

19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **3 009 013**

51 Int. Cl.:

G01S 17/42	(2006.01)	H04N 23/698	(2013.01)
G01S 17/86	(2010.01)		
G01S 17/89	(2010.01)		
G01S 7/481	(2006.01)		
G03B 17/56	(2011.01)		
G03B 37/02	(2011.01)		
H04N 23/51	(2013.01)		
H04N 23/55	(2013.01)		
H04N 23/58	(2013.01)		
H04N 23/695	(2013.01)		

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

- 86 Fecha de presentación y número de la solicitud internacional: **30.12.2020 PCT/US2020/067474**
- 87 Fecha y número de publicación internacional: **08.07.2021 WO21138427**
- 96 Fecha de presentación y número de la solicitud europea: **30.12.2020 E 20909333 (5)**
- 97 Fecha y número de publicación de la concesión europea: **26.02.2025 EP 4085302**

54 Título: **Sistema y método de captura y generación de imágenes tridimensionales panorámicas**

30 Prioridad:

30.12.2019 US 201962955414 P

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

25.03.2025

73 Titular/es:

**MATTERPORT, INC. (100.00%)
352 East Java Drive
Sunnyvale, CA 94089, US**

72 Inventor/es:

**GAUSEBECK, DAVID ALAN;
STROMBERG, KIRK;
MARZANO, LOUIS D.;
PROCTOR, DAVID;
SAKAKIBARA, NAOTO;
TRIEU, SIMEON;
KANE, KEVIN y
WYNN, SIMON**

74 Agente/Representante:

VALLEJO LÓPEZ, Juan Pedro

ES 3 009 013 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

Sistema y método de captura y generación de imágenes tridimensionales panorámicas

5 **Antecedentes**

1. Campo de la una o más invenciones

10 Las realizaciones de la una o más presentes invenciones se refieren en general a la captura y unión de imágenes panorámicas de escenas en un entorno físico.

2. Técnica relacionada

15 La popularidad de proporcionar imágenes panorámicas tridimensionales (3D) del mundo físico ha creado muchas soluciones que tienen la capacidad de capturar múltiples imágenes bidimensionales (2D) y crear una imagen 3D basándose en las imágenes 2D capturadas. Existen soluciones de hardware y aplicaciones de software (o "apps") capaces de capturar múltiples imágenes 2D y unir las en una imagen panorámica.

20 Existen tecnologías para capturar y generar datos 3D de un edificio.

Sin embargo, las tecnologías existentes son generalmente incapaces de capturar y generar una representación 3D de un área con luz brillante. Una ventana con el sol brillando a fondo o un área de un suelo o pared con una luz brillante generalmente aparece como un agujero en la representación 3D, lo que puede requerir la realización de trabajo de postproducción adicional. Esto aumenta el tiempo de respuesta y la autenticidad de la Representación 3D. Además, el entorno exterior también proporciona un desafío para muchos dispositivos de captura 3D existentes porque la luz de la estructura puede no utilizarse para capturar imágenes 3D.

Otras limitaciones de las tecnologías existentes para capturar y generar datos 3D incluyen la cantidad de tiempo requerida para capturar y procesar las imágenes digitales requeridas para producir una imagen panorámica 3D.

30 El documento US 2010/134596 A1 divulga un escáner láser configurado para escanear un haz alrededor de un eje horizontal y vertical. Una cámara lineal se sujeta a un brazo pivotante soportado por la carcasa del escáner láser de modo que la cámara se pueda colocar para tener su lente dispuesta a lo largo del eje vertical para minimizar los errores de paralaje. De acuerdo con esta disposición, la cámara lineal está dispuesta en un tubo fuera de la carcasa del escáner láser.

40 El documento WO 2014/043461 divulga un escáner láser configurado para escanear un haz láser alrededor de un eje horizontal y vertical usando un cabezal de medición que rota alrededor del eje vertical, y que comprende un LiDAR y un espejo que rota alrededor del eje horizontal. Una cámara interna para capturar imágenes a color 2D puede tener un eje óptico que está alineado con el eje óptico del sistema de escaneado 3D de modo que se reducen o minimizan los errores de paralaje.

45 El documento US 2015/116691 A1 divulga un escáner láser basado en triangulación configurado para escanear horizontalmente un haz alrededor de un eje vertical y que tiene una cámara con una lente dispuesta a lo largo del eje vertical.

Sumario

50 Un dispositivo de captura de imágenes de acuerdo con la invención comprende una carcasa, un primer motor, una lente gran angular, un sensor de imagen, una montura, un LiDAR, un segundo motor y un espejo. La carcasa tiene un lado delantero y un lado trasero. El primer motor está acoplado a la carcasa en una primera posición entre el lado delantero y el lado trasero de la carcasa, estando el primer motor configurado para girar horizontalmente el dispositivo de captura de imágenes sustancialmente 270 grados alrededor de un eje vertical. La lente gran angular se coloca dentro de la carcasa en una segunda posición entre el lado delantero y el lado trasero de la carcasa a lo largo del eje vertical, siendo la segunda posición un punto sin paralaje y teniendo la lente gran angular un campo de visión lejos del lado delantero de la carcasa. El sensor de imagen está acoplado a la carcasa y configurado para generar señales de imagen a partir de la luz recibida por la lente gran angular. El soporte está acoplado al primer motor. El LiDAR está acoplado a la carcasa en una tercera posición, el LiDAR configurado para generar pulsos láser y generar señales de profundidad. El segundo motor está acoplado a la carcasa. El espejo está acoplado al segundo motor, el segundo motor está configurado para rotar el espejo alrededor de un eje horizontal, incluyendo el espejo una superficie en ángulo configurada para recibir los pulsos láser del LiDAR y dirigir los pulsos láser alrededor del eje horizontal.

65 En algunas realizaciones, el sensor de imagen está configurado para generar una primera pluralidad de imágenes a diferentes exposiciones cuando el dispositivo de captura de imágenes está estacionario y apuntado en una primera dirección. El primer motor puede configurarse para girar el dispositivo de captura de imágenes alrededor del eje vertical después de que se genere la primera pluralidad de imágenes. En diversas realizaciones, el sensor de imagen no

genera imágenes mientras el primer motor gira el dispositivo de captura de imágenes y en donde el LiDAR genera señales de profundidad basándose en los pulsos láser mientras el primer motor gira el dispositivo de captura de imágenes. El sensor de imagen puede configurarse para generar una segunda pluralidad de imágenes en las diferentes exposiciones cuando el dispositivo de captura de imágenes está estacionario y apunta en una segunda dirección y el primer motor está configurado para girar el dispositivo de captura de imágenes 90 grados alrededor del eje vertical después de la se genera una segunda pluralidad de imágenes. El sensor de imagen puede configurarse para generar una tercera pluralidad de imágenes en las diferentes exposiciones cuando el dispositivo de captura de imágenes está estacionario y apunta en una tercera dirección y el primer motor está configurado para girar el dispositivo de captura de imágenes 90 grados alrededor del eje vertical después de la se genera una tercera pluralidad de imágenes. El sensor de imagen puede configurarse para generar una cuarta pluralidad de imágenes en las diferentes exposiciones cuando el dispositivo de captura de imágenes está estacionario y apunta en una cuarta dirección y el primer motor está configurado para girar el dispositivo de captura de imágenes 90 grados alrededor del eje vertical después de la se genera una cuarta pluralidad de imágenes.

En algunas realizaciones, el sistema puede comprender adicionalmente un procesador configurado para combinar fotogramas de la primera pluralidad de imágenes antes de que el sensor de imagen genere la segunda pluralidad de imágenes. Un dispositivo digital remoto puede estar en comunicación con el dispositivo de captura de imágenes y configurado para generar una visualización 3D basándose en la primera, segunda, tercera y cuarta pluralidad de imágenes y las señales de profundidad, estando el dispositivo digital remoto configurado para generar la visualización 3D usando no más imágenes que la primera, segunda, tercera y cuarta pluralidad de imágenes. En algunas realizaciones, la primera, segunda, tercera y cuarta pluralidad de imágenes se generan entre giros que combinan giros que giran el dispositivo de captura de imágenes 270 grados alrededor del eje vertical. La velocidad o rotación del espejo alrededor del eje horizontal aumenta a medida que el primer motor gira el dispositivo de captura de imágenes. La superficie en ángulo del espejo puede ser de 90 grados. En algunas realizaciones, el LiDAR emite los pulsos láser en una dirección que es opuesta al lado delantero de la carcasa.

Un método de acuerdo con la invención comprende recibir la luz de una lente gran angular de un dispositivo de captura de imágenes, estando la lente gran angular dentro de una carcasa del dispositivo de captura de imágenes, recibiendo la luz en un campo de visión de la lente gran angular, extendiéndose el campo de visión lejos de un lado delantero de la carcasa, generar una primera pluralidad de imágenes mediante un sensor de imagen de un dispositivo de captura de imágenes utilizando la luz de la lente gran angular, estando el sensor de imagen acoplado a la carcasa, estando la primera pluralidad de imágenes a diferentes exposiciones, girar horizontalmente el dispositivo de captura de imágenes mediante un primer motor sustancialmente 270 grados alrededor de un eje vertical, estando el primer motor acoplado a la carcasa en una primera posición entre el lado delantero y el lado trasero de la carcasa, estando el objetivo gran angular en una segunda posición a lo largo del eje vertical, siendo la segunda posición un punto de no paralaje, rotar un espejo con una superficie en ángulo alrededor de un eje horizontal mediante un segundo motor, estando el segundo motor acoplado a la carcasa, generar pulsos láser mediante un LiDAR, estando el LiDAR acoplado a la carcasa en una tercera posición, dirigiéndose el pulso láser al espejo rotativo mientras el dispositivo de captura de imágenes gira horizontalmente, y generar señales de profundidad por el LiDAR basándose en los pulsos láser.

La generación de la primera pluralidad de imágenes por el sensor de imagen puede tener lugar antes de que el dispositivo de captura de imágenes gire horizontalmente. En algunas realizaciones, el sensor de imagen no genera imágenes mientras el primer motor gira el dispositivo de captura de imágenes y en donde el LiDAR genera las señales de profundidad basándose en los pulsos láser mientras el primer motor gira el dispositivo de captura de imágenes.

El método puede comprender además generar una segunda pluralidad de imágenes a las diferentes exposiciones por el sensor de imagen cuando el dispositivo de captura de imágenes está estacionario y apuntado en una segunda dirección y girar el dispositivo de captura de imágenes 90 grados alrededor del eje vertical mediante el primer motor después de generar una segunda pluralidad de imágenes.

En algunas realizaciones, el método puede comprender además generar una tercera pluralidad de imágenes a las diferentes exposiciones por el sensor de imagen cuando el dispositivo de captura de imágenes está estacionario y apuntado en una tercera dirección y girar el dispositivo de captura de imágenes 90 grados alrededor del eje vertical mediante el primer motor después de generar una tercera pluralidad de imágenes. El método puede comprender adicionalmente generar una cuarta pluralidad de imágenes a las diferentes exposiciones por el sensor de imagen cuando el dispositivo de captura de imágenes está estacionario y apuntado en una cuarta dirección. El método puede comprender generar una visualización 3D usando la primera, segunda, tercera y cuarta pluralidad de imágenes y basándose en las señales de profundidad, la generación de la visualización 3D sin usar ninguna otra imagen.

En algunas realizaciones, el método puede comprender adicionalmente combinar fotogramas de la primera pluralidad de imágenes antes de que el sensor de imagen genere la segunda pluralidad de imágenes. La primera, segunda, tercera y cuarta pluralidad de imágenes pueden generarse entre giros que combinan giros que giran el dispositivo de captura de imágenes 270 grados alrededor del eje vertical. En algunas realizaciones, una velocidad o rotación del espejo alrededor del eje horizontal aumenta a medida que el primer motor gira el dispositivo de captura de imágenes.

Breve descripción de los dibujos

La Figura 1a representa una vista de casa de muñecas de un entorno de ejemplo, tal como una casa, de acuerdo con algunas realizaciones.

La Figura 1b representa una vista en planta del primer piso de la casa de acuerdo con algunas realizaciones.

5 La Figura 2 representa una vista a nivel de los ojos de ejemplo de la sala de estar que puede ser parte de un recorrido virtual.

La Figura 3 representa un ejemplo de un sistema de captura del entorno de acuerdo con algunas realizaciones.

La Figura 4 representa una representación de un sistema de captura del entorno en algunas realizaciones.

10 La Figura 5 es una representación de los pulsos láser del LiDAR sobre el sistema de captura del entorno en algunas realizaciones.

La Figura 6a representa una vista lateral del sistema de captura del entorno.

La Figura 6b representa una vista desde arriba del sistema de captura del entorno en algunas realizaciones.

La Figura 7 representa una representación de los componentes de un ejemplo del sistema de captura del entorno de acuerdo con algunas realizaciones.

15 La Figura 8a representa dimensiones de lente de ejemplo en algunas realizaciones.

La Figura 8b representa una especificación del diseño de lente de ejemplo en algunas realizaciones.

La Figura 9a representa un diagrama de bloques de un sistema de captura del entorno de acuerdo con algunas realizaciones.

20 La Figura 9b representa un diagrama de bloques de un PCBA de SOM de ejemplo del sistema de captura del entorno de acuerdo con algunas realizaciones.

Las Figuras 10a-10c representan un proceso para el sistema de captura del entorno para tomar imágenes en algunas realizaciones.

La Figura 11 representa un diagrama de bloques de un entorno de ejemplo capaz de capturar y unir imágenes para formar visualizaciones 3D de acuerdo con algunas realizaciones.

25 La Figura 12 representa un diagrama de bloques de un ejemplo del sistema de alineación y unión de acuerdo con algunas realizaciones.

La Figura 13 representa un diagrama de flujo de un proceso de captura y generación de imágenes panorámicas 3D de acuerdo con algunas realizaciones.

30 La Figura 14 representa un diagrama de flujo de un proceso de captura y unión 3D y panorámico de acuerdo con algunas realizaciones.

La Figura 15 representa un diagrama de flujo que muestra detalles adicionales de una etapa del proceso de captura y unión 3D y panorámico de la Figura 14.

La Figura 16 representa un diagrama de bloques de un dispositivo digital de ejemplo de acuerdo con algunas realizaciones.

35

Descripción detallada

Muchas de las innovaciones descritas en el presente documento se realizan con referencia a los dibujos. Los números de referencia similares se usan para referirse a elementos similares. En la siguiente descripción, para fines de explicación, se exponen numerosos detalles específicos para proporcionar una comprensión completa. Podrá ser evidente, sin embargo, que las diferentes innovaciones pueden ponerse en práctica sin estos detalles específicos. En otros casos, se muestran estructuras y componentes bien conocidos en forma de diagrama de bloques para facilitar la descripción de las innovaciones.

40

45 Diversas realizaciones del aparato proporcionan a los usuarios imágenes panorámicas 3D de entornos interiores así como exteriores. En algunas realizaciones, el aparato puede proporcionar a los usuarios de manera eficiente y rápida imágenes panorámicas 3D de entornos interiores y exteriores usando una única lente de campo de visión amplio (FOV) y una única luz y sensores de detección y alcance (sensor LiDAR).

50

Lo siguiente es un caso de uso de ejemplo de un aparato de ejemplo descrito en el presente documento. El siguiente caso de uso es de una de las realizaciones. Diferentes realizaciones del aparato, como se analiza en el presente documento, pueden incluir una o más características y capacidades similares a las del caso de uso.

55

La Figura 1a representa una vista de casa de muñecas 100 de un entorno de ejemplo, tal como una casa, de acuerdo con algunas realizaciones. La vista de casa de muñecas 100 proporciona una vista general del entorno de ejemplo capturado por un sistema de captura del entorno (analizado en el presente documento). Un usuario puede interactuar con la vista de casa de muñecas 100 en un sistema de usuario alternando entre diferentes vistas del entorno de ejemplo. Por ejemplo, el usuario puede interactuar con el área 110 para activar una vista en planta del primer piso de la casa, como se ve en la Figura 1b. En algunas realizaciones, el usuario puede interactuar con iconos en la vista de casa de muñecas 100, tal como los iconos 120, 130 y 140, para proporcionar una vista de recorrido (por ejemplo, para un recorrido 3D), una vista de plano de planta o una vista de medición, respectivamente.

60

La Figura 1b representa una vista en planta del primer piso de la casa de acuerdo con algunas realizaciones. La vista en planta es una vista de arriba hacia abajo del primer piso de la casa. El usuario puede interactuar con áreas de la vista de plano de planta, tal como el área 150, para activar una vista a nivel de los ojos de una porción particular del plano de planta, tal como una sala de estar. Un ejemplo de la vista a nivel de los ojos de la sala de estar se puede

65

encontrar en la Figura 2 que puede ser parte de un recorrido virtual.

El usuario puede interactuar con una porción del plano de planta 200 correspondiente al área 150 de la Figura 1b. El usuario puede mover una vista alrededor de la habitación como si el usuario estuviera realmente en la sala de estar. Además de una vista horizontal de 360° de la sala de estar, el usuario también puede ver o navegar por el suelo o el techo de la sala de estar. Además, el usuario puede atravesar la sala de estar a otras partes de la casa interactuando con áreas particulares de la porción del plano de planta 200, tales como las áreas 210 y 220. Cuando el usuario interactúa con el área 220, el sistema de captura del entorno puede proporcionar una transición de estilo de caminar entre el área de la casa que corresponde sustancialmente a la región de la casa representada por el área 150 a un área de la casa que corresponde sustancialmente a la región de la casa representada por el área 220.

La Figura 3 representa un ejemplo de un sistema de captura 300 del entorno de acuerdo con algunas realizaciones. El sistema de captura 300 del entorno incluye la lente 310, una carcasa 320, un accesorio de montura 330 y una cubierta desplazable 340.

Cuando está en uso, el sistema de captura 300 del entorno puede colocarse en un entorno tal como una habitación. El sistema de captura 300 del entorno puede colocarse sobre un soporte (por ejemplo, un trípode). La cubierta desplazable 340 puede moverse para revelar un LiDAR y un espejo giratorio. Una vez activado, el sistema de captura 300 del entorno puede tomar una ráfaga de imágenes y luego girar usando un motor. El sistema de captura 300 del entorno puede encender el accesorio de montura 330. Mientras gira, el LiDAR puede tomar mediciones (mientras gira, el sistema de captura del entorno puede no tomar imágenes). Una vez dirigido a una nueva dirección, el sistema de captura del entorno puede tomar otra ráfaga de imágenes antes de girar a la siguiente dirección.

Por ejemplo, una vez colocado, un usuario puede ordenar al sistema de captura 300 del entorno que inicie un barrido. El barrido puede ser como sigue:

- (1) Estimación de exposición y luego tomar imágenes RGB HDR
Rotar 90 grados capturando datos de profundidad
- (2) Estimación de exposición y luego tomar imágenes RGB HDR
Rotar 90 grados capturando datos de profundidad
- (3) Estimación de exposición y luego tomar imágenes RGB HDR
Rotar 90 grados capturando datos de profundidad
- (4) Estimación de exposición y luego tomar imágenes RGB HDR
Rotar 90 grados (total 360) capturando datos de profundidad

Para cada ráfaga, puede haber cualquier número de imágenes a diferentes exposiciones. El sistema de captura del entorno puede combinar cualquier número de las imágenes de una ráfaga juntas mientras espera otro fotograma y/o espera la siguiente ráfaga.

La carcasa 320 puede proteger los componentes electrónicos del sistema de captura 300 del entorno y puede proporcionar una interfaz para la interacción del usuario, con un botón de encendido, un botón de escaneo y otros. Por ejemplo, la carcasa 320 puede incluir la cubierta desplazable 340, que puede ser desplazable para descubrir el LiDAR. Además, la carcasa 320 puede incluir interfaces electrónicas, tales como un adaptador de alimentación y luces indicadoras. En algunas realizaciones, la carcasa 320 es una carcasa de plástico moldeado. En diversas realizaciones, la carcasa 320 es una combinación de uno o más de plástico, metal y polímero.

La lente 310 puede ser parte de un conjunto de lentes. Pueden describirse detalles adicionales del conjunto de lente en la descripción de la Figura 7. La lente 310 se coloca estratégicamente en un centro de un eje de rotación 305 del sistema de captura ambiental 300. En este ejemplo, el eje de rotación 305 está en el plano x-y. Al colocar la lente 310 en el centro del eje de rotación 305, puede eliminarse o reducirse un efecto de paralaje. El paralaje es un error que surge debido a la rotación del dispositivo de captura de imágenes alrededor de un punto que no es un punto sin paralaje (NPP). En este ejemplo, el NPP se puede encontrar en el centro de la pupila de entrada del 1'ns.

Por ejemplo, suponiendo que se genera una imagen panorámica del entorno físico basándose en cuatro imágenes capturadas por el sistema de captura 300 del entorno con un 25 % de solapamiento entre imágenes de la imagen panorámica. Si no hay paralaje, entonces el 25 % de una imagen puede solaparse exactamente con otra imagen de la misma área del entorno físico. Eliminar o reducir el efecto de paralaje de las múltiples imágenes capturadas por un sensor de imagen a través de la lente 310 puede ayudar a unir múltiples imágenes en una imagen panorámica 2D.

La lente 310 puede incluir un gran campo de visión (por ejemplo, la lente 310 puede ser una lente de ojo de pez). En algunas realizaciones, la lente puede tener un FOV horizontal (HFOV) de al menos 148 grados y un FOV vertical (VFOV) de al menos 94 grados.

5 El accesorio de montura 330 puede permitir que el sistema de captura 300 del entorno se una a una montura. La montura puede permitir que el sistema de captura 300 del entorno se acople con un trípode, superficie plana o montura motorizada (por ejemplo, para mover el sistema de captura 300 del entorno). En algunas realizaciones, la montura puede permitir que el sistema de captura 300 del entorno rote a lo largo de un eje horizontal.

10 En algunas realizaciones, el sistema de captura 300 del entorno puede incluir un motor para girar el sistema de captura 300 del entorno horizontalmente alrededor del accesorio de montura 330.

En algunas realizaciones, una montura motorizada puede mover el sistema de captura 300 del entorno a lo largo de un eje horizontal, un eje vertical o ambos. En algunas realizaciones, el montura motorizada puede rotar o moverse en el plano x-y. El uso de un accesorio de montura 330 puede permitir que el sistema de captura 300 del entorno se acople a una montura motorizada, trípode o similar para estabilizar el sistema de captura 300 del entorno para reducir o minimizar la agitación. En otro ejemplo, el accesorio de montura 330 puede acoplarse a una montura motorizada que permite que el sistema de captura 300 del entorno y 3D rote a una velocidad constante y conocida, lo que ayuda al LiDAR a determinar las coordenadas (x, y, z) de cada pulso láser del LiDAR.

20 La Figura 4 representa una representación de un sistema de captura 400 del entorno en algunas realizaciones. La representación muestra el sistema de captura 400 del entorno (que puede ser un ejemplo del sistema de captura 300 del entorno de la Figura 3) desde una variedad de vistas, tal como una vista frontal 410, una vista superior 420, una vista lateral 430 y una vista posterior 440. En estas representaciones, el sistema de captura 400 del entorno puede incluir una porción hueca opcional representada en la vista lateral 430.

En algunas realizaciones, el sistema de captura 400 del entorno tiene una anchura de 75 mm, una altura de 180 mm y una profundidad de 189 mm. Se apreciará que el sistema de captura 400 del entorno puede tener cualquier anchura, altura o profundidad. En diversas realizaciones, la relación de anchura a altura a profundidad en el primer ejemplo se mantiene independientemente de las mediciones específicas.

30 La carcasa del sistema de captura 400 del entorno y 3D puede proteger los componentes electrónicos del sistema de captura 400 del entorno y puede proporcionar una interfaz (por ejemplo, pantalla en vista posterior 440) para la interacción del usuario. Además, la carcasa puede incluir interfaces electrónicas, tales como un adaptador de alimentación y luces indicadoras. En algunas realizaciones, la carcasa es una carcasa de plástico moldeado. En diversas realizaciones, la carcasa es una combinación de uno o más de plástico, metal y polímero. El sistema de captura 400 del entorno puede incluir una cubierta desplazable, que puede moverse para descubrir el LiDAR y proteger el LiDAR de los elementos cuando no está en uso.

40 La lente representada en la vista frontal 410 puede ser parte de un conjunto de lentes. Al igual que el sistema de captura 300 del entorno, la lente del sistema de captura 400 del entorno se coloca estratégicamente en un centro de un eje de rotación. La lente puede incluir un gran campo de visión. En diversas realizaciones, la lente representada en la vista frontal 410 está rebajada y la carcasa está ensanchada de tal manera que la lente gran angular está directamente en el punto sin paralaje (por ejemplo, directamente por encima de un punto medio de la montura y/o motor) pero aún puede tomar imágenes sin interferencia de la carcasa.

El accesorio de montura en la base del sistema de captura 400 del entorno puede permitir que el sistema de captura del entorno se fije a una montura. La montura puede permitir que el sistema de captura 400 del entorno se acople con un trípode, superficie plana o montura motorizada (por ejemplo, para mover el sistema de captura 400 del entorno). En algunas realizaciones, la montura puede acoplarse a un motor interno para girar el sistema de captura 400 del entorno alrededor de la montura.

55 En algunas realizaciones, la montura puede permitir que el sistema de captura 400 del entorno rote a lo largo de un eje horizontal. En diversas realizaciones, una montura motorizada puede mover el sistema de captura 400 del entorno a lo largo de un eje horizontal, un eje vertical o ambos. El uso de un accesorio de montura puede permitir que el sistema de captura 400 del entorno se acople a una montura motorizada, trípode o similar para estabilizar el sistema de captura 400 del entorno para reducir o minimizar la agitación. En otro ejemplo, el accesorio de montura puede acoplarse a una montura motorizada que permite que el sistema de captura 400 del entorno rote a una velocidad constante y conocida, lo que ayuda al LiDAR en la determinación de las coordenadas (x, y, z) de cada pulso láser del LiDAR.

60 En la vista 430, se revela un espejo 450. Un LiDAR puede emitir un pulso láser al espejo (en una dirección que es opuesta a la vista de la lente). El pulso láser puede golpear el espejo 450 que puede estar en ángulo (por ejemplo, en un ángulo de 90 grados). El espejo 450 puede estar acoplado a un motor interno que gira el espejo de tal manera que los pulsos láser del LiDAR pueden emitirse y/o recibirse en muchos ángulos diferentes alrededor del sistema de captura 400 del entorno.

La Figura 5 es una representación de los pulsos láser del LiDAR sobre el sistema de captura 400 del entorno en algunas realizaciones. En este ejemplo, los pulsos láser se emiten en el espejo giratorio 450. Los pulsos láser pueden emitirse y recibirse perpendiculares a un eje horizontal 602 (véase Figura 6) del sistema de captura 400 del entorno.
 5 El espejo 450 puede estar en ángulo de manera que los pulsos láser del LiDAR se dirijan lejos del sistema de captura 400 del entorno. En algunos ejemplos, el ángulo de la superficie en ángulo del espejo puede ser de 90 grados o estar entre 60 grados y 120 grados.

