

19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 730 952**

51 Int. Cl.:

B25J 9/16 (2006.01)

G06T 7/11 (2007.01)

G06T 7/194 (2007.01)

G06T 7/174 (2007.01)

G06K 9/78 (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

86 Fecha de presentación y número de la solicitud internacional: **12.10.2011 PCT/FI2011/050884**

87 Fecha y número de publicación internacional: **26.04.2012 WO12052614**

96 Fecha de presentación y número de la solicitud europea: **12.10.2011 E 11833918 (3)**

97 Fecha y número de publicación de la concesión europea: **13.03.2019 EP 2629939**

54 Título: **Procedimiento para el filtrado de imágenes de objetos objetivo en un sistema robótico**

30 Prioridad:

21.10.2010 FI 20106090

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

13.11.2019

73 Titular/es:

ZENROBOTICS OY (100.0%)

Vilhonkatu 5 A

00100 Helsinki, FI

72 Inventor/es:

VALPOLA, HARRI

74 Agente/Representante:

DURAN-CORRETJER, S.L.P

ES 2 730 952 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

Procedimiento para el filtrado de imágenes de objetos objetivo en un sistema robótico

5 ANTECEDENTES DE LA INVENCION

Sector de la invención:

10 La presente invención se refiere a sistemas y procedimientos utilizados para manipular objetos físicos con un brazo robótico y un elemento de agarre. En particular, la presente invención se refiere a un procedimiento para el filtrado de imágenes de objetos objetivo en un sistema robótico.

Descripción de la técnica relacionada:

15 Se puede utilizar un sistema robótico en la ordenación y clasificación de una variedad de objetos físicos, tales como componentes de fabricación, partes de máquinas y material para reciclar. La ordenación y la clasificación requieren que los objetos físicos puedan reconocerse con suficiente probabilidad. En aplicaciones tales como el reciclaje y la gestión de residuos, es importante que la pureza de un grupo de objetos ordenados sea alta, concretamente que el menor número posible de objetos de un tipo incorrecto terminen en los grupos de objetos ordenados.

20 En los sistemas robóticos, el reconocimiento de objetos físicos que van a moverse o manipularse puede comprender dos fases. En la primera fase, se reconoce un objeto objetivo que va a agarrarse utilizando un elemento de agarre, garra o abrazadera u otro dispositivo similar conectado a un brazo robótico entre una serie de objetos. En la segunda fase, el objeto objetivo se ha agarrado satisfactoriamente y se puede inspeccionar más de cerca. La inspección se realiza habitualmente utilizando una serie de sensores, que normalmente comprenden, una cámara o un sensor de infrarrojos. Se puede conectar una cámara al brazo robótico o al elemento de agarre. La inspección se puede realizar contra un fondo en blanco o por lo demás limpio que no contenga ningún objeto que pueda interferir en el proceso de reconocimiento. En el entorno, desde donde se agarra el objeto objetivo, habitualmente hay otros objetos que incluso pueden cubrir el objeto parcial o totalmente, haciendo de ese modo que sea difícil reconocer y clasificar el objeto objetivo en su entorno de fondo. Un entorno de este tipo puede denominarse un escenario no estructurado.

35 En general, el problema de reconocer objetos se aborda habitualmente utilizando algoritmos de reconocimiento de patrones, que buscan objetos en datos sensoriales, como imágenes de cámaras digitales. Tales algoritmos constituyen un sector activamente estudiado. Aunque hay muchos algoritmos que incluso pueden reconocer objetos contra un fondo desigual, los algoritmos de reconocimiento de patrones en general funcionan mejor cuando el fondo es uniforme y predeterminado. Anteriormente, se han buscado objetos de tipo predeterminado de un área de funcionamiento despejada y se han seleccionado del área de funcionamiento como objetos reconocidos. Se pueden realizar conjuntos de acciones en un objeto seleccionado de un tipo conocido. El conjunto de acciones se puede elegir en base al tipo de objeto, por ejemplo, colocando diferentes tipos de objetos en diferentes compartimentos.

45 Es posible mover un objeto objetivo utilizando el brazo robótico a un fondo uniforme de modo que pueda inspeccionarse el objeto objetivo con el sensor sin la presencia de objetos que interfieran. Sin embargo, el problema con este enfoque es que cada objeto potencial de interés para una operación de ordenación debe moverse por separado a una ubicación donde pueda verse contra el fondo de inspección. Esto requiere mucho tiempo y consume mucha energía debido al hecho de que el brazo robótico tiene que moverse con frecuencia. Por tanto, sería beneficioso poder evitar la etapa de mover un objeto objetivo a una ubicación de inspección específica.

50 La publicación 2010 IEEE ICRA, 3 de mayo de 2010, páginas 2012 - 2019 de Kai Welke y otros da a conocer una disposición para la adquisición autónoma de representaciones de objetos que permite el reconocimiento, la localización y el agarre de objetos en el entorno es una tarea exigente, que ha demostrado ser difícil. La publicación presenta un sistema para la adquisición autónoma de representaciones de objetos visuales, que dota a un robot humanoide de la capacidad de enriquecer su representación de objetos interna y permite la realización de tareas visuales complejas.

55 El documento US2008/310677 da a conocer un procedimiento y un sistema para detectar ópticamente un objeto dentro de un campo de visión donde la detección es difícil debido al desorden de fondo dentro del campo de visión que oculta el objeto. Una cámara hace una toma panorámica con movimiento del objeto para estabilizar el movimiento del objeto respecto al desorden de fondo mientras toma una serie de fotogramas de imagen del objeto. Se realiza un análisis fotograma a fotograma para determinar varianzas en la intensidad de cada píxel, a lo largo del tiempo, con respecto a los fotogramas recopilados. A partir de este análisis, se construye una imagen de varianza que incluye un valor de varianza de intensidad para cada píxel. Los píxeles que representan el desorden de fondo normalmente variarán de manera considerable en intensidad de fotograma a fotograma, mientras que los píxeles que constituyen el objeto variarán poco o nada. Entonces se aplica una prueba de umbral binario a cada valor de varianza y los resultados se utilizan para construir una imagen final. La imagen final puede ser una imagen en blanco y negro que muestra claramente el objeto como una silueta.

