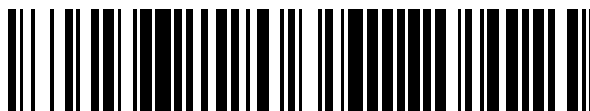


19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 756 677**

51 Int. Cl.:

G06T 7/194 (2007.01)

G06K 9/32 (2006.01)

G06T 7/11 (2007.01)

G06T 7/174 (2007.01)

G06K 9/00 (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

96 Fecha de presentación y número de la solicitud europea: **06.09.2017 E 17001482 (3)**

97 Fecha y número de publicación de la concesión europea: **06.11.2019 EP 3300022**

54 Título: **Aparato de procesamiento de imágenes, procedimiento de procesamiento de imágenes y programa**

30 Prioridad:

26.09.2016 JP 2016187494

28.09.2016 JP 2016190052

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

27.04.2020

73 Titular/es:

CANON KABUSHIKI KAISHA (100.0%)
30-2, Shimomaruko 3-chome, Ohta-ku
Tokyo, JP

72 Inventor/es:

ITAKURA, KINA

74 Agente/Representante:

DURAN-CORRETJER, S.L.P

ES 2 756 677 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

Aparato de procesamiento de imágenes, procedimiento de procesamiento de imágenes y programa

5 ESTADO DE LA TÉCNICA ANTERIOR

Sector técnico de la invención

La presente invención se refiere a una técnica para extraer un área de primer plano de una imagen capturada.

Descripción de la técnica relacionada

Convencionalmente, como un procedimiento para extraer un área de primer plano de una imagen capturada, existe un procedimiento diferencial de segundo plano. En el procedimiento diferencial de segundo plano, se extrae un área de primer plano basándose en la diferencia para cada píxel entre el valor de píxel de una imagen capturada en la que están fotografiados un objeto de primer plano y un segundo plano y el valor de píxel de una imagen de segundo plano en la que sólo está fotografiado el segundo plano. En este momento, en el caso en que una imagen capturada con antelación bajo condiciones específicas se utiliza como una imagen de segundo plano, con la condición de que el segundo plano cambie debido a un cambio en la luz solar que acompaña a un lapso de tiempo, existe un problema en que la precisión de la extracción de un área de primer plano se reduce.

Con respecto al problema descrito anteriormente, la Patente japonesa a inspección pública .número 2012-104053 describe la extracción de un área de primer plano utilizando una imagen de segundo plano creada basándose en una pluralidad de imágenes cuyos tiempos de captura de imagen son diferentes.

Además, la Patente japonesa a inspección pública número. 2014-230180 describe la extracción de un área de primer plano utilizando una imagen de segundo plano creada basándose en una pluralidad de imágenes capturadas desde diferentes puntos de visión en el mismo instante de tiempo.

Otra técnica relacionada es Xu, M. et al., "A Multiview Approach to Robust Detection in the Presence of Cast Shadows", Actas de la 6ª Conferencia Internacional del IEEE sobre Imágenes y Gráficos, páginas 494-499, 2011, y la Patente US 2010/0153449 A1.

35 CARACTERÍSTICAS DE LA INVENCION

No obstante, con la técnica convencional, existe la posibilidad de que se reduzca la precisión de la extracción de un área de primer plano. Por ejemplo, en la Patente japonesa a inspección pública .número 2012-104053, en el caso en el que el área de primer plano no se mueve y permanece estática, se determina erróneamente que esta área de primer plano es un segundo plano y, por tanto, no es posible crear la imagen de segundo plano con alta precisión. Debido a esto, existe un problema en el que se reduce la precisión de la extracción de un área de primer plano.

Además, en la Patente japonesa a inspección pública .número 2014-230180, se crea una imagen de segundo plano al componer la información del segundo plano que no se ve desde un único punto de visión mediante información obtenida desde otro punto de visión, pero en un área o similar donde se superpone una pluralidad de áreas de primer plano que existen dentro de una escena no es posible crear una imagen de segundo plano con alta precisión. Debido a esto, existe un problema en el que se reduce la precisión de la extracción de un área de primer plano.

En consecuencia, la presente invención da a conocer un aparato de procesamiento de imágenes para extraer un área de primer plano con alta precisión en vista del problema descrito anteriormente.

La presente invención en su primer aspecto da a conocer un aparato de procesamiento de imágenes en las reivindicaciones 1 a 15. La presente invención en su segundo aspecto da a conocer un procedimiento para el procesamiento de imágenes en la reivindicación 16. La presente invención en su tercer aspecto da a conocer un programa en la reivindicación 17.

Otras características de la presente invención serán evidentes a partir de la siguiente descripción de realizaciones a modo de ejemplo, con referencia a los dibujos adjuntos.

60 BREVE DESCRIPCIÓN DE LOS DIBUJOS

La figura 1 es un diagrama de bloques que muestra una configuración de hardware de un aparato de procesamiento de imágenes en las realizaciones primera a tercera;

la figura 2 es un diagrama de bloques que muestra una configuración de función de un aparato de procesamiento de imágenes en la primera realización;

la figura 3 es un diagrama de flujo que muestra un flujo de un procesamiento para extraer un área de primer plano en la primera realización;

5 la figura 4 es un diagrama que explica un esquema del procesamiento para extraer un área de primer plano en la primera realización;

la figura 5 es un diagrama que explica la conversión de imagen en la primera realización;

10 la figura 6 es un diagrama que explica los resultados de la primera realización;

la figura 7 es un diagrama de bloques que muestra una configuración de función del aparato de procesamiento de imágenes en la segunda realización;

15 la figura 8 es un diagrama de flujo que muestra un flujo del procesamiento para extraer un área de primer plano en la segunda realización;

la figura 9 es un diagrama que explica un procedimiento de cálculo de continuidad en la segunda realización;

20 la figura 10 es un diagrama que explica los resultados de la segunda realización;

la figura 11 es un diagrama de bloques que muestra una configuración de función del aparato de procesamiento de imágenes en la tercera realización;

25 la figura 12 es un diagrama de flujo que muestra un flujo de procesamiento para extraer un área de primer plano en la tercera realización;

la figura 13 es un diagrama que explica los resultados de la tercera realización;

30 la figura 14 es un diagrama de bloques que muestra una configuración de función de un aparato de procesamiento de imágenes en una cuarta realización;

la figura 15 es un diagrama de flujo que muestra un flujo de procesamiento para extraer un área de primer plano en la cuarta realización;

35 la figura 16 es un diagrama que explica un esquema del procesamiento para extraer un área de primer plano en la cuarta realización;

la figura 17 es un diagrama que explica los resultados de la cuarta realización;

40 la figura 18 es un diagrama de bloques que muestra una configuración de función de un aparato de procesamiento de imágenes en una quinta realización;

la figura 19 es un diagrama que muestra las relaciones de las figuras 19A y 19B;

45 la figura 19A es un diagrama de flujo que muestra un flujo de procesamiento para extraer un área de primer plano en la quinta realización;

la figura 19B es un diagrama de flujo que muestra un flujo de procesamiento para extraer un área de primer plano en la quinta realización; y

50 la figura 20 es un diagrama que explica los resultados de la quinta realización.

DESCRIPCIÓN DE LAS REALIZACIONES

55 A continuación, se explican las realizaciones de la presente invención con referencia a los dibujos. No obstante, las siguientes realizaciones no pretenden limitar la presente invención y todas las combinaciones de las características explicadas en las siguientes realizaciones no son necesariamente indispensables para la solución de la presente invención. La explicación se proporciona adjuntando el mismo símbolo al mismo componente.

60 (Primera realización)

En una primera realización, basándose en una imagen que incluye un primer plano y un segundo plano en una pluralidad de diferentes puntos de visión (en adelante, imagen de referencia), se crea una imagen de segundo plano que no incluye un primer plano en un punto de visión de interés (en adelante, imagen con el primer plano eliminado).

65 A continuación, utilizando la imagen con el primer plano eliminado, se extrae un área de primer plano de una imagen objetivo de procesamiento. En la presente invención, en vista de las circunstancias, se utiliza en la explicación el

término “imagen con el primer plano eliminado”, pero cabe señalar que existe una posibilidad de que esté fotografiada un área de primer plano en parte de la imagen con el primer plano eliminado. Además, el punto de visión de interés puede ser un punto de visión de una cámara seleccionada entre una pluralidad de cámaras o puede ser un punto de visión virtual. El punto de visión virtual es un punto de visión que es establecido libremente (o bajo restricciones fijas) virtualmente por un usuario, independientemente de la posición de la cámara.

<Acerca del esquema de procesamiento para extraer el área de primer plano>

A continuación, se explica un esquema del procesamiento para extraer un área de primer plano en la presente realización utilizando la figura 4. En la presente realización, en primer lugar, se adquieren imágenes 401 de referencia en una pluralidad de diferentes puntos de visión. La imagen de referencia que se adquiere en este caso no tiene por qué ser una imagen que no incluya en absoluto una imagen de primer plano, pero es deseable que sea una imagen capturada en un momento próximo a aquel en que se captura la imagen objetivo de la que se debe extraer un área de primer plano. Se supone que una imagen en el mismo punto de visión que un punto de visión 402 desde el que se ha capturado una imagen que es un objetivo de extracción de un área de primer plano está incluida en una pluralidad de imágenes 401 de referencia a adquirir. En adelante, una imagen de la que se debe extraer un área de primer plano se denomina una imagen (datos) objetivo, las imágenes basadas en la fotografía desde una pluralidad de puntos de visión se denominan imágenes (datos) de referencia, y un punto de visión desde el que se captura una imagen (datos) objetivo se denomina un punto de visión de interés.

A continuación, al convertir la imagen 401 de referencia adquirida en una imagen en el caso en el que la imagen 401 de referencia se ve desde el punto de visión de interés 402 con una superficie del terreno como base para cada punto de visión, se crea una imagen 403 de referencia en el punto de visión de interés. El número de imágenes 403 de referencia a crear en este caso es el mismo que el número de imágenes 401 de referencia. No obstante, los números no tienen que ser necesariamente los mismos en todo momento. En adelante, la imagen 403 de referencia que se obtiene al convertir la imagen 401 de referencia se denomina la imagen 403 de referencia convertida. En la presente realización, se explica principalmente el ejemplo en el que la conversión de una imagen se realiza con la superficie del terreno como base, pero esto no es limitativo. Por ejemplo, es asimismo posible tomar la superficie del agua como base y, en el caso en el que la altura de la superficie del terreno es diferente dependiendo de la posición, puede ser posible asimismo convertir una imagen tomando la altura promediada como base para cada área. Es decir, en la presente realización, al convertir la imagen 401 de referencia basándose en una superficie base predeterminada, se crea la imagen 403 de referencia convertida.

En este caso, el primer plano (objeto) significa un objeto que satisface las condiciones predeterminadas de entre los objetos incluidos en la imagen capturada. A menos que se describan en concreto, en adelante, el objeto de primer plano y el primer plano se utilizan con el mismo significado. Por ejemplo, en el caso en que el objetivo de captura de la imagen es una escena de juego competitivo, tal como un deporte, personas, tal como un jugador y un árbitro, instalaciones, tal como una portería, y equipamiento, tal como un balón, puede ser un objeto de primer plano. Habitualmente, el objeto de primer plano incluye lo que sigue en movimiento en una pluralidad de las imágenes capturadas continuamente a lo largo de una serie temporal. Por otra parte, el segundo plano corresponde al área distinta del objeto de primer plano de la imagen capturada. Por ejemplo, en el caso en el que el objetivo de captura de la imagen es una escena de un juego competitivo, tal como un deporte, el terreno formado por césped y tierra, el suelo de un gimnasio, etc., puede ser el segundo plano. Muchos segundos planos son estáticos casi todo el tiempo en una pluralidad de imágenes capturadas continuamente a lo largo de una serie temporal. No obstante, existe un caso en el que el área en la que hay movimiento, tal como una grada de espectadores, es reconocida como segundo plano.

En la presente realización, se supone un caso en el que el objeto de primer plano tiene una altura desde la superficie del terreno, pero, por otra parte, el segundo plano no tiene una altura desde la superficie del terreno. En consecuencia, al detectar un objeto de primer plano que tiene una altura desde la superficie del terreno utilizando una pluralidad de imágenes 403 de referencia convertidas y eliminando de una imagen objetivo el objeto de primer plano detectado, se crea una imagen con el primer plano eliminado en el punto de visión de interés 402. Concretamente, para la pluralidad de imágenes 403 de referencia convertidas que incluyen la imagen en el punto de visión de interés 402, se calcula para cada píxel un grado de coincidencia entre píxeles de interés y se detecta un píxel cuyo grado de coincidencia es bajo como un píxel del área de la imagen del objeto de primer plano. Es posible denominar el grado de coincidencia con otro término, tal como grado de similitud y grado de diferencia. Tal como se ha descrito anteriormente, la imagen 403 de referencia convertida es una imagen obtenida convirtiendo la imagen 401 de referencia en una imagen en el caso en el que la imagen 401 de referencia se ve desde el punto de visión de interés 402 con la superficie del terreno como base. En consecuencia, las coordenadas de las áreas 405 a 407 en la imagen 401 de referencia, que corresponden a un área 404 de la superficie del suelo que existe en la superficie del terreno y no tiene altura, se convierten respectivamente en las coordenadas de un área 408 que existe en la misma posición en común en todas las imágenes 403 de referencia convertidas. Por otra parte, las coordenadas de las áreas 410 a 412 en la imagen 401 de referencia, que corresponden a un jugador 409 que tiene una altura, se convierten respectivamente en las coordenadas de las áreas 413 a 415, cuyas posiciones son diferentes para diferentes puntos de visión. En consecuencia, en la pluralidad de imágenes 403 de referencia convertidas, el píxel cuyo grado de coincidencia entre píxeles de interés es alto se considera un píxel del área de la imagen del segundo

plano que no tiene altura y un píxel cuyo grado de coincidencia es bajo se considera un píxel del área de la imagen del objeto de primer plano que tiene una altura. De este modo, se crea una imagen con el primer plano eliminado. A continuación, al comparar la imagen con el primer plano eliminado creada en el punto de visión de interés 402 con la imagen objetivo, se extrae un área de primer plano de la imagen objetivo.

Lo anterior es el esquema del procesamiento que se realiza en la presente realización. La imagen objetivo no está limitada al ejemplo anteriormente descrito y es posible utilizar una variedad de imágenes, tales como datos de una imagen capturada por una cámara de supervisión. Además, aquí, se explica el caso en el que la imagen en el punto de visión de interés está incluida en las imágenes 401 de referencia, pero también es posible aplicar la presente realización al caso en el que una imagen en el punto de visión de interés no está incluida en las imágenes de referencia, y más adelante se describirá un procedimiento de procesamiento concreto.

<Acerca de la configuración de hardware del aparato de procesamiento de imágenes>

A continuación, se describe una configuración de hardware de un aparato de procesamiento de imágenes de la presente realización. La figura 1 es un diagrama de bloques que muestra un ejemplo de una configuración de hardware de un aparato de procesamiento de imágenes de la presente realización. Un aparato 100 de procesamiento de imágenes de la presente realización incluye una CPU 101, una RAM 102, una ROM 103, un dispositivo 104 de almacenamiento secundario, una interfaz 105 de entrada y una interfaz 106 de salida, y estos componentes están conectados entre sí a través de un bus 107 de sistema. Además, el aparato 100 de procesamiento de imágenes está conectado a un dispositivo 108 de almacenamiento externo a través de la interfaz 105 de entrada y está conectado con el dispositivo 108 de almacenamiento externo y un dispositivo 109 de visualización a través de la interfaz 106 de salida.

La CPU 101 ejecuta programas almacenados en la ROM 103 utilizando la RAM 102 como una memoria de trabajo y controla de manera centralizada cada componente del aparato 100 de procesamiento de imágenes a través del bus 107 de sistema. Debido a esto, se realizan diversos tipos de procesamiento, que se describirán más adelante.

El dispositivo 104 de almacenamiento secundario es un dispositivo de almacenamiento que almacena diversos tipos de datos manipulados en el aparato 100 de procesamiento de imágenes, y en la presente realización se utiliza un HDD. Es posible que la CPU 101 escriba datos en el dispositivo 104 de almacenamiento secundario y lea datos almacenados en el dispositivo 104 de almacenamiento secundario a través del bus 107 de sistema. En cuanto al dispositivo 104 de almacenamiento secundario, es posible utilizar una variedad de dispositivos de almacenamiento, tal como un disco duro óptico y una memoria flash, además de un HDD.

La interfaz 105 de entrada es una interfaz de bus serie, por ejemplo, tal como un USB e IEEE 1394, y la entrada de datos, una instrucción, etc., al aparato 100 de procesamiento de imágenes desde un dispositivo externo se realiza a través de la interfaz 105 de entrada. El aparato 100 de procesamiento de imágenes obtiene datos desde el dispositivo 108 de almacenamiento externo (por ejemplo, medio de almacenamiento tal como un disco duro, tarjeta de memoria, tarjeta de CF, tarjeta de SD y memoria USB) a través de la interfaz 105 de entrada. Es asimismo posible conectar un dispositivo de entrada (no mostrado esquemáticamente) para entradas de usuario, tal como un ratón y un teclado, a la interfaz 105 de entrada. La interfaz 106 de salida incluye un terminal de salida de vídeo, por ejemplo, tal como DVI y HDMI (marca registrada), además de la interfaz de bus serie, tal como USB e IEEE 1394, similar a la interfaz 105 de entrada. Se entregan datos desde el aparato 100 de procesamiento de imágenes a un dispositivo externo a través de la interfaz 106 de salida. El aparato 100 de procesamiento de imágenes produce la visualización de una imagen entregando una imagen procesada y similar al dispositivo 109 de visualización (diversos tipos de dispositivo de visualización de imágenes, tal como una pantalla de cristal líquido) a través de la interfaz 106 de salida. Existen componentes del aparato 100 de procesamiento de imágenes distintos de los descritos anteriormente, pero no son el propósito principal de la presente invención y, por tanto, se omite su explicación.

<Acerca del procesamiento para extraer el área de primer plano>

A continuación, utilizando la figura 2 y la figura 3, se explica el procesamiento para extraer un área de primer plano que es realizado por el aparato 100 de procesamiento de imágenes en la presente realización. La figura 2 es un diagrama de bloques que muestra una configuración de función del aparato 100 de procesamiento de imágenes y la figura 3 es un diagrama de flujo que muestra un flujo del procesamiento para extraer un área de primer plano. La CPU 101 del aparato 100 de procesamiento de imágenes funciona como cada componente mostrado en la figura 2 y realiza una serie de procesamiento mostrada en la figura 3 mediante la ejecución de programas almacenados en la ROM 103 utilizando la RAM 102 como memoria de trabajo. Todo el procesamiento mostrado a continuación no tiene por qué ser realizado por la CPU 101 y puede ser posible asimismo conformar el aparato 100 de procesamiento de imágenes de modo que parte o la totalidad del procesamiento se realice mediante uno o una pluralidad de circuitos de procesamiento distintos de la CPU 101.

A continuación, se explica un flujo de procesamiento que es realizado por cada componente. En la etapa S301, una unidad 201 de adquisición de imágenes objetivo adquiere una imagen objetivo desde el dispositivo 108 de almacenamiento externo a través de la interfaz 105 de entrada, o desde el dispositivo 104 de almacenamiento

secundario. Tal como se ha descrito anteriormente, la imagen objetivo es una imagen que es un objetivo del que se extrae un área de primer plano. Además, la unidad 201 de adquisición de imágenes objetivo determina el punto de visión de una cámara que ha capturado la imagen objetivo que será el punto de visión de interés. Además, en la presente realización, se explica el caso en el que existe una imagen objetivo, pero el número de imágenes objetivo puede ser de dos o más. Además, la unidad 201 de adquisición de imágenes objetivo adquiere parámetros de una cámara (en adelante, parámetros de la cámara) que ha capturado la imagen objetivo junto con la imagen objetivo. En este caso, los parámetros de la cámara son parámetros que permiten un cálculo para proyectar un punto en el espacio tridimensional sobre una imagen capturada por una cámara e incluyen parámetros externos que representan la posición y la postura de una cámara y parámetros internos que representan la longitud focal y el centro óptico. Puede ser posible asimismo utilizar los valores medidos y los valores de diseño almacenados con antelación en la memoria como parámetros de la cámara. La unidad 201 de adquisición de imágenes objetivo entrega la imagen objetivo a una unidad 207 de extracción del primer plano y entrega los parámetros de la cámara a una unidad 203 de conversión de imágenes.

En la etapa S302, una unidad 202 de adquisición de imágenes de referencia adquiere una pluralidad de imágenes de referencia en una pluralidad de puntos de visión diferentes desde el dispositivo 108 de almacenamiento externo a través de la interfaz 105 de entrada, o desde el dispositivo 104 de almacenamiento secundario. En este caso, la imagen de referencia es una imagen basada en fotografías en un entorno (tiempo meteorológico, zona horaria, etc.) sustancialmente igual que el entorno en el momento de captura de la imagen objetivo. Tal como se ha descrito anteriormente, la imagen de referencia que se adquiere en esta etapa no tiene por qué ser una imagen de segundo plano que no incluya en absoluto una imagen de primer plano. Además, en la imagen de referencia que se adquiere en la etapa S302, puede estar incluida o puede no estar incluida una imagen basada en el punto de visión de interés.

En la presente realización, se crea una imagen de referencia en cada punto de visión al realizar un procesamiento de filtro utilizando un valor medio para una pluralidad de imágenes correspondientes a una pluralidad de tiempos diferentes adquiridas al capturar de forma continua las imágenes de una escena desde el mismo punto de visión a lo largo de una serie temporal. No obstante, el procedimiento de crear una imagen de referencia no está limitado a este procedimiento. Por ejemplo, también puede ser posible crear una imagen de referencia utilizando otro filtro, tal como un filtro de valor promedio o crear una imagen de referencia al realizar un procesamiento de agrupación para una pluralidad de imágenes. Además, también puede ser posible utilizar una imagen de referencia adquirida al realizar una captura de imagen con antelación en la situación en la que no existe un objeto de primer plano para cada punto de visión.

La unidad 202 de adquisición de imágenes de referencia adquiere los parámetros de la cámara correspondientes a cada imagen de referencia junto con la imagen de referencia. Además, la unidad 202 de adquisición de imágenes de referencia almacena cada imagen de referencia en asociación con un número para distinguir el punto de visión de una cámara de otro (en adelante, número del punto de visión de la cámara) para distinguir una imagen de referencia de otra en la pluralidad de imágenes de referencia. La unidad 202 de adquisición de imágenes de referencia entrega la imagen de referencia y los parámetros de la cámara a la unidad 203 de conversión de imágenes y entrega solo la imagen de referencia a una unidad 206 de corrección. 2

En la etapa S303, la unidad 203 de conversión de imágenes convierte la imagen de referencia adquirida desde la unidad 202 de adquisición de imágenes de referencia en una imagen en el caso en el que la imagen de referencia se ve desde el punto de visión de interés utilizando los parámetros de la cámara adquiridos desde la unidad 201 de adquisición de imágenes objetivo y la unidad 202 de adquisición de imágenes de referencia. Concretamente, al realizar una conversión de proyección para cada imagen de referencia con la superficie del terreno como base, se obtiene la imagen en el caso en el que la imagen de referencia se ve desde el punto de visión de interés. La imagen (datos) de referencia obtenida por la conversión de imagen en esta etapa se denomina imagen (datos) de referencia convertida. En este caso, el procedimiento de conversión de imagen en esta etapa se explica utilizando la figura 5.

Tal como se muestra en la figura 5, en el caso en el que un punto 501 en el espacio tridimensional se proyecta sobre una imagen de una cámara 502, un punto 504 que es una intersección de una línea recta que conecta el punto 501 y la cámara 502, y un plano 503 de la imagen es una imagen proyectada del punto 501 en el espacio tridimensional sobre el plano 503 de la imagen. De modo similar, en el caso de una cámara 505 (cámara con un punto de visión diferente) existente en una posición diferente de la de la cámara 502, un punto 507 que es una intersección de una línea recta que conecta el punto 501 y la cámara 505, y un plano 506 de la imagen es una imagen proyectada del punto 501 sobre el plano 506 de la imagen. Aquí, se analiza un caso en el que todos los puntos del espacio tridimensional proyectados sobre el plano 503 de la imagen y el plano 506 de la imagen que incluye el punto 501 existen en el mismo plano, que es la superficie del terreno. En este caso, al utilizar una matriz homográfica 3×3 H_{01} calculada por los parámetros de cámara de la cámara 502 y de la cámara 505, las coordenadas arbitrarias (u_0, v_0) del plano 503 de la imagen se convierten en las coordenadas (u_1, v_1) del plano 506 de la imagen mediante la expresión (1).

$$\begin{bmatrix} u_1 \\ v_1 \\ 1 \end{bmatrix} = H_{01} \begin{bmatrix} u_0 \\ v_0 \\ 1 \end{bmatrix} \quad \dots \text{expresión (1)}$$

En la etapa S303, se realiza para cada imagen de referencia una conversión de proyección tomando la cámara con el punto de visión correspondiente a la imagen de referencia adquirida desde la unidad 202 de adquisición de imágenes de referencia para que sea la cámara 502 descrita anteriormente y la cámara con el punto de visión de interés determinado por la unidad 201 de adquisición de imágenes objetivo para que sea la cámara 505. Debido a esto, el número de imágenes de referencia convertidas adquiridas en esta etapa es el mismo que el número de imágenes de referencia adquiridas por la unidad 202 de adquisición de imágenes de referencia. Además, cada una de las imágenes de referencia convertidas es almacenada en asociación con el número del punto de visión de cada imagen de referencia adquirida por la unidad 202 de adquisición de imágenes de referencia. La unidad 203 de conversión de imágenes entrega la imagen de referencia convertida a una unidad 204 de cálculo del grado de coincidencia y a la unidad 206 de corrección.

