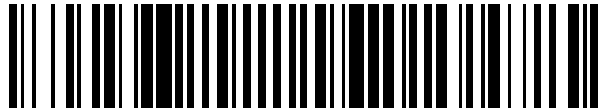


19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 402 229**

51 Int. Cl.:

G06T 7/00 (2006.01)

G06T 5/00 (2006.01)

G06K 9/20 (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

96 Fecha de presentación y número de la solicitud europea: **27.07.2009 E 09812625 (3)**

97 Fecha y número de publicación de la concesión europea: **16.01.2013 EP 2328125**

54 Título: **Método y dispositivo de fusión de imágenes**

30 Prioridad:

10.09.2008 CN 200810212183

45 Fecha de publicación y mención en BOPI de la traducción de la patente:
29.04.2013

73 Titular/es:

**HUAWEI DEVICE CO., LTD (100.0%)
Building 2, Zone B Huawei Industrial Base
Bantian, Longgang District
Shenzhen, Guangdong 518129 , CN**

72 Inventor/es:

**LI, KAI y
LIU, YUAN**

74 Agente/Representante:

LEHMANN NOVO, María Isabel

ES 2 402 229 T3

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

DESCRIPCIÓN

Método y dispositivo de fusión de imágenes

5 CAMPO DE LA INVENCION

La presente invención se refiere a una tecnología de procesamiento de imágenes y en particular, a un método y aparato de fusión de imágenes.

10 ANTECEDENTES DE LA INVENCION

Utilizando un sistema de reflexión / refracción y dispositivos de hardware, tales como lentes de distorsión, 'ojo de pez', se pueden establecer directamente imágenes de visión amplia sin discontinuidades. Sin embargo, dichos dispositivos de hardware capturan tanta información como sea posible en un plano de generación de imágenes limitado y en consecuencia, las imágenes de visión amplia, sin discontinuidades, sufren una distorsión importante.

Un dispositivo de creación de imágenes digitales fotografía imágenes en un formato digital. Múltiples imágenes digitales se fusionan juntas para formar una imagen panorámica digital de visión amplia. La imagen panorámica digital final implica poca distorsión. Las tecnologías de fusión de imágenes predominantes son como sigue:

(1) Múltiples imágenes se obtienen fotografiando desde diferentes direcciones de visión en un punto de vista fijo y luego, se calcula la relación de proyección punto a punto (también conocida como homografía) entre las imágenes y la relación se refleja principalmente como panorámica, inclinada, afín y de transformación de proyección. Más adelante, todas las imágenes se proyectan sobre un sistema de coordenadas uniforme en función de la relación de proyección y se fusionan en una imagen de visión amplia o en una imagen de 360 grados; por último, las discontinuidades en la zona de solapamiento formada en el proceso de fusión de imágenes se eliminan utilizando una función de alisado gráfico.

Este método de fusión de imágenes se suele aplicar a la itinerancia virtual en Internet. Quienes visiten un sitio web pueden explorar las imágenes panorámicas de objetos tal como se presentan en el mundo real. Sin embargo, en este método de fusión, solamente las imágenes obtenidas mediante fotografías en un punto de vista fijo tienen sentido. Todas las imágenes han de obtenerse fotografiando en la misma posición estática. De no ser así, la imagen fusionada tiene imágenes fantasma debido al paralaje y la zona de fusión no es natural. En la práctica, resulta difícil obtener imágenes fotografiando en la misma posición estática, esto es, difícil girar la cámara alrededor del centro óptico para hacer coincidir los centros ópticos de todas las imágenes.

(2) Después de que las imágenes se obtengan desde diferentes puntos de vista, algunos puntos intermedios continuos, entre los puntos de vista, se utilizan como nuevos puntos de vista para reconstruir imágenes y luego, se encuentra la zona de solapamiento entre la imagen antigua y la imagen reconstruida y las imágenes son objeto de fusión en conformidad con la zona de solapamiento.

El método para utilizar puntos intermedios como puntos de vista para reconstruir imágenes puede eliminar los 'fantasmas' de la imagen fusionada, que tiene dichos 'fantasmas' causados por el paralaje entre dos puntos de vista. Sin embargo, si el paralaje de las imágenes originales es demasiado grande, necesitan generarse numerosas imágenes de puntos intermedios para eliminar los 'fantasmas' de la imagen fusionada y se requieren cálculos de gran magnitud. Cuando se fusionan imágenes de más de dos puntos de vista, se generan demasiadas imágenes de puntos intermedios y la fusión de imágenes requiere demasiado tiempo.

(3) La longitud focal de la cámara, las coordenadas de puntos principales, el factor de distorsión, el ángulo de rotación y la posición de rotación se obtienen mediante la calibración de la cámara, con el fin de calcular la relación de proyección para las imágenes obtenidas por la cámara.

El tercer método anterior implica la calibración de la cámara, lo que es difícil en el campo de la informática. Asimismo, en este método, necesitan establecerse los fondos de las imágenes obtenidas como un fondo único para ignorar el paralaje de los fondos en la compensación del paralaje.

El documento de POLLEFEYS M ET AL: "Reconstrucción en 3D urbana en tiempo real detallada a partir de señales de vídeo" (International Journal of Computer Vision, Kluwer Academic Publisher, BO, vol. 78, nº 2- 3, 20 octubre 2007 (2007-10-20), páginas 143-167, XP019581883, ISSN: 1573-1450) da a conocer un sistema para la reconstrucción en 3D en tiempo real, geo-registrada, automática a partir de señales de vídeo de escenas urbanas. El sistema recoge flujos continuos de señales de vídeo así como mediciones de GPS con el fin de colocar el modelo reconstruido en coordenadas geo-registradas y utiliza el estado actual de los módulos en tiempo real para la etapa de procesamiento y diseña el sistema y las etapas de los conductos de procesamiento.

65 SUMARIO DE LA INVENCION

Las formas de realización de la presente invención proporcionan un método y aparato de fusión de imágenes para

eliminar los 'fantasmas' de la imagen fusionada, sin imponer estrictas restricciones sobre el escenario operativo de aplicación. El problema técnico objetivo se resuelve por las reivindicaciones del método contenidas en la reivindicación 1 y sus reivindicaciones subordinadas, por la reivindicación del dispositivo de la reivindicación 5 y sus reivindicaciones subordinadas.

5 Utilizando el método y aparato, en el proceso de fusión de dos imágenes, la zona de solapamiento entre las imágenes se encuentra primero y luego, se obtiene una imagen de profundidad de la zona de solapamiento. La imagen de profundidad refleja la distancia desde cada punto de píxel a la cámara, con precisión. Si los fondos fotográficos de las imágenes no son un fondo único, la información en las imágenes genera paralaje debido a las diferentes distancias a la cámara y los píxel en las imágenes se sitúan en diferentes capas. El método y aparato dados a conocer en la presente invención pueden calcular el paralaje entre las capas, de forma precisa, utilizando la imagen de profundidad y los píxel de la zona de solapamiento se pueden procesar, con precisión, en el proceso de fusión de imágenes para eliminar los 'fantasmas' de la imagen fusionada. Por lo tanto, el método y aparato aquí dados a conocer pueden eliminar los 'fantasmas' de la imagen fusionada, sin necesidad de imponer estrictas restricciones sobre el entorno operativo de aplicación.

Asimismo, el proceso de fusión, según la presente invención, implica solamente dos imágenes y una imagen de profundidad de la zona de solapamiento y no implica un nuevo cálculo de las imágenes de los puntos de vista intermedios, con lo que se reduce, en gran medida, la carga de cálculo.