CARACTERÍSTICAS DE LA INVENCION:

5 La invención se refiere a un procedimiento que comprende: agarrar un objeto con un elemento de agarre unido a un brazo robótico; capturar, como mínimo, dos imágenes de origen de un área que comprende el objeto; calcular una imagen promedio de las, como mínimo, dos imágenes de origen; calcular una imagen de varianza de las, como mínimo, dos imágenes de origen; formar una imagen de filtrado a partir de la imagen de varianza; y obtener una imagen de resultado enmascarando la imagen promedio utilizando la imagen de filtrado como una máscara de bits.

10 La invención también se refiere a un aparato, que comprende: medios para controlar un elemento de agarre y un brazo robótico para agarrar un objeto; medios para obtener, como mínimo, dos imágenes de origen de un área que comprende el objeto; medios para calcular una imagen promedio de las, como mínimo, dos imágenes de origen; medios para calcular una imagen de varianza de las, como mínimo, dos imágenes de origen; medios para formar una imagen de filtrado a partir de la imagen de varianza; y medios para obtener una imagen de resultado enmascarando la imagen promedio utilizando la imagen de filtrado como una máscara de bits.

15 La invención también se refiere a un programa informático implementado en un medio legible por ordenador, comprendiendo el programa informático código para controlar un procesador para ejecutar un procedimiento que comprende: controlar un elemento de agarre y un brazo robótico para agarrar un objeto; obtener como mínimo dos imágenes de origen de un área que comprende el objeto con un sensor de imagen; calcular una imagen promedio de las, como mínimo, dos imágenes de origen; calcular una imagen de varianza de las, como mínimo, dos imágenes de origen; formar una imagen de filtrado a partir de la imagen de varianza; y obtener una imagen de resultado enmascarando la imagen promedio utilizando la imagen de filtrado como una máscara de bits.

20 La invención también se refiere a un producto de programa informático que comprende: controlar un elemento de agarre y un brazo robótico para agarrar un objeto; obtener como mínimo dos imágenes de origen de un área que comprende el objeto con un sensor de imagen; calcular una imagen promedio de las, como mínimo, dos imágenes de origen; calcular una imagen de varianza de las, como mínimo, dos imágenes de origen; formar una imagen de filtrado a partir de la imagen de varianza; y obtener una imagen de resultado enmascarando la imagen promedio utilizando la imagen de filtrado como una máscara de bits.

25 La invención también se refiere a un aparato que comprende una memoria y, como mínimo, un procesador configurado para controlar un elemento de agarre y un brazo robótico para agarrar un objeto, para obtener como mínimo dos imágenes de origen de un área que comprende el objeto con un sensor de imagen, para calcular una imagen promedio de las, como mínimo, dos imágenes de origen, para calcular una imagen de varianza de las, como mínimo, dos imágenes de origen, para formar una imagen de filtrado a partir de la imagen de varianza y para obtener una imagen de resultado enmascarando la imagen promedio utilizando la imagen de filtrado como una máscara de bits.

30 La invención también se refiere a un procedimiento, que comprende: agarrar un objeto con un elemento de agarre unido a un brazo robótico; capturar como mínimo dos imágenes de origen que comprenden el objeto con un sensor de imagen; registrar un movimiento del elemento de agarre durante la captura de las, como mínimo, dos imágenes de origen; determinar como mínimo un primer vector de movimiento para un movimiento entre las, como mínimo, dos imágenes de origen en base al movimiento registrado del elemento de agarre; dividir como mínimo una de las, como mínimo, dos imágenes de origen en una serie de áreas de imagen; y determinar como mínimo un segundo vector de movimiento en base a una comparación de datos de imagen en las, como mínimo, dos imágenes de origen, representando el, como mínimo, un segundo vector de movimiento el movimiento de un área de imagen; y hacer coincidir el, como mínimo, un segundo vector de movimiento con el, como mínimo, un primer vector de movimiento, para obtener como mínimo un área de imagen para la clasificación de objetos.

35 La invención también se refiere a un aparato, que comprende: medios para agarrar un objeto con un elemento de agarre unido a un brazo robótico; medios para capturar como mínimo dos imágenes de origen que comprenden el objeto con un sensor de imagen; medios para registrar un movimiento del elemento de agarre durante la captura de las, como mínimo, dos imágenes de origen; medios para determinar como mínimo un primer vector de movimiento para un movimiento entre las, como mínimo, dos imágenes de origen en base al movimiento registrado del elemento de agarre; medios para dividir como mínimo una de las, como mínimo, dos imágenes de origen en una serie de áreas de imagen; y medios para determinar como mínimo un segundo vector de movimiento en base a una comparación de datos de imagen en las, como mínimo, dos imágenes de origen, representando el, como mínimo, un segundo vector de movimiento el movimiento de un área de imagen; y medios para hacer coincidir el, como mínimo, un segundo vector de movimiento con el, como mínimo, un primer vector de movimiento, para obtener como mínimo un área de imagen para la clasificación de objetos.

40 La invención también se refiere a un producto de programa informático o a un programa informático, que se implementa en un medio legible por ordenador. El programa informático o producto de programa informático comprende código para controlar un procesador para ejecutar un procedimiento que comprende: agarrar un objeto con un elemento de agarre unido a un brazo robótico; capturar, como mínimo, dos imágenes de origen que

comprenden el objeto; registrar un movimiento del elemento de agarre durante la captura de las, como mínimo, dos imágenes de origen; determinar, como mínimo, un primer vector de movimiento para un movimiento entre las, como mínimo, dos imágenes de origen en base al movimiento registrado del elemento de agarre; dividir, como mínimo, una de las, como mínimo, dos imágenes de origen en una serie de áreas de imagen; y determinar, como mínimo, un segundo vector de movimiento en base a una comparación de datos de imagen en las, como mínimo, dos imágenes de origen, representando el, como mínimo, un segundo vector de movimiento el movimiento de un área de imagen; y hacer coincidir el, como mínimo, un segundo vector de movimiento con el, como mínimo, un primer vector de movimiento, para obtener, como mínimo, un área de imagen para la clasificación de objetos.