En la etapa S304, la unidad 203 de conversión de imágenes determina que una imagen entre las imágenes de referencia adquiridas desde la unidad 202 de adquisición de imágenes de referencia, que corresponde al punto de visión más próximo al punto de visión de la cámara (punto de visión de interés) desde el que se captura la imagen objetivo, es una imagen de referencia de base. Concretamente, se calcula para cada punto de visión una distancia entre las coordenadas (X_0, Y_0, Z_0) del punto de visión de interés y las coordenadas (X_i, Y_i, Z_i) del punto de visión correspondiente a la imagen de referencia adquirida desde la unidad 202 de adquisición de imágenes de referencia. En este caso, i representa el número del punto de visión y se tiene que $1 \leq i < \text{número de puntos de visión} + 1$. Entonces, se detecta el punto de visión (punto de visión base) cuya distancia calculada es la más corta y la imagen (datos) de referencia correspondiente al punto de visión base se toma como la imagen (datos) de referencia base. La unidad 203 de conversión de imágenes entrega el número del punto de visión correspondiente a la imagen de referencia de base a la unidad 204 de cálculo del grado de coincidencia y a la unidad 206 de corrección. En la presente realización, el número del punto de visión correspondiente a la imagen de referencia de base se denomina el número del punto de visión base. Puede ser posible asimismo que el punto de visión de interés y el punto de visión de la imagen de referencia de base coincidan entre sí perfectamente.

En la etapa S305, la unidad 204 de cálculo del grado de coincidencia determina un píxel de interés en la imagen de referencia convertida, que será el objetivo de la determinación del grado de coincidencia del valor de píxel en la pluralidad de imágenes de referencia convertidas. En la presente realización, en primer lugar, se selecciona el píxel superior-izquierdo de la imagen de referencia de base como el píxel de interés y, después de esto, se seleccionan secuencialmente píxeles sin procesar como el píxel de interés. Siempre que la determinación del grado de coincidencia del valor de píxel en la pluralidad de imágenes de referencia convertidas se realice para todos los píxeles de la imagen de referencia de base, se puede determinar el píxel de interés en cualquier orden. Además, en la presente realización, el ejemplo del caso se explica principalmente cuando los datos relacionados con el grado de coincidencia se obtienen para todos los píxeles de la imagen de referencia de base, pero el ejemplo no se limita a esto. Por ejemplo, en el caso en el que se determina con antelación el área que es el objetivo de la extracción de un objeto de primer plano, es suficiente obtener datos relacionados con el grado de coincidencia solo para los píxeles que pertenecen al área. Por ejemplo, en el caso en el que no es necesario extraer el objeto de primer plano de la grada de espectadores de fútbol, no es necesario realizar el procesamiento para obtener datos relacionados con el grado de coincidencia para el área de la grada de espectadores.

En la etapa S306, la unidad 204 de cálculo del grado de coincidencia calcula el grado de coincidencia en el píxel de interés entre la imagen de referencia convertida (imagen de referencia de base) correspondiente al número del punto de visión base y otra imagen de referencia convertida utilizando la pluralidad de imágenes de referencia convertidas adquiridas desde la unidad 203 de conversión de imágenes. A continuación, se explica concretamente el procedimiento de cálculo del grado de coincidencia.

En primer lugar, la unidad 204 de cálculo del grado de coincidencia adquiere un valor de píxel $B_j(u_2, v_2)$ de una pluralidad de imágenes de referencia convertidas en coordenadas (u_2, v_2) del píxel de interés en la imagen de referencia de base. En este caso, j representa un subíndice para distinguir una imagen de referencia convertida de otra en la pluralidad de imágenes de referencia convertidas y la unidad 204 de cálculo del grado de coincidencia adquiere valores de píxel en el número correspondiente al número de imágenes de referencia convertidas. A continuación, la unidad 204 de cálculo del grado de coincidencia calcula un valor medio de todos los valores de píxel adquiridos. Este valor medio se utiliza como un valor base M en el momento del cálculo del grado de coincidencia. El valor base no está limitado a esto y también puede ser posible utilizar un valor arbitrario, tal como un valor promedio, que refleje la naturaleza estadística de una pluralidad de valores de píxel como un valor base.

A continuación, la unidad 204 de cálculo del grado de coincidencia calcula el grado de coincidencia en el píxel de interés a partir de la expresión (2) utilizando un valor de píxel $B_0(u_2, v_2)$ del píxel de interés en la imagen de referencia convertida (imagen de referencia de base) correspondiente al número del punto de visión base y un valor base calculado $M(u_2, v_2)$.

$$D(u_2, v_2) = \sum_{k=1}^3 (B_0^k(u_2, v_2) - M^k(u_2, v_2))^2 \dots \text{expresión (2)}$$

En este caso, k representa un subíndice para identificar los tres canales de RGB. Un grado de coincidencia D que se calcula mediante la expresión (2) se hace menor en el caso en el que las fluctuaciones en el valor de píxel en la pluralidad de imágenes de referencia convertidas son menores. El grado de coincidencia que se utiliza no está limitado a esto y también puede ser posible utilizar un valor arbitrario que indica una diferencia entre píxeles. Por ejemplo, también puede ser posible utilizar la suma total de diferencias entre el valor de píxel $B_0(u_2, v_2)$ del píxel de interés en la imagen de referencia de base y el valor de píxel del píxel de interés en otra imagen de referencia convertida como el grado de coincidencia. Es decir, la unidad 204 de cálculo del grado de coincidencia obtiene los datos relacionados con el grado de coincidencia del valor de píxel de la pluralidad de imágenes de referencia convertidas en las etapas S305 a S307. En la presente realización, se muestra el ejemplo en el que los datos relacionados con el grado de coincidencia del píxel de interés en la imagen de referencia de base se calcula utilizando los valores de píxel de todas las imágenes de referencia convertidas, pero el ejemplo no está limitado a esto. Por ejemplo, también puede ser posible obtener los datos relacionados con el grado de coincidencia al no utilizar el valor de píxel de la imagen de referencia convertida en la que el píxel de interés no está fotografiado entre la pluralidad de imágenes de referencia convertidas. Además, en la presente realización, se muestra el ejemplo en el que los datos relacionados con el grado de coincidencia se calculan para cada píxel, pero también puede ser posible calcular los datos relacionados con el grado de coincidencia para cada área, tal como un bloque. En el caso en el que el grado de coincidencia se calcula para cada píxel, es posible mejorar adicionalmente la precisión de extracción de un objeto de primer plano y en el caso en el que el grado de coincidencia se calcula para cada área, es posible reducir la carga del procesamiento de extracción de un objeto de primer plano.

En la etapa S307, la unidad 204 de cálculo del grado de coincidencia determina si el procesamiento en la etapa S305 y la etapa S306 se ha realizado para todos los píxeles de la imagen de referencia convertida. En el caso en el que los resultados de la determinación en la etapa S307 son afirmativos, la unidad 204 de cálculo del grado de coincidencia entrega los grados de coincidencia calculados de todos los píxeles a una unidad 205 de determinación de la corrección y entrega el valor base calculado a la unidad 206 de corrección y el procesamiento avanza a la etapa S308. Por otra parte, en el caso en el que los resultados de la determinación en la etapa S307 son negativos, el procesamiento vuelve a la etapa S305.

En la etapa S308, la unidad 205 de determinación de la corrección inicializa el mapa de indicadores, es decir, establece los valores de píxel de todos los píxeles del mapa de indicadores a 0. El mapa de indicadores que se inicializa en esta etapa se utiliza para determinar el píxel que es el objetivo del procesamiento de corrección en el momento de la corrección del píxel de la imagen de referencia convertida (imagen de referencia de base) correspondiente al número del punto de visión base en la etapa S311. En este mapa de indicadores, 1 se sustituye por el valor de píxel correspondiente al píxel del objetivo del procesamiento de corrección y 0 se sustituye para el valor de píxel correspondiente al píxel que no es el objetivo del procesamiento de corrección. Mediante la inicialización en esta etapa, todos los píxeles de la imagen de referencia convertida correspondiente al número del punto de visión base se ajustan a los que no son el objetivo del procesamiento de corrección.

En la etapa S309, la unidad 205 de determinación de la corrección actualiza el mapa de indicadores basándose en el grado de coincidencia adquirido desde la unidad 204 de cálculo del grado de coincidencia. Concretamente, la unidad 205 de determinación de la corrección cambia el valor de píxel del mapa de indicadores a 1, que corresponde al píxel considerado que tiene una alta probabilidad de ser el píxel del área de la imagen del objeto de primer plano en la imagen de referencia convertida (imagen de referencia de base) correspondiente al número del punto de visión base. En la presente realización, se determina que el píxel cuyo grado de coincidencia D es mayor o igual que un valor umbral determinado con antelación tiene una alta probabilidad de ser el píxel del objeto de primer plano debido a que el grado de coincidencia entre el píxel de la imagen de referencia de base y el píxel de otra imagen de referencia convertida es bajo. Por otra parte, se determina que el píxel cuyo grado de coincidencia D es menor que el valor umbral tiene una alta probabilidad de ser el píxel del área de la imagen del segundo plano debido a que el grado de coincidencia entre el píxel de la imagen de referencia de base y el píxel de otra imagen de referencia convertida es alto. El valor umbral que se utiliza en esta etapa se determina basándose en el valor máximo o similar del valor de píxel y el valor umbral se determina utilizando un valor menor que el 20 % del valor máximo, por ejemplo, un valor arbitrario dentro de un intervalo del 1 % al 5 % del valor máximo. Es decir, en el caso en el que se toma un valor arbitrario como a, en la expresión (2), la suma de los cuadrados de las diferencias se utiliza como el grado de coincidencia y, por tanto, el valor umbral será a x a x 3. En el caso en el que la suma total de las diferencias se utiliza como el grado de coincidencia, el valor umbral será a x 3. Tomando el valor umbral como un valor variable según se ha descrito anteriormente, la precisión de la extracción de un objeto de primer plano mejora aún más. No obstante, también puede ser posible ajustar el valor umbral a un valor fijo. En la presente realización, la determinación de si el píxel de interés es el píxel del área de la imagen del objeto de primer plano se realiza para cada píxel. No obstante, esto no está limitado y también puede ser posible realizar la determinación para cada área, tal como un bloque. Al hacer esto, es posible reducir la carga de procesamiento relacionada con la extracción de un objeto de primer plano. La unidad 205 de determinación de la corrección entrega el mapa de indicadores para los que se ha completado la actualización a la unidad 206 de corrección.

En la etapa S310, la unidad 206 de corrección determina un píxel de interés en la imagen de referencia de base. En la presente realización, en primer lugar, el píxel superior izquierdo de la imagen de referencia de base se selecciona como el píxel de interés y, después de esto, los píxeles sin procesar son seleccionados secuencialmente como el píxel de interés. Siempre que la actualización (etapa S311) del valor de píxel basada en el mapa de indicadores se realice para todos los píxeles de la imagen de referencia de base, el píxel de interés puede ser determinado en cualquier orden. Además, no se requiere necesariamente realizar el procesamiento en la etapa S310 para todos los píxeles de la imagen de referencia de base. Por ejemplo, en el caso en el que el área en la que el objeto de primer plano no puede existir es seleccionada con antelación por un usuario, no es necesario generar un mapa de indicadores para el área seleccionada y no es necesario realizar el procesamiento en la etapa S310.

En la etapa S311, la unidad 206 de corrección corrige el valor de píxel del píxel de interés en la imagen de referencia convertida correspondiente al número del punto de visión base basándose en el mapa de indicadores obtenido desde la unidad 205 de determinación de la corrección. En la presente realización, en el caso en el que el valor de píxel del mapa de indicadores correspondiente al píxel de interés en la imagen de referencia convertida correspondiente al número del punto de visión base es 1, el valor de píxel del píxel de interés se sustituye con el valor base calculado por la unidad 204 de cálculo del grado de coincidencia. Por otra parte, en el caso en el que el valor de píxel del mapa de indicadores correspondiente al píxel de interés en la imagen de referencia convertida correspondiente al número del punto de visión base es 0, el valor de píxel del píxel de interés no cambia. El valor base es, por ejemplo, un valor medio o un valor promedio de los valores de píxel de una pluralidad de imágenes de referencia, o un valor arbitrario que refleja otra naturaleza estadística. El procedimiento de corrección de un valor de píxel no está limitado a esto, y puede ser posible asimismo utilizar otro procedimiento, tal como un procedimiento de sustitución de un valor de píxel con otro valor de píxel de una imagen de segundo plano correspondiente al punto de visión adyacente al punto de visión base.

En la etapa S312, la unidad 206 de corrección determina si el procesamiento en la etapa S310 y la etapa 311 se ha realizado para todos los píxeles de la imagen de referencia de base. En el caso en el que los resultados de la determinación en la etapa S312 son afirmativos, la unidad 206 de corrección entrega la imagen de referencia de base para la que se ha completado la corrección a la unidad 207 de extracción del primer plano y el procesamiento avanza a la etapa S313. Por otra parte, en el caso en el que los resultados de la determinación son negativos, el procesamiento vuelve a la etapa S310.

En la etapa S313, la unidad 207 de extracción del primer plano extrae un objeto de primer plano de una imagen objetivo (tomada como I) utilizando una imagen de referencia de base (tomada como imagen con el primer plano eliminado I_b) adquirida de la unidad 206 de corrección, para la que se ha completado la corrección. Concretamente, tal como se expresa mediante la expresión (3), la suma de los cuadrados de las diferencias se calcula para cada píxel entre la imagen con el primer plano eliminado I_b y la imagen objetivo I y considerando el píxel cuya suma de los cuadrados de las diferencias es mayor o igual que un valor umbral como el píxel del área de la imagen del objeto de primer plano, se crea una imagen I_f en la que se extrae el objeto de primer plano. La imagen I_f es una imagen binaria y 1 es sustituido por el valor de píxel correspondiente al píxel del área de la imagen del objeto de primer plano y 0 es sustituido por el valor de píxel correspondiente al píxel del área de la imagen del primer plano.

$$I_f(x, y) = \begin{cases} 1 & \text{si } \sum_{k=1}^3 (I_b^k(x, y) - I^k(x, y))^2 \geq Th \\ 0 & \text{si } \sum_{k=1}^3 (I_b^k(x, y) - I^k(x, y))^2 < Th \end{cases} \quad \dots \text{expresión (3)}$$

En este caso, Th representa un valor umbral y k representa un subíndice para identificar los tres canales de RGB. El valor umbral que se utiliza en este caso se determina basándose en el valor máximo o similar del valor de píxel y también puede ser posible encontrar el valor umbral utilizando un valor menor que el 20 % del valor máximo del valor de píxel, por ejemplo, utilizando un valor arbitrario dentro de un intervalo del 1 % al 5 % del valor máximo. El procedimiento de encontrar el valor umbral es el mismo que en el caso de la expresión (2). La unidad 207 de extracción del primer plano entrega la imagen I_f creada al dispositivo 104 de almacenamiento secundario, al dispositivo 108 de almacenamiento externo y al dispositivo 109 de visualización y, posteriormente, se completa la serie de procesamiento. Lo anterior es el procesamiento para extraer un área de primer plano, que es realizado por el aparato 100 de procesamiento de imágenes en la presente realización.

<Acerca de los resultados de la presente realización>

En lo siguiente, se explican los resultados de la presente realización utilizando la figura 6. En la figura 6, una imagen 601 es una imagen de segundo plano en un punto de visión 602, que se crea basándose en una pluralidad de imágenes capturadas continuamente a lo largo de una serie temporal de acuerdo con un procedimiento convencional. En la imagen 601 de segundo plano, se fotografían un objeto 603 de primer plano (portero), un objeto 604 de primer plano (portería), etc. La razón es que los objetos 603 y 604 de primer plano continúan existiendo en la misma posición y no se mueven mientras se capturan imágenes continuas para crear una imagen de referencia, y como resultado de esto, los objetos 603 y 604 de primer plano se consideran erróneamente como el segundo plano en el momento de crear una imagen de segundo plano. En el caso en el que se extrae el área de primer plano de

una imagen objetivo 605 utilizando la imagen 601 de segundo plano, se adquiere una imagen 606 de primer plano. En la imagen 606 de primer plano, los objetos de primer plano que se están desplazando, distintos de los objetos 603 y 604 de primer plano, están casi extraídos. No obstante, el área de los objetos 603 y 604 de primer plano que son estáticos no se ha extraído.

Además, una imagen 607 es una imagen de segundo plano en el punto de visión 602, que se crea basándose en una pluralidad de imágenes capturadas desde una pluralidad de puntos de visión diferentes en el mismo instante de tiempo que aquel en que es capturada la imagen objetivo 605, de acuerdo con el procedimiento convencional. En la imagen 607 de segundo plano, los objetos de primer plano, tales como el objeto 603 de primer plano (portero) y el objeto 604 de primer plano (portería), no son fotografiados, pero falta parte del segundo plano. La razón es que los objetos de primer plano se agrupan entre sí dentro de la escena cuya imagen es capturada para crear una imagen de segundo plano y, por tanto, parte del objeto de segundo plano no es visto desde ningún punto de visión. En el caso en el que el área de primer plano es extraída de la imagen objetivo 605 utilizando la imagen 607 de segundo plano, se adquiere una imagen 608 de primer plano. En la imagen 608 de primer plano, los objetos de primer plano que tienen una altura con respecto a la superficie del terreno son casi extraídos. No obstante, los objetos 609 de primer plano que no se ven desde una pluralidad de puntos de visión debido a que los objetos de primer plano se agrupan entre sí no son extraídos.

En contraste con esto, en la presente realización, al utilizar las imágenes de referencia (por ejemplo, la imagen 601 de referencia y similares) en una pluralidad de puntos de visión diferentes, se crea una imagen 610 de referencia, que es una imagen con el primer plano eliminado. En el caso en el que el área de primer plano se extrae de la imagen objetivo 605 utilizando la imagen 610 de referencia, se adquiere una imagen 611 de primer plano. En la imagen 611 de primer plano, el área de los objetos 603 y 604 de primer plano estáticos, y los objetos de primer plano que no se ven desde una pluralidad de puntos de visión, se extraen con alta precisión. Como anteriormente, de acuerdo con la presente realización, es posible extraer objetos de primer plano con alta precisión tanto si hay o no un cambio (movimiento o similar) de los objetos de primer plano acompañando a un lapso de tiempo, y tanto si los objetos de primer plano se agrupan entre sí o no. A

(Segunda realización)

Una segunda realización se explica centrando la atención en las diferencias con respecto a la primera realización. En la primera realización, en el momento de crear una imagen con el primer plano eliminado basada en una pluralidad de imágenes de referencia, se utilizan datos que indican el grado de coincidencia en el píxel de interés de una imagen de referencia convertida que difiere según el punto de visión. Por otro lado, en la presente realización, en el momento de crear una imagen con el primer plano eliminado basada en una pluralidad de imágenes de referencia, además de los datos que indican el grado de coincidencia, se utiliza el grado de suavidad del cambio en el valor de píxel, es decir, la llamada continuidad, en el píxel de interés de una imagen de referencia convertida que difiere según el punto de visión. La misma configuración y procesamiento que los de la primera realización reciben los mismos símbolos que los de la primera realización, y se omite su explicación.

<Acerca del esquema de procesamiento para extraer el área de primer plano>

A continuación, se explica un esquema de procesamiento para extraer un área de primer plano en la presente realización. En la presente realización, al usar imágenes de referencia convertidas obtenidas al convertir imágenes de referencia en una pluralidad de puntos de visión diferentes en imágenes en el caso en que las imágenes de referencia son ven desde el punto de visión de interés, se calcula la continuidad en el valor de píxel entre puntos de visión. La continuidad en el valor de píxel es el grado de suavidad del cambio en el valor de píxel entre la imagen de referencia convertida en el punto de visión de interés y la imagen de referencia convertida en el punto de visión adyacente al punto de visión de interés.

Concretamente, se comparan el valor de píxel del píxel de interés en la imagen de referencia convertida correspondiente al número del punto de visión base y el valor de píxel del píxel de interés en la imagen de referencia convertida en el punto de visión adyacente al punto de visión base, y la suma total de las diferencias entre los valores de píxel se calcula como continuidad. Después de esto, al usar el grado de coincidencia explicado en la primera realización y la continuidad calculada en la presente realización, se considera que el píxel cuyo grado de coincidencia es bajo y cuyo cambio en el valor de píxel no es uniforme tiene una alta probabilidad de ser el píxel del área de la imagen del objeto de primer plano y el píxel se detecta como un píxel objetivo de corrección. A continuación, al corregir la imagen de referencia convertida actualizando el valor de píxel del píxel objetivo de corrección detectado, se crea una imagen con el primer plano eliminado. Finalmente, la imagen con el primer plano eliminado creada y la imagen objetivo se comparadas y, de este modo, se extrae un área de primer plano.

En la primera realización, usando solo el grado de coincidencia entre los valores de píxel calculados basándose en las imágenes de referencia en todos los puntos de visión, se determina si el píxel de interés es el píxel del área de la imagen del objeto de primer plano. Debido a esto, el píxel del área de la imagen del segundo plano cuyo valor de píxel difiere debido a que el aspecto de color cambia según el punto de visión se considera también que tiene una alta probabilidad de ser el píxel del área de la imagen del objeto de primer plano y, por tanto, el píxel se detecta

como un píxel objetivo de corrección. Como resultado de esto, el píxel que no tiene por qué ser corregido también se corrige y, por tanto, se produce un error en la imagen de referencia convertida después de la corrección y existe la posibilidad de que se cree una imagen con el primer plano eliminado, incluyendo un objeto de primer plano. Como segundo plano cuya aspecto cambia según el punto de visión, se menciona el césped que se ha cortado con direccionalidad, que existe en una imagen capturada de una escena de juego competitivo, tal como un deporte. La aspecto de color del césped cortado con direccionalidad difiere según la dirección en la que se ve el césped y, como resultado, el valor de píxel de césped cambia según el punto de visión, aunque el césped se encuentre en la misma posición. En el caso en el que la primera realización se aplica a la escena cuyo segundo plano es césped como este, el grado de coincidencia entre los píxeles en la pluralidad de imágenes de referencia convertidas disminuye y, por tanto, el píxel del área de la imagen del césped, que es el segundo plano, puede determinarse erróneamente como el píxel del área de la imagen del objeto de primer plano. Para evitar tal determinación errónea, en la presente realización, si el píxel de interés es el píxel del área de la imagen del objeto de primer plano se determina usando la continuidad, además del grado de coincidencia. En general, para un objeto cuyo aspecto de color cambia según el punto de visión, existe un caso en el que surge una diferencia notable en el aspecto de color entre puntos de visión distantes entre sí, pero el cambio en el aspecto de color entre puntos de visión cercanos entre sí es gradual. Debido a esto, en la presente realización, el píxel del área de la imagen del segundo plano cuyo valor de píxel ha cambiado debido a la diferencia en el aspecto de color se distingue del píxel del área de la imagen del objeto de primer plano cuyo valor de píxel ha cambiado porque posee una altura con respecto a la superficie del terreno. Como resultado de esto, es posible crear una imagen con el primer plano eliminado corrigiendo la imagen de referencia convertida con alta precisión y, por tanto, es posible extraer un objeto de primer plano de una imagen objetivo con alta precisión. El objeto cuyo aspecto de color cambia según el punto de visión no se limita al ejemplo del césped descrito anteriormente, y existen una variedad de objetos, tal como el suelo de un gimnasio.

<Acerca del procesamiento para extraer el área de primer plano>

A continuación, el procesamiento para extraer un área de primer plano que se realiza mediante el aparato 100 de procesamiento de imágenes en la presente realización se explica usando la figura 7 y la figura 8. La figura 7 es un diagrama de bloques que muestra una configuración de función del aparato 100 de procesamiento de imágenes en la presente realización y la figura 8 es un diagrama de flujo que muestra un flujo del procesamiento para extraer un área de primer plano en la presente realización. La CPU 101 del aparato 100 de procesamiento de imágenes funciona como cada componente mostrado en la figura 7 y realiza una serie de procesamiento mostrada en la figura 8 ejecutando programas almacenados en la ROM 103, utilizando la RAM 102 como memoria de trabajo. Todo el procesamiento que se muestra a continuación no tiene por qué ser realizado por la CPU 101 y también puede ser posible conformar el aparato 100 de procesamiento de imágenes de modo que parte o la totalidad del procesamiento sea realizado por uno o una pluralidad de circuitos de procesamiento que no sean la CPU 101.

En la etapa S801, una unidad 701 de cálculo de continuidad determina un píxel de interés en la imagen de segundo plano base, que es el objetivo para el que se calcula la continuidad. En la presente realización, en primer lugar, el píxel superior izquierdo de la imagen de segundo plano base se selecciona como el píxel de interés y, después de esto, los píxeles sin procesar son seleccionados secuencialmente como el píxel de interés. Siempre que el cálculo de continuidad se realice para todos los píxeles de la imagen de referencia convertida, el píxel de interés puede determinarse en cualquier orden.

En la etapa S802, la unidad 701 de cálculo de continuidad calcula la continuidad en el valor de píxel del píxel de interés en la imagen de segundo plano base usando una pluralidad de imágenes de referencia convertidas (imágenes de referencia convertidas correspondientes al punto de visión base y los puntos de visión en la periferia del mismo) adquiridas desde la unidad 203 de conversión de imágenes. En este caso, el procedimiento de cálculo de continuidad en esta etapa se explica usando la figura 9.

En primer lugar, se detectan las cámaras 902 y 903 adyacentes a una cámara 901 correspondiente al número del punto de visión base determinado por la unidad 203 de conversión de imágenes y se adquieren los números de punto de visión correspondientes a estas cámaras. En lo sucesivo, el número del punto de visión adquirido se denomina número del punto de visión adyacente. En este caso, la cámara adyacente a la cámara 901 correspondiente al número del punto de visión base se determina basándose en la distancia a la cámara 901, que se calcula a partir de las coordenadas en el espacio tridimensional de la cámara. En la presente realización, la cámara 902 cuya distancia a la cámara 901 es la más corta de entre las cámaras existentes en el lado izquierdo de la cámara 901, y la cámara 903 cuya distancia a la cámara 901 es la más corta de entre las cámaras existentes en el lado derecho de la cámara 901 son detectadas como las cámaras adyacentes a la cámara 901. No obstante, la cámara más próxima a la cámara correspondiente al número del punto de visión base no se selecciona necesariamente como la cámara del punto de visión adyacente. Por ejemplo, también puede ser posible seleccionar una cámara del punto de visión adyacente de acuerdo con la dirección de fotografiado de cada cámara o con varios parámetros, tales como la resolución y el enfoque, de la cámara.

