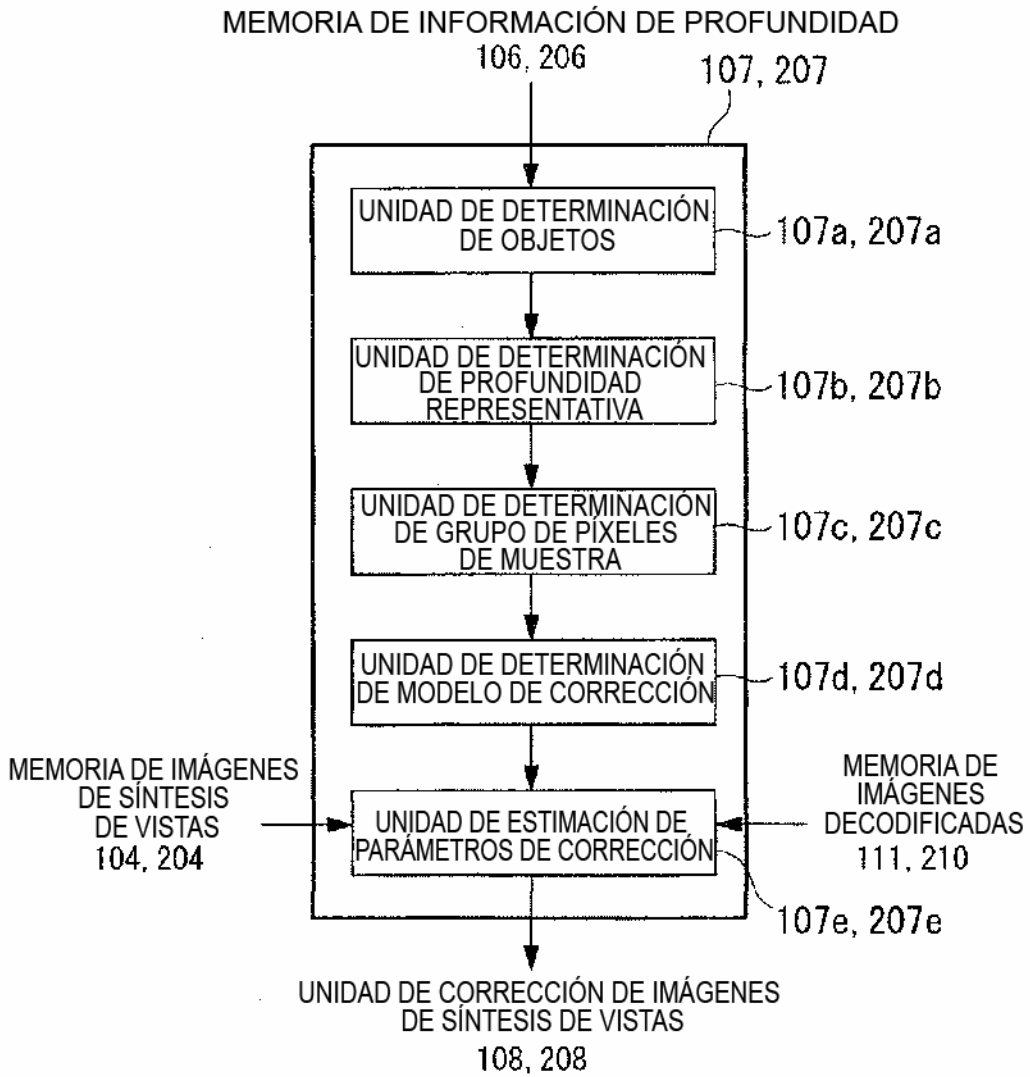


FIG. 9

