

19



OFICINA ESPAÑOLA DE  
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 813 066**

51 Int. Cl.:

**G06T 7/246** (2007.01)

**G06T 7/73** (2007.01)

**G06T 3/00** (2006.01)

**G06T 3/40** (2006.01)

**H04N 5/225** (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

86 Fecha de presentación y número de la solicitud internacional: **29.09.2010 PCT/JP2010/067003**

87 Fecha y número de publicación internacional: **08.03.2012 WO12029193**

96 Fecha de presentación y número de la solicitud europea: **29.09.2010 E 10856738 (9)**

97 Fecha y número de publicación de la concesión europea: **15.07.2020 EP 2592820**

54 Título: **Dispositivo de conversión de imágenes, procedimiento y medio de registro legible por ordenador**

30 Prioridad:

**30.08.2010 JP 2010192927**

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

**22.03.2021**

73 Titular/es:

**RAKUTEN, INC. (100.0%)  
1-14-1, Tamagawa, Setagaya-ku  
Tokyo 158-0094, JP**

72 Inventor/es:

**HIRANO, HIROMI**

74 Agente/Representante:

**VIDAL GONZÁLEZ, Maria Ester**

ES 2 813 066 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**

Dispositivo de conversión de imágenes, procedimiento y medio de registro legible por ordenador

**5 Campo técnico**

La presente invención se refiere a un dispositivo de conversión de imágenes, un procedimiento, un medio de registro de información y a un sistema de conversión de imágenes que facilitan a un usuario capturar una secuencia de imágenes de toda la superficie exterior visible de un producto.

10

**Técnica antecedente**

Una secuencia de imágenes de un objeto que usa una pluralidad de imágenes obtenidas mediante la formación de imágenes secuenciales de toda la superficie exterior visible de un objeto que se puede aplicar a diversos fines, tal como la generación de un modelo de forma tridimensional del objeto (véase, por ejemplo, la Bibliografía 1 diferente de patente) y exposición aparente de la forma externa de un producto a un consumidor.

15

Las imágenes de toda la superficie exterior visible de un objeto se pueden obtener capturando sucesivamente imágenes de las apariencias, incluido un objeto en el mundo real, mientras se mueve una cámara, de manera que la dirección de captura de la misma se dirige al objeto en el transcurso de una trayectoria circular completa en torno al objeto.

20

Por lo tanto, para capturar imágenes de todo el entorno de un objeto, es necesario establecer la posición y la dirección de la cámara para satisfacer la técnica explicada anteriormente y, por lo tanto, típicamente son necesarios instrumentos especiales.

25

Por el contrario, en los últimos días, las cámaras digitales baratas se han hecho populares. Las lentes de las cámaras digitales capturan la luz reflectante del mundo externo, incluso de un objeto, y un sensor de imagen, como un CCD (dispositivo de carga acoplada), forma una imagen del mundo externo.

30

Cuando un usuario presiona un obturador, se crea una imagen en un formato de archivo electrónico basándose en la intensidad, la longitud de onda, etc., de la luz detectada por el sensor de imagen dentro de un cierto tiempo después de la presión del obturador. El archivo electrónico corresponde a una película revelada y un papel impreso revelado en el caso de cámaras ópticas convencionales.

35

Además, las tecnologías a menudo denominadas vista previa en vivo, y una vista en vivo también se han hecho populares. De acuerdo con la vista previa en vivo, una imagen que actualmente es detectada por un sensor de imagen se muestra directamente en una pantalla como una pantalla de cristal líquido en tiempo real (véase, por ejemplo, Bibliografía 2 no perteneciente a patente). Es decir, de acuerdo con la vista previa en vivo, se utiliza un buscador electrónico a través de una pantalla de cristal líquido, etc.

40

Una imagen mostrada en una pantalla de buscador por la vista previa en vivo corresponde a "una imagen a obtener si un usuario presiona un obturador actualmente en este momento".

45

Por lo tanto, si se utiliza la vista previa en vivo, es posible que el usuario verifique una composición, una exposición, etc., de un objeto antes de la obtención de imágenes.

**Documentos de la técnica anterior****50 Literatura de Patentes**

Documento de Patente 1: JP2005 049999 A (RICOH KK) divulga un procedimiento para generar una imagen tridimensional desde diferentes puntos de vista del mismo objeto.

**55 Literatura no Perteneciente a Patentes**

Literatura no Perteneciente a Patente 1: Kazutaka YASUDA, Takeshi MIGITA, Masato AOYAMA, Masayuki KUREKI, Naoki ASADA, "Generation of Dense Three-dimensional Shape Model from Rough Image Sequence of Entire Surroundings (Generación de un modelo de forma tridimensional densa a partir de una secuencia de imágenes aproximadas de todo el entorno)", Research Report of Information Processing Society of Japan, CVIM, Computer Vision and Image Media, 2003-CVIM-138 (11), páginas 73 a 80, <http://harp.lib.hiroshima-u.ac.jp/bitstream/harp/6507>, 8 de mayo de 2003.

60

Literatura no Perteneciente a Patente 2: Wikipedia, The Free Encyclopedia, Live Preview, [http://en.wikipedia.org/wiki/Live\\_preview](http://en.wikipedia.org/wiki/Live_preview), 15 de junio de 2010.

65

## Divulgación de la invención

### Problema a resolver por la invención

5 Sin embargo, cuando es deseable que un propietario de una tienda pequeña o un expositor individual de una subasta tome imágenes de todo el entorno del objeto para presentar un producto, la preparación de los instrumentos especiales explicados anteriormente no es realista.

10 Por consiguiente, existe demanda de una tecnología que facilite la adquisición de una secuencia de imágenes de todo el entorno de un objeto o la adquisición de una secuencia de imágenes que se pueda convertir fácilmente en una secuencia de imágenes de todo el entorno de un objeto después de la captura de imágenes.

15 La presente invención trata de abordar los problemas técnicos explicados anteriormente, y es un objeto de la presente invención proporcionar un dispositivo de conversión de imágenes, un procedimiento, un medio de registro de información y un sistema de conversión de imágenes que permita de forma apropiada que un usuario capture fácilmente una secuencia de imágenes de todo el entorno de un producto.

### Medios para resolver el problema

20 Se proporciona un dispositivo de conversión de imágenes, un procedimiento y un medio de registro de información no transitorio legible por ordenador de acuerdo con las reivindicaciones adjuntas.

### Efecto de la invención

25 De acuerdo con la presente invención, se puede proporcionar un dispositivo de conversión de imágenes, un procedimiento, un medio de registro de información y un sistema de conversión de imágenes que permiten de forma apropiada que un usuario pueda capturar fácilmente una secuencia de imágenes de todo el entorno de un producto.

### Breve descripción de los dibujos

30 La Figura 1 es un diagrama explicativo que muestra una configuración general de un sistema de procesamiento de imágenes;

35 La Figura 2 es un diagrama explicativo para explicar una configuración general de un dispositivo de imágenes de producto;

La Figura 3 es un diagrama explicativo que muestra una relación entre una imagen capturada, una imagen de vista previa en vivo y una imagen mostrada en una pantalla de buscador;

La Figura 4 es un diagrama explicativo que muestra una relación entre una imagen capturada, una imagen de vista previa en vivo y una imagen mostrada en una pantalla de buscador;

40 La Figura 5 es un diagrama de flujo que muestra un flujo de un control para un proceso de imágenes de producto ejecutado por el dispositivo de imágenes de producto;

La Figura 6 es un diagrama explicativo que muestra una configuración general de un dispositivo de conversión de imágenes de acuerdo con una realización de la presente invención;

45 La Figura 7 es un diagrama de flujo que muestra un flujo de un control para un proceso de conversión de imágenes ejecutado por el dispositivo de conversión de imágenes de acuerdo con la realización;

La Figura 8 es un diagrama explicativo que muestra una relación entre cada unidad del dispositivo de procesamiento de imágenes y cada unidad de un dispositivo terminal; y

50 La Figura 9 es un diagrama de flujo que muestra un flujo de un control para un proceso de imágenes ejecutado por cada unidad del dispositivo de procesamiento de imágenes y cada unidad del dispositivo terminal.

### Mejor modo de realizar la invención

55 Una realización de la presente invención se explicará a continuación. La realización es para explicar la invención, y no para limitar el alcance de la presente invención, que se define por las reivindicaciones adjuntas.

#### 1. Sistema de procesamiento de imágenes

60 La Figura 1 es un diagrama explicativo que muestra una configuración general de un sistema de procesamiento de imágenes. La explicación a continuación se dará con referencia a esta figura.

Un sistema de procesamiento de imágenes 101 incluye un dispositivo de imágenes de producto 121, un dispositivo de conversión de imágenes 141, un dispositivo de procesamiento de imágenes 161 y un dispositivo terminal 181.

65

5 El dispositivo de imágenes de producto 121 se usa por un propietario de una tienda o un expositor de una subasta, etc., que intenta presentar un producto para capturar imágenes del producto. Típicamente, el dispositivo de imágenes de producto 121 es una cámara digital, un teléfono celular, o un teléfono inteligente, etc., que tiene una función de procesamiento de información, pero puede ser un ordenador personal, etc., que tiene una cámara conectada al mismo a través de un USB (Bus Universal en Serie).

10 Una persona que intenta capturar imágenes utiliza el dispositivo de imágenes de producto 121 de tal manera que el producto está dispuesto en el centro de un campo visual tanto como sea posible y una distancia entre el producto y el dispositivo de imágenes de producto 121 se mantiene constante tanto como sea posible. A continuación, la persona que intenta capturar imágenes repite la captura de imágenes mientras recorre el producto dando una vuelta o rotando el propio producto, obteniendo de este modo una pluralidad de imágenes de producto.

15 El dispositivo de conversión de imágenes 141 convierte la pluralidad de imágenes de producto tomadas por el dispositivo de imágenes de producto 121 en una secuencia de imágenes de todo el entorno del producto. La secuencia de imágenes de toda la superficie exterior visible externa es un conjunto de imágenes obtenidas cuando se supone que una cámara que se mueve alrededor de la circunferencia en torno a un eje de rotación que atraviesa todo el producto captura imágenes del producto con una dirección del eje óptico dirigida hacia el producto.

20 Por lo tanto, las imágenes basadas en la presunción de que dichas imágenes se capturan, siendo el ángulo del eje de rotación con respecto a la cámara a través del producto constante, siendo la distancia entre el eje de rotación y la cámara constante, y siendo la distancia entre el producto y la cámara también constante, son las imágenes de toda la superficie exterior visible externa.

25 Como se ha explicado anteriormente, la persona que intenta capturar imágenes utilizando el dispositivo de imágenes de producto 121 no tiene experiencia en tal técnica de obtención de imágenes especial, y no utiliza ningún instrumento especial. Por lo tanto, se realizan diversas ayudas para facilitar la adquisición de la secuencia de imágenes en toda la superficie exterior visible externa tanto como sea posible cuando se usa el dispositivo de imágenes de producto 121.

30 Es decir, en el momento de la captura de imágenes, se utiliza un buscador electrónico como una pantalla de cristal líquido para una vista previa en vivo. La pantalla de buscador del buscador electrónico muestra la apariencia actual del producto tal como se ve a través del dispositivo de imágenes de producto 121, y también la apariencia pasada del producto ya capturado de manera superpuesta.

35 Por lo tanto, la persona que intenta tomar imágenes puede tomar imágenes del producto mientras las compara con los registros anteriores en cuanto al tamaño del producto en la imagen y en qué ángulo está el producto en la imagen. Por consiguiente, el ángulo del eje de rotación con respecto al dispositivo de imágenes de producto 121 a través del producto y la distancia entre el producto y el dispositivo de imágenes de producto 121 se puede establecer constante tanto como sea posible.

40 El dispositivo de conversión de imágenes 141 convierte la pluralidad de imágenes de producto en la secuencia de imágenes de todo el entorno, pero el resultado de la conversión obtenido a través de esta realización es un tipo de aproximación, y una característica de esta realización es un punto de que es innecesario un cálculo complejo divulgado en la Bibliografía 1 no perteneciente a patente.

45 El dispositivo de conversión de imágenes 141 obtiene la secuencia de imágenes de todo el entorno a partir de las imágenes de producto, y el dispositivo de procesamiento de imágenes 161 gestiona la secuencia de imágenes de todo el entorno del producto, pero típicamente, ambos dispositivos están configurados por un dispositivo servidor para la venta al por menor por Internet de tal producto.

50 Es decir, el propietario de la tienda, etc., carga la pluralidad de imágenes de productos tomadas por él mismo a través del dispositivo de imágenes de producto 121 como una cámara digital, en el dispositivo servidor.

55 El dispositivo servidor funciona como el dispositivo de conversión de imágenes 141 para convertir las imágenes de producto en la secuencia de imágenes de todo el entorno.

60 Además, el dispositivo servidor funciona como el dispositivo de procesamiento de imágenes 161 para permitir que un usuario que está planeando adquirir el producto a través de un pedido por correo a través de Internet o una subasta de Internet, explore la secuencia de imágenes de todo el entorno. En este momento, otra característica es ejecutar un procesamiento de imágenes apropiado para atraer al usuario, no proporcionar directamente la secuencia de imágenes de todo el entorno al usuario.

65 Con respecto al dispositivo de procesamiento de imágenes 161 que sirve como servidor web, el dispositivo

terminal 181 funciona como un terminal web, y solicita al dispositivo de procesamiento de imágenes 161 que proporcione adecuadamente la secuencia de imágenes de todo el entorno. Es decir, el dispositivo terminal 181 es, por ejemplo, un ordenador personal, un teléfono celular o un teléfono inteligente que se puede conectar a una red de comunicación informática como la Web.

5

La secuencia de imágenes de todo el entorno es una serie de imágenes sucesivas dispuestas para mostrar el aspecto del producto cuando se avanza alrededor del mismo con un giro.

10

El dispositivo terminal 181 notifica al dispositivo de procesamiento de imágenes 161 de la imagen en qué ángulo en el giro el usuario desea ver el producto.

15

Tras esta notificación, el dispositivo de procesamiento de imágenes 161 ejecuta un proceso de interpolación de manera apropiada en la secuencia de imágenes de todo el entorno, y genera una imagen interpolada que muestra la apariencia externa del producto de acuerdo con tal ángulo.

20

A continuación, el dispositivo de procesamiento de imágenes 161 transmite la imagen interpolada al dispositivo terminal 181, y el dispositivo terminal 181 muestra la imagen interpolada en la pantalla del monitor.

Es posible emplear una configuración en la que el usuario puede introducir directamente un ángulo desde el dispositivo terminal 181 a través de un teclado o un panel táctil. Sin embargo, algunos teléfonos celulares y teléfonos inteligentes, etc., tienen una función de brújula, una función de GPS (sistema de posicionamiento global), y una función de detección de inclinación usando un sensor de aceleración, etc.

25

Por lo tanto, si el ángulo obtenido a través de dichas funciones se usa directamente o se usa un ángulo multiplicado por una determinada constante, la dirección del producto expresada en la imagen interpolada se puede cambiar simplemente cambiando la dirección del dispositivo terminal 181. Esta es la otra característica de esta realización en la que el usuario puede observar la apariencia externa del producto con un giro a través de una sencilla operación.

30

El dispositivo de procesamiento de imágenes 161 y el dispositivo terminal 181 están configurados por dispositivos separados, pero ambos dispositivos pueden ser configurados conjuntamente por un ordenador personal, un teléfono inteligente o un teléfono celular, etc. En este caso, la transmisión y recepción de información entre el dispositivo terminal 181 y el dispositivo de procesamiento de imágenes 161 se puede realizar mediante el intercambio de señales eléctricas en los circuitos electrónicos de un dispositivo.

35

Por ejemplo, la secuencia de imágenes de todo el entorno del producto se descarga a un teléfono inteligente desde el servidor web que configura el dispositivo de conversión de imágenes 141. A continuación, el teléfono inteligente funciona como el dispositivo de procesamiento de imágenes 161, y el teléfono inteligente genera una imagen interpolada de la secuencia de imágenes de todo el entorno de acuerdo con la inclinación del teléfono inteligente, y muestra la imagen interpolada en la pantalla del monitor.

40

Si bien las funciones respectivas para el procesamiento de imágenes se realizan mediante diversos esquemas servidor-cliente de manera distribuida, o se realizan mediante un dispositivo solo, se pueden cambiar según sea necesario, como se ha explicado anteriormente, dependiendo del propósito y el rendimiento de un dispositivo de procesamiento de información que realiza el procesamiento de imágenes, como un ordenador.

45

A continuación, los dispositivos respectivos del sistema de procesamiento de imágenes 101 se explicarán con más detalle a continuación.