A continuación, a partir de una imagen 904 de referencia convertida correspondiente al número del punto de visión base y las imágenes 905 y 906 de referencia convertidas correspondientes a los números de punto de visión

adyacentes, se adquieren los valores de píxel de los píxeles 907, 908 y 909 en las coordenadas (u2, v2) de los píxeles de interés, y utilizando los valores de píxel adquiridos se calcula la continuidad mediante la expresión (4).

$$C(u_2, v_2) = \sum_{k=1}^a |B_{901}^k(u_2, v_2) - B_{902}^k(u_2, v_2)| + |B_{901}^k(u_2, v_2) - B_{903}^k(u_2, v_2)| \quad \dots \text{expresión (4)}$$

En este caso, cada uno de $B_{901}(u_2, v_2)$, $B_{902}(u_2, v_2)$ y $B_{903}(u_2, v_2)$ representa el valor de píxel de los píxeles de interés 907, 908 y 909 en las imágenes 904, 905 y 906 de referencia convertidas correspondientes a las cámaras 901, 902 y 903. Además, k representa un subíndice para identificar los tres canales de RGB. El valor de C calculado por la expresión (4) se hace menor a medida que el cambio en el valor de píxel entre puntos de visión es más suave. La continuidad que se usa no se limita a C calculada por la expresión (4) y también puede ser posible usar cualquier valor que indique continuidad en el valor de píxel entre puntos de visión, tal como el diferencial secundario obtenido de valores discretos. Además, en la presente realización, se explica el caso en el que se usan las cámaras 902 y 903 adyacentes a la cámara 901 correspondiente al número del punto de visión base, pero las cámaras que se usan no están limitadas a estas y también puede ser posible usar otra cámara dependiendo del aspecto de un objeto. Por ejemplo, también puede ser posible usar la cámara cuya distancia a la cámara 901 es la segunda más corta, junto a la cámara 902, en el lado izquierdo de la cámara 901 correspondiente al número del punto de visión base, en lugar de la cámara 902. Esto también es cierto con la cámara que se usa en el lado derecho de la cámara 901.

En la etapa S803, la unidad 701 de cálculo de continuidad determina si el procesamiento en la etapa 801 y la etapa S802 se ha realizado para todos los píxeles de la imagen de referencia de base. En el caso en que los resultados de la determinación en la etapa S803 son afirmativos, la unidad 701 de cálculo de continuidad entrega la continuidad calculada de todos los píxeles a una unidad 702 de determinación de la corrección y el procesamiento avanza a la etapa S308. Por otro lado, en el caso en que los resultados de la determinación son negativos, el procesamiento vuelve a la etapa S801.

En la etapa S804, la unidad 702 de determinación de la corrección actualiza el mapa de indicadores basándose en el grado de coincidencia adquirido de la unidad 204 de cálculo del grado de coincidencia y la continuidad adquirida de la unidad 701 de cálculo de continuidad. Concretamente, la unidad 702 de determinación de la corrección cambia el valor de píxel del mapa de indicadores a 1, que corresponde al píxel que se considera tiene una alta probabilidad de ser el píxel del área de la imagen del objeto de primer plano en la imagen de referencia de base (imagen de referencia convertida correspondiente al número del punto de visión base). En la presente realización, en el caso en el que el grado de coincidencia D calculado es mayor o igual que un valor umbral determinado de antemano y la continuidad calculada C es mayor o igual que un valor umbral determinado de antemano, se determina que el grado de coincidencia y el grado de suavidad del cambio en el píxel de interés entre la imagen de referencia convertida correspondiente al número del punto de visión base y otra imagen de referencia convertida son bajos. Es decir, se determina que la posibilidad de que el píxel de interés sea el píxel del área de la imagen del objeto de primer plano es alta. Por otro lado, en el caso en que no se cumplen estas condiciones, se determina que la posibilidad de que el píxel de interés sea el píxel del área de la imagen del segundo plano es alta. El valor umbral que se usa en esta etapa se determina basándose en el valor máximo o similar del valor de píxel y también puede ser posible encontrar el valor umbral utilizando un valor menor que el 20 % del valor máximo, por ejemplo, utilizando un valor arbitrario dentro de un intervalo del 1 % al 5 % del valor máximo. El procedimiento para encontrar el valor umbral es el mismo que el de la primera realización. Además, la determinación de si el píxel de interés es el píxel del área de la imagen del objeto de primer plano se realiza para cada píxel. No obstante, que no necesariamente se requiera realizar la determinación para cada píxel es lo mismo que se ha descrito también en la primera realización. La unidad 702 de determinación de la corrección entrega el mapa de indicadores para el que se ha completado la actualización a la unidad 206 de corrección. Lo anterior es el procesamiento para extraer un área de primer plano, que es realizado por el aparato 100 de procesamiento de imágenes en la presente realización.

<Acerca de los resultados de la presente realización>

A continuación, los resultados de la presente realización se explican usando la figura 10. Una imagen 1002 es una imagen de referencia convertida adquirida al convertir la imagen de referencia, para cada punto de visión, en una imagen en el caso en el que la imagen de referencia se ve desde un punto de visión 1001 con la superficie del terreno como base. En este caso, se supone que el punto de visión 1001 es el punto de visión de interés y también es el punto de visión base. Además, 1003 indica un segundo plano (por ejemplo, césped) cuyo aspecto de color cambia según el punto de visión y 1005 indica un objeto de primer plano.

En el caso en el que se crea una imagen con el primer plano eliminado basándose en imágenes de referencia en una pluralidad de puntos de visión, aplicando la primera realización a una escena mostrada en la figura 10, se adquiere una imagen 1004 de referencia. En la imagen 1004 de referencia, el objeto 1005 de primer plano se elimina, pero el segundo plano 1003 no es fotografiado correctamente. La razón es que, en el momento de crear la imagen con el primer plano eliminado, se determina si el píxel de interés es el píxel del área de la imagen del objeto de primer plano utilizando el grado de coincidencia, pero no utilizando continuidad y, por tanto, se determina que el píxel del área de la imagen del segundo plano 1003 es el píxel del objeto de primer plano. Como resultado de esto, se crea la imagen 1004 de referencia en la que se ha corregido el píxel del área de la imagen del segundo plano

1003. Incluso en el caso en el que se intenta extraer el área de primer plano de la imagen objetivo utilizando la imagen 1004 de referencia, no es posible extraer el área de primer plano con alta precisión.

En contraste con esto, en la presente realización, en el momento de crear la imagen con el primer plano eliminado basándose en imágenes de referencia en una pluralidad de puntos de visión, se determina si el píxel de interés es el píxel del área de la imagen del objeto de primer plano basándose en el grado de coincidencia y continuidad. Como resultado de esto, no se determina que el píxel del área de la imagen del segundo plano 1003 sea el píxel del área de la imagen del objeto de primer plano y, por tanto, se crea una imagen 1006 de referencia en la que el píxel del área de la imagen del segundo plano 1003 no se ha corregido. En la imagen 1006 de referencia, el objeto 1005 de primer plano se elimina y el segundo plano 1003 se fotografía correctamente. Al extraer el área de primer plano de la imagen objetivo utilizando la imagen 1006 de referencia, es posible extraer el área de primer plano con alta precisión. Tal como se describió anteriormente, de acuerdo con la presente realización, incluso en el caso en el que el segundo plano es un objeto cuyo aspecto de color cambia dependiendo del punto de visión, es posible extraer un objeto de primer plano con alta precisión.

(Tercera realización)

En la primera realización y la segunda realización, se extrae un objeto de primer plano creando una imagen con el primer plano eliminado basada en imágenes de referencia en una pluralidad de puntos de visión diferentes y comparando la imagen con el primer plano eliminado creada con una imagen objetivo. Por otro lado, en la presente realización, se extrae un área de primer plano que no incluye un área de sombra usando imágenes de primer plano imperfectas en una pluralidad de puntos de visión diferentes. En este caso, la imagen de primer plano imperfecta significa una imagen en la que el área de un objeto de primer plano y una sombra que acompaña al objeto de primer plano se extraen como un área de primer plano.

En la presente realización, al convertir la imagen de primer plano imperfecta para cada punto de visión en una imagen en el caso en el que la imagen de primer plano imperfecta se ve desde el punto de visión de interés con la superficie del terreno como base, se adquiere una pluralidad de imágenes de primer plano convertidas y se calcula el grado de coincidencia entre píxeles en la pluralidad adquirida de imágenes de primer plano convertidas. Como se explicó en la primera realización, el objeto de primer plano tiene una altura con respecto a la superficie del terreno, pero la sombra que acompaña al objeto de primer plano no tiene una altura con respecto a la superficie del terreno. En consecuencia, en la presente realización, se detecta el píxel cuyo grado de coincidencia entre píxeles es alto en la pluralidad de imágenes de primer plano convertidas y el píxel detectado se corrige considerando que el píxel detectado tiene una alta probabilidad de ser el píxel del área de sombra que no tiene altura. Como resultado, es posible crear una imagen de primer plano en la que solo se extrae el objeto de primer plano que tiene una altura como el área de primer plano sin extraer el área de sombra. En lo sucesivo, la imagen en la que se extrae el objeto de primer plano que tiene una altura sin extraer el área de sombra se denomina imagen de primer plano con sombra eliminada. La misma configuración y procesamiento que los de las realizaciones descritas anteriormente reciben los mismos símbolos que los de las realizaciones descritas anteriormente y se omite su explicación.

<Acerca del procesamiento para extraer el área de primer plano>

A continuación, el procesamiento para extraer un área de primer plano que es realizado por el aparato 100 de procesamiento de imágenes en la presente realización se explica usando la figura 11 y la figura 12. La figura 11 es un diagrama de bloques que muestra una configuración de función del aparato 100 de procesamiento de imágenes en la presente realización y la figura 12 es un diagrama de flujo que muestra un flujo del procesamiento para extraer un área de primer plano en la presente realización. La CPU 101 del aparato 100 de procesamiento de imágenes funciona como cada componente mostrado en la figura 11 y realiza una serie de procesamiento mostrada en la figura 12 ejecutando programas almacenados en la ROM 103, utilizando la RAM 102 como memoria de trabajo. Todo el procesamiento que se muestra a continuación no tiene por qué ser realizado por la CPU 101 y también puede ser posible conformar el aparato 100 de procesamiento de imágenes de modo que parte o la totalidad del procesamiento sea realizado por uno o una pluralidad de circuitos de procesamiento que no sean la CPU 101.

En la etapa S1201, una unidad 1101 de adquisición de parámetros de la cámara adquiere los parámetros de cámara de una cámara que ha capturado una imagen objetivo desde el dispositivo 108 de almacenamiento externo a través de la interfaz 105 de entrada, o desde el dispositivo 104 de almacenamiento secundario. Además, la unidad 1101 de adquisición de parámetros de la cámara determina el punto de visión de la cámara que ha capturado la imagen objetivo como el punto de visión de interés. Los parámetros de la cámara que se adquieren en esta etapa son los mismos que los parámetros de la cámara explicados en la primera realización. La unidad 1101 de adquisición de los parámetros de la cámara entrega los parámetros de la cámara a una unidad 1103 de conversión de imágenes.

En la etapa S1202, una unidad 1102 de adquisición de imágenes de primer plano adquiere una pluralidad de imágenes de primer plano en una pluralidad de puntos de visión diferentes desde el dispositivo 108 de almacenamiento externo a través de la interfaz 105 de entrada, o desde el dispositivo 104 de almacenamiento secundario. La imagen de primer plano que se adquiere en esta etapa es una imagen en la que se extrae un objeto de primer plano y se supone que un área de sombra está incluida en el área extraída. En la presente realización,

esta imagen de primer plano se crea basándose en la imagen capturada y la imagen de segundo plano capturada de antemano. A continuación, se explica concretamente el procedimiento para crear una imagen de primer plano. La imagen capturada que se usa en este caso es una imagen obtenida al capturar el objeto de primer plano y el segundo plano en la imagen objetivo en un entorno sustancialmente igual al entorno en el momento de capturar la imagen objetivo. Además, la imagen de segundo plano es una imagen obtenida al capturar el segundo plano en la imagen objetivo en un entorno sustancialmente igual al entorno en el momento de capturar la imagen objetivo. En la presente realización, se crea una imagen binaria para cada punto de visión comparando, para cada punto de visión, el valor de píxel de la imagen capturada y el valor de píxel de la imagen de segundo plano para cada píxel y estableciendo el valor de píxel del píxel en las coordenadas donde estos valores de píxel son iguales a 0 y estableciendo el valor de píxel de los otros píxeles a 1. Esta imagen binaria es una imagen de primer plano. El procedimiento para crear una imagen de primer plano no se limita a esto y la imagen de primer plano que se crea no se limita a una imagen binaria y puede ser una imagen multivaluada. Además, la unidad 1102 de adquisición de imágenes de primer plano adquiere los parámetros de la cámara correspondientes a cada imagen de primer plano junto con la imagen de primer plano. Además, la unidad 1102 de adquisición de imágenes de primer plano almacena cada imagen de primer plano en asociación con el número del punto de visión de la cámara para distinguir una imagen de primer plano de otra en la pluralidad de imágenes de primer plano. La unidad 1102 de adquisición de imágenes de primer plano entrega las imágenes de primer plano y los parámetros de la cámara a la unidad 1103 de conversión de imágenes.

En la etapa S1203, la unidad 1103 de conversión de imágenes convierte la imagen de primer plano obtenida de la unidad 1102 de adquisición de imágenes de primer plano en una imagen en el caso en el que la imagen de primer plano se ve desde el punto de visión de interés usando los parámetros de la cámara obtenidos de la unidad 1101 de adquisición de los parámetros de la cámara y la unidad 1102 de adquisición de imágenes de primer plano. La conversión en esta etapa es la misma que en la etapa S303 de la primera realización y la imagen en el caso en el que la imagen de primer plano se ve desde el punto de visión de interés se obtiene realizando una conversión de proyección para la imagen de primer plano con la superficie del terreno como base para cada punto de visión. La imagen (datos) de primer plano obtenida por la conversión de imagen en esta etapa se denomina imagen (datos) de primer plano convertida. La unidad 1103 de conversión de imágenes entrega la imagen de primer plano convertida a una unidad 1104 de cálculo del grado de coincidencia.

En la etapa 1204, la unidad 1103 de conversión de imágenes determina la imagen correspondiente al punto de visión más próximo al punto de visión de la cámara (punto de visión de interés) desde el que se captura la imagen objetivo para ser una imagen de primer plano base entre las imágenes de primer plano adquiridas desde la unidad 1102 de adquisición de imágenes de primer plano. Concretamente, la distancia entre las coordenadas del punto de visión de interés y las coordenadas del punto de visión correspondiente a la imagen de primer plano se calcula para cada punto de visión. A continuación, la imagen (datos) de primer plano correspondiente al punto de visión (punto de visión base) cuya distancia calculada es la más corta se considera la imagen (datos) de primer plano base. La unidad 1103 de conversión de imágenes entrega el número del punto de visión correspondiente a la imagen de primer plano base a una unidad 1105 de corrección. En la presente realización, el número del punto de visión correspondiente a la imagen de primer plano base se denomina número del punto de visión base. También puede ser posible que el punto de visión de interés y el punto de visión de la imagen de primer plano base coincidan perfectamente entre sí.

En la etapa S1205, la unidad 1104 de cálculo del grado de coincidencia determina un píxel de interés en la imagen de primer plano convertida, que es el objetivo de la determinación del grado de coincidencia del valor de píxel en una pluralidad de imágenes de primer plano convertidas. En la presente realización, en primer lugar, el píxel superior izquierdo de la imagen de primer plano convertida se selecciona como el píxel de interés y, después de esto, los píxeles sin procesar se seleccionan secuencialmente como el píxel de interés. Siempre que la determinación de si los píxeles coinciden en la pluralidad de imágenes de primer plano convertidas se realice para todos los píxeles de la imagen de primer plano convertida, el píxel de interés puede determinarse en cualquier orden.

En la etapa S1206, la unidad 1104 de cálculo del grado de coincidencia calcula el grado de coincidencia en el píxel de interés entre la imagen de primer plano convertida correspondiente al número del punto de visión base y otra imagen de primer plano convertida utilizando la pluralidad de imágenes de primer plano convertidas adquiridas de la unidad 1103 de conversión de imágenes. A continuación, se explica concretamente el procedimiento para calcular el grado de coincidencia.

En primer lugar, la unidad 1104 de cálculo del grado de coincidencia adquiere un valor de píxel $F_1(u_2, v_2)$ de la imagen de primer plano convertida en las coordenadas (u_2, v_2) del píxel de interés determinado. En este caso, 1 representa un subíndice para distinguir una imagen de primer plano convertida de otra en una pluralidad de imágenes de primer plano convertidas, y la unidad 1104 de cálculo del grado de coincidencia adquiere valores de píxel en el número correspondiente al número de imágenes de primer plano convertidas. A continuación, la unidad 1104 de cálculo del grado de coincidencia calcula un valor promedio de todos los valores de píxel adquiridos. En la presente realización, este valor promedio se usa como el grado de coincidencia. El grado de coincidencia no se limita a esto y también puede ser posible utilizar un valor que refleje la naturaleza estadística de una pluralidad de valores de píxel como el grado de coincidencia.

En la etapa S1207, la unidad 1104 de cálculo del grado de coincidencia determina si el procesamiento en la etapa S1205 y la etapa S1206 se ha realizado para todos los píxeles de la imagen de primer plano convertida. En el caso en el que los resultados de la determinación en la etapa S1207 son afirmativos, la unidad 1104 de cálculo del grado de coincidencia entrega los grados de coincidencia calculados de todos los píxeles a la unidad 1105 de corrección y el procesamiento avanza a la etapa S1208. Por otro lado, en el caso en el que los resultados de la determinación en la etapa S1207 son negativos, el procesamiento vuelve a la etapa S1205.

En la etapa S1208, la unidad 1105 de corrección determina un píxel de interés en la imagen de primer plano base (imagen de primer plano convertida correspondiente al número del punto de visión base). En la presente realización, en primer lugar, el píxel superior izquierdo de la imagen de primer plano base se selecciona como el píxel de interés y los píxeles sin procesar se seleccionan secuencialmente como el píxel de interés. Siempre que la actualización (etapa S1209) del valor de píxel basada en el grado de coincidencia se realice para todos los píxeles de la imagen de primer plano base, el píxel de interés puede determinarse en cualquier orden.

En la etapa S1209, la unidad 1105 de corrección detecta un píxel que tiene una alta probabilidad de ser el píxel del área de sombra en la imagen de primer plano base basándose en el grado de coincidencia adquirido de la unidad 1104 de cálculo del grado de coincidencia. A continuación, la unidad 1105 de corrección elimina el área de sombra de la imagen de primer plano incompleta cambiando el valor de píxel del píxel detectado a 0. En la presente realización, en el caso en el que el grado de coincidencia calculado es mayor o igual a un valor umbral determinado de antemano, el grado de coincidencia entre los píxeles de interés en todos los puntos de visión es elevado y, por tanto, se determina que la posibilidad de que el píxel de interés sea el píxel del área de sombra que no tiene una altura es alta. A continuación, el valor de píxel del píxel de interés en la imagen de primer plano base se cambia a 0. Por otro lado, en el caso en el que el grado de coincidencia calculado es menor que el valor umbral, el grado de coincidencia entre los píxeles de interés en todos los puntos de visión es bajo y, por tanto, se determina que la posibilidad de que el píxel de interés sea el píxel del objeto de primer plano que tiene una altura es alta. En este caso, el valor de píxel del píxel de interés en la imagen de primer plano base no cambia. En la presente realización, como valor umbral, se usa 0,8, pero el valor del valor umbral no se limita a esto.

En la etapa S1210, la unidad 1105 de corrección determina si el procesamiento de la etapa S1208 y la etapa S1209 se ha realizado para todos los píxeles de la imagen de primer plano base. En el caso en el que los resultados de la determinación en la etapa S1210 son afirmativos, la unidad 1105 de corrección entrega la imagen de primer plano base para la cual se ha completado la corrección a la unidad 104 de almacenamiento secundario, al dispositivo 108 de almacenamiento externo y al dispositivo 109 de visualización y se completa la serie de procesamiento. Por otro lado, en el caso en el que los resultados de la determinación en la etapa S1210 son negativos, el procesamiento vuelve a S1208. Lo anterior es el procesamiento para extraer un área de primer plano, que es realizado por el aparato 100 de procesamiento de imágenes en la presente realización.

<Acerca de los resultados de la presente realización>

A continuación, se explican los resultados de la presente realización usando la figura 13. El símbolo de referencia 1301 indica un objeto de primer plano cuya propia sombra 1302 existe en una superficie del terreno 1303. Las imágenes 1304 son imágenes de primer plano en una pluralidad de puntos de visión diferentes, en los que el área del objeto de primer plano 1301 y la sombra 1302 que lo acompaña se extrae como un área de primer plano. En la presente realización, se detecta un píxel del área de sombra que no tiene una altura desde la superficie del terreno basándose en el grado de coincidencia entre los píxeles de interés en una pluralidad de imágenes de primer plano convertidas obtenidas al convertir las imágenes 1304 en imágenes en el caso donde las imágenes 1304 se ven desde un punto de visión de interés 1305. A continuación, al corregir el píxel detectado, se crea una imagen 1306 de primer plano. En la imagen 1306 de primer plano, se elimina el área de la sombra 1302 que acompaña al objeto 1301 de primer plano que tiene una altura y solo se extrae el área del objeto 1301 de primer plano. Tal como se describió anteriormente, de acuerdo con la presente realización, incluso en el caso en que existe una sombra que acompañe a un objeto de primer plano que tenga una altura, es posible extraer solo el objeto de primer plano con alta precisión sin extraer el área de sombra.

En la presente realización, como imagen de primer plano incompleta, se usa la imagen de primer plano creada basándose en la imagen capturada y la imagen de segundo plano capturada de antemano, pero también puede ser posible usar la imagen de primer plano creada por la primera realización o la segunda realización. En esta ocasión, es posible extraer un objeto de primer plano con alta precisión en comparación con el caso en el que la primera realización, la segunda realización y la tercera realización se realizan por separado.

(Otras realizaciones)

Las realizaciones de la presente invención no se limitan a las realizaciones primera a tercera descritas anteriormente y puede haber una variedad de realizaciones. Por ejemplo, en las realizaciones primera a tercera descritas anteriormente, se explica el caso donde el tamaño de la imagen de referencia y el tamaño de la imagen objetivo son iguales, pero estos tamaños tienen por qué ser iguales. En esta ocasión, la imagen de segundo plano se convierte

en una imagen en el caso en el que la imagen de segundo plano se ve desde el punto de visión base, que es el punto de visión en el caso en el que la superficie del terreno se ve desde arriba. A continuación, corrigiendo la imagen de segundo plano utilizando la imagen convertida y convirtiendo la imagen de segundo plano corregida en una imagen en el caso en el que la imagen de segundo plano corregida se ve desde el píxel de interés, se crea una imagen de referencia correspondiente a la imagen objetivo.

Además, en las realizaciones primera a tercera descritas anteriormente, en el momento del cálculo del grado de coincidencia y de la extracción de un primer plano, se usa un valor de píxel en el espacio RGB, pero la información que se usa no está limitada a esto. Por ejemplo, también puede ser posible calcular el grado de coincidencia y extraer un primer plano utilizando un valor de píxel en un espacio de color diferente, tal como HSV y Lab.

Además, en las realizaciones primera a tercera descritas anteriormente, en el momento de realizar la conversión de proyección de una imagen, solo un plano de la superficie del terreno se toma como base, pero también puede ser posible usar una pluralidad de planos paralelos a la superficie del terreno como base. Por ejemplo, también puede ser posible calcular el grado de coincidencia estableciendo una pluralidad de planos dividiendo por igual el espacio entre una altura de 0 cm y una altura de 1 cm desde la superficie del terreno y utilizando todas las imágenes convertidas obtenidas por conversión de proyección con cada uno de los planos establecidos como base. Al hacerlo, mejoran las propiedades robustas para un error en los parámetros de la cámara.

(Cuarta realización)

A continuación, se explica una cuarta realización centrando la atención en las diferencias con respecto a la primera y la tercera realizaciones. En la presente realización, al usar imágenes de primer plano en una pluralidad de puntos de visión diferentes, se detecta un área que tiene una alta probabilidad de ser un área de sombra, y basándose en una diferencia en el color entre la imagen de segundo plano y la imagen capturada en el área detectada, se extrae un objeto de primer plano que no incluye una sombra. El área del objeto de primer plano se denomina área de primer plano.

<Acerca del esquema de la presente realización>

A continuación, se explica un esquema del procesamiento para extraer un área de primer plano en la presente realización usando la figura 16. En la presente realización, en primer lugar, se adquieren imágenes 1601 de primer plano en una pluralidad de puntos de visión diferentes. La imagen de primer plano que se adquiere en este caso es una imagen en la que el área de un objeto 1609 de primer plano y una sombra 1604 que lo acompaña se extrae como un área de primer plano. Se supone que en la pluralidad de imágenes 1601 de primer plano que se adquirirán, se incluye una imagen cuyo punto de visión es el mismo que un punto de visión de interés 1602 desde el que se captura una imagen, que es un objetivo de extracción de solo el área del objeto 1609 de primer plano que no incluye el área de la sombra 1604 como un área de primer plano. En lo sucesivo, la imagen de la que solo el área del objeto 1609 de primer plano que no incluye el área de la sombra 1604 se extrae como un área de primer plano se denomina imagen objetivo y el punto de visión desde el que se captura la imagen objetivo se denomina punto de visión de interés.