La invención también se refiere a un procedimiento o a un aparato configurado para realizar el procedimiento o a un programa informático que comprende las etapas del procedimiento, comprendiendo el procedimiento: agarrar un objeto con un elemento de agarre, que está unido a un brazo robótico o montado por separado; capturar, utilizando un sensor de imagen, una serie de imágenes de origen de un área que comprenden el objeto; seleccionar elementos de imagen móviles de la serie de imágenes de origen en base a la correspondencia con el movimiento registrado del elemento de agarre durante el tiempo de captura de la serie de imágenes de origen; producir una imagen de resultado utilizando información sobre los elementos de información móviles seleccionados; y utilizar la imagen de resultado para clasificar el objeto agarrado. En una realización de la invención, la selección de los elementos de imagen móviles puede comprender calcular una imagen promedio de las, como mínimo, dos imágenes de origen, calcular una imagen de varianza de las, como mínimo, dos imágenes de origen y formar una imagen de filtrado a partir de la imagen de varianza. En una realización de la invención, la producción de la imagen de resultado puede comprender obtener una imagen de resultado enmascarando la imagen promedio utilizando la imagen de filtrado como una máscara de bits. En una realización de la invención, la selección de los elementos de imagen móviles puede comprender determinar, como mínimo, un primer vector de movimiento para un movimiento entre las, como mínimo, dos imágenes de origen en base al movimiento registrado del elemento de agarre, dividir, como mínimo, una de las, como mínimo, dos imágenes de origen en una serie de áreas de imagen, determinar, como mínimo, un segundo vector de movimiento en base a una comparación de datos de imagen en las, como mínimo, dos imágenes de origen, representando el, como mínimo, un segundo vector de movimiento el movimiento de un área de imagen y hacer coincidir el, como mínimo, un segundo vector de movimiento con el, como mínimo, un primer vector de movimiento.

En una realización de la invención, el sensor de imagen está configurado para moverse junto con el elemento de agarre, por ejemplo, el sensor de imagen puede estar unido al elemento de agarre o al brazo robótico.

En una realización de la invención, el sensor de imagen está situado en una posición que permite la obtención de, como mínimo, dos imágenes de origen de un área que comprende el objeto mientras el objeto se está moviendo.

En una realización de la invención, el aparato está configurado para reconocer una imagen del elemento de agarre en las, como mínimo, dos imágenes de origen. El aparato calcula como mínimo un desplazamiento para una imagen del elemento de agarre entre una primera imagen de origen y una segunda imagen de origen y determina una colocación mutua de la primera imagen de origen y la segunda imagen de origen como mínimo para las etapas de calcular una imagen promedio y calcular una imagen de varianza en base al desplazamiento. El desplazamiento puede utilizarse para desplazar la segunda imagen de origen para superponer de manera precisa las imágenes del objeto agarrado en las imágenes de origen primera y segunda. La imagen real del objeto agarrado puede eliminarse. El desplazamiento puede ser sólo lógico y utilizarse solo como un valor delta de desplazamiento en el cálculo de las imágenes promedio y de varianza.

En una realización de la invención, el aparato determina una colocación mutua de las, como mínimo, dos imágenes de origen, cuyo movimiento corresponde a un movimiento registrado del elemento de agarre, como mínimo, para las etapas de calcular una imagen promedio y calcular una imagen de varianza en base al desplazamiento.

En una realización de la invención, el aparato determina como mínimo un área móvil en las, como mínimo, dos imágenes de origen, cuyo movimiento corresponde a un movimiento registrado del elemento de agarre y el aparato filtra la, como mínimo, un área móvil a partir de las, como mínimo, dos imágenes de origen. El filtrado puede comprender ajustar los píxeles en las áreas a un valor predefinido tal como cero o uno. El área móvil puede estar comprendida en un bloque, un contorno o una forma o puede ser un solo píxel.

En una realización de la invención, se obtiene como mínimo un vector de movimiento que corresponde al movimiento del elemento de agarre entre dos imágenes de origen subsiguientes, entre las, como mínimo, dos imágenes de origen, a partir de la información de movimiento registrada para el brazo robótico. En una realización de la invención, el vector de movimiento puede obtenerse traducido en una función de correlación entre movimientos en las coordenadas del brazo robótico y las coordenadas de la imagen. El vector de movimiento puede utilizarse junto con la información de movimiento de píxel o bloque obtenida comparando las dos imágenes de origen subsiguientes. Estos bloques o píxeles con vectores de movimiento correspondientes al movimiento conocido del elemento de agarre pueden filtrarse de las, como mínimo, dos imágenes de origen. De este modo, puede ser posible eliminar totalmente el fondo móvil de las, como mínimo, dos imágenes de origen y restringirlo solo a objetos que se mueven con el elemento de agarre en las etapas de cálculo de la imagen promedio y de varianza. En el ejemplo de

un bloque móvil está un macrobloque de grupo de expertos de imágenes en movimiento (MPEG).

5 En una realización de la invención, se utiliza como mínimo una característica visual del elemento de agarre para que la entidad de reconocedor de objetos en el aparato reconozca el elemento de agarre en las, como mínimo, dos imágenes de origen. El movimiento de como mínimo parte de la imagen del elemento de agarre dentro de las, como mínimo, dos imágenes de origen se utiliza para obtener un vector de movimiento para el elemento de agarre que indica la magnitud y la dirección del movimiento del elemento de agarre. El vector de movimiento del elemento de agarre también es el vector de movimiento para el objeto agarrado, como mínimo parte del objeto del elemento de agarre, debido al hecho de que está sujetado por el elemento de agarre. Debe indicarse que en el caso de un objeto largo, el objeto puede tener partes que tiemblan, vibran o se quedan atrás en relación con la parte que se mantiene en contacto directo con el elemento de agarre. El vector de movimiento puede utilizarse para desplazar las, como mínimo, dos imágenes de origen en una relación entre sí que corresponde al vector inverso del vector de movimiento del elemento de agarre entre dos imágenes de origen respectivas. Los procedimientos adicionales que comprenden como mínimo calcular una imagen promedio de las, como mínimo, dos imágenes de origen, calcular una imagen de varianza de las, como mínimo, dos imágenes de origen, formar una imagen de filtrado a partir de la imagen de varianza, y obtener una imagen de resultado enmascarando la imagen promedio utilizando la imagen de filtrado como una máscara de bits pueden realizarse entonces para las, como mínimo, dos imágenes de origen que se han desplazado en proporción al vector inverso del vector de movimiento.