50

## 2. Dispositivo de imágenes de producto

La Figura 2 es un diagrama explicativo para explicar una configuración general del dispositivo de imágenes de producto 121. Se dará una explicación con referencia a esta figura.

55

El dispositivo de imágenes de producto 121 incluye una unidad de sensor de imagen 201, una unidad de recepción de instrucciones 202, una unidad de memoria 203 y una unidad de visualización de buscador 204. Además, el dispositivo de imágenes de producto puede configurarse para tener además una unidad de sensor de inclinación 205.

60

Es decir, la unidad de sensor de imagen 201 detecta la luz incidente del mundo externo donde se dispone el producto, y emite una imagen que representa el resultado de la detección.

Típicamente, la unidad de sensor de imagen 201 es un sensor de imagen como un CCD. La luz detectada por la unidad de sensor de imagen 201 llega al mismo a través de lentes y diversas aberturas.

65

Por el contrario, la unidad de recepción de instrucciones 202 recibe una instrucción para una captura de imagen del usuario. La unidad de recepción de instrucciones 202 corresponde a un llamado botón obturador.

5 Además, la unidad de memoria 203 almacena la imagen detectada por la unidad de sensor de imagen 201 tras la recepción la instrucción de captura de imagen.

10 La unidad de memoria 203 es típicamente un medio de memoria de información no volátil, tal como una tarjeta de memoria flash, como una EEPROM (memoria de solo lectura programable y borrable eléctricamente), o un disco duro, que graba un archivo de imagen tomada.

La unidad de visualización de buscador 204 muestra la imagen detectada por la unidad de sensor de imagen 201 en la pantalla de buscador.

15 La persona que intenta tomar imágenes ajusta la posición y la dirección del dispositivo de imágenes de producto 121 y ajusta el enfoque y la apertura al referirse a la imagen mostrada en la pantalla de buscador.

Las funciones de las unidades respectivas son las mismas que las de una cámara digital construida con una función de vista previa en vivo.

20 La unidad de visualización de buscador 204 tiene una característica para sintetizar la imagen detectada por la unidad de sensor de imagen 201 con la imagen almacenada en la unidad de memoria 203 cuando la imagen detectada por la unidad de sensor de imagen se muestra en la pantalla de buscador.

25 En dicha síntesis, la imagen que representa el campo visual de la captura de imagen pasada y la imagen que representa el campo visual para una captura de imagen presente se disponen de forma superpuesta en la pantalla de buscador. Por lo tanto, la persona que intenta tomar imágenes puede ajustar la posición y la dirección del dispositivo de imágenes de producto 121 mientras compara ambas imágenes en tiempo real.

30 La técnica de síntesis más sencilla para superponer las dos imágenes es hacer que la imagen (una imagen ya tomada) almacenada en la unidad de memoria 203 sea semitransparente, superponer en síntesis con la imagen (una imagen de vista previa en vivo) detectada por la unidad de sensor de imagen 201, y mostrar la imagen de vista previa en vivo en la pantalla de buscador.

35 Las Figuras 3 y 4 son diagramas explicativos que muestran una relación entre la imagen ya capturada, la imagen de vista previa en vivo, y la imagen que se muestra en la pantalla de buscador. La explicación se dará con referencia a estas figuras. En estas figuras, se usa una línea de puntos según sea necesario para expresar una condición semitransparente.

40 La Figura 3 muestra una imagen ya capturada 311 donde un producto 301 está dispuesto en un centro sustancial, y una imagen de vista previa en vivo 321 donde el producto 301 también está dispuesto en un centro sustancial. Ambas imágenes miran el producto 301 desde un ángulo ligeramente diferente.

45 De acuerdo con una imagen de buscador 331 obtenida al hacer que la imagen ya capturada 311 sea semitransparente y superponiendo y sintetizando tal imagen en la imagen de vista previa en vivo 321, las dos imágenes del producto 301 se muestran de manera superpuesta sustancialmente al mismo tamaño. Por lo tanto, si se presiona el obturador en esta condición actual, se puede reducir un error en el momento de la conversión en la secuencia de imágenes de todo el entorno.

50 Por el contrario, en la Figura 4, en la imagen ya tomada 311, el producto 301 está dispuesto en un centro sustancial, pero en la imagen de vista previa en vivo 321, la posición del producto 301 y el tamaño de la misma difieren en gran medida de los de la imagen ya capturada 311.

55 En este caso, de acuerdo con la imagen de buscador 331 obtenida al hacer que la imagen ya capturada 311 sea semitransparente, y superponiendo y sintetizando con la imagen de vista previa en vivo 321, las dos imágenes del producto 301 difícilmente se superponen entre sí. Por lo tanto, incluso si se presiona el obturador en esta condición actual y se obtiene una nueva imagen capturada 311 que representa el mismo aspecto que la imagen de vista previa en vivo 321, el error en el momento de la conversión en la secuencia de imágenes de todo el entorno posiblemente se vuelve grande.

60 Es decir, como se muestra en la Figura 3, resulta adecuado si la persona que intenta capturar imágenes ajusta la posición del dispositivo de imágenes de producto 121 de tal manera que la imagen del producto 301 capturada en el pasado y la imagen del producto 301 vista actualmente tengan el mismo tamaño en la medida de lo posible, pero con un ángulo de visión ligeramente diferente, y proporciona una entrada de instrucciones para la captura de imágenes.

65

Incluso si la persona que intenta capturar imágenes mueve el dispositivo de imágenes de producto 121, la posición del producto 301 en la imagen ya tomada 311 y la del fondo no cambian, pero la posición del producto 301 en la imagen de vista previa en vivo 321 y la del fondo cambian junto con el movimiento de la persona que intenta capturar las imágenes.

5 Por lo tanto, es fácil para la persona que intenta capturar imágenes distinguir qué parte de la imagen de buscador 331 se origina a partir de la imagen ya tomada 311 y qué parte de la imagen de buscador se origina a partir de la imagen de vista previa en vivo 321.

10 Además, para aclarar aún más la distinción de ambas imágenes, se puede aplicar un esquema que aplica un filtro de color o un filtro de extracción de contorno a la imagen ya tomada 311 y sintetiza tal imagen con la imagen de vista previa en vivo 321.

15 Otro esquema posible es mostrar alternativamente la imagen ya tomada 311 y la imagen de vista previa en vivo 321 en la pantalla de buscador. Como se ha explicado anteriormente, se puede distinguir si la imagen que se muestra en la pantalla de buscador es la imagen de vista previa en vivo 321 o la imagen ya tomada 311 basándose en si la visualización de la pantalla de buscador sigue el movimiento de la persona que intenta capturar las imágenes o es fija.

20 Por consiguiente, el periodo de tiempo para visualizar ambas imágenes se puede establecer de forma arbitraria. Sin embargo, en general, el tiempo de cambio de visualización de ambas imágenes se configura de tal manera que el tiempo de visualización de la imagen de vista previa en vivo 321 sea mayor que el tiempo de visualización de la imagen ya tomada 311.

25 La imagen en el momento del comienzo de la captura de imagen del producto 301 se aplica típicamente como la imagen ya tomada 311 sometida a síntesis, pero está bien si la persona que intenta capturar las imágenes puede seleccionar cualesquiera imágenes según sea necesario. Además, la imagen capturada justo antes se puede aplicar como la imagen ya tomada 311 para la síntesis.

30 La unidad de sensor de inclinación 205 detecta la inclinación del dispositivo de imágenes de producto 121, y tras la recepción de la instrucción de captura de imagen, la unidad de memoria 203 almacena la imagen detectada por la unidad de sensor de imagen 201 en asociación con la inclinación detectada por unidad de sensor de inclinación.

35 Cuando las imágenes se capturan mientras se recorre el producto en un giro, es típico que el eje de rotación sea la dirección de la gravedad. En este caso, la unidad de sensor de inclinación 205 puede configurarse mediante un sensor de aceleración. Junto con tal configuración, la inclinación se puede detectar a través de las funciones de un GPS y una brújula, etc.

40 Cuando la inclinación del dispositivo de imágenes de producto 121 con respecto al eje de la dirección gravitacional se detecta a través del sensor de aceleración, y el nivel de rotación del dispositivo de imágenes de producto 121 en torno al eje de la dirección gravitacional se detecta a través del GPS, y la brújula, etc., la unidad de sensor de inclinación 205 puede detectar toda la inclinación del dispositivo de imágenes de producto 121 en las direcciones de tres ejes. Esto hace posible que el usuario capte completamente la postura del dispositivo de imágenes de producto 121 en el mundo real. La información sobre la inclinación en las direcciones de los tres ejes se remite por el dispositivo de conversión de imágenes 141 que se analizará más adelante para normalizar la imagen ya tomada 311 y obtener la secuencia de imágenes de todo el entorno.

50 Como se ha explicado anteriormente, para obtener una secuencia de imágenes apropiada de todo el entorno, es deseable que el ángulo del eje de rotación con respecto al dispositivo de imágenes de producto 121 a través del producto 301 sea constante, es decir, la inclinación con respecto a la gravedad debe ser constante.

55 Por lo tanto, si la información auxiliar que incluye la inclinación del dispositivo de imágenes de producto 121 se detecta actualmente, es decir, la información sobre la inclinación de la imagen de vista previa en vivo 321 e información sobre la inclinación en el momento de tomar la imagen ya tomada 311 se muestra en la imagen de buscador 331, la persona que intenta capturar imágenes puede capturar imágenes del producto 301 en un ángulo lo más constante posible y, por lo tanto, se puede reducir el error en el tiempo de conversión en la secuencia de imágenes de todo el entorno.

60 Como se muestra en las Figuras 3 y 4, un gráfico de barras 322 que representa la inclinación de la imagen de vista previa en vivo 321 y un gráfico de barras 312 que representa la inclinación de la imagen ya tomada 311 se muestran en la imagen de buscador 331. Es apropiado para la persona que intenta capturar imágenes capturar las imágenes mientras intenta igualar las longitudes respectivas de ambos gráficos de barras 312 y 322.

65 El esquema para indicar la inclinación no se limita a los gráficos de barras 312 y 322, y se puede adoptar

cualquier gráfico que represente el nivel de consistencia/inconsistencia. Además, el valor de la inclinación se puede mostrar a través de información de textura.

5 La Figura 5 es un diagrama de flujo que muestra un flujo de un control para un proceso de imágenes de producto ejecutado por el dispositivo de imágenes de producto 121. En lo sucesivo en el presente documento, la explicación se dará con referencia a esta figura.

10 Este proceso se inicia por el usuario que selecciona comenzar a tomar imágenes del producto que son los materiales de la secuencia de imágenes de todo el entorno a través de una cámara digital, etc., configurando el dispositivo de imágenes de producto 121.

15 En primer lugar, en el dispositivo de imágenes de producto 121, la unidad de sensor de imagen 201 detecta la luz incidente del mundo externo donde se dispone el producto, y emite una imagen que representa el resultado de la detección (etapa S501). La imagen de salida es la imagen de vista previa en vivo 321.

A continuación, la unidad de sensor de inclinación 205 obtiene la inclinación actual del dispositivo de imágenes de producto 121 (etapa S502).

20 Posteriormente, la unidad de recepción de instrucciones 202 determina si se da o no una instrucción de captura de imagen del usuario (etapa S503).

25 Cuando el usuario ha dado las instrucciones de captura de imagen (etapa S503: Sí), la unidad de memoria 203 almacena la imagen de vista previa en vivo 321 en asociación con la inclinación detectada en la etapa S502 (etapa S504). Posteriormente, el control regresa a la etapa S501.

30 Por el contrario, cuando el usuario no ha dado la instrucción de captura de imagen (etapa S502: NO), se verifica si hay imágenes almacenadas en la unidad de memoria 203 después del inicio de este proceso, es decir, las imágenes ya capturadas 311 (etapa S505).

35 Cuando hay imágenes ya capturadas 311 (etapa S505: Sí), se selecciona una cualquiera de las imágenes ya capturadas 311 (etapa S506). Como se ha explicado anteriormente, con respecto a esta selección, la imagen capturada al principio puede seleccionarse, el usuario puede cambiar la imagen según corresponda durante el proceso, o puede seleccionarse la imagen capturada justo antes.

40 A continuación, la imagen ya capturada seleccionada 311 se hace semitransparente, y se sintetiza con la imagen de vista previa en vivo 321 (etapa S507). En este momento, como se ha explicado anteriormente, la síntesis se puede realizar después de que se aplique un filtro de color o un filtro de extracción de contorno a la imagen ya capturada 311. Además, como se ha explicado anteriormente, cualquier imagen se puede seleccionar en forma de tiempo compartido y la imagen seleccionada puede tomarse como resultado de la síntesis.

45 Además, los gráficos de barras 312 y 322 que indican la inclinación asociada con la imagen ya capturada seleccionada 311 y la inclinación actual del dispositivo de imágenes de producto 121 obtenida en la etapa S502 se trazan sobre el resultado de la síntesis en la etapa S506, obteniendo de este modo la imagen de buscador 331 (etapa S508).

A continuación, la unidad de visualización de buscador 204 muestra la imagen de buscador 331 en la pantalla de buscador configurada por una pantalla de cristal líquido, etc. (etapa S509), y el proceso regresa a la etapa S501.

50 Por el contrario, cuando no hay una imagen ya capturada (etapa S505: NO), el gráfico de barras 322 que indica la inclinación actual del dispositivo de imágenes de producto 121 obtenida en la etapa S502 se traza sobre la imagen de vista previa en vivo 321, obteniendo de este modo la imagen de buscador 331 (etapa S510).

55 A continuación, la unidad de visualización de buscador 204 muestra la imagen de buscador 331 en la pantalla de buscador configurada por la pantalla de cristal líquido, etc. (etapa S511), y el proceso regresa a la etapa S501.

De acuerdo con un ejemplo en el que se omite la unidad de sensor de inclinación 205, el proceso relacionado con la inclinación también se omite en consecuencia.

60 La inclinación del dispositivo de imágenes del producto 121 con relación a la gravedad se traza como los gráficos de barras 312 y 322 en la imagen de buscador 311.

65 Como se ha explicado anteriormente, de acuerdo con el dispositivo de captura de imágenes 121, la imagen del producto capturado previamente y la imagen de vista en vivo actual del producto se visualizan en la pantalla de buscador de manera superpuesta y, por lo tanto, es posible mantener la distancia entre el producto y el dispositivo de captura de imágenes 121 lo más constante que sea posible.

Además, de acuerdo con el ejemplo de detección de la inclinación, el dispositivo de captura de imágenes 121 se puede mover para que gire en una postura constante tanto como sea posible con respecto al eje de rotación predeterminado.

- 5 Por lo tanto, el usuario puede capturar fácilmente imágenes que se pueden convertir en la secuencia de imágenes de todo el entorno del producto con una alta calidad.

### 3. Dispositivo de conversión de imágenes

- 10 Como se ha explicado anteriormente, el uso del dispositivo de captura de imágenes 121 permite al usuario capturar sucesivamente imágenes del producto 301 mientras mueve el dispositivo de captura de imágenes 121 alrededor del producto 301 en una vuelta y mantiene la distancia con respecto al producto 301 y la inclinación con respecto a la dirección gravitacional tan constante como sea posible.

- 15 Cuando la pluralidad de imágenes ya capturadas 311 se obtienen a través de esta técnica de captura de imágenes, estas imágenes se normalizan para obtener la secuencia de imágenes de todo el entorno del producto 301. En este momento, se usa el dispositivo de conversión de imágenes 141 de esta realización.

- 20 La secuencia de imágenes de todo el entorno del producto 301 es una serie de imágenes obtenidas basándose en la suposición de que las imágenes del producto 301 se toman sucesivamente mientras se gira la cámara en torno al eje de rotación, manteniendo el ángulo desde el eje de rotación con respecto a la cámara a través del producto 301 constante, y manteniendo la distancia entre el producto 301 y la cámara constante con el eje de rotación que se presume que atraviesa el producto 301.

- 25 Por lo tanto, es probable que la pluralidad de imágenes ya capturadas 311 tomadas por el dispositivo de captura de imágenes 121 sea "la secuencia de imágenes de todo el entorno con un gran error". De acuerdo con el dispositivo de captura de imágenes 121, el ángulo y la distancia explicados anteriormente se mantienen tan constantes como sea posible para capturar imágenes del producto 301, pero la captura de imágenes por un ser humano causa un error inherente en el momento de la captura de imágenes.

- 30 El dispositivo de conversión de imágenes 141 convierte la pluralidad de imágenes ya capturadas 311, y obtiene la secuencia de imágenes de todo el entorno con un error lo más pequeño posible.