A continuación, al convertir, para cada punto de visión, la imagen 1601 de primer plano adquirida en una imagen en el caso en el que la imagen 1601 de primer plano se ve desde el punto de visión de interés 1602 con la superficie del terreno como base, se crea una imagen 1603 de primer plano en el punto de visión de interés. El número de imágenes 1603 de primer plano creadas en este caso es el mismo que el número de imágenes 1601 de primer plano. De aquí en adelante, la imagen 1603 de primer plano obtenida al convertir la imagen 1601 de primer plano se denomina imagen 1603 de primer plano convertida.

Tal como se describe asimismo en las realizaciones primera a tercera, casi todos los objetos de primer plano, tales como una persona y equipamiento, tienen una altura con respecto a la superficie del terreno. En contraste con esto, una sombra que acompaña al objeto de primer plano normalmente existe en la superficie del terreno y no tiene una altura con respecto a la superficie del terreno. Debido a esto, en la presente realización, al usar la imagen 1603 de primer plano convertida, se detecta un objeto de primer plano que no tiene una altura con respecto a la superficie del terreno y se considera que el objeto de primer plano correspondiente al área detectada tiene una alta probabilidad de ser una sombra. Concretamente, se determina para cada píxel si el píxel es el área de primer plano (en adelante, área del primer plano común) en común en la pluralidad de imágenes 1603 de primer plano convertidas, y el píxel determinado como el área del primer plano común se detecta como un candidato del área de sombra. Tal como se ha descrito anteriormente, la imagen 1603 de primer plano convertida se obtiene convirtiendo la imagen 1601 de primer plano en una imagen en el caso en el que la imagen 1601 de primer plano se ve desde el punto de visión de interés 1602 con la superficie del terreno como un plano base. Debido a esto, las coordenadas de las áreas 1605 a 1607 de sombra en las imágenes 1601 de primer plano, que corresponden a la sombra 1604 existente en la superficie del terreno y que no tiene una altura, se convierten en las coordenadas de un área 1608 del primer plano común existente en la misma posición en común en todas las imágenes 1603 de primer plano convertidas. Por otro lado, las coordenadas de las áreas 1610 a 1612 en las imágenes 1601 de primer plano, que corresponden al objeto 1609 que tiene una altura, se convierten en las coordenadas de las áreas 1613 a 1615 cuyas posiciones difieren

según el punto de visión. Debido a esto, en el caso en que existe un área del primer plano común en la pluralidad de imágenes 1603 de primer plano convertidas, el píxel del área del primer plano común se detecta al considerarlo como un candidato del píxel del área de sombra y, por otro lado, el píxel del área de primer plano, que no es el área del primer plano común, se considera el píxel del objeto de primer plano que tiene una altura.

A continuación, los candidatos del píxel detectado del área de sombra se comparan entre la imagen de segundo plano y la imagen capturada en el punto de visión de interés 1602 y el píxel cuya diferencia de color es pequeña se determina como el píxel del área de sombra. En las imágenes 1601 de primer plano en el punto de visión de interés 1602, en las que el área del objeto 1609 y la sombra 1604 que lo acompaña se extrae como un área de primer plano, cambiando finalmente el valor de píxel del píxel determinado como el píxel del área de sombra, se extrae un área de primer plano que no incluye el área de sombra. En el caso en el que se produce una sombra que acompaña a un objeto, tal como una persona y equipamiento, en la imagen capturada, el área en la que existe una sombra en el segundo plano y sobre la superficie del terreno aparece como un área oscura en comparación con el caso donde no existe una sombra. Por otro lado, en la imagen capturada, la imagen de un objeto, tal como una persona y equipamiento, normalmente se dibuja en un color diferente del color del segundo plano y la superficie del terreno que aparece en el caso en que el objeto no existe. Debido a esto, en el área que tiene una alta probabilidad de ser una sombra, es posible considerar el píxel cuya diferencia de color entre la imagen de segundo plano y la imagen capturada es pequeña como el píxel del área de sombra.

Lo anterior es el esquema del procesamiento que se realiza en la presente realización. La imagen objetivo que se utiliza no se limita al ejemplo descrito anteriormente y también puede ser posible usar varios tipos de datos de imagen, tales como datos cuya imagen ha sido capturada por una cámara de monitorización.

<Acerca de la configuración de hardware del aparato de procesamiento de imágenes>

La configuración de hardware del aparato de procesamiento de imágenes de la presente realización es similar a la de la primera realización (véase la figura 1).

<Acerca del procesamiento para extraer el área de primer plano>

A continuación, el procesamiento para extraer un área de primer plano que se realiza mediante el aparato 100 de procesamiento de imágenes en la presente realización se explica usando la figura 14 y la figura 15. La figura 14 es un diagrama de bloques que muestra una configuración de función del aparato 100 de procesamiento de imágenes y la figura 15 es un diagrama de flujo que muestra un flujo del procesamiento para extraer un área de primer plano. La CPU 101 del aparato 100 de procesamiento de imágenes funciona como cada componente mostrado en la figura 14 y realiza una serie de procesamiento mostrada en la figura 15 ejecutando los programas almacenados en la ROM 103, utilizando la RAM 102 como memoria de trabajo. Todo el procesamiento que se muestra a continuación no tiene por qué ser realizado por la CPU 101 y también puede ser posible conformar el aparato 100 de procesamiento de imágenes de modo que parte o la totalidad del procesamiento sea realizado por uno o una pluralidad de circuitos de procesamiento que no sean la CPU 101.

A continuación, se explica el flujo de procesamiento que realiza cada componente. En la etapa S1501, una unidad 1401 de adquisición de imágenes objetivo adquiere una imagen objetivo desde el dispositivo 108 de almacenamiento externo a través de la interfaz 105 de entrada, o desde la unidad 104 de almacenamiento secundario. Tal como se describió anteriormente, la imagen objetivo es una imagen que es un objetivo de extracción de un objeto de primer plano. Además, la unidad 1401 de adquisición de imágenes objetivo determina el punto de visión de una cámara que ha capturado la imagen objetivo como el punto de visión de interés. Además, la unidad 1401 de adquisición de imágenes objetivo adquiere los parámetros (en adelante, parámetros de la cámara) de la cámara que ha capturado la imagen objetivo, junto con la imagen objetivo. En este caso, los parámetros de la cámara son parámetros que permiten que un cálculo proyecte un punto en el espacio tridimensional sobre una imagen capturada por la cámara e incluye parámetros externos que representan la posición y la postura de la cámara y los parámetros internos que representan la distancia focal y el centro óptico. También puede ser posible utilizar valores medidos y valores de diseño almacenados de antemano en una memoria como parámetros de la cámara. La unidad 1401 de adquisición de imágenes objetivo entrega la imagen objetivo a una unidad 1406 de cálculo del grado de similitud en el color y los parámetros de la cámara de la imagen objetivo a una unidad 1404 de conversión de imágenes. En este caso, se explica el caso en el que el número de imágenes objetivo es uno, pero también es posible aplicar la presente realización al caso en el que el número de imágenes objetivo es de dos o más.

En la etapa S1502, una unidad 1402 de adquisición de imágenes de segundo plano adquiere la imagen de segundo plano en el punto de visión de interés desde el dispositivo 108 de almacenamiento externo a través de la interfaz 105 de entrada, o desde la unidad 104 de almacenamiento secundario. La imagen de segundo plano en la presente realización es una imagen en la que solo se fotografía el segundo plano en la imagen objetivo. En la presente realización, la imagen de segundo plano se adquiere realizando la captura de imágenes de antemano en la situación en la que el objeto de primer plano no existe y solo existe el segundo plano. En detalle, la captura de imágenes se realiza utilizando una cámara que tiene los mismos parámetros de cámara que los parámetros de cámara de la

cámara que ha capturado la imagen objetivo en un entorno sustancialmente igual que el entorno (tiempo meteorológico, zona horaria, etc.) del momento de captura de la imagen objetivo. El procedimiento para adquirir una imagen de segundo plano no se limita a este procedimiento. Por ejemplo, también puede ser posible crear una imagen de segundo plano realizando un procesamiento de filtro usando un filtro de valor medio o un filtro de valor promedio para una pluralidad de imágenes correspondientes a una pluralidad de tiempos diferentes, que se obtiene capturando continuamente las imágenes de una escena desde el mismo punto de visión a lo largo de una serie temporal. Alternativamente, también puede ser posible crear una imagen de segundo plano realizando un procesamiento de agrupación para la pluralidad de imágenes. La unidad 1402 de adquisición de imágenes de segundo plano entrega la imagen de segundo plano a la unidad 1406 de cálculo del grado de similitud en el color.

En la etapa S1503, una unidad 1403 de adquisición de imágenes de primer plano adquiere una pluralidad de imágenes de primer plano en una pluralidad de puntos de visión diferentes como imágenes de referencia desde el dispositivo 108 de almacenamiento externo a través de la interfaz 105 de entrada, o desde la unidad 104 de almacenamiento secundario. Se supone que la imagen de primer plano adquirida en esta etapa es una imagen (por ejemplo, la imagen 1601 de primer plano en la figura 16) en la que se ha extraído el objeto de primer plano, y en el área de primer plano extraída, se incluye un área de sombra. En la presente realización, esta imagen de primer plano se crea basándose en la imagen capturada y la imagen de segundo plano. A continuación, se explica específicamente el procedimiento de creación de una imagen de primer plano. La imagen capturada que se utiliza en este caso es una imagen obtenida al capturar la imagen del objeto de primer plano y el segundo plano en la imagen objetivo en un entorno sustancialmente igual al entorno en el momento de capturar la imagen objetivo. Además, la imagen de segundo plano es una imagen obtenida al capturar la imagen del segundo plano en la imagen objetivo en un entorno sustancialmente igual al entorno en el momento de capturar la imagen objetivo. En la presente realización, para cada punto de visión, el valor de píxel de la imagen capturada y el valor de píxel de la imagen de segundo plano se comparan para cada píxel, y al ajustar a 0 el valor de píxel del píxel en las coordenadas en las que estos valores de píxel son iguales y ajustar a 1 el valor de píxel de los otros píxeles, se crea una imagen binaria para cada punto de visión. Esta imagen binaria es la imagen de primer plano. Se supone que, en una pluralidad de imágenes de primer plano así creadas, está incluido una imagen cuyo punto de visión es el mismo que el punto de visión de interés. El procedimiento de creación de una imagen de primer plano no se limita a este procedimiento y la imagen de primer plano que se crea puede ser una imagen multivaluada, no limitada a una imagen binaria. Además, la unidad 1403 de adquisición de imágenes de primer plano adquiere los parámetros de la cámara correspondientes a cada imagen de primer plano, junto con la imagen de primer plano. Además, la unidad 1403 de adquisición de imágenes de primer plano almacena cada imagen de primer plano en asociación con el número del punto de visión de la cámara para distinguir una imagen de primer plano de otra en la pluralidad de imágenes de primer plano. La unidad 1403 de adquisición de imágenes de primer plano entrega la imagen de primer plano y los parámetros de la cámara a la unidad 1404 de conversión de imágenes y entrega solamente la imagen de primer plano a una unidad 1408 de modificación de datos de imagen de primer plano.

En la etapa S1504, la unidad 1404 de conversión de imágenes convierte la pluralidad de imágenes de primer plano en imágenes en el caso en el que las imágenes de primer plano son vistas desde el punto de visión de interés, respectivamente, usando los parámetros de la cámara adquiridos de la unidad 1403 de adquisición de imágenes de primer plano. Específicamente, al realizar una conversión de proyección para cada imagen de primer plano con la superficie del terreno como base, se obtiene una imagen en el caso en el que la imagen de primer plano se ve desde el punto de visión de interés. La imagen (datos) obtenida mediante la conversión de imágenes en esta etapa se denomina imagen (datos) de primer plano convertida. En este caso, el procedimiento de conversión de imágenes en esta etapa es tal como se describió anteriormente usando la figura 5.

En la etapa S1504, se realiza para cada imagen de primer plano la conversión de proyección que toma la cámara con el punto de visión correspondiente a la imagen de primer plano adquirida de la unidad 1403 de adquisición de imágenes de primer plano como la cámara 502 en la figura 5 y la cámara con el punto de visión de interés determinado por la unidad 1401 de adquisición de imágenes objetivo como la cámara 505. Debido a esto, el número de imágenes de primer plano convertidas adquiridas en esta etapa es el mismo que el número de imágenes de primer plano adquiridas por la unidad 1403 de adquisición de imágenes de primer plano. Además, cada una de las imágenes de primer plano convertidas se almacena en asociación con el número del punto de visión de cada imagen de primer plano adquirida por la unidad 1403 de adquisición de imágenes de primer plano. La unidad 1404 de conversión de imágenes entrega la imagen de primer plano convertida a una unidad 1405 de detección del primer plano común.

En la etapa S1505, la unidad 1405 de detección del primer plano común inicializa el mapa de indicadores, es decir, ajusta a 0 los valores de píxel de todos los píxeles del mapa de indicadores. El mapa de indicadores que se inicializa en esta etapa se usa para detectar un área de sombra en la imagen de primer plano del punto de visión de interés, que es adquirida por la unidad 1403 de adquisición de imágenes de primer plano. En este mapa de indicadores, 1 se sustituye por el valor de píxel correspondiente al píxel determinado como el píxel del área de sombra (o el píxel que tiene una alta probabilidad de ser el píxel del área de sombra) y 0 se sustituye por el valor de píxel correspondiente al píxel que se ha determinado que no es el píxel del área de sombra. Mediante la inicialización en esta etapa, se determina que toda el área de la imagen de primer plano del punto de visión de interés no es el área de sombra.

En la etapa S1506, la unidad 1405 de detección del primer plano común determina un píxel de interés en la imagen de primer plano convertida, que es el objetivo para el que se realiza la determinación sobre si el píxel de interés es el píxel del área del primer plano común. En la presente realización, en primer lugar, se selecciona el píxel superior izquierdo de la imagen de primer plano convertida como el píxel de interés y después de esto, los píxeles sin procesar se seleccionan secuencialmente como el píxel de interés. Siempre que la determinación de si el píxel de interés es el píxel del área del primer plano común se realice para todos los píxeles de la imagen de primer plano convertida, el píxel de interés puede determinarse en cualquier orden.

En la etapa S1507, la unidad 1405 de detección del primer plano común calcula un valor de evaluación que se utiliza en el momento de determinar si el píxel de interés es el píxel del área del primer plano común basándose en la pluralidad de imágenes de primer plano convertidas adquiridas de la unidad 1404 de conversión de imágenes. A continuación, se explica específicamente el procedimiento de cálculo de un valor de evaluación.

En primer lugar, la unidad 1405 de detección del primer plano común adquiere un valor de píxel $F_i(u_2, v_2)$ de la imagen de primer plano convertida en las coordenadas (u_2, v_2) del píxel de interés. En este caso, i representa un subíndice para distinguir una imagen de primer plano convertida de otra en la pluralidad de imágenes de primer plano convertidas y la unidad 1405 de detección del primer plano común adquiere valores de píxel en el número correspondiente al número de imágenes de primer plano convertidas. A continuación, la unidad 1405 de detección del primer plano común calcula un valor promedio de todos los valores de píxel adquiridos. En la presente realización, este valor promedio se usa como un valor de evaluación en el momento de determinar si el píxel de interés es el píxel de la imagen de primer plano común. El procedimiento de cálculo de un valor de evaluación no se limita a esto, y también puede ser posible utilizar como valor de evaluación un valor arbitrario que represente la naturaleza estadística de una pluralidad de valores de píxel, tal como un valor medio.

En la etapa S1508, la unidad 1405 de detección del primer plano común determina si el píxel de interés es o no el píxel del área del primer plano común basándose en el valor de evaluación y actualiza el valor de píxel del mapa de indicadores de acuerdo con los resultados de la determinación. Específicamente, tal como se expresa mediante la expresión (5), al considerar el píxel cuyo valor de evaluación es mayor o igual que un valor umbral como el píxel del área del primer plano común, el valor de píxel del mapa de indicadores correspondiente al píxel se ajusta a 1. En cambio, al considerar el píxel cuyo valor de evaluación es menor que el valor umbral como un píxel que no es el píxel del área del primer plano común y el valor de píxel del mapa de indicadores correspondiente al píxel se ajusta a 0.

$$M_f(x, y) = \begin{cases} 1 & \text{si } V(x, y) \geq Th \\ 0 & \text{si } V(x, y) < Th \end{cases} \quad \dots \text{expresión (5)}$$

En este caso, $V(x, y)$ representa un valor de evaluación en las coordenadas (x, y) y Th representa un valor umbral. El umbral que se utiliza en este caso se determina basándose en el valor de píxel máximo que puede tomar el píxel de la imagen de primer plano o en el número de imágenes de primer plano convertidas (es decir, el número de puntos de visión). Por ejemplo, en el caso de la figura 16, como imagen de primer plano, se utiliza una imagen binaria y, por tanto, el valor de píxel máximo es 1, el número de puntos de visión es diez y también puede ser posible utilizar un valor de 0,6 como valor umbral, lo que indica que el píxel de interés es el píxel de la imagen de primer plano en la mayoría de los puntos de visión.

Tal como se ha descrito anteriormente, en el caso en que existe un área del primer plano común en una pluralidad de imágenes de primer plano convertidas, es posible considerar el píxel del área del primer plano común como un candidato del píxel del área de sombra. Debido a esto, al actualizar el valor de píxel del mapa de indicadores, que corresponde al píxel cuyo valor de evaluación se determina que es mayor o igual que el valor umbral, a 1 en esta etapa, se indica que la probabilidad de que el píxel sea el píxel del área de sombra es alta.

En la etapa S1509, la unidad 1405 de detección del primer plano común determina si el procesamiento en la etapa S1506 a la etapa S1508 se ha realizado para todos los píxeles de la imagen de primer plano convertida. En el caso en que los resultados de la determinación en la etapa S1509 son afirmativos, la unidad 1405 de detección del primer plano común entrega el mapa de indicadores actualizado a una unidad 1407 de determinación del área de sombra y el procesamiento avanza a la etapa S1510. Por otro lado, en el caso en que los resultados de la determinación son negativos, el procesamiento vuelve a la etapa S1506.

En la etapa S1510, la unidad 1406 de cálculo del grado de similitud en el color determina un píxel de interés en la imagen objetivo, que es el objetivo del cálculo del grado de similitud en el color entre la imagen objetivo adquirida de la unidad 1401 de adquisición de imágenes objetivo y la imagen de segundo plano adquirida de la unidad 1402 de adquisición de imágenes de segundo plano. En la presente realización, en primer lugar, el píxel superior izquierdo de la imagen objetivo se selecciona como el píxel de interés y después de esto, los píxeles sin procesar se seleccionan secuencialmente como el píxel de interés. Siempre que se calcule el grado de similitud en el color para todos los píxeles de la imagen objetivo, el píxel de interés puede determinarse en cualquier orden.

En la etapa S1511, la unidad 1406 de cálculo del grado de similitud en el color calcula el grado de similitud en el color en el píxel de interés entre la imagen objetivo adquirida de la unidad 1401 de adquisición de imágenes objetivo y la imagen de segundo plano adquirida de la unidad 1402 de adquisición de imágenes de segundo plano. En la presente realización, tal como se expresa en la expresión (6), el error cuadrático medio del valor de píxel se usa como el grado de similitud en el color.

$$C(x, y) = \frac{1}{3} \sum_{k=1}^3 (I^k(x, y) - I_b^k(x, y))^2 \quad \dots \text{expresión (6)}$$

En este caso, $I(x, y)$ e $I_b(x, y)$ representan el valor de píxel de la imagen objetivo y el valor de píxel de la imagen de segundo plano, respectivamente, en las coordenadas (x, y) y k representa un subíndice para identificar los tres canales de RGB. El valor del grado de similitud en el color C que se calcula mediante la expresión (6) se hace menor a medida que el color del píxel de interés en la imagen objetivo y el color del píxel en las mismas coordenadas que las del píxel de interés en la imagen de segundo plano se asemejan más. El grado de similitud en el color que se usa no se limita al error cuadrático medio y se puede usar un valor arbitrario que indique el grado de similitud en el color entre píxeles. En el momento de calcular el grado de similitud en el color en el píxel de interés, también puede ser posible calcular el grado de similitud en el color usando el píxel de interés y los píxeles periféricos del mismo.

En la etapa S1512, la unidad 1406 de cálculo del grado de similitud en el color determina si el procesamiento en la etapa S1510 y la etapa S1511 se ha realizado para todos los píxeles de la imagen objetivo. En el caso en que los resultados de la determinación en la etapa S1512 son afirmativos, la unidad 1406 de cálculo del grado de similitud en el color entrega los grados de similitud en el color calculados de todos los píxeles a la unidad 1407 de determinación del área de sombra y el procesamiento avanza a la etapa S1513. Por otro lado, en el caso en que los resultados de la determinación son negativos, el procesamiento vuelve a la etapa 1510.

En la etapa S1513, la unidad 1407 de determinación del área de sombra determina un píxel de interés en el mapa de indicadores adquirido de la unidad 1405 de detección del primer plano común, que corresponde al píxel objetivo para el que se determina si el píxel de interés es el píxel correspondiente al píxel del área de sombra. En la presente realización, en primer lugar, el píxel superior izquierdo del mapa de indicadores se selecciona como el píxel de interés y, después de esto, los píxeles sin procesar se seleccionan secuencialmente como el píxel de interés. Siempre que la actualización (etapa S1514) del valor de píxel basada en el valor de píxel del mapa de indicadores y el grado de similitud en el color se realice para todos los píxeles del mapa de indicadores, el píxel de interés puede determinarse en cualquier orden.

En la etapa S1514, la unidad 1407 de determinación del área de sombra determina si el píxel de interés es el píxel correspondiente al píxel del área de sombra basándose en el mapa de indicadores adquirido de la unidad 1405 de detección del primer plano común y el grado de similitud en el color adquirido de la unidad 1406 de cálculo del grado de similitud en el color. Específicamente, en el caso en que se cumplen las dos condiciones siguientes, se determina que el píxel de interés es el píxel correspondiente al píxel del área de sombra y el valor de píxel del píxel de interés en el mapa de indicadores se ajusta a 1. Por otro lado, en el caso en que no se cumplen las dos condiciones, se determina que el píxel de interés no es el píxel correspondiente al píxel del área de sombra y el valor de píxel del píxel de interés en el mapa de indicadores se ajusta a 0.

Primera condición: el valor de píxel del mapa de indicadores es 1 ($M_F(u_2, v_2) = 1$)

Segunda condición: el grado de similitud en el color es menor o igual que un valor umbral ($C(u_2, v_2) \leq \text{Thr}$).

Tal como se ha descrito anteriormente, es posible considerar el píxel cuya diferencia de color entre la imagen de segundo plano y la imagen capturada es pequeña como el píxel del área de sombra en el área que tiene una alta probabilidad de ser una sombra. Debido a esto, el píxel que satisface las dos condiciones descritas anteriormente, es decir, el píxel cuyo valor de píxel del mapa de indicadores es 1 y cuyo grado de similitud en el color es menor o igual que el valor umbral, (es decir, el píxel cuyo color es similar entre la imagen de segundo plano y la imagen capturada) se considera como el píxel del área de sombra.

En la etapa S1515, la unidad 1407 de determinación del área de sombra determina si el procesamiento en la etapa S1513 y la etapa S1514 se ha realizado para todos los píxeles del mapa de indicadores. En el caso en el que los resultados de la determinación en la etapa S1515 son afirmativos, la unidad 1407 de determinación del área de sombra entrega el mapa de indicadores actualizado a la unidad 1408 de modificación de los datos de la imagen de primer plano y el procesamiento avanza a la etapa S1516. Por otro lado, en el caso en el que los resultados de la determinación son negativos, el procesamiento vuelve a la etapa S1513.

En la etapa S1516, la unidad 1408 de modificación de los datos de la imagen de primer plano modifica la imagen de primer plano en el punto de visión de interés adquirida de la unidad 1403 de adquisición de imágenes de primer plano basándose en el mapa de indicadores adquirido de la unidad 1407 de determinación del área de sombra. Específicamente, para cada píxel de la imagen de primer plano en el punto de visión de interés, se determina si el valor de píxel es 1 y el valor de píxel del píxel en las mismas coordenadas en el mapa de indicadores es 1 (lo que

indica que el píxel de la imagen de primer plano es el píxel del área de sombra). A continuación, el valor de píxel del píxel de la imagen de primer plano en el punto de visión de interés, que satisface estas condiciones, se cambia a 0. Por otro lado, el valor de píxel del píxel que no satisface estas condiciones no se cambia. Mediante esta etapa, es posible modificar la imagen de primer plano en el punto de visión de interés y extraer el área de primer plano que no incluye el área de sombra. La unidad 1408 de modificación de los datos de la imagen de primer plano entrega la imagen de primer plano modificada al dispositivo 104 de almacenamiento secundario, al dispositivo 108 de almacenamiento externo y al dispositivo 109 de visualización y la serie de procesamiento se completa. Lo anterior es el procesamiento para extraer un área de primer plano, que es realizado por el aparato 100 de procesamiento de imágenes en la presente realización.

<Acerca de los resultados de la presente realización>

A continuación, se explican los resultados de la presente realización usando la figura 17. En la figura 17, una imagen 1702 es una imagen capturada que se captura desde un punto de visión de interés 1701 y una imagen 1703 es una imagen de primer plano en el punto de visión de interés 1701 en la que el área de los objetos de primer plano 1706 y 1707 y las sombras que los acompañan se extrae como un área de primer plano. Un área 1704 y un área 1705 de la imagen 1703 de primer plano son áreas correspondientes a las formas que existen cerca de la superficie del terreno entre formas tridimensionales restauradas usando imágenes de primer plano en una pluralidad de puntos de visión, y se considera que tienen una alta probabilidad de ser el área de sombra. Aquí, en el caso en el que el tamaño del vóxel utilizado para la restauración de la forma tridimensional es grande, en el área 1705 que tiene una alta probabilidad de ser el área de sombra del objeto de primer plano 1706, está incluida parte del área del objeto de primer plano 1707. En el área 1705 como esta, en el caso en el que la imagen 1703 de primer plano se modifica considerando el píxel cuya diferencia de color entre la imagen capturada 1702 y una imagen 1708 de segundo plano correspondiente a la misma como el píxel del área de sombra, se obtiene una imagen 1709 de primer plano. En la imagen 1709 de primer plano, el área de los objetos de primer plano 1706 y 1707 se extrae sin extraer el área de sombra debido a las sombras que acompañan a los objetos de primer plano 1706 y 1707. Sin embargo, en la imagen 1709 de primer plano, parte del área del objeto de primer plano 1707, es decir, el área que está incluida en el área que tiene una alta probabilidad de ser el área de sombra y en la que el color del objeto de primer plano 1707 y el color del segundo plano son similares, no se extrae como el área de primer plano.