En algunas realizaciones, mientras el sistema de captura 400 del entorno está estacionario y en funcionamiento, el sistema de captura 400 del entorno puede tomar una ráfaga de imágenes a través de la lente. El sistema de captura 400 del entorno puede encender un motor horizontal entre ráfagas de imágenes. Mientras gira a lo largo de la montura, el LiDAR del sistema de captura 400 del entorno puede emitir y/o recibir pulsos láser que golpean el espejo giratorio 450. El LiDAR puede generar señales de profundidad a partir de las reflexiones de pulso láser recibidas y/o generar datos de profundidad.
 10

En algunas realizaciones, los datos de profundidad pueden asociarse con coordenadas alrededor del sistema de captura 400 del entorno. De manera similar, los píxeles o partes de imágenes pueden asociarse con las coordenadas sobre el sistema de captura 400 del entorno para permitir la creación de la visualización 3D (por ejemplo, una imagen desde diferentes direcciones, un recorrido 3D o similares) que se generará usando las imágenes y los datos de profundidad.
 15

Como se muestra en la Figura 5, los pulsos de LiDAR pueden bloquearse por la porción inferior del sistema de captura 400 del entorno. Se apreciará que el espejo 450 puede girar de manera consistente mientras el sistema de captura 400 del entorno se mueve alrededor de la montura o el espejo 450 puede girar más lentamente cuando el sistema de captura 400 del entorno comienza a moverse y nuevamente cuando el sistema de captura 400 del entorno se ralentiza para detenerse (por ejemplo, manteniendo una velocidad constante entre el arranque y la parada del motor de montura).
 20

El LiDAR puede recibir datos de profundidad de los pulsos. Debido al movimiento del sistema de captura 400 del entorno y/o al aumento o disminución de la velocidad del espejo 450, la densidad de datos de profundidad sobre el sistema de captura 400 del entorno puede ser inconsistente (por ejemplo, más densos en algunas áreas y menos densos en otras).
 25

La Figura 6a representa una vista lateral del sistema de captura 400 del entorno. En esta vista, se representa el espejo 450 y puede girar alrededor de un eje horizontal. El pulso 604 puede emitirse por el LiDAR en el espejo giratorio 450 y puede emitirse perpendicular al eje horizontal 602. De manera similar, el pulso 604 puede recibirse por el LiDAR de una manera similar.
 30

Aunque los pulsos de LiDAR se analizan como perpendiculares al eje horizontal 602, se apreciará que los pulsos de LiDAR pueden estar en cualquier ángulo con respecto al eje horizontal 602 (por ejemplo, el ángulo de espejo puede estar en cualquier ángulo, incluido entre 60 y 120 grados). En diversas realizaciones, el LiDAR emite pulsos opuestos a un lado delantero (por ejemplo, el lado delantero 604) del sistema de captura 400 del entorno (por ejemplo, en una dirección opuesta al centro del campo de visión de la lente o hacia el lado trasero 606).
 35

Como se analiza en el presente documento, el sistema de captura 400 del entorno puede girar alrededor del eje vertical 608. En diversas realizaciones, el sistema de captura 400 del entorno toma imágenes y luego gira 90 grados, tomando de este modo un cuarto conjunto de imágenes cuando el sistema de captura 400 del entorno completa el giro de 270 grados desde la posición inicial original donde se tomó el primer conjunto de imágenes. Como tal, el sistema de captura 400 del entorno puede generar cuatro conjuntos de imágenes entre giros que suman un total de 270 grados (por ejemplo, suponiendo que el primer conjunto de imágenes se tomó antes del giro inicial del sistema de captura 400 del entorno). En diversas realizaciones, las imágenes de un único barrido (por ejemplo, los cuatro conjuntos de imágenes) del sistema de captura 400 del entorno (por ejemplo, tomadas en una única rotación completa o una rotación de 270 grados alrededor del eje vertical) son suficientes junto con la datos de profundidad adquiridos durante el mismo barrido para generar la visualización 3D sin ningún barrido o giro adicional del sistema de captura 400 del entorno.
 40

Se apreciará que, en este ejemplo, los pulsos de LiDAR son emitidos y dirigidos por el espejo giratorio en una posición que está distante del punto de rotación del sistema de captura 400 del entorno. En este ejemplo, la distancia desde el punto de rotación de la montura es 608 (por ejemplo, la lente puede estar en el punto sin paralaje mientras que la lente puede estar en una posición detrás de la lente con respecto al frente del sistema de captura 400 del entorno). Dado que los pulsos de LiDAR son dirigidos por el espejo 450 en una posición que está fuera del punto de rotación, el LiDAR puede no recibir datos de profundidad desde un cilindro que se extiende desde por encima del sistema de captura 400 del entorno hasta por debajo del sistema de captura 400 del entorno. En este ejemplo, el radio del cilindro (por ejemplo, siendo el cilindro una falta de información de profundidad) puede medirse desde el centro del punto de rotación de la montura de motor hasta el punto donde el espejo 450 dirige los pulsos de LiDAR.
 45

Además, en la Figura 6b, se representa la cavidad 610. En este ejemplo, el sistema de captura 400 del entorno incluye
 50

5 el espejo giratorio dentro del cuerpo de la carcasa del sistema de captura 400 del entorno. Hay una sección recortada de la carcasa. Los pulsos láser pueden reflejarse por el espejo fuera de la carcasa y, a continuación, los reflejos pueden recibirse por el espejo y dirigirse de vuelta al LiDAR para permitir que el LiDAR cree señales de profundidad y/o datos de profundidad. La base del cuerpo del sistema de captura 400 del entorno por debajo de la cavidad 610 puede bloquear algunos de los pulsos láser. La cavidad 610 puede definirse por la base del sistema de captura 400 del entorno y el espejo giratorio. Como se representa en la Figura 6b, aún puede haber un espacio entre un borde del espejo en ángulo y la carcasa del sistema de captura 400 del entorno que contiene el LiDAR.

10 En diversas realizaciones, el LiDAR está configurado para dejar de emitir pulsos láser si la velocidad de rotación del espejo cae por debajo de un umbral de seguridad de rotación (por ejemplo, si hay un fallo del motor que hace girar el espejo o si el espejo se mantiene en su lugar). De esta manera, el LiDAR puede configurarse por seguridad y reducir la posibilidad de que un pulso láser continúe emitiéndose en la misma dirección (por ejemplo, en los ojos de un usuario).

15 La Figura 6b representa una vista desde arriba del sistema de captura 400 del entorno en algunas realizaciones. En este ejemplo, la parte delantera del sistema de captura 400 del entorno se representa con la lente rebajada y por encima directamente sobre el centro del punto de rotación (por ejemplo, encima del centro de la montura). La parte delantera de la cámara está rebajada para los extremos y la parte delantera de la carcasa está ensanchada para permitir que el campo de visión del sensor de imagen no se vea obstruido por la carcasa. El espejo 450 se representa apuntando hacia arriba.

20 La Figura 7 representa una representación de los componentes de un ejemplo del sistema de captura 300 del entorno de acuerdo con algunas realizaciones. El sistema de captura 700 del entorno incluye una cubierta delantera 702, un conjunto de lente 704, un marco estructural 706, un LiDAR 708, una carcasa delantera 710, un conjunto de espejo 712, una antena de GPS 714, una carcasa trasera 716, un motor vertical 718, una pantalla 720, un paquete de baterías 722, una montura 724 y un motor horizontal 726.

30 En diversas realizaciones, el sistema de captura 700 del entorno puede configurarse para escanear, alinear y crear mallas 3D en exteriores a pleno sol, así como en interiores. Esto elimina una barrera para la adopción de otros sistemas que son una herramienta solo para interiores. El sistema de captura 700 del entorno puede escanear grandes espacios más rápidamente que otros dispositivos. El sistema de captura 700 del entorno puede, en algunas realizaciones, proporcionar una precisión de profundidad mejorada mejorando la precisión de profundidad de exploración única a 90 m.

35 En algunas realizaciones, el sistema de captura ambiental 700 puede pesar 1 kg o aproximadamente 1 kg. En un ejemplo, el sistema de captura ambiental 700 puede pesar entre 1-3 kg.

40 La cubierta delantera 702, la carcasa delantera 710 y la carcasa trasera 716 forman parte de la carcasa. En un ejemplo, la cubierta delantera puede tener una anchura, w, de 75 mm.

45 El conjunto de lente 704 puede incluir una lente de cámara que enfoca la luz sobre un dispositivo de captura de imágenes. El dispositivo de captura de imágenes puede capturar una imagen de un entorno físico. El usuario puede colocar el sistema de captura 700 del entorno para capturar una porción de un piso de un edificio, tal como el segundo edificio 422 de la Figura 1 para obtener una imagen panorámica de la una porción del suelo. El sistema de captura 700 del entorno puede moverse a otra porción del suelo del edificio para obtener una imagen panorámica de otra porción del suelo. En un ejemplo, la profundidad de campo del dispositivo de captura de imágenes es de 0,5 metros al infinito. La Figura 8a representa dimensiones de lente de ejemplo en algunas realizaciones.

50 En algunas realizaciones, el dispositivo de captura de imágenes es un sensor de imagen complementario de metal-óxido-semiconductor (CMOS) (por ejemplo, un sensor MIPI CMOS Sony IMX283 de ~20 megapíxeles con Nano SOM NVidia Jetson). En diversas realizaciones, el dispositivo de captura de imágenes es un dispositivo acoplado cargado (CCD). En un ejemplo, el dispositivo de captura de imágenes es un sensor rojo-verde-azul (RGB). En una realización, el dispositivo de captura de imágenes es un sensor de infrarrojos (IR). El conjunto de lente 704 puede proporcionar al dispositivo de captura de imágenes un amplio campo de visión.

55 El sensor de imagen puede tener muchas especificaciones diferentes. En un ejemplo, el sensor de imagen incluye lo siguiente:

Píxeles por columna	píxeles	5496
Píxeles por fila	píxeles	3694
Resolución	MP	> 20
Diámetro del círculo de imagen	mm	15,86 mm
Distancia entre píxeles	um	2,4 um
Píxeles por grado (PPD)	PPD	> 37

ES 3 009 013 T3

Ángulo de rayo principal a altura total	grado s	
---	---------	--

(continuación)

Interfaz de salida	-	MIPI
Sensibilidad a color verde	V/lux*s	> 1,7
SNR (100 lux, ganancia lx)	dB	> 65
Rango dinámico	dB	> 70

Las especificaciones de ejemplo pueden ser las siguientes:

número F	-	2,8
Diámetro del círculo de imagen	mm	15,86
Distancia mínima del objeto	mm	500
Distancia máxima del objeto	mm	Infinito
Ángulo de rayo principal a altura total del sensor	grad	3,0
diámetro L1	mm	< 60
Longitud total de seguimiento (TTL)	mm	<= 80
Distancia focal posterior (BFL)	mm	-
Longitud focal efectiva (EFL)	mm	-
Iluminación relativa	%	> 50
Distorsión máxima	%	< 5
52 lp/mm (en el eje)	%	> 85
104 lp/mm (en el eje)	%	> 66
208 lp/mm (en el eje)	%	> 45
52 lp/mm (83 % de campo)	%	> 75
104 lp/mm (83 % de campo)	%	> 41
208 lp/mm (83 % de campo)	%	> 25

5 En diversas realizaciones, al observar la MTF en el campo relativo F0 (es decir, el centro), el desplazamiento de enfoque puede variar de +28 micrómetros a 0,5 m a -25 micrómetros en el infinito para un desplazamiento de enfoque pasante total de 53 micrómetros.

10 La Figura 8b representa especificaciones del diseño de lente de ejemplo en algunas realizaciones.

En algunos ejemplos, el conjunto de lente 704 tiene un HFOV de al menos 148 grados y un VFOV de al menos 94 grados. En un ejemplo, el conjunto de lente 704 tiene un campo de visión de 150°, 180°, o está dentro de un intervalo de 145° a 180°. La captura de imagen de una vista de 360° alrededor del sistema de captura 700 del entorno puede obtenerse, en un ejemplo, con tres o cuatro capturas de imagen separadas del dispositivo de captura de imágenes del sistema de captura 700 del entorno. En diversas realizaciones, el dispositivo de captura de imágenes puede tener una resolución de al menos 37 píxeles por grado. En algunas realizaciones, el sistema de captura 700 del entorno incluye una tapa de lente (no mostrada) para proteger el conjunto de lente 704 cuando no está en uso. La salida del conjunto de lente 704 puede ser una imagen digital de un área del entorno físico. Las imágenes capturadas por el conjunto de lente 704 pueden unirse para formar una imagen panorámica 2D del entorno físico. Puede generarse una panorámica 3D combinando los datos de profundidad capturados por el LiDAR 708 con la imagen panorámica 2D generada uniendo juntas múltiples imágenes del conjunto de lente 704. En algunas realizaciones, las imágenes capturadas por el sistema de captura 402 del entorno se unen juntas por el sistema de procesamiento de imágenes 406. En diversas realizaciones, el sistema de captura 402 del entorno genera una versión de "vista previa" o "miniatura" de una imagen panorámica 2D. La versión de vista previa o miniatura de la imagen panorámica 2D puede presentarse en un sistema de usuario 1110 tal como un iPad, ordenador personal, teléfono inteligente o similar. En algunas realizaciones, el sistema de captura 402 del entorno puede generar un minimapa de un entorno físico que representa un área del entorno físico. En diversas realizaciones, el sistema de procesamiento de imágenes 406 genera el minimapa que representa el área del entorno físico.

30 Las imágenes capturadas por el conjunto de lente 704 pueden incluir datos de ubicación de dispositivo de captura que identifican o indican una ubicación de captura de una imagen 2D. Por ejemplo, en algunas implementaciones, los datos de ubicación de dispositivo de captura pueden incluir unas coordenadas de sistema de posicionamiento global (GPS) asociadas con una imagen 2D. En otras implementaciones, los datos de ubicación del dispositivo de captura pueden incluir información de posición que indica una posición relativa del dispositivo de captura (por ejemplo, la cámara y/o un sensor 3D) a su entorno, tal como una posición relativa o calibrada del dispositivo de captura a un objeto en el entorno, otra cámara en el entorno, otro dispositivo en el entorno o similares. En algunas implementaciones, este tipo de datos de ubicación pueden determinarse por el dispositivo de captura (por ejemplo, la cámara y/o un dispositivo acoplado operativamente a la cámara que comprende hardware y/o software de posicionamiento) en asociación con la captura de una imagen y recibidos con la imagen. La colocación del conjunto de lente 704 no es únicamente por diseño. Al colocar el conjunto de lente 704 en el centro, o sustancialmente en el centro, del eje de rotación, puede reducirse el efecto de paralaje.

5 En algunas realizaciones, el marco estructural 706 mantiene el conjunto de lente 704 y el LiDAR 708 en una posición particular y puede ayudar a proteger los componentes del ejemplo del sistema de captura del entorno. El marco estructural 706 puede servir para ayudar a montar rígidamente el LiDAR 708 y colocar el LiDAR 708 en una posición fija. Además, la posición fija del conjunto de lente 704 y el LiDAR 708 posibilitan una relación fija para alinear los datos de profundidad con la información de imagen para ayudar a crear las imágenes 3D. Los datos de imagen 2D y los datos de profundidad capturados en el entorno físico pueden alinearse con relación a un espacio de coordenadas 3D común para generar un modelo 3D del entorno físico.

10 En diversas realizaciones, el LiDAR 708 captura información de profundidad de un entorno físico. Cuando el usuario coloca el sistema de captura 700 del entorno en una porción de un piso del segundo edificio, el LiDAR 708 puede obtener información de profundidad de objetos. El LiDAR 708 puede incluir un módulo de detección óptica que puede medir la distancia a un objetivo u objetos en una escena utilizando pulsos de un láser para irradiar un objetivo o escena y medir el tiempo que tardan los fotones en viajar al objetivo y volver al LiDAR 708. La medición puede transformarse a continuación en un sistema de coordenadas de cuadrícula usando información derivada de un tren de accionamiento horizontal del sistema de captura 700 del entorno.

20 En algunas realizaciones, el LiDAR 708 puede devolver puntos de datos de profundidad cada 10 usegundos con una indicación de tiempo (de un reloj interno). El LiDAR 708 puede muestrear una esfera parcial (pequeños orificios en la parte superior e inferior) cada 0,25 grados. En algunas realizaciones, con un punto de datos cada 10 us y 0,25 grados, puede haber 14,40 milisegundos por "disco" de puntos y 1440 discos para hacer una esfera que es nominalmente de 20,7 segundos. Debido a que cada disco captura hacia delante y hacia atrás, la esfera podría capturarse en un barrido de 180°.

25 En un ejemplo, la especificación LiDAR 708 puede ser como sigue:

Alcance (10 % de reflectancia)	m	90
Alcance (20 % de reflectancia)	m	130
Alcance (100 % de reflectancia)	m	260
Precisión de alcance (1σ a 20 m)	cm	2
Longitud de onda	nm	905
Seguridad de láser	--	Clase 1
Velocidad de puntos	puntos/s	100.000
Divergencia de haz	grado s	0,28 x 0,03
Resolución angular	grad	0,1
Haz colimado		
Dimensiones (a 10 cm)	mm	14,71 x 8,46
Temperatura de funcionamiento	grad C	-20 a 65
Potencia (modo normal, activo)	W	4,83
Potencia (modo normal, inactivo)	W	4,38
Potencia (modo de espera)	W	4,07
Tiempo para activación desde apagado	s	3,898
Tiempo para activación desde modo de espera	s	0,289
Tiempo para activación desde modo normal inactivo	s	0,003
Tensión	v	10-15,6
Sincronización de datos	—	Pulso por segundo (PPS)
Dimensiones	mm	60 x 58 x 56
Peso	g	230
Latencia de datos	ms	2
Índice de falsas alarmas (a 100 klx)	%	<0,01 %

30 Una ventaja de utilizar LiDAR es que con un LiDAR a la longitud de onda más baja (por ejemplo, 905 nm, 900-940 nm o similar) puede permitir que el sistema de captura 700 del entorno determine información de profundidad para un entorno exterior o un entorno interior con luz brillante.

35 La colocación del conjunto de lente 704 y el LiDAR 708 puede permitir que el sistema de captura 700 del entorno o un dispositivo digital en comunicación con el sistema de captura 700 del entorno genere una imagen panorámica 3D usando los datos de profundidad del LiDAR 708 y el conjunto de lente 704. En algunas realizaciones, las imágenes panorámicas 2D y 3D no se generan en el sistema de captura 402 del entorno.

La salida del LiDAR 708 puede incluir atributos asociados con cada pulso láser enviado por el LiDAR 708. Los atributos

incluyen la intensidad del pulso láser, el número de retornos, el número de retorno actual, el punto de clasificación, los valores de RGC, el tiempo de GPS, el ángulo de exploración, la dirección de exploración o cualquier combinación en los mismos. La profundidad de campo puede ser (0,5 m; infinito), (1 m; infinito), o similares. En algunas realizaciones, la profundidad de campo es de 0,2 m a 1 m e infinito.

5 En algunas realizaciones, el sistema de captura 700 del entorno captura cuatro imágenes de RBG separadas usando el conjunto de lente 704 mientras el sistema de captura 700 del entorno está estacionario. En diversas realizaciones, el LiDAR 708 captura datos de profundidad en cuatro instancias diferentes mientras el sistema de captura 700 del entorno está en movimiento, moviéndose desde una posición de captura de imagen de RBG a otra posición de captura de imagen de RBG. En un ejemplo, la imagen panorámica 3D se captura con una rotación de 360° del sistema de captura 700 del entorno, que puede denominarse barrido. En diversas realizaciones, la imagen panorámica 3D se captura con una rotación de menos de 360° del sistema de captura 700 del entorno. La salida del barrido puede ser una lista de barrido (SWL), que incluye datos de imagen del conjunto de lente 704 y datos de profundidad del LiDAR 708 y propiedades del barrido, incluyendo la ubicación de GPS y una indicación de tiempo de cuándo tuvo lugar el barrido. En diversas realizaciones, un único barrido (por ejemplo, un único giro de 360 grados del sistema de captura 700 del entorno) captura suficiente imagen e información de profundidad para generar una visualización 3D (por ejemplo, por el dispositivo digital en comunicación con el sistema de captura 700 del entorno que recibe las imágenes y los datos de profundidad del sistema de captura 700 del entorno y crea la visualización 3D usando solo las imágenes y los datos de profundidad del sistema de captura 700 del entorno capturados en el único barrido).

20 En algunas realizaciones, las imágenes capturadas por el sistema de captura 402 del entorno pueden combinarse, unirse entre sí y combinarse con los datos de profundidad del LiDAR 708 mediante un sistema de unión y procesamiento de imágenes analizado en el presente documento.

25 En diversas realizaciones, el sistema de captura 402 del entorno y/o una aplicación en el sistema de usuario 1110 pueden generar una vista previa o versión en miniatura de una imagen panorámica 3D. La versión de vista previa o miniatura de la imagen panorámica 3D puede presentarse en el sistema de usuario 1110 y puede tener una resolución de imagen más baja que la imagen panorámica 3D generada por el sistema de procesamiento de imágenes 406. Después de que el conjunto de lente 704 y el LiDAR 708 capturen las imágenes y los datos de profundidad del entorno físico, el sistema de captura 402 del entorno puede generar un minimapa que representa un área del entorno físico que se ha capturado por el sistema de captura 402 del entorno. En algunas realizaciones, el sistema de procesamiento de imágenes 406 genera el minimapa que representa el área del entorno físico. Después de capturar imágenes y datos de profundidad de una sala de estar de una casa usando el sistema de captura 402 del entorno, el sistema de captura 402 del entorno puede generar una vista de arriba hacia abajo del entorno físico. Un usuario puede usar esta información para determinar áreas del entorno físico en las que el usuario no ha capturado o generado imágenes panorámicas 3D.

40 En una realización, el sistema de captura 700 del entorno puede intercalar captura de imagen con el dispositivo de captura de imágenes del conjunto de lente 704 con captura de información de profundidad con el LiDAR 708. Por ejemplo, el dispositivo de captura de imágenes puede capturar una imagen de la sección 1605, como se ve en la Figura 16, del entorno físico con el dispositivo de captura de imágenes, y luego el LiDAR 708 obtiene información de profundidad de la sección 1605. Una vez que el LiDAR 708 obtiene información de profundidad de la sección 1605, el dispositivo de captura de imágenes puede moverse para capturar una imagen de otra sección 1610, y a continuación el LiDAR 708 obtiene información de profundidad de la sección 1610, intercalando de ese modo la captura de imagen y la captura de información de profundidad.

50 En algunas realizaciones, el LiDAR 708 puede tener un campo de visión de al menos 145°, la información de profundidad de todos los objetos en una vista de 360° del sistema de captura 700 del entorno puede obtenerse por el sistema de captura 700 del entorno en tres o cuatro exploraciones. En otro ejemplo, el LiDAR 708 puede tener un campo de visión de al menos 150°, 180°, o entre 145° a 180°.

55 Un aumento en el campo de visión de la lente reduce la cantidad de tiempo requerido para obtener información visual y de profundidad del entorno físico alrededor del sistema de captura 700 del entorno. En diversas realizaciones, el LiDAR 708 tiene un alcance de profundidad mínimo de 0,5 m. En una realización, el LiDAR 708 tiene un alcance de profundidad máximo de más de 8 metros.

60 El LiDAR 708 puede utilizar el conjunto de espejo 712 para dirigir el láser en diferentes ángulos de exploración. En una realización, el motor vertical opcional 718 tiene la capacidad de mover el conjunto de espejo 712 verticalmente. En algunas realizaciones, el conjunto de espejo 712 puede ser un espejo dieléctrico con un revestimiento o capa hidrófoba. El conjunto de espejo 712 puede acoplarse al motor vertical 718 que gira el conjunto de espejo 712 cuando está en uso.

El espejo del conjunto de espejo 712 puede incluir, por ejemplo, las siguientes especificaciones:

65

ES 3 009 013 T3

Reflectividad a 905 nm	%	>99
Absorción a longitudes de onda visibles (380-700 nm)**	%	>60 %
Apertura libre	%	>=85
Umbral de daño por láser a 905 nm	uJ	>=0,45
Ángulo de incidencia (AOI)	grad	45 ± 1

El espejo del conjunto de espejo 712 puede incluir, por ejemplo, la siguiente especificación para materiales y revestimientos:

S1L1	material	Dieléctrico	
S1L2	material	Hidrófobo	
S2L1	material	Emulsión de pintura negra	Polvo suspendido en pintura
Sustrato	material	Schott B270I	

5 El revestimiento hidrófobo del espejo del conjunto de espejo 712 puede incluir, por ejemplo, un ángulo de contacto >105 grad.

10 El espejo del conjunto de espejo 712 puede incluir las siguientes especificaciones de calidad:

Estándar de arañazos/socavado	3
HTS: 80 C, 50 h.	3
LTS: -30 C, 1000 h	3
THS: 60 C/90 % HR, 1000 h	3
TC: -30 a 70 C, 50 cic (30 min/5 min/30 min)	3
(Resistencia al disolvente), lado a, 50 toallitas con etanol, alcohol, 300 g	
(Resistencia al disolvente), lado b, 10 toallitas con etanol, alcohol, 200 g	3
(Resistencia a la abrasión), lado a, 50 toallitas, 300 g	
(Resistencia a la abrasión), lado b, 10 toallitas, 200 g	3
(Durabilidad), Lado a, 10 desprendimientos de cinta (CT-18)	
(Durabilidad), Lado b, 5 desprendimientos de cinta (CT-18)	3
Resistencia a los rayos UV (entorno exterior)	3
simulación, 340 nm, irradiancia de 0,35 KV/m ² /nm, 306 min de luz a 125 C BTP 54 min, luz y pulverización de agua desionizada (temperatura no controlada) 6 h de oscuridad a 95 % de HR, 24 C (aire))	
Rugosidad superficial	>=10
Ángulo de contacto hidrófobo	>=10

El motor vertical puede incluir, por ejemplo, las siguientes especificaciones:

15 **Precisión 1 tiempo de revolución**

Velocidad máxima	RPM	4000 y 6500
Aceleración máxima	grad/s ²	300
Durabilidad	Ciclo	70000
Accionador de motor		
Precisión 1 tiempo de revolución desviación estándar de la varianza de <5 μ s		

Debido al dispositivo de captura de RGB y al LiDAR 708, el sistema de captura 700 del entorno puede capturar imágenes en el exterior con luz solar brillante o en el interior con luces brillantes o deslumbramiento de luz solar desde las ventanas. En sistemas que utilizan diferentes dispositivos (por ejemplo, dispositivos de luz estructurada), estos pueden no ser capaces de operar en entornos brillantes, ya sea dentro o fuera. Esos dispositivos a menudo se limitan a su uso solo en interiores y solo durante el amanecer o el atardecer para controlar la luz. De lo contrario, los puntos brillantes en una habitación crean alteraciones de imagen o "agujeros" en las imágenes que deben rellenarse o corregirse. El sistema de captura ambiental 700, sin embargo, puede utilizarse con luz solar brillante tanto en el interior como en el exterior. El dispositivo de captura y el LiDAR 708 pueden capturar datos de imagen y profundidad en entornos brillantes sin alteraciones de imagen u orificios causados por deslumbramiento o luz brillante.

En una realización, la antena de GPS 714 recibe datos del sistema de posicionamiento global (GPS). Los datos de GPS pueden usarse para determinar la ubicación del sistema de captura 700 del entorno en cualquier momento dado.

En diversas realizaciones, la pantalla 720 permite que el sistema de captura 700 del entorno proporcione un estado actual del sistema, tal como actualización, calentamiento, exploración, exploración completa, error y similares.

El paquete de baterías 722 proporciona energía al sistema de captura 700 del entorno. El paquete de baterías 722 puede ser extraíble y recargable, permitiendo de este modo que un usuario coloque un paquete de baterías nuevo 722 mientras carga un paquete de baterías agotado. En algunas realizaciones, el paquete de baterías 722 puede permitir al menos 1000 SWL o al menos 250 SWL de uso continuo antes de la recarga. El sistema de captura 700 del entorno puede utilizar un enchufe USB-C para recargar.

En algunas realizaciones, la montura 724 proporciona un conector para que el sistema de captura 700 del entorno se conecte a una plataforma tal como un trípode o montura. El motor horizontal 726 puede rotar el sistema de captura 700 del entorno alrededor de un plano x-y. En algunas realizaciones, el motor horizontal 726 puede proporcionar información a un sistema de coordenadas de cuadrícula para determinar las coordenadas (x, y, z) asociadas con cada pulso láser. En diversas realizaciones, debido al amplio campo de visión de la lente, la colocación de la lente alrededor del eje de rotación y el dispositivo LiDAR, el motor horizontal 726 puede posibilitar que el sistema de captura 700 del entorno explore rápidamente.

El motor horizontal 726 puede tener las siguientes especificaciones en un ejemplo:

Velocidad máxima	grad/s	60
Aceleración máxima	grad/s ²	300
Par de torsión máximo	Nm	0,5
Resolución de posición angular	grad	<0,125 a < 0,025
Precisión de posición angular	grad	<0,1
Resolución del codificador	CPR	4096
Durabilidad	Ciclo	70.000

En diversas realizaciones, la montura 724 puede incluir un adaptador de liberación rápida. El par de retención puede ser, por ejemplo, >2,0 Nm y la durabilidad de la operación de captura puede ser de hasta o más de 70.000 ciclos.

Por ejemplo, el sistema de captura 700 del entorno puede posibilitar la construcción de una malla 3D de una casa estándar con una distancia entre barridos mayor que 8 m. Un tiempo para capturar, procesar y alinear un barrido interior puede ser inferior a 45 segundos. En un ejemplo, un marco de tiempo desde el inicio de una captura de barrido hasta cuando el usuario puede mover el sistema de captura 700 del entorno puede ser inferior a 15 segundos.

En diversas realizaciones, estos componentes proporcionan al sistema de captura 700 del entorno la capacidad de alinear posiciones de escaneo tanto en exteriores como en interiores y, por lo tanto, crear experiencias de recorrido sin interrupciones entre interiores y exteriores (esto puede ser una alta prioridad para hoteles, alquileres vacacionales, bienes raíces, documentación de construcción, CRE y modelado y verificación según construcción. El sistema de captura 700 del entorno también puede crear una "casa de muñecas al aire libre" o minimapa al aire libre. El sistema de captura 700 del entorno, como se muestra en el presente documento, también puede mejorar la precisión de la reconstrucción 3D, principalmente desde una perspectiva de medición. Para la densidad de exploración, la capacidad para que el usuario la sintonice también puede ser una ventaja. Estos componentes también pueden posibilitar que el sistema de captura 700 del entorno tenga la capacidad de capturar amplios espacios vacíos (por ejemplo, mayor alcance). Para generar un modelo 3D de amplios espacios vacíos puede requerirse que el sistema de captura del entorno escanee y capture datos 3D y datos de profundidad desde un intervalo de distancia mayor que generar un modelo 3D de espacios más pequeños.

En diversas realizaciones, estos componentes posibilitan que el sistema de captura 700 del entorno alinee las SWL y reconstruya el modelo 3D de una manera similar para uso en interiores así como en exteriores. Estos componentes

también pueden habilitar que el sistema de captura 700 del entorno realice la geolocalización de modelos 3D (que puede facilitar la integración a Google Street View y ayudar a alinear panoramas exteriores si es necesario).

5 El dispositivo de captura de imágenes del sistema de captura 700 del entorno puede proporcionar una imagen similar a DSLR con calidad imprimible a 8,5" x 11" para VFOV de 70° y un estilo de imagen RGB.