- 35 De acuerdo con esta realización, a diferencia de la Bibliografía 1 no perteneciente a patente, que establece un modelo tridimensional de un objeto para la captura de imágenes, la secuencia de imágenes de todo el entorno se puede obtener a través de cálculos sencillos. Esto se explicará a continuación con más detalle.

- 40 La Figura 6 es un diagrama explicativo que muestra una configuración general del dispositivo de conversión de imágenes 141 de acuerdo con esta realización. La explicación se dará con referencia a esta figura.

- El dispositivo de conversión de imágenes 141 de acuerdo con esta realización incluye una unidad de recepción de imágenes 601, y una unidad convertidora 602.

- 45 La unidad de recepción de imágenes 601 recibe la pluralidad de imágenes ya capturadas 311 tomadas y almacenadas por el dispositivo de imágenes de producto 121 de todo el entorno.

- 50 Cuando el dispositivo de conversión de imágenes 141 está configurado por un servidor web, etc., la persona que intenta capturar las imágenes carga las imágenes ya capturadas 311 tomadas por tal persona a través de Internet, etc. La unidad de recepción de imágenes 601 recibe esta carga.

- 55 Se puede llevar a cabo una realización en la que el dispositivo de imágenes de producto 121 transmite automáticamente las imágenes ya capturadas 311 al dispositivo de conversión de imágenes 141. De acuerdo con esta realización, el dispositivo de imágenes de producto 121 transmite automáticamente las imágenes ya capturadas 311 cada vez y se almacena un número predeterminado de imágenes ya capturadas 311 o no se realiza ninguna captura de imagen durante un tiempo predeterminado. Esta realización es especialmente adecuada cuando el dispositivo de imágenes de producto 121 se realiza por un ordenador portátil, tal como un PDA (asistente de datos personales) o un teléfono inteligente que ejecuta un programa predeterminado.

- 60 Cuando el dispositivo de captura de imágenes 121 está configurado por una cámara digital, etc., y el dispositivo de conversión de imágenes 141 está configurado por un ordenador personal, etc., la carga/descarga de una tarjeta de memoria entre estos dispositivos permite la transferencia de las imágenes ya capturadas 311.

- 65 Por el contrario, la unidad convertidora 602 convierte las imágenes múltiples ya capturadas 311 recibidas en la secuencia de imágenes de todo el entorno del producto 301.

La unidad convertidora 602 de esta realización incluye además una unidad de extracción 611, una unidad de estimación 612, una unidad de aproximación 613 y una unidad de salida 614.

5 La unidad de extracción 611 obtiene las imágenes ya capturadas 311 dos por dos sucesivamente. A continuación, la unidad de extracción extrae los puntos característicos correspondientes de las dos imágenes ya capturadas obtenidas 311. En este momento, se pueden aplicar diversos algoritmos de reconocimiento de imagen.

10 Cuando se obtienen los puntos característicos correspondientes, basándose en las ubicaciones de los puntos característicos en las dos imágenes ya capturadas 311, las posiciones y direcciones del dispositivo de imágenes de producto 121 cuando se tomaron las dos imágenes ya capturadas 311 se estiman, y se obtiene una matriz de transformación afín para normalizar cada una de las dos imágenes ya capturadas 311. La transformación afín aplicada en este proceso incluye la transformación de la perspectiva a través de la perspectiva de un punto además de la escala, rotación y transformación oblicua. Por lo tanto, cuando se expresa la matriz, se utiliza una coordenada homogénea según sea necesario.

15 En la siguiente explicación, para facilitar la comprensión a través de una simplificación de la notación, en la medida en que no se produce una confusión, una matriz original, vector y la matriz, teniendo el vector la dimensión de una constante aumentada y añadida para que sea la coordenada homogénea, se representan por el mismo símbolo en algunos casos.

20 A continuación, la unidad de aproximación 613 aplica la transformación afín a las imágenes ya capturadas 311 basándose en las matrices de transformación afín obtenidas para cada una de las dos imágenes ya capturadas 311 para normalizar estas imágenes. La imagen normalizada se aproxima a la imagen que debe incluirse en la secuencia de imágenes de todo el entorno.

La unidad de salida 614 genera las dos imágenes normalizadas obtenidas y cuánto difiere el ángulo de captura de imagen en torno al eje de rotación entre las dos imágenes.

30 El flujo de un control para un proceso de conversión de imágenes ejecutado por el dispositivo de conversión de imágenes 141 se explicará a continuación con más detalle.

35 La Figura 7 es un diagrama de flujo que muestra un flujo de un control para el proceso de conversión de imágenes ejecutado por el dispositivo de conversión de imágenes 141 de acuerdo con esta realización. La explicación se dará con referencia a esta figura.

40 En primer lugar, la unidad de recepción de imágenes 601 recibe la pluralidad de imágenes ya capturadas 311 tomadas y almacenadas por el dispositivo de imágenes de producto 121 y la inclinación asociada con cada imagen ya capturada 311 (etapa S701). En lo sucesivo en el presente documento, para facilitar la comprensión, se presume que el número de imágenes ya capturadas 311 es  $N$ , y cada imagen ya capturada se expresará como  $G[0]$ ,  $G[1]$ , ... y  $G[N-1]$ . Además, la inclinación cuando se tomó cada imagen  $G[0]$ ,  $G[1]$ , ... y  $G[N-1]$ , se expresará como  $R[0]$ ,  $R[1]$ , ... y  $R[N-1]$ .

45 Se presume a continuación que el índice expresado por los corchetes se expresa mediante el conjunto complementario de  $N$ . Es decir, cuando el número entero del índice es negativo y es igual o mayor de  $N$ , un resto obtenido al dividir tal número entero por  $N$  se utiliza como el valor del índice. Por ejemplo, la inclinación  $R[N]$  tiene el mismo significado que la inclinación  $R[0]$ , y la imagen  $G[-1]$  tiene el mismo significado que la imagen  $R[N-1]$ .

50 La inclinación  $R[0]$ ,  $R[1]$ , ... y  $R[N-1]$  es información para determinar la dirección de cada eje de coordenadas del sistema de coordenadas de la cámara que está fijado al dispositivo de imágenes de producto 121. Por lo tanto, el grado de libertad para cada inclinación  $R[0]$ ,  $R[1]$ , ... y  $R[N-1]$  es 6.

55 Sin embargo, es típico expresar mediante una matriz de  $3 \times 3$  que tiene vectores unitarios que representan direcciones de los ejes de coordenadas respectivos dispuestos.

A continuación, se cambia un número entero  $k$  entre 0 y  $N$  para repetir el siguiente proceso (etapa S702).

60 Es decir, la unidad de extracción 611 extrae los puntos característicos correspondientes de la imagen  $G[k]$  y la imagen  $G[k+1]$  (etapa S703).

65 A continuación, la unidad de estimación 612 selecciona tres conjuntos cualesquiera entre los conjuntos de puntos característicos extraídos (etapa S704). La selección de los tres conjuntos es opcional, pero cuando la selección se hace de tal forma que la distancia entre los puntos correspondientes seleccionados es excelente en la imagen  $G[k]$  y tal distancia entre los puntos correspondientes seleccionados también es mayor en la imagen  $G[k+1]$ , un

error puede suprimirse. Se pueden seleccionar tres puntos característicos cualesquiera de los puntos característicos cerca del lado superior, el lado inferior, el centro del lado izquierdo y el centro del lado derecho del producto.

5 En la siguiente explicación, los tres puntos característicos se conocen como P, Q y W, respectivamente. Las posiciones donde estos puntos característicos se trazan en la imagen G[k] se expresan como p[k], q[k] y w[k], y las posiciones donde estos puntos característicos se trazan en la imagen G[k+1] se expresan como p[k+1], q[k+1] y w[k+1].

10 La unidad de estimación 612 estima lo siguiente (etapa S705):

(x) La posición T[k] del dispositivo de imágenes de producto 121 cuando se tomó la imagen G[k]; y  
 (y) La posición T[k+1] del dispositivo de imágenes de producto 121 cuando se tomó la imagen G[k+1], desde:

15

- (1) Las posiciones p[k] y p[k+1] del punto característico p en las imágenes G[k] y G[k+1];
- (2) las posiciones q[k] y q[k+1] del punto característico Q en las imágenes G[k] y G[k+1];
- (3) las posiciones w[k] y w[k+1] del punto característico W en las imágenes G[k] y G[k+1];
- (4) la inclinación R[k] del dispositivo de imágenes de producto 121 cuando se tomó la imagen G[k]; y
- (5) la inclinación R[k+1] del dispositivo de imágenes de producto 121 cuando se tomó la imagen G[k+1].

20

El detalle de este algoritmo de estimación se explicará a continuación con más detalle.

25 En general, la captura de imágenes a través de una cámara utiliza una transformación de coordenadas entre un sistema de coordenadas mundo fijo con respecto al mundo real y un sistema de coordenadas de cámara fijado a la cámara. El sistema de coordenadas mundo y el sistema de coordenadas de cámara tienen tres ejes de coordenadas ortogonales entre sí.

30 Los vectores direccionales de los tres ejes de coordenadas en el sistema de coordenadas de cámara se representarán como rx, ry y rz, respectivamente (Fórmula 1 a Fórmula 3).

[Fórmula 1]

35

$$\mathbf{rx} = \begin{bmatrix} rx1 \\ rx2 \\ rx3 \end{bmatrix}$$

40

[Fórmula 2]

45

$$\mathbf{ry} = \begin{bmatrix} ry1 \\ ry2 \\ ry3 \end{bmatrix}$$

50

[Fórmula 3]

55

$$\mathbf{rz} = \begin{bmatrix} rz1 \\ rz2 \\ rz3 \end{bmatrix}$$

60

65

Los valores respectivos  $rx_1$ ,  $rx_2$ , ... y  $rz_3$  son todos valores de coordenadas en el sistema de coordenadas mundo, y representan los vectores  $rx$ ,  $ry$  y  $rz$ .

Además, un vector posicional en el origen del sistema de coordenadas de cámara en relación con el origen del sistema de coordenadas mundo se representará como  $T$  (Fórmula 4).

[Fórmula 4]

$$\mathbf{T} = \begin{bmatrix} t_1 \\ t_2 \\ t_3 \end{bmatrix}$$

Los valores  $t_1$ ,  $t_2$  y  $t_3$  son valores de coordenadas del vector posicional  $T$  en el sistema de coordenadas mundo.

En dichas condiciones, se presume que un vector  $pc$  que dispone los valores de coordenadas de un punto  $P$  en el sistema de coordenadas de cámara se expresa como se indica a continuación (Fórmula 5).

[Fórmula 5]

$$pc = \begin{bmatrix} pc_1 \\ pc_2 \\ pc_3 \end{bmatrix}$$

En este caso, un vector  $pw$  que dispone los valores de coordenadas del punto  $P$  en el sistema de coordenadas mundo puede expresarse como se indica a continuación (Fórmula 6).

[Fórmula 6]

$$\begin{aligned} pw &= \begin{bmatrix} pw_1 \\ pw_2 \\ pw_3 \end{bmatrix} \\ &= \begin{bmatrix} rx & ry & rz \end{bmatrix} pc + T \\ &= R pc + T \end{aligned}$$

Una matriz  $R$  es una matriz que expresa la dirección del eje de coordenadas del sistema de coordenadas de cámara. Es decir, esta matriz se puede obtener a partir de los resultados de detección del sensor de gravedad y la brújula, etc., de la unidad de sensor de inclinación 205 cuando se realizó la captura de imágenes a través del dispositivo de imágenes de producto 121.

De acuerdo con esta fórmula, un movimiento paralelo se expresa mediante la adición, pero cuando se usa una expresión de coordenadas homogénea, solo se puede expresar mediante la multiplicación de la matriz (Fórmula 7).

[Fórmula 7]

$$\begin{aligned} \begin{bmatrix} pw \\ 1 \end{bmatrix} &= \begin{bmatrix} R & T \\ 0^t & 1 \end{bmatrix} \begin{bmatrix} pc \\ 1 \end{bmatrix} = D \begin{bmatrix} pc \\ 1 \end{bmatrix}, \\ \begin{bmatrix} pc \\ 1 \end{bmatrix} &= D^{-1} \begin{bmatrix} pw \\ 1 \end{bmatrix} = \begin{bmatrix} R^t & -R^t T \\ 0^t & 1 \end{bmatrix} \begin{bmatrix} pw \\ 1 \end{bmatrix} \end{aligned}$$

En la fórmula anterior,  $R^t$  significa una matriz transpuesta de la matriz  $R$ . Dado que  $R$  es una matriz unitaria,  $R^{-1} = R^t$  se cumple. Además,  $0^t$  es un vector nulo cero lateral tridimensional.

Se supone que existe un plano de proyección ortogonal al eje  $z$  del sistema de coordenadas de cámara y distante del origen en una distancia focal  $f$ . Un resultado proyectado en el plano de proyección corresponde a la imagen ya capturada 311. Por lo tanto, un vector  $u$  que dispone los valores de coordenadas en un sistema de coordenadas bidimensional fijo con respecto al plano de proyección se expresa como se indica a continuación (Fórmula 8).

[Fórmula 8]

$$pu = \begin{bmatrix} pu1 \\ pu2 \end{bmatrix}$$

En este caso, la ubicación donde se proyecta el punto  $P$  en el plano de proyección se puede expresar como se indica a continuación (Fórmula 9).

[Fórmula 9]

$$\begin{aligned} pu1 &= f \frac{pc1}{pc3}, \\ pu2 &= f \frac{pc2}{pc3} \end{aligned}$$

Esto se reescribirá como se indica a continuación (Fórmula 10).

[Fórmula 10]

$$\begin{aligned} pc3 \begin{bmatrix} pu1 \\ pu2 \\ 1 \end{bmatrix} &= \begin{bmatrix} f & 0 & 0 & 0 \\ 0 & f & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} pc1 \\ pc2 \\ pc3 \\ 1 \end{bmatrix}, \\ pc3 \begin{bmatrix} pu \\ 1 \end{bmatrix} &= A \begin{bmatrix} pc \\ 1 \end{bmatrix} \end{aligned}$$

Una condición en la que dos vectores están en una condición proporcional por una constante proporcional distinta de cero se indicará por  $\equiv$  (Fórmula 11).

[Fórmula 11]

$$\begin{bmatrix} pu \\ 1 \end{bmatrix} \equiv A \begin{bmatrix} pw \\ 1 \end{bmatrix}$$

5 Cuando la imagen del punto P que tiene las coordenadas pw en el sistema de coordenadas mundo se toma por el dispositivo de imágenes de producto 121 en la posición T y la inclinación R, las coordenadas pu donde se proyecta tal punto en la imagen ya capturada 311 se puede expresar como se indica a continuación (Fórmula 12).

[Fórmula 12]

$$\begin{bmatrix} pu \\ 1 \end{bmatrix} \equiv A \begin{bmatrix} R^t & -R^t T \\ 0^t & 1 \end{bmatrix} \begin{bmatrix} pw \\ 1 \end{bmatrix} = M \begin{bmatrix} pw \\ 1 \end{bmatrix}$$

20 En la fórmula anterior, M es una matriz de 3 x 4.

Este plano de proyección se basa en la suposición de una cámara estenopeica ideal, y se fija al sistema de coordenadas de cámara. El primer eje del sistema de coordenadas bidimensional es paralelo al eje x, y el segundo eje del sistema de coordenadas bidimensional es paralelo al eje y.

25 Cuando se realiza una proyección general que no siempre satisface esta condición, se puede usar una matriz A dependiendo de la situación.

30 Cuando, por ejemplo, el plano de proyección y el eje z del sistema de coordenadas de cámara tienen un ángulo  $\psi$ , la matriz A se puede expresar como se indica a continuación usando las constantes apropiadas k1, k2, uc1 y uc2 (Fórmula 13).

[Fórmula 13]

$$A = \begin{bmatrix} f k1 & -f k1 \cot \psi & uc1 & 0 \\ 0 & f k2 / \sin \psi & uc2 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}$$

En un caso en el que, por ejemplo, se usa una lente ojo de pez, se puede obtener una matriz de transformación afín M similar a la matriz explicada anteriormente definiendo la matriz A de manera apropiada.

45 Se dará una explicación de una técnica para obtener las posiciones de captura de imagen T[k] y T[k+1] basándose en la relación de captura de imagen explicada anteriormente. Cada una de estas posiciones tiene tres elementos y, por lo tanto, el número de incógnitas es seis en total.

50 En primer lugar, una matriz de transformación afín con respecto a la posición de captura de imágenes T[k] y la inclinación R[k] se define como M[k] y una matriz de transformación afín con respecto a la posición de captura de imágenes T[k+1] y la inclinación R[k+1] se define como M[k+1]. En este caso, se cumple la siguiente relación (Fórmula 14).