20 En una realización de la invención, una entidad de controlador de brazo en el aparato detecta un agarre satisfactorio del objeto. El aparato está conectado al elemento de agarre y al brazo robótico. En respuesta a la detección, la entidad de controlador de brazo indica esto a una entidad de control de la cámara en el aparato, que controla el sensor de imagen para capturar las, como mínimo, dos imágenes de origen. La entidad de control de la cámara obtiene las, como mínimo, dos imágenes de origen capturadas para el aparato a partir del sensor de imagen.

25 En una realización de la invención, una entidad de reconocedor de objetos en el aparato elimina como mínimo un área de imagen de la imagen de resultado que comprende características visuales de, como mínimo, uno del elemento de agarre y el brazo robótico.

30 En una realización de la invención, la entidad de reconocedor de objetos clasifica el objeto en la imagen de resultado en base a como mínimo una característica visual en la imagen de resultado e indica a la entidad de control del brazo que haga que el brazo robótico mueva el objeto a una ubicación correspondiente a la clasificación.

35 En una realización de la invención, la etapa de calcular una imagen de varianza comprende calcular una varianza de cada píxel respectivo en las, como mínimo, dos imágenes de origen; y ajustar la varianza calculada como el valor del píxel respectivo en la imagen de varianza.

40 En una realización de la invención, por una imagen de varianza se entiende una imagen que para cada píxel tiene un valor que es proporcional a la variación de valores en ese píxel en las, como mínimo, dos imágenes de origen. Un modo de medir la variación del valor en un píxel en las, como mínimo, dos imágenes de origen es calcular la varianza estadística.

45 En una realización de la invención, la etapa de formar una imagen de filtrado comprende ajustar cada píxel respectivo a 1 en la imagen de filtrado para la que el píxel respectivo en la imagen de varianza tiene un valor por debajo de un valor de umbral predefinido.

50 En una realización de la invención, la etapa de obtener una imagen de resultado comprende seleccionar cada píxel respectivo para la imagen de resultado a partir de la imagen promedio solo si el valor del píxel respectivo en la imagen de filtrado es 1. La imagen de filtrado puede entenderse como una máscara de bits bidimensional que se utiliza para permitir la selección de píxeles a partir de la imagen promedio.

55 En una realización de la invención, las, como mínimo, dos imágenes se toman contra un fondo que comprende otros objetos que interfieren potencialmente en el reconocimiento del objeto agarrado. El fondo del objeto puede ser un escenario no estructurado, es decir, un entorno o, en general, un espacio tridimensional, que no está predeterminado en una o varias de sus características, tales como color y geometría del fondo, y que puede incluir, además de los objetos de interés, otros objetos de características desconocidas. Por ejemplo, un montón de residuos podría constituir un escenario no estructurado, es decir, un espacio de funcionamiento de un robot. Un escenario no estructurado también puede cambiar a medida que avanza el tiempo. Por ejemplo, cuando los fragmentos de residuos se eliminan del montón, los fragmentos de residuos pueden mezclarse, moverse o caerse. También pueden añadirse residuos nueva al montón.

60 En una realización de la invención, el sensor de imagen es como mínimo uno de una cámara, una cámara de infrarrojos y un escáner láser.

65 En una realización de la invención, las etapas de calcular una imagen promedio, calcular una imagen de varianza y formar una imagen de filtrado se realizan por separado para cada canal de color de píxel. En una realización de la

invención, las, como mínimo, dos imágenes de origen se convierten en escala de grises y las etapas de calcular una imagen promedio, calcular una imagen de varianza y formar una imagen de filtrado se realizan para la escala de grises.

5 En una realización de la invención, el sensor de imagen está unido al elemento de agarre o forma una parte del elemento de agarre. La cámara también puede estar unida al extremo del brazo robótico directamente o a través de un eje.

10 En una realización de la invención, la invención incluye un brazo robótico controlado por una unidad de control e instalado de modo que pueda alcanzar objetos que se encuentran en un escenario no estructurado.

15 En una realización de la invención, el sistema incluye además un elemento de agarre unido al brazo robótico y controlado por la unidad de control. El elemento de agarre puede ser, por ejemplo, un dispositivo que agarra objetos encerrándolos, de un modo que se asemeja a una mano o una garra.

En una realización de la invención, el sistema incluye además como mínimo un dispositivo sensor que puede utilizarse para producir datos sensoriales sobre el escenario no estructurado. Un dispositivo sensor de este tipo puede ser, por ejemplo, una cámara digital orientada para visualizar el escenario no estructurado.

20 En una realización de la invención, el elemento de agarre incluye sensores que pueden utilizarse para medir si el elemento de agarre está en contacto con objetos en el escenario no estructurado, por ejemplo cuando el elemento de agarre se mueve y choca contra un objeto, o cuando se agarra un objeto.

25 En una realización de la invención, el éxito de la operación de agarre se determina utilizando datos procedentes de los sensores. Si el agarre no es satisfactorio, el brazo robótico se mueve entonces a una ubicación diferente para otro intento.

30 En una realización de la invención, el sistema se mejora adicionalmente utilizando sistemas de aprendizaje, que pueden ejecutarse en el aparato.

35 En una realización de la invención, el programa informático se almacena en un medio legible por ordenador. El medio legible por ordenador puede ser una tarjeta de memoria extraíble, un módulo de memoria extraíble, un disco magnético, un disco óptico, una memoria holográfica o una cinta magnética. Un módulo de memoria extraíble puede ser, por ejemplo, una memoria extraíble USB, una tarjeta PCMCIA o una tarjeta de memoria inteligente.

40 Las realizaciones de la invención descritas anteriormente en el presente documento pueden utilizarse en cualquier combinación entre sí. Varias de las realizaciones pueden combinarse entre sí para formar una realización adicional de la invención. Un procedimiento, un sistema, un aparato, un programa informático o un producto de programa informático al que se refiere la invención puede comprender como mínimo una de las realizaciones de la invención descrita anteriormente en el presente documento.

45 Los beneficios de la invención se refieren a la calidad mejorada en la selección de objetos de un espacio de funcionamiento de un robot. La invención también puede utilizarse para simplificar procedimientos de reconocimiento de objetos objetivo adicionales que se utilizan, por ejemplo, para reconocer la forma o la textura de un objeto objetivo. La invención también reduce los movimientos requeridos de un brazo robótico al evitar el movimiento de una imagen a un fondo en blanco y, por tanto, puede conducir a un consumo de energía reducido por un brazo robótico.