10 En algunas realizaciones, el sistema de captura 700 del entorno puede tomar una imagen RGB con el dispositivo de captura de imágenes (por ejemplo, usando la lente gran angular) y luego mover la lente antes de tomar la siguiente imagen RGB (para un total de cuatro movimientos usando el motor). Mientras el motor horizontal 726 rota el sistema de captura del entorno 90 grados, el LiDAR 708 puede capturar datos de profundidad. En algunas realizaciones, el LiDAR 708 incluye una matriz de APD.

15 En algunas realizaciones, los datos de imagen y profundidad pueden enviarse a continuación a una aplicación de captura (por ejemplo, un dispositivo en comunicación con el sistema de captura 700 del entorno, tal como un dispositivo inteligente o un sistema de captura de imagen en una red). En algunas realizaciones, el sistema de captura 700 del entorno puede enviar la imagen y los datos de profundidad al sistema de procesamiento de imágenes 406 para procesar y generar la imagen panorámica 2D o la imagen panorámica 3D. En diversas realizaciones, el sistema de captura 700 del entorno puede generar una lista de barrido de la imagen RGB capturada y los datos de profundidad a partir de una revolución de 360 grados del sistema de captura 700 del entorno. La lista de barrido puede enviarse al sistema de procesamiento de imágenes 406 para su unión y alineación. La salida del barrido puede ser una SWL que incluye datos de imagen del conjunto de lente 704 y datos de profundidad del LiDAR 708 y propiedades del barrido, incluyendo la ubicación de GPS y una indicación de tiempo de cuándo tuvo lugar el barrido.

20 En diversas realizaciones, el LiDAR, el espejo vertical, la lente RGB, la montura de trípode y el accionamiento horizontal están montados rigidamente dentro de la carcasa para permitir que la carcasa se abra sin requerir que el sistema se recalibre.

30 La Figura 9a representa un diagrama de bloques 900 de un sistema de captura del entorno de acuerdo con algunas realizaciones. El diagrama de bloques 900 incluye una fuente de alimentación 902, un convertidor de potencia 904, un conjunto de placa de circuito impreso (PCBA) de entrada/salida (E/S), un PCBA de sistema en módulo (SOM), una interfaz de usuario 910, un LiDAR 912, un motor de corriente continua sin escobillas (BLCC) de espejo 914, un tren de accionamiento 916, una lente de FOV ancho (WFOV) 918 y un sensor de imagen 920.

35 La fuente de alimentación 902 puede ser el paquete de baterías 722 de la Figura 7. La fuente de alimentación puede ser una batería recargable extraíble, tal como una batería de ion litio (por ejemplo, 4x celda de ion litio 18650) capaz de proporcionar alimentación al sistema de captura ambiental.

40 El convertidor de potencia 904 puede cambiar el nivel de tensión de la fuente de potencia 902 a un nivel de tensión más bajo o más alto para que pueda ser utilizado por los componentes electrónicos del sistema de captura ambiental. El sistema de captura del entorno puede utilizar 4x celdas de ion litio 18650 en configuración 4S1P, o cuatro conexiones en serie y una configuración de conexión en paralelo.

45 En algunas realizaciones, el PCBA de E/S 906 puede incluir elementos que proporcionan IMU, Wi-Fi, GPS, Bluetooth, unidad de medición inercial (IMU), accionadores de motor y microcontroladores. En algunas realizaciones, el PCBA de E/S 906 incluye un microcontrolador para controlar el motor horizontal y codificar controles de motor horizontal así como controlar el motor vertical y codificar controles de motor vertical.

50 El PCBA de SOM 908 puede incluir una unidad de procesamiento central (CPU) y/o unidad de procesamiento de gráficos (GPU), memoria e interfaz desplazable. El PCBA de SOM 908 puede controlar el LiDAR 912, el sensor de imagen 920 y el PCBA de E/S 906. El PCBA de SOM 908 puede determinar las coordenadas (x, y, z) asociadas con cada pulso láser del LiDAR 912 y almacenar las coordenadas en un componente de memoria del PCBA de SOM 908. En algunas realizaciones, el PCBA de SOM 908 puede almacenar las coordenadas en el sistema de procesamiento de imágenes del sistema de captura 400 del entorno. Además de las coordenadas asociadas con cada pulso láser, el PCBA de SOM 908 puede determinar atributos adicionales asociados con cada pulso láser, incluyendo la intensidad del pulso láser, el número de retornos, el número de retorno actual, el punto de clasificación, los valores de RGC, el tiempo de GPS, ángulo de escaneo y la dirección de escaneo.

60 En algunas realizaciones, el PCBA de SOM 908 incluye una PCBA de SOM de Nvidia con CPU/GPU, DDR, eMMC, Ethernet.

65 La interfaz de usuario 910 puede incluir botones físicos o conmutadores con los que el usuario puede interactuar. Los botones o conmutadores pueden proporcionar funciones tales como encender y apagar el sistema de captura del entorno, escanear un entorno físico y otras. En algunas realizaciones, la interfaz de usuario 910 puede incluir una pantalla tal como la pantalla 720 de la Figura 7.

En algunas realizaciones, el LiDAR 912 captura información de profundidad del entorno físico. El LiDAR 912 incluye

ES 3 009 013 T3

- un módulo de detección óptica que puede medir la distancia a un objetivo u objetos en una escena irradiando el objetivo o la escena con luz, usando pulsos de un láser. El módulo de detección óptica del LiDAR 912 mide el tiempo que tardan los fotones en viajar a dicho objetivo u objeto y regresar después de la reflexión a un receptor en el LiDAR 912, dando de este modo una distancia del LiDAR desde el objetivo u objeto. Junto con la distancia, el PCBA de SOM 908 puede determinar las coordenadas (x, y, z) asociadas con cada pulso láser. El LiDAR 912 puede encajar dentro de un ancho de 58 mm, una altura de 55 mm y una profundidad de 60 mm.
- El LiDAR 912 puede incluir un alcance (10 % de reflectancia) de 90 m, alcance (20 % de reflectancia) 130 m, alcance (100 % de reflectancia) 260 m, una precisión de alcance (1σ a 900 m) de 2 cm, una longitud de onda 1705 nm y divergencia de haz de $0,28 \times 0,03$ grados.
- El PCBA de SOM 908 puede determinar las coordenadas basándose en la ubicación del tren de accionamiento 916. En diversas realizaciones, el LiDAR 912 puede incluir uno o más dispositivos LiDAR. Pueden utilizarse múltiples dispositivos LiDAR para aumentar la resolución LiDAR.
- El motor de corriente continua sin escobillas de espejo (BLCC) 914 puede controlar el conjunto de espejo 712 de la Figura 7.
- En algunas realizaciones, el tren de accionamiento 916 puede incluir el motor horizontal 726 de la Figura 7. El tren de accionamiento 916 puede proporcionar la rotación del sistema de captura del entorno cuando está montado en una plataforma tal como un trípode. El tren de accionamiento 916 puede incluir un motor paso a paso Nema 14, tren de accionamiento de rueda helicoidal y de plástico, embrague, cojinete de buje y un mecanismo de prevención de holgura. En algunas realizaciones, el sistema de captura del entorno puede ser capaz de completar una exploración en menos de 17 segundos. En diversas realizaciones, el tren de accionamiento 916 tiene una velocidad máxima de 60 grados/segundo, una aceleración máxima de 300 grados/segundo², un par de torsión máximo de 0,5 nm, una precisión de posición angular de menos de 0,1 grados y una resolución del codificador de aproximadamente 4096 recuentos por revolución.
- En algunas realizaciones, el tren de accionamiento 916 incluye un espejo monógono vertical y un motor. En este ejemplo, el tren de accionamiento 916 puede incluir un motor BLCC, un sensor de efecto Hall externo, un imán (emparejado con el sensor de efecto Hall), un soporte de espejo y un espejo. El tren de accionamiento 916 en este ejemplo puede tener una velocidad máxima de 4.000 RPM y una aceleración máxima de 300 grados/s². En algunas realizaciones, el espejo monógono es un espejo dieléctrico. En una realización, el espejo monógono incluye un revestimiento o capa hidrófoba.
- La colocación de los componentes del sistema de captura del entorno es tal que el conjunto de lente y el LiDAR se colocan sustancialmente en un centro de un eje de rotación. Esto puede reducir el paralaje de imagen que se produce cuando un sistema de captura de imágenes no se coloca en el centro del eje de rotación.
- En algunas realizaciones, la lente de WFOV 918 puede ser la lente del conjunto de lente 704 de la Figura 7. La lente de WFOV 918 enfoca la luz sobre un dispositivo de captura de imágenes. En algunas realizaciones, la lente de WFOV puede tener un FOV de al menos 145 grados. Con un FOV tan amplio, puede obtenerse una captura de imagen de una vista de 360 grados alrededor del sistema de captura del entorno con tres capturas de imagen separadas del dispositivo de captura de imágenes. En algunas realizaciones, la lente de WFOV 918 puede tener aproximadamente ~60 mm de diámetro y ~80 mm de longitud total de seguimiento (TTL). En un ejemplo, la lente de WFOV 918 puede incluir un campo de visión horizontal que es mayor o igual que 148,3 grados y un campo de visión vertical que es mayor o igual que 94 grados.
- Un dispositivo de captura de imágenes puede incluir la lente de WFOV 918 y el sensor de imagen 920. El sensor de imagen 920 puede ser un sensor de imagen CMOS. En una realización, el sensor de imagen 920 es un dispositivo acoplado cargado (CCD). En algunas realizaciones, el sensor de imagen 920 es un sensor rojo-verde-azul (RGB). En una realización, el sensor de imagen 920 es un sensor IR. En diversas realizaciones, el dispositivo de captura de imágenes puede tener una resolución de al menos 35 píxeles por grado (PPD).
- En algunas realizaciones, el dispositivo de captura de imágenes puede incluir un número F de $f/2,4$, diámetro del círculo de imagen de 15,86 mm, distancia entre píxeles de 2,4 μm , HFOV > 148,3°, VFOV > 94,0°, píxeles por grado > 38,0 PPD, ángulo de rayo principal a altura total de 3,0°, distancia mínima del objeto 1300 mm, distancia máxima del objeto, infinita, iluminación relativa > 130 %, distorsión máxima < 90 % y variación de transmisión espectral ≤ 5 %.
- En algunas realizaciones, la lente puede incluir número F 2,8, diámetro del círculo de imagen 15,86 mm, píxeles por grado > 37, ángulo de rayo principal a altura total del sensor 3,0, diámetro L1 < 60 mm, TTL < 80 mm e iluminación relativa > 50 %.
- La lente puede incluir 52 lp/mm (en el eje) > 85 %, 104 lp/mm (en el eje) > 66 %, 1308 lp/mm (en el eje) > 45 %, 52 lp/mm (83 % de campo) > 75 %, 104 lp/mm (83 % de campo) > 41 % y 1308 lp/mm (83 % de campo) > 25 %.

El sistema de captura del entorno puede tener una resolución de >20 MP, sensibilidad a color verde >1,7 V/lux*s, SNR (100 lux, ganancia lx) >65 dB y un rango dinámico de >70 dB.

5 La Figura 9b representa un diagrama de bloques de un PCBA de SOM 908 de ejemplo del sistema de captura del entorno de acuerdo con algunas realizaciones. El PCBA de SOM 908 puede incluir un componente de comunicación 922, un componente de control de LiDAR 924, un componente de ubicación de LiDAR 926, un componente de interfaz de usuario 928, un componente de clasificación 930, un almacenamiento de datos de LiDAR 932 y un almacenamiento de datos de imagen capturada 934.

10 En algunas realizaciones, el componente de comunicación 922 puede enviar y recibir solicitudes o datos entre cualquiera de los componentes del PCBA de SOM 1008 y los componentes del sistema de captura del entorno de la Figura 9a.

15 En diversas realizaciones, el componente de control de LiDAR 924 puede controlar diversos aspectos del LiDAR. Por ejemplo, el componente de control de LiDAR 924 puede enviar una señal de control al LiDAR 912 para comenzar a enviar un pulso láser. La señal de control enviada por el componente de control de LiDAR 924 puede incluir instrucciones sobre la frecuencia de los pulsos láser.

20 En algunas realizaciones, el componente de ubicación de LiDAR 926 puede utilizar datos de GPS para determinar la ubicación del sistema de captura del entorno. En diversas realizaciones, el componente de ubicación de LiDAR 926 utiliza la posición del conjunto de espejo para determinar el ángulo de exploración y las coordenadas (x, y, z) asociadas con cada pulso láser. El componente de ubicación de LiDAR 926 también puede utilizar la IMU para determinar la orientación del sistema de captura del entorno.

25 El componente de interfaz de usuario 928 puede facilitar la interacción de usuario con el sistema de captura del entorno. En algunas realizaciones, el componente de interfaz de usuario 928 puede proporcionar uno o más elementos de interfaz de usuario con los que un usuario puede interactuar. La interfaz de usuario proporcionada por el componente de interfaz de usuario 928 puede enviarse al sistema de usuario 1110. Por ejemplo, el componente de interfaz de usuario 928 puede proporcionar al sistema de usuario (por ejemplo, un dispositivo digital) una representación visual de un área de un plano de planta de un edificio. A medida que el usuario coloca el sistema de captura del entorno en diferentes partes del piso del edificio para capturar y generar imágenes panorámicas 3D, el sistema de captura del entorno puede generar la representación visual del plano de planta. El usuario puede colocar el sistema de captura del entorno en un área del entorno físico para capturar y generar imágenes panorámicas 3D en esa región de la casa. Una vez que la imagen panorámica 3D del área ha sido generada por el sistema de procesamiento de imágenes, el componente de interfaz de usuario puede actualizar la vista en planta con una vista de arriba hacia abajo del área de la sala de estar representada en la Figura 1b. En algunas realizaciones, la vista en planta 200 puede generarse por el sistema de usuario 1110 después de que se haya capturado un segundo barrido de la misma casa, o planta de un edificio.

40 En diversas realizaciones, el componente de clasificación 930 puede clasificar el tipo de entorno físico. El componente de clasificación 930 puede analizar objetos en las imágenes u objetos en imágenes para clasificar el tipo de entorno físico que fue capturado por el sistema de captura del entorno. En algunas realizaciones, el sistema de procesamiento de imágenes puede ser responsable de clasificar el tipo de entorno físico que fue capturado por el sistema de captura 400 del entorno.

50 El almacenamiento de datos de LiDAR 932 puede ser cualquier estructura y/o estructuras adecuadas para datos de LiDAR capturados (por ejemplo, una base de datos activa, una base de datos relacional, una base de datos autorreferencial, una tabla, una matriz, una disposición, un archivo plano, un sistema de almacenamiento orientado, un sistema No-SQL no relacional, un sistema de gestión de FTS tal como Lucene/Solar y/o similares). El almacenamiento de datos de imagen 408 puede almacenar los datos de LiDAR capturados. Sin embargo, el almacenamiento de datos de LiDAR 932 puede utilizarse para almacenar en caché los datos de LiDAR capturados en casos donde la red de comunicación 404 no es funcional. Por ejemplo, en los casos donde el sistema de captura 402 del entorno y el sistema de usuario 1110 están en una ubicación remota sin red celular o en una región sin Wi-Fi, el almacenamiento de datos de LiDAR 932 puede almacenar los datos de LiDAR capturados hasta que puedan transferirse al almacenamiento de datos de imágenes 934.

60 De manera similar al almacenamiento de datos de LiDAR, el almacenamiento de datos de imágenes capturadas 934 puede ser cualquier estructura y/o estructuras adecuadas para imágenes capturadas (por ejemplo, una base de datos activa, una base de datos relacional, una base de datos autorreferencial, una tabla, una matriz, una disposición, un archivo plano, un sistema de almacenamiento orientado, un sistema No-SQL no relacional, un sistema de gestión de FTS tal como Lucene/Solar y/o similares). El almacenamiento de datos de imagen 934 puede almacenar las imágenes capturadas.

65 Las Figuras 10a-10c representan un proceso para el sistema de captura 400 del entorno para tomar imágenes en algunas realizaciones. Como se representa en la Figura 10a-10c, el sistema de captura 400 del entorno puede tomar

una ráfaga de imágenes a diferentes exposiciones. Una ráfaga de imágenes puede ser un conjunto de imágenes, cada una con diferentes exposiciones. La primera ráfaga de imagen tiene lugar en el tiempo 0,0. El sistema de captura 400 del entorno puede recibir el primer fotograma y a continuación evaluar el fotograma mientras espera el segundo fotograma. La Figura 10a indica que el primer fotograma se combina antes de que llegue el segundo fotograma. En algunas realizaciones, el sistema de captura 400 del entorno puede procesar cada fotograma para identificar píxeles, color y similares. Una vez que llega el siguiente fotograma, el sistema de captura 400 del entorno puede procesar el fotograma recibido recientemente y a continuación combinar los dos fotogramas juntos.

En diversas realizaciones, el sistema de captura 400 del entorno realiza procesamiento de imagen para combinar el sexto fotograma y evaluar adicionalmente los píxeles en el fotograma combinado (por ejemplo, el fotograma que puede incluir elementos de cualquier número de los fotogramas de la ráfaga de imagen). Durante la última etapa antes de o durante el movimiento (por ejemplo, giro) del sistema de captura 400 del entorno, el sistema de captura 400 del entorno puede transferir opcionalmente la imagen combinada de la unidad de procesamiento de gráficos a la memoria de CPU.

El proceso continúa en la Figura 10b. Al comienzo de la Figura 10b, el sistema de captura 400 del entorno realiza otra ráfaga. El sistema de captura 400 del entorno puede comprimir los fotogramas combinados y/o todos o partes de los fotogramas capturados usando JxR). Como la Figura 10a, una ráfaga de imágenes puede ser un conjunto de imágenes, cada una con diferentes exposiciones (la duración de exposición para cada fotograma del conjunto puede ser la misma y en el mismo orden que otras ráfagas cubiertas en las Figuras 10a y 10c). La segunda ráfaga de imagen tiene lugar en el tiempo 2 segundos. El sistema de captura 400 del entorno puede recibir el primer fotograma y a continuación evaluar el fotograma mientras espera el segundo fotograma. La Figura 10b indica que el primer fotograma se combina antes de que llegue el segundo fotograma. En algunas realizaciones, el sistema de captura 400 del entorno puede procesar cada fotograma para identificar píxeles, color y similares. Una vez que llega el siguiente fotograma, el sistema de captura 400 del entorno puede procesar el fotograma recibido recientemente y a continuación combinar los dos fotogramas juntos.

En diversas realizaciones, el sistema de captura 400 del entorno realiza procesamiento de imagen para combinar el sexto fotograma y evaluar adicionalmente los píxeles en el fotograma combinado (por ejemplo, el fotograma que puede incluir elementos de cualquier número de los fotogramas de la ráfaga de imagen). Durante la última etapa antes de o durante el movimiento (por ejemplo, giro) del sistema de captura 400 del entorno, el sistema de captura 400 del entorno puede transferir opcionalmente la imagen combinada de la unidad de procesamiento de gráficos a la memoria de CPU.

Después de girar, el sistema de captura 400 del entorno puede continuar el proceso realizando otra ráfaga de color (por ejemplo, después de girar 180 grados) en aproximadamente un tiempo de 3,5 segundos. El sistema de captura 400 del entorno puede comprimir los fotogramas combinados y/o todos o partes de los fotogramas capturados usando JxR). La ráfaga de imágenes puede ser un conjunto de imágenes, cada una con diferentes exposiciones (la duración de exposición para cada fotograma del conjunto puede ser la misma y en el mismo orden que otras ráfagas cubiertas en las Figuras 10a y 10c). El sistema de captura 400 del entorno puede recibir el primer fotograma y a continuación evaluar el fotograma mientras espera el segundo fotograma. La Figura 10b indica que el primer fotograma se combina antes de que llegue el segundo fotograma. En algunas realizaciones, el sistema de captura 400 del entorno puede procesar cada fotograma para identificar píxeles, color y similares. Una vez que llega el siguiente fotograma, el sistema de captura 400 del entorno puede procesar el fotograma recibido recientemente y a continuación combinar los dos fotogramas juntos.

En diversas realizaciones, el sistema de captura 400 del entorno realiza procesamiento de imagen para combinar el sexto fotograma y evaluar adicionalmente los píxeles en el fotograma combinado (por ejemplo, el fotograma que puede incluir elementos de cualquier número de los fotogramas de la ráfaga de imagen). Durante la última etapa antes de o durante el movimiento (por ejemplo, giro) del sistema de captura 400 del entorno, el sistema de captura 400 del entorno puede transferir opcionalmente la imagen combinada de la unidad de procesamiento de gráficos a la memoria de CPU.

La última ráfaga ocurre en un tiempo de 5 segundos en la Figura 10c. El sistema de captura 400 del entorno puede comprimir los fotogramas combinados y/o todos o partes de los fotogramas capturados usando JxR). La ráfaga de imágenes puede ser un conjunto de imágenes, cada una con diferentes exposiciones (la duración de exposición para cada fotograma del conjunto puede ser la misma y en el mismo orden que otras ráfagas cubiertas en las Figuras 10a y 10b). El sistema de captura 400 del entorno puede recibir el primer fotograma y a continuación evaluar el fotograma mientras espera el segundo fotograma. La Figura 10c indica que el primer fotograma se combina antes de que llegue el segundo fotograma. En algunas realizaciones, el sistema de captura 400 del entorno puede procesar cada fotograma para identificar píxeles, color y similares. Una vez que llega el siguiente fotograma, el sistema de captura 400 del entorno puede procesar el fotograma recibido recientemente y a continuación combinar los dos fotogramas juntos.

En diversas realizaciones, el sistema de captura 400 del entorno realiza procesamiento de imagen para combinar el sexto fotograma y evaluar adicionalmente los píxeles en el fotograma combinado (por ejemplo, el fotograma que puede incluir elementos de cualquier número de los fotogramas de la ráfaga de imagen). Durante la última etapa antes de o durante el movimiento (por ejemplo, giro) del sistema de captura 400 del entorno, el sistema de captura 400 del entorno puede transferir opcionalmente la imagen combinada de la unidad de procesamiento de gráficos a la memoria de CPU.

El rango dinámico de un dispositivo de captura de imágenes es una medida de cuánta luz puede capturar un sensor de imagen. El rango dinámico es la diferencia entre el área más oscura y el área más brillante de una imagen. Hay muchas formas de aumentar el rango dinámico del dispositivo de captura de imágenes, una de las cuales es capturar múltiples imágenes del mismo entorno físico usando diferentes exposiciones. Una imagen capturada con una exposición corta capturará áreas más brillantes del entorno físico, mientras que una exposición larga capturará áreas de entorno físico más oscuras. En algunas realizaciones, el sistema de captura del entorno puede capturar múltiples imágenes con seis tiempos de exposición diferentes. Algunas o todas las imágenes capturadas por el sistema de captura del entorno se usan para generar imágenes 2D con alto rango dinámico (HDR). Una o más de las imágenes capturadas pueden usarse para otras funciones tales como detección de luz ambiental, detección de parpadeo y similares.

Puede generarse una imagen panorámica 3D del entorno físico basándose en cuatro capturas de imagen separadas del dispositivo de captura de imágenes y cuatro capturas de datos de profundidad separadas del dispositivo LiDAR del sistema de captura del entorno. Cada una de las cuatro capturas de imagen separadas puede incluir una serie de capturas de imagen de diferentes tiempos de exposición. Puede usarse un algoritmo de combinación para combinar la serie de capturas de imagen con los diferentes tiempos de exposición para generar una de cuatro capturas de imagen RGB, que puede utilizarse para generar una imagen panorámica 2D. Por ejemplo, el sistema de captura del entorno puede usarse para capturar una imagen panorámica 3D de una cocina. Las imágenes de una pared de la cocina pueden incluir una ventana, una imagen con una imagen capturada con una exposición más corta puede proporcionar la vista por la ventana pero puede dejar el resto de la cocina subexpuesta. En contraste, otra imagen capturada con una exposición más larga puede proporcionar la vista del interior de la cocina. El algoritmo de combinación puede generar una imagen RGB combinada combinando la vista por la ventana de la cocina desde una imagen con el resto de la vista de la cocina desde otra imagen.

En diversas realizaciones, la imagen panorámica 3D puede generarse basándose en tres capturas de imagen separadas del dispositivo de captura de imágenes y cuatro capturas de datos de profundidad separadas del dispositivo LiDAR del sistema de captura del entorno. En algunas realizaciones, el número de capturas de imagen y el número de capturas de datos de profundidad pueden ser los mismos. En una realización, el número de capturas de imagen y el número de capturas de datos de profundidad pueden ser diferentes.

Después de capturar una primera de una serie de imágenes con un tiempo de exposición, un algoritmo de combinación recibe la primera de la serie de imágenes, calcula ponderaciones de intensidad iniciales para esa imagen, y establece esa imagen como una imagen de línea de base para combinar las imágenes recibidas posteriormente. En algunas realizaciones, el algoritmo de combinación puede utilizar una rutina de procesamiento de imagen de unidad de procesamiento gráfico (GPU) tal como una rutina de "núcleo de combinación". El algoritmo de combinación puede recibir imágenes posteriores que pueden combinarse con imágenes recibidas previamente. En algunas realizaciones, el algoritmo de combinación puede utilizar una variación de la rutina de procesamiento de imágenes de GPU de núcleo de combinación.

En una realización, el algoritmo de fusión utiliza otros métodos de fusión de múltiples imágenes, tales como determinar la diferencia entre la parte más oscura y más brillante, o contraste, de la imagen de línea de base para determinar si la imagen de línea de base puede estar sobreexpuesta o subexpuesta. Por ejemplo, un valor de contraste menor que un umbral de contraste predeterminado significa que la imagen de línea de base está sobreexpuesta o subexpuesta. En una realización, el contraste de la imagen de línea de base puede calcularse tomando un promedio de la intensidad de luz de la imagen o un subconjunto de la imagen. En algunas realizaciones, el algoritmo de combinación calcula una intensidad de luz promedio para cada fila o columna de la imagen. En algunas realizaciones, el algoritmo de combinación puede determinar un histograma de cada una de las imágenes recibidas desde el dispositivo de captura de imágenes y analizar el histograma para determinar las intensidades de luz de los píxeles que componen cada una de las imágenes.

En diversas realizaciones, la combinación puede implicar muestrear colores dentro de dos o más imágenes de la misma escena, incluyendo a lo largo de objetos y apariencias. Si hay una diferencia significativa en el color entre las dos imágenes (por ejemplo, dentro de un umbral predeterminado de color, tono, brillo, saturación y/o similares), un módulo de combinación (por ejemplo, en el sistema de captura 400 del entorno o el usuario el dispositivo 1110) puede combinar un tamaño predeterminado de ambas imágenes a lo largo de la posición donde existe la diferencia. En algunas realizaciones, cuanto mayor sea la diferencia en color o imagen en una posición en la imagen, mayor será la cantidad de espacio alrededor o cerca de la posición que puede combinarse.

En algunas realizaciones, después de la combinación, el módulo de combinación (por ejemplo, en el sistema de captura 400 del entorno o el dispositivo de usuario 1110) puede volver a escanear y muestrear colores a lo largo de la imagen o imágenes para determinar si hay otras diferencias en la imagen o el color que superar el umbral predeterminado de color, tono, brillo, saturación y/o similares. Si es así, el módulo de combinación puede identificar las porciones dentro de la imagen o imágenes y continuar combinando esa porción de la imagen. El módulo de combinación puede continuar remuestreando las imágenes a lo largo de la costura hasta que no haya porciones adicionales de las imágenes para combinar (por ejemplo, cualquier diferencia en el color está por debajo del umbral o umbrales predeterminados).

sistema de captura y unión

La Figura 11 representa un diagrama de bloques de un entorno 1100 de ejemplo capaz de capturar y unir imágenes para formar visualizaciones 3D de acuerdo con algunas realizaciones. El entorno 1100 de ejemplo incluye un sistema de captura 3D y panorámico 1102, una red de comunicación 1104, un sistema de unión y procesador de imágenes 1106, un almacenamiento de datos de imagen 1108, un sistema de usuario 1110 y una primera escena de un entorno físico 1112. El sistema de unión y captura 3D y panorámico 1102 y/o el sistema de usuario 1110 pueden incluir un dispositivo de captura de imágenes (por ejemplo, el sistema de captura 400 del entorno) que puede usarse para capturar imágenes de un entorno (por ejemplo, el entorno físico 1112).

El sistema de unión y captura 3D y panorámico 1102 y el sistema de unión y procesador de imágenes 1106 pueden ser una parte del mismo sistema (por ejemplo, parte de uno o más dispositivos digitales) que están acoplados, en comunicación, al sistema de captura 400 del entorno. En algunas realizaciones, una o más de la funcionalidad de los componentes del sistema de unión y captura 3D y panorámico 1102 y del sistema de unión y procesador de imágenes 1106 puede realizarse por el sistema de captura 400 del entorno. De manera similar o como alternativa, el sistema de unión y captura 3D y panorámico 1102 y el sistema de unión y procesador de imágenes 1106 pueden ser realizados por el sistema de usuario 1110 y/o el sistema de unión y procesador de imágenes 1106

El sistema de unión y captura panorámica 3D 1102 puede utilizarse por un usuario para capturar múltiples imágenes 2D de un entorno, tal como el interior de un edificio y/o y fuera del edificio. Por ejemplo, el usuario puede utilizar el sistema de unión y captura 3D y panorámico 1102 para capturar múltiples imágenes 2D de la primera escena del entorno físico 1112 proporcionada por el sistema de captura 400 del entorno. El sistema de captura y unión 3D y panorámico 1102 puede incluir un sistema de alineación y unión 1114. Como alternativa, el sistema de usuario 1110 puede incluir el sistema de alineación y unión 1114.