[Fórmula 14]

$$\begin{aligned} M[k]^{-1} \begin{bmatrix} p[k] \\ 1 \end{bmatrix} &\equiv M[k+1]^{-1} \begin{bmatrix} p[k+1] \\ 1 \end{bmatrix}, \\ M[k]^{-1} \begin{bmatrix} q[k] \\ 1 \end{bmatrix} &\equiv M[k+1]^{-1} \begin{bmatrix} q[k+1] \\ 1 \end{bmatrix}, \\ M[k]^{-1} \begin{bmatrix} w[k] \\ 1 \end{bmatrix} &\equiv M[k+1]^{-1} \begin{bmatrix} w[k+1] \\ 1 \end{bmatrix} \end{aligned}$$

Sin embargo, dado que  $M[k]$  y  $M[k+1]$  son cada uno una matriz de  $3 \times 4$ ,  $M[k]^{-1}$  y  $M[k+1]^{-1}$  son matrices inversas generalizadas para  $M[k]$  y  $M[k+1]$ , respectivamente. La matriz inversa generalizada a veces se llama "matriz pseudoinversa de Moore y Penrose".

5 El número de incógnitas es nueve en el total, que son seis valores de coordenadas de las posiciones de captura de imagen  $T[k]$  y  $T[k+1]$  y tres coeficientes proporcionales del lado izquierdo y el lado derecho en cada condición  $\Xi$ .

10 Por el contrario, la ecuación matricial tiene tres incógnitas tridimensionales y, por lo tanto, su número total es nueve. Por lo tanto, todas las incógnitas se pueden obtener resolviendo estas ecuaciones simultáneas.

15 Al resolver las ecuaciones simultáneas, resulta adecuado si una solución se obtiene a través de una manipulación de fórmula basada en la manipulación de un símbolo, o se pueden aplicar varios solucionadores aproximados. Por ejemplo, se puede aplicar una técnica que dé valores iniciales apropiados a  $T[k]$ ,  $T[k+1]$ , y la constante proporcional y obtenga una solución a través del procedimiento de descenso más pronunciado.

20 Además, cuando aumenta el número de puntos característicos, el número de ecuaciones simultáneas se puede aumentar adicionalmente. En este caso, estas incógnitas se pueden obtener realizando una estimación de máxima verosimilitud para minimizar la suma cuadrada de un error obtenido al restar el lado derecho de cada ecuación simultánea del lado izquierdo de la misma.

25 Como se ha explicado anteriormente, las posiciones de captura de imagen  $T[k]$  y  $T[k+1]$  se pueden identificar a través de la selección de los tres puntos característicos correspondientes en la imagen. Por otro lado, si se identifican las posiciones de captura de imagen  $T[k]$  y  $T[k+1]$ , pueden obtenerse las matrices de transformación afín  $M[k]$  y  $M[k+1]$  en la proyección en perspectiva por perspectiva de un punto.

30 Cuando esos procesos se repiten (etapa S707), los valores estimados respectivos de  $T[0]$ ,  $T[1]$ ,  $T[2]$ , ... y  $T[N-1]$  se pueden obtener a través de la etapa S705 (x), y los valores estimados respectivos de  $T[1]$ ,  $T[2]$ , ...  $T[N-1]$ , y  $T[0]$  se pueden obtener a través de la etapa S705 (y).

Por lo tanto, seleccionando uno de estos u obteniendo un promedio de estos, las posiciones  $T[0]$ ,  $T[1]$ ,  $T[2]$ , ... y  $T[N-1]$  del dispositivo de imágenes de producto 121 pueden identificarse (etapa S707).

35 Además, se puede identificar una posición H del producto 301 en el mundo real a través de la siguiente técnica (etapa S708).

40 Es decir, para cada uno de los números enteros  $k = 0, 1, 2, \dots$  y  $N-1$ , la dirección del eje óptico de la captura de imagen se obtiene a partir de la inclinación  $R[k]$ . A continuación, se simulan las líneas rectas que atraviesan la posición de captura de imagen  $T[k]$  y se dirigen hacia la dirección del eje óptico. Estos números N de líneas rectas deben atravesar de la posición H del producto 301 o las proximidades del mismo.

45 Por lo tanto, la posición H del producto 301 se puede obtener a través de la técnica de mínimos cuadrados que minimiza la suma cuadrada de la distancia entre la posición H del producto 301 y el número N de líneas rectas.

50 Cuando la captura de imágenes explicada anteriormente se realiza a través del dispositivo de imágenes de producto 121, las posiciones de captura de imagen  $T[0]$ ,  $T[1]$ ,  $T[2]$ , ... y  $T[N-1]$  posiblemente estén posicionados circularmente en torno al eje central que atraviesa la posición H del producto 301. Además, cuando un ser humano captura imágenes del producto, no siempre es cierto que las imágenes se capturan desde el lado izquierdo del producto, y las imágenes a menudo se capturan oblicuamente por encima.

55 Es decir, en un caso ideal, las posiciones de captura de imagen estimadas  $T[0]$ ,  $T[1]$ ,  $T[2]$ , ... y  $T[N-1]$  están dispuestas sobre la cara inferior de una proyección cónica circular que tiene la posición H como vértice. Además, cuando las imágenes se capturan desde el lado derecho, las posiciones de captura de imagen estimadas  $T[0]$ ,  $T[1]$ ,  $T[2]$ , ... y  $T[N-1]$  están dispuestas sobre la circunferencia que tiene la posición H como centro.

60 Sin embargo, la captura de imágenes se realiza por un ser humano, lo que con frecuencia causa errores. En consecuencia, al normalizar las posiciones de captura de imagen  $T[0]$ ,  $T[1]$ ,  $T[2]$ , ... y  $T[N-1]$ , las posiciones normalizadas  $S[0]$ ,  $S[1]$ ,  $S[2]$ , ... y  $S[N-1]$  se identifican sobre una circunferencia precisa que es la parte inferior de una proyección cónica circular verdadera que tiene la posición H como centro (etapa S709). Cuando el ángulo del vértice de esta proyección cónica circular verdadera se establece en 180 grados, es posible hacer frente a un caso en el que las posiciones normalizadas  $S[0]$ ,  $S[1]$ ,  $S[2]$ , ... y  $S[N-1]$  están dispuestas sobre la circunferencia precisa alrededor de la posición H.

65 Para esta identificación, más específicamente, se realizan los siguientes cálculos.

En primer lugar, las distancias de las respectivas posiciones de captura de imagen  $T[0]$ ,  $T[1]$ ,  $T[2]$ , ... y  $T[N-1]$  distantes de la posición  $H$  se hacen equivalentes entre sí. Es decir, para cada uno de los números enteros  $k = 0, 1, 2, \dots$  y  $N-1$ , el coeficiente de escala  $sc[k]$  para igualar la distancia y una posición  $U[k] = T[k] + sc[k](T[k] - H)/|T[k]-H|$  se calculan para que  $sc[0]$ ,  $sc[1]$ ,  $sc[2]$ , ... y  $sc[N-1]$  sean lo más pequeños posible.

Es decir, la posición  $U[k]$  está ubicada sobre una línea recta que interconecta la posición  $H$  del producto 301 con la posición de captura de imagen  $T[k]$ , y la posición  $U[k]$  y la captura de imagen la posición  $T[k]$  están separadas entre sí por  $sc[k]$ . En este momento, la minimización de  $sc[0]$ ,  $sc[1]$ ,  $sc[2]$ , ... y  $sc[N-1]$  se toma como condición de restricción.

De la forma más sencilla, una distancia promedio  $D$  a las respectivas posiciones de captura de imagen  $T[0]$ ,  $T[1]$ ,  $T[2]$ , ... y  $T[N-1]$  se obtiene a partir de

$$D = \sum_{k=0}^{N-1} |T[k]-H|/N,$$

y cuando se establece que sea

$$sc[k] = D - |T[k]-H|,$$

cada  $|U[k]-H|$  se vuelve igual a la distancia  $D$ .

A continuación, se calculan una posición  $J$  del centro de la circunferencia que es la parte inferior de una proyección cónica circular y un radio  $r$  de la misma. Es más sencillo si la posición central  $J$  de la circunferencia se toma como el centro promediado de la posición  $U[k]$  como

$$J = \sum_{k=0}^{N-1} U[k]/N.$$

A continuación, es más sencillo si el radio  $r$  de la circunferencia se toma como el promedio de las distancias entre la posición  $J$  y el  $U[k]$  respectivo como

$$r = \sum_{k=0}^{N-1} |U[k]-J|/N.$$

La posición del producto 301 y la trayectoria de la circunferencia cuando se toma la secuencia de imágenes de todo el entorno del producto 301 se definen de esta manera. Esta circunferencia tiene una línea recta que interconecta la posición  $H$  con la posición  $J$  como un eje de rotación.

Por lo tanto, para cada  $k = 0, 1, 2, \dots$  y  $N-1$ , a continuación, se identifica una posición  $S[k]$  sobre la circunferencia con respecto a la posición de captura de imágenes  $T[k]$ . La posición  $S[k]$  obtenida aquí corresponde a una posición de captura de imagen virtual obtenida al normalizar la posición de captura de imagen  $T[k]$ .

La técnica más sencilla es tomar la posición más cercana a la posición de captura de imagen  $T[k]$  sobre la circunferencia como  $S[k]$ .

En primer lugar, a partir de un vector  $H-J$  y un vector  $T[k]-J$ , se calcula un vector  $OP[k] = ((H-J) \times (T[k]-J)) \times (H-J)$ .

El vector  $(H-J \times T[k]-J)$  es ortogonal tanto al vector  $H-J$  como al vector  $T[k]-J$ .

Por consiguiente, el vector  $OP[k] = ((H-J) - (T[k]-J)) \times (H-J)$  es paralelo a un plano formado por el vector  $H-J$  y el vector  $T[k]-J$ , y es ortogonal al vector  $H-J$ .

Es decir, una línea media que se extiende hasta una parte de tramo extendida desde  $T[k]$  con respecto al plano formado por la circunferencia desde la posición  $J$  y el vector  $OP[k]$  transcurre en la misma dirección en paralelo entre sí.

Por lo tanto, la posición normalizada  $S[k]$  de un punto sobre la circunferencia y más cercana a la posición de captura de imagen  $T[k]$  se puede obtener a partir de:

$$S[k] = J + r OP[k]/|OP[k]|.$$

Como se ha explicado anteriormente, las posiciones normalizadas  $S[0]$ ,  $S[1]$ , ... y  $S[N-1]$  están dispuestas sobre la circunferencia predeterminada. Por consiguiente, también es posible obtener cuánto gira cada uno de  $S[0]$ ,  $S[1]$ , ... y  $S[N-1]$  con respecto al centro  $J$  sobre la circunferencia, tal como se ve desde  $S[0]$ .

Es decir, se supone que los ángulos de rotación respectivos de las posiciones normalizadas S[0], S[1], ... y S[N-1] son  $\theta[0]$ ,  $\theta[1]$ , ... y  $\theta[N-1]$ . Para cada uno de los números enteros  $k = 0, 1, \dots$  y  $N-1$ ,  $\theta[k]$  es un ángulo de un vector S[k]-J con relación a un vector S[0]-J. Es decir, cada ángulo de rotación  $\theta[k]$  se puede obtener a partir de:

$$\begin{aligned} \cos(\theta[k]) &= \langle\langle S[0]-J, S[k]-J \rangle\rangle / (|S[0]-J| \cdot |S[k]-J|), \\ \text{sen}(\theta[k]) &= |(S[0]-J) \times (S[k]-J)| / (|S[0]-J| \cdot |S[k]-J|). \end{aligned}$$

"x, y" es un producto interno de un vector x y un vector y.

Si se supone que  $\theta[k]$  obtenido de esta manera están dispuestos en orden ascendente como

$$0 = \theta[0] \leq \theta[1] \leq \theta[2] \leq \dots \leq \theta[N-2] < 2\pi,$$

no se pierde ninguna generalidad. Cuando, para una cierta k,  $\theta[k] > \theta[k+1]$  se satisface y se cambia el orden,

- (a) La imagen G[k], la inclinación R[k], la posición de captura de imagen T[k], la posición normalizada S[k] y el ángulo de rotación  $\theta[k]$  y
- (b) La imagen G[k+1], la inclinación R[k+1], la posición de captura de imagen T[k+1], la posición normalizada S[k+1] y el ángulo de rotación  $\theta[k+1]$

pueden intercambiarse en su conjunto.

A continuación, para cada uno de los números enteros  $k = 0, 1, 2, \dots$  y  $N-1$ , se calcula una inclinación V[k] de una captura de imagen virtual en una posición virtualizada S[k] (etapa S710). Más específicamente, es típico que un vector H-S[k] dirigido a la posición H del producto 301 desde la posición virtualizada S[k] se toma como una dirección óptica (eje z) de la captura de imagen, una dirección de la línea tangente de la circunferencia en la posición virtualizada S[k] se toma como la dirección horizontal (eje x) de la captura de imagen, y un componente de un vector J-S[k] dirigido a la posición central J de la circunferencia desde la posición virtualizada S[k] y ortogonal a H-S[k] se toma como una dirección vertical (eje y) de la captura de imagen. Al establecer esas dos direcciones, se determinan los seis grados de libertad para el sistema de coordenadas de cámara. La inclinación V[k] de la captura de imagen en la posición virtualizada S[k] se determina automáticamente.

A continuación, como se ha explicado anteriormente, el punto P dispuesto en una posición pw en el espacio real se proyecta con respecto a una posición pu de la imagen G[k] que satisface la siguiente condición cuando se toma en la posición de captura de imagen T[k] en una dirección R[k] (Fórmula 15).

[Fórmula 15]

$$\begin{bmatrix} pu \\ 1 \end{bmatrix} \equiv M[k] \begin{bmatrix} pw \\ 1 \end{bmatrix}$$

La constante proporcional de esta fórmula es, como también se expresa en (Fórmula 10), el valor de coordenadas pc3 del punto P a lo largo del eje z en el sistema de coordenadas de cámara, es decir, el eje óptico de la dirección de captura de imagen.

Por lo tanto, cuando el dispositivo de imágenes de producto 121 captura imágenes del producto 301 que tiene el punto P dispuesto sobre una superficie, la distancia entre el producto 301 y el dispositivo de imágenes de producto 121 se puede tomar como un valor aproximado de pc3 (Fórmula 16) como

$$d[k] = |T[k]-H|.$$

[Fórmula 16]

$$d[k] \begin{bmatrix} pu \\ 1 \end{bmatrix} \approx M[k] \begin{bmatrix} pw \\ 1 \end{bmatrix}$$

Por el contrario, se puede obtener una posición  $pw$  del punto P proyectado en la posición  $pu$  en la imagen  $G[k]$  en el espacio real a través del siguiente cálculo aproximado (Fórmula 17).

[Fórmula 17]

5

$$\begin{bmatrix} pw \\ 1 \end{bmatrix} \approx \frac{1}{d[k]} M[k]^{-1} \begin{bmatrix} pu \\ 1 \end{bmatrix}$$

10

Asimismo, se supone que el punto P dispuesto en la posición  $pc$  en el mundo real se proyecta en una posición  $pu'$  en una imagen normalizada  $B[k]$  en el ángulo de rotación  $\theta[k]$ .

15

La posición en la que se toma la imagen normalizada  $B[k]$  es la parte inferior de la proyección cónica circular verdadera, y la distancia  $d = |S[k]-H|$  entre cada posición normalizada  $S[k]$  y un vértice H es constante independientemente de k, y se satisface la siguiente relación (Fórmula 18).

[Fórmula 18]

20

$$d \begin{bmatrix} pu' \\ 1 \end{bmatrix} \approx L[k] \begin{bmatrix} pw \\ 1 \end{bmatrix},$$

$$\begin{bmatrix} pw \\ 1 \end{bmatrix} \approx dL[k]^{-1} \begin{bmatrix} pu' \\ 1 \end{bmatrix}$$

25

30

Es decir, la posición  $pu'$  en la imagen normalizada  $B[k]$  corresponde a la posición  $pu$  en la imagen ya capturada  $G[k]$ , pero la posición  $pu$  se puede aproximar de la siguiente manera (Fórmula 19).