BREVE DESCRIPCIÓN DE LOS DIBUJOS:

50 Los dibujos adjuntos, que se incluyen para proporcionar una comprensión adicional de la invención y constituyen una parte de esta memoria descriptiva, ilustran realizaciones de la invención, y junto con la descripción, ayudan a explicar los principios de la invención. En los dibujos:

55 la **figura 1** es un diagrama de bloques que ilustra un sistema robótico que realiza el filtrado de imágenes de objetos objetivo en una realización de la invención;

la **figura 2A** es un diagrama de bloques que ilustra el filtrado de imágenes de objetos objetivo en un sistema robótico;

60 la **figura 2B** es un diagrama de bloques que ilustra el desplazamiento de imágenes de objetos objetivo en base a un vector de movimiento de un elemento de agarre o brazo robótico en un sistema robótico, tal como el sistema robótico en una realización de la invención;

la **figura 3** es un diagrama de flujo que ilustra un procedimiento para el filtrado de imágenes de objetos objetivo en un sistema robótico en una realización de la invención; y

65 la **figura 4** es un diagrama de flujo que ilustra un procedimiento para un procedimiento de reconocimiento de objetos basado en el movimiento de objetos en un sistema robótico en una realización de la invención.

DESCRIPCIÓN DETALLADA DE LAS REALIZACIONES:

Ahora se hará referencia en detalle a las realizaciones de la presente invención, cuyos ejemplos se ilustran en los dibujos adjuntos.

5 La figura 1 es un diagrama de bloques que ilustra un sistema robótico que realiza el filtrado de imágenes de objetos objetivo en una realización de la invención. En la figura 1, el sistema robótico 100 comprende un robot 110, por ejemplo, un robot industrial que comprende un brazo robótico 116. Al brazo robótico 116 está conectado un elemento de agarre 112, que también puede ser una pinza o una garra. El brazo robótico 116 puede mover el elemento de agarre 112 dentro de un área de funcionamiento 102. El brazo robótico 116 puede comprender varios motores, por ejemplo, servomotores que permiten que se controle la rotación, la elevación y el agarre de los brazos robóticos. Pueden efectuarse diversos movimientos del brazo robótico 116 y el elemento de agarre 112 mediante dispositivos de accionamiento. A modo de ejemplo, los dispositivos de accionamiento pueden ser eléctricos, neumáticos o hidráulicos, o cualquier combinación de ellos. Los dispositivos de accionamiento pueden mover o rotar diversos elementos del robot 110. Puede utilizarse un conjunto de controladores eléctricos para convertir las señales de proceso de datos, dicho de otro modo, las instrucciones procedentes del aparato 120 en niveles apropiados de tensión y energía para controlar los dispositivos de accionamiento del brazo robótico 116. En respuesta a las señales de control procedentes del aparato 120, los dispositivos de accionamiento realizan diversas funciones mecánicas incluyendo, pero sin limitarse necesariamente a: situar el elemento de agarre 112 en una ubicación específica dentro del área de funcionamiento 102, bajar o subir el elemento de agarre 112, y cerrar y abrir el elemento de agarre 112. El robot 110 puede comprender diversos sensores. A modo de ejemplo, los sensores comprenden diversos sensores de posición (no mostrados) que indican la posición del brazo robótico 116 y el elemento de agarre 112, así como el estado abierto/cerrado del elemento de agarre 112. El estado abierto/cerrado del elemento de agarre no está restringido a un simple bit de sí/no. En una realización de la invención, el elemento de agarre 112 puede indicar un estado abierto/cerrado de múltiples bits con respecto a cada uno de sus dedos, por lo que puede obtenerse una indicación del tamaño y/o la forma del objeto u objetos en el elemento de agarre. Además de los sensores de posición, el conjunto de sensores puede comprender sensores de deformación, también conocidos como galgas extensiométricas o sensores de retroalimentación de fuerza, que indican la deformación experimentada por diversos elementos del brazo robótico 116 y el elemento de agarre 112. En un ejemplo de implementación ilustrativo pero no restrictivo, los sensores de deformación comprenden resistencias variables cuya resistencia varía dependiendo de la tensión de compresión que se les aplique. Debido a que los cambios en la resistencia son pequeños en comparación con el valor absoluto de la resistencia, las resistencias variables normalmente se miden en una configuración de puente de Wheatstone.

35 Al elemento de agarre 112 o al brazo robótico 116 está conectada una cámara 114, que se dirige para tener en su campo visual objetos agarrados por el elemento de agarre 112, como mínimo parcialmente. En la figura 1, se ilustra que la cámara está dentro del elemento de agarre 112. La cámara también puede estar ubicada en un eje independiente conectado al brazo robótico 116 y estar situada de modo que los objetos agarrados por el elemento de agarre 112 queden dentro del campo visual de la cámara 114. La cámara también puede estar ubicada en una posición alejada independiente del brazo robótico 116. El robot 110 ordena los objetos contenidos en un escenario no estructurado 102, es decir, en su escenario de funcionamiento. El escenario 102 comprende varios objetos tales como los objetos 103, 104 y 105. En la figura 1 se muestra que el robot 110 ha realizado una operación de agarre en el objeto 105 objetivo y lo sujeta en el elemento de agarre 112.

45 El robot 110 está conectado a un aparato de proceso de datos 120, en resumen, un aparato. Las funciones internas del aparato 120 se ilustran con el recuadro 140. El aparato 120 comprende como mínimo un procesador 142, una memoria de acceso aleatorio (RAM) 148 y un disco duro 146. El uno o varios procesadores 142 controlan el brazo robótico ejecutando las entidades de software 150, 152, 154 y 156. El aparato 120 también comprende como mínimo una interfaz periférica 145 de la cámara y una interfaz 144 del robot para controlar el robot 110. La interfaz periférica 145 puede ser un bus, por ejemplo, un bus de serie universal (USB). Al aparato 120 también está conectado un terminal 130, que comprende como mínimo una pantalla y un teclado. El terminal 130 puede ser un ordenador portátil conectado utilizando una red de área local al aparato 120.