20 BREVE DESCRIPCIÓN DE LOS DIBUJOS

Para ilustrar la solución técnica bajo la presente invención o la técnica anterior, con mayor claridad, a continuación se describen los dibujos adjuntos implicados en la descripción de las formas de realización de la presente invención o de la técnica anterior. Evidentemente, los dibujos adjuntos, descritos a continuación, son a modo de ejemplo solamente y no exhaustivos. Los expertos en esta materia pueden derivar otros dibujos a partir de dichos dibujos adjuntos, sin necesidad de un gran esfuerzo creativo.

30 La Figura 1 es un diagrama de flujo de un método de fusión de imágenes según la primera forma de realización de la presente invención;

La Figura 2 representa una estructura de un aparato de fusión de imágenes según la primera forma de realización de la presente invención;

35 La Figura 3 es un diagrama de flujo de un método de fusión de imágenes según la segunda forma de realización de la presente invención;

La Figura 4 representa una relación geométrica entre profundidad y paralaje según la segunda forma de realización de la presente invención;

40 La Figura 5 representa una estructura de un aparato de fusión de imágenes, según la segunda forma de realización de la presente invención,

45 La Figura 6 representa una estructura de una segunda unidad de obtención según la segunda forma de realización de la presente invención;

La Figura 7 es un diagrama de flujo de un método de fusión de imágenes según la tercera forma de realización de la presente invención y

50 La Figura 8 representa una estructura de un aparato de fusión de imágenes según la tercera forma de realización de la presente invención.

DESCRIPCIÓN DETALLADA DE LAS FORMAS DE REALIZACIÓN DE LA INVENCION

55 La solución técnica, bajo la presente invención, se describe en detalle haciendo referencia a las formas de realización y a los dibujos adjuntos. Evidentemente, estas formas de realización son meramente representativas de formas de realización particulares de la presente invención y habrán de considerarse como ilustrativas en su naturaleza y no exhaustivas ni restrictivas. Todas las demás formas de realización, que pueden derivarse por los expertos en esta materia a partir de las formas de realización aquí dadas sin ningún esfuerzo creativo, estarán dentro del alcance de protección de la presente invención.

Primera forma de realización

65 Según se ilustra en la Figura 1, un método de fusión de imágenes, en esta forma de realización, incluye las etapas siguientes:

101. Obtener dos imágenes mediante fotografías en cámaras en tiempo real o mediante la recuperación a partir de una biblioteca de imágenes existente.

5 102. Buscar la zona de solapamiento de las dos imágenes. Si las dos imágenes son adecuadas para la fusión, los objetos en las dos imágenes deben coincidir parcialmente.

10 La zona de solapamiento puede encontrarse utilizando: el método de detección de puntos característicos de la denominada Transformación de Características Invariantes a la Escala (SIFT), un método de detección de puntos característicos de Harris, un método de detección de puntos característicos de Susan, un método de adaptación de coincidencia en estéreo u otros métodos de detección de puntos característicos. En algunos casos, los puntos característicos pueden seleccionarse manualmente para obtener la zona de solapamiento.

15 El método de detección de puntos característicos SIFT es actualmente predominante porque es un invariante –afín. Por lo tanto, el método de detección de puntos característicos SIFT se utiliza aquí para encontrar la zona de solapamiento.

103. Obtener una imagen de profundidad de la zona de solapamiento para representar la información tridimensional de la zona de solapamiento. La información de tridimensional de la zona de solapamiento se requiere para una fusión de imágenes precisa.

20 104. Fusionar las dos imágenes en una sola imagen en función de la imagen de profundidad para evitar la presencia de ‘fantasmas’ en la zona de solapamiento de la imagen fusionada.

25 Según se indica en la Figura 2, un aparato de fusión de imágenes, dado a conocer según la primera forma de realización de la presente invención, incluye una primera unidad de obtención 21, una unidad de búsqueda 22, una segunda unidad de obtención 23 y una unidad de fusión 24; en donde

30 la primera unidad de obtención 21 está configurada para obtener dos imágenes, en donde: la primera unidad de obtención 21 es un par de cámaras si las dos imágenes se obtiene fotografiando en cámaras en tiempo real o la primera unidad de obtención 21 es una interfaz de entrada si las dos imágenes se recuperan desde una biblioteca de imágenes existente; la unidad de búsqueda 22 está configurada para buscar la zona de solapamiento de las dos imágenes, en donde la unidad de búsqueda 22 puede utilizarse con los principios de trabajo descritos en la etapa 102 anterior; la segunda unidad de obtención 23 está configurada para obtener la imagen de profundidad de la zona de solapamiento encontrada por la unidad de búsqueda y la unidad de fusión 24 está configurada para fusionar las dos imágenes en una sola imagen en función de la imagen de profundidad.

35 Utilizando el método y aparato de fusión de imágenes según esta forma de realización, si los fondos fotografiados de las imágenes no son un fondo único, los píxeles en las imágenes están situados en numerosas capas. El método y aparato, según esta forma de realización, puede calcular el paralaje entre las capas, con precisión, utilizando la imagen de profundidad y los píxeles de la zona de solapamiento pueden procesarse, con precisión, en el proceso de fusión de imágenes para eliminar los ‘fantasmas’ de la imagen fusionada.

Segunda forma de realización

45 Según se ilustra en la Figura 3, un método de fusión de imágenes, en esta forma de realización, incluye las etapas siguientes:

301. Utilizar dos cámaras para fotografiar objetos y obtener dos imágenes. Una zona de solapamiento existe entre los objetos de las dos imágenes.

50 302. Buscar la zona de solapamiento de las dos imágenes. En la práctica, la zona de solapamiento puede encontrarse utilizando: un método de detección de puntos característicos de SIFT, un método de detección de puntos característicos de Harris, un método de detección de puntos característicos de Susan, un método de adaptación por coincidencia de imágenes tridimensionales u otros métodos de detección de puntos característicos. En algunos casos, los puntos característicos pueden seleccionarse manualmente para obtener la zona de solapamiento.

55 303. Utilizar una cámara de profundidad para obtener directamente una imagen de profundidad de la zona de solapamiento. La cámara de profundidad mide la distancia desde el objeto a la cámara de profundidad directamente por medios físicos, a modo de ejemplo, emitiendo una onda electromagnética o una radiación infrarroja al objeto. La distancia desde el objeto a la cámara de profundidad se determina en función del tiempo requerido para reflejar la onda electromagnética o el rayo de infrarrojos desde el objeto. Es decir, en esta forma de realización, dos cámaras ordinarias, mencionadas en la etapa 301, necesitan obtener las dos imágenes y la cámara de profundidad necesita obtener la imagen de profundidad.

60 Lo anteriormente descrito es un método, a modo de ejemplo, para obtener la imagen de profundidad de la zona de solapamiento. En la práctica, el método para obtener la imagen de profundidad depende de la escena, a modo de ejemplo, la imagen de profundidad se determina realizando el cálculo del paralaje entre las dos imágenes.

5 El paralaje significa una distancia existente entre las dos cámaras y la distancia (profundidad) desde un punto objeto a las cámaras no es de magnitud infinita. Por lo tanto, la posición de las imágenes creadas del mismo punto objeto es diferente en las dos cámaras, lo que se conoce como paralaje. En términos generales, una mayor profundidad da lugar a un paralaje más pequeño y una menor profundidad da lugar a un mayor paralaje.