El sistema de alineación y unión 1114 puede ser un software, hardware o una combinación de ambos configurados para proporcionar orientación al usuario de un sistema de captura de imágenes (por ejemplo, en el sistema de unión y captura 3D y panorámico 1102 o el sistema de usuario 1110) y/o procesar imágenes para permitir que se realicen fotografías panorámicas mejoradas (por ejemplo, a través de unión, alineación, recorte y/o similares). El sistema de alineación y unión 1114 puede estar en un medio legible por ordenador (descrito en el presente documento). En algunas realizaciones, el sistema de alineación y unión 1114 puede incluir un procesador para realizar funciones.

Un ejemplo de la primera escena del entorno físico 1112 puede ser cualquier habitación, bienes inmuebles o similares (por ejemplo, una representación de una sala de estar). En algunas realizaciones, el sistema de unión y captura 3D y panorámico 1102 se utiliza para generar imágenes panorámicas 3D de entornos interiores. El sistema de unión y captura panorámica 3D 1102 puede ser, en algunas realizaciones, el sistema de captura 400 del entorno analizado con respecto a la Figura 4.

En algunas realizaciones, el sistema de unión y captura panorámica 3D 1102 puede estar en comunicación con un dispositivo para capturar imágenes y datos de profundidad así como software (por ejemplo, el sistema de captura 400 del entorno). Todo o parte del software puede instalarse en el sistema de unión y captura 3D panorámico 1102, el sistema de usuario 1110, el sistema de captura 400 del entorno, o ambos. En algunas realizaciones, el usuario puede interactuar con el sistema de unión y captura 3D y panorámico 1102 a través del sistema de usuario 1110.

El sistema de unión y captura 3D y panorámico 1102 o el sistema de usuario 1110 pueden obtener múltiples imágenes 2D. El sistema de unión y captura 3D y panorámico 1102 o el sistema de usuario 1110 puede obtener datos de profundidad (por ejemplo, desde un dispositivo LiDAR o similar).

En diversas realizaciones, una aplicación en el sistema de usuario 1110 (por ejemplo, un dispositivo inteligente del usuario tal como un teléfono inteligente o tableta) o una aplicación en el sistema de captura 400 del entorno puede proporcionar una guía visual o auditiva al usuario para tomar imágenes con el sistema de captura 400 del entorno. La guía gráfica puede incluir, por ejemplo, una flecha flotante en una pantalla del sistema de captura 400 del entorno (por ejemplo, en un visor o pantalla LED en la parte posterior del sistema de captura 400 del entorno) para guiar al usuario sobre dónde colocar y/o apuntar un dispositivo de captura de imágenes. En otro ejemplo, la aplicación puede proporcionar una guía de audio sobre dónde colocar y/o apuntar el dispositivo de captura de imágenes.

En algunas realizaciones, la guía puede permitir que el usuario capture múltiples imágenes del entorno físico sin la ayuda de una plataforma estabilizadora tal como un trípode. En un ejemplo, el dispositivo de captura de imágenes puede ser un dispositivo personal tal como un teléfono inteligente, tableta, tableta de medios, ordenador portátil y similares. La aplicación puede proporcionar dirección en la posición para cada barrido, para aproximar el punto sin paralaje basándose en la posición del dispositivo de captura de imágenes, información de ubicación del dispositivo de captura de imágenes y/o imagen previa del dispositivo de captura de imágenes.

En algunas realizaciones, la guía visual y/o auditiva permite la captura de imágenes que pueden unirse para formar panorámicas sin un trípode y sin información de posicionamiento de la cámara (por ejemplo, indicando una ubicación, posición y/u orientación de la cámara desde un sensor, dispositivo GPS o similar).

El sistema de alineación y unión 1114 puede alinear o unir imágenes 2D (por ejemplo, capturadas por el sistema de usuario 1110 o el sistema de unión y captura panorámica 3D 1102) para obtener una imagen panorámica 2D.

5 En algunas realizaciones, el sistema de alineación y unión 1114 utiliza un algoritmo de aprendizaje automático para alinear o unir múltiples imágenes 2D en una imagen panorámica 2D. Los parámetros del algoritmo de aprendizaje automático pueden gestionarse por el sistema de alineación y unión 1114. Por ejemplo, el sistema de unión y captura 3D y panorámico 1102 y/o el sistema de alineación y unión 1114 pueden reconocer objetos dentro de las imágenes 2D para ayudar a alinear las imágenes en una imagen panorámica 2D.

10 En algunas realizaciones, el sistema de alineación y unión 1114 puede utilizar datos de profundidad y la imagen panorámica 2D para obtener una imagen panorámica 3D. La imagen panorámica 3D puede proporcionarse al sistema de unión 3D y panorámico 1102 o al sistema de usuario 1110. En algunas realizaciones, el sistema de alineación y unión 1114 determina mediciones 3D/de profundidad asociadas con objetos reconocidos dentro de una imagen panorámica 3D y/o envía una o más imágenes 2D, datos de profundidad, imagen o imágenes panorámicas 2D, imagen o imágenes panorámicas 3D al sistema de unión y procesador de imágenes 106 para obtener una imagen panorámica 2D o una imagen panorámica 3D con una resolución de píxel que es mayor que la imagen panorámica 2D o la imagen panorámica 3D proporcionada por el sistema de unión y captura 3D y panorámico 1102.

20 La red de comunicación 1104 puede representar una o más redes informáticas (por ejemplo, LAN, WAN o similares) u otros medios de transmisión. La red de comunicación 1104 puede proporcionar comunicación entre los sistemas 1102, 1106-1110 y/u otros sistemas descritos en el presente documento. En algunas realizaciones, la red de comunicación 104 incluye uno o más dispositivos digitales, enrutadores, cables, buses y/u otras topologías de red (por ejemplo, malla y similares). En algunas realizaciones, red de comunicación 1104 puede ser cableada y/o inalámbrica.

25 En diversas realizaciones, la red de comunicación 1104 puede incluir Internet, una o más redes de área extensa (WAN) o redes de área local (LAN), una o más redes que pueden ser públicas, privadas, basadas en IP, no basadas en IP, y así sucesivamente.

30 El sistema de unión y procesador de imágenes 1106 puede procesar imágenes 2D capturadas por el dispositivo de captura de imágenes (por ejemplo, el sistema de captura 400 del entorno o un dispositivo de usuario tal como un teléfono inteligente, ordenador personal, tableta de medios o similar) y unirlas en un 2D imagen panorámica. La imagen panorámica 2D procesada por el sistema de unión y procesador de imágenes 106 puede tener una resolución de píxel más alta que la imagen panorámica obtenida por el sistema de captura y unión 3D y panorámico 1102.

35 En algunas realizaciones, el sistema de unión y procesador de imágenes 1106 recibe y procesa la imagen panorámica 3D para crear una imagen panorámica 3D con una resolución de píxel que es mayor que la de la imagen panorámica 3D recibida. Las imágenes panorámicas de resolución de píxeles más alta pueden proporcionarse a un dispositivo de salida con una resolución de pantalla más alta que el sistema de usuario 1110, tal como una pantalla de ordenador, pantalla de proyector y similares. En algunas realizaciones, las imágenes panorámicas de mayor resolución de píxeles pueden proporcionar al dispositivo de salida una imagen panorámica con mayor detalle y pueden ampliarse.

45 El almacenamiento de datos de imágenes 1108 puede ser cualquier estructura y/o estructuras adecuadas para imágenes y/o datos de profundidad capturados (por ejemplo, una base de datos activa, una base de datos relacional, una base de datos autorreferencial, una tabla, una matriz, una disposición, un archivo plano, un sistema de almacenamiento orientado, un sistema No-SQL no relacional, un sistema de gestión de FTS tal como Lucene/Solar y/o similares). El almacenamiento de datos de imágenes 1108 puede almacenar imágenes capturadas por el dispositivo de captura de imágenes del sistema de usuario 1110. En diversas realizaciones, el almacenamiento de datos de imagen 1108 almacena datos de profundidad capturados por uno o más sensores de profundidad del sistema de usuario 1110. En diversas realizaciones, el almacenamiento de datos de imagen 1108 almacena propiedades asociadas con el dispositivo de captura de imágenes o propiedades asociadas con cada una de las múltiples capturas de imagen o capturas de profundidad usadas para determinar la imagen panorámica 2D o 3D. En algunas realizaciones, el almacenamiento de datos de imagen 1108 almacena imágenes panorámicas 2D o 3D panorámicas. Las imágenes panorámicas 2D o 3D pueden determinarse por el sistema de unión y captura 3D y panorámico 1102 o el sistema de procesador y unión de imágenes 106.

55 El sistema de usuario 1110 puede comunicarse entre usuarios y otros sistemas asociados. En algunas realizaciones, el sistema de usuario 1110 puede ser o incluir uno o más dispositivos desplazables (por ejemplo, teléfonos inteligentes, teléfonos celulares, relojes inteligentes o similares).

60 El sistema de usuario 1110 puede incluir uno o más dispositivos de captura de imágenes. El uno o más dispositivos de captura de imágenes pueden incluir, por ejemplo, cámaras RGB, cámaras HDR, cámaras de vídeo, cámaras IR y similares.

65 El sistema de unión y captura 3D y panorámico 1102 y/o el sistema de usuario 1110 pueden incluir dos o más dispositivos de captura que pueden estar dispuestos en posiciones relativas entre sí en o dentro de la misma carcasa desplazable de modo que sus campos de visión colectivos abarquen hasta 360°. En algunas realizaciones, pueden

usarse pares de dispositivos de captura de imágenes capaces de generar pares de imágenes estéreo (por ejemplo, con campos de visión ligeramente desplazados pero parcialmente superpuestos). El sistema de usuario 1110 puede incluir dos dispositivos de captura de imágenes con campos de visión de desplazamiento estéreo vertical capaces de capturar pares de imágenes estéreo verticales. En otro ejemplo, el sistema de usuario 1110 puede comprender dos dispositivos de captura de imágenes con campos de visión de desplazamiento estéreo vertical capaces de capturar pares de imágenes estéreo verticales.

En algunas realizaciones, el sistema de usuario 1110, el sistema de captura 400 del entorno, o el sistema de unión y captura 3D y panorámico 1102 pueden generar y/o proporcionar información de ubicación y posición de captura de imagen. Por ejemplo, el sistema de usuario 1110 o el sistema de unión y captura 3D y panorámico 1102 puede incluir una unidad de medición inercial (IMU) para ayudar a determinar datos de posición en asociación con uno o más dispositivos de captura de imágenes que capturan las múltiples imágenes 2D. El sistema de usuario 1110 puede incluir un sensor de posicionamiento global (GPS) para proporcionar información de coordenadas de GPS en asociación con las múltiples imágenes 2D capturadas por uno o más dispositivos de captura de imágenes.

En algunas realizaciones, los usuarios pueden interactuar con el sistema de alineación y unión 1114 usando una aplicación desplazable instalada en el sistema de usuario 1110. El sistema de unión y captura 3D y panorámico 1102 puede proporcionar imágenes al sistema de usuario 1110. Un usuario puede utilizar el sistema de alineación y unión 1114 en el sistema de usuario 1110 para ver imágenes y vistas previas.

En diversas realizaciones, el sistema de alineación y unión 1114 puede configurarse para proporcionar o recibir una o más imágenes panorámicas 3D desde el sistema de captura y unión 3D y panorámico 1102 y/o el sistema de unión y procesador de imágenes 1106. En algunas realizaciones, el sistema de unión y captura 3D y panorámico 1102 puede proporcionar una representación visual de una porción de un plano de planta de un edificio, que ha sido capturado por el sistema de unión y captura 3D y panorámico 1102 al sistema de usuario 1110.

El usuario del sistema 1110 puede navegar por el espacio alrededor del área y ver diferentes habitaciones de la casa. En algunas realizaciones, el usuario del sistema de usuario 1110 puede visualizar las imágenes panorámicas 3D, tal como la imagen panorámica 3D de ejemplo, a medida que el sistema de unión y procesador de imágenes 1106 completa la generación de la imagen panorámica 3D. En diversas realizaciones, el sistema de usuario 1110 genera una vista previa o miniatura de la imagen panorámica 3D. La imagen panorámica 3D de vista previa puede tener una resolución de imagen que es menor que una imagen panorámica 3D generada por el sistema de unión y captura 3D y panorámico 1102.

La Figura 12 representa un diagrama de bloques de un ejemplo del sistema de alineación y unión 1114 de acuerdo con algunas realizaciones. El sistema de alineación y unión 1114 incluye un módulo de comunicación 1202, un módulo de posición de captura de imagen 1204, un módulo de unión 1206, un módulo de recorte 1208, un módulo de corte gráfico 1210, un módulo de combinación 1211, un generador de imágenes 3D 1214, una imagen 2D capturada el almacenamiento de datos 1216, un almacenamiento de datos de imagen panorámica 3D 1218 y un módulo de guiado 220. Puede apreciarse que puede haber cualquier número de módulos del sistema de alineación y unión 1114 que realizan una o más funciones diferentes como se describe en el presente documento.

En algunas realizaciones, el sistema de alineación y unión 1114 incluye un módulo de captura de imágenes configurado para recibir imágenes desde uno o más dispositivos de captura de imágenes (por ejemplo, cámaras). El sistema de alineación y unión 1114 también puede incluir un módulo de profundidad configurado para recibir datos de profundidad desde un dispositivo de profundidad tal como un LiDAR si está disponible.

El módulo de comunicación 1202 puede enviar y recibir solicitudes, imágenes o datos entre cualquiera de los módulos o almacenamientos de datos del sistema de alineación y unión 1114 y los componentes del entorno 1100 de ejemplo de la Figura 11. De manera similar, el sistema de alineación y unión 1114 puede enviar y recibir solicitudes, imágenes o datos a través de la red de comunicación 1104 a cualquier dispositivo o sistema.

En algunas realizaciones, el módulo de posición de captura de imagen 1204 puede determinar datos de posición de dispositivo de captura de imágenes de un dispositivo de captura de imágenes (por ejemplo, una cámara que puede ser una cámara independiente, teléfono inteligente, tableta de medios, ordenador portátil o similares). Los datos de posición del dispositivo de captura de imágenes pueden indicar una posición y orientación de un dispositivo de captura de imágenes y/o lente. En un ejemplo, el módulo de posición de captura de imagen 1204 puede utilizar la IMU del sistema de usuario 1110, cámara, dispositivo digital con una cámara, o el sistema de captura y unión 3D y panorámico 1102 para generar datos de posición del dispositivo de captura de imágenes. El módulo de posición de captura de imagen 1204 puede determinar la dirección, ángulo o inclinación actual de uno o más dispositivos de captura de imágenes (o lentes). El módulo de posición de captura de imagen 1204 también puede utilizar el GPS del sistema de usuario 1110 o el sistema de unión y captura 3D y panorámico 1102.

Por ejemplo, cuando un usuario desea usar el sistema de usuario 1110 para capturar una vista de 360° del entorno físico, tal como una sala de estar, el usuario puede sostener el sistema de usuario 1110 frente a él o ella a la altura de los ojos para comenzar a capturar una de un múltiplo de imágenes que eventualmente se convertirán en una imagen

panorámica 3D. Para reducir la cantidad de paralaje a la imagen y capturar imágenes más adecuadas para unir y generar imágenes panorámicas 3D, puede ser preferible si uno o más dispositivos de captura de imágenes rotan en el centro del eje de rotación. El sistema de alineación y unión 1114 puede recibir información de posición (por ejemplo, desde la IMU) para determinar la posición del dispositivo o lente de captura de imágenes. El sistema de alineación y unión 1114 puede recibir y almacenar un campo de visión de la lente. El módulo de guiado 1220 puede proporcionar información visual y/o de audio con respecto a una posición inicial recomendada del dispositivo de captura de imágenes. El módulo de guiado 1220 puede hacer recomendaciones para posicionar el dispositivo de captura de imágenes para imágenes posteriores. En un ejemplo, el módulo de guiado 1220 puede proporcionar una guía al usuario para rotar y posicionar el dispositivo de captura de imágenes de manera que el dispositivo de captura de imágenes rota cerca de un centro de rotación. Además, el módulo de guiado 1220 puede proporcionar guía al usuario para rotar y posicionar el dispositivo de captura de imágenes de tal manera que las imágenes posteriores se alinean sustancialmente basándose en características del campo de visión y/o dispositivo de captura de imágenes.

El módulo de guiado 1220 puede proporcionar al usuario una guía visual. Por ejemplo, el módulo de guiado 1220 puede colocar marcadores o una flecha en un visor o visualizador en el sistema de usuario 1110 o el sistema de captura y unión 3D y panorámico 1102. En algunas realizaciones, el sistema de usuario 1110 puede ser un ordenador de teléfono inteligente o tableta con una pantalla. Cuando se toman una o más fotografías, el módulo de guiado 1220 puede colocar uno o más marcadores (por ejemplo, marcadores de color diferentes o los mismos marcadores) en un dispositivo de salida y/o en un visor. El usuario puede usar a continuación los marcadores en el dispositivo de salida y/o visor para alinear la siguiente imagen.

Existen numerosas técnicas para guiar al usuario del sistema de usuario 1110 o el sistema de unión y captura 3D y panorámico 1102 para tomar múltiples imágenes para facilitar la unión de las imágenes en un panorama. Cuando se toma una panorámica a partir de múltiples imágenes, las imágenes pueden unirse entre sí. Para mejorar el tiempo, la eficiencia y la efectividad de unir las imágenes con una necesidad reducida de corregir alteraciones de imagen o desalineaciones, el módulo de posición de captura de imagen 1204 y el módulo de guiado 1220 pueden ayudar al usuario a tomar múltiples imágenes en posiciones que mejoran la calidad, la eficiencia de tiempo y la efectividad de la unión de imágenes para el panorama deseado.

Por ejemplo, después de tomar la primera instantánea, la pantalla del sistema de usuario 1110 puede incluir dos o más objetos, tales como círculos. Dos círculos pueden parecer estacionarios con respecto al entorno y dos círculos pueden moverse con el sistema de usuario 1110. Cuando los dos círculos estacionarios están alineados con los dos círculos que se mueven con el sistema de usuario 1110, el dispositivo de captura de imágenes y/o el sistema de usuario 1110 pueden estar alineados para la siguiente imagen.

En algunas realizaciones, después de que se toma una imagen por un dispositivo de captura de imágenes, el módulo de posición de captura de imagen 1204 puede tomar una medición de sensor de la posición del dispositivo de captura de imágenes (por ejemplo, incluyendo orientación, inclinación y similares). El módulo de posición de captura de imagen 1204 puede determinar uno o más bordes de la imagen que se tomó calculando la ubicación del borde de un campo de visión basándose en la medición del sensor. Adicionalmente, o como alternativa, el módulo de posición de captura de imagen 1204 puede determinar uno o más bordes de la imagen escaneando la imagen tomada por el dispositivo de captura de imágenes, identificando objetos dentro de esa imagen (por ejemplo, usando modelos de aprendizaje automático analizados en el presente documento), determinando uno o más bordes de la imagen, y colocar objetos (por ejemplo, círculos u otras formas) en el borde de una pantalla en el sistema de usuario 1110.

El módulo de posición de captura de imagen 1204 puede visualizar dos objetos dentro de un visualizador del sistema de usuario 1110 que indica el posicionamiento del campo de visión para la siguiente instantánea. Estos dos objetos pueden indicar posiciones en el entorno que representan dónde hay un borde de la última imagen. El módulo de posición de captura de imagen 1204 puede continuar recibiendo mediciones de sensor de la posición del dispositivo de captura de imágenes y calcular dos objetos adicionales en el campo de visión. Los dos objetos adicionales pueden estar separados por la misma anchura que los dos objetos anteriores. Mientras que los dos primeros objetos pueden representar un borde de la imagen tomada (por ejemplo, el borde más a la derecha de la imagen), los siguientes dos objetos adicionales que representan un borde del campo de visión pueden estar en el borde opuesto (por ejemplo, el borde más a la izquierda del campo de visión). Al hacer que el usuario alinee físicamente los dos primeros objetos en el borde de la imagen con los dos objetos adicionales en el borde opuesto del campo de visión, el dispositivo de captura de imágenes se puede colocar para tomar otra imagen que se puede unir de manera más efectiva sin un trípode. Este proceso puede continuar para cada imagen hasta que el usuario determina que se ha capturado la panorámica deseada.

Aunque se analizan múltiples objetos en el presente documento, se apreciará que el módulo de posición de captura de imagen 1204 puede calcular la posición de uno o más objetos para posicionar el dispositivo de captura de imágenes. Los objetos pueden tener cualquier forma (por ejemplo, circular, oblonga, cuadrada, emoji, flechas o similares). En algunas realizaciones, los objetos pueden tener diferentes formas.

En algunas realizaciones, puede haber una distancia entre los objetos que representan el borde de una imagen capturada y la distancia entre los objetos de un campo de visión. El usuario puede ser guiado para moverse hacia

adelante para alejarse para permitir que haya suficiente distancia entre los objetos. Como alternativa, el tamaño de los objetos en el campo de visión puede cambiar para coincidir con un tamaño de los objetos que representan un borde de una imagen capturada a medida que el dispositivo de captura de imágenes se acerca a la posición correcta (por ejemplo, acercándose o alejándose de una posición que permitirá tomar la siguiente imagen en una posición que mejorará la unión de imágenes.

En algunas realizaciones, el módulo de posición de captura de imagen 1204 puede utilizar objetos en una imagen capturada por el dispositivo de captura de imágenes para estimar la posición del dispositivo de captura de imágenes. Por ejemplo, el módulo de posición de captura de imagen 1204 puede utilizar coordenadas de GPS para determinar la ubicación geográfica asociada con la imagen. El módulo de posición de captura de imagen 1204 puede usar la posición para identificar puntos de referencia que pueden capturarse por el dispositivo de captura de imágenes.

El módulo de posición de captura de imagen 1204 puede incluir un modelo de aprendizaje automático 2D para convertir imágenes 2D en imágenes panorámicas 2D. El módulo de posición de captura de imagen 1204 puede incluir un modelo de aprendizaje automático 3D para convertir imágenes 2D en representaciones 3D. En un ejemplo, puede utilizarse una representación 3D para visualizar un recorrido o visualización tridimensional de un entorno interior y/o exterior.

El modelo de aprendizaje automático 2D puede entrenarse para unir o ayudar a unir dos o más imágenes 2D juntas para formar una imagen panorámica 2D. El modelo de aprendizaje automático 2D puede ser, por ejemplo, una red neuronal entrenada con imágenes 2D que incluyen objetos físicos en las imágenes así como información de identificación de objeto para entrenar el modelo de aprendizaje automático 2D para identificar objetos en imágenes 2D posteriores. Los objetos en las imágenes 2D pueden ayudar a determinar la posición o posiciones dentro de una imagen 2D para ayudar a determinar bordes de la imagen 2D, deformar en la imagen 2D, y ayudar en la alineación de la imagen. Además, los objetos en las imágenes 2D pueden ayudar a determinar alteraciones de imagen en la imagen 2D, a la combinación de una alteración de imagen o borde entre dos imágenes, a las posiciones para cortar imágenes y/o recortar las imágenes.

En algunas realizaciones, el modelo de aprendizaje automático 2D puede ser, por ejemplo, una red neuronal entrenada con imágenes 2D que incluyen información de profundidad (por ejemplo, desde un dispositivo LiDAR o dispositivo de luz estructurada del sistema de usuario 1110 o el sistema de captura y unión 3D y panorámico 1102) del entorno, así como incluir objetos físicos en las imágenes para identificar los objetos físicos, la posición de los objetos físicos y/o la posición del dispositivo de captura de imágenes/campo de visión. El modelo de aprendizaje automático 2D puede identificar objetos físicos así como su profundidad en relación con otros aspectos de las imágenes 2D para ayudar en la alineación y posición de dos imágenes 2D para la unión (o para unir las dos imágenes 2D).

El modelo de aprendizaje automático 2D puede incluir cualquier número de modelos de aprendizaje automático (por ejemplo, cualquier número de modelos generados por redes neuronales o similares).

El modelo de aprendizaje automático 2D puede almacenarse en el sistema de unión y captura 3D y panorámico 1102, el sistema de unión y procesador de imágenes 1106 y/o el sistema de usuario 1110. En algunas realizaciones, el modelo de aprendizaje automático 2D puede ser entrenado por el sistema de unión y procesador de imágenes 1106.

El módulo de posición de captura de imagen 1204 puede estimar la posición del dispositivo de captura de imágenes (una posición del campo de visión del dispositivo de captura de imágenes) basándose en una costura entre dos o más imágenes 2D del módulo de unión 1206, deformándose la imagen desde el módulo de recorte 1208 y/o por el corte gráfico del módulo de corte gráfico 1210.

El módulo de unión 1206 puede combinar dos o más imágenes 2D para generar una panorámica 2D. Basándose en la unión entre dos o más imágenes 2D del módulo de unión 1206, la deformación de la imagen del módulo de recorte 1208 y/o un corte gráfico, que tiene un campo de visión que es mayor que el campo de visión de cada una de las dos o más imágenes.

El módulo de unión 1206 puede estar configurado para alinear o "unir entre sí" dos imágenes 2D diferentes que proporcionan diferentes perspectivas del mismo entorno para generar una imagen 2D panorámica del entorno. Por ejemplo, el módulo de unión 1206 puede emplear información conocida o derivada (por ejemplo, usando técnicas descritas en el presente documento) con respecto a las posiciones de captura y orientaciones de respectivas imágenes 2D para ayudar a unir dos imágenes entre sí.

El módulo de unión 1206 puede recibir dos imágenes 2D. La primera imagen 2D puede haberse tomado inmediatamente antes de la segunda imagen o dentro de un período de tiempo predeterminado. En diversas realizaciones, el módulo de unión 1206 puede recibir información de posicionamiento del dispositivo de captura de imágenes asociado con la primera imagen y a continuación información de posicionamiento asociada con la segunda imagen. La información de posicionamiento puede asociarse con una imagen basándose, en el momento en que se tomó la imagen, en los datos de posicionamiento de la IMU, GPS y/o información proporcionada por el usuario.

En algunas realizaciones, el módulo de unión 1206 puede utilizar un módulo de aprendizaje de máquina 2D para

escanear ambas imágenes para reconocer objetos dentro de ambas imágenes, incluyendo objetos (o partes de objetos) que pueden compartirse por ambas imágenes. Por ejemplo, el módulo de unión 1206 puede identificar una esquina, patrón en una pared, mueble o similar compartido en bordes opuestos de ambas imágenes.

5 El módulo de unión 1206 puede alinear bordes de las dos imágenes 2D basándose en el posicionamiento de los objetos compartidos (o partes de objetos), datos de posicionamiento de la IMU, datos de posicionamiento del GPS y/o información proporcionada por el usuario y luego combinar los dos bordes de las imágenes (es decir, "unirlas" entre sí). En algunas realizaciones, el módulo de unión 1206 puede identificar una porción de las dos imágenes 2D que se superponen entre sí y unir las imágenes en la posición que se superpone (por ejemplo, usando los datos de
10 posicionamiento y/o los resultados del modelo de aprendizaje automático 2D.

En diversas realizaciones, el modelo de aprendizaje automático 2D puede entrenarse para usar los datos de posicionamiento de la IMU, datos de posicionamiento del GPS y/o información proporcionada por el usuario para combinar o unir los dos bordes de las imágenes. En algunas realizaciones, el modelo de aprendizaje automático 2D
15 puede entrenarse para identificar objetos comunes en ambas imágenes 2D para alinear y posicionar las imágenes 2D y a continuación combinar o unir los dos bordes de las imágenes. En realizaciones adicionales, el modelo de aprendizaje automático 2D puede entrenarse para usar los datos de posicionamiento y el reconocimiento de objetos para alinear y posicionar las imágenes 2D y a continuación unir los dos bordes de las imágenes para formar toda o
20 parte de la imagen 2D panorámica.

El módulo de unión 1206 puede utilizar información de profundidad para las imágenes respectivas (por ejemplo, píxeles en las imágenes respectivas, objetos en las imágenes respectivas o similares) para facilitar la alineación de las imágenes 2D respectivas entre sí en asociación con la generación de una única imagen panorámica 2D del entorno.

25 El módulo de recorte 1208 puede resolver problemas con dos o más imágenes 2D donde el dispositivo de captura de imágenes no se mantuvo en la misma posición cuando se capturaron imágenes 2D. Por ejemplo, mientras se captura una imagen, el usuario puede colocar el sistema de usuario 1110 en una posición vertical. Sin embargo, mientras captura otra imagen, el usuario puede colocar el sistema de usuario en un ángulo. Las imágenes resultantes pueden no estar alineadas y pueden sufrir efectos de paralaje. Pueden producirse efectos de paralaje cuando los objetos de
30 primer plano y de fondo no se alinean de la misma manera en la primera imagen y en la segunda imagen.