En contraste con esto, en la presente realización, usando imágenes de primer plano (que incluyen una imagen 1710 en el punto de visión de interés 1701) en una pluralidad de puntos de visión diferentes, se detecta un área del primer plano común en una pluralidad de imágenes de primer plano convertidas. A continuación, al detectar un área de sombra 1711 basándose en el área del primer plano común y cambiar el valor de píxel del píxel del área de sombra 1711, se crea una imagen 1712 de primer plano. Una comparación entre la imagen 1709 de primer plano y la imagen 1712 de primer plano indica que el área de los objetos de primer plano 1706 y 1707 se extrae con alta precisión en la imagen 1712 de primer plano. Tal como se ha descrito anteriormente, según la presente realización, es posible extraer un objeto de primer plano que no incluye una sombra con alta precisión independientemente del estado de una escena, tal como un estado en el que los objetos de primer plano están agrupados. Además, la presente realización es el procesamiento de imágenes de una imagen bidimensional que no requiere la restauración de una forma tridimensional y, por tanto, es posible extraer un objeto de primer plano que no incluya una sombra con una pequeña cantidad de cálculo.

(Quinta realización)

En la cuarta realización, al usar imágenes de primer plano en una pluralidad de puntos de visión diferentes, se detecta un área (área del primer plano común) que tiene una alta probabilidad de ser un área de sombra y, basándose en la diferencia de color entre una imagen de segundo plano y una imagen capturada en el área, se extrae un objeto de primer plano que no incluye una sombra. En contraste con esto, en la presente realización, se extrae un objeto de primer plano que no incluye una sombra basándose en la diferencia de color entre la imagen de segundo plano y la imagen capturada y la diferencia de textura entre la imagen de segundo plano y la imagen capturada en el área que tiene una alta probabilidad de ser una sombra. La misma configuración y procesamiento que los de la cuarta realización reciben los mismos símbolos y se omite su explicación.

<Acerca del esquema de la presente realización>

A continuación, en la presente realización se explica un esquema de procesamiento para extraer un área de primer plano. En la presente realización, se calcula para cada píxel un grado de similitud en la textura entre la imagen de segundo plano y la imagen capturada en el punto de visión de interés. Específicamente, el grado de similitud en la textura se calcula estableciendo un área (en adelante, bloque) que consiste en una pluralidad de píxeles en la imagen de segundo plano y la imagen capturada, y calculando la suma de la magnitud de cambio en el valor de píxel entre el segundo plano y la imagen capturada para los píxeles dentro del bloque. Después de esto, al usar el grado de similitud en el color explicado en la cuarta realización y el grado de similitud en la textura calculado en la presente realización, se detecta un píxel similar en color y similar en textura entre la imagen de segundo plano y la imagen capturada, y el píxel detectado se determina como el píxel del área de sombra. Finalmente, al cambiar el valor de píxel del píxel determinado como el píxel del área de sombra en la imagen de primer plano en la que se extrae el

área de primer plano que incluye el área de sombra, se extrae el área de primer plano que no incluye el área de sombra.

En la cuarta realización, en el área que tiene una alta probabilidad de ser una sombra, el píxel de color similar entre la imagen de segundo plano y la imagen capturada se considera como el píxel del área de sombra. En consecuencia, en el caso en que un objeto de primer plano de color similar entre una imagen de entrada y la imagen de segundo plano está incluido en el área que tiene una alta probabilidad de ser una sombra, el objeto de primer plano se considera erróneamente como una sombra. Como resultado de esto, no es posible extraer el objeto de primer plano considerado erróneamente como una sombra y, por tanto, se reduce la precisión de la extracción de un objeto de primer plano que no incluye una sombra. Teniendo en cuenta este problema, en la quinta realización, usando el grado de similitud en la textura, además del grado de similitud en el color, se determina si el píxel de interés es el píxel de un área de sombra.

En general, un área sombreada mantiene el mismo patrón que antes de estar sombreada. Por ejemplo, en el caso en que una persona está de pie en el césped, en el área donde una persona proyecta una sombra sobre el césped, el color cambia según la presencia/ausencia de una sombra, pero se mantiene el patrón del césped. En consecuencia, en el área que tiene una alta probabilidad de ser una sombra, es posible distinguir el píxel de un objeto de primer plano, tal como una persona, cuyo color es similar entre el segundo plano y el primer plano, del píxel de un área de sombra que acompaña el objeto basándose en el grado de similitud en la textura. En vista de esto, al usar el grado de similitud en la textura en el momento de determinar un área de sombra, la precisión de la determinación de un área de sombra mejora y es posible extraer con alta precisión un objeto de primer plano que no incluye una sombra. El área en la que se mantiene la textura en el caso en que se proyecta una sombra no se limita al ejemplo de césped descrito anteriormente y es posible aplicar la presente realización a varios tipos de datos de imagen.

<Acerca del procesamiento para extraer el área de primer plano>

A continuación, el procesamiento para extraer un área de primer plano que se realiza mediante el aparato 100 de procesamiento de imágenes en la presente realización se explica usando la figura 18, la figura 19A y la figura 19B. La figura 18 es un diagrama de bloques que muestra una configuración de función del aparato 100 de procesamiento de imágenes de la presente realización y la figura 19A y la figura 19B son diagramas de flujo que muestran un flujo del procesamiento para extraer un área de primer plano en la presente realización. La CPU 101 del aparato 100 de procesamiento de imágenes funciona como cada componente mostrado en la figura 18 y realiza una serie de procesamiento mostrada en la figura 19A y la figura 19B ejecutando los programas almacenados en la memoria ROM 103 utilizando la memoria RAM 102 como memoria de trabajo. Todo el procesamiento que se muestra a continuación no tiene por qué ser realizado por la CPU 101 y también puede ser posible conformar el aparato 100 de procesamiento de imágenes de modo que parte o la totalidad del procesamiento sea realizado por uno o una pluralidad de circuitos de procesamiento que no sean la CPU 101.

En la etapa S1901, una unidad 1801 de cálculo del grado de similitud en la textura determina un píxel de interés en una imagen objetivo. El píxel de interés que se determina en esta etapa es un píxel que es un objetivo para el que el grado de similitud en la textura se calcula entre la imagen objetivo adquirida desde la unidad 1401 de adquisición de imágenes objetivo y la imagen de segundo plano adquirida desde la unidad 1402 de adquisición de imágenes de segundo plano. En la presente realización, en primer lugar, el píxel superior izquierdo de la imagen objetivo se selecciona como el píxel de interés y, después de esto, los píxeles sin procesar son seleccionados secuencialmente como el píxel de interés. Siempre que se calcule el grado de similitud en la textura para todos los píxeles de la imagen objetivo, el píxel de interés puede determinarse en cualquier orden.

En la etapa S1902, la unidad 1801 de cálculo del grado de similitud en la textura calcula el grado de similitud en la textura en el píxel de interés entre la imagen objetivo adquirida desde la unidad 1401 de adquisición de imágenes objetivo y la imagen de segundo plano adquirida desde la unidad 1402 de adquisición de imágenes de segundo plano. A continuación, se explica específicamente el procedimiento de cálculo del grado de similitud en la textura.

En primer lugar, en la imagen de segundo plano, se determina un píxel de segundo plano que se compara con el píxel de interés de la imagen objetivo. Específicamente, el píxel de la imagen de segundo plano, cuyas coordenadas son las mismas que las coordenadas (u_2 , v_2) del píxel de interés, se considera un píxel de referencia. Además, se determinan los píxeles periféricos del píxel de interés, los cuales se utilizan en el momento de calcular el grado de similitud en la textura. En la presente realización, se define un bloque con el píxel de interés como centro y se determina que los píxeles incluidos en el bloque son píxeles que se usan en el momento de calcular el grado de similitud en la textura. El tamaño del bloque se determina de antemano de acuerdo con el tamaño de imagen de la imagen objetivo. Por ejemplo, en el caso en el que el tamaño de imagen de la imagen objetivo sea FHD (1920 x 1080), el tamaño del bloque se ajusta a 9 x 9.

A continuación, el bloque con el píxel de interés de la imagen objetivo como centro se compara con el bloque con la imagen de referencia de la imagen de segundo plano como centro, y se calcula el grado de similitud en la textura. En la presente realización, tal como se expresa en la expresión (7), el error cuadrático medio del valor de píxel entre los bloques, objetivos de la comparación, se usa como el grado de similitud en la textura.

$$W(x, y) = \frac{1}{3} \sum_{k=1}^3 \frac{1}{|B|} \sum_{(x,y) \in B} (I_b^k(x, y) - I(x, y))^2 \quad \dots \text{expresión (7)}$$

Aquí, B representa un conjunto de píxeles incluidos en el bloque con las coordenadas (x, y) del píxel que se utiliza para ajustar como centro y |B| representa el número de píxeles incluidos en el conjunto B. Además, I(x, y) e I_b(x, y) representan el valor de píxel de la imagen objetivo y el valor de píxel de la imagen de segundo plano, respectivamente, en las coordenadas (x, y) y k representa un subíndice para identificar los tres canales de RGB. El valor de un grado de similitud en la textura W que se calcula mediante la expresión (7) se hace menor a medida que la textura se asemeja más entre el píxel de interés en la imagen objetivo y el píxel de referencia en la imagen de segundo plano. El grado de similitud en la textura que se usa no se limita al error cuadrático medio expresado en la expresión (7) y se puede usar un valor arbitrario que indica el grado de similitud en la textura entre píxeles. Además, también puede ser posible calcular el grado de similitud en la textura utilizando imágenes convertidas creando las imágenes convertidas en las que se extrae el contorno realizando un procesamiento de filtro o similar para la imagen objetivo y la imagen de segundo plano. Además, también puede ser posible usar una diferencia calculada como el grado de similitud en la textura calculando la diferencia en la magnitud de características de cada bloque entre la imagen objetivo y la imagen de segundo plano después de calcular la magnitud de características en cada bloque de la imagen objetivo y la imagen de segundo plano.

En la etapa S1903, la unidad 1801 de cálculo del grado de similitud en la textura determina si el procesamiento en la etapa 1901 y la etapa S1902 se ha realizado para todos los píxeles de la imagen objetivo. En el caso en el que los resultados de la determinación en la etapa 1903 son afirmativos, la unidad 1801 de cálculo del grado de similitud en la textura entrega los grados de similitud en la textura calculados de todos los píxeles a una unidad 1802 de determinación del área de sombra y el procesamiento avanza a la etapa S1513. Por otro lado, en el caso en el que los resultados de la determinación son negativos, el procesamiento vuelve a la etapa S1901.

En la etapa S1904, la unidad 1802 de determinación del área de sombra determina si el píxel de interés es el píxel correspondiente al píxel del área de sombra. Esta determinación se realiza basándose en el mapa de indicadores adquirido desde la unidad 1405 de detección del primer plano común, el grado de similitud en el color adquirido desde la unidad 1406 de cálculo del grado de similitud en el color, y el grado de similitud en la textura adquirido desde la unidad 1801 de cálculo del grado de similitud en la textura. Específicamente, en el caso en que se cumplen las siguientes tres condiciones, se determina que el píxel de interés en el mapa de indicadores es el píxel correspondiente al píxel del área de sombra y el valor de píxel del píxel de interés se ajusta a 1.

Primera condición: el valor de píxel del mapa de indicadores es 1 ($M_F(u_2, v_2) = 1$)

Segunda condición: el grado de similitud en el color es menor o igual que el valor umbral ($C(u_2, v_2) \leq \text{Thr}$)

Tercera condición: el grado de similitud en la textura es menor o igual que un valor umbral ($W(u_2, v_2) \leq \text{Thr2}$). Por otro lado, en el caso en que estas tres condiciones no se cumplen, se determina que el píxel de interés en el mapa de indicadores no es el píxel correspondiente al píxel del área de sombra, y el valor de píxel del píxel de interés se ajusta a 0.

Tal como se ha descrito anteriormente, en los candidatos que tienen una alta probabilidad de ser una sombra, el píxel cuya diferencia de color es pequeña y cuya diferencia de textura es pequeña entre la imagen de segundo plano y la imagen capturada puede considerarse como el píxel del área de sombra. Debido a esto, el píxel que satisface las tres condiciones descritas anteriormente se considera el píxel del área de sombra. El píxel que satisface las tres condiciones descritas anteriormente es, en otras palabras, el píxel cuyo valor de píxel del mapa de indicadores es 1, cuyo grado de similitud en el color es menor o igual que el valor umbral, es decir, el color es similar entre la imagen de segundo plano y la imagen capturada, y cuyo grado de similitud en la textura es inferior o igual que el valor umbral, es decir, la textura es similar entre la imagen de segundo plano y la imagen capturada. Lo anterior es el procesamiento para extraer un área de primer plano, que es realizado por el aparato 100 de procesamiento de imágenes en la presente realización.

<Acerca de los resultados de la presente realización>

A continuación, los resultados de la presente realización se explican usando la figura 20. En la figura 20, una imagen 2004 es una imagen de segundo plano capturada desde un punto de visión de interés 2001 y una imagen 2005 es una imagen capturada, obtenida al capturar las imágenes de un objeto 2002 acompañado de una sombra y un objeto 2003 que existe en la superficie del terreno desde el punto de visión de interés 2001. Una imagen 2006 es una imagen que representa un área que tiene una alta probabilidad de ser una sombra en la imagen 2005, que es detectada por el aparato 100 de procesamiento de imágenes de la cuarta realización o el aparato 100 de procesamiento de imágenes de la presente realización. En la imagen 2006, el píxel del área que tiene una alta probabilidad de ser una sombra se representa como un píxel blanco.

La aplicación de la cuarta realización al caso de la figura 20 creará una imagen 2007 de primer plano en la que el objeto de primer plano que no incluye una sombra se extrae considerando el píxel cuya diferencia de color es pequeña entre la imagen de segundo plano y la imagen capturada como el píxel del área de sombra en el área que tiene una alta probabilidad de ser una sombra. En la imagen 2007 de primer plano, el área del objeto 2002 de primer plano se extrae sin extraer el área de sombra que acompaña al objeto 2002. Sin embargo, en la imagen 2007 de primer plano, el color del objeto 2003 que existe en la superficie del terreno y el color de la imagen 2004 de segundo plano son similares y, por tanto, se determina erróneamente que el objeto 2003 es una sombra y, como resultado de esto, no se extrae el área del objeto 2003 de primer plano.

En contraste con esto, en la presente realización, se crea una imagen 2008 de primer plano considerando el píxel cuya diferencia de color entre la imagen de segundo plano y la imagen capturada es pequeña y cuya diferencia de textura entre la imagen de segundo plano y la imagen capturada es pequeña, como el píxel del área de sombra en el área que tiene una alta probabilidad de ser una sombra. En la imagen 2008 de primer plano, se extrae el área del objeto 2003 mientras se extrae el área del objeto 2002 de primer plano sin extraer el área de la sombra que acompaña al objeto 2002. Tal como se ha descrito anteriormente, de acuerdo con la presente realización, es posible extraer con alta precisión un objeto de primer plano que no incluye una sombra, independientemente del color.

(Otras realizaciones)

Las realizaciones de la presente invención pueden ser una variedad de realizaciones, sin limitarse a las realizaciones cuarta y quinta descritas anteriormente. Por ejemplo, en las realizaciones cuarta y quinta descritas anteriormente, el valor de píxel en el espacio RGB de la imagen objetivo y la imagen de segundo plano se usa para calcular el grado de similitud en el color, pero la información que se usa no se limita a esto. Por ejemplo, también puede ser posible calcular el grado de similitud en el color utilizando un valor de píxel en un espacio de color diferente, tal como HSV y Lab.

Además, en las realizaciones cuarta y quinta descritas anteriormente, en el momento de realizar la conversión de proyección de una imagen, solo se toma como base un plano de la superficie del terreno, pero también puede ser posible usar una pluralidad de planos paralelos a la superficie del terreno como base. En este momento, la superficie del terreno puede estar o no incluida en la pluralidad de planos que se utilizarán como base. Por ejemplo, también puede ser posible detectar un área que tiene una alta probabilidad de ser una sombra estableciendo una pluralidad de planos dividiendo por igual el espacio entre una altura de 0 cm y una altura de 1 cm desde la superficie del terreno y usando todas las imágenes convertidas obtenidas por conversión de proyección con cada uno de los planos establecidos como base. Al hacerlo, mejoran las propiedades robustas para un error en los parámetros de la cámara.

Además, en las realizaciones cuarta y quinta descritas anteriormente, un área del primer plano común que tiene una alta probabilidad de ser una sombra se detecta usando todas las imágenes de primer plano adquiridas en una pluralidad de puntos de visión, pero también puede ser posible detectar un área del primer plano común utilizando solo una parte de las imágenes de primer plano adquiridas.

La realización o realizaciones de la presente invención también pueden realizarse mediante un ordenador de un sistema o aparato que lee y ejecuta instrucciones ejecutables por ordenador (por ejemplo, uno o varios programas) grabados en un medio de almacenamiento (al que también se puede hacer referencia de manera más completa como un 'medio de almacenamiento no transitorio legible por ordenador') para realizar las funciones de una o varias de las realizaciones descritas anteriormente y/o que incluye uno o varios circuitos (por ejemplo, circuito integrado de aplicación específica (ASIC, application specific integrated circuit)) para realizar las funciones de una o varias de las realizaciones descritas anteriormente, y mediante un procedimiento realizado por el ordenador del sistema o aparato, por ejemplo, leyendo y ejecutando las instrucciones ejecutables del ordenador desde el medio de almacenamiento para realizar las funciones de una o varias de las realizaciones descritas anteriormente y/o controlar el uno o varios circuitos para realizar las funciones de una o varias de las realizaciones descritas anteriormente. El ordenador puede comprender uno o varios procesadores (por ejemplo, unidad central de procesamiento (CPU, central processing unit), unidad de microprocesador (MPU, micro processing unit)) y puede incluir una red de ordenadores independientes o procesadores independientes para leer y ejecutar las instrucciones ejecutables por ordenador. Las instrucciones ejecutables por ordenador se pueden proporcionar al ordenador, por ejemplo, desde una red o desde el medio de almacenamiento. El medio de almacenamiento puede incluir, por ejemplo, uno o varios de un disco duro, una memoria de acceso aleatorio (RAM), una memoria de solo lectura (ROM), un almacenamiento de sistemas informáticos distribuidos, un disco óptico (como un disco compacto (CD), disco versátil digital (DVD) o Blu-ray Disc (BD)™), un dispositivo de memoria flash, una tarjeta de memoria y similares.

De acuerdo con la presente realización, es posible extraer con alta precisión un objeto de primer plano.

Si bien la presente invención se ha descrito con referencia a realizaciones a modo de ejemplo, debe entenderse que la invención no se limita a las realizaciones a modo de ejemplo descritas. El alcance de las siguientes reivindicaciones debe ser conforme a la interpretación más amplia para abarcar todas las modificaciones y estructuras y funciones equivalentes que quedan dentro del alcance de las reivindicaciones adjuntas.

REIVINDICACIONES

1. Aparato de procesamiento de imágenes, que comprende:

- 5 una unidad (201) de adquisición de imágenes objetivo configurada para adquirir una imagen objetivo que es un objetivo de extracción de un área de primer plano;
- una unidad (202) de adquisición de imágenes de referencia configurada para adquirir una pluralidad de imágenes de referencia que incluyen una imagen cuyo punto de visión es diferente del de la imagen objetivo;
- 10 una unidad (203) de conversión configurada para convertir la pluralidad de imágenes de referencia adquiridas por la unidad de adquisición de imágenes de referencia basándose en un punto de visión correspondiente a la imagen objetivo; y
- una unidad (207) de extracción configurada para extraer un área de primer plano de la imagen objetivo utilizando datos relacionados con un grado de coincidencia calculado en la pluralidad de imágenes de referencia convertidas por la unidad de conversión.

2. Aparato de procesamiento de imágenes, según la reivindicación 1, en el que

la unidad de conversión realiza una conversión de proyección para convertir la pluralidad de imágenes de referencia en imágenes respectivas en un caso en el que la pluralidad de imágenes de referencia se ve desde un punto de visión de la imagen objetivo con una superficie del terreno como base.

3. Aparato de procesamiento de imágenes, según la reivindicación 1, en el que

la unidad de conversión realiza una conversión de proyección para convertir la pluralidad de imágenes de referencia en imágenes respectivas en un caso en el que la pluralidad de imágenes de referencia se ve desde el punto de visión de interés con una pluralidad de planos paralelos a una superficie del terreno como base, además de la superficie del terreno.

4. Aparato de procesamiento de imágenes, según la reivindicación 1, que comprende, además:

una unidad de especificación configurada para especificar un punto de visión de la imagen objetivo; y
una unidad de selección configurada para seleccionar una imagen de referencia correspondiente a un punto de visión cuya distancia desde un punto de visión especificado por la unidad de especificación es el más corto de los puntos de visión de la pluralidad de imágenes de referencia como una imagen de referencia de base, en el que
35 la unidad de extracción corrige una imagen de referencia de base seleccionada por la unidad de selección utilizando datos relacionados con el grado de coincidencia y extrae un área de primer plano de la imagen objetivo utilizando la imagen de referencia de base ya corregida.

5. Aparato de procesamiento de imágenes, según la reivindicación 1, que comprende, además:

una unidad de especificación configurada para especificar un punto de visión de la imagen objetivo; y
una unidad de generación configurada para generar una imagen que muestra un aspecto desde un punto de visión especificado por la unidad de especificación desde una pluralidad de imágenes de referencia, en el que
45 la unidad de extracción corrige una imagen generada por la unidad de generación utilizando datos relacionados con el grado de coincidencia y extrae un área de primer plano de la imagen objetivo utilizando la imagen corregida.

6. Aparato de procesamiento de imágenes, según la reivindicación 4, en el que

la unidad de corrección corrige, en un caso en el que los datos relacionados con el grado de coincidencia indican que un grado de coincidencia entre píxeles de interés en la pluralidad de imágenes de referencia ya convertidas por la unidad de conversión es menor que un valor umbral predeterminado, un valor del píxel de interés en la imagen de referencia de base y
50 la unidad de corrección no corrige, en un caso en el que los datos relacionados con el grado de coincidencia indican que un grado de coincidencia entre los píxeles de interés en la pluralidad de imágenes de referencia ya convertidas por la unidad de conversión es mayor o igual que el valor umbral predeterminado, un valor del píxel de interés en la imagen de referencia de base.

7. Aparato de procesamiento de imágenes, según la reivindicación 4, que comprende, además:

una unidad de cálculo de continuidad configurada para calcular la continuidad indicando un grado de suavidad de un cambio en el valor de píxel entre píxeles de interés en una imagen de referencia ya convertida por la unidad de conversión; y
una unidad de detección configurada para detectar un píxel que es un objetivo de corrección mediante la unidad de corrección basándose en los datos relacionados con el grado de coincidencia y la continuidad.

8. Aparato de procesamiento de imágenes, según la reivindicación 1, en el que

las imágenes de referencia adquiridas por la unidad de adquisición de imágenes de referencia son imágenes obtenidas al extraer un área de primer plano de una pluralidad de imágenes cuyos puntos de visión son diferentes de los de la imagen objetivo,

- 5 la unidad de conversión convierte la pluralidad de imágenes de referencia adquiridas por la unidad de adquisición de imágenes de referencia basándose en un punto de visión correspondiente a la imagen objetivo, y
la unidad de extracción extrae un área de primer plano de la imagen objetivo utilizando datos relacionados con un grado de coincidencia calculado en la pluralidad de imágenes de referencia convertidas por la unidad de conversión.

- 10 9. Aparato de procesamiento de imágenes, según la reivindicación 8, que comprende, además:

una unidad de detección configurada para detectar un área de primer plano que existe en la misma posición en común en la pluralidad de imágenes de referencia convertidas por la unidad de conversión como un área del primer plano común, en el que

- 15 la unidad de extracción extrae un área de primer plano de la imagen objetivo modificando una imagen de referencia de un punto de visión correspondiente a la imagen objetivo basándose en el área del primer plano común.

10. Aparato de procesamiento de imágenes, según la reivindicación 9, en el que

- 20 la unidad de detección crea un mapa de indicadores en el que 1 es sustituido por un valor de píxel correspondiente a un píxel del área del primer plano común y 0 es sustituido por un valor de píxel correspondiente a un píxel que no es un píxel del área del primer plano común, y
la unidad de extracción extrae un área de primer plano de la imagen objetivo utilizando el mapa de indicadores.

- 25 11. Aparato de procesamiento de imágenes, según la reivindicación 10, que comprende, además:

una unidad de cálculo configurada para calcular un grado de similitud en el color entre un píxel de la imagen objetivo y un píxel de una imagen de segundo plano cuyo punto de visión es sustancialmente el mismo que el de la imagen objetivo y que no incluye el área de primer plano; y

- 30 una unidad de determinación configurada para determinar si un píxel del área del primer plano común es un píxel del área de sombra basándose en un valor de píxel del mapa de indicadores y el grado de similitud en el color, en el que la unidad de extracción extrae un área de primer plano de la imagen objetivo corrigiendo una imagen de referencia correspondiente a un punto de visión de la imagen objetivo basándose en los resultados de la determinación de la unidad de determinación.