Según se ilustra en la Figura 4, el paralaje p está relacionado con la profundidad z_p como sigue:

$$\left\{ \begin{array}{l} \frac{x_L}{D} = \frac{x_p}{D - z_p} \\ \frac{x_R - x_B}{D} = \frac{x_p - x_B}{D - z_p} \end{array} \right. \Rightarrow \frac{x_L - x_R + x_B}{D} = \frac{x_B}{D - z_p} \Rightarrow p = x_B \left(1 - \frac{D}{D - z_p} \right) = x_B \left(\frac{1}{\frac{z_p}{D} - 1} + 1 \right)$$

10 $X_L - X_R$ es el paralaje, D es la longitud focal de la cámara y x_B es la distancia entre cámaras. Cuando la profundidad es pequeña, la relación entre paralaje p y la profundidad z_p es equivalente a $p \approx \frac{x_B}{D} z_p$.

15 Por lo tanto, la imagen de profundidad puede calcularse en función del paralaje después de que se calcule el paralaje a través de la coincidencia de imágenes. En esta forma de realización, los dos métodos de cálculo siguientes son aplicables:

20 (1) Calcular el paralaje entre los píxeles correspondientes en la zona de solapamiento de las dos imágenes directamente y calcular la imagen de profundidad de la zona de solapamiento en función del paralaje.

(2) Obtener dos nuevas imágenes en la zona de solapamiento en dos puntos de vista, calcular el paralaje entre los píxeles correspondientes en las dos nuevas imágenes y calcular la imagen de profundidad en la zona de solapamiento en función del paralaje.

25 El primer método anterior no necesita obtener, por separado, imágenes para calcular el paralaje, lo que reduce las cámaras requeridas y el segundo método necesita una cámara binocular adicional.

30 Una vez obtenida la imagen de profundidad de la zona de solapamiento, las imágenes pueden fusionarse. El proceso de fusión se describe a continuación.

304. Calcular la matriz homográfica entre las dos imágenes. Cada punto de píxel, en las dos imágenes, tiene un elemento de matriz correspondiente en la matriz homográfica.

35 305. Proyectar la primera imagen sobre el sistema de coordenadas objetivo, en función de la relación de proyección entre el punto de píxel y el elemento en la matriz homográfica, porque cada píxel, en cada imagen, corresponde a un elemento en la matriz homográfica.

40 En términos generales, la proyección es equivalente a proyectar completamente todos los puntos de píxel en la primera imagen sobre el sistema de coordenadas objetivo. Por lo tanto, el sistema de coordenadas objetivo puede ser simplemente el sistema de coordenadas de la primera imagen. De este modo, la operación de proyección compleja no se requiere para la primera imagen y la función de proyección, en esta etapa, se puede realizar en tanto que la primera imagen sea reservada de forma completa.

45 306. Calcular el paralaje entre píxeles correspondientes de las dos imágenes en la zona de solapamiento, respectivamente, en función del valor de profundidad en la imagen de profundidad, considerando que el paralaje p está

relacionado con la profundidad z_p como $p \approx \frac{x_B}{D} z_p$, en donde D es la longitud focal de la cámara y x_B es la distancia entre cámaras, según se indicó en la etapa 304.

50 307. Juzgar si el paralaje calculado supera el margen de paralaje preestablecido. Si es así, proseguir con la etapa 309; si el paralaje calculado no excede el margen de paralaje preestablecido, proseguir con la etapa 308. Considerando la precisión requerida para la imagen fusionada, el margen de paralaje preestablecido es 0 píxel a 1 píxel, preferentemente.

308. Desechar los puntos de píxel de la segunda imagen y proseguir con la etapa 310.

55 309. Proyectar los puntos de píxel de la segunda imagen sobre las coordenadas de los puntos de píxeles correspondientes de la primera imagen en el sistema de coordenadas objetivo. Es decir, calcular el valor de crominancia media ponderada y el valor de luminancia media ponderada de los puntos de píxeles correspondientes en las dos

imágenes, utilizar el valor de la crominancia media ponderada y el valor de la luminancia media ponderada como el valor de crominancia y el valor de luminancia de los puntos de píxeles correspondientes en la primera imagen, en el sistema de coordenadas objetivo, respectivamente y luego, proseguir con la etapa 310.

5 310. Juzgar si los píxeles de la zona de solapamiento están procesados completamente. Si es así, proseguir con la etapa 311; si los píxeles de la zona de solapamiento no están procesados completamente, volver a la etapa 306.

10 311. Proyectar los píxeles, exteriores a la zona de solapamiento en la segunda imagen, sobre el sistema de coordenadas objetivo en función de la relación de proyección entre el píxel en la segunda imagen y el elemento en la matriz homográfica. Esta etapa es equivalente a proyectar los píxeles, exteriores a la zona de solapamiento, sobre el sistema de coordenadas objetivo directamente. El sistema de coordenadas objetivo, en esta forma de realización, es el sistema de coordenadas de la primera imagen.

15 En correspondencia con el método de fusión de imágenes anterior, un aparato de fusión de imágenes se da a conocer en la segunda forma de realización de la presente invención. Según se ilustra en la Figura 5, el aparato de fusión de imágenes incluye: una primera unidad de obtención 51, una unidad de búsqueda 52, una segunda unidad de obtención 53 y una unidad de fusión 54.

20 La primera unidad de obtención 51 está configurada para obtener dos imágenes. En esta forma de realización, las dos imágenes se obtienen fotografiando con cámaras en tiempo real. Por lo tanto, la primera unidad de obtención 51 es un par de cámaras ordinarias. La unidad de búsqueda 52 está configurada para buscar la zona de solapamiento de las dos imágenes.

25 Una vez fusionada la zona de solapamiento, se fusionan las dos imágenes. Para reducir los 'fantasmas' generados en la fusión de la zona de solapamiento, necesita obtenerse primero una imagen de profundidad de la zona de solapamiento. La segunda unidad de obtención 53, en esta forma de realización, está configurada para obtener la imagen de profundidad de la zona de solapamiento. La segunda unidad de obtención 53 puede funcionar en las formas siguientes:

30 (1) La segunda unidad de obtención 53 sirve directamente como una cámara de profundidad.

(2) La segunda unidad de obtención 53 incluye un módulo de cálculo 531. El módulo de cálculo 531 está configurado para calcular el paralaje entre los píxeles correspondientes en la zona de solapamiento de las dos imágenes y para calcular la imagen de profundidad de la zona de solapamiento en función del paralaje. El paralaje p está relacionado con

35 la profundidad z_p como
$$p \approx \frac{x_B}{D} z_p$$
, en donde D es la longitud focal de la cámara y x_B es la distancia entre cámaras.

(3) Según se ilustra en la Figura 6, la segunda unidad de obtención incluye un módulo de obtención 61 y un módulo de cálculo 62.

40 El módulo de obtención 61 está configurado para obtener dos nuevas imágenes en la zona de solapamiento en dos puntos de vista y es equivalente a una cámara binocular.

El módulo de cálculo 62 está configurado para calcular el paralaje entre los píxeles correspondientes en las dos nuevas imágenes y para calcular la imagen de profundidad, en la zona de solapamiento, en función del paralaje. El paralaje p

45 está relacionado con la profundidad z_p como
$$p \approx \frac{x_B}{D} z_p$$
, en donde D es la longitud focal de la cámara y x_B es la distancia entre cámaras.