El módulo de recorte 1208 puede utilizar el modelo de aprendizaje automático 2D (aplicando información de posicionamiento, información de profundidad y/o reconocimiento de objetos) para detectar cambios en la posición del dispositivo de captura de imágenes en dos o más imágenes y luego medir la cantidad de cambio en posición del
35 dispositivo de captura de imágenes. El módulo de recorte 1208 puede deformar una o múltiples imágenes 2D para que las imágenes puedan alinearse juntas para formar una imagen panorámica cuando las imágenes se unen, y al mismo tiempo conservar ciertas características de las imágenes, tales como mantener la rectitud de una línea recta.

La salida del módulo de recorte 1208 puede incluir el número de columnas y filas de píxeles para desplazar cada píxel de la imagen para enderezar la imagen. La cantidad de desplazamiento para cada imagen puede emitirse en forma de una matriz que representa el número de columnas de píxeles y filas de píxeles para desplazar cada píxel de la
40 imagen.

En algunas realizaciones, el módulo de recorte 1208 puede determinar la cantidad de deformación de imagen a realizar en una o más de las múltiples imágenes 2D capturadas por los dispositivos de captura de imágenes del sistema de usuario 1110 basándose en una o más de la posición de captura de imágenes del módulo de posición de captura de imagen 1204 o la costura entre dos o más imágenes 2D del módulo de unión 1206, el corte gráfico del módulo de corte gráfico 1210 o la combinación de colores del módulo de combinación 1211.

50 El módulo de corte gráfico 1210 puede determinar dónde cortar o segmentar una o más de las imágenes 2D capturadas por el dispositivo de captura de imágenes. Por ejemplo, el módulo de corte gráfico 1210 puede utilizar el modelo de aprendizaje automático 2D para identificar objetos en ambas imágenes y determinar que son el mismo objeto. El módulo de posición de captura de imagen 1204, el módulo de recorte 1208 y/o el módulo de corte gráfico 1210 pueden determinar que las dos imágenes no pueden alinearse, incluso si están deformadas. El módulo de corte gráfico 1210
55 puede utilizar la información del modelo de aprendizaje automático 2D para identificar secciones de ambas imágenes que pueden unirse (por ejemplo, cortando una parte de una o ambas imágenes para ayudar a su alineación y posicionamiento). En algunas realizaciones, las dos imágenes 2D pueden superponerse en al menos una porción del mundo físico representado en las imágenes. El módulo de corte gráfico 1210 puede identificar un objeto, tal como la misma silla, en ambas imágenes. Sin embargo, las imágenes de la silla pueden no alinearse para generar una panorámica que no está distorsionada y no representaría correctamente la porción del mundo físico, incluso después del posicionamiento de captura de imagen y la deformación de imagen realizada por el módulo de recorte 1208. El módulo de corte gráfico 1210 puede seleccionar una de las dos imágenes de la silla para que sea la representación correcta (por ejemplo, basándose en la desalineación, posicionamiento y/o alteraciones de imagen de una imagen en comparación con la otra) y cortar la silla de la imagen con desalineación, errores de posicionamiento y/o alteraciones
60 de imagen. El módulo de unión 1206 puede unir posteriormente las dos imágenes juntas.
65

5 El módulo de corte gráfico 1210 puede probar ambas combinaciones, por ejemplo, cortar la imagen de la silla de la primera imagen y unir la primera imagen, eliminar la silla a la segunda imagen, para determinar qué corte gráfico genera una imagen panorámica más precisa. La salida del módulo de corte gráfico 1210 puede ser una ubicación para cortar una o más de las múltiples imágenes 2D que corresponden al corte gráfico, que genera una imagen panorámica más precisa.

10 El módulo de corte gráfico 1210 puede determinar cómo cortar o segmentar una o más de las imágenes 2D capturadas por el dispositivo de captura de imágenes basándose en una o más posiciones de captura de imágenes desde el módulo de posición de captura de imagen 1204, cosiendo o uniendo entre dos o más 2D imágenes del módulo de unión 1206, la deformación de imagen del módulo de recorte 1208 y el corte gráfico del módulo de corte gráfico 1210.

15 El módulo de combinación 1211 puede colorear en las costuras (por ejemplo, unión) entre dos imágenes de modo que las costuras sean invisibles. La variación en la iluminación y las sombras puede hacer que el mismo objeto o superficie se emita en colores o sombras ligeramente diferentes. El módulo de combinación puede determinar la cantidad de combinación de colores requerida basándose en una o más posiciones de captura de imagen desde el módulo de posición de captura de imagen 1204, unión, colores de imagen a lo largo de las costuras de ambas imágenes, la deformación de imagen del módulo de recorte 1208, y/o el corte gráfico del módulo de corte gráfico 1210.

20 En diversas realizaciones, el módulo de combinación 1211 puede recibir una panorámica a partir de una combinación de dos imágenes 2D y a continuación muestrear colores a lo largo de la unión de las dos imágenes 2D. El módulo de combinación 1211 puede recibir información de ubicación de unión desde el módulo de posición de captura de imagen 1204 para posibilitar que el módulo de combinación 1211 muestree colores a lo largo de la costura y determine diferencias. Si hay una diferencia significativa en el color a lo largo de una unión entre las dos imágenes (por ejemplo, dentro de un umbral predeterminado de color, tono, brillo, saturación y/o similares), el módulo de combinación 1211 puede combinar un tamaño predeterminado de ambas imágenes a lo largo de la costura en la posición donde existe la diferencia. En algunas realizaciones, cuanto mayor sea la diferencia en color o imagen a lo largo de la costura, mayor será la cantidad de espacio a lo largo de la costura de las dos imágenes que pueden combinarse.

30 En algunas realizaciones, después de la combinación, el módulo de combinación 1211 puede volver a escanear y muestrear colores a lo largo de la costura para determinar si hay otras diferencias en la imagen o el color que superar el umbral predeterminado de color, tono, brillo, saturación y/o similares. Si es así, el módulo de combinación 1211 puede identificar las porciones a lo largo de la costura y continuar combinando esa porción de la imagen. El módulo de combinación 1211 puede continuar remuestreando las imágenes a lo largo de la costura hasta que no haya porciones adicionales de las imágenes para combinar (por ejemplo, cualquier diferencia en el color está por debajo del umbral o umbrales predeterminados).

40 El generador de imágenes 3D 1214 puede recibir imágenes panorámicas 2D y generar representaciones 3D. En diversas realizaciones, el generador de imágenes 3D 1214 utiliza un modelo de aprendizaje automático 3D para transformar las imágenes panorámicas 2D en representaciones 3D. El modelo de aprendizaje automático 3D puede entrenarse usando imágenes panorámicas 2D y datos de profundidad (por ejemplo, desde un sensor LiDAR o dispositivo de luz estructurada) para crear representaciones 3D. Las representaciones 3D pueden probarse y revisarse para su conservación y retroalimentación. En algunas realizaciones, el modelo de aprendizaje automático 3D puede usarse con imágenes panorámicas 2D y datos de profundidad para generar las representaciones 3D.

45 En diversas realizaciones, la precisión, velocidad de representación y calidad de la representación 3D generada por el generador de imágenes 3D 1214 se mejoran enormemente utilizando los sistemas y métodos descritos en el presente documento. Por ejemplo, al representar una representación 3D a partir de imágenes panorámicas 2D que se han alineado, posicionado y unido usando los métodos descritos en el presente documento (por ejemplo, mediante información de alineación y posicionamiento proporcionada por hardware, mediante posicionamiento mejorado provocado por la guía proporcionada al usuario durante la captura de imágenes, recortando y cambiando la deformación de imágenes, cortando imágenes para evitar alteraciones de imagen y superar la deformación, combinando imágenes y/o cualquier combinación), se mejoran la precisión, la velocidad de representación y la calidad de la representación 3D. Además, se apreciará que utilizando imágenes panorámicas 2D que se han alineado, posicionado y unido usando métodos descritos en el presente documento, puede mejorarse enormemente el entrenamiento del modelo de aprendizaje automático 3D (por ejemplo, en términos de velocidad y precisión). Además, en algunas realizaciones, el modelo de aprendizaje automático 3D puede ser más pequeño y menos complejo debido a la reducción de procesamiento y aprendizaje que se habría usado para superar desalineaciones, errores de posicionamiento, deformaciones, corte de gráficos deficiente, combinación deficiente, alteraciones de imagen y similares para generar representaciones 3D razonablemente precisas.

60 El modelo de aprendizaje automático 3D entrenado puede almacenarse en el sistema de unión y captura 3D y panorámico 1102, el sistema de unión y procesador de imágenes 106 y/o el sistema de usuario 1110.

65 En algunas realizaciones, el modelo de aprendizaje automático 3D puede entrenarse usando múltiples imágenes 2D y datos de profundidad desde el dispositivo de captura de imágenes del sistema de usuario 1110 y/o el sistema de unión y captura 3D y panorámico 1102. Además, el generador de imágenes 3D 1214 puede entrenarse usando

información de posición de captura de imagen asociada con cada una de las múltiples imágenes 2D del módulo de posición de captura de imagen 1204, ubicaciones de unión para alinear o unir cada una de las múltiples imágenes 2D del módulo de unión 1206, desplazamiento o desplazamientos de píxel para cada una de las múltiples imágenes 2D del módulo de recorte 1208, y/o el corte gráfico del módulo de corte gráfico 1210. En algunas realizaciones, el modelo de aprendizaje automático 3D puede usarse con imágenes panorámicas 2D, datos de profundidad, información de posición de captura de imagen asociada con cada una de las múltiples imágenes 2D del módulo de posición de captura de imagen 1204, ubicaciones de unión para alinear o unir cada una de las múltiples imágenes 2D del módulo de unión 1206, desplazamiento o desplazamientos de píxel para cada una de las múltiples imágenes 2D del módulo de recorte 1208, y/o el corte gráfico del módulo de corte gráfico 1210 para generar las representaciones 3D.

El módulo de unión 1206 puede ser una parte de un modelo 3D que convierte múltiples imágenes 2D en imágenes panorámicas 2D o panorámicas 3D. En algunas realizaciones, el modelo 3D es un algoritmo de aprendizaje automático, tal como un modelo de red neuronal de predicción 3D a partir de 2D. El módulo de recorte 1208 puede ser una parte de un modelo 3D que convierte múltiples imágenes 2D en imágenes panorámicas 2D o panorámicas 3D. En algunas realizaciones, el modelo 3D es un algoritmo de aprendizaje automático, tal como un modelo de red neuronal de predicción 3D a partir de 2D. El módulo de corte gráfico 1210 puede ser una parte de un modelo 3D que convierte múltiples imágenes 2D en imágenes panorámicas 2D o panorámicas 3D. En algunas realizaciones, el modelo 3D es un algoritmo de aprendizaje automático, tal como un modelo de red neuronal de predicción 3D a partir de 2D. El módulo de combinación 1211 puede ser una parte de un modelo de aprendizaje automático 3D que convierte múltiples imágenes 2D en imágenes panorámicas 2D o panorámicas 3D. En algunas realizaciones, el modelo 3D es un algoritmo de aprendizaje automático, tal como un modelo de red neuronal de predicción 3D a partir de 2D.

El generador de imágenes 3D 1214 puede generar una ponderación para cada uno del módulo de posición de captura de imagen 1204, el módulo de recorte 1208, el módulo de corte gráfico 1210 y el módulo de combinación 1211, que puede representar la fiabilidad o una "fortaleza" o "debilidad" del módulo. En algunas realizaciones, la suma de las ponderaciones de los módulos es igual a 1.

En casos donde los datos de profundidad no están disponibles para las múltiples imágenes 2D, el generador de imágenes 3D 1214 puede determinar datos de profundidad para uno o más objetos en las múltiples imágenes 2D capturadas por el dispositivo de captura de imágenes del sistema de usuario 1110. En algunas realizaciones, el generador de imágenes 3D 1214 puede derivar los datos de profundidad basándose en imágenes capturadas por pares de imágenes estéreo. El generador de imágenes 3D puede evaluar pares de imágenes estéreo para determinar datos acerca de la calidad de coincidencia fotométrica entre las imágenes a diversas profundidades (un resultado más intermedio), en lugar de determinar datos de profundidad a partir de un algoritmo estéreo pasivo.

El generador de imágenes 3D 1214 puede ser una parte de un modelo 3D que convierte múltiples imágenes 2D en imágenes panorámicas 2D o panorámicas 3D. En algunas realizaciones, el modelo 3D es un algoritmo de aprendizaje automático, tal como un modelo de red neuronal de predicción 3D a partir de 2D.

El almacenamiento de datos de imagen 2D capturada 1216 puede ser cualquier estructura y/o estructuras adecuadas para imágenes y/o datos de profundidad capturados (por ejemplo, una base de datos activa, una base de datos relacional, una base de datos autorreferencial, una tabla, una matriz, una disposición, un archivo plano, un sistema de almacenamiento orientado, un sistema No-SQL no relacional, un sistema de gestión de FTS tal como Lucene/Solar y/o similares). El almacenamiento de datos de imagen 2D capturada 1216 puede almacenar imágenes capturadas por el dispositivo de captura de imágenes del sistema de usuario 1110. En diversas realizaciones, el almacenamiento de datos de imagen 2D capturada 1216 almacena datos de profundidad capturados por uno o más sensores de profundidad del sistema de usuario 1110. En diversas realizaciones, el almacenamiento de datos de imagen 2D capturada 1216 almacena parámetros del dispositivo de captura de imágenes asociados con el dispositivo de captura de imágenes, o propiedades de captura asociadas con cada una de las múltiples capturas de imagen, o capturas de profundidad usadas para determinar la imagen panorámica 2D. En algunas realizaciones, el almacenamiento de datos de imagen 1108 almacena imágenes panorámicas 2D. Las imágenes panorámicas 2D pueden determinarse por el sistema de unión y captura 3D y panorámico 1102 o el sistema de procesador y unión de imágenes 106. Los parámetros del dispositivo de captura de imágenes pueden incluir iluminación, color, longitud focal de la lente de captura de imágenes, apertura máxima, ángulo de inclinación y similares. Las propiedades de captura pueden incluir resolución de píxeles, distorsión de lente, iluminación y otros metadatos de imagen.

El almacenamiento de datos de imagen panorámica 3D 1218 puede ser cualquier estructura y/o estructuras adecuadas para imágenes panorámicas 3D (por ejemplo, una base de datos activa, una base de datos relacional, una base de datos autorreferencial, una tabla, una matriz, una disposición, un archivo plano, un sistema de almacenamiento orientado, un sistema No-SQL no relacional, un sistema de gestión de FTS tal como Lucene/Solar y/o similares). El almacenamiento de datos de imagen panorámica 3D 1218 puede almacenar imágenes panorámicas 3D generadas por el sistema de captura y unión 3D y panorámico 1102. En diversas realizaciones, el almacenamiento de datos de imagen panorámica 3D 1218 almacena propiedades asociadas con el dispositivo de captura de imagen o propiedades asociadas con cada una de las múltiples capturas de imagen o capturas de profundidad usadas para determinar la imagen panorámica 3D. En algunas realizaciones, el almacenamiento de datos de imagen panorámica 3D 1218 almacena las imágenes panorámicas 3D. Las imágenes panorámicas 2D o 3D pueden determinarse por el sistema de

unión y captura 3D y panorámico 1102 o el sistema de procesador y unión de imágenes 106.

La Figura 13 representa un diagrama de flujo 1300 de un proceso de captura y generación de imágenes panorámicas 3D de acuerdo con algunas realizaciones. En la etapa 1302, el dispositivo de captura de imágenes puede capturar múltiples imágenes 2D usando el sensor de imagen 920 y la lente WFOV 918 de la Figura 9. El FOV más ancho significa que el sistema de captura 402 del entorno requerirá menos exploraciones para obtener una vista de 360°. La lente de WFOV 918 también puede ser más ancha tanto horizontal como verticalmente. En algunas realizaciones, el sensor de imagen 920 captura imágenes de RGB. En una realización, el sensor de imagen 920 captura imágenes en blanco y negro.

En la etapa 1304, el sistema de captura del entorno puede enviar las imágenes 2D capturadas al sistema de unión y procesador de imágenes 1106. El sistema de unión y procesador de imágenes 1106 puede aplicar un algoritmo de modelado 3D a las imágenes 2D capturadas para generar una imagen 2D panorámica. En algunas realizaciones, el algoritmo de modelado 3D es un algoritmo de aprendizaje automático para unir las imágenes 2D capturadas en una imagen 2D panorámica. En algunas realizaciones, la etapa 1304 puede ser opcional.

En la etapa 1306, el LiDAR 912 y la lente WFOV 918 de la Figura 9 pueden capturar datos de LiDAR. El FOV más ancho significa que el sistema de captura 400 del entorno requerirá menos exploraciones para obtener una vista de 360°.

En la etapa 1308, los datos de LiDAR pueden enviarse al sistema de unión y procesador de imágenes 1106. El sistema de unión y procesador de imágenes 1106 puede introducir los datos de LiDAR y la imagen 2D capturada en el algoritmo de modelado 3D para generar la imagen panorámica 3D. El algoritmo de modelado 3D es un algoritmo de aprendizaje automático.

En la etapa 1310, el sistema de unión y procesador de imágenes 1106 genera la imagen panorámica 3D. La imagen panorámica 3D puede almacenarse en el almacenamiento de datos de imagen 408. En una realización, la imagen panorámica 3D generada por el algoritmo de modelado 3D se almacena en el sistema de unión y procesador de imágenes 1106. En algunas realizaciones, el algoritmo de modelado 3D puede generar una representación visual del plano de planta del entorno físico a medida que se utiliza el sistema de captura del entorno para capturar diversas partes del entorno físico.

En la etapa 1312, el sistema de unión y procesador de imágenes 1106 puede proporcionar al menos una porción de la imagen panorámica 3D generada al sistema de usuario 1110. El sistema de unión y procesador de imágenes 1106 puede proporcionar la representación visual del plano de planta del entorno físico.

El orden de una o más etapas del diagrama de flujo 1300 puede cambiarse sin afectar al producto final de la imagen panorámica 3D. Por ejemplo, el sistema de captura del entorno puede intercalar la captura de imagen con el dispositivo de captura de imágenes con datos de LiDAR o captura de información de profundidad con el LiDAR 912. Por ejemplo, el dispositivo de captura de imágenes puede capturar una imagen de la sección del entorno físico con el dispositivo de captura de imágenes, y luego el LiDAR 912 obtiene información de profundidad de la sección 1605. Una vez que el LiDAR 912 obtiene información de profundidad de la sección, el dispositivo de captura de imágenes puede moverse para capturar una imagen de otra sección, y a continuación el LiDAR 912 obtiene información de profundidad de la sección, intercalando de ese modo la captura de imagen y la captura de información de profundidad.

En algunas realizaciones, los dispositivos y/o sistemas analizados en el presente documento emplean un dispositivo de captura de imágenes para capturar imágenes de entrada 2D. En algunas realizaciones, el uno o más dispositivos de captura de imágenes 1116 pueden representar un único dispositivo de captura de imágenes (o lente de captura de imágenes). De acuerdo con algunas de estas realizaciones, el usuario del dispositivo desplazable que aloja el dispositivo de captura de imágenes puede configurarse para rotar alrededor de un eje para generar imágenes en diferentes orientaciones de captura con respecto al entorno, en donde los campos de visión colectivos de las imágenes abarcan hasta 360° horizontalmente.

En diversas realizaciones, los dispositivos y/o sistemas analizados en el presente documento pueden emplear dos o más dispositivos de captura de imágenes para capturar imágenes de entrada 2D. En algunas realizaciones, los dos o más dispositivos de captura de imágenes pueden disponerse en posiciones relativas entre sí en o dentro de la misma carcasa desplazable de tal manera que sus campos de visión colectivos abarcan hasta 360°. En algunas realizaciones, pueden usarse pares de dispositivos de captura de imágenes capaces de generar pares de imágenes estéreo (por ejemplo, con campos de visión ligeramente desplazados pero parcialmente superpuestos). Por ejemplo, el sistema de usuario 1110 (por ejemplo, el dispositivo que comprende el uno o más dispositivos de captura de imágenes usados para capturar las imágenes de entrada 2D) puede comprender dos dispositivos de captura de imágenes con campos de visión de desplazamiento estéreo horizontales capaces de capturar pares de imágenes estéreo. En otro ejemplo, el sistema de usuario 1110 puede comprender dos dispositivos de captura de imágenes con campos de visión de desplazamiento estéreo vertical capaces de capturar pares de imágenes estéreo verticales. De acuerdo con cualquiera de estos ejemplos, cada una de las cámaras puede tener campos de visión que abarcan hasta 360. En este sentido, en una realización, el sistema de usuario 1110 puede emplear dos cámaras panorámicas con desplazamientos estéreo

verticales capaces de capturar pares de imágenes panorámicas que forman pares estéreo (con desplazamientos estéreo verticales).

5 El componente de posicionamiento 1118 puede incluir cualquier hardware y/o software configurado para capturar datos de posición de sistema de usuario y/o datos de ubicación de sistema de usuario. Por ejemplo, el componente de posicionamiento 1118 incluye una IMU para generar los datos de posición del sistema de usuario 1110 en asociación con el uno o más dispositivos de captura de imágenes del sistema de usuario 1110 usados para capturar las múltiples imágenes 2D. El componente de posicionamiento 1118 puede incluir una unidad de GPS para proporcionar información de coordenadas de GPS en asociación con las múltiples imágenes 2D capturadas por uno o más dispositivos de captura de imágenes. En algunas realizaciones, el componente de posicionamiento 1118 puede correlacionar datos de posición y datos de ubicación del sistema de usuario con respectivas imágenes capturadas usando el uno o más dispositivos de captura de imágenes del sistema de usuario 1110.

15 Diversas realizaciones del aparato proporcionan a los usuarios imágenes panorámicas 3D de entornos interiores así como exteriores. En algunas realizaciones, el aparato puede proporcionar a los usuarios de manera eficiente y rápida imágenes panorámicas 3D de entornos interiores y exteriores usando una única lente de campo de visión amplio (FOV) y una única luz y sensores de detección y alcance (sensor LiDAR).

20 Lo siguiente es un caso de uso de ejemplo de un aparato de ejemplo descrito en el presente documento. El siguiente caso de uso es de una de las realizaciones. Diferentes realizaciones del aparato, como se analiza en el presente documento, pueden incluir una o más características y capacidades similares a las del caso de uso.

25 La Figura 14 representa un diagrama de flujo de un proceso de captura y unión 3D y panorámico 1400 de acuerdo con algunas realizaciones. El diagrama de flujo de la Figura 14 se refiere al sistema de unión y captura 3D y panorámico 1102 como que incluye el dispositivo de captura de imágenes, pero, en algunas realizaciones, el dispositivo de captura de datos puede ser el sistema de usuario 1110.

30 En la etapa 1402, el sistema de unión y captura 3D y panorámico 1102 puede recibir múltiples imágenes 2D desde al menos un dispositivo de captura de imágenes. El dispositivo de captura de imágenes del sistema de unión y captura 3D y panorámico 1102 puede ser o incluir un sensor de imagen de semiconductor de óxido de metal (CMOS) complementario. En diversas realizaciones, el dispositivo de captura de imágenes es un dispositivo acoplado cargado (CCD). En un ejemplo, el dispositivo de captura de imágenes es un sensor rojo-verde-azul (RGB). En una realización, el dispositivo de captura de imágenes es un sensor IR. Cada una de las múltiples imágenes 2D puede tener campos de visión parcialmente superpuestos con al menos otra imagen de las múltiples imágenes 2D. En algunas realizaciones, al menos algunas de las múltiples imágenes 2D se combinan para crear una vista de 360° del entorno físico (por ejemplo, interior, exterior o ambos).

40 En algunas realizaciones, todas las múltiples imágenes 2D se reciben desde el mismo dispositivo de captura de imágenes. En diversas realizaciones, al menos una porción de las múltiples imágenes 2D se recibe desde dos o más dispositivos de captura de imágenes del sistema de unión y captura 3D y panorámico 1102. En un ejemplo, las múltiples imágenes 2D incluyen un conjunto de imágenes RGB y un conjunto de imágenes IR, donde las imágenes IR proporcionan datos de profundidad al sistema 1102 de captura y unión 3D y panorámico. En algunas realizaciones, cada imagen 2D puede asociarse con datos de profundidad proporcionados desde un dispositivo LiDAR. Cada una de las imágenes 2D puede, en algunas realizaciones, estar asociada con datos de posicionamiento.

45 En la etapa 1404, el sistema de unión y captura 3D y panorámico 1102 puede recibir parámetros de captura y parámetros de dispositivo de captura de imágenes asociados con cada una de las múltiples imágenes 2D recibidas. Los parámetros del dispositivo de captura de imágenes pueden incluir iluminación, color, longitud focal de la lente de captura de imágenes, apertura máxima, un campo de visión y similares. Las propiedades de captura pueden incluir resolución de píxeles, distorsión de lente, iluminación y otros metadatos de imagen. El sistema de unión y captura 3D y panorámico 1102 también puede recibir los datos de posicionamiento y los datos de profundidad.

50 En la etapa 1406, el sistema de captura y unión 3D y panorámico 1102 puede tomar la información recibida de las etapas 1402 y 1404 para unir las imágenes 2D para formar una imagen panorámica 2D. El proceso de unión de las imágenes 2D se analiza adicionalmente con respecto al diagrama de flujo de la Figura 15.

60 En la etapa 1408, el sistema de unión y captura 3D y panorámico 1102 puede aplicar un modelo de aprendizaje automático 3D para generar una representación 3D. La representación 3D puede almacenarse en un almacenamiento de datos de imagen panorámica 3D. En diversas realizaciones, la representación 3D es generada por el sistema de unión y procesador de imágenes 1106. En algunas realizaciones, el modelo de aprendizaje automático 3D puede generar una representación visual del plano de planta del entorno físico a medida que el sistema de captura del entorno se utiliza para capturar diversas partes del entorno físico.

65 En la etapa 1410, el sistema de unión y captura 3D y panorámico 1102 puede proporcionar al menos una porción de la representación o modelo 3D generado al sistema de usuario 1110. El sistema de usuario 1110 puede proporcionar la representación visual del plano de planta del entorno físico.

En algunas realizaciones, el sistema de usuario 1110 puede enviar las múltiples imágenes 2D, parámetros de captura y parámetros de captura de imagen al sistema de unión y procesador de imágenes 1106. En diversas realizaciones, el sistema de captura y unión 3D y panorámico 1102 puede enviar las múltiples imágenes 2D, parámetros de captura y parámetros de captura de imagen al sistema de unión y procesador de imágenes 1106.

El sistema de unión y procesador de imágenes 1106 puede procesar las múltiples imágenes 2D capturadas por el dispositivo de captura de imágenes del sistema de usuario 1110 y unir las imágenes en una imagen panorámica 2D. La imagen panorámica 2D procesada por el sistema de unión y procesador de imágenes 1106 puede tener una resolución de píxel más alta que la imagen panorámica 2D obtenida por el sistema de captura y unión 3D y panorámico 1102.

En algunas realizaciones, el sistema de unión y procesador de imágenes 106 puede recibir la representación panorámica 3D y emitir una imagen panorámica 3D con una resolución de píxel que es mayor que la de la imagen panorámica 3D recibida. Las imágenes panorámicas de resolución de píxeles más alta pueden proporcionarse a un dispositivo de salida con una resolución de pantalla más alta que el sistema de usuario 1110, tal como una pantalla de ordenador, pantalla de proyector y similares. En algunas realizaciones, las imágenes panorámicas de mayor resolución de píxeles pueden proporcionar al dispositivo de salida una imagen panorámica con mayor detalle y pueden ampliarse.

La Figura 15 representa un diagrama de flujo que muestra detalles adicionales de una etapa del proceso de captura y unión 3D y panorámico de la Figura 14. En la etapa 1502, el módulo de posición de captura de imagen 1204 puede determinar datos de posición de dispositivo de captura de imágenes asociados con cada imagen capturada por el dispositivo de captura de imágenes. El módulo de posición de captura de imagen 1204 puede utilizar la IMU del sistema de usuario 1110 para determinar los datos de posición del dispositivo de captura de imágenes (o el campo de visión de la lente del dispositivo de captura de imágenes). Los datos de posición pueden incluir la dirección, ángulo o inclinación de uno o más dispositivos de captura de imágenes cuando se toman una o más imágenes 2D. Uno o más del módulo de recorte 1208, el módulo de corte gráfico 1210 o el módulo de combinación 1212 pueden utilizar la dirección, el ángulo o la inclinación asociados con cada una de las múltiples imágenes 2D para determinar cómo deformar, cortar y/o combinar las imágenes.

En la etapa 1504, el módulo de recorte 1208 puede deformar una o más de las múltiples imágenes 2D para que dos imágenes puedan alinearse juntas para formar una imagen panorámica y, al mismo tiempo, conservar características específicas de las imágenes, tales como mantener la rectitud de una línea recta. La salida del módulo de recorte 1208 puede incluir el número de columnas y filas de píxeles para desplazar cada píxel de la imagen para enderezar la imagen. La cantidad de desplazamiento para cada imagen puede emitirse en forma de una matriz que representa el número de columnas de píxeles y filas de píxeles para desplazar cada píxel de la imagen. En esta realización, el módulo de recorte 1208 puede determinar la cantidad de deformación que requiere cada una de las múltiples imágenes 2D basándose en la estimación de postura de captura de imagen de cada una de las múltiples imágenes 2D.