[Fórmula 19]

35

$$\begin{bmatrix} pu \\ 1 \end{bmatrix} \approx \frac{d}{d[k]} M[k] L[k]^{-1} \begin{bmatrix} pu' \\ 1 \end{bmatrix}$$

40

Por consiguiente, para cada posición en la imagen normalizada  $B[k]$ ,

45

- (1) la posición de un píxel de la imagen ya capturada  $G[k]$  se obtiene a partir de la relación explicada anteriormente,
- (2) se obtiene un valor de píxel de la posición obtenida en la imagen ya capturada  $G[k]$ , y
- (3) el valor de píxel obtenido se traza como el valor de píxel de esa posición en la imagen normalizada  $B[k]$ ,

obteniendo de este modo la imagen normalizada  $B[k]$ .

50

A través de tal algoritmo, la unidad de aproximación 613 calcula, para cada uno de los números enteros  $k = 0, 1, 2, \dots$  y  $N-1$ , una matriz de transformación afín  $L[k]$  que transforma la posición en el real mundo en la posición en la imagen cuando la captura de imagen se realiza en la posición virtualizada  $S[k]$  en una inclinación  $V[k]$  (etapa S711).

A continuación, la transformación por una matriz de transformación afín

55

$$(d/d[k])M[k]L[k]^{-1}$$

60

se realiza en la imagen  $G[k]$ , y se obtiene la imagen normalizada  $B[k]$  con respecto al ángulo de rotación  $\theta[k]$  incluido en la secuencia de imágenes de toda la superficie exterior visible (etapa S712).

65

Cuando la secuencia de imágenes de toda la superficie exterior visible tiene las imágenes normalizadas  $B[0], B[1], B[2], \dots$  y  $B[N-1]$  dispuestas y los ángulos de rotación  $\theta[0], \theta[1], \theta[2], \dots$  y  $\theta[N-1]$  en relación con las imágenes normalizadas respectivas se obtienen de esta manera, la unidad de salida 614 envía éstas como información sobre la secuencia de imágenes de toda la superficie exterior visible (etapa S713), y finaliza el proceso.

Es deseable en la etapa S719 generar, como parte de la información en la secuencia de imágenes de toda la superficie exterior visible, las matrices de transformación afín  $(1/d)L[0]$ ,  $(1/d)L[1]$ ,  $(1/d)L[2]$ , ... y  $(1/d)L[N-1]$  que transforman las posiciones en el mundo real en las posiciones en las imágenes normalizadas.

5 Los fragmentos de información en la secuencia de imágenes de salida de toda la superficie exterior visible, el ángulo de rotación, y la matriz de transformación afín se almacenan típicamente en el disco duro, etc., del dispositivo servidor, pero se pueden descargar al teléfono inteligente, el teléfono celular, y el ordenador personal, etc., desde tal dispositivo servidor.

10 De acuerdo con la explicación anterior, las imágenes del producto 301 se toman mientras el dispositivo de imágenes de producto 121 se mueve en la circunferencia sustancial de la parte inferior de la proyección cónica circular sustancial que tiene el producto 301 como vértice, obteniendo de este modo las imágenes capturadas.

15 A continuación, el error contenido en los resultados de captura de imagen se corrige, y las imágenes normalizadas que posiblemente se obtienen cuando se toman las imágenes del producto 301 mientras el dispositivo de imágenes de producto 121 se mueve alrededor de la verdadera circunferencia de la parte inferior de la proyección cónica circular verdadera que tiene el producto 301 como el vértice se obtienen como resultado.

20 Para simplificar el proceso, está bien suponer que el dispositivo de imágenes de producto 121 se mueve alrededor de la circunferencia sustancial alrededor del producto 301, y la imagen normalizada representa una apariencia del producto como se ve desde la circunferencia verdadera alrededor del producto 301. De acuerdo con este esquema, la condición de restricción se simplifica notablemente en comparación con la explicación anterior, y se espera que la velocidad de cálculo aumente adicionalmente.

25 Además, la posición del producto 301 y la forma de la proyección cónica circular se pueden definir a través de diversas técnicas de estimación de máxima verosimilitud.

30 En la siguiente explicación, se explicará una técnica para obtener las posiciones de captura de imagen,  $T[k]$  y  $T[k+1]$ , la inclinación  $R[k]$  y  $R[k+1]$ , y las matrices de transformación afín  $M[k]$  y  $M[k+1]$  a través de un cálculo de aproximación más sencillo adicional cuando se usa el dispositivo de imágenes de producto 121 de esta realización.

35 Como se ha explicado anteriormente, en la transformación de la (Fórmula 10) en la (Fórmula 11), se elimina la constante proporcional  $pc3$ . La constante proporcional  $pc3$  representa el valor de coordenadas de un punto de un objeto de captura de imagen en el eje óptico de la dirección de captura de imagen, es decir, una profundidad del punto del objeto de captura de imagen de la cámara.

40 Por el contrario, cuando se usa el dispositivo de imágenes de producto 121 de esta realización, la captura de imágenes se realiza de tal manera que el tamaño del producto 301 se vuelve sustancialmente constante a lo largo de las imágenes ya capturadas. Por consiguiente, se puede pensar que la profundidad del producto 301 es sustancialmente igual a la vista desde cualquier posición de captura de imagen.

45 En este caso, la relación proporcional "≡" en la (Fórmula 14) puede reemplazarse con el signo igual "=". Después, se hace innecesario obtener la constante proporcional que es desconocida.

Por lo tanto, siempre que la profundidad del producto 301 sea igual, desde dos posiciones características de los puntos  $p$  y  $q$ , las posiciones de captura de imagen  $T[k]$  y  $T[k+1]$ , la inclinación  $R[k]$  y  $R[k+1]$ , y las matrices de transformación afín  $M[k]$  y  $M[k+1]$  se pueden obtener a través de las siguientes ecuaciones simultáneas.

50 [Fórmula 20]

$$\begin{aligned}
 M[k]^{-1} \begin{bmatrix} p[k] \\ 1 \end{bmatrix} &= M[k+1]^{-1} \begin{bmatrix} p[k+1] \\ 1 \end{bmatrix}, \\
 M[k]^{-1} \begin{bmatrix} q[k] \\ 1 \end{bmatrix} &= M[k+1]^{-1} \begin{bmatrix} q[k+1] \\ 1 \end{bmatrix}
 \end{aligned}$$

60 Es decir, el número de incógnitas es seis que son los valores de coordenadas de las posiciones de captura de imagen  $T[k]$  y  $T[k+1]$ , y hay dos ecuaciones matriciales de tres dimensiones, y el número total es seis. Por consiguiente, las posiciones de captura de imagen  $T[k]$  y  $T[k+1]$  se pueden obtener por cálculo.

65 Con respecto al cálculo específico, se pueden aplicar técnicas similares a las de la realización explicada anteriormente.

Cuanto más aumenta el número de puntos característicos a remitirse para obtener las posiciones de captura de imagen, más aumenta la carga de cálculo, pero más aumenta la precisión de cálculo. Además, cuando el número de puntos característicos se establece para que sea igual o superior a tres, las posiciones de captura de imagen  $T[k]$  y  $T[k+1]$ , la inclinación  $R[k]$  y  $R[k+1]$ , y las matrices de transformación afín  $M[k]$  y  $M[k+1]$  pueden obtenerse sin ninguna precondition específica para la posición del producto 301 en relación con cada posición de captura de imagen  $T[k]$ .

Por el contrario, como esta realización, cuando se puede pensar que la distancia entre el producto 301 y cada posición de captura de imagen  $T[k]$  es sustancialmente igual, si el número de puntos característicos se establece en dos, la carga de cálculo se puede reducir, mientras que, al mismo tiempo, las posiciones de captura de imagen  $T[k]$  y  $T[k+1]$ , la inclinación  $R[k]$  y  $R[k+1]$ , y las matrices de transformación afín  $M[k]$  y  $M[k+1]$  pueden obtenerse con gran precisión aunque se aplique la aproximación.

#### 4. Dispositivo de procesamiento de imágenes

Las imágenes normalizadas  $B[0]$ ,  $B[1]$ ,  $B[2]$ , ... y  $B[N-1]$  incluidas en la secuencia de imágenes de toda la superficie exterior visible están asociadas con los ángulos de rotación  $\theta[0]$ ,  $\theta[1]$ ,  $\theta[2]$ , ... y  $\theta[N-1]$  alrededor del producto 301, respectivamente. Como se ha explicado anteriormente, si se establece una condición

$$0 = \theta[0] \leq \theta[1] \leq \theta[2] \leq \dots \leq \theta[N-1] < 2\pi$$

no se pierde ninguna generalidad.

Cuando el usuario da un ángulo de rotación arbitrario  $\phi$ , el dispositivo de procesamiento de imágenes 161 tiene la función de interpolar la imagen del producto 301 según se ve desde la posición correspondiente al ángulo de rotación  $\phi$  utilizando la información explicada anteriormente, y de presentar la imagen interpolada al usuario. La técnica de interpolación más sencilla es como se indica a continuación.

Como se ha explicado anteriormente, desde el ángulo de rotación  $\phi$  en la parte inferior de la proyección cónica circular, una posición  $S(\phi)$  sobre la circunferencia de la parte inferior de la proyección cónica circular, y una inclinación  $V(\phi)$  de la captura de imagen en esta posición se pueden obtener de forma única. Por lo tanto, una matriz de transformación afín  $(1/d)L(\phi)$  se puede obtener de forma única a partir de la posición  $S(\phi)$ , la inclinación  $V(\phi)$  y una longitud  $d$  del lado oblicuo de la proyección cónica circular.

A continuación, entre los ángulos de rotación  $\theta[0]$ ,  $\theta[1]$ ,  $\theta[2]$ , ... y  $\theta[N-1]$ , se selecciona uno más cercano al ángulo de rotación  $\phi$ . Se asume que el ángulo de rotación seleccionado es  $\theta[k]$ .

De acuerdo con esta técnica, una imagen interpolada  $B(\phi)$  con respecto al ángulo de rotación  $\phi$  se obtiene a partir de la imagen normalizada  $B[k]$ .

El esquema más sencillo es utilizar directamente la imagen normalizada  $B[k]$  como la imagen interpolada  $B(\phi)$ . Esto es eficaz para suprimir la cantidad de cálculo cuando  $N$  es suficientemente grande (por ejemplo, igual o superior a 20).

Además, como se ha explicado anteriormente, la imagen normalizada  $B[k]$  está asociada con la matriz de transformación afín  $(1/d)L[k]$  para proyectar el mundo real en la imagen y la matriz de transformación afín inversa  $L[k]^{-1}$ .

Por lo tanto, cuando una transformación afín

$$(1/d)L(\phi)d L[k]^{-1} = L(\phi)L[k]^{-1}$$

se realiza en la imagen normalizada  $B[k]$ , se obtiene la imagen interpolada  $B(\phi)$ .

De acuerdo con esta técnica, solamente se usa una imagen normalizada  $B[k]$  para obtener una imagen interpolada  $B(\phi)$ , pero es posible obtener una imagen interpolada  $B(\phi)$  de la pluralidad de imágenes normalizadas  $B[k]$ . Por ejemplo, para dos imágenes normalizadas  $B[k]$  y  $B[k+1]$  que tienen el ángulo de rotación que satisface

$$\theta[k] \leq \phi < \theta[k+1],$$

la transformación afín se realiza como se ha explicado anteriormente.

A continuación, esta técnica extrae el lado izquierdo  $(\phi - \theta[k]) / (\theta[k+1] - \theta[k])$  de la imagen obtenida al realizar la transformación afín en  $B[k]$ , y extrae el lado derecho  $(\theta[k+1] - \phi) / (\theta[k+1] - \theta[k])$  de la imagen obtenida al realizar la

transformación afín en  $B[k+1]$ . Estos se disponen lado a lado para obtener la imagen interpolada  $B(\varphi)$ . En este caso,  $\varphi$  y  $\theta$  tienen direcciones establecidas para aumentar cuando un lado cercano al usuario gira de izquierda a derecha, pero en el caso de la dirección inversa, si se cambian las direcciones, se puede aplicar una técnica similar.

5

Además, al disponer un mayor número de imágenes normalizadas sin disponer las dos imágenes normalizadas, como en el caso explicado anteriormente, la imagen normalizada se corta en tiras verticales con respecto a la dirección de rotación, y las tiras se extraen y se disponen para obtener la imagen interpolada  $B(\varphi)$ .

10

A continuación, la Figura 8 es un diagrama explicativo que muestra una relación entre cada unidad del dispositivo de procesamiento de imágenes 161 y cada unidad del dispositivo terminal 181. La Figura 9 es un diagrama de flujo que muestra un flujo de un control para un procesamiento de imágenes por cada unidad del dispositivo de procesamiento de imágenes 161 y cada unidad del dispositivo terminal 181. En lo sucesivo en el presente documento, la explicación se dará con referencia a estas figuras.

15

Como se ha explicado anteriormente, el dispositivo de procesamiento de imágenes 161 y el dispositivo terminal 181 pueden configurarse como un solo dispositivo, o pueden configurarse como dispositivos separados. También es opcional cambiar en qué dispositivo se dispone cada unidad. Por lo tanto, las figuras ilustran la configuración más típica.

20

Como se muestra en las figuras, el dispositivo de procesamiento de imágenes 161 incluye una unidad de recepción de ángulo 801, una unidad de aceptación de ángulo 802, una unidad de interpolación 803 y una unidad transmisora de imágenes 804.

25

Por el contrario, el dispositivo terminal 181 incluye una unidad de sensor de ángulo 811, una unidad transmisora de ángulo 812, una unidad de recepción de imágenes 813 y una unidad de visualización de monitor 814.

Cuando se inicia el procesamiento de imágenes, el dispositivo de procesamiento de imágenes 161 ejecuta un proceso de servidor, y el dispositivo terminal 181 ejecuta un proceso de terminal.

30

En primer lugar, en el dispositivo terminal 181, la unidad de sensor de ángulo 811 detecta un ángulo  $\varphi$  del dispositivo terminal 181 en torno a un eje predeterminado (etapa S901).

35

Cuando el dispositivo de imágenes de producto 121 captura una imagen del producto, es típico que se detecten inclinaciones a lo largo de tres ejes, pero en el dispositivo terminal 181, es suficiente si solo se usa un sensor de aceleración y se detecta una inclinación con respecto a la dirección gravitacional. No hace falta decir que está bien si se detecta una inclinación a lo largo de tres ejes.

40

A continuación, la unidad transmisora de ángulo 812 transmite el ángulo detectado  $\varphi$  al dispositivo de procesamiento de imagen 161 (etapa S902).

La unidad de recepción de ángulo 801 en el dispositivo de procesamiento de imágenes 161 recibe el ángulo  $\varphi$  transmitido desde el dispositivo terminal 181, y hace que la unidad de aceptación de ángulo 802 acepte el ángulo recibido (etapa S903).

45

A continuación, la unidad de interpolación 803 obtiene la imagen interpolada  $B(\varphi)$  basándose en el algoritmo explicado anteriormente desde el ángulo aceptado  $\varphi$  (etapa S904).

50

Además, la unidad transmisora de imágenes 804 transmite la imagen interpolada obtenida  $B(\varphi)$  al dispositivo terminal 161 (etapa S905) y el proceso regresa a la etapa S903.

En el dispositivo terminal 181, la unidad de recepción de imágenes 813 recibe la imagen interpolada  $B(\varphi)$  transmitida desde el dispositivo de procesamiento de imágenes 161 (etapa S906).

55

A continuación, la unidad de visualización de monitor 814 muestra la imagen interpolada recibida  $B(\varphi)$  en la pantalla del monitor (etapa S907), y el proceso regresa a la etapa S901.

60

Como se ha explicado anteriormente, simplemente inclinando el dispositivo terminal 181, el producto 301 que se muestra en la pantalla del monitor del dispositivo terminal 181 también gira de acuerdo con la inclinación, y el usuario puede ver fácilmente la apariencia del producto 301 desde todo el entorno.

65

Cuando el dispositivo de procesamiento de imágenes 161 y el dispositivo terminal 181 están configurados como un solo dispositivo, el ángulo  $\varphi$  detectado por la unidad de sensor de ángulo 811 se entrega a la unidad de interpolación 803, la imagen interpolada  $B(\varphi)$  obtenida por la unidad de interpolación 803 se entrega a la unidad de visualización de monitor 814, y la imagen interpolada  $B(\varphi)$  se muestra en la pantalla del monitor. Es decir, el

dispositivo terminal 181 incluye la unidad de sensor de ángulo 811, la unidad de interpolación 803, y la unidad de visualización de monitor 814 en este caso.