55 Con el fin de comunicarse con dispositivos externos, tales como el robot 110, el aparato 120 comprende o utiliza un conjunto de circuitos externos de recepción/transmisión tal como la interfaz 144 del robot, que comprende un conjunto de circuitos de transmisión, un conjunto de circuitos de recepción y puede comprender una antena interna o externa (no mostrada). El aparato 120 puede utilizar varias tecnologías de interconexión diferentes para comunicarse con el mundo físico, que en el presente ejemplo comprende el robot 110, el elemento de agarre 112 y la cámara 114. Las redes de área local inalámbricas (WLAN) y las interfaces inalámbricas de corto alcance, tales como por infrarrojos, radio o Bluetooth, son ejemplos ilustrativos pero no restrictivos de tales conjuntos de circuitos inalámbricos de recepción/transmisión. En lugar de tales tecnologías de comunicación inalámbricas o además de ellas, el aparato de proceso de datos puede utilizar conexiones por cable, tales como un USB, cualquier interfaz paralela o en serie, u otros tipos de interfaces convencionales en la industria o patentadas.

65 Además, la memoria 140 del aparato 120 contiene un conjunto de programas o, en general, entidades de software que se ejecutan por el, como mínimo, un procesador 142. Existe una entidad 150 de controlador de brazo que emite

instrucciones a través de la interfaz 144 del robot al robot 110 con el fin de controlar la rotación, la elevación y el agarre del brazo robótico 116 y el elemento de agarre 112. La entidad 150 de controlador de brazo también puede recibir datos de sensor referidos a la rotación, la elevación y el agarre medidos del brazo robótico 116 y el elemento de agarre 112. El controlador del brazo puede activar el brazo con nuevas instrucciones emitidas en base a la retroalimentación recibida al aparato 120 a través de la interfaz 144. El controlador 150 del brazo está configurado para emitir instrucciones al robot 110 para realizar operaciones de alto nivel bien definidas. Un ejemplo de una operación de alto nivel es mover el brazo robótico a una posición especificada. El controlador 150 del brazo puede utilizar diversos controladores de software, rutinas o bibliotecas de enlace dinámico para convertir la operación de alto nivel en una serie de operaciones de bajo nivel, tales como la emisión de una secuencia apropiada de señales de salida a través de controladores eléctricos a dispositivos de accionamiento del robot 110. La entidad 152 de controlador de cámara se comunica con la cámara 114 utilizando la interfaz 145. La entidad de controlador de cámara hace que la cámara 114 tome varias fotografías a intervalos de tiempo predefinidos comenzando en un momento en el tiempo indicado por la entidad 152 de controlador de cámara. En general, el controlador 152 de la cámara puede emitir una instrucción a la cámara 114 para que tome una fotografía en cualquier momento en el tiempo. La entidad 152 de controlador de cámara obtiene las imágenes tomadas por la cámara 114 a través de la interfaz 145 y almacena las imágenes en la memoria 140. La entidad 154 de extractor de objetos está configurada para extraer un objeto objetivo de un número predefinido de imágenes de origen (no mostradas). En una realización de la invención, la entidad 154 de extractor de objetos utiliza las imágenes de origen para calcular una imagen promedio 160 y una imagen de varianza 162. En una realización de la invención, a partir de la imagen de varianza 152, la entidad 154 de extractor de objetos calcula una imagen de máscara (no mostrada), que se utiliza para formar una imagen de resultado 164 enmascarando la imagen promedio 160 con un mapa de bits bidimensional constituido por la imagen de máscara. El objeto objetivo filtrado a partir de las imágenes de origen tal como aparece en la imagen de resultado 164 se proporciona adicionalmente a un reconocedor de objetos 156, que puede realizar análisis adicionales en el objeto objetivo en base a diversas características visuales del objeto objetivo tal como una forma, un color, una textura, varios coeficientes de transformada discreta del coseno (DCT), varios coeficientes de transformada de tren de ondas, un macrobloque de MPEG y un contorno. El reconocedor de objetos también puede comprender información sobre las características visuales de una imagen del elemento de agarre 112 y puede utilizar esa información para eliminar partes del elemento de agarre 112 visibles en la imagen de resultado 164 con el fin de producir una imagen de resultado incluso mejor para el análisis adicional. Las características visuales de una imagen del elemento de agarre 112 pueden comprender como mínimo una de una forma, un color, una textura, varios coeficientes de transformada discreta del coseno (DCT), varios coeficientes de transformada de tren de ondas, un macrobloque de MPEG y un contorno. En base a los resultados en el reconocimiento de objetos objetivo, el reconocedor de objetos 156 puede clasificar el objeto objetivo e indicar al robot 110 a través del controlador 150 del brazo que mueva el objeto objetivo a una ubicación específica que corresponde, por ejemplo, a un compartimento objetivo.

El filtrado realizado por la entidad 154 de extractor de objetos se explica a continuación en el presente documento en una realización a modo de ejemplo. En una realización de la invención, en la vista de la cámara 114, se visualiza el objeto objetivo 105 contra el fondo del entorno de agarre 102. Sin embargo, puesto que se conoce el movimiento del brazo robótico 116, ya sea porque es una operación predeterminada o porque el sistema de control del robot lo mide con precisión, se conoce el movimiento del elemento de agarre 112 y, por tanto, el movimiento del objeto objetivo 105 en la imagen de cámara, mientras que el fondo de la vista del sensor de la cámara se está moviendo de un modo diferente. Entonces puede reconocerse con más precisión el objeto objetivo 112 a partir de la imagen de la cámara seleccionando de los datos de imagen solo aquellas áreas que cambian en el modo correspondiente al movimiento conocido del brazo robótico. Una implementación de ejemplo sencilla de esto tiene la cámara 114 unida al elemento de agarre 112. En una configuración de este tipo, los objetos que se mueven con el elemento de agarre parecen estacionarios en la imagen de cámara, mientras que el fondo parece estar en movimiento. El reconocimiento de objetos a partir de los datos de imagen de la cámara puede implementarse de un modo "en línea". A medida que pasa el tiempo se reciben nuevas imágenes de cámara procedentes de la cámara. Dependiendo del tipo de cámara utilizada, podrían construirse nuevas imágenes de cámara como resultado de que el sistema de reconocimiento de objetos solicita nuevos datos de imagen de la cámara, o la cámara puede construir nuevas imágenes a una determinada velocidad interna a la cámara y el sistema de reconocimiento de objetos o bien puede solicitar entonces la última imagen, que podría ser la misma imagen que la solicitada anteriormente, o bien el sistema de reconocimiento de objetos podría recibir una indicación de la cámara cuando están disponibles nuevos datos de imagen. A medida que se reciben nuevas imágenes, se calculan el promedio de ejecución y la varianza de los valores de píxel en la nueva imagen y las imágenes recibidas previamente, que puede pensarse que son la formación de dos imágenes, la imagen promedio y la imagen de varianza. El número de imágenes recibidas previamente utilizadas puede fijarse como parámetro para adaptarse mejor a la aplicación en cuestión. Cuando se ha recibido y procesado un número suficiente de imágenes de cámara de esta manera, se calcula una imagen de máscara ajustando el valor de píxel a uno en aquellos píxeles en los que el píxel correspondiente en la imagen de varianza es más pequeño que un umbral predefinido y a cero en caso contrario. Entonces, la imagen promedio se utiliza para realizar el reconocimiento de objetos, dando como resultado una imagen de característica, que incluye aquellos píxeles en la imagen promedio que se considera que tienen alguna característica visual que el sistema de reconocimiento de objetos está configurado para reconocer. Por ejemplo, si el sistema de reconocimiento de objetos está configurado para reconocer objetos rojos, la imagen de característica incluiría aquellos píxeles de la imagen promedio que el sistema de reconocimiento de objetos considera que son "rojos". Una vez que se ha calculado la