En la etapa 1506, el módulo de corte gráfico 1210 determina dónde cortar o segmentar una o más de las múltiples imágenes 2D. En esta realización, el módulo de corte gráfico 1210 puede determinar dónde cortar o segmentar cada una de las múltiples imágenes 2D basándose en la estimación de postura de captura de imagen y la deformación de imagen de cada una de las múltiples imágenes 2D.

En la etapa 1508, el módulo de unión 1206 puede unir dos o más imágenes juntas usando los bordes de las imágenes y/o los cortes de las imágenes. El módulo de unión 1206 puede alinear y/o posicionar imágenes basándose en objetos detectados dentro de las imágenes, deformación, corte de la imagen y/o similares.

En la etapa 1510, el módulo de combinación 1212 puede ajustar el color en las uniones (por ejemplo, unión de dos imágenes) o la ubicación en una imagen que toca o se conecta a otra imagen. El módulo de combinación 1212 puede determinar la cantidad de combinación de colores requerida basándose en una o más posiciones de captura de imagen desde el módulo de posición de captura de imagen 1204, la deformación de imagen del módulo de recorte 1208, y el corte gráfico del módulo de corte gráfico 1210.

El orden de una o más etapas del proceso de captura y unión 3D y panorámico 1400 puede cambiarse sin afectar al producto final de la imagen panorámica 3D. Por ejemplo, el sistema de captura del entorno puede intercalar la captura de imagen con el dispositivo de captura de imágenes con datos de LiDAR o captura de información de profundidad. Por ejemplo, el dispositivo de captura de imágenes puede capturar una imagen de una sección 1605 de la Figura 16 del entorno físico con el dispositivo de captura de imágenes, y luego el LiDAR 612 obtiene información de profundidad de la sección 1605. Una vez que el LiDAR obtiene información de profundidad de la sección 1605, el dispositivo de captura de imágenes puede moverse para capturar una imagen de otra sección 1610, y luego

El LiDAR 612 obtiene información de profundidad de la sección 1610, intercalando de este modo la captura de imagen y la captura de información de profundidad.

La Figura 16 representa un diagrama de bloques de un dispositivo digital 1602 de ejemplo de acuerdo con algunas

- realizaciones. Cualquiera del sistema de usuario 1110, el sistema de unión y captura panorámica 3D 1102, y el sistema de unión y procesador de imágenes puede comprender una instancia del dispositivo digital 1602. El dispositivo digital 1602 comprende un procesador 1604, una memoria 1606, un almacenamiento 1608, un dispositivo de entrada 1610, una interfaz de red de comunicación 1612, un dispositivo de salida 1614, un dispositivo de captura de imágenes 1616 y un componente de posicionamiento 1618. El procesador 1604 está configurado para ejecutar instrucciones ejecutables (por ejemplo, programas). En algunas realizaciones, el procesador 1604 comprende circuitería o cualquier procesador capaz de procesar las instrucciones ejecutables.
- La memoria 1606 almacena datos. Algunos ejemplos de memoria 1606 incluyen dispositivos de almacenamiento, tales como RAM, ROM, caché de RAM, memoria virtual, etc. En diversas realizaciones, los datos de trabajo se almacenan dentro de la memoria 1606. Los datos dentro de la memoria 1606 pueden borrarse o transferirse finalmente al almacenamiento 1608.
- El almacenamiento 1608 incluye cualquier almacenamiento configurado para recuperar y almacenar datos. Algunos ejemplos de almacenamiento 1608 incluyen unidades flash, unidades de disco duro, unidades ópticas y/o cinta magnética. Cada uno de la memoria 1606 y el almacenamiento 1608 comprende un medio legible por ordenador, que almacena instrucciones o programas ejecutables por el procesador 1604.
- El dispositivo de entrada 1610 es cualquier dispositivo que introduce datos (por ejemplo, teclado táctil, lápiz óptico). El dispositivo de salida 1614 emite datos (por ejemplo, altavoz, pantalla, casco de realidad virtual). Se apreciará que el almacenamiento 1608, el dispositivo de entrada 1610 y un dispositivo de salida 1614. En algunas realizaciones, el dispositivo de salida 1614 es opcional. Por ejemplo, los enrutadores/conmutadores pueden comprender el procesador 1604 y la memoria 1606 así como un dispositivo para recibir y emitir datos (por ejemplo, una interfaz de red de comunicación 1612 y/o dispositivo de salida 1614).
- La interfaz de red de comunicación 1612 puede acoplarse a una red (por ejemplo, la red de comunicación 104) a través de la interfaz de red de comunicación 1612. La interfaz de red de comunicación 1612 puede soportar comunicación a través de una conexión de Ethernet, una conexión en serie, una conexión en paralelo y/o una conexión de ATA. La interfaz de red de comunicación 1612 también puede soportar comunicación inalámbrica (por ejemplo, 802.16 a/b/g/n, WIMAX, LTE, Wi-Fi). Será evidente que la interfaz de red de comunicación 1612 puede soportar muchas normas alámbricas e inalámbricas.
- Un componente puede ser hardware o software. En algunas realizaciones, el componente puede configurar uno o más procesadores para realizar funciones asociadas con el componente. Aunque se analizan diferentes componentes en el presente documento, se apreciará que el sistema de servidor puede incluir cualquier número de componentes que realizan cualquiera o todas las funcionalidades analizadas en el presente documento.
- El dispositivo digital 1602 puede incluir uno o más dispositivos de captura de imágenes 1616. El uno o más dispositivos de captura de imágenes 1616 pueden incluir, por ejemplo, cámaras RGB, cámaras HDR, cámaras de vídeo y similares. El uno o más dispositivos de captura de imágenes 1616 también pueden incluir una cámara de vídeo capaz de capturar vídeo de acuerdo con algunas realizaciones. En algunas realizaciones, uno o más dispositivos de captura de imágenes 1616 pueden incluir un dispositivo de captura de imágenes que proporciona un campo de visión relativamente estándar (por ejemplo, alrededor de 75°). En otras realizaciones, el uno o más dispositivos de captura de imágenes 1616 pueden incluir cámaras que proporcionan un campo de visión relativamente amplio (por ejemplo, desde aproximadamente 120° hasta 360°), tal como una cámara de ojo de pez, y similares (por ejemplo, el dispositivo digital 1602 puede incluir o estar incluido en el sistema de captura 400 del entorno).
- Un componente puede ser hardware o software. En algunas realizaciones, el componente puede configurar uno o más procesadores para realizar funciones asociadas con el componente. Aunque se analizan diferentes componentes en el presente documento, se apreciará que el sistema de servidor puede incluir cualquier número de componentes que realizan cualquiera o todas las funcionalidades analizadas en el presente documento.

REIVINDICACIONES

1. Un dispositivo de captura de imágenes (300, 700) que comprende:

5 una carcasa (320, 710, 716), teniendo la carcasa un lado delantero (410) y un lado trasero (440);
 un primer motor (726) acoplado a la carcasa (320, 710, 716) en una primera posición entre el lado delantero (410)
 y el lado trasero (440) de la carcasa, estando el primer motor (726) configurado para girar horizontalmente el
 dispositivo de captura de imágenes (300, 700) sustancialmente 270 grados alrededor de un eje vertical (305);
 10 una lente gran angular (310, 704) dentro de la carcasa (320, 710, 716) del dispositivo de captura de imágenes en
 una segunda posición entre el lado delantero (410) y el lado trasero (440) de la carcasa a lo largo del eje vertical
 (305), siendo la segunda posición un punto sin paralaje, teniendo la lente gran angular (310, 704) un campo de
 visión alejado del lado delantero (410) de la carcasa;
 un sensor de imagen acoplado a la carcasa (320, 710, 716) del dispositivo de captura de imágenes (300, 700) y
 configurado para generar señales de imagen a partir de la luz recibida por la lente gran angular (310, 704);
 15 una montura (330) acoplada con el primer motor (726);
 un LiDAR (708) acoplado a la carcasa (320, 710, 716) del dispositivo de captura de imágenes (300, 700) en una
 tercera posición, el LiDAR (708) configurado para generar pulsos láser y generar señales de profundidad;
 un segundo motor (718) acoplado a la carcasa (320, 710, 716) del dispositivo de captura de imágenes; y
 un espejo (712) acoplado al segundo motor (718), el segundo motor (718) configurado para rotar el espejo (712)
 20 alrededor de un eje horizontal, incluyendo el espejo (712) una superficie en ángulo configurada para recibir los
 pulsos láser del LiDAR y dirigir los pulsos láser alrededor del eje horizontal.

2. El dispositivo de captura de imágenes de la reivindicación 1, en donde el sensor de imagen está configurado para
 25 generar una primera pluralidad de imágenes a diferentes exposiciones cuando el dispositivo de captura de imágenes
 está estacionario y apuntado en una primera dirección.

3. El dispositivo de captura de imágenes de la reivindicación 2, en donde el primer motor está configurado para girar
 el dispositivo de captura de imágenes alrededor del eje vertical después de que se genera la primera pluralidad de
 30 imágenes.

4. El dispositivo de captura de imágenes de la reivindicación 3, en donde el sensor de imagen no genera imágenes
 mientras el primer motor gira el dispositivo de captura de imágenes y en donde el LiDAR genera señales de profundidad
 basándose en los pulsos láser mientras el primer motor gira el dispositivo de captura de imágenes.

5. El dispositivo de captura de imágenes de la reivindicación 3, en donde el sensor de imagen está configurado para
 35 generar una segunda pluralidad de imágenes en las diferentes exposiciones cuando el dispositivo de captura de
 imágenes está estacionario y apunta en una segunda dirección y el primer motor está configurado para girar el
 dispositivo de captura de imágenes 90 grados alrededor del eje vertical después de que se genera una segunda
 pluralidad de imágenes.

6. El dispositivo de captura de imágenes de la reivindicación 5, en donde el sensor de imagen está configurado para
 40 generar una tercera pluralidad de imágenes en las diferentes exposiciones cuando el dispositivo de captura de
 imágenes está estacionario y apunta en una tercera dirección y el primer motor está configurado para girar el
 dispositivo de captura de imágenes 90 grados alrededor del eje vertical después de que se genera una tercera
 45 pluralidad de imágenes.

7. El dispositivo de captura de imágenes de la reivindicación 6, en donde el sensor de imagen está configurado para
 generar una cuarta pluralidad de imágenes en las diferentes exposiciones cuando el dispositivo de captura de
 imágenes está estacionario y apunta en una cuarta dirección y el primer motor está configurado para girar el dispositivo
 50 de captura de imágenes 90 grados alrededor del eje vertical después de que se genera una cuarta pluralidad de
 imágenes.

8. El dispositivo de captura de imágenes de la reivindicación 7, que comprende además un dispositivo digital remoto
 en comunicación con el dispositivo de captura de imágenes y configurado para generar una visualización 3D
 55 basándose en la primera, segunda, tercera y cuarta pluralidad de imágenes y las señales de profundidad, estando el
 dispositivo digital remoto configurado para generar la visualización 3D usando no más imágenes que la primera,
 segunda, tercera y cuarta pluralidad de imágenes.

9. Un método que comprende:
 60 recibir luz desde una lente gran angular de un dispositivo de captura de imágenes, estando la lente gran angular
 dentro de una carcasa del dispositivo de captura de imágenes, recibéndose la luz en un campo de visión de la
 lente gran angular, extendiéndose el campo de visión lejos de un lado delantero de la carcasa;
 generar una primera pluralidad de imágenes mediante un sensor de imagen de un dispositivo de captura de
 65 imágenes usando la luz de la lente gran angular, estando el sensor de imagen acoplado a la carcasa del dispositivo
 de captura de imágenes, estando la primera pluralidad de imágenes a diferentes exposiciones;

- girar horizontalmente el dispositivo de captura de imágenes mediante un primer motor sustancialmente 270 grados alrededor de un eje vertical, estando el primer motor acoplado a la carcasa en una primera posición entre el lado delantero y un lado trasero de la carcasa, estando la lente gran angular en una segunda posición a lo largo del eje vertical, siendo la segunda posición un punto sin paralaje;
- 5 rotar un espejo con una superficie en ángulo alrededor del eje horizontal mediante un segundo motor, estando el segundo motor acoplado a la carcasa del dispositivo de captura de imágenes;
- generar pulsos láser mediante un LiDAR, estando el LiDAR acoplado a la carcasa del dispositivo de captura de imágenes en una tercera posición, dirigiéndose el pulso láser al espejo giratorio mientras el dispositivo de captura de imágenes gira horizontalmente; y
- 10 generar señales de profundidad por el LiDAR basándose en los pulsos láser.
10. El método de la reivindicación 9, en donde la generación de la primera pluralidad de imágenes por el sensor de imagen tiene lugar antes de que el dispositivo de captura de imágenes gire horizontalmente.
- 15 11. El método de la reivindicación 10, en donde el sensor de imagen no genera imágenes mientras el primer motor gira el dispositivo de captura de imágenes y en donde el LiDAR genera las señales de profundidad basándose en los pulsos láser mientras el primer motor gira el dispositivo de captura de imágenes.
- 20 12. El método de la reivindicación 11, que comprende además:
- generar una segunda pluralidad de imágenes a las diferentes exposiciones por el sensor de imagen cuando el dispositivo de captura de imágenes está estacionario y apuntado en una segunda dirección; y
- girar el dispositivo de captura de imágenes 90 grados alrededor del eje vertical mediante el primer motor después de generar la segunda pluralidad de imágenes.
- 25 13. El método de la reivindicación 12, que comprende además:
- generar una tercera pluralidad de imágenes a las diferentes exposiciones por el sensor de imagen cuando el dispositivo de captura de imágenes está estacionario y apuntado en una tercera dirección; y
- 30 girar el dispositivo de captura de imágenes 90 grados alrededor del eje vertical mediante el primer motor después de generar la tercera pluralidad de imágenes.
14. El método de la reivindicación 13, que comprende además:
- 35 generar una cuarta pluralidad de imágenes a las diferentes exposiciones por el sensor de imagen cuando el dispositivo de captura de imágenes está estacionario y apuntado en una cuarta dirección.
15. El método de la reivindicación 14, que comprende además generar una visualización 3D usando la primera, segunda, tercera y cuarta pluralidad de imágenes y basándose en las señales de profundidad, la generación de la visualización 3D sin usar ninguna otra imagen.
- 40