50 Una vez obtenida la imagen de profundidad anterior, la unidad de fusión efectúa la fusión de las dos imágenes en una sola imagen en función de la imagen de profundidad. La unidad de fusión 54 incluye un primer módulo de cálculo 541, un módulo de proyección 542 y un segundo módulo de cálculo 543.

55 El primer módulo de cálculo 541 está configurado para calcular la matriz homográfica entre dos imágenes. Cada píxel, en las dos imágenes, tiene un elemento correspondiente en la matriz homográfica. Si dos puntos de píxel, en las dos imágenes, corresponden al mismo punto en el objeto, los dos puntos de píxel corresponden al mismo elemento en la matriz homográfica. El módulo de proyección 542 está configurado para proyectar la primera imagen sobre el sistema de coordenadas objetivo, según la relación de proyección entre la primera imagen y la matriz homográfica. El segundo módulo de cálculo 543 está configurado para calcular el paralaje entre los píxeles correspondientes de las dos imágenes, en la zona de solapamiento, en función de la imagen de profundidad. Cuando el paralaje cae dentro del margen de paralaje preestablecido, el módulo de proyección 542 desecha los puntos de píxel de la segunda imagen directamente; cuando el paralaje es superior al margen de paralaje preestablecido, el módulo de proyección 542 efectúa la proyección de los puntos de píxel de la segunda imagen sobre las coordenadas de los puntos de píxeles correspondientes de la primera imagen en el sistema de coordenadas objetivo. Es decir, el módulo de proyección 542 calcula el valor de crominancia media ponderada y el valor de luminancia media ponderada de los puntos de píxeles correspondientes en las

dos imágenes y utiliza el valor de crominancia media ponderada y el valor de luminancia media ponderada como el valor de crominancia y el valor de luminancia de los puntos de píxels correspondientes en la primera imagen, en el sistema de coordenadas objetivo, respectivamente. Considerando la precisión requerida para la imagen fusionada, el margen de paralaje preestablecido es 0 píxel a 1 píxel, preferentemente. Una vez realizada la proyección de la zona de solapamiento, el módulo de proyección 542, en esta forma de realización, necesita proyectar los píxels, exteriores a la zona de solapamiento en la segunda imagen, sobre el sistema de coordenadas objetivo en función de la relación de proyección entre el píxel en la segunda imagen y el elemento en la matriz homográfica. De este modo, la imagen fusionada está completa y no se pierde ningún dato original.

Utilizando el método y aparato de fusión de imágenes, según esta forma de realización, si los fondos fotográficos de las imágenes no son un fondo único, los píxels en las imágenes se sitúan en numerosas capas. El método y aparato, en esta forma de realización, puede calcular el paralaje entre las capas, con precisión, utilizando la imagen de profundidad y los píxels de la zona de solapamiento se pueden procesar, con precisión, en el proceso de fusión de imágenes para eliminar los 'fantasmas' de la imagen fusionada. Además, esta forma de realización no implica ninguna calibración de la cámara, con lo que se simplifica la aplicación de esta forma de realización.

Puesto que esta forma de realización está libre de calcular numerosas imágenes de punto de vista intermedios, esta forma de realización ahorra los recursos de cálculo en comparación con el método para fusionar imágenes basado en puntos de vista intermedios. En el método para fusionar imágenes basado en puntos de vista intermedios, cuando un objeto se observa desde dos puntos de vista, el objeto es visible desde un punto de vista y no es visible desde el otro punto de vista y pueden producirse sombras o cavidades en el momento de generación de las imágenes de puntos intermedios. Por lo tanto, la relación de sombras necesita calcularse con precisión y necesita rellenarse las cavidades. En esta forma de realización, sin embargo, se utiliza una imagen de profundidad para expresar la información tridimensional de la imagen. Por lo tanto, en la imagen fusionada, la relación de sombras se determina directamente en función de la información de profundidad y no se genera ninguna cavidad.

En una aplicación práctica de esta forma de realización, se utilizan dos cámaras para fotografiar un solo objeto simultáneamente, que no requieren la coincidencia de los centros ópticos de las cámaras. Por lo tanto, esta forma de realización permite que los centros ópticos de las cámaras estén en diferentes puntos en el proceso de realización de fotografías.

Tercera forma de realización

Según se ilustra en la Figura 7, un método de fusión de imágenes en esta forma de realización incluye las etapas siguientes:

701. Utilizar dos cámaras para fotografiar objetos y obtener dos imágenes. Una zona de solapamiento existe entre los objetos de j xs imágenes.

702. Buscar la zona de solapamiento de las dos imágenes. En la práctica, la zona de solapamiento puede encontrarse utilizando: un método de detección de puntos característicos de SIFT, un método de detección de puntos característicos de Harris, un método de detección de puntos característicos de Susan, un método de coincidencia de imágenes tridimensionales u otros métodos de detección de puntos característicos. En algunos casos, los puntos característicos pueden seleccionarse manualmente para obtener la zona de solapamiento.

703. Utilizar una cámara de profundidad para obtener una imagen de profundidad de la zona de solapamiento directamente. La cámara de profundidad mide la distancia desde el objeto a la cámara de profundidad directamente por medios físicos, a modo de ejemplo, emitiendo una onda electromagnética o un rayo de infrarrojos al objeto. La distancia desde el objeto a la cámara de profundidad se determina según el tiempo requerido para reflejar la onda electromagnética o el rayo de infrarrojos desde el objeto. Es decir, en esta forma de realización, necesitan obtenerse dos cámaras ordinarias, mencionadas en la etapa 701, y la cámara de profundidad necesita obtener la imagen de profundidad.

Lo anteriormente descrito es un método, a modo de ejemplo, para obtener la imagen de profundidad de la zona de solapamiento. En la práctica, el método para obtener la imagen de profundidad depende de la escena, a modo de ejemplo, la imagen de profundidad se determina realizando el cálculo del paralaje entre las dos imágenes.

704. Generar imágenes de puntos de vista virtuales intermedios continuas entre los puntos de vistas de las dos imágenes en función de la imagen de profundidad. Más concretamente, calcular el paralaje entre la imagen del punto de vista virtual y la imagen conocida utilizando la fórmula siguiente:

$$d=A1O1'-A2O2'=f*(CO1-CO2)/Z=f* B_{p0}/Z.$$

En la fórmula, d es el paralaje entre la imagen de punto de vista virtual y la imagen de punto de vista conocida, f es la longitud focal de la cámara, B_{p0} es la distancia entre los puntos de vista virtuales y Z es la profundidad de la imagen.

El paralaje anteriormente calculado y la imagen conocida se pueden utilizar para generar la imagen de puntos de vista virtuales intermedios utilizando la fórmula siguiente:

5 $x_p = d \cdot \frac{B_{p0}}{B} + x_0$, en donde x_0 es una coordenada X del punto de píxel de la imagen conocida, x_p es la coordenada X correspondiente del punto de píxel de la imagen conocida en la imagen de punto de vista virtual intermedio, B es la distancia entre dos puntos fotografiados originales y B_{p0} es la distancia desde el punto de vista virtual intermedio a uno de los puntos fotografiados originales.

10 705. Fusionar las dos imágenes y las imágenes de puntos de vista virtuales intermedios en una sola imagen. Más concretamente, disponer todas las imágenes en el orden de puntos de vista y luego, fusionar cada dos imágenes.