imagen de característica, se calcula una imagen final enmascarando la imagen de característica con la imagen de máscara. Esto se lleva a cabo seleccionando de la imagen de característica solo aquellos píxeles para los que los píxeles correspondientes en la imagen de máscara tienen el valor de píxel de 1. Por tanto, la parte de la imagen correspondiente al objeto agarrado se reconoce fácilmente a partir de la imagen final. Una vez que se ha reconocido la parte de los datos de imagen de la cámara que corresponde al objeto agarrado, el sistema de reconocimiento de objetos puede utilizar los datos para determinar el tipo del objeto agarrado mientras el objeto se está moviendo. Esta nueva información de tipo puede utilizarse entonces en combinación con la información de tipo determinada previamente para seleccionar la operación realizada en el objeto.

Cuando como mínimo un procesador ejecuta entidades funcionales asociadas con la invención, una memoria comprende entidades tales como la entidad 150 de controlador de brazo, la entidad 152 de controlador de cámara, la entidad 154 de extractor de objetos y la entidad 156 de reconocedor de objetos. Las entidades funcionales dentro del aparato 120 ilustrado en la figura 1 pueden implementarse de varios modos. Pueden implementarse como procesos ejecutados bajo el sistema operativo nativo del nodo de red. Las entidades pueden implementarse como procesos o subprocesos independientes o de modo que se implementen varias entidades diferentes por medio de un proceso o subproceso. Un proceso o un subproceso puede ser la instancia de un bloque de programa que comprende varias rutinas, es decir, por ejemplo, procedimientos y funciones. Las entidades funcionales pueden implementarse como programas informáticos independientes o como un solo programa informático que comprende varias rutinas o funciones que implementan las entidades. Los bloques de programa se almacenan como mínimo en un medio legible por ordenador tal como, por ejemplo, un circuito de memoria, tarjeta de memoria, disco magnético u óptico. Algunas entidades funcionales pueden implementarse como módulos de programa asociados a otra entidad funcional. Las entidades funcionales en la figura 1 también pueden almacenarse en memorias independientes y ejecutarse por procesadores independientes, que se comunican, por ejemplo, a través de un bus de mensajes o una red interna dentro del nodo de red. Un ejemplo de un bus de mensajes de este tipo es el bus de interconexión de componentes periféricos (PCI).

En una realización de la invención, las entidades de software 150 - 156 pueden implementarse como entidades de software independientes tales como, por ejemplo, subrutinas, procesos, subprocesos, procedimientos, objetos, módulos y secuencias de código de programa. También pueden ser solo funcionalidades lógicas dentro del software en el aparato 120, que no se han agrupado en ninguna subrutina, proceso, subproceso, procedimiento, objeto, módulo y secuencia de código de programa independiente específico. Sus funciones pueden distribuirse por todo el software del aparato 120. Algunas funciones pueden realizarse en el sistema operativo del aparato 120.

En un ejemplo ilustrativo pero no restrictivo, el escenario no estructurado 102 es una cinta transportadora, o la parte de la cinta transportadora que cruza el área de funcionamiento del robot. El aparato 120 tiene poca o ninguna información a priori sobre los objetos 103, 104 y 105 dentro del escenario no estructurado 102, tal como el tamaño, la forma y/o el color de los objetos de interés. En algunas realizaciones de la invención, el aparato 120 puede tener alguna información a priori sobre los objetos de interés, o puede haber obtenido información sobre los objetos mediante el aprendizaje, pero como mínimo a priori normalmente se desconoce el fondo (otros objetos), la posición y la orientación de los objetos de interés. Es decir, los objetos 103, 104 y 105 pueden estar en posiciones y orientaciones aleatorias en el escenario no estructurado 102, y los objetos pueden superponerse entre sí.

Las realizaciones de la invención descritas anteriormente en el presente documento con respecto a la figura 1 pueden utilizarse en cualquier combinación entre sí. Varias de las realizaciones pueden combinarse entre sí para formar una realización adicional de la invención.