- 35 12. Aparato de procesamiento de imágenes, según la reivindicación 11, en el que

la unidad de determinación determina que un píxel de interés en una imagen de referencia convertida por la unidad de conversión es un píxel del área de sombra en un caso en el que un valor de píxel del mapa de indicadores, que

- 40 corresponde al píxel de interés, es 1 y la imagen objetivo y la imagen de segundo plano son similares en color.

13. Aparato de procesamiento de imágenes, según la reivindicación 12, que comprende, además:

una unidad de cálculo configurada para calcular un grado de similitud en el color entre un píxel de la imagen objetivo y un píxel de la imagen de segundo plano; y

- 45 una unidad de cálculo configurada para calcular un grado de similitud en la textura entre un píxel de la imagen objetivo y un píxel de la imagen de segundo plano, en el que
la unidad de determinación determina si un píxel del área del primer plano común es un píxel del área de sombra basándose en un valor de píxel del mapa de indicadores, el grado de similitud en el color y el grado de similitud en la

- 50 textura.

14. Aparato de procesamiento de imágenes, según la reivindicación 13, en el que

la unidad de determinación determina que un píxel de interés en una imagen de referencia convertida por la unidad de conversión es un píxel del área de sombra en un caso en el que un valor de píxel del mapa de indicadores, que

- 55 corresponde al píxel de interés, es 1, la imagen objetivo y la imagen de segundo plano son similares en color, y la imagen objetivo y la imagen de segundo plano son similares en textura.

15. Aparato de procesamiento de imágenes, según la reivindicación 11, en el que

- 60 la imagen de segundo plano es una imagen creada basándose en una pluralidad de imágenes correspondientes a una pluralidad de momentos diferentes capturados de forma continua desde el punto de visión de interés a lo largo de una serie temporal, o una imagen capturada en una situación en la que no existe el área de primer plano.

- 65 16. Procedimiento de procesamiento de imágenes, que comprende:

una etapa (S301) de adquisición de imágenes objetivo para adquirir una imagen objetivo que es un objetivo de extracción de un área de primer plano;

una etapa (S302) de adquisición de imágenes de referencia para adquirir una pluralidad de imágenes de referencia que incluyen una imagen cuyo punto de visión es diferente del de la imagen objetivo;

5 una etapa (S303) de conversión para convertir la pluralidad de imágenes de referencia adquiridas en la etapa de adquisición de imágenes de referencia basándose en un punto de visión correspondiente a la imagen objetivo; y
una etapa (S313) de extracción para extraer un área de primer plano de la imagen objetivo utilizando datos relacionados con un grado de coincidencia calculado en la pluralidad de imágenes de referencia convertidas en la etapa de conversión.

10 17. Programa para hacer que un ordenador realice un procedimiento de procesamiento de imágenes, comprendiendo el procedimiento:

una etapa (S301) de adquisición de imágenes objetivo para adquirir una imagen objetivo que es un objetivo de extracción de un área de primer plano;

15 una etapa (S302) de adquisición de imágenes de referencia para adquirir una pluralidad de imágenes de referencia que incluye una imagen cuyo punto de visión es diferente del de la imagen objetivo;

una etapa (S303) de conversión para convertir la pluralidad de las imágenes de referencia adquiridas en la etapa de adquisición de imágenes de referencia basándose en un punto de visión correspondiente a la imagen objetivo; y

20 una etapa (S313) de extracción para extraer un área de primer plano de la imagen objetivo utilizando datos relacionados con un grado de coincidencia calculado en la pluralidad de imágenes de referencia convertidas en la etapa de conversión.

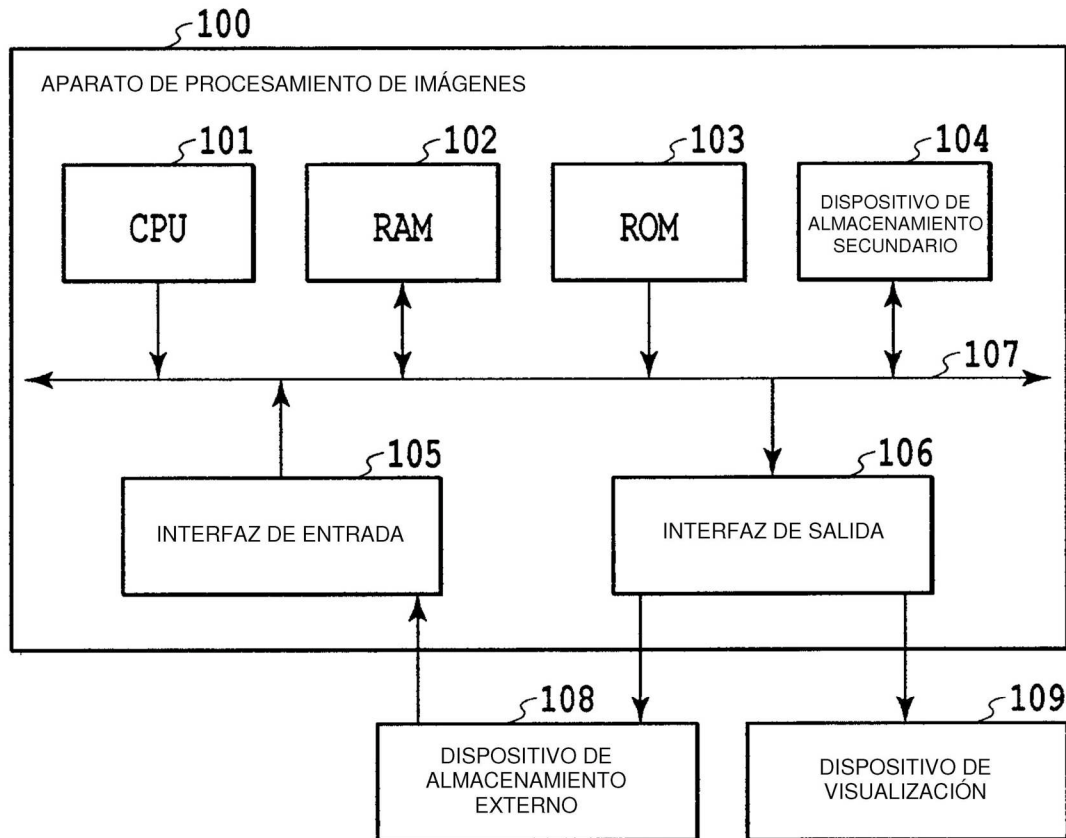


FIG.1

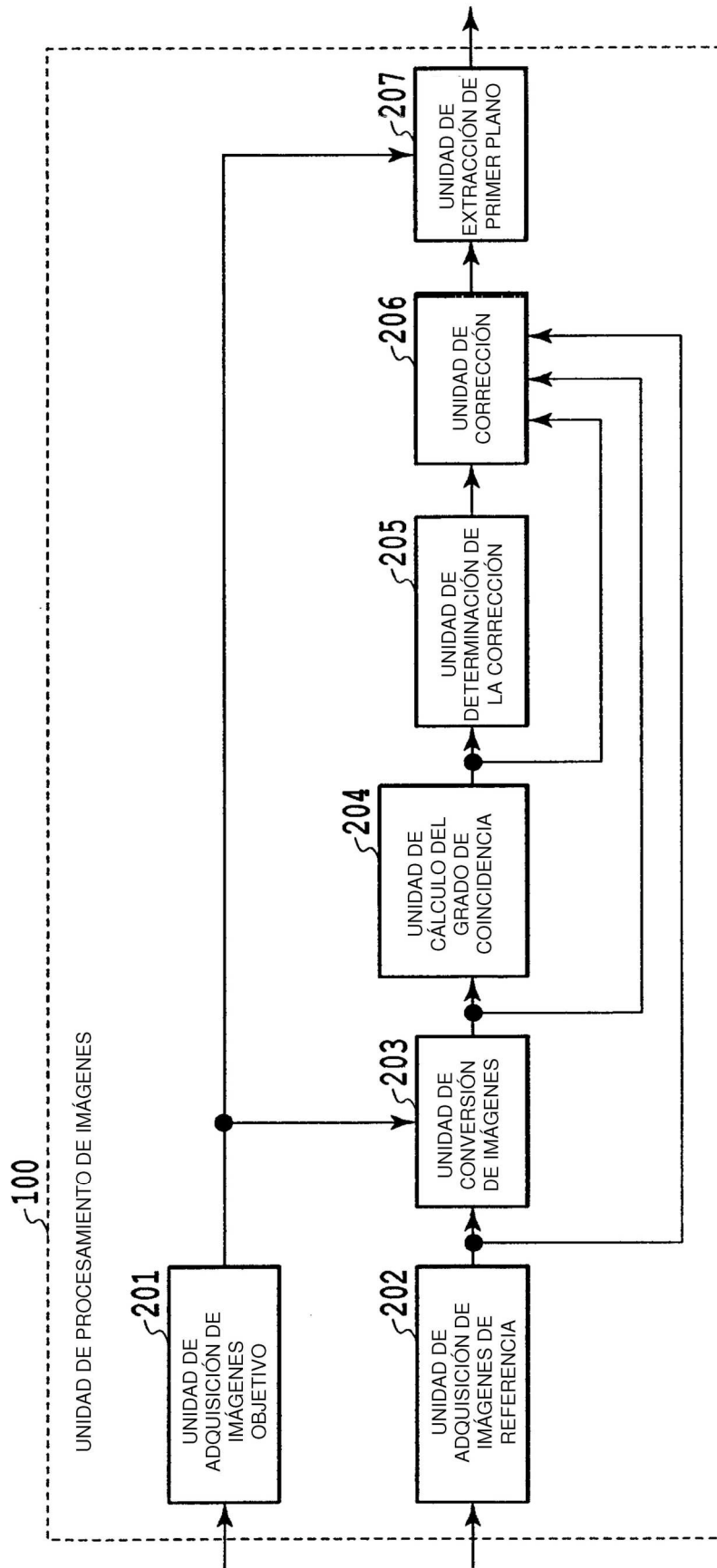
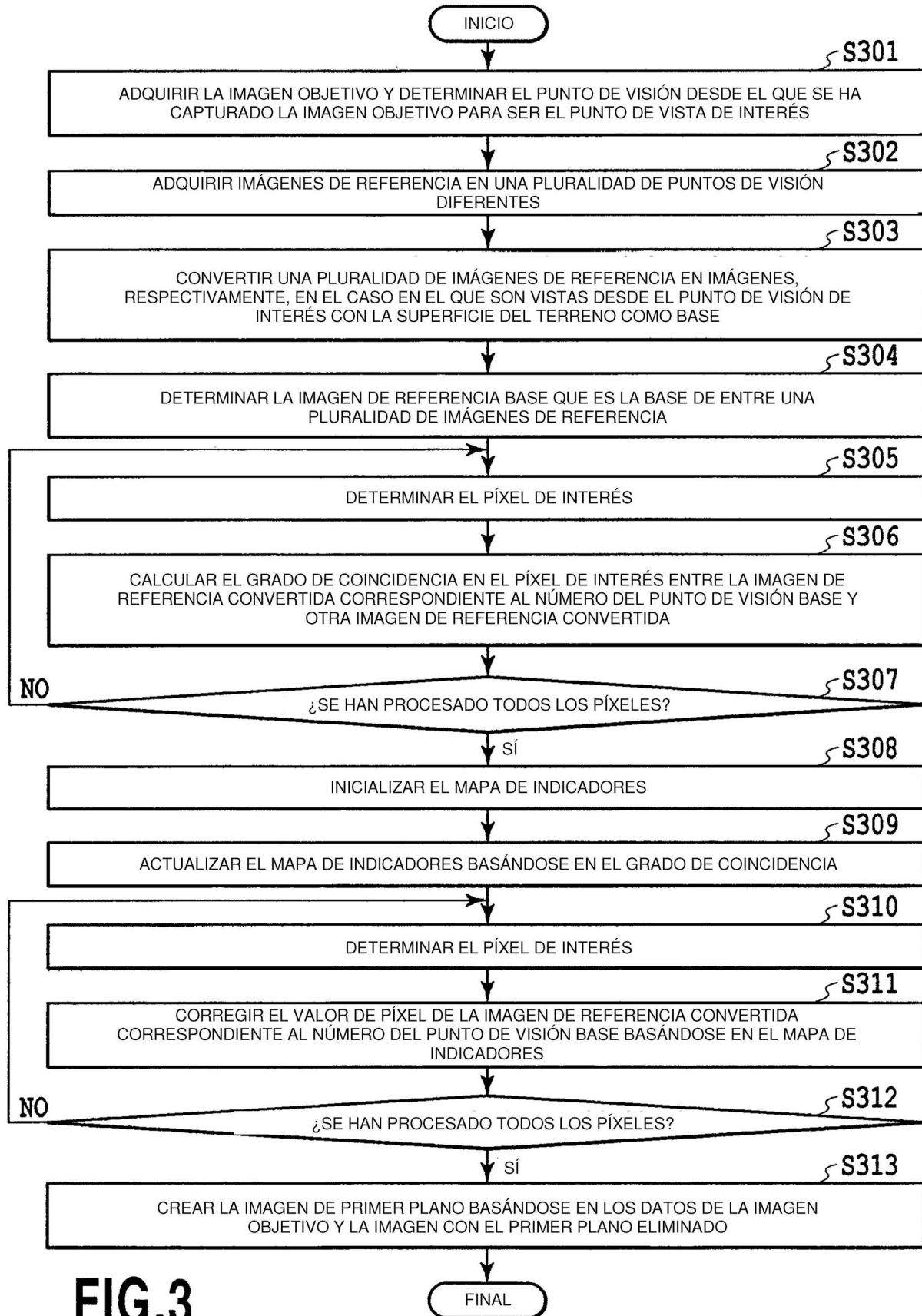