15 En correspondencia con el método de fusión de imágenes anterior, un aparato de fusión de imágenes se da a conocer en la tercera forma de realización de la presente invención. Según se indica en la Figura 8, el aparato de fusión de imágenes incluye: una primera unidad de obtención 81, una unidad de búsqueda 82, una segunda unidad de obtención 83 y una unidad de fusión 84.

20 La primera unidad de obtención 81 está configurada para obtener dos imágenes. En esta forma de realización, las dos imágenes se obtienen fotografiando con cámaras en tiempo real. Por lo tanto, la primera unidad de obtención 81 es un par de cámaras ordinarias. La unidad de búsqueda 82 está configurada para buscar la zona de solapamiento de las dos imágenes. La segunda unidad de obtención 83 es una cámara de profundidad y está configurada para obtener una imagen de profundidad de la zona de solapamiento.

25 La unidad de fusión 84 incluye un módulo de generación 841 y un módulo de fusión 842.

El módulo de generación 841 está configurado para generar imágenes de puntos de vista virtuales intermedios continuas entre puntos de vista de las dos imágenes en función de la imagen de profundidad y más concretamente, calcular el paralaje entre la imagen del punto de vista virtual intermedio y la imagen conocida utilizando la fórmula siguiente:

30
$$d = A1O1' - A2O2' = f \cdot (CO1 - CO2) / Z = f \cdot B_{p0} / Z.$$

En la fórmula d es el paralaje entre la imagen de punto de vista virtual y la imagen de punto de vista conocida, f es la longitud focal de la cámara, B_{p0} es la distancia entre puntos de vista virtuales y Z es la profundidad de la imagen.

35 El paralaje anteriormente calculado y la imagen conocida se pueden utilizar para generar la imagen del punto de vista virtual intermedio utilizando la fórmula siguiente:

40 $x_p = d \cdot \frac{B_{p0}}{B} + x_0$, en donde: x_0 es una coordenada X del punto de píxel de la imagen conocida, x_p es la coordenada X correspondiente del punto de píxel de la imagen conocida en la imagen del punto de vista virtual intermedio, B es la distancia entre dos puntos fotografiados originales y B_{p0} es la distancia desde el punto de vista virtual intermedio a uno de los puntos de fotografía originales.

45 El módulo de fusión 842 está configurado para fusionar las dos imágenes y las imágenes de puntos de vista virtuales intermedios en una sola imagen. Más concretamente, el módulo de fusión 842 dispone todas las imágenes en el orden de puntos de vista y luego, fusiona cada dos imágenes.

50 En esta forma de realización, el método y el aparato de fusión utilizan una imagen de profundidad para calcular directamente la imagen de punto de vista virtual intermedio y luego, fusionar la imagen de punto de vista virtual intermedio y las dos imágenes originales en una sola imagen. En comparación con la técnica anterior, que utiliza dos imágenes para calcular la imagen de punto de vista virtual intermedio, el método para utilizar una imagen de profundidad para calcular la imagen de punto de vista virtual intermedio directamente, en esta forma de realización, puede poner en práctica cálculos más rápidos.

55 En esta forma de realización, se obtienen dos imágenes desde diferentes puntos de vista, se calcula una imagen de profundidad de la zona de solapamiento y las dos imágenes se fusionan en función de la imagen de profundidad. Por lo tanto, el método de fusión es aplicable a más de las imágenes obtenidas desde el mismo punto de vista. Si dos imágenes obtenidas, desde el mismo punto de vista, tienen una zona de solapamiento, el método de fusión, dado a conocer en esta forma de realización, es también aplicable.

Las formas de realización anteriores utilizan solamente dos imágenes para la descripción. En la práctica, en el caso de fusión de más de dos imágenes, se fusionan dos imágenes en una sola imagen primero, que luego se fusiona a una tercera imagen de la misma manera y así sucesivamente, para obtener una imagen de visión amplia.

- 5 El método y aparato, según la presente invención, son aplicables a videoconferencias multiescena y las ocasiones de obtener imágenes o vídeos de visión amplia.

10 Después de la lectura de las formas de realización anteriores, los expertos en esta materia tienen un claro conocimiento de que las formas de realización de la presente invención se pueden poner en práctica mediante hardware solamente o en una forma de realización preferida, a través de software en adición al hardware universal necesario. Por lo tanto, la solución técnica bajo la presente invención o su novedad, respecto a la técnica anterior, pueden materializarse en un producto de software. El producto de software se memoriza en un medio de almacenamiento legible por ordenador tal como un disco flexible de ordenador, un disco duro y una memoria de solamente lectura-disco compacto (CR-ROM) e incorpora varias instrucciones para dar instrucciones a un dispositivo informático (a modo de ejemplo, un ordenador personal, un servidor o un dispositivo de red) para ejecutar el método dado a conocer en cualquier forma de realización de la presente invención.

15 Las anteriores descripciones son realizaciones simplemente preferidas de la presente invención, pero no pretenden limitar el alcance de protección de la presente invención. Cualesquiera modificaciones, variaciones o sustituciones que puedan fácilmente derivarse por los expertos en esta materia debe caer dentro del alcance de protección de la presente invención. Por lo tanto, el alcance de la presente invención está sujeto a las reivindicaciones adjuntas.

20

REIVINDICACIONES

1. Un método de fusión de imágenes, que comprende:

- 5 la obtención de dos imágenes entre las que existe una zona de solapamiento;
la búsqueda de la zona de solapamiento de las dos imágenes;
la obtención de una imagen de profundidad de la zona de solapamiento y
10 la fusión de las dos imágenes en una sola imagen en conformidad con la imagen de profundidad,
caracterizado porque: la etapa de fusión de las dos imágenes en una sola imagen, en conformidad con la imagen de
profundidad, comprende:
15 el cálculo de una matriz homográfica entre las dos imágenes;
la proyección de una primera imagen sobre un sistema de coordenadas objetivo en conformidad con una relación de
proyección entre la primera imagen y la matriz homográfica;
20 el cálculo de un paralaje entre píxels correspondientes de las dos imágenes en la zona de solapamiento en conformidad
con la imagen de profundidad;
la eliminación de puntos de píxel de una segunda imagen si el paralaje se sitúa en el interior de una margen de paralaje
preestablecido;
25 la proyección de puntos de píxel de la segunda imagen, en la zona de solapamiento, sobre coordenadas de puntos de
píxels correspondientes de la primera imagen, en el sistema de coordenadas objetivo, si el paralaje supera el margen de
paralaje preestablecido y
30 la proyección de píxels de la segunda imagen, exteriores a la zona de solapamiento, sobre el sistema de coordenadas
objetivo, en conformidad con una relación de proyección entre la segunda imagen y la matriz homográfica.

35 2. El método de fusión de imágenes según la reivindicación 1, en donde la etapa de obtener la imagen de profundidad
de la zona de solapamiento comprende:

- la obtención de la imagen de profundidad de la zona de solapamiento mediante la utilización directa de una cámara de
profundidad o el cálculo de un paralaje entre píxels correspondientes en la zona de solapamiento de las dos imágenes y
el cálculo de la imagen de profundidad de la zona de solapamiento en función del paralaje o
40 la obtención de dos nuevas imágenes, en la zona de solapamiento, desde dos puntos de vista, el cálculo de un paralaje
entre píxels correspondientes en las dos nuevas imágenes y el cálculo de la imagen de profundidad en la zona de
solapamiento en función del paralaje.

45 3. El método de fusión de imágenes según la reivindicación 1, en donde:

el sistema de coordenadas objetivo es un sistema de coordenadas de la primera imagen.