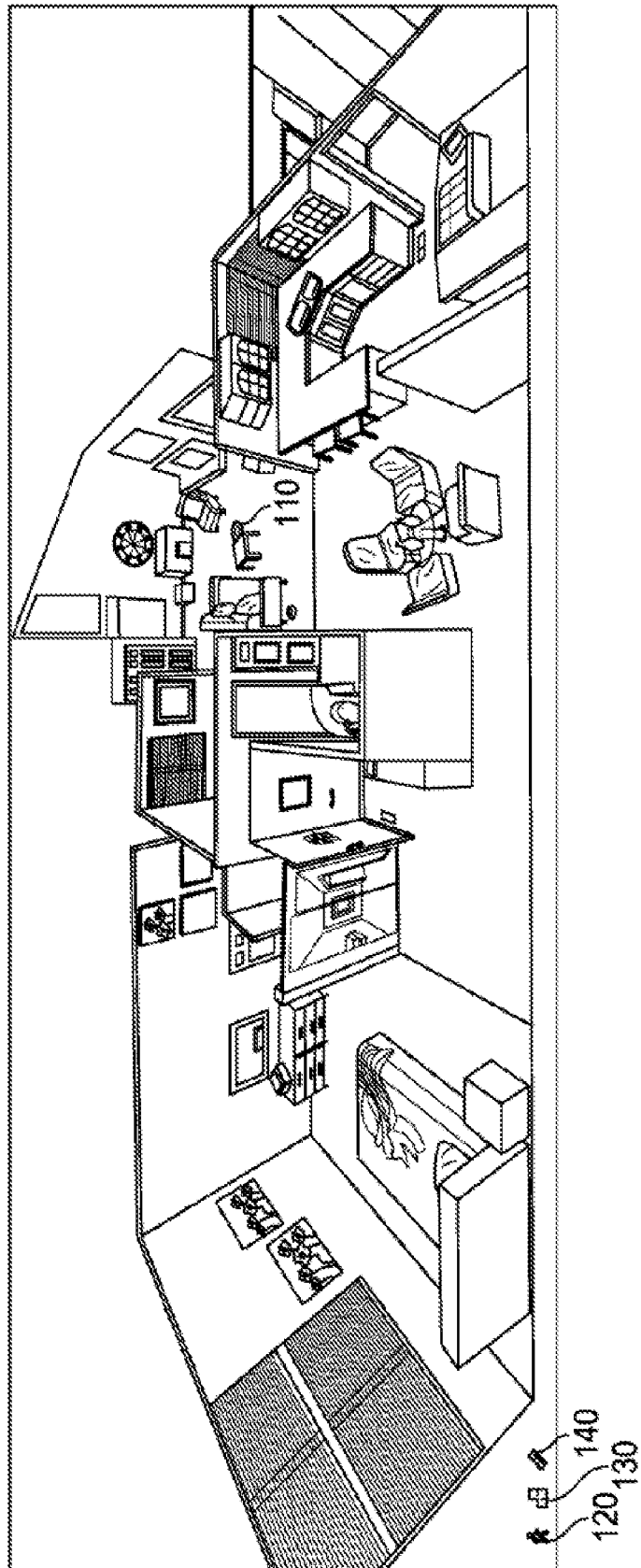


FIG. 1A

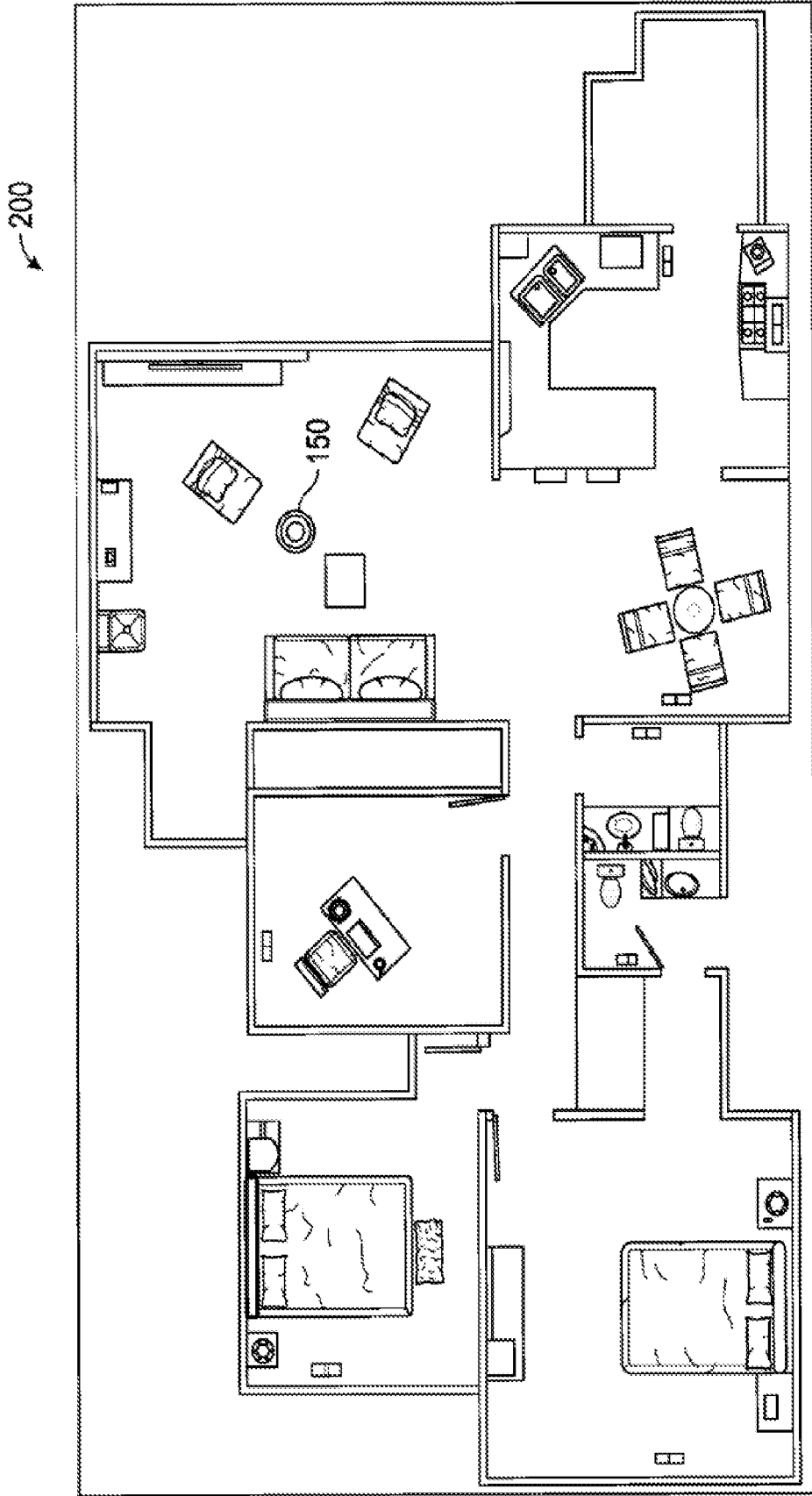


FIG. 1B

200

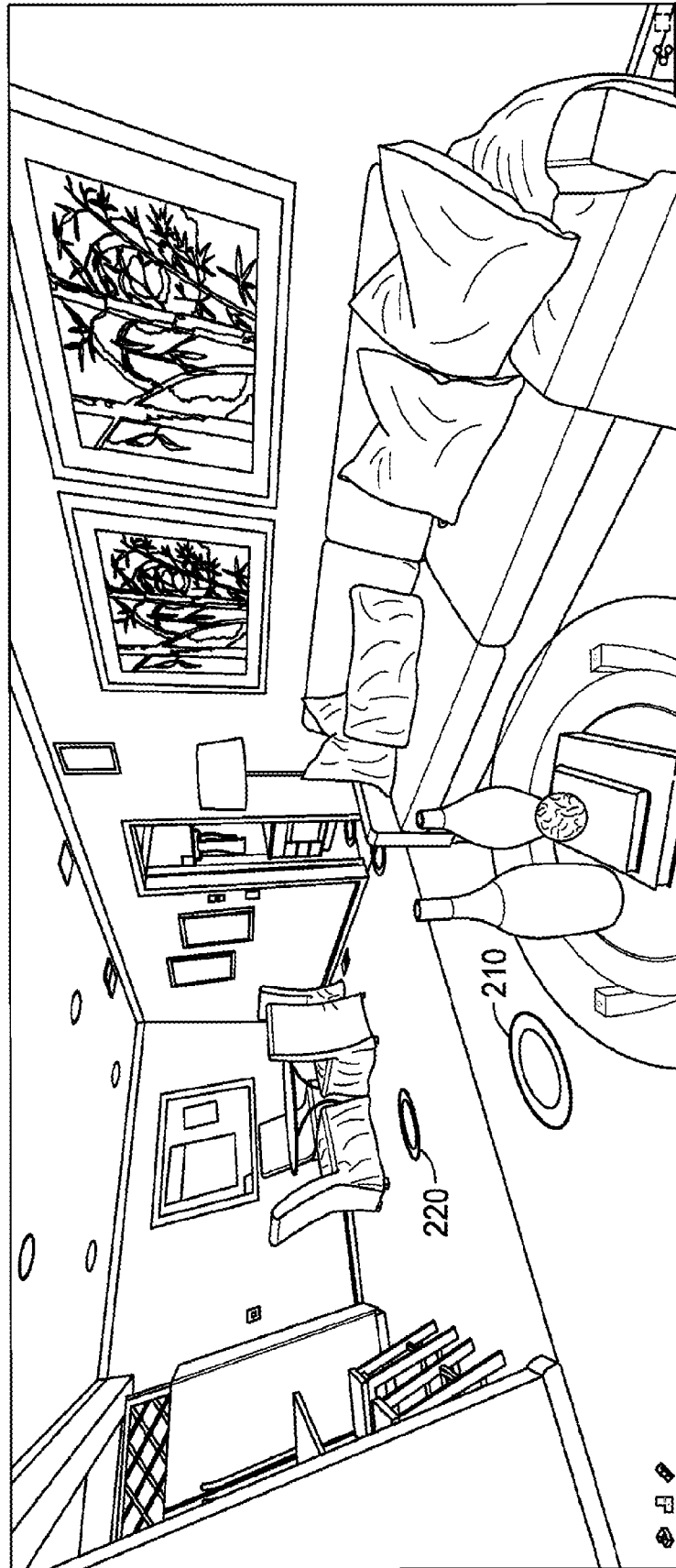


FIG. 2

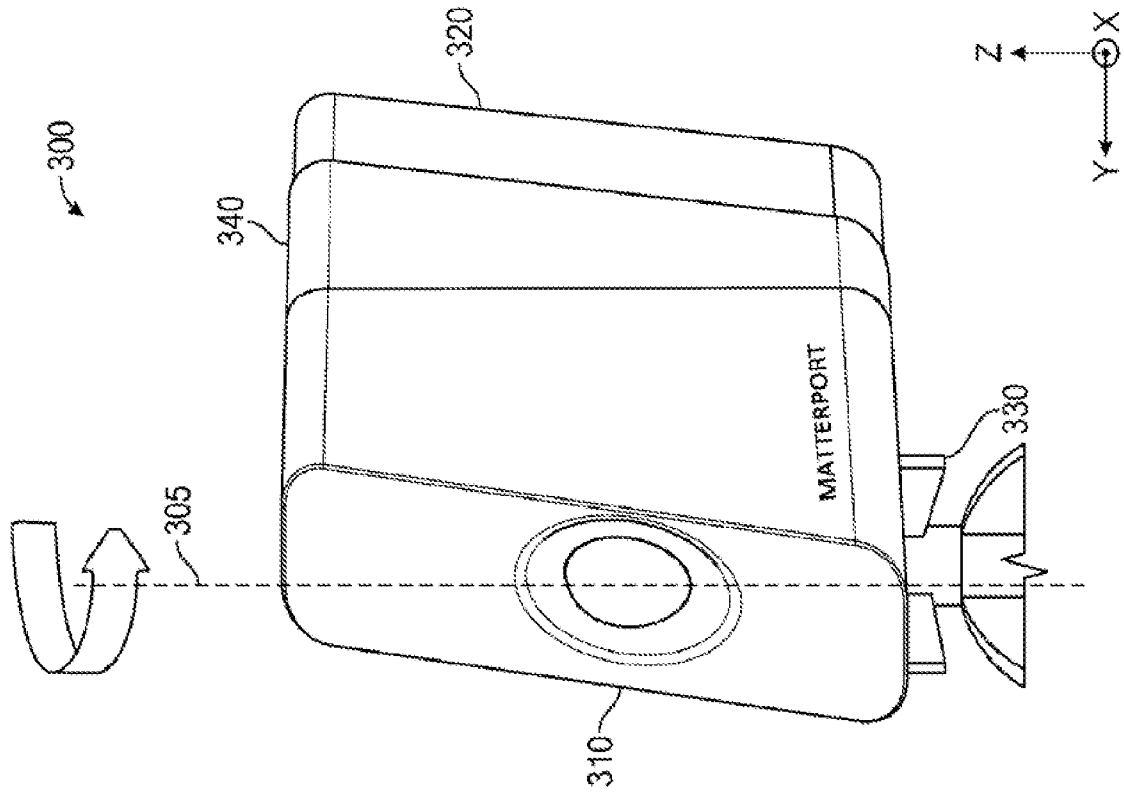


FIG. 3

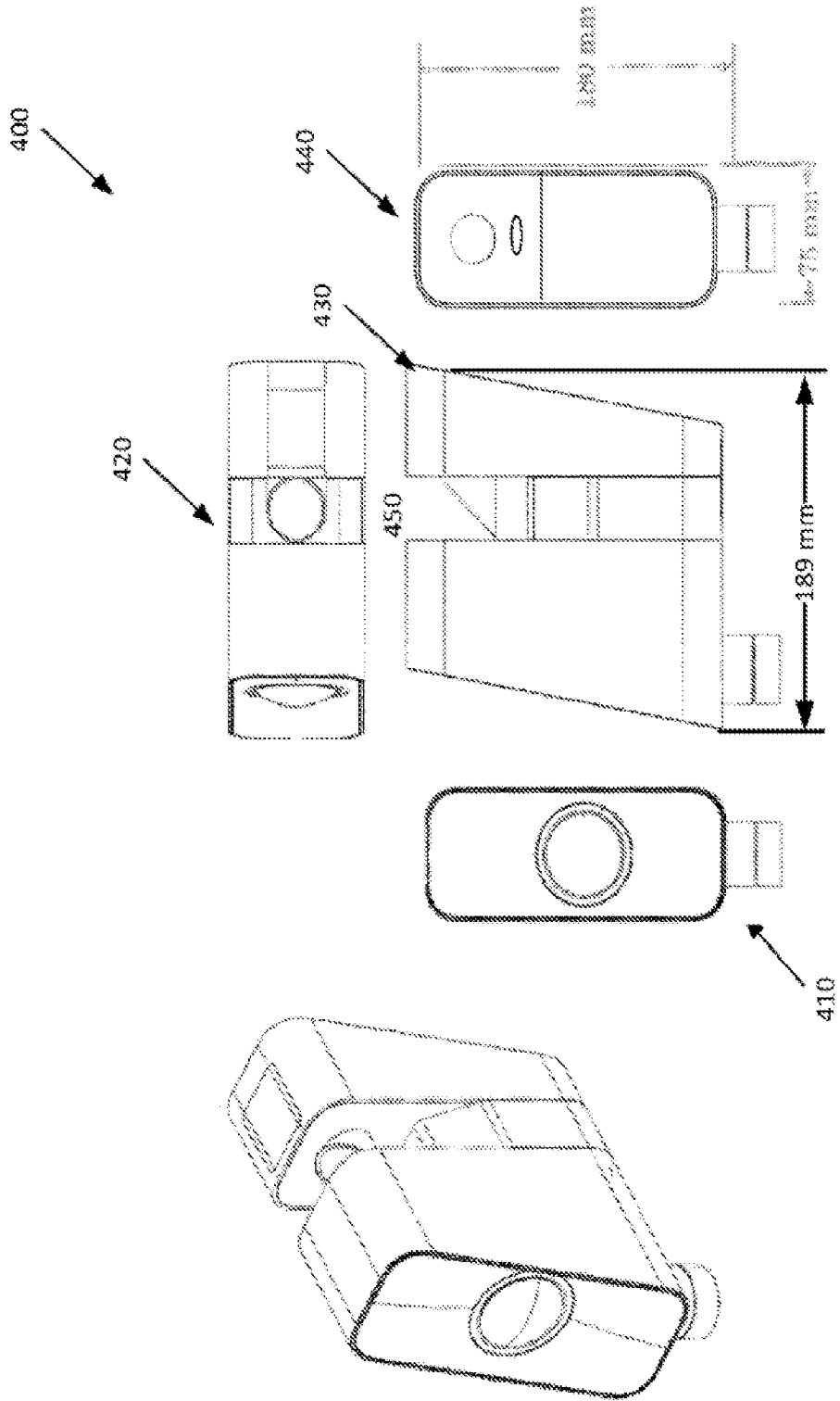


FIG. 4

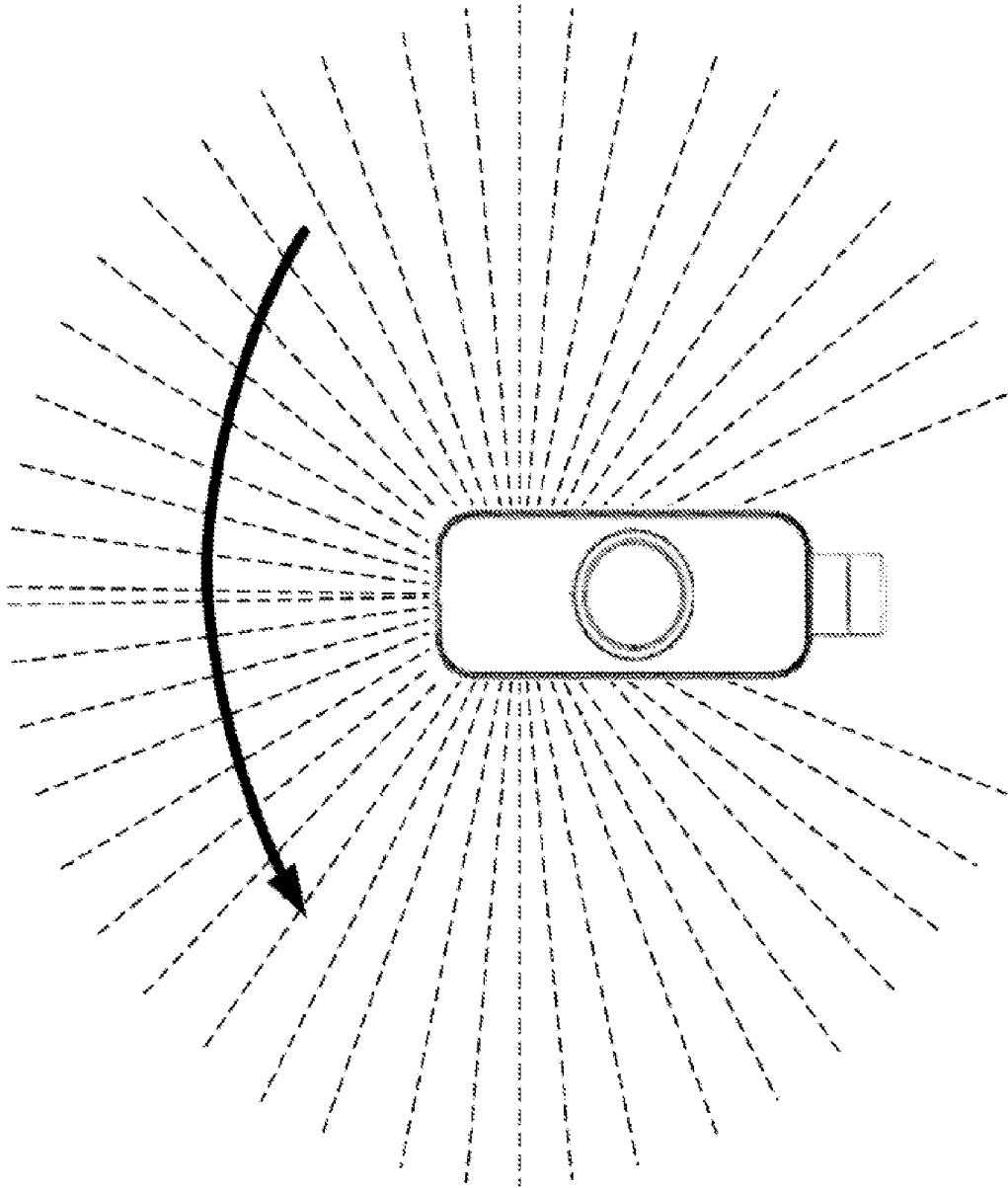


FIG. 5

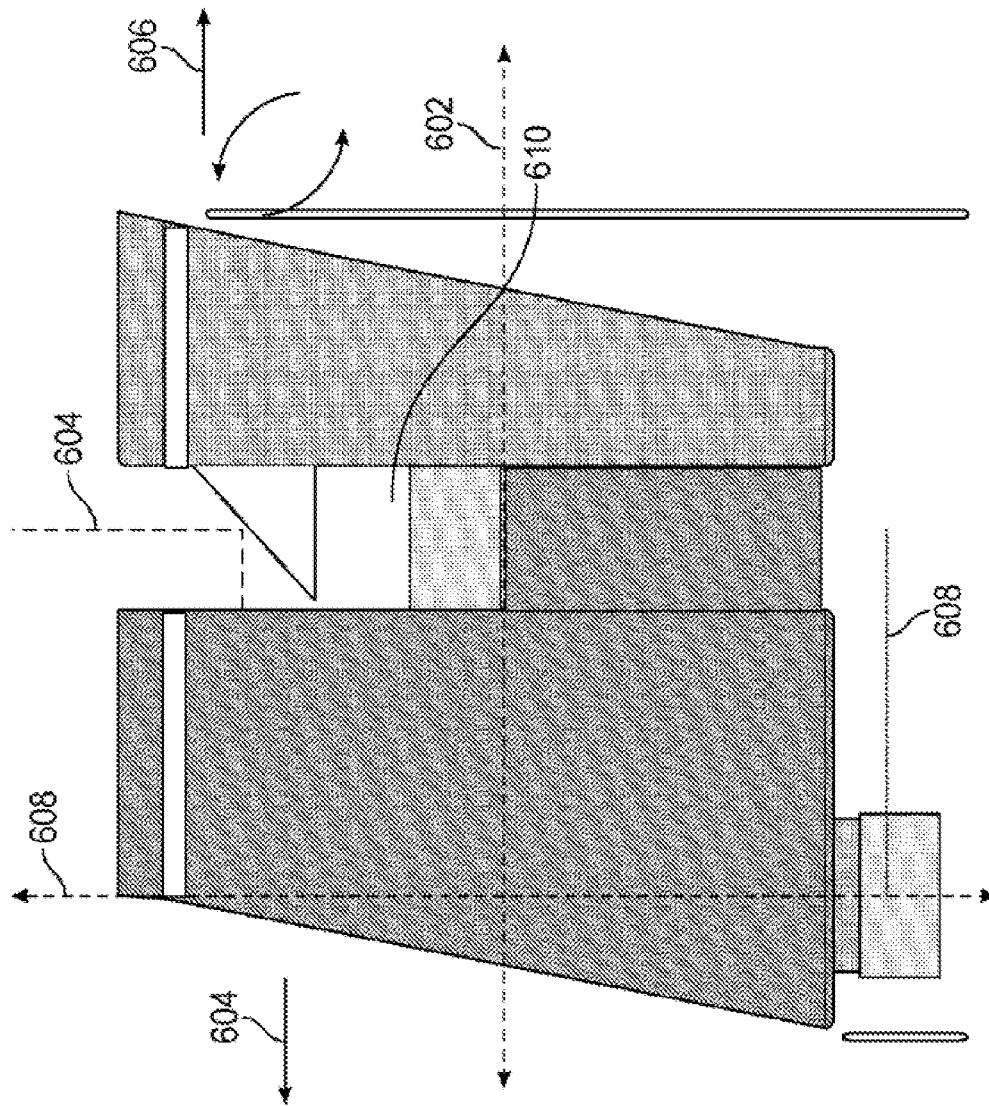


FIG. 6A

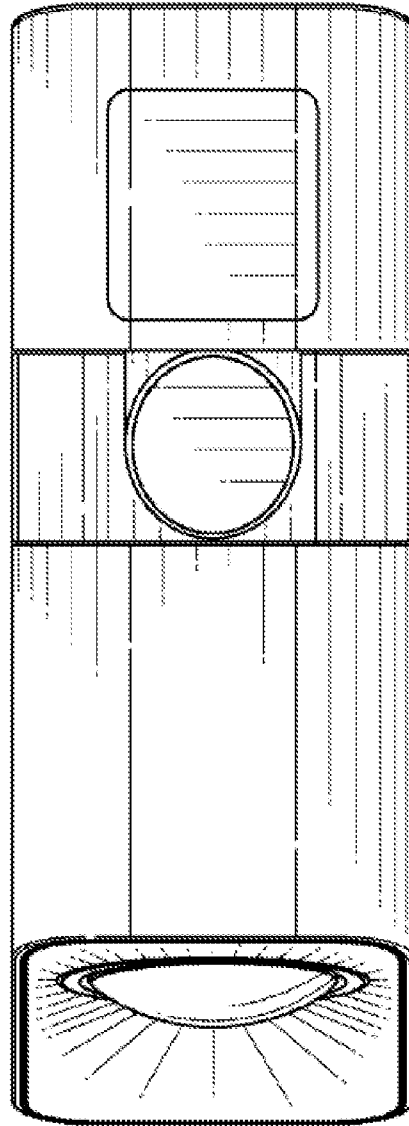


FIG. 6B

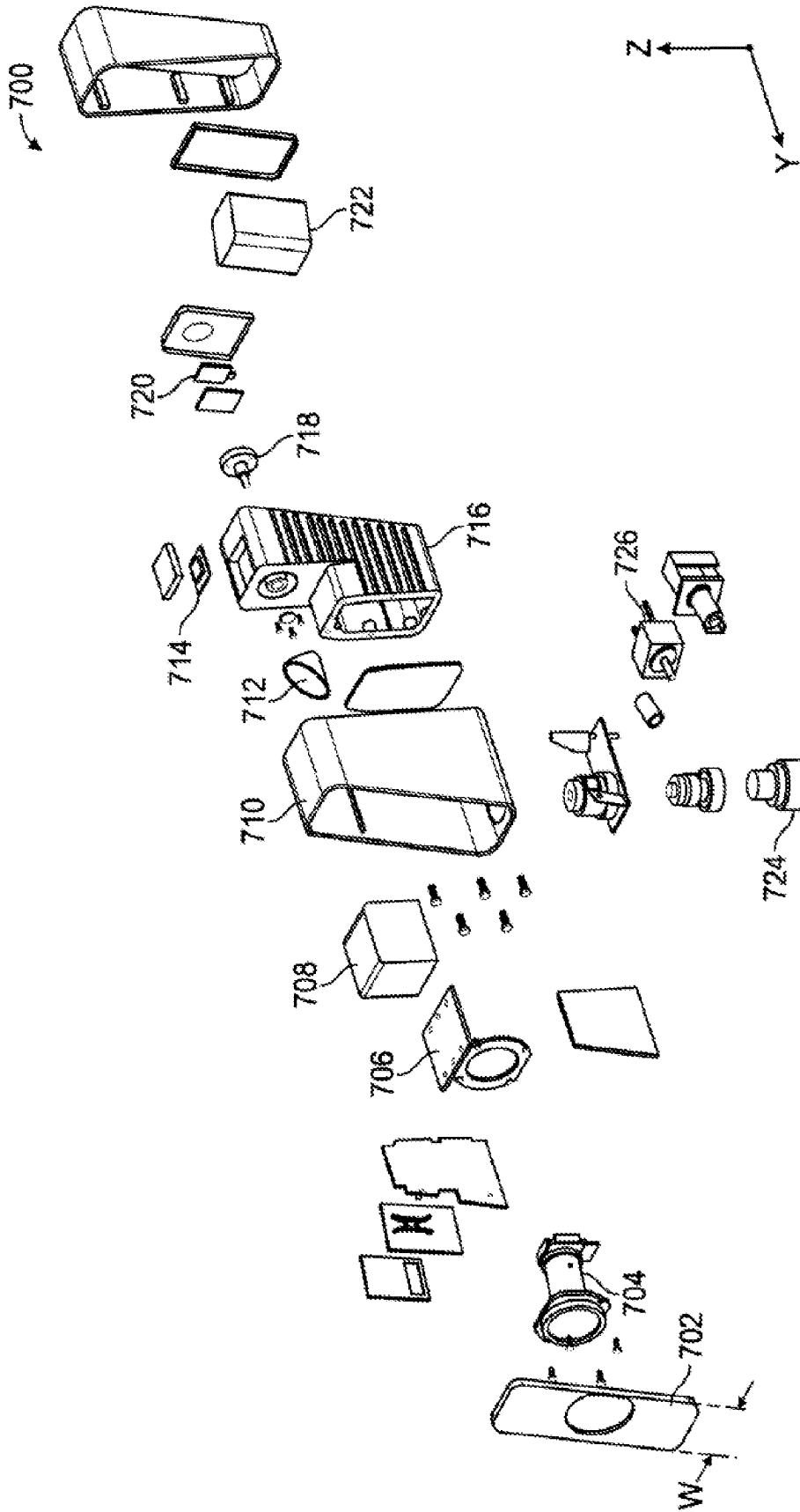


FIG. 7

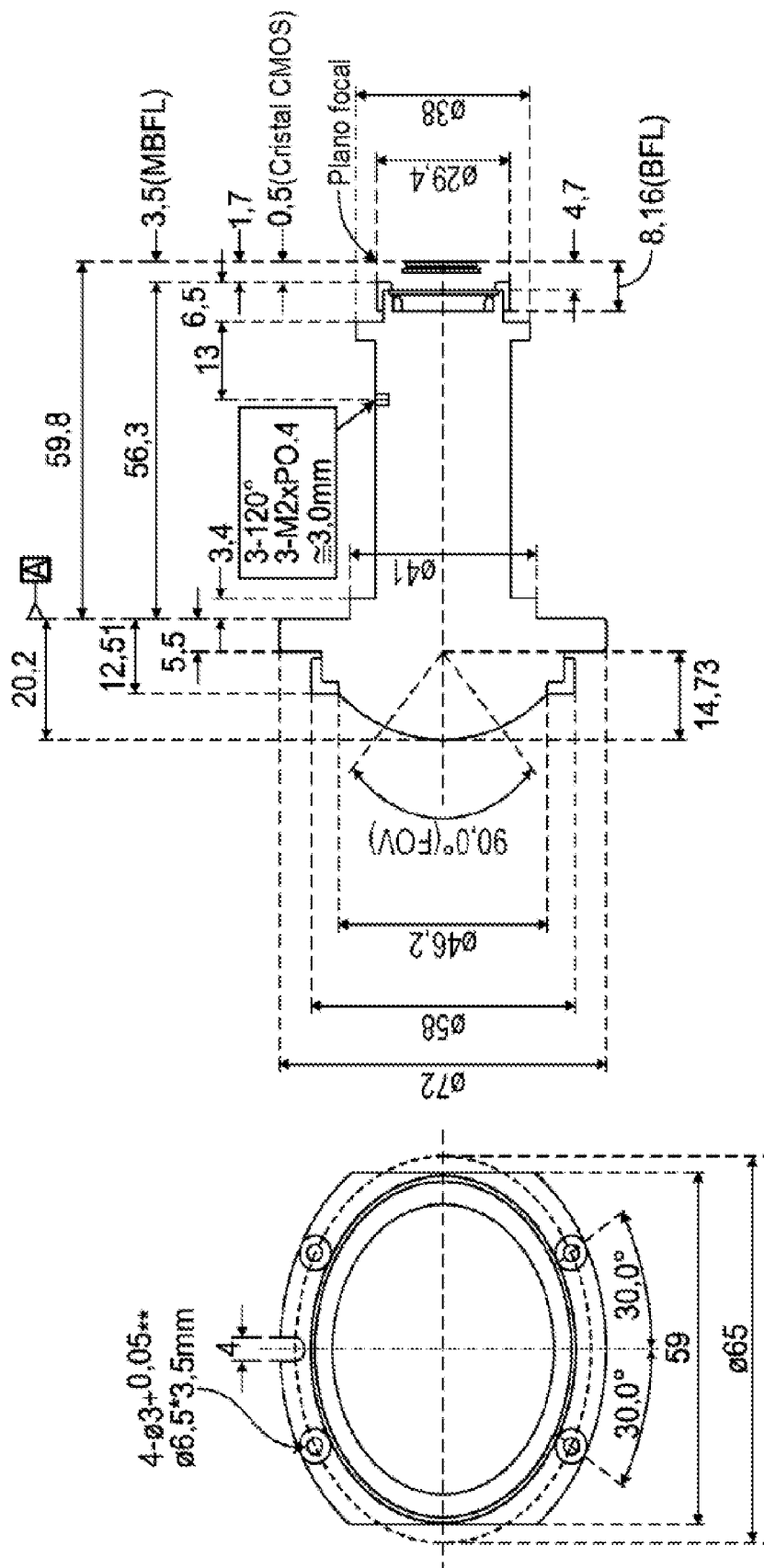


FIG. 8A