FIG.2



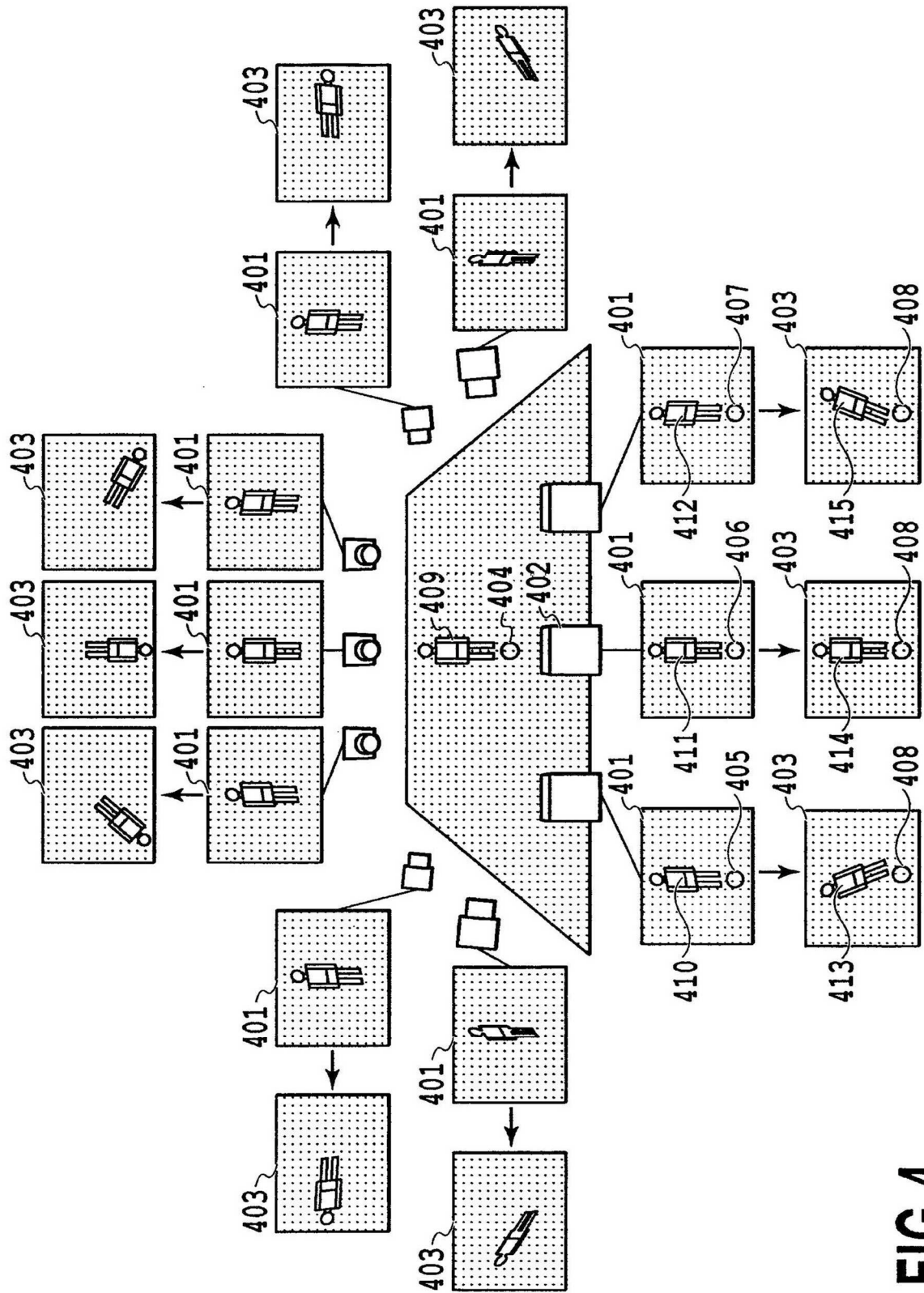


FIG. 4

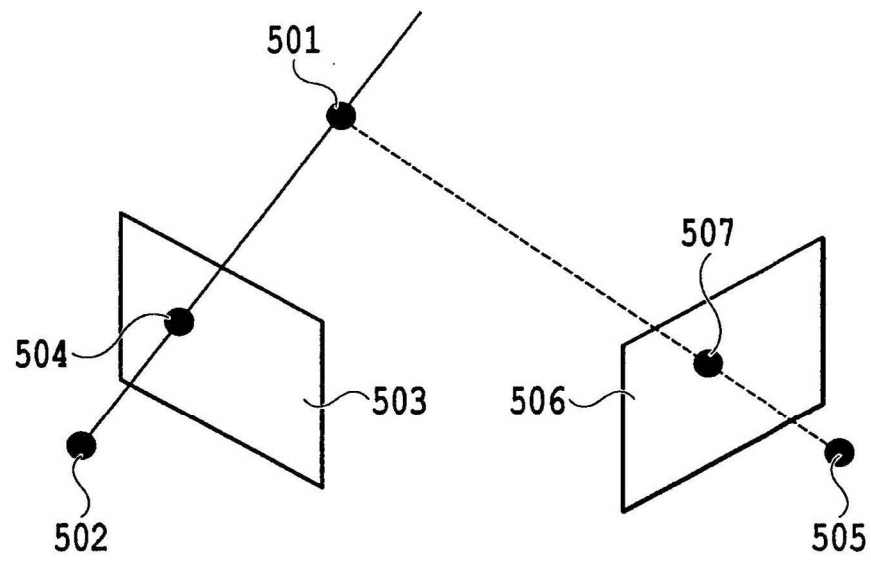


FIG.5

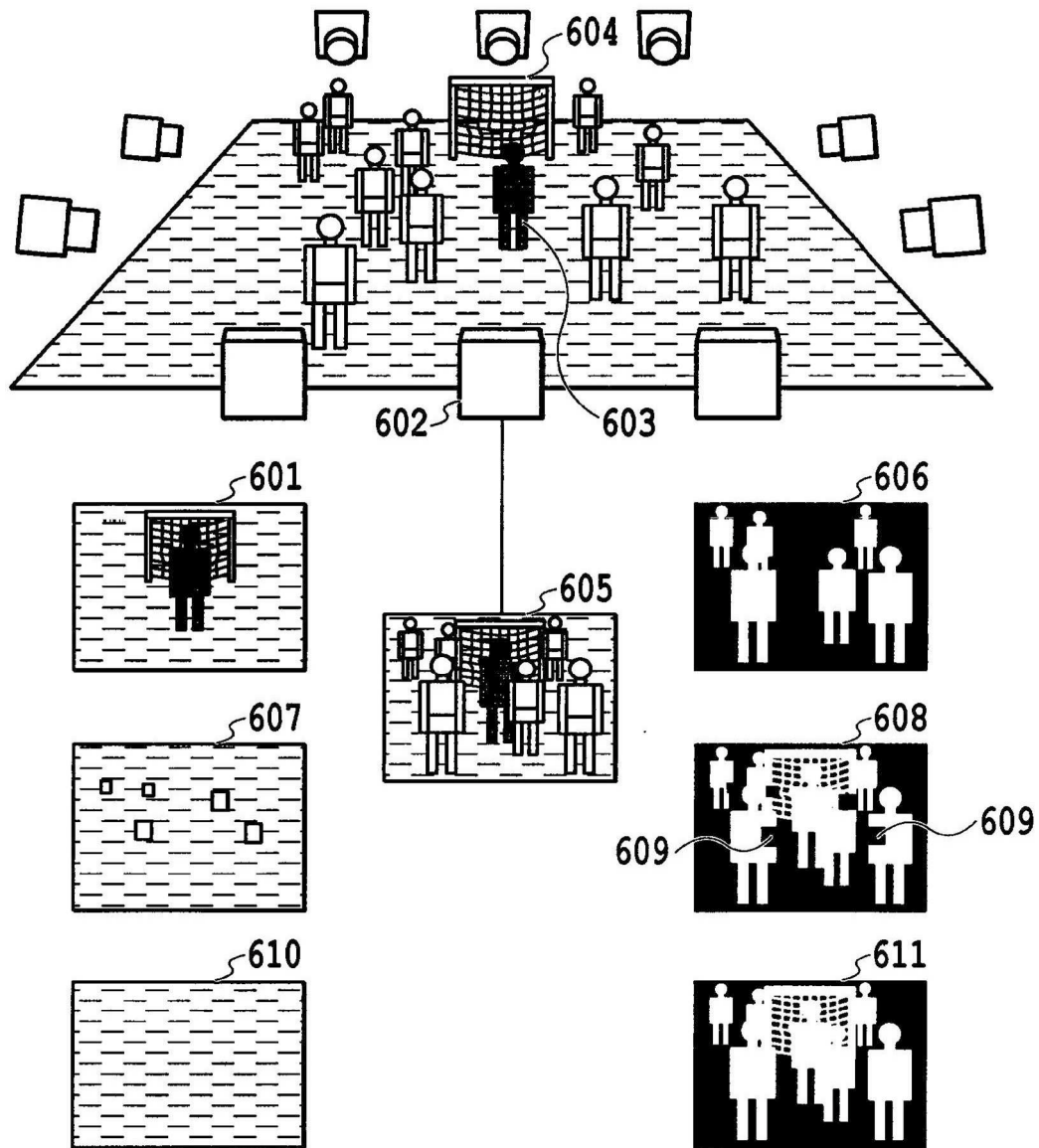


FIG.6

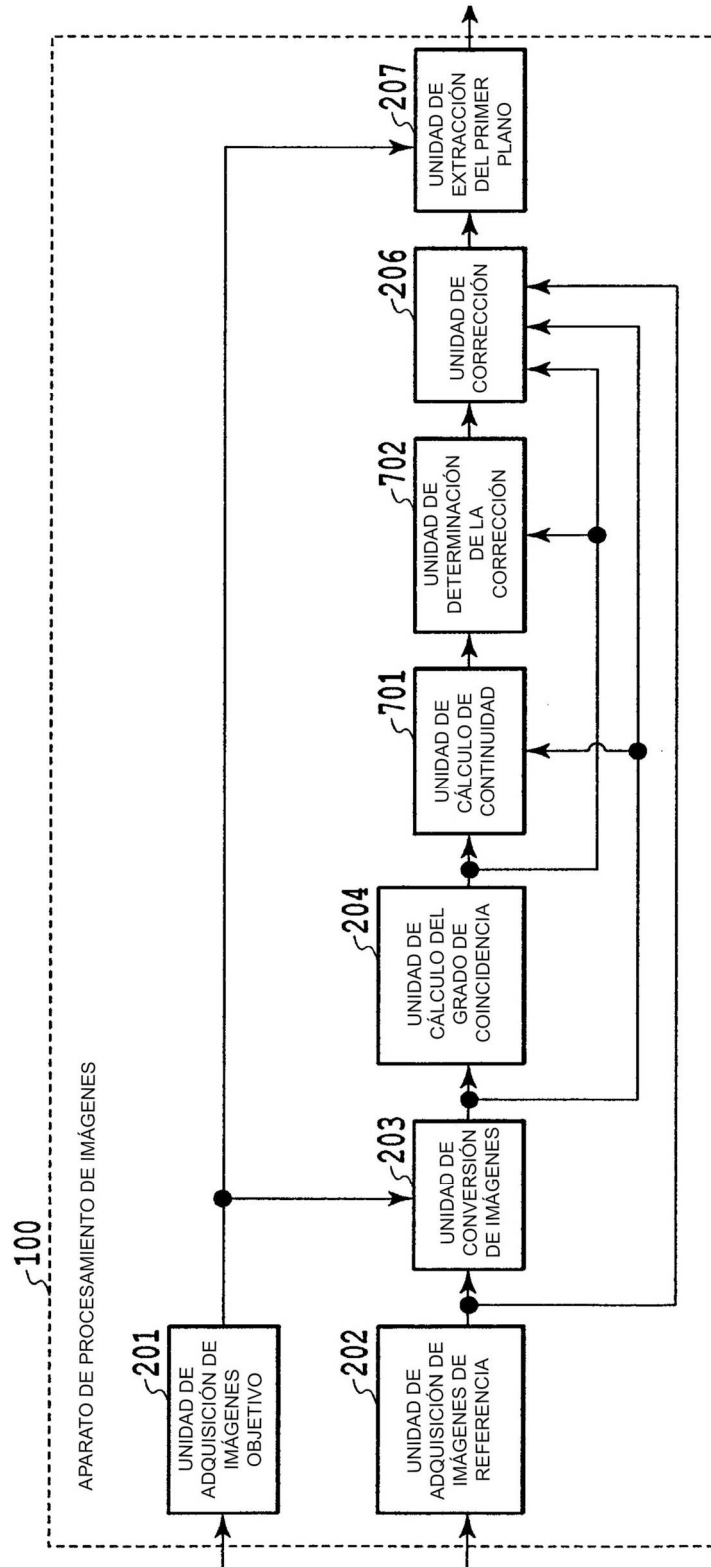
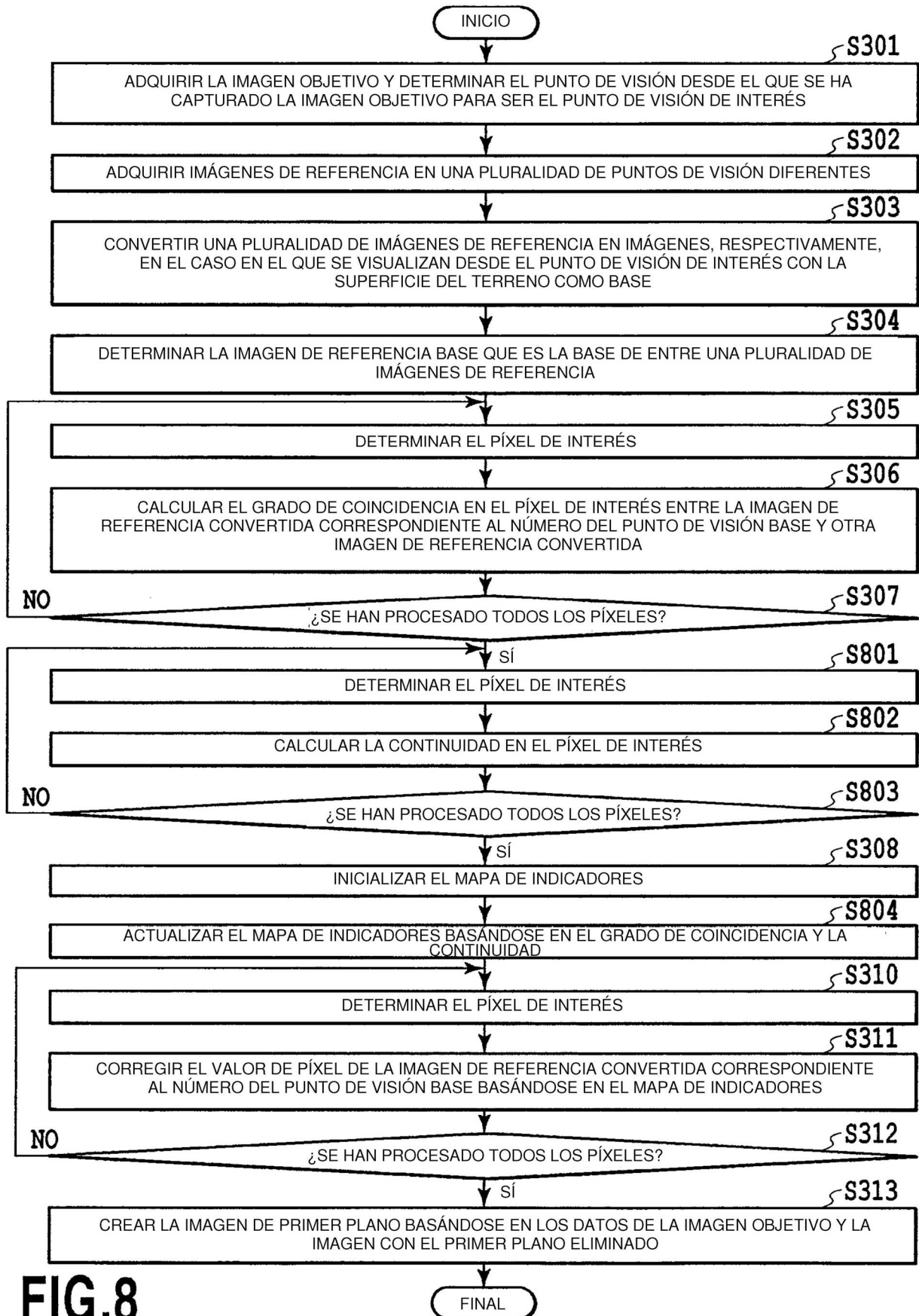


FIG.7



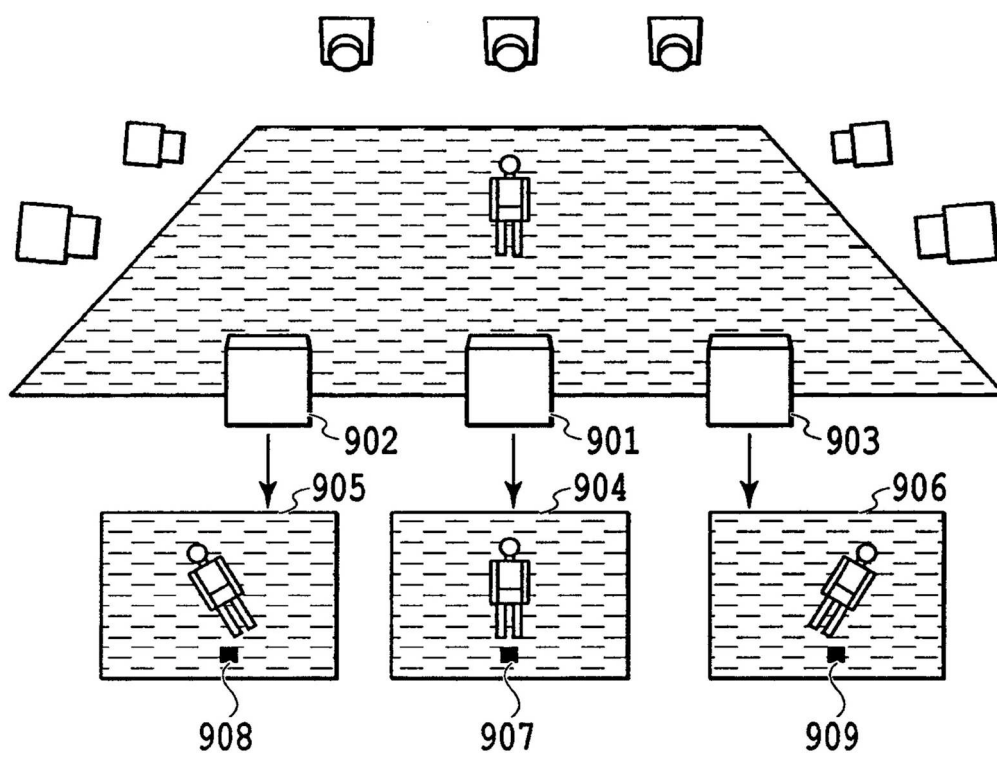


FIG.9

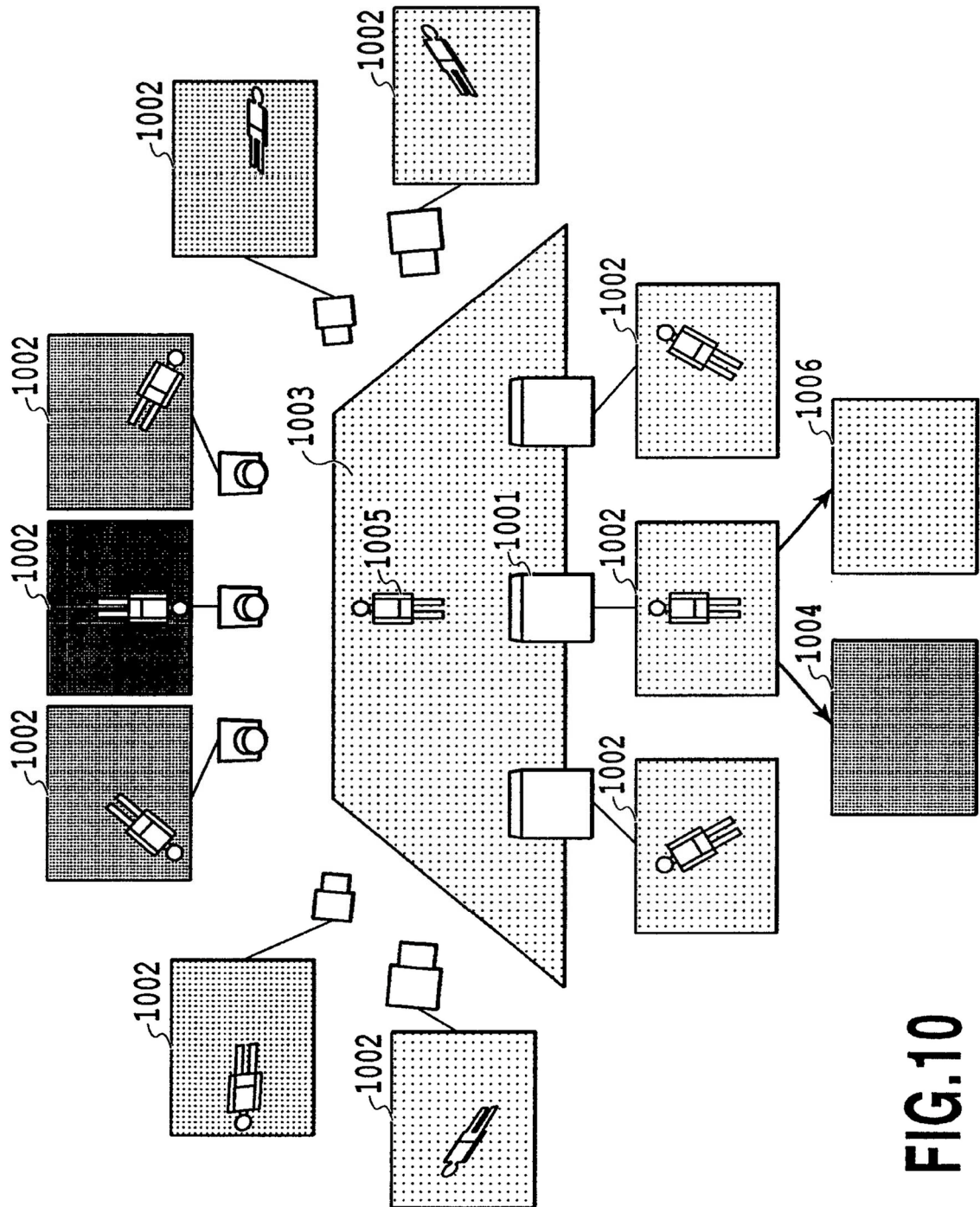


FIG.10

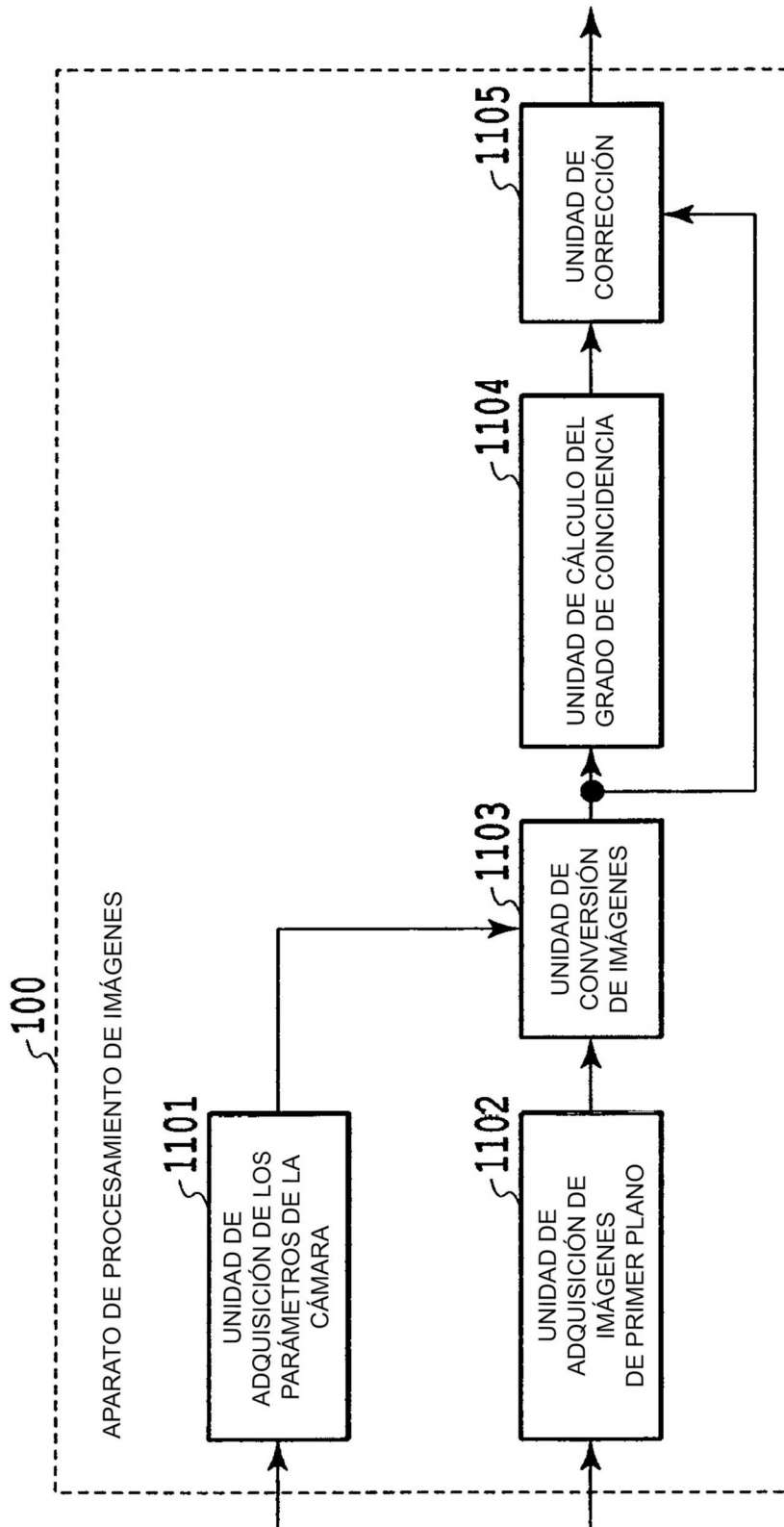


FIG.11

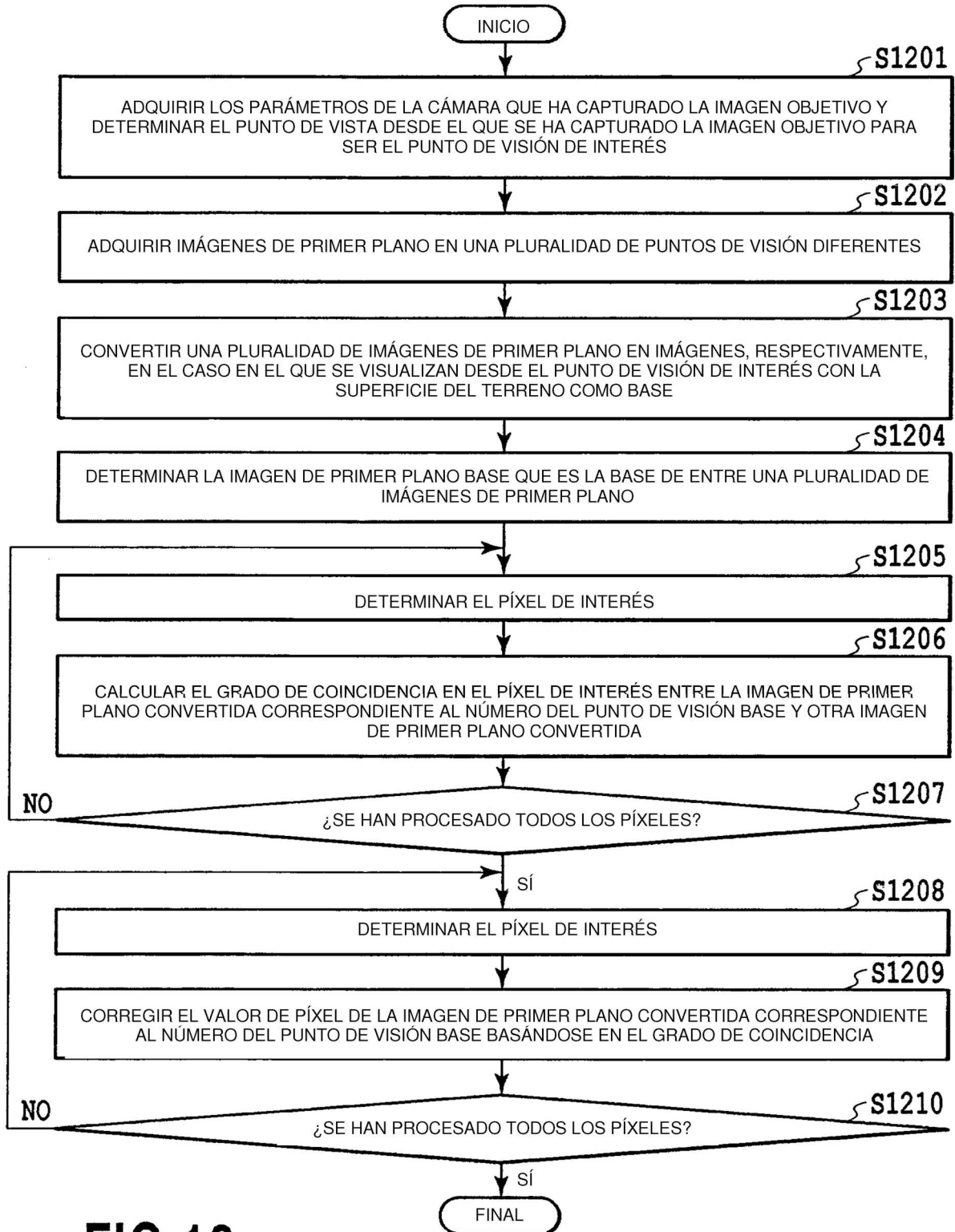


FIG.12

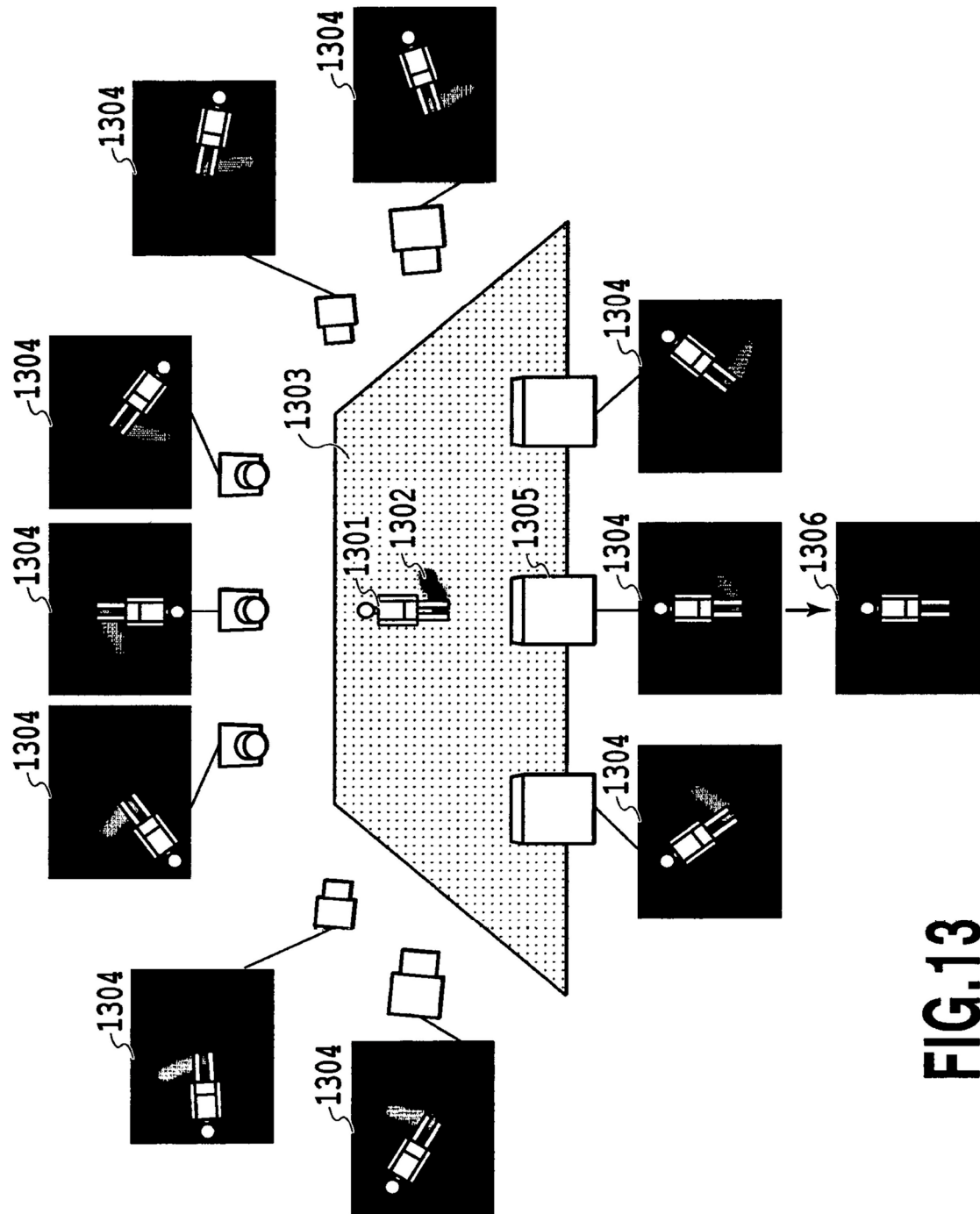


FIG.13

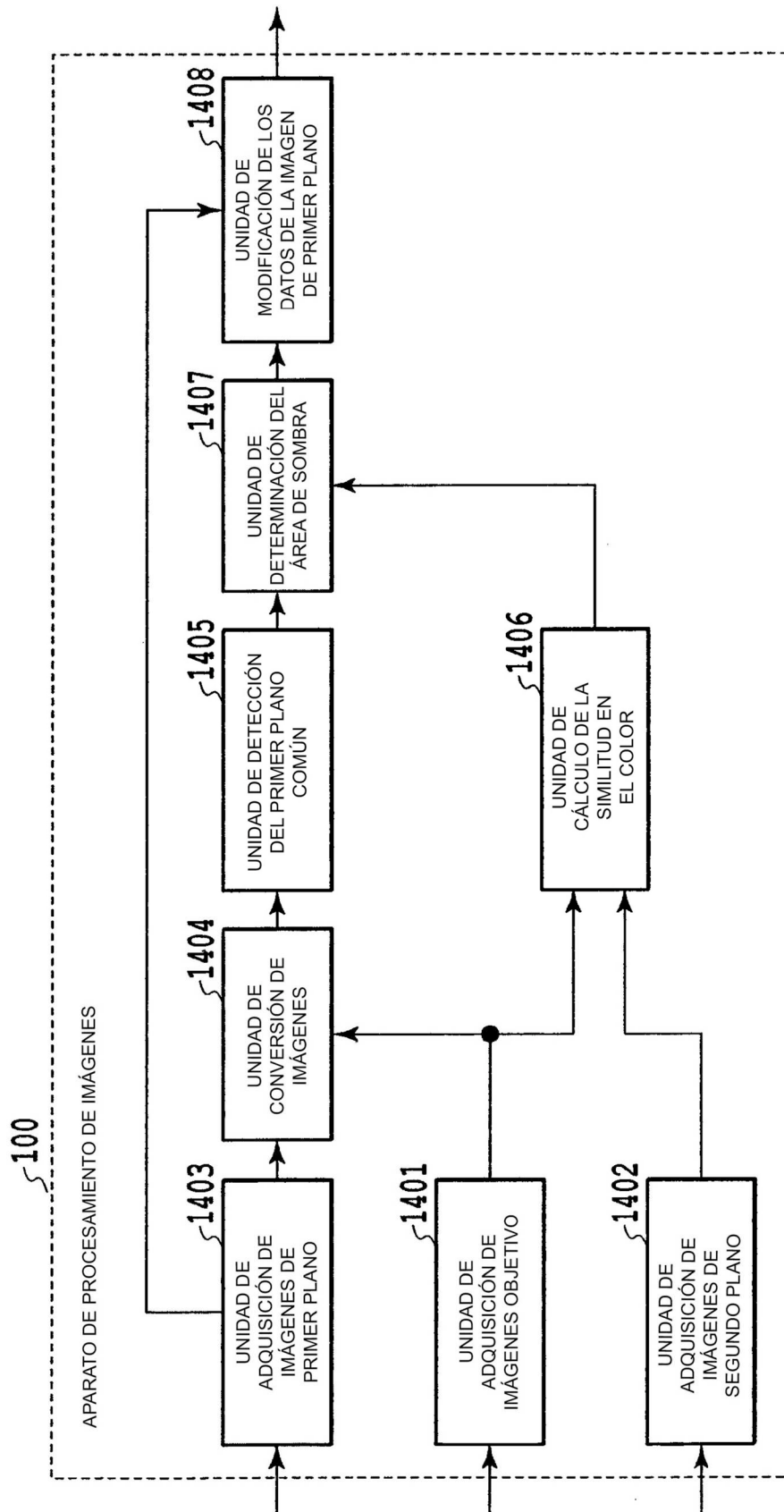
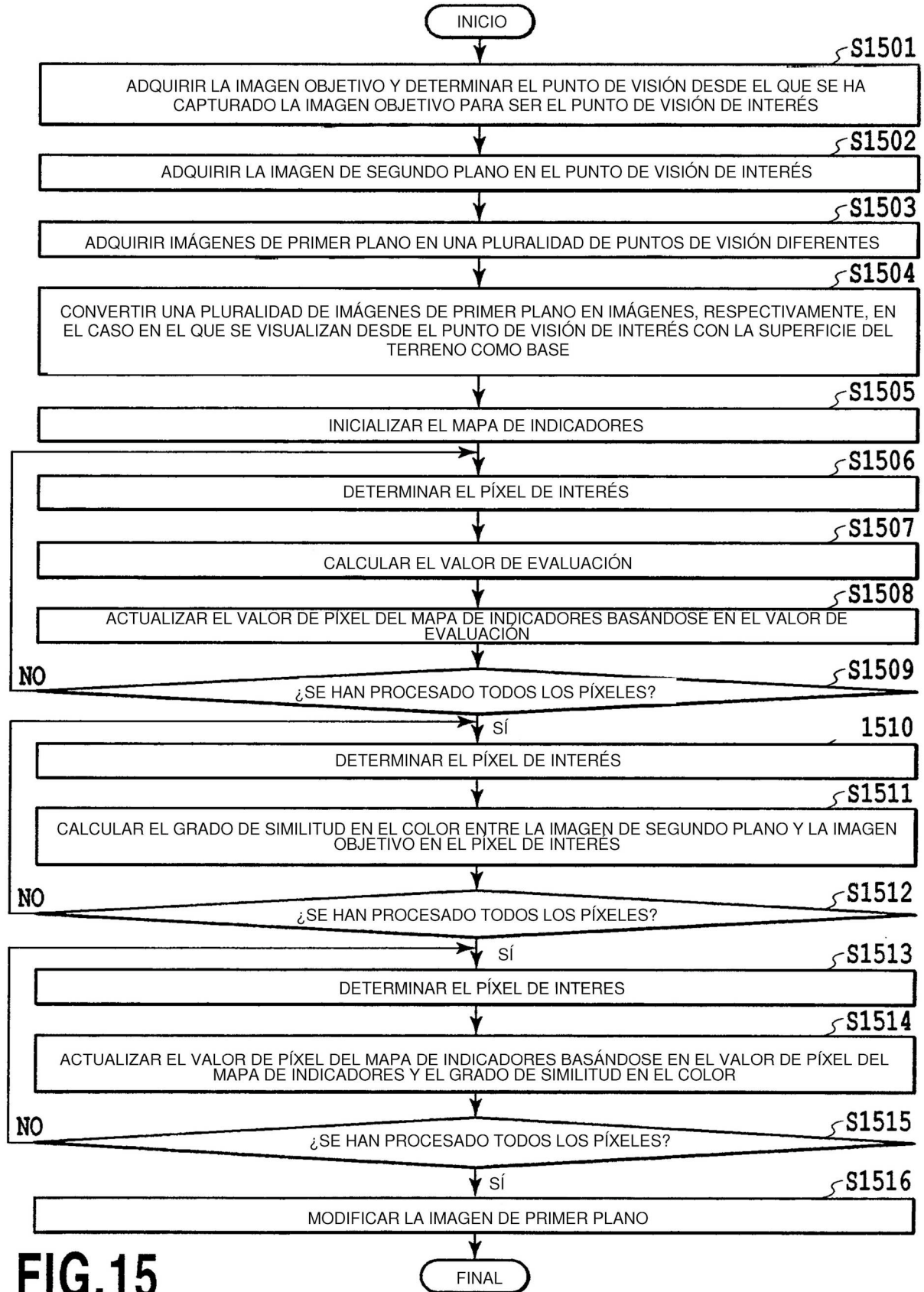