50 4. El método de fusión de imágenes según la reivindicación 1, en donde:

- la etapa de proyección de los puntos de píxel de la segunda imagen en la zona de solapamiento sobre las coordenadas
de los puntos de píxels correspondientes de la primera imagen en el sistema de coordenadas objetivo comprende:
el cálculo de un valor de crominancia media ponderada y de un valor de luminancia media ponderada de los puntos de
píxels correspondientes en las dos imágenes y la utilización del valor de crominancia media ponderada y del valor de
luminancia media ponderada como un valor de crominancia y un valor de luminancia de los correspondientes puntos de
píxel de la primera imagen, en el sistema de coordenadas objetivo, respectivamente.

60 5. Un aparato de fusión de imágenes, que comprende:

- una primera unidad de obtención, configurada para obtener dos imágenes entre las que existe una zona de
solapamiento;
una unidad de búsqueda, configurada para buscar la zona de solapamiento de las dos imágenes;
65 una segunda unidad de obtención, configurada para obtener una imagen de profundidad de la zona de solapamiento y

una unidad de fusión, configurada para realizar la fusión de las dos imágenes en una sola imagen en conformidad con la imagen de profundidad,

5 caracterizado porque la unidad de fusión comprende:

un primer módulo de cálculo, configurado para calcular una matriz homográfica entre las dos imágenes;

10 un segundo módulo de cálculo, configurado para calcular un paralaje entre píxels correspondientes de las dos imágenes, en la zona de solapamiento, en conformidad con la imagen de profundidad y

un módulo de proyección, configurado para: proyectar una primera imagen sobre un sistema de coordenadas objetivo en función de una relación de proyección entre la primera imagen y la matriz homográfica;

15 proyectar puntos de píxel de una segunda imagen, en la zona de solapamiento, sobre coordenadas de puntos de píxels correspondientes de la primera imagen, en el sistema de coordenadas objetivo, si el paralaje supera un margen de paralaje preestablecido y proyectar píxel de la segunda imagen, exteriores a la zona de solapamiento, sobre el sistema de coordenadas objetivo, en función de una relación de proyección entre la segunda imagen y la matriz homográfica.

20 **6.** El aparato de fusión de imágenes según la reivindicación 5, en donde:

la segunda unidad de obtención es una cámara de profundidad.

25 **7.** El aparato de fusión de imágenes según la reivindicación 5, en donde la segunda unidad de obtención comprende:

un módulo de cálculo, configurado para calcular el paralaje entre píxels correspondientes en la zona de solapamiento de las dos imágenes y calcular la imagen de profundidad, en la zona de solapamiento, en conformidad con el paralaje.

30 **8.** El aparato de fusión de imágenes según la reivindicación 5, en donde la segunda unidad de obtención comprende:

un módulo de obtención, configurado para obtener dos nuevas imágenes en la zona de solapamiento en dos puntos de vista y

35 un módulo de cálculo, configurado para calcular el paralaje entre píxels correspondientes en las dos nuevas imágenes y calcular la imagen de profundidad en la zona de solapamiento en función del paralaje.

9. El aparato de fusión de imágenes según la reivindicación 5, en donde:

40 el módulo de proyección está configurado para calcular un valor de crominancia media ponderada y un valor de luminancia media ponderada de los puntos de píxels correspondientes, en las dos imágenes, si el paralaje supera el margen de paralaje preestablecido y utilizar el valor de crominancia media ponderada y el valor de luminancia media ponderada como un valor de crominancia y un valor de luminancia de los puntos de píxels correspondientes de la primera imagen en el sistema de coordenadas objetivos, respectivamente.