5 Además, cuando el dispositivo terminal 181 se configura utilizando un ordenador personal general que no tiene sensor de inclinación, se puede emplear una configuración en la que el usuario ingrese el ángulo  $\varphi$  a través de un dispositivo de entrada como un teclado o un ratón, sin detectar el ángulo  $\varphi$  utilizando la unidad de sensor de ángulo 811. Por ejemplo, puede emplearse un sistema operativo que aumenta  $\varphi$  cuando se presiona un primer botón del teclado y disminuye  $\varphi$  cuando se presiona un segundo botón, o un sistema operativo que vincula el desplazamiento del ratón con el valor de  $\varphi$ .

10 Además, de acuerdo con el ejemplo explicado anteriormente, el ángulo detectado  $\varphi$  se utiliza directamente para mostrar la imagen interpolada  $B(\varphi)$ . Sin embargo, al aplicar una constante  $k$  mayor de 1, se puede mostrar una imagen interpolada  $B(k\varphi)$  relativa al ángulo detectado  $\varphi$ . Es decir, la unidad de aceptación de ángulo 802 multiplica el ángulo aceptado por  $k$ .

15 De acuerdo con la técnica de visualización de la imagen interpolada  $B(\varphi)$ , para observar el producto 301 desde todo el entorno, es necesario que el usuario gire el propio dispositivo terminal 181 en un giro de 360 grados.

20 Por el contrario, de acuerdo con la técnica de visualización de la imagen interpolada  $B(k\varphi)$ , la cantidad rotacional del dispositivo terminal 181 necesaria para observar el producto 301 desde todo el entorno es solo de  $(360/k)$  grados.

25 Por ejemplo, cuando  $k = 4$ , si el usuario cambia la inclinación del dispositivo terminal 181 en 90 grados, es posible observar el producto 301 desde todo el entorno.

El proceso para multiplicar el ángulo por  $k$  puede ejecutarse, por ejemplo, mediante el dispositivo terminal 181. Por ejemplo, la unidad transmisora de ángulo 812 multiplica el ángulo detectado  $\varphi$  por  $k$ , y transmite un ángulo  $k\varphi$ .

30 A continuación, se proporcionará una explicación de otra técnica para obtener la imagen interpolada  $B(\varphi)$  de las imágenes normalizadas  $B[0]$ ,  $B[1]$ ,  $B[2]$ , ... y  $B[N-1]$ , los ángulos de rotación  $\theta[0]$ ,  $\theta[1]$ ,  $\theta[2]$ , ... y  $\theta[N-1]$ , y el ángulo de rotación dado.

35 Es decir, se supone que el producto 301 es una esfera, y las figuras desarrolladas se obtienen a partir de las imágenes normalizadas  $B[0]$ ,  $B[1]$ ,  $B[2]$ , ... y  $B[N-1]$  a través de la proyección policónica normal interrumpida. A continuación, se forma una esfera a partir de las figuras desarrolladas, y se crea la imagen interpolada  $B(\varphi)$  obtenida al observar la esfera en una posición  $S(\varphi)$  y una inclinación  $V(\varphi)$ .

40 De acuerdo con la proyección policónica normal interrumpida en el momento de crear un globo, se corta una superficie del globo por cada longitud de 30 grados para formar 12 formas de barco. El número de formas de barco es  $N$ . Además, como se ha explicado anteriormente, se supone que  $\theta[0] = 0$ . A continuación,

- 45 (0) las formas de barco desde la longitud de  $(\theta[N-1]+2\pi)/2$  hasta la longitud de  $2\pi$  y desde la longitud de 0 a la longitud de  $\theta[1]/2$  se forman a partir de la imagen normalizada  $B[0]$ ,
- (1) una forma de barco desde la longitud de  $(\theta[0]+\theta[1])/2$  a la longitud de  $(\theta[1]+\theta[2])/2$  se forma a partir de la imagen normalizada  $B[1]$ ,
- (2) una forma de barco desde la longitud de  $(\theta[1]+\theta[2])/2$  a la longitud de  $(\theta[2]+\theta[2])/3$  se forma a partir de la imagen normalizada  $B[2]$ ,
- 50 (k) una forma de barco desde la longitud de  $(\theta[k-1]+\theta[k])/2$  a la longitud de  $(\theta[k]+\theta[k+1])/2$  se forma a partir de la imagen normalizada  $B[k]$ , y
- (N-1) una forma de barco desde la longitud de  $(\theta[N-2]+\theta[N-1])/2$  a la longitud de  $(\theta[N-1]+2\pi)/2$  se forma a partir de la imagen normalizada  $B[N-1]$ .

55 La técnica más sencilla para formar las formas de barco es cortar directamente una posición correspondiente a la forma de barco de cada imagen normalizada. Esta técnica es adecuada cuando toda la forma del producto 301 es similar a una esfera.

60 Uniendo el número  $N$  de formas de barco formadas de esta manera, se puede formar virtualmente una esfera sustancial. Posteriormente, se crea la imagen interpolada  $B(\varphi)$  obtenida cuando se observa la esfera sustancial en la posición  $S(\varphi)$  y la inclinación  $V(\varphi)$ .

65 Además, al igual que la aproximación de poliedros que se analizará más adelante, también existe una técnica para ajustar la posición  $S[k]$  como una posición de punto visual, pegar formas de barco en la superficie de una esfera para formar un plano de proyección, y proyectar la imagen normalizada  $B[k]$  en la forma de barco. De acuerdo con esta técnica, la superficie de la esfera que cubre el producto 301 se divide en una pluralidad de

áreas de forma de barco, y la imagen normalizada se proyecta en cada área de forma de barco.

Como se ha explicado anteriormente, existe otra técnica para aproximar el producto 301 con un poliedro convexo virtual. El poliedro convexo similar a toda la forma del producto se selecciona en esta técnica.

5

A continuación, se forma un modelo de poliedro convexo del producto 301 realizando el siguiente proceso en cada polígono dispuesto sobre la superficie del poliedro convexo.

10

Primero, entre las imágenes normalizadas B[0], B[1], B[2], ... y B[N-1], se selecciona una imagen normalizada B[k] que tiene la inclinación V[0], V[1], V[2], ... o V[N-1] en el momento de la generación más cercana a la dirección de la línea normal del polígono.

15

La imagen normalizada B[k] está virtualmente dispuesta en la posición del producto 301 en el ángulo de rotación  $\theta[k]$ . A continuación, la posición S[k] se toma como la posición del punto visual, y el polígono se toma como el plano de proyección, proyectando de este modo la imagen normalizada B[k] en el polígono.

Una vez que se forman dicho modelo sustancialmente esférico y un modelo de poliedro convexo, es posible ver estos modelos desde diversas direcciones.

20

La presente solicitud reivindica prioridad de la Solicitud de Patente Japonesa N.º 2010-192927 presentada el 30 de agosto de 2010, cuya memoria descriptiva completa, reivindicaciones y dibujos se incorporan por referencia en la presente memoria descriptiva.

#### Aplicabilidad industrial

25

De acuerdo con la presente invención, es posible proporcionar un dispositivo de conversión de imágenes, un procedimiento, un medio de registro de información y un sistema de conversión de imágenes que permita apropiadamente a un usuario capturar fácilmente una secuencia de imágenes de toda la superficie exterior visible de un producto.

30

#### Descripción de los números de referencia

35

101 Sistema de procesamiento de imágenes  
121 Dispositivo de imagen del producto  
141 Dispositivo de conversión de imágenes  
161 Dispositivo de procesamiento de imágenes  
181 Dispositivo terminal

40

201 Unidad de sensor de imagen  
202 Unidad de recepción de instrucciones  
203 Unidad de memoria  
204 Unidad de visualización de buscador  
205 Unidad de sensor de inclinación  
301 Producto

45

311 Imagen ya capturada  
312 Gráfico de barras que indica una inclinación cuando se capturó una imagen ya capturada  
321 Imagen de vista previa en vivo  
322 Gráfico de barras que indica la inclinación actual  
331 Imagen de buscador

50

601 Unidad de recepción de imágenes  
602 Unidad convertidora  
611 Unidad de extracción  
612 Unidad de estimación  
613 Unidad de aproximación  
614 Unidad de salida

55

801 Unidad de recepción de ángulo  
802 Unidad de aceptación de ángulo  
803 Unidad de interpolación  
804 Unidad transmisora de imágenes

60

811 Unidad de sensor de ángulo  
812 Unidad transmisora de ángulo  
813 Unidad de recepción de imágenes  
814 Unidad de visualización de monitor

65

**REIVINDICACIONES**

1. Un dispositivo de conversión de imágenes (141) que comprende:

5 una unidad de recepción (601) que recibe una imagen G[0] de un producto (301) capturado por un dispositivo de captura de imágenes (121), una inclinación R[0] del dispositivo de captura de imágenes (121) en un sistema de coordenadas mundial cuando la imagen G[0] fue capturada, una imagen G[1] del producto (301) capturada por el dispositivo de captura de imágenes (121) y una inclinación R[1] del dispositivo de captura de imágenes (121) en el sistema de coordenadas mundial cuando la imagen G[1] fue capturada, habiendo sido capturadas la imagen G[0] y la imagen G[1] de tal manera que el tamaño del producto (301) es sustancialmente constante en la imagen G[0] y en la imagen G[1];  
 10 una unidad de extracción (611) que extrae dos puntos característicos correspondientes de cada una de la imagen G[0] y la imagen G[1];  
 15 una unidad de estimación (612) que estima una posición de captura de imagen T[0] relativa a un origen del sistema de coordenadas mundial donde la imagen G[0] fue capturada y una posición de captura de imagen T[1] relativa al origen del sistema de coordenadas mundial donde la imagen G[1] fue capturada en función de las posiciones p[0] y q[0] de los dos puntos característicos extraídos en la imagen G[0], las posiciones p[1] y q[1] de los dos puntos característicos extraídos en la imagen G[1], la inclinación R[0] y la inclinación R[1], y una posición de producto del producto (301) relativa al origen del sistema de coordenadas mundial;  
 20 una unidad de aproximación (613) que asocia las respectivas posiciones de captura de imagen estimadas para la imagen G[0] y la imagen G[1] con las respectivas posiciones normalizadas a una distancia igual de un eje central que pasa por la posición del producto y a una distancia igual desde la posición del producto estimada, y que transforma cada una de las imágenes G[0] y la imagen G[1] en una imagen normalizada que es una imagen que se aproxima a una imagen del producto capturada desde la posición normalizada asociada y con una inclinación dirigida hacia la posición del producto; y  
 25 una unidad de salida (614) que emite la pluralidad de imágenes normalizadas, ángulos de rotación respectivos de la pluralidad de posiciones normalizadas alrededor del eje central, y una inclinación respectiva desde la pluralidad de posiciones normalizadas hacia el producto, estando las imágenes normalizadas, los ángulos de rotación y la inclinación asociados entre sí,  
 30 la unidad de estimación (612) resuelve, para una matriz de transformación M[0] de una coordenada homogénea que transforma una posición de un objetivo de captura de imagen cuando el objetivo de captura de imagen fue capturado en la posición de captura de imagen T[0] y en la inclinación R[0] a una posición en la imagen G[0], una matriz de transformación M[1] de una coordenada homogénea que transforma una posición del objetivo de captura de imagen cuando el objetivo de captura de imagen fue capturado en la posición de captura de imagen T[1] y en la inclinación R[1] en una posición en la imagen G[1], un vector vertical (p[0], 1)<sup>t</sup> que organiza los elementos respectivos de la posición p[0] y 1 en una dirección vertical, un vector vertical (p[1], 1)<sup>t</sup> que organiza los elementos respectivos de la posición p[1] y 1 en la dirección vertical, un vector vertical (q[0], 1)<sup>t</sup> que organiza los elementos respectivos de la posición q[0] y 1 en la dirección vertical, y un vector vertical (q[1], 1)<sup>t</sup> que organiza los elementos respectivos de la posición q[1] y 1 en la dirección vertical, las siguientes ecuaciones simultáneas:

$$M[0]^{-1}(p[0], 1)^t = M[1]^{-1}(p[1], 1)^t$$

y

$$M[0]^{-1}(q[0], 1)^t = M[1]^{-1}(q[1], 1)^t,$$

50 estimando de este modo la posición de captura de imagen T[0] y la posición de captura de imagen T[1].

2. El dispositivo de conversión de imágenes (141) de acuerdo la reivindicación 1, en el que:

55 la unidad de recepción (601) recibe una entrada de N número de imágenes ya capturadas, incluyendo la imagen G[0], la imagen G[1], ... una imagen G[N-1], del producto (301) capturado desde diferentes posiciones y una inclinación, incluyendo la inclinación R[0], la inclinación R[1], ... R[N-1] en el sistema de coordenadas mundial cuando se capturó cada una del N número de imágenes ya capturadas habiendo sido capturada cada una de las imágenes G[0], G[1] ... G[N-1] de tal manera que el tamaño del producto (301) es sustancialmente constante a través de las imágenes;  
 60 la unidad de extracción (611) extrae, para cada uno de los números enteros k = 0, 1, 2, ... y N-2, dos puntos característicos correspondientes de una k-ésima imagen ya capturada entre el número N recibido de imágenes ya capturadas y una k+1-ésima imagen ya capturada;  
 65 la unidad de estimación (612) estima, desde las posiciones p[k] y q[k] de los dos puntos característicos

extraídos en la k-ésima imagen ya capturada, las posiciones  $p[k+1]$  y  $q[k+1]$  de los dos puntos característicos extraídos en la k+1-ésima imagen ya capturada, una inclinación  $R[k]$  asociada con la k-ésima imagen ya capturada, y una inclinación  $R[k+1]$  asociada con la k+1-ésima imagen ya capturada, una posición de captura de imagen  $T[k]$  relativa al sistema de coordenadas mundial donde se capturó la k-ésima imagen ya capturada, una posición de captura de imagen  $T[k+1]$  donde se capturó la k+1-ésima imagen ya capturada, y una posición del producto del producto (301) relativa al origen del sistema de coordenadas mundial,

en el que, en el dispositivo de conversión de imágenes (141):

la unidad de aproximación (613) asocia las respectivas posiciones de captura de imagen estimadas para el número N de imágenes ya capturadas con respectivas posiciones normalizadas a una distancia igual de un eje central que pasa por la posición del producto y a una distancia igual de la posición del producto estimada, y transforma cada una del N número de imágenes ya capturadas en una imagen normalizada que es una imagen que se aproxima a una imagen del producto capturada desde la posición normalizada asociada y en una inclinación dirigida hacia la posición del producto, y en el que la unidad de estimación (612) resuelve, para una matriz de transformación  $M[k]$  de una coordenada homogénea que transforma una posición de un objetivo de captura de imagen en una posición en la k-ésima imagen, una matriz de transformación  $M[k+1]$  de una coordenada homogénea que transforma la posición del objetivo de captura de imagen en una posición en la k+1-ésima imagen, un vector vertical  $(p[k], 1)^t$  que organiza los elementos respectivos de la posición  $p[k]$  y 1 en una dirección vertical, un vector vertical  $(q[k], 1)^t$  que organiza los elementos respectivos de la posición  $q[k]$  y 1 en la dirección vertical, un vector vertical  $(p[k+1], 1)^t$  que organiza los elementos respectivos de la posición  $p[k+1]$  y 1 en la dirección vertical, y un vector vertical  $(q[k+1], 1)^t$  que organiza los elementos respectivos de la posición  $q[k+1]$  y 1 en la dirección vertical, las siguientes ecuaciones simultáneas:

$$M[k]^{-1}(p[k], 1)^t = M[k+1]^{-1}(p[k+1], 1)^t$$

y

$$M[k]^{-1}(q[k], 1)^t = M[k+1]^{-1}(q[k+1], 1)^t,$$

estimando de este modo la posición de captura de imagen  $T[k]$  y la posición de captura de imagen  $T[k+1]$ .