FIG.14



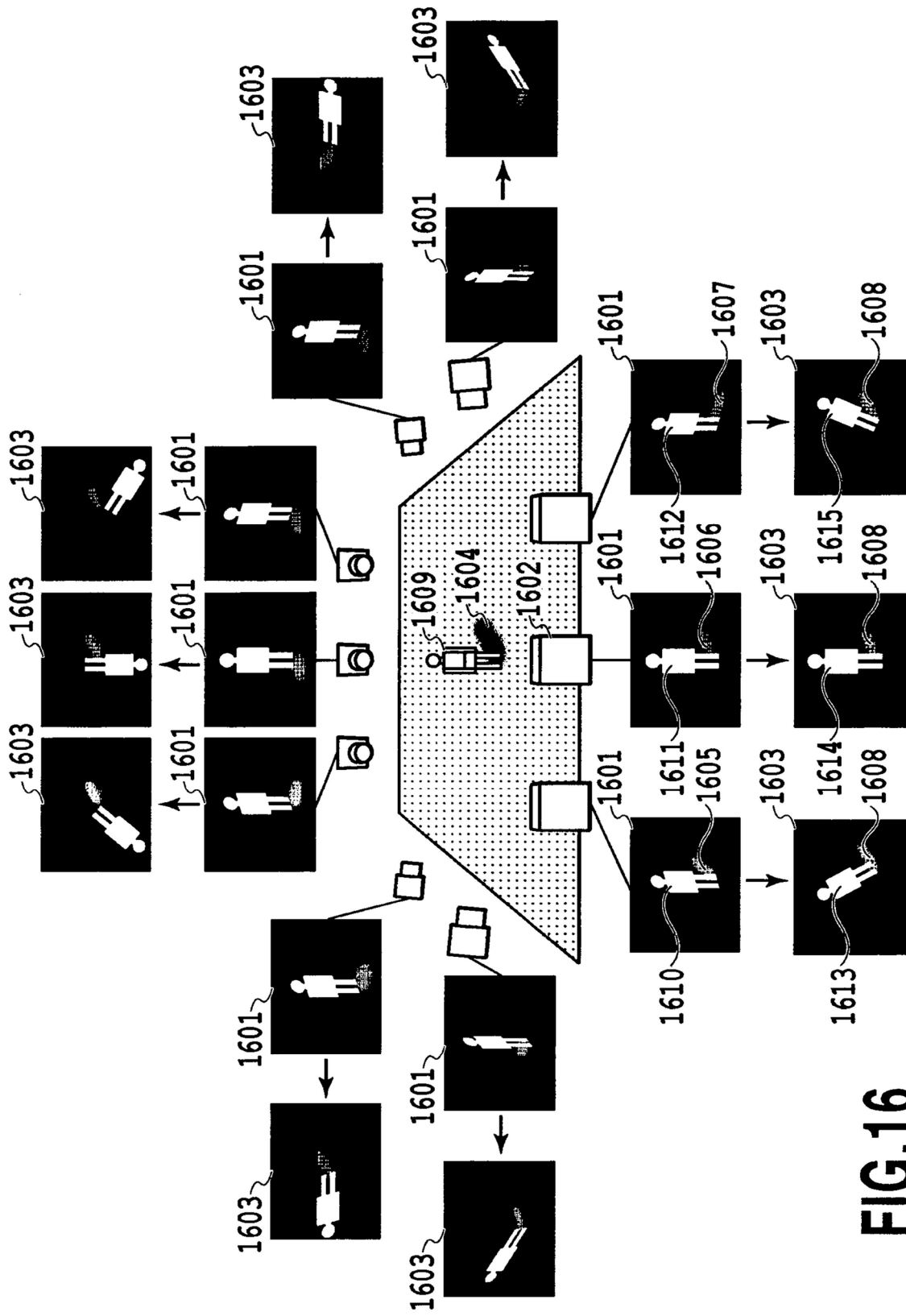


FIG. 16

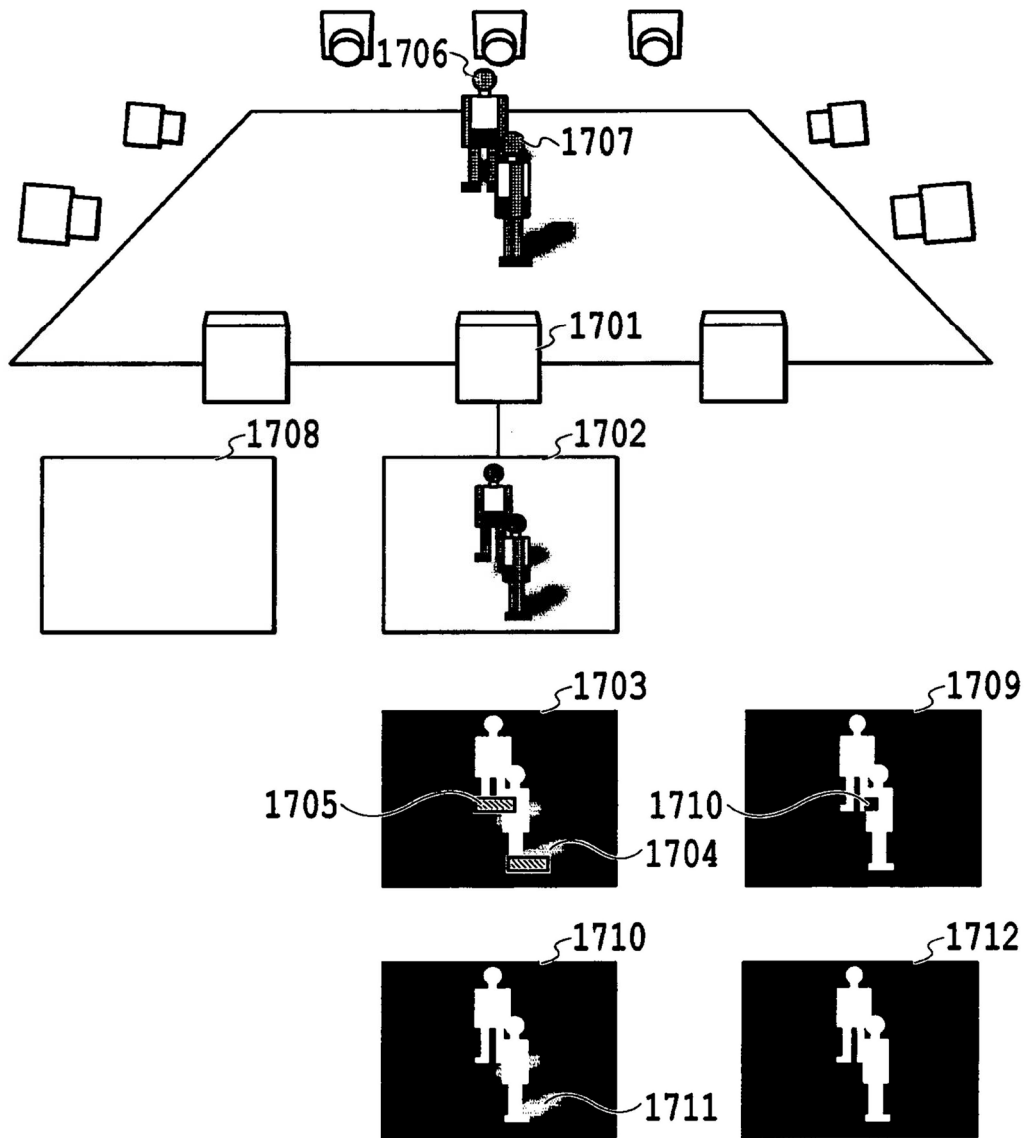


FIG.17

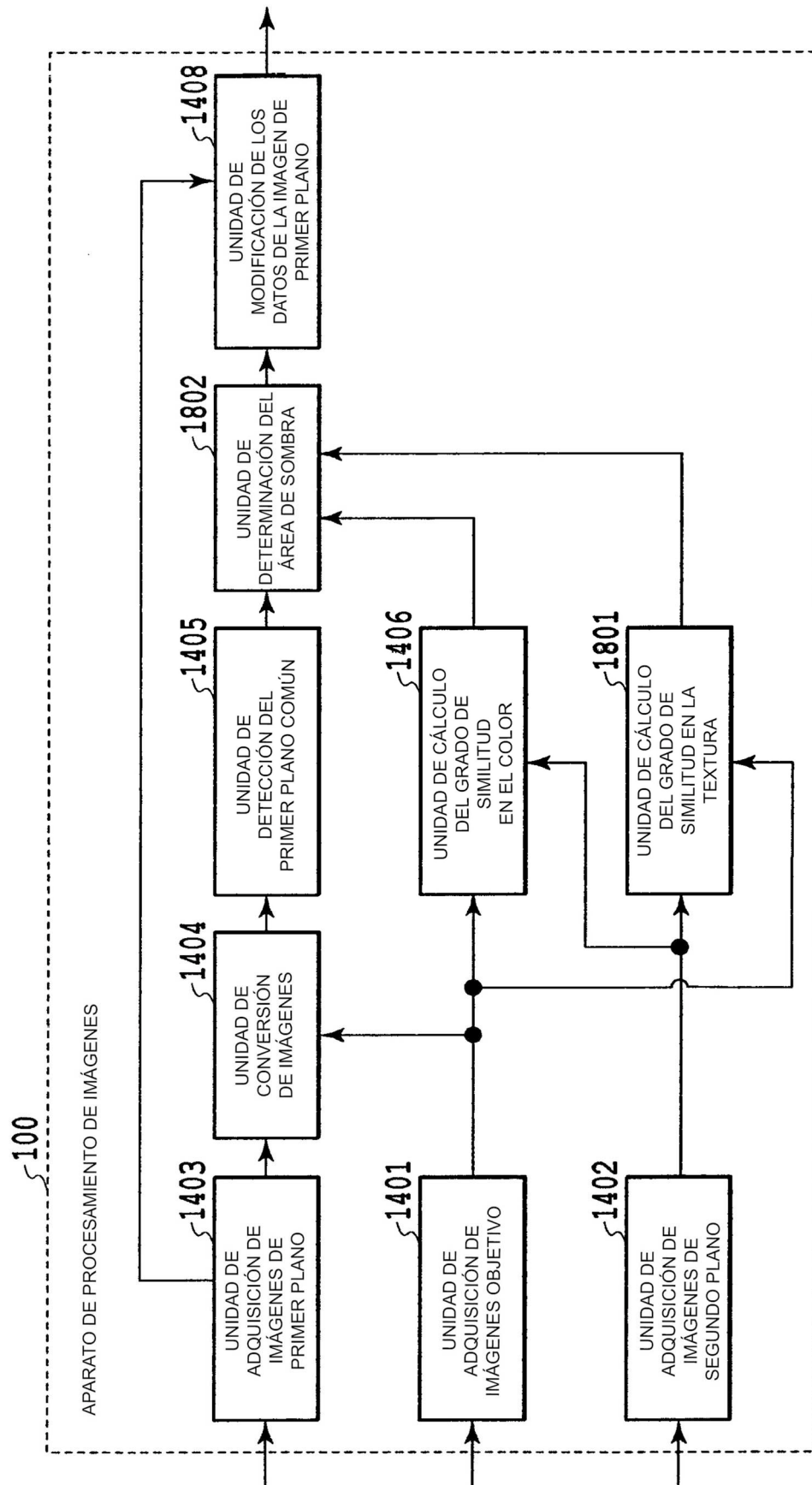


FIG.18

FIG.19

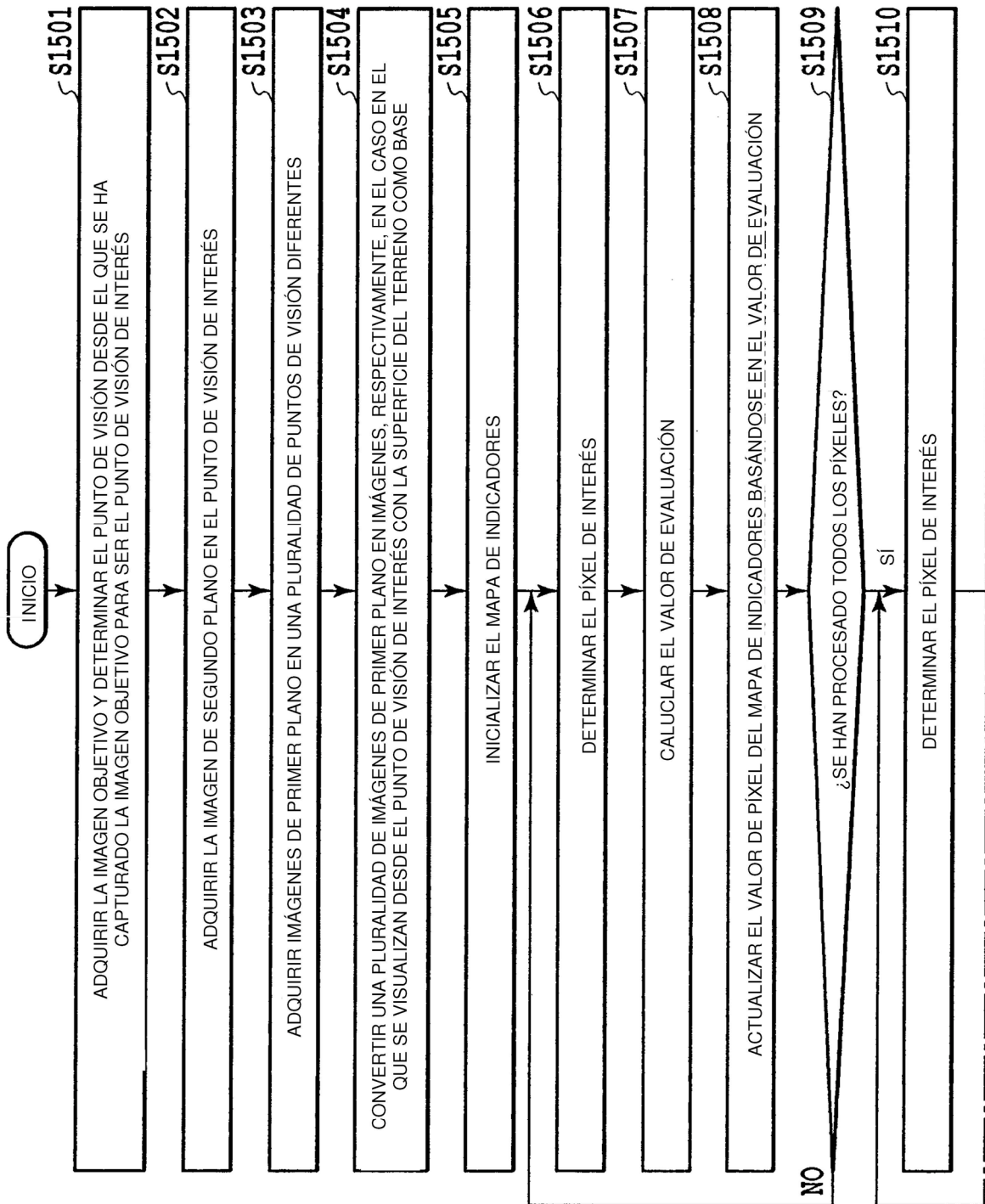
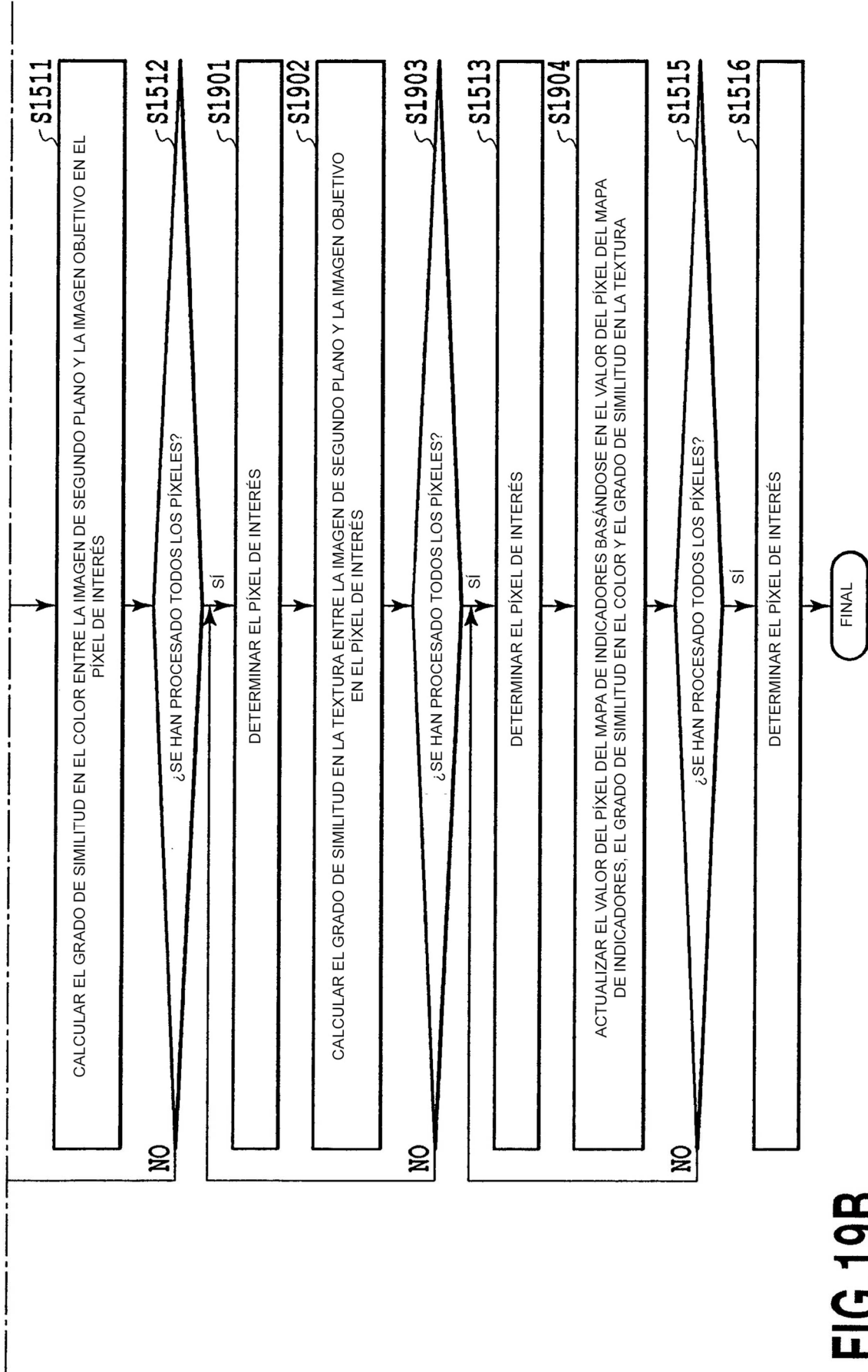


FIG.19A



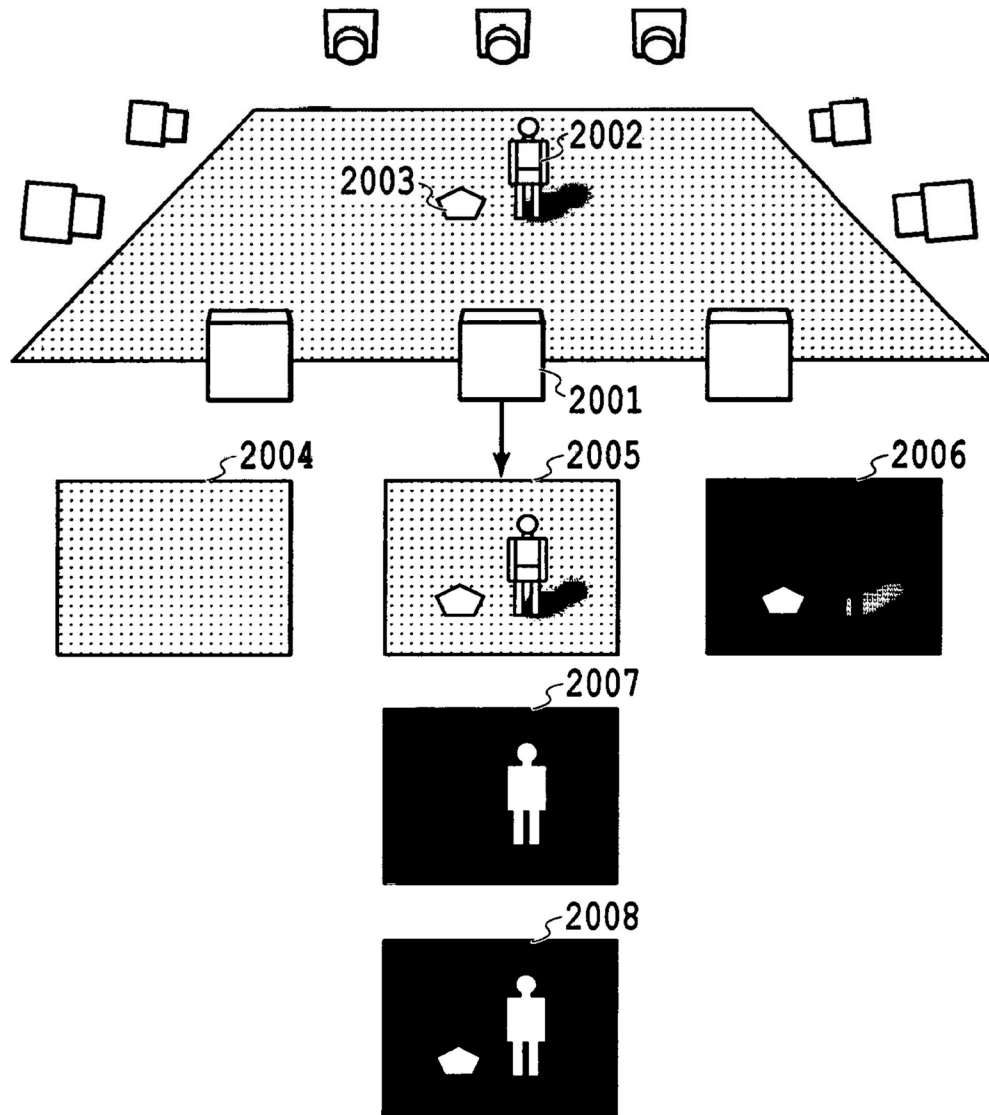


FIG.20

REFERENCIAS CITADAS EN LA DESCRIPCIÓN

Esta lista de referencias citada por el solicitante es únicamente para mayor comodidad del lector. No forman parte del documento de la Patente Europea. Incluso teniendo en cuenta que la compilación de las referencias se ha efectuado con gran cuidado, los errores u omisiones no pueden descartarse; la EPO se exime de toda responsabilidad al respecto.

Documentos de patentes citados en la descripción

- JP 2012104053 A
- JP 2014230180 A
- US 20100153449 A1

Literatura no patente citada en la descripción

- **XU, M. et al.** A Multiview Approach to Robust Detection in the Presence of Cast Shadows. Proc. Sixth IEEE Intl. Conf. on Images and Graphics, 2011, 494-499