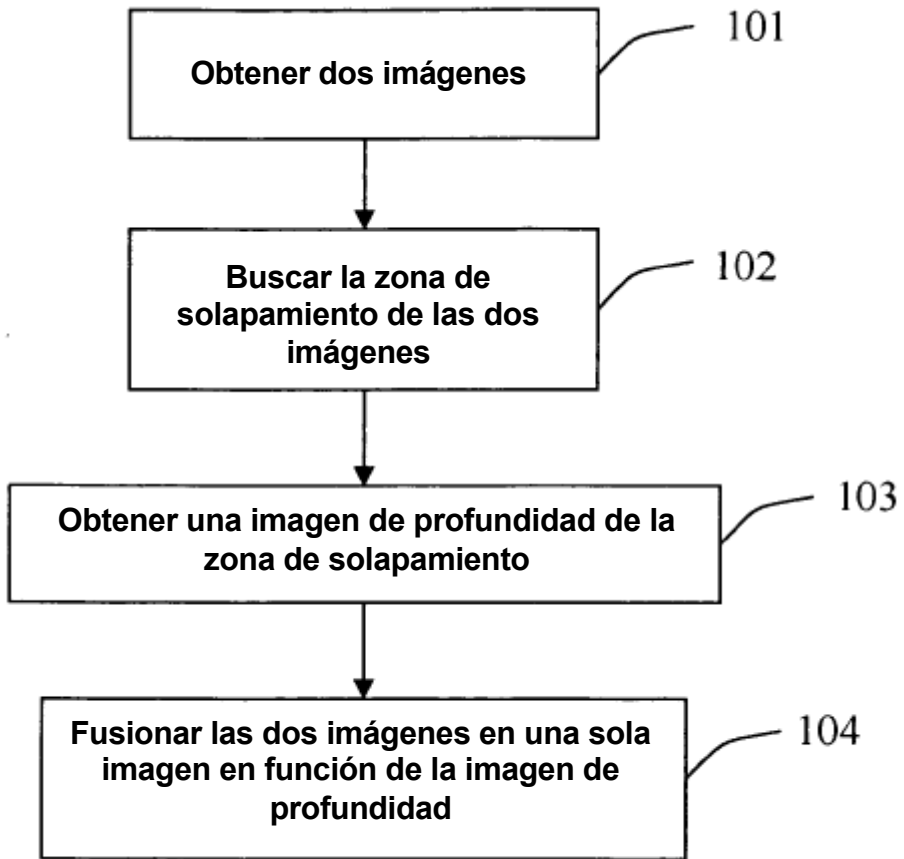


FIG. 1

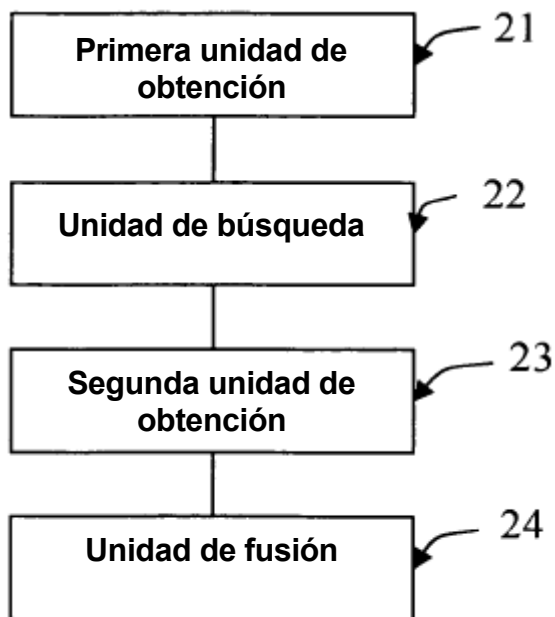


FIG. 2

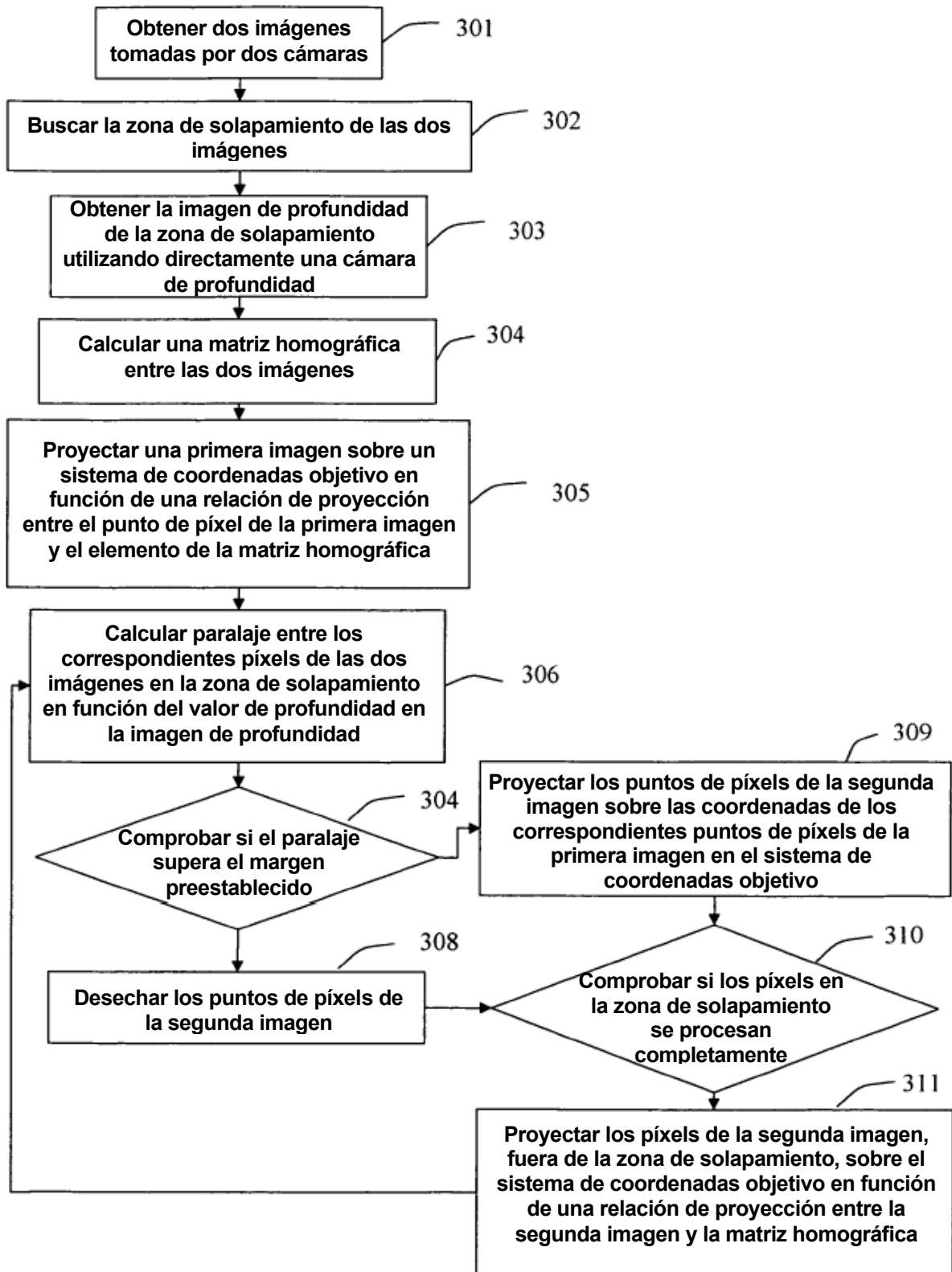


Fig 3

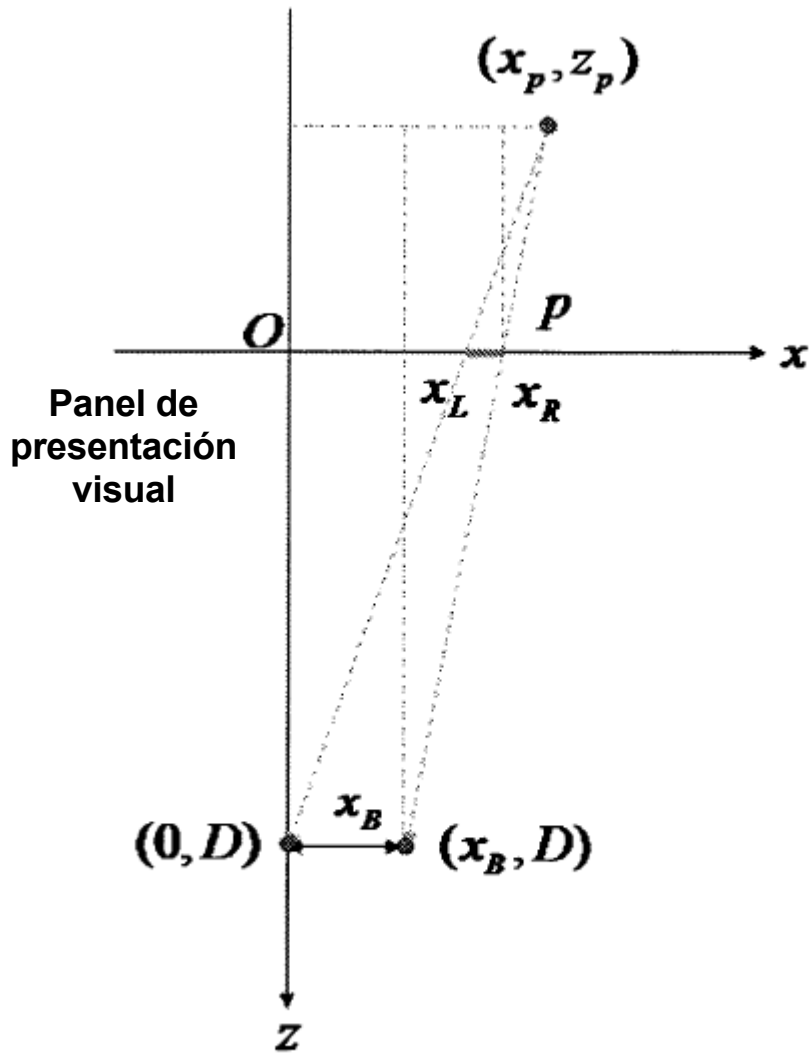


FIG. 4.