3. El dispositivo de conversión de imágenes (141) de acuerdo con la reivindicación 2, en el que la unidad de aproximación escala cada una del N número de imágenes ya capturadas en función de una relación de una distancia entre la posición del producto estimada y la posición normalizada y una distancia entre la posición del producto estimada y la posición de captura de imagen para obtener la imagen normalizada transformada.
4. El dispositivo de conversión de imágenes (141) de acuerdo la reivindicación 3, en el que la unidad de aproximación (613) utiliza una transformación afín de una coordenada homogénea definida en función de la posición de captura de imagen y una inclinación en la posición de captura de imagen y una transformación afín de una coordenada homogénea definida en función de la posición normalizada y una inclinación en la posición normalizada para transformar cada una del N número de imágenes ya capturadas en la imagen normalizada.
5. Un procedimiento que comprende:

una etapa de recepción (S701) de recibir una imagen  $G[0]$  de un producto (301) capturado por un dispositivo de captura de imágenes, una inclinación  $R[0]$  del dispositivo de captura de imágenes (121) en un sistema de coordenadas mundial cuando una imagen  $G[0]$  fue capturada, una imagen  $G[1]$  del producto (301) capturada por el dispositivo de captura de imágenes (121) y una inclinación  $R[1]$  del dispositivo de captura de imágenes (121) en el sistema de coordenadas mundial cuando la imagen  $G[1]$  fue capturada, habiendo sido capturadas la imagen  $G[0]$  y la imagen  $G[1]$  de tal manera que el tamaño del producto (301) es sustancialmente constante en la imagen  $G[0]$  y la imagen  $G[1]$ ;

una etapa de extracción (S703) de extraer dos puntos característicos correspondientes de cada una de la imagen  $G[0]$  y la imagen  $G[1]$ ;

una etapa de estimación (S705) de estimar una posición de captura de imagen  $T[0]$  relativa a un origen del sistema de coordenadas mundial donde la imagen  $G[0]$  fue capturada y una posición de captura de imagen  $T[1]$  relativa al origen del sistema de coordenadas mundial donde la imagen  $G[1]$  fue capturada en función de las posiciones  $p[0]$  y  $q[0]$  de los dos puntos característicos extraídos en la imagen  $G[0]$ , las posiciones  $p[1]$  y  $q[1]$  de los dos puntos característicos extraídos en la imagen  $G[1]$ , la inclinación  $R[0]$  y la inclinación  $R[1]$ , y una posición del producto del producto (301) relativa al origen del sistema de coordenadas mundial;

una etapa de aproximación de asociar (S709) respectivas posiciones de captura de imagen estimadas para la imagen G[0] y la imagen G[1] con las respectivas posiciones normalizadas a una distancia igual de un eje central que pasa por la posición del producto y a una distancia igual de la posición del producto estimada, y de transformar (S712) cada una de la imagen G[0] y la imagen G[1] en una imagen normalizada que es una imagen que se aproxima a una imagen del producto capturada desde la posición normalizada asociada y a un inclinación dirigida hacia la posición del producto; y una etapa de salida (S713) de emisión de la pluralidad de imágenes normalizadas, de ángulos de rotación respectivos de la pluralidad de posiciones normalizadas alrededor del eje central, y de una inclinación respectiva desde la pluralidad de posiciones normalizadas hacia el producto, estando las imágenes normalizadas, los ángulos de rotación y la inclinación asociados entre sí, en la etapa de estimación (S705), para una matriz de transformación M[0] de una coordenada homogénea que transforma una posición de un objetivo de captura de imagen cuando el objetivo de captura de imagen se capturó en la posición de captura de imagen T[0] y en la inclinación R[0] a una posición en la imagen G[0], una matriz de transformación M[1] de una coordenada homogénea que transforma una posición del objetivo de captura de imagen cuando el objetivo de captura de imagen se capturó en la posición de captura de imagen T[1] y en la inclinación R[1] en una posición en la imagen G[1], un vector vertical  $(p[0], 1)^t$  que organiza los elementos respectivos de la posición p[0] y 1 en dirección vertical, un vector vertical  $(p[1], 1)^t$  que organiza los elementos respectivos de la posición p[1] y 1 en la dirección vertical, un vector vertical  $(q[0], 1)^t$  que organiza los elementos respectivos de la posición q[0] y 1 en la dirección vertical, y un vector vertical  $(q[1], 1)^t$  que organiza los elementos respectivos de la posición q[1] y 1 en la dirección vertical, las siguientes ecuaciones simultáneas:

$$M[0]^{-1}(p[0], 1)^t = M[1]^{-1}(p[1], 1)^t$$

y

$$M[0]^{-1}(q[0], 1)^t = M[1]^{-1}(q[1], 1)^t,$$

estimando de este modo la posición de captura de imagen T[0] y la posición de captura de imagen T[1].

6. Un medio de registro de información no transitorio legible por computadora que registra en el mismo un programa que permite que un ordenador funcione como:

una unidad de recepción (601) que recibe una imagen G[0] de un producto (301) capturado por un dispositivo de captura de imágenes (121) en un sistema de coordenadas mundial, una inclinación R[0] del dispositivo de captura de imágenes (121) cuando la imagen G[0] fue capturada, una imagen G[1] del producto (301) capturada por el dispositivo de captura de imágenes (121) y una inclinación R[1] del dispositivo de captura de imágenes en el sistema de coordenadas mundial cuando la imagen G[1] fue capturada, habiendo sido capturadas la imagen G[0] y la imagen G[1] de tal manera que el tamaño del producto (301) es sustancialmente constante en la imagen G[0] y en la imagen G[1];

una unidad de extracción (611) que extrae dos puntos característicos correspondientes de cada una de la imagen G[0] y la imagen G[1];

una unidad de estimación (612) que estima una posición de captura de imagen T[0] relativa a un origen del sistema de coordenadas mundial donde la imagen G[0] fue capturada y una posición de captura de imagen T[1] relativa al origen del sistema de coordenadas mundial donde la imagen G[1] fue capturada en función de las posiciones p[0] y q[0] de los dos puntos característicos extraídos en la imagen G[0], las posiciones p[1] y q[1] de los dos puntos característicos extraídos en la imagen G[1], la inclinación R[0] y la inclinación R[1], y una posición de producto del producto (301) relativa al origen del sistema de coordenadas mundial;

una unidad de aproximación (613) que asocia las respectivas posiciones de captura de imagen estimadas para la imagen G[0] y la imagen G[1] con las respectivas posiciones normalizadas a una distancia igual de un eje central que pasa por la posición del producto y a una distancia igual desde la posición del producto estimada, y que transforma cada una de las imágenes G[0] y la imagen G[1] en una imagen normalizada que es una imagen que se aproxima a una imagen del producto capturada desde la posición normalizada asociada y con una inclinación dirigida hacia la posición del producto; y una unidad de salida (614) que emite la pluralidad de imágenes normalizadas, ángulos de rotación respectivos de la pluralidad de posiciones normalizadas alrededor del eje central, y una inclinación respectiva desde la pluralidad de posiciones normalizadas hacia el producto, estando las imágenes normalizadas, los ángulos de rotación y la inclinación asociados entre sí,

la unidad de estimación (612) resuelve, para una matriz de transformación M[0] de una coordenada homogénea que transforma una posición de un objeto de captura de imagen cuando el objeto de captura de imagen fue capturado en la posición de captura de imagen T[0] y en la inclinación R[0] a

una posición en la imagen G[0], una matriz de transformación M[1] de una coordenada homogénea que transforma una posición del objeto de captura de imagen cuando el objeto de captura de imagen fue capturado en la posición de captura de imagen T[1] y en la inclinación R[1] en una posición en la imagen G[1], un vector vertical  $(p[0], 1)^t$  que organiza los elementos respectivos de la posición p[0] y 1 en una dirección vertical, un vector vertical  $(p[1], 1)^t$  que organiza los elementos respectivos de la posición p[1] y 1 en la dirección vertical, un vector vertical  $(q[0], 1)^t$  que organiza los elementos respectivos de la posición q[0] y 1 en la dirección vertical, y un vector vertical  $(q[1], 1)^t$  que organiza los elementos respectivos de la posición q[1] y 1 en la dirección vertical, las siguientes ecuaciones simultáneas:

$$M[0]^{-1}(p[0], 1)^t = M[1]^{-1}(p[1], 1)^t$$

y

$$M[0]^{-1}(q[0], 1)^t = M[1]^{-1}(q[1], 1)^t,$$

estimando de este modo la posición de captura de imagen T[0] y la posición de captura de imagen T[1].

7. Un sistema de conversión de imágenes que comprende:

el dispositivo de conversión de imágenes (141) de acuerdo con una cualquiera de las reivindicaciones 1 a 4; y

un dispositivo de imágenes de producto que comprende:

una unidad de sensor de imagen que detecta la luz incidente de un mundo externo donde se dispone un producto y que emite una imagen que representa un resultado de la detección;

una unidad de recepción de instrucciones que recibe una instrucción de captura de imagen de un usuario;

una unidad de memoria que almacena, tras la recepción de la instrucción de captura de imagen, la imagen detectada por la unidad de sensor de imagen; y

una unidad de visualización de buscador que sintetiza una primera imagen almacenada en la unidad de memoria con una segunda imagen detectada actualmente por la unidad de sensor de imagen y que muestra una imagen sintetizada en una pantalla de buscador, en el que la unidad de visualización de buscador aplica un filtro de imagen predeterminado en la primera imagen almacenada en la unidad de memoria para hacer una imagen semitransparente, y superpone la imagen semitransparente en la segunda imagen detectada actualmente por la unidad de sensor de imagen para obtener la imagen sintetizada.