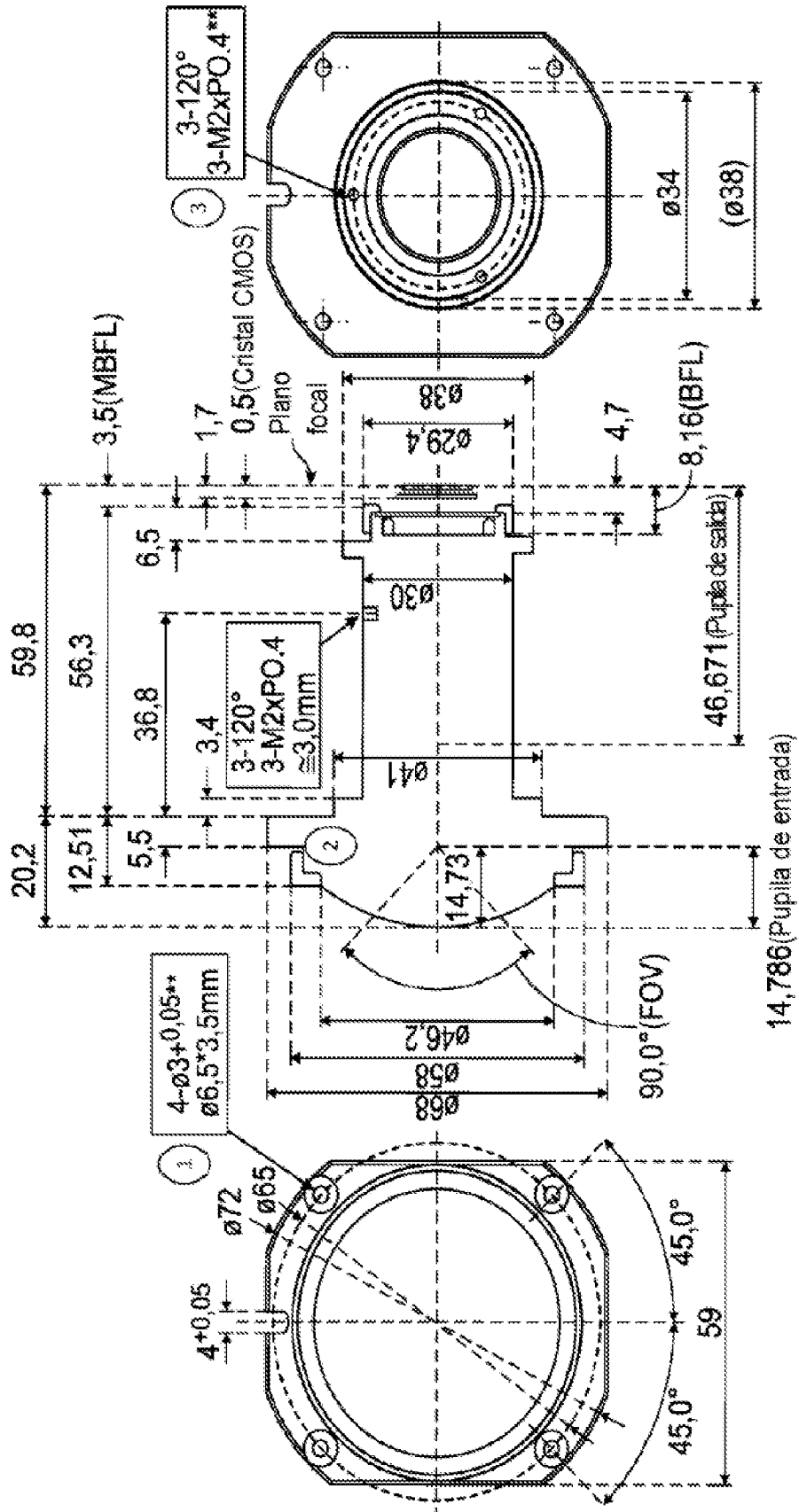


FIG. 8B

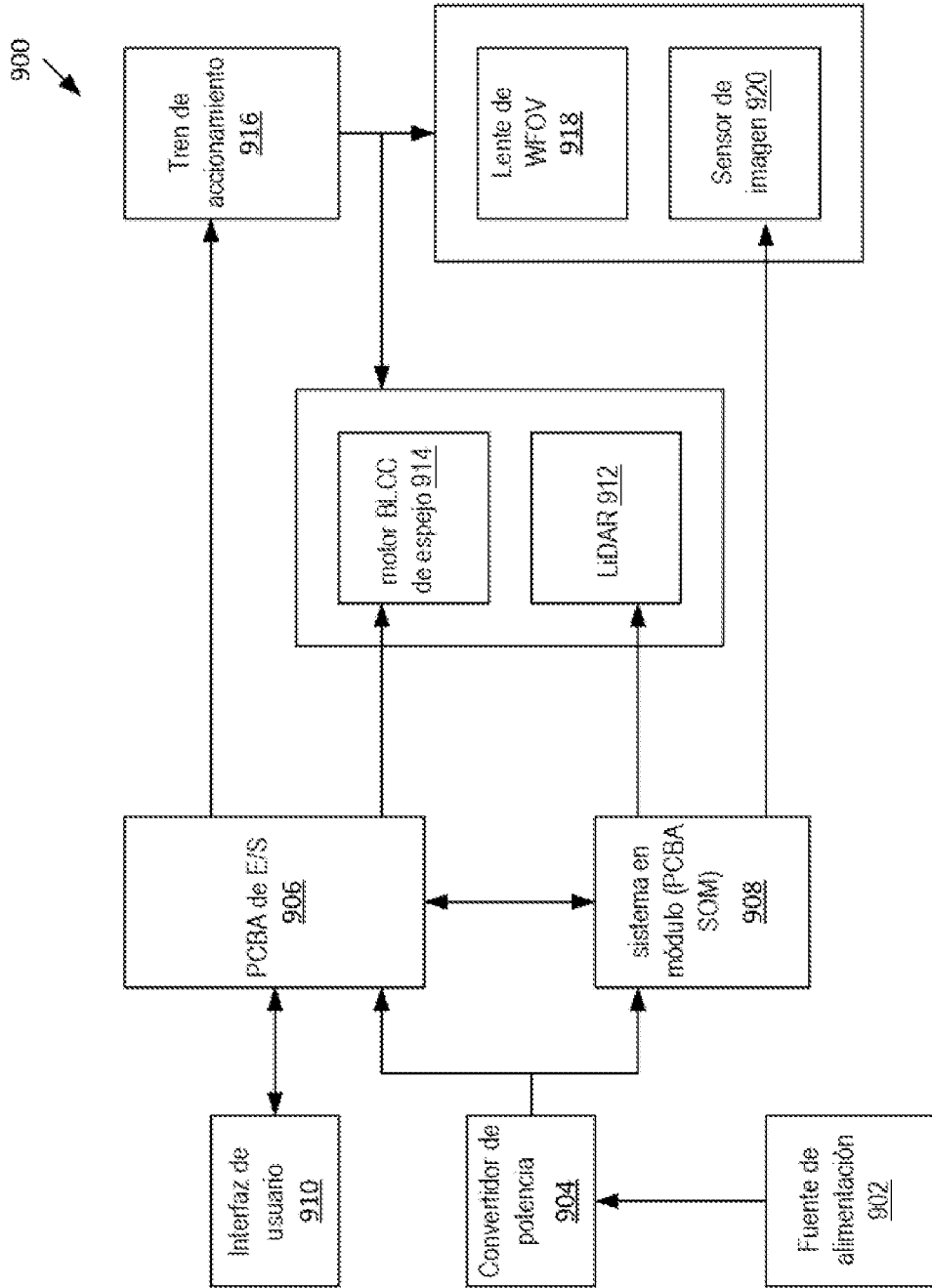


FIG. 9a

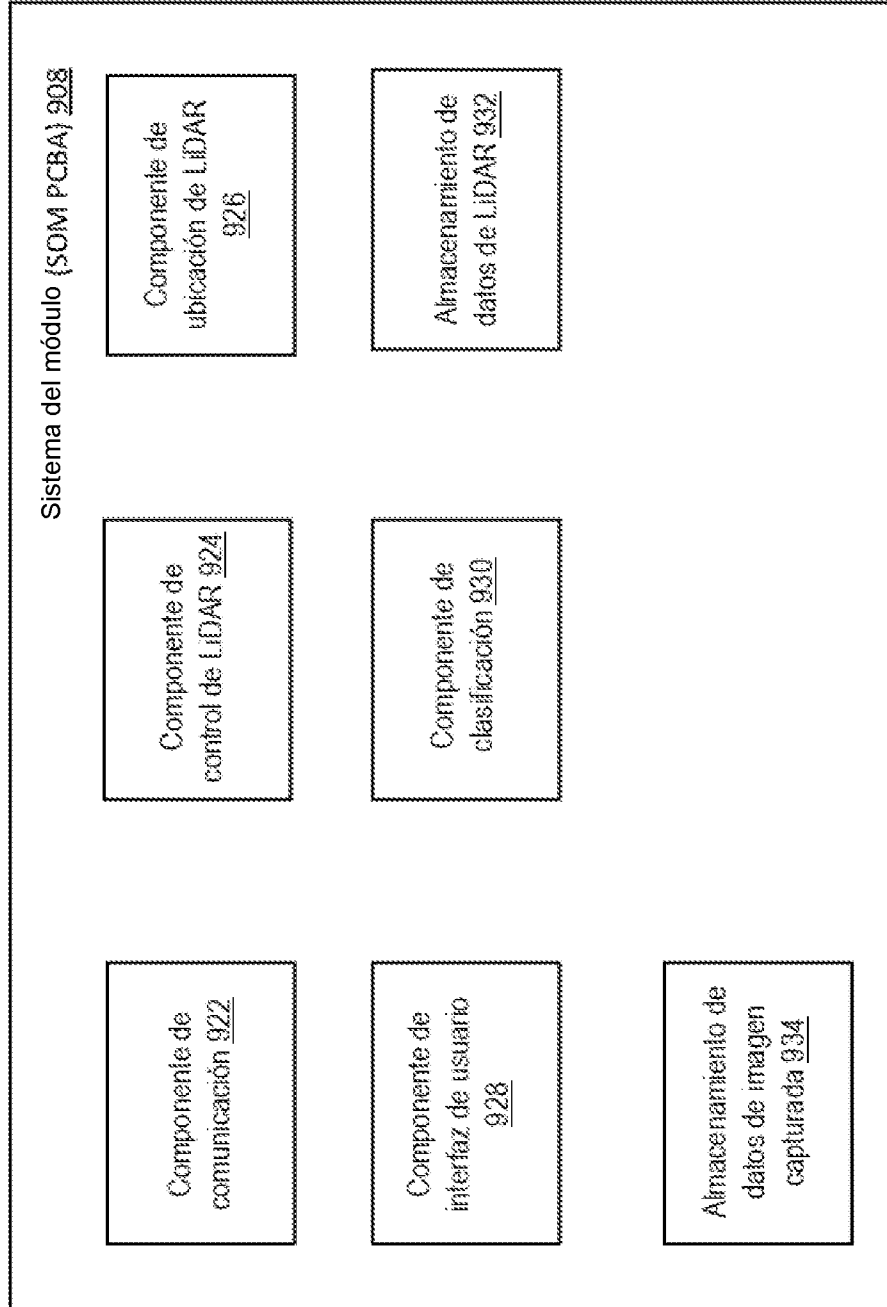


FIG. 9b

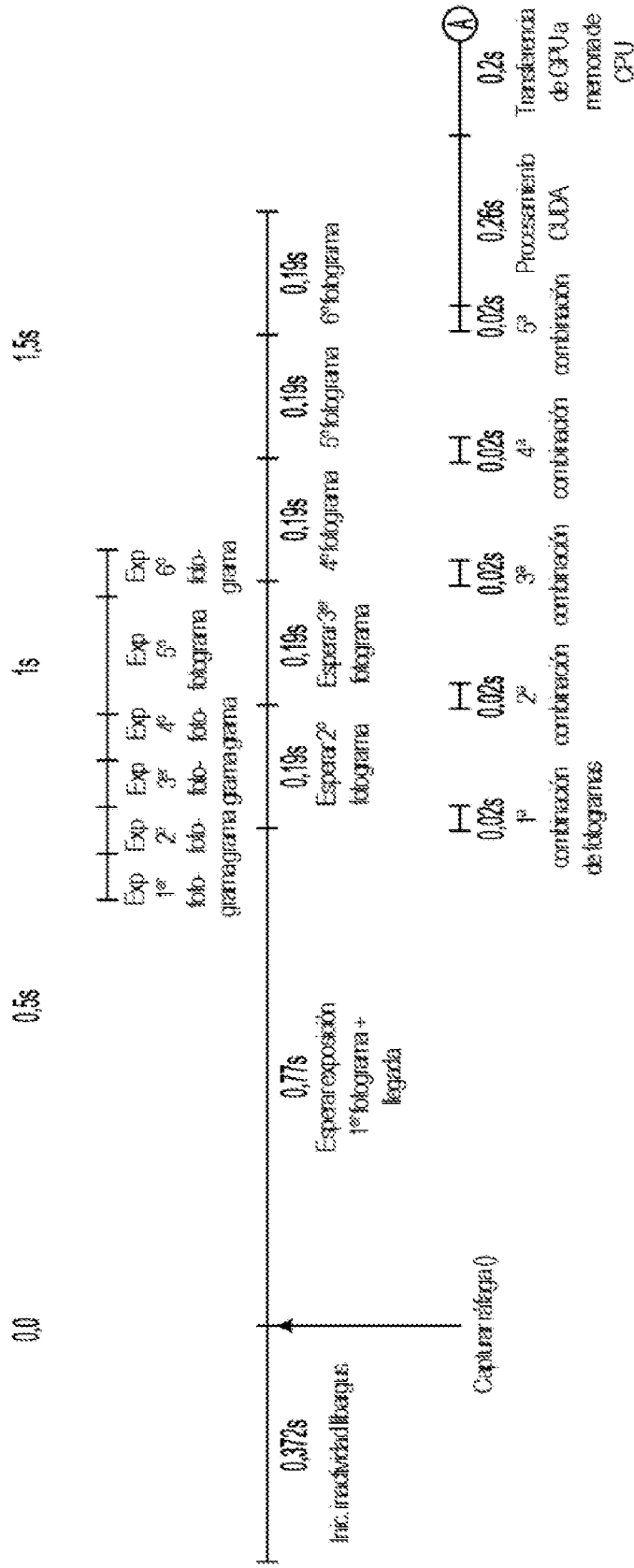


FIG. 10a

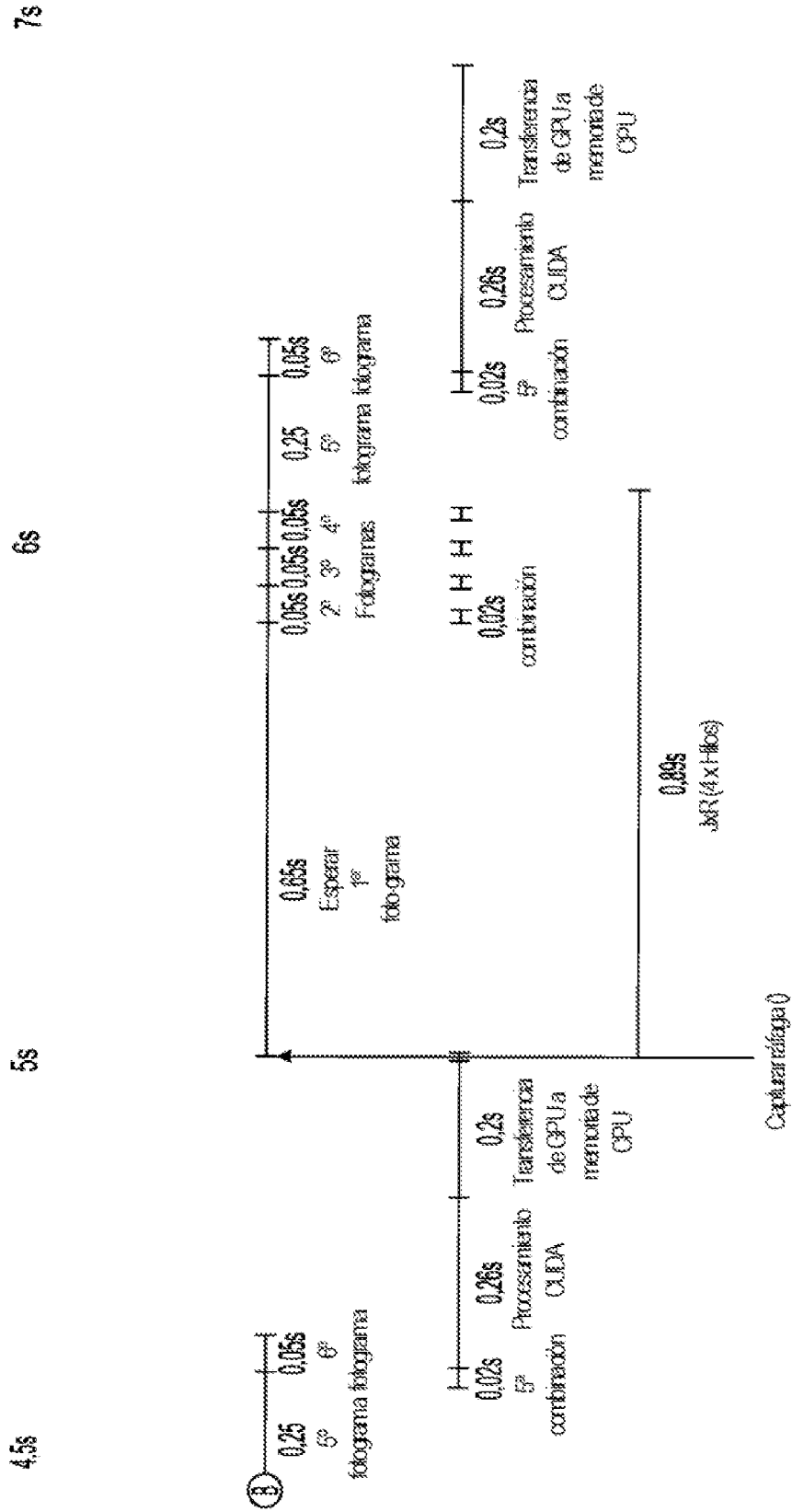


FIG. 10c

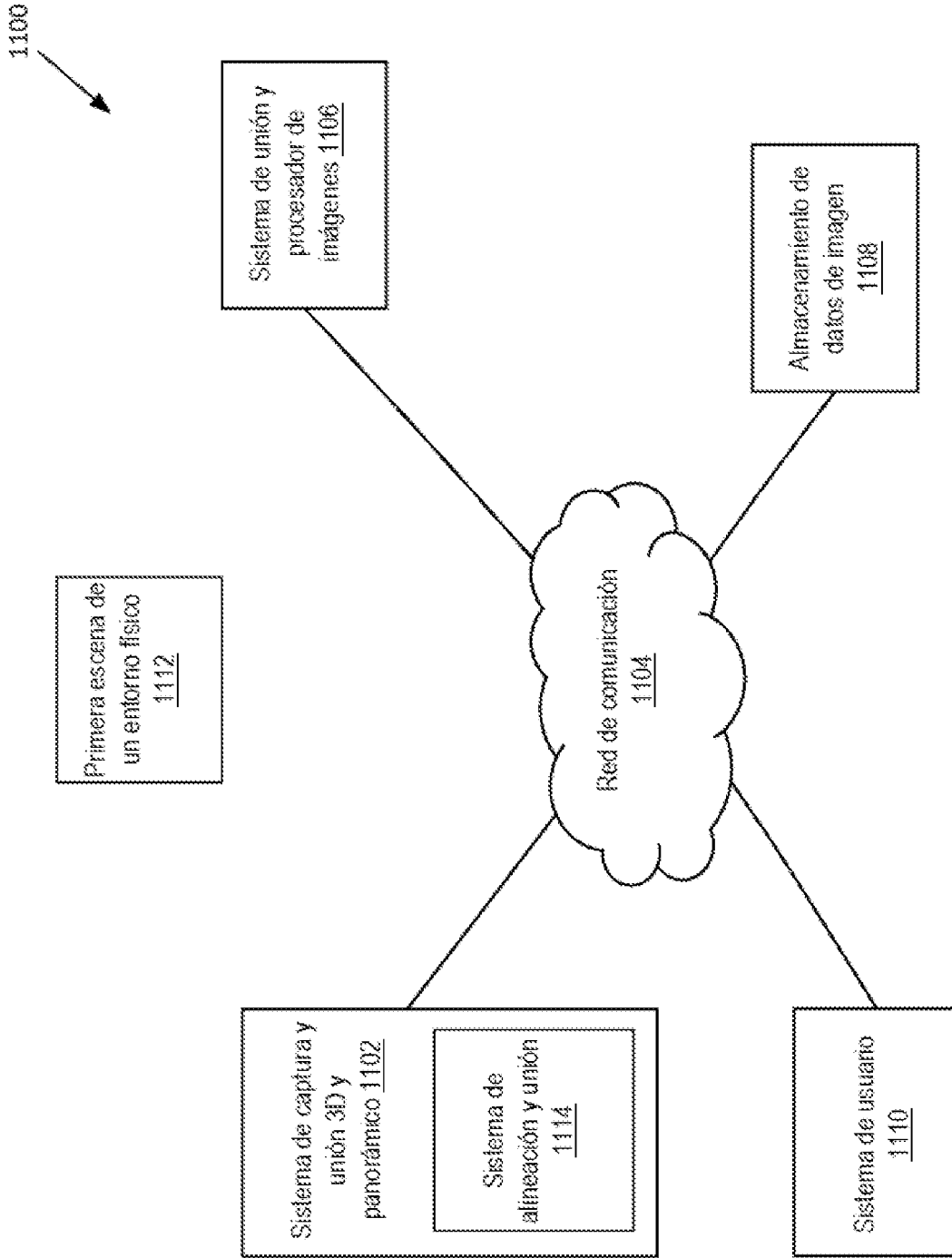


FIG. 11

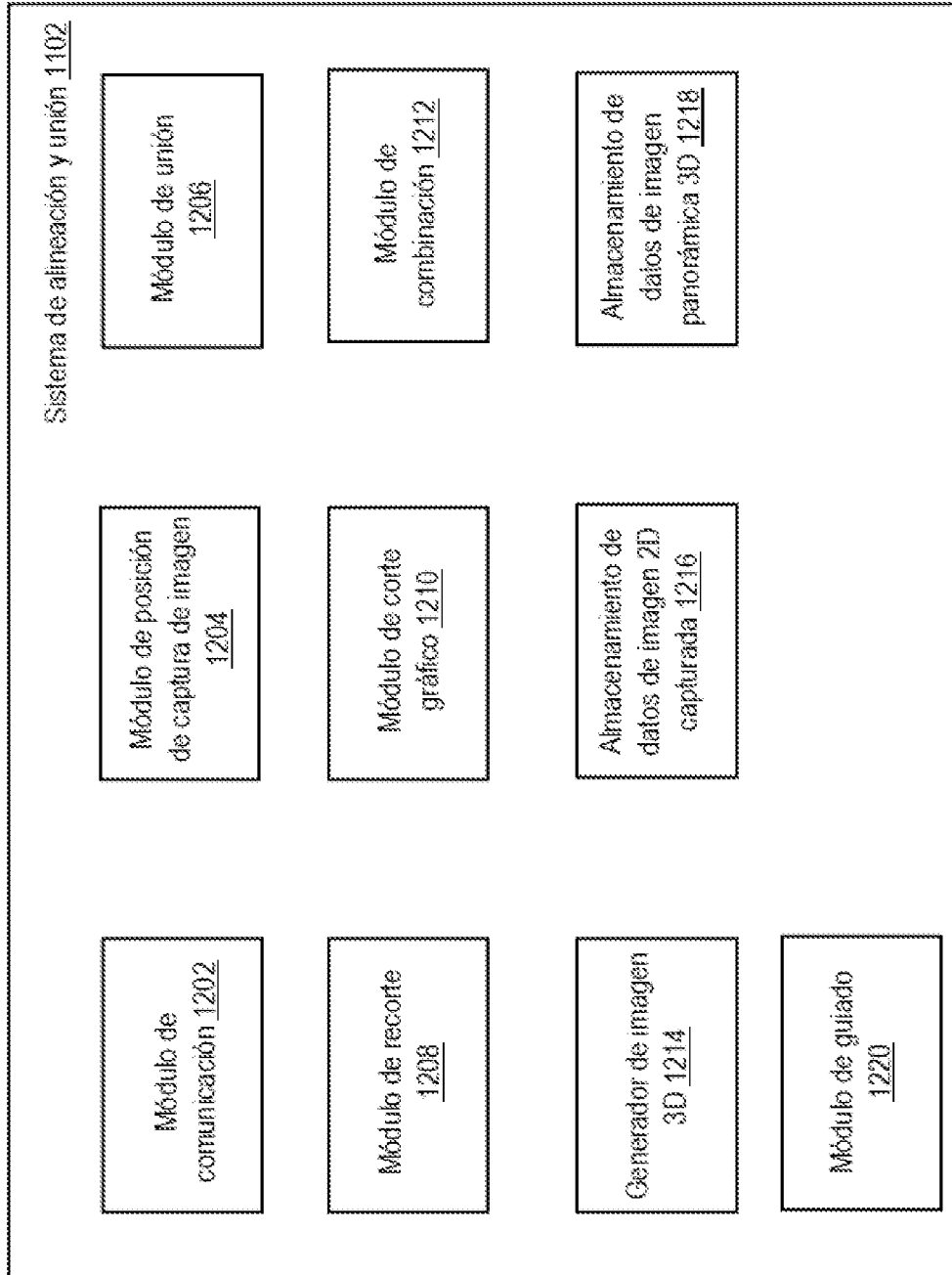


FIG. 12

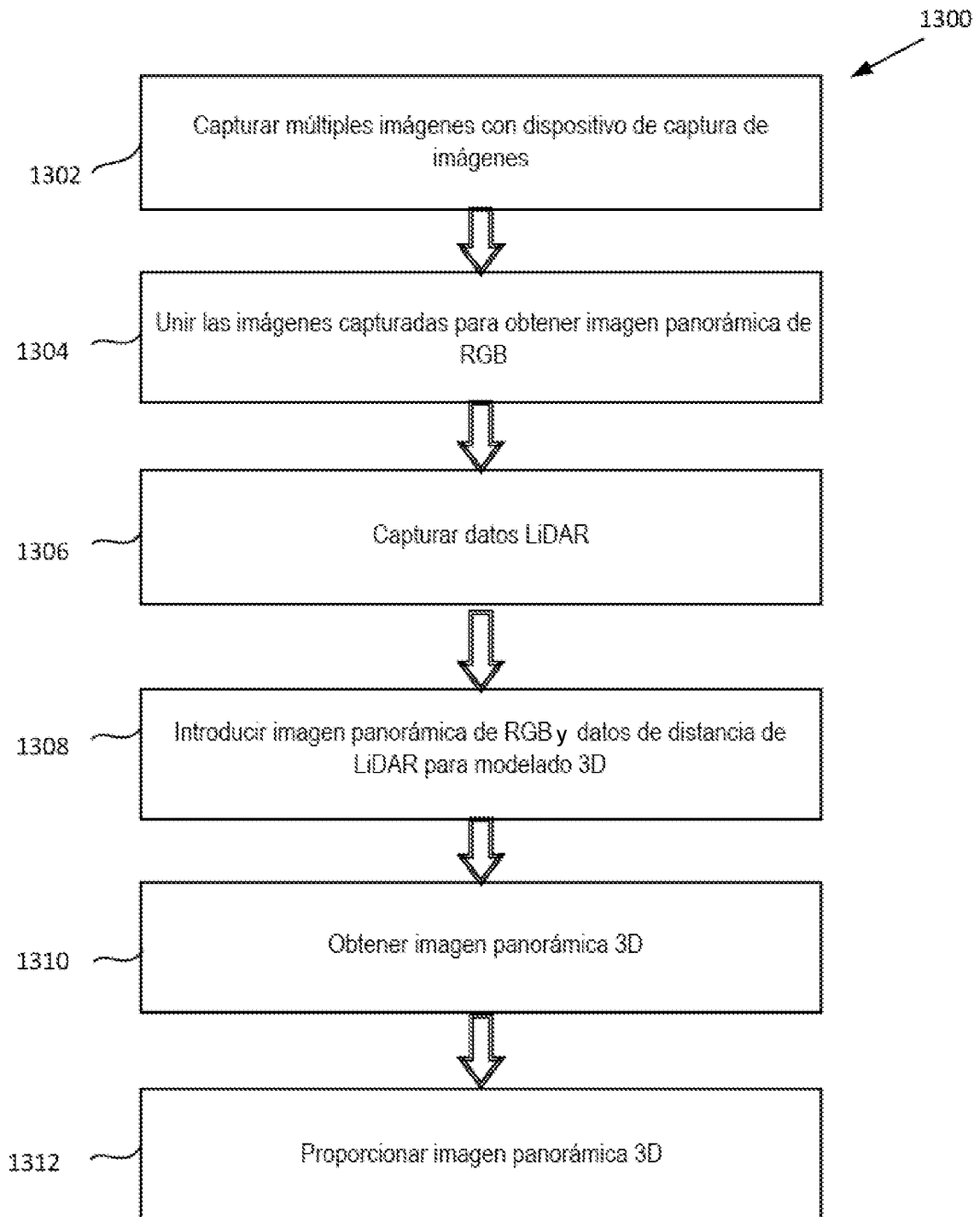
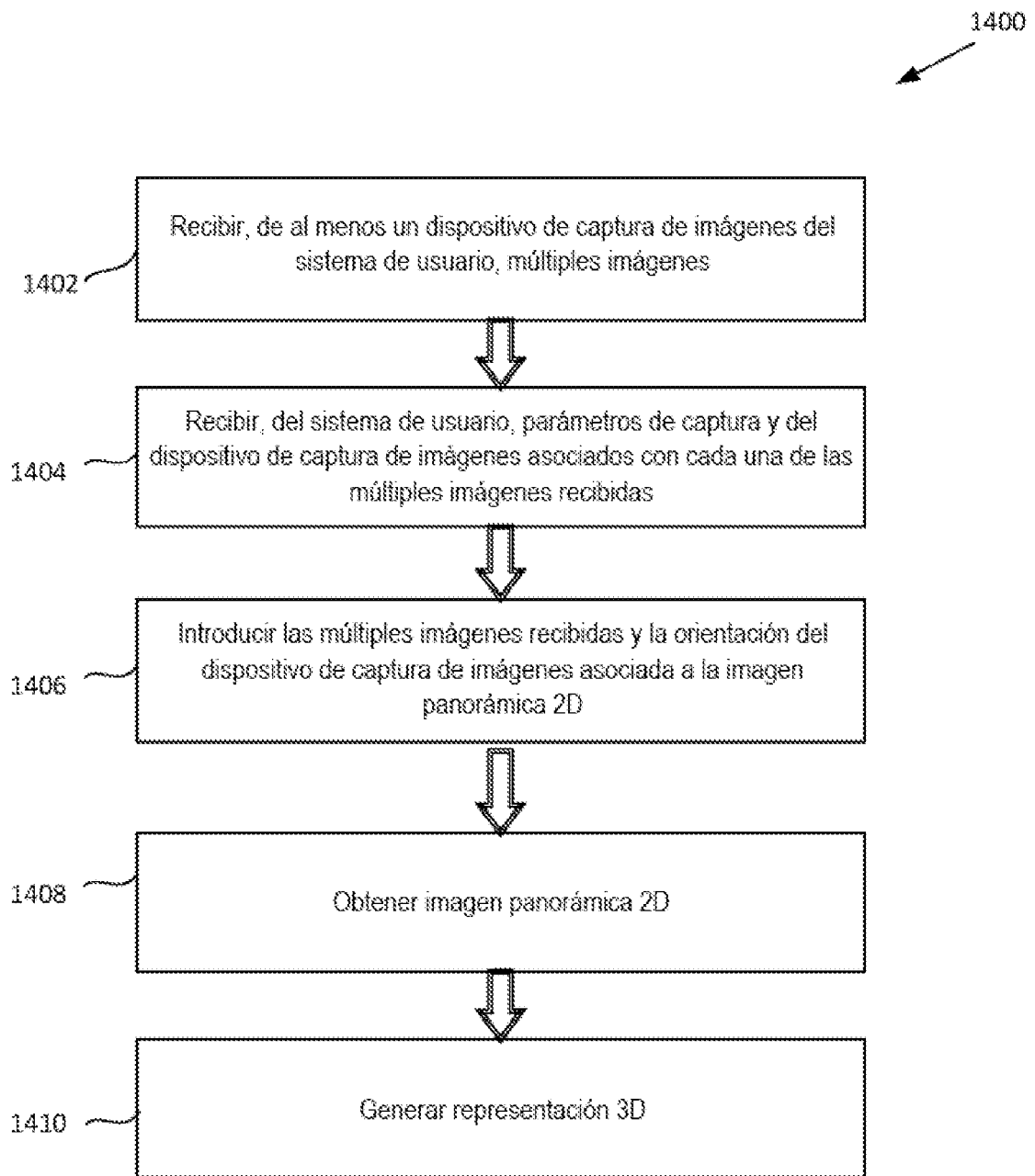


FIG. 13



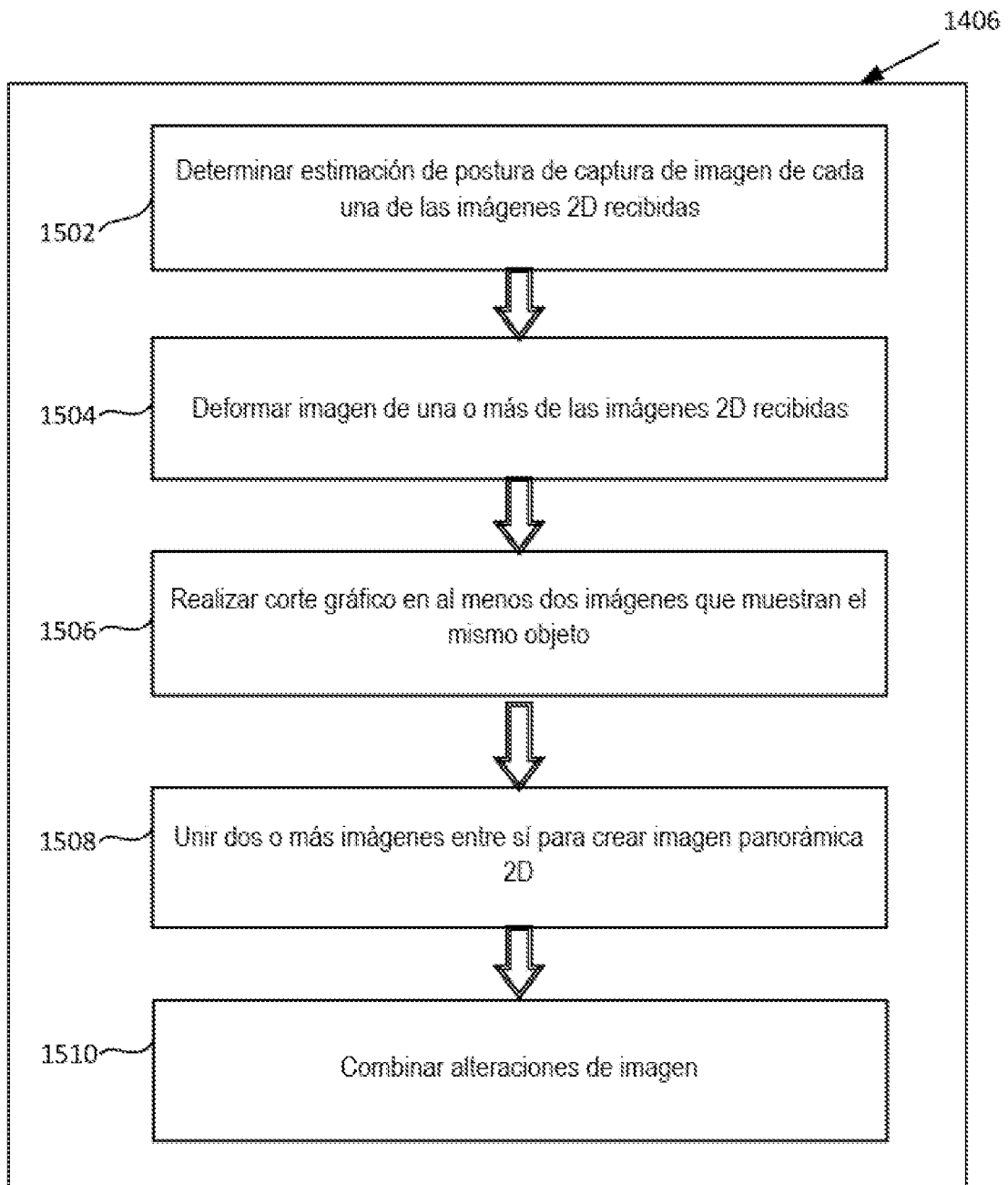


FIG. 15

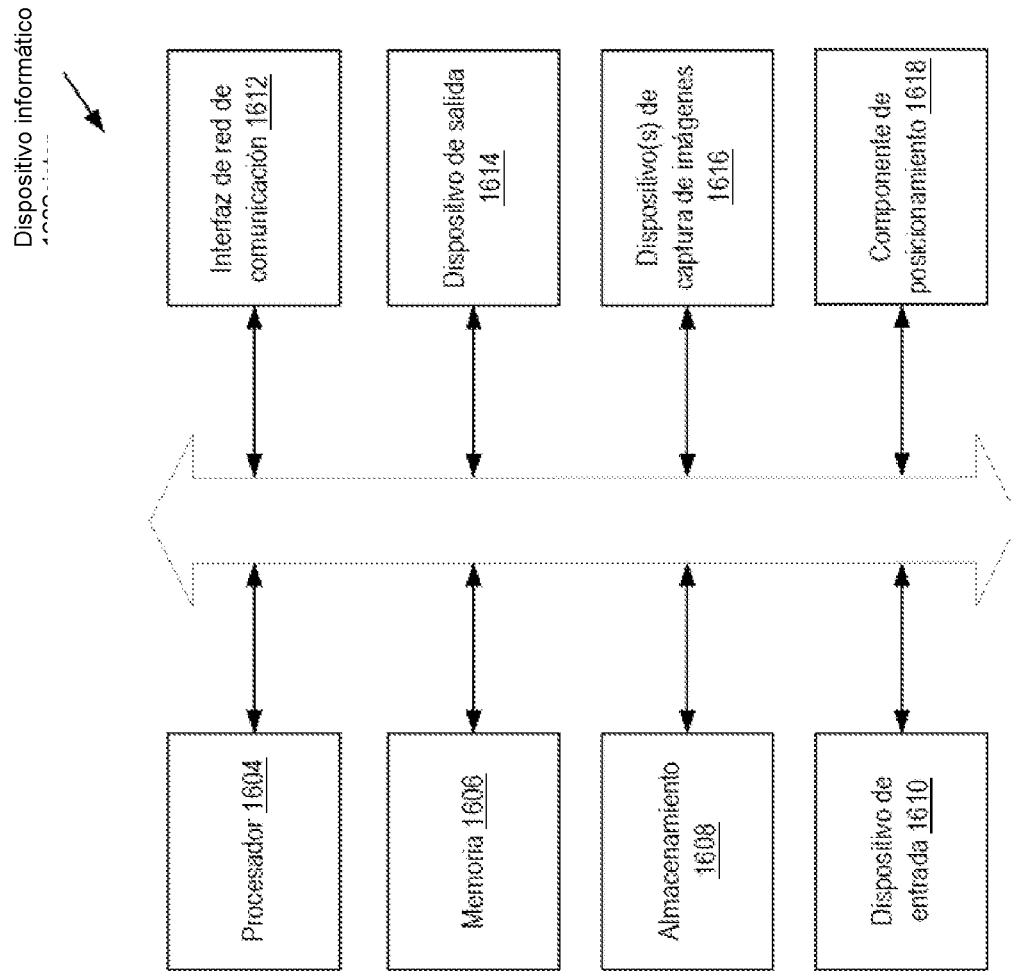


FIG. 16