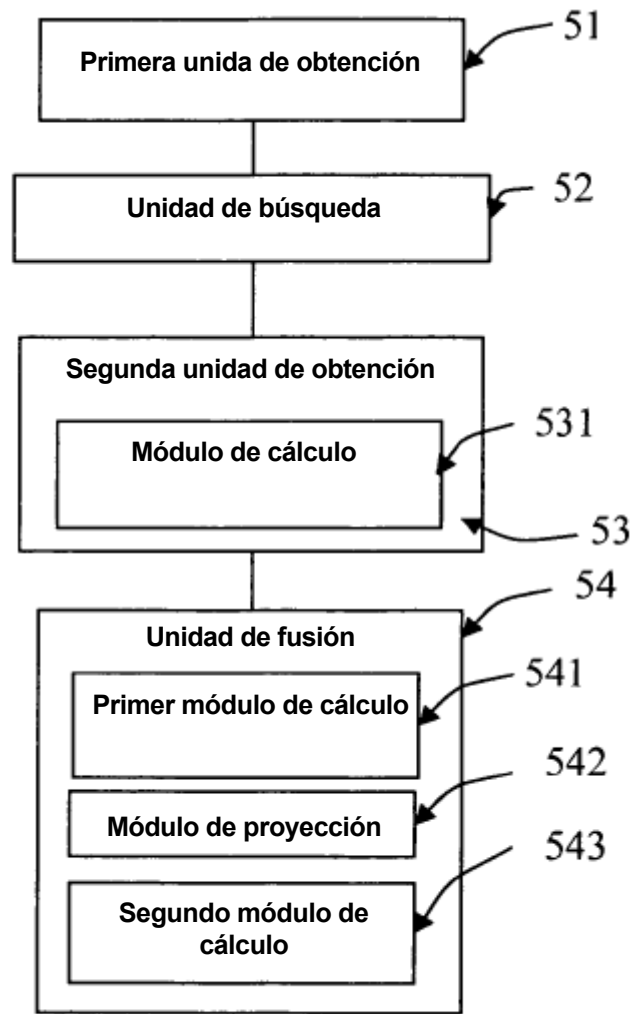


FIG. 5

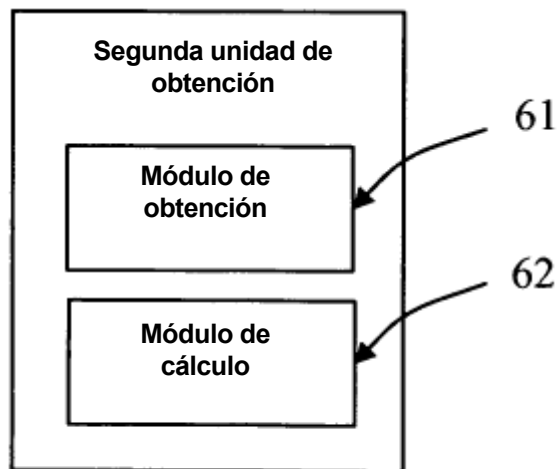


FIG. 6

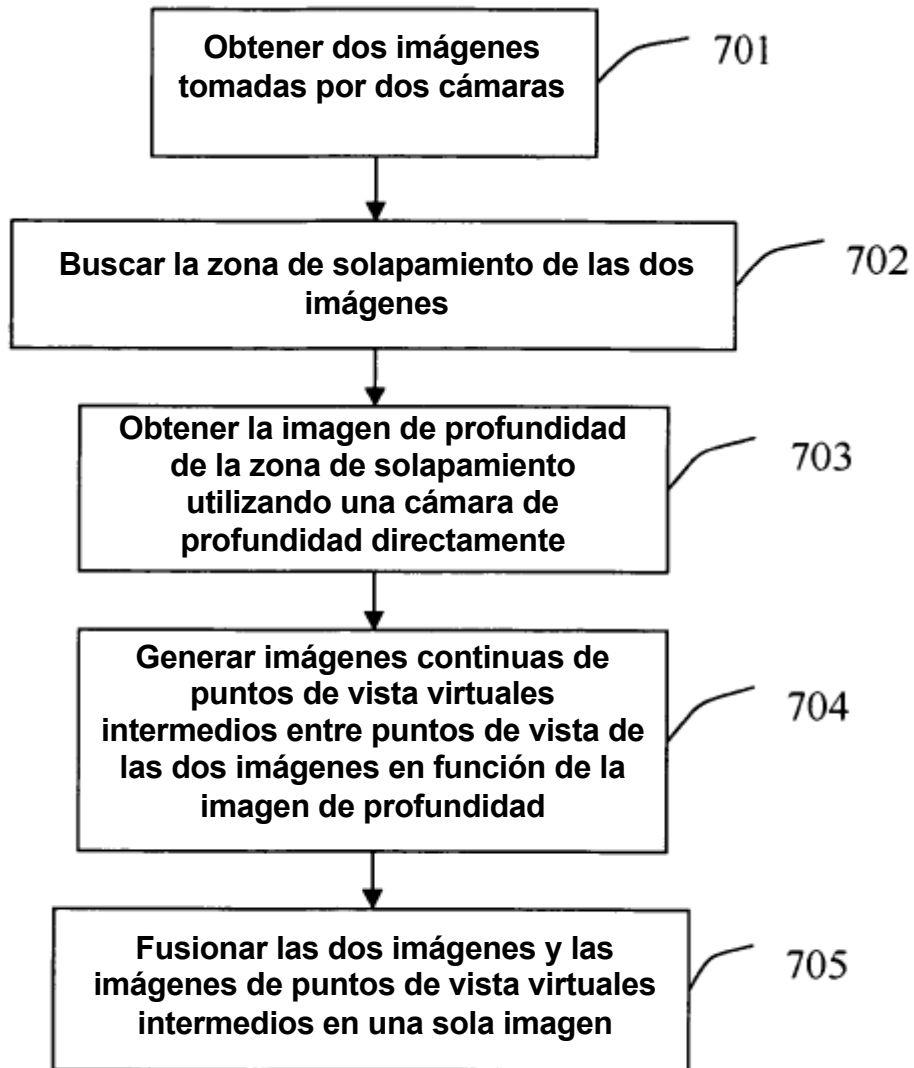


FIG. 7

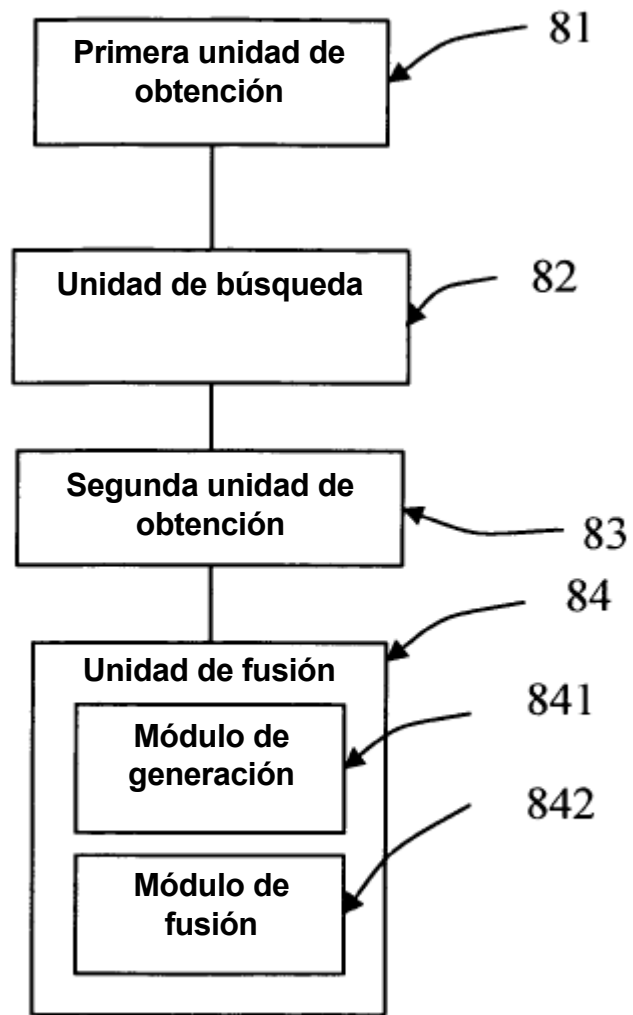


FIG. 8