FIG. 1

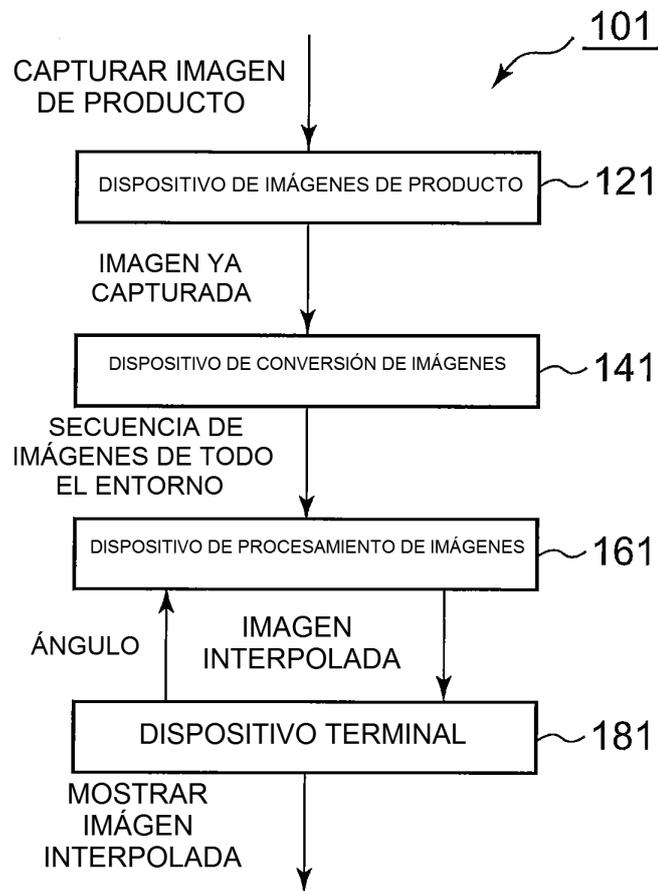


FIG. 2

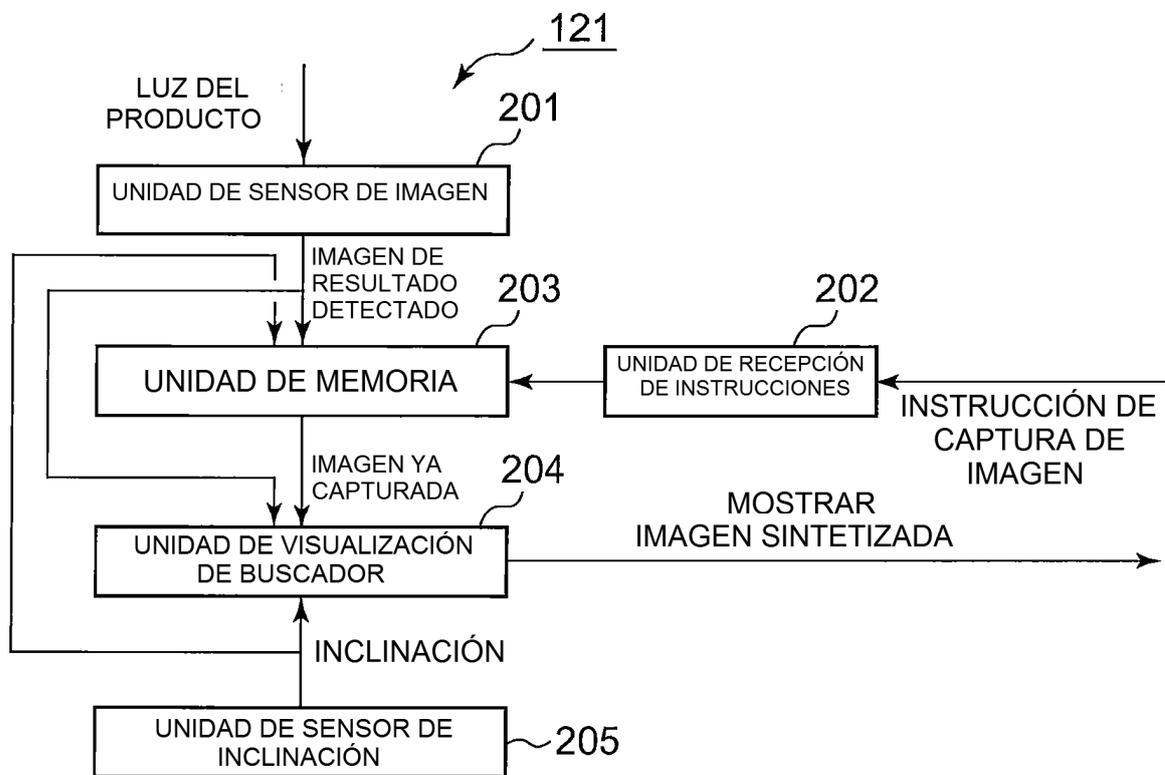


FIG. 3

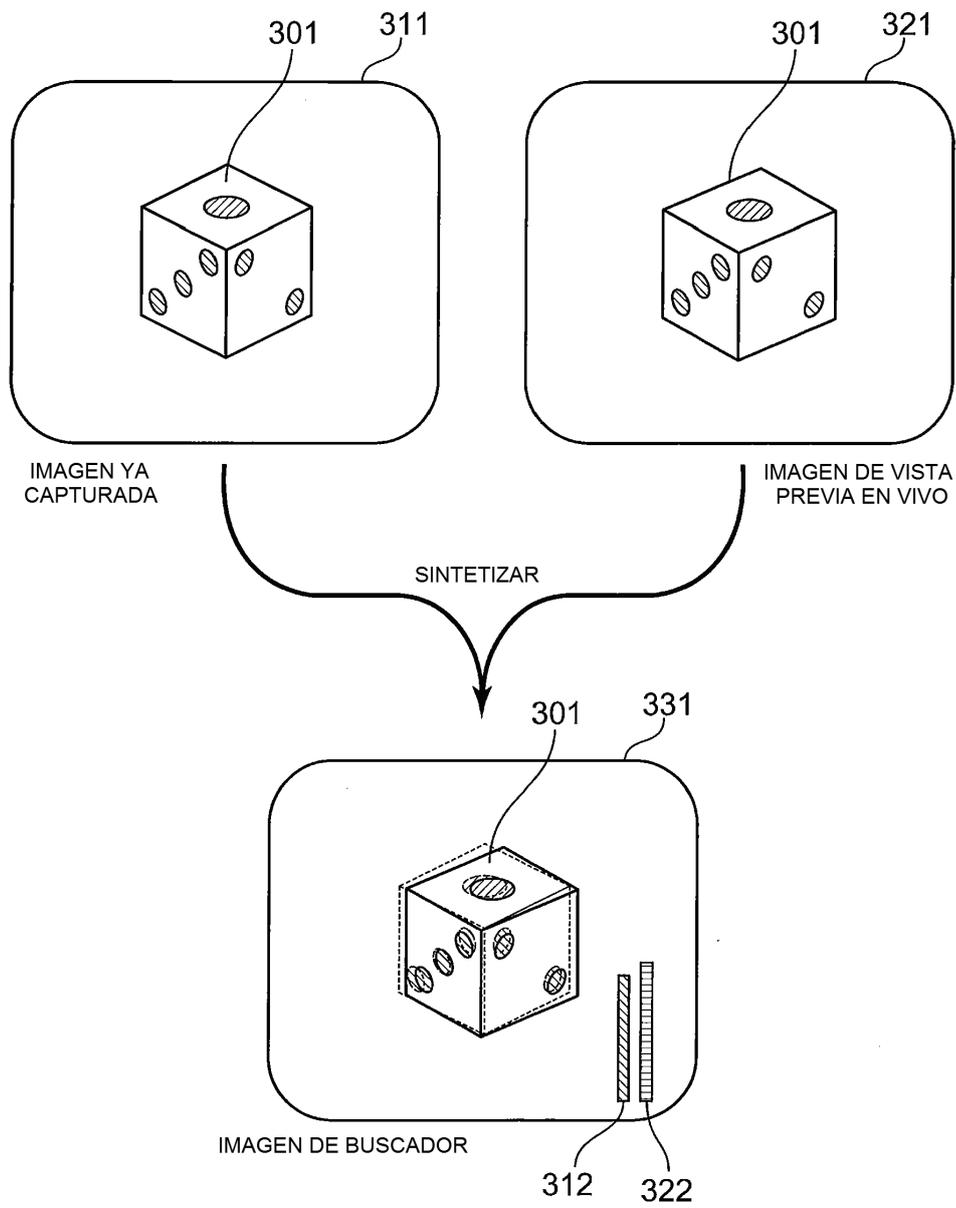


FIG. 4

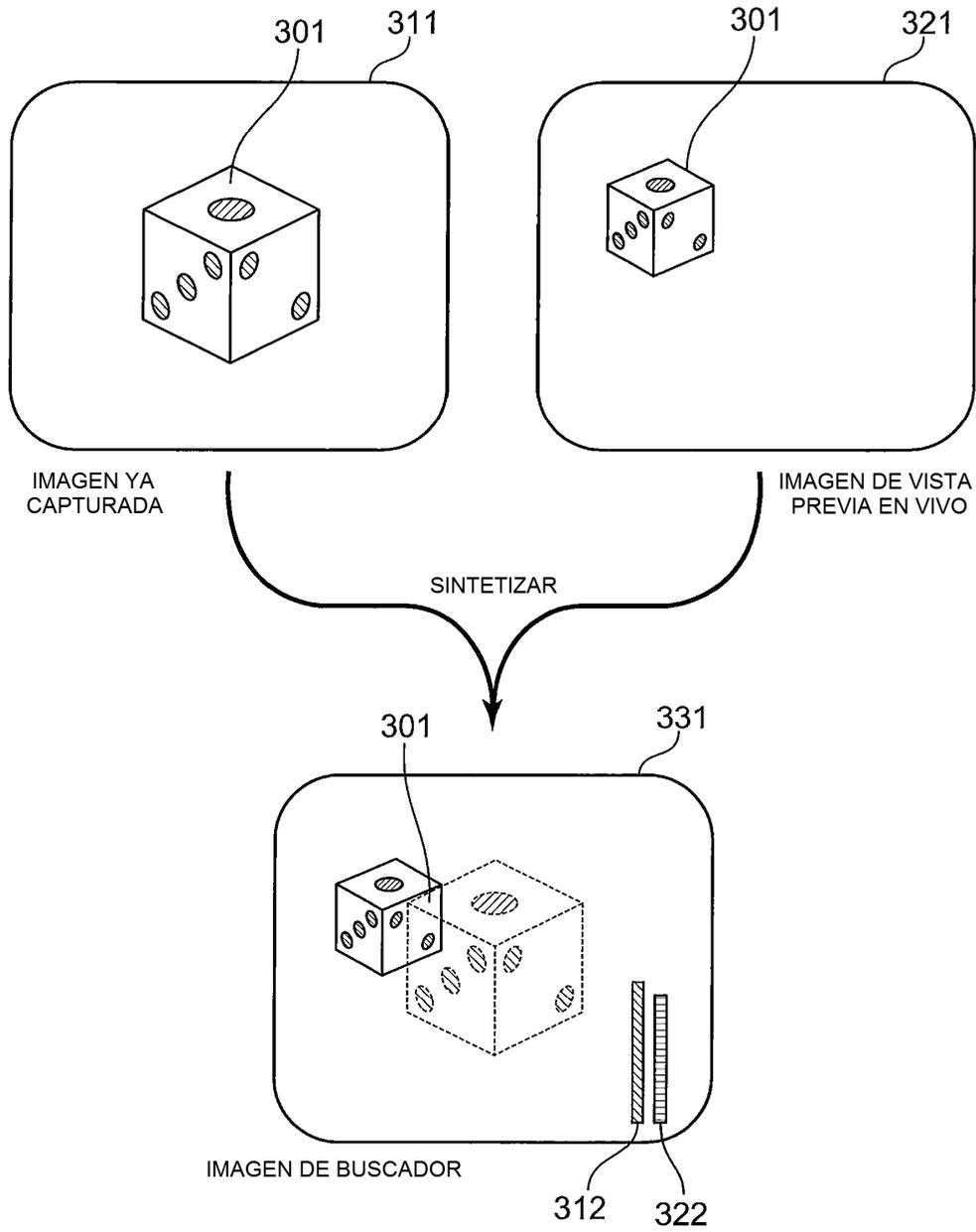


FIG. 5

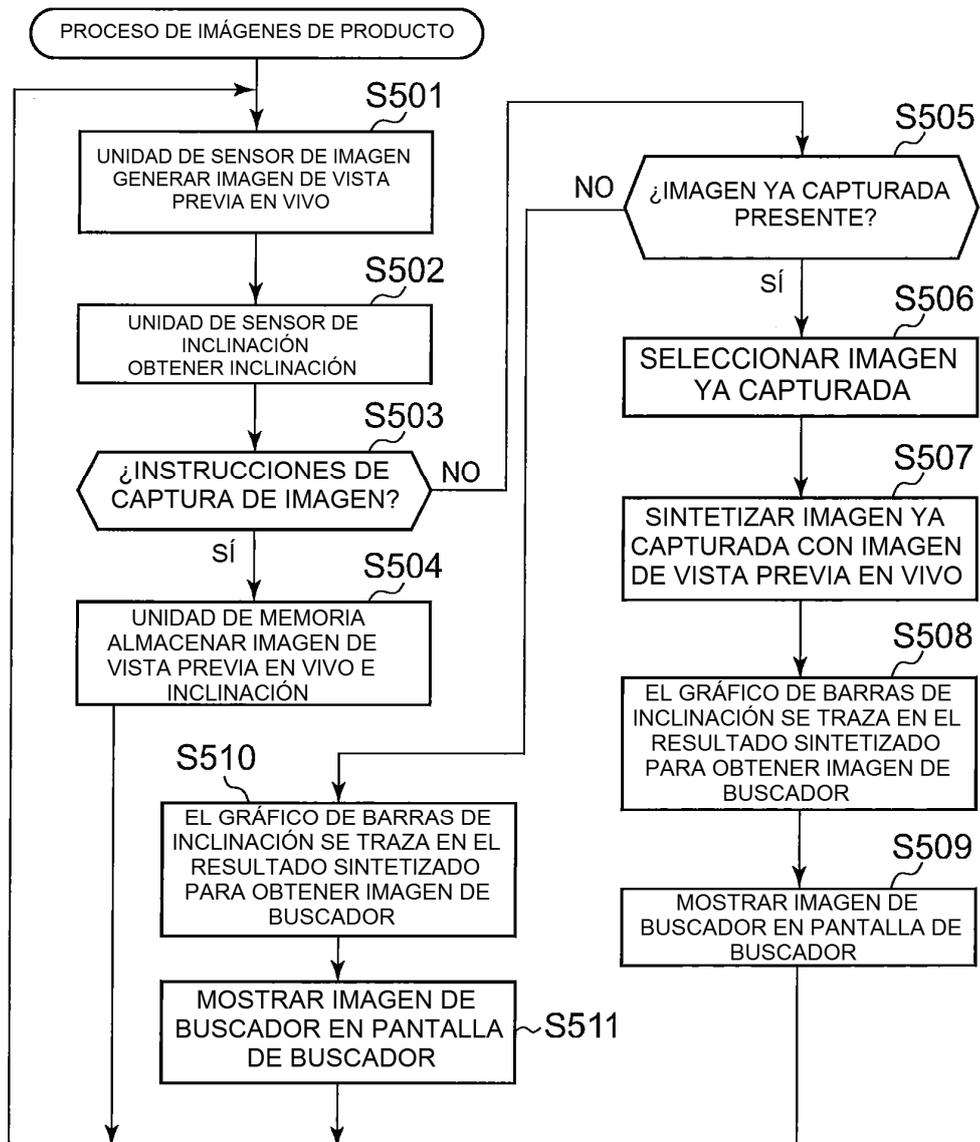


FIG. 6

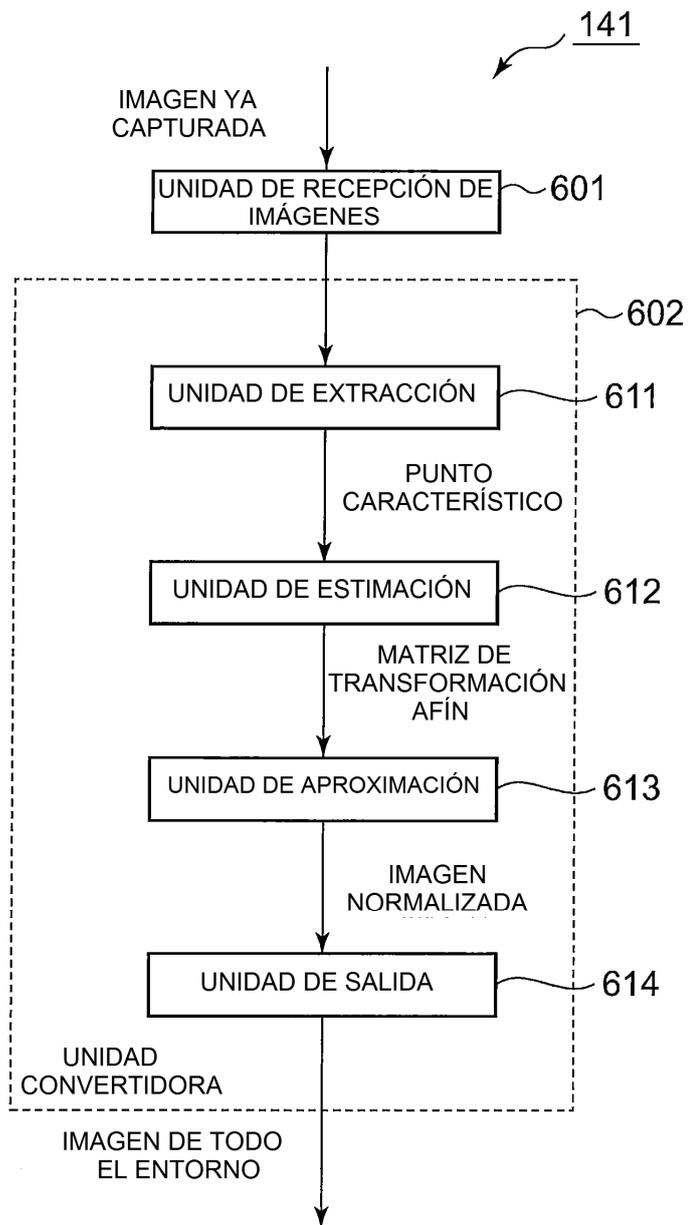


FIG. 7

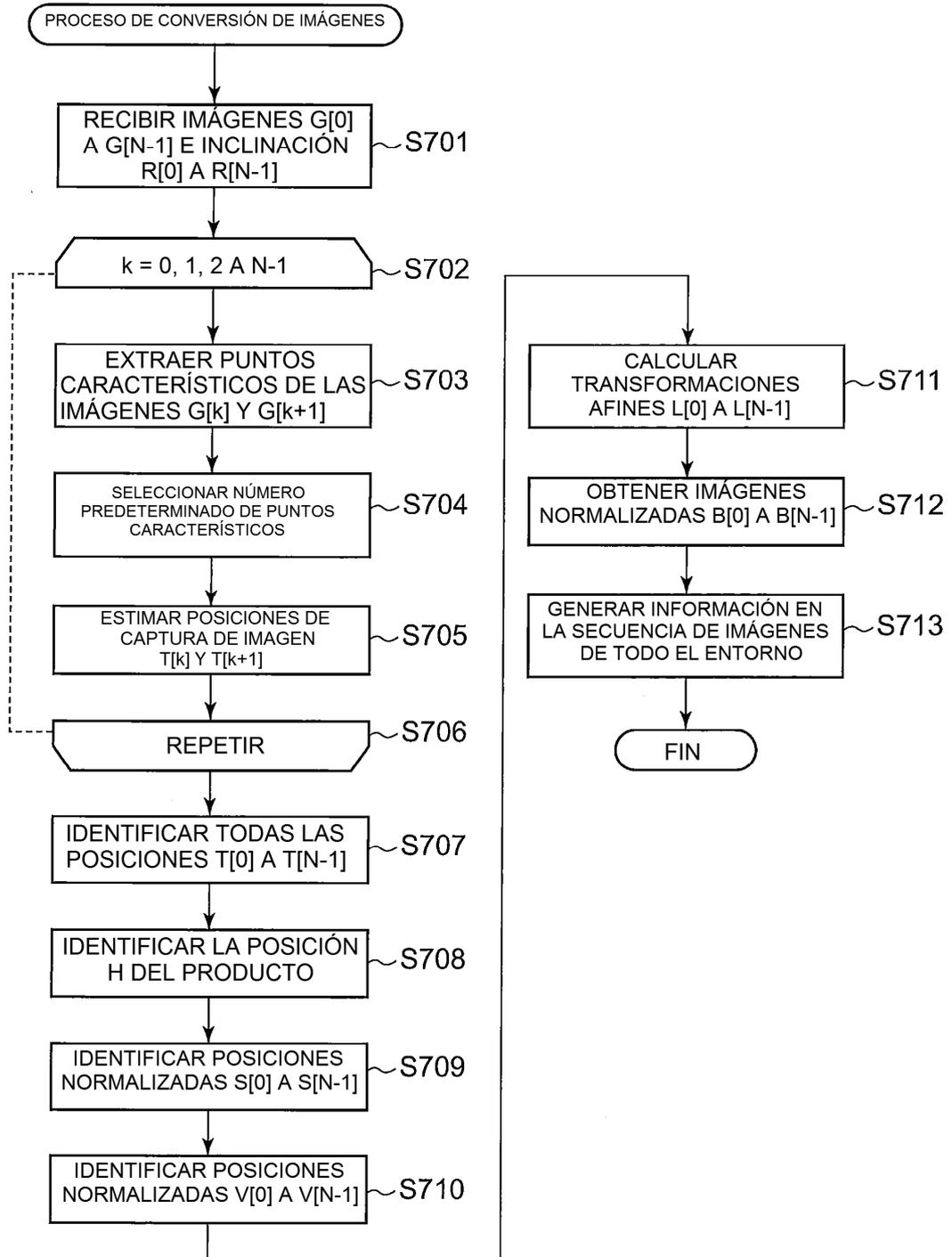


FIG. 8

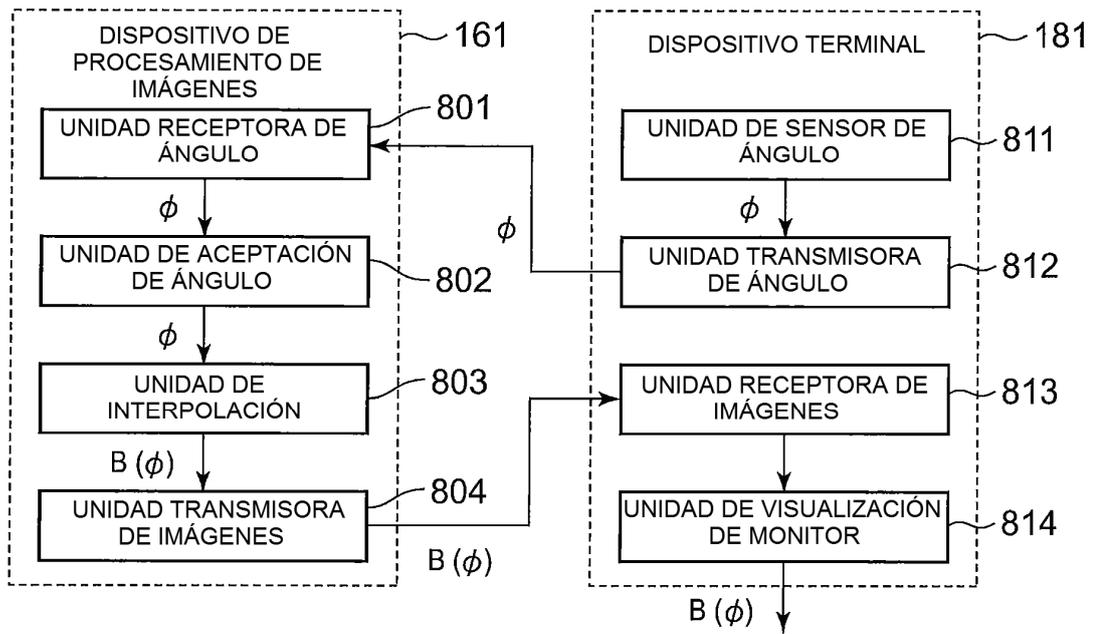


FIG. 9