La figura 2A es un diagrama de bloques que ilustra el filtrado de imágenes de objetos objetivo en una realización de la invención, en un sistema robótico tal como el sistema robótico 100 ilustrado en la figura 1. El punto de partida en la figura 2 es que el elemento de agarre 112 conectado al brazo robótico 116 ha agarrado satisfactoriamente un objeto, tal como el objeto 105 ilustrado en la figura 1. A continuación, el brazo robótico 116 comienza a mover el objeto 105 hacia una dirección dada. Durante el movimiento del brazo, la cámara 114 toma una secuencia de cinco imágenes, concretamente, las imágenes 250, 252, 254, 256 y 258 de cámara, en este orden. El agarre satisfactorio del objeto 105 y el inicio del movimiento del brazo robótico 116 pueden actuar como desencadenante del inicio de la toma de una secuencia de imágenes de cámara. Las imágenes se toman a intervalos de tiempo predefinidos que pueden ser, por ejemplo, de 100 milisegundos a 5 segundos. Por ejemplo, el intervalo de tiempo puede ser de 0,5 segundos. La cámara 114 se coloca de modo que el objeto 105 cabe en su campo visual. El objeto 105 es visible en la imagen 250 de cámara como el recuadro 280. Dos dedos del elemento de agarre 112 son visibles en la imagen 250 de cámara como los rectángulos 282 y 284. En la imagen 250 de cámara también son visibles objetos de fondo tales como el objeto 285. El movimiento del brazo robótico es visible como cambios en las imágenes 250 - 258 de cámara y se ilustra con la flecha 286. La dirección es hacia abajo en relación con los objetos de fondo, tales como el objeto 285. La velocidad es de un píxel por imagen, lo que totaliza un movimiento de cuatro píxeles en la secuencia de imágenes que consiste en las imágenes 250, 252, 254, 256 y 258 de cámara. Debe indicarse que las imágenes de cámara ilustradas en la figura 2 están altamente simplificadas en comparación con las imágenes de cámara reales. Las imágenes de cámara se ilustran para resaltar el procedimiento de la invención.

Tal como se ilustra con las flechas marcadas 201, el aparato 120, por ejemplo, utilizando la entidad 154 de extractor

de objetos, calcula una imagen promedio 260 a partir de las imágenes 250, 252, 254, 256 y 258 de cámara. La fórmula para un valor de píxel p dado en las coordenadas x,y en la imagen promedio se calcula utilizando

$$P_{(x,y)} = \frac{\sum_{i=1}^n P_{i(x,y)}}{n} ,$$

en la que i es un índice para las imágenes de cámara y n es el número de imágenes de cámara. En el caso de la figura 2, se supone que n = 5. El número n puede tener cualquier valor de número entero, por ejemplo, $3 < n < 20$. Las escalas de grises en la figura 2 ilustran los valores de píxel calculados.

Tal como se ilustra con las flechas marcadas 202, el aparato 120, por ejemplo, utilizando la entidad 154 de extractor de objetos, calcula una imagen de varianza 262 a partir de las imágenes 250, 252, 254, 256 y 258 de cámara. La fórmula para un valor de píxel p dado en las coordenadas x,y en la imagen de varianza se calcula utilizando

$$P_{(x,y)} = \frac{1}{n} \left(\sum_{i=1}^n P_{i(x,y)}^2 \right) - \left(\frac{1}{n} \left(\sum_{i=1}^n P_{i(x,y)} \right) \right)^2 ,$$

en la que i es un índice para las imágenes de cámara y n es el número de imágenes de cámara. Las escalas de grises en la figura 2 ilustran los valores de píxel calculados.

En una realización de la invención, se forman una imagen promedio independiente y una imagen de varianza independiente para cada canal de color, es decir, por ejemplo, para los canales R, G y B. En una realización de la invención, las imágenes de cámara se convierten en imágenes en escala de grises y solo se forman imágenes promedio y de varianza únicas.

A continuación, la entidad 154 de extractor de objetos calcula una imagen de máscara 264, tal como se ilustra con la flecha 203. La imagen de máscara 264 se obtiene ajustando el valor de píxel en un píxel dado $p_{(x,y)}$ a 1 en la imagen de máscara 264, si el valor de ese píxel está por debajo de un umbral predefinido. En caso contrario, el valor en ese píxel se ajusta a 0.

A continuación, la entidad 154 de extractor de objetos utiliza la imagen de máscara 264 para eliminar de la imagen promedio 260 aquellos píxeles $p_{(x,y)}$ que tenían un valor de 0 en la ubicación x, y en la imagen de máscara 264, tal como se ilustra con las flechas 204A y 204B. Los píxeles $p_{(x,y)}$ eliminados se ajustan a cero en la imagen de resultado 266. Otros píxeles respectivos en la imagen de resultado 266 se ajustan a los valores de píxeles respectivos obtenidos a partir de la imagen promedio 260. El resultado de la operación de enmascaramiento se almacena como la imagen de resultado 266. Finalmente, la entidad 156 de reconocedor de objetos puede realizar un proceso adicional para la imagen de resultado 266. Por ejemplo, la entidad 156 de reconocedor de objetos puede eliminar las características visuales pertenecientes a los dedos del elemento de agarre 112, en base a la textura conocida de los dedos del elemento de agarre.

La figura 2B es un diagrama de bloques que ilustra el desplazamiento de imágenes de objetos objetivo en base a un vector de movimiento de un elemento de agarre o un brazo robótico en un sistema robótico, tal como el sistema robótico 100 ilustrado en la figura 1, en una realización de la invención.

En la figura 2B existe una cámara montada de manera independiente del brazo robótico y el elemento de agarre en una posición donde puede capturar una secuencia temporal de imágenes de cámara que comprende como mínimo las imágenes 291, 292, 293 y 294 de cámara de un objeto 231 que está portándose por un elemento de agarre tal como el elemento de agarre 112. El orden numérico de los números de referencia 291 - 294 indica un posible orden de la captura de las imágenes 291 - 294 de cámara. Después de la captura de las imágenes 291 - 294 de cámara, se utiliza la información predefinida sobre como mínimo una característica visual que comprende como mínimo uno de un color, una textura, varios coeficientes de transformada discreta del coseno (DCT), varios coeficientes de transformada de tren de ondas, un macrobloque de MPEG y un contorno para reconocer una imagen 230 de un elemento de agarre, tal como el elemento de agarre 112 en la imágenes 291 - 294 de cámara mediante, por ejemplo, la entidad 154 de extractor de objetos en el aparato 120. El movimiento de la imagen 230 del elemento de agarre dentro de las imágenes 291 - 294 de cámara se utiliza para obtener los vectores de movimiento 221, 222 y 223 para la imagen 230 del elemento de agarre. Los vectores de movimiento 221 - 223 de la imagen 230 del elemento de agarre también son los vectores de movimiento para el objeto agarrado 231, es decir, como mínimo parte del objeto agarrado 231, en una realización de la invención. Debe indicarse que en el caso de un objeto largo, el objeto puede tener partes que tiemblan, vibran o se quedan atrás. Los vectores de movimiento 221, 222 y 223 se utilizan para obtener vectores de movimiento inversos 224, 225 y 226 respectivos. La imagen 230 del elemento de agarre puede filtrarse a partir de las imágenes 291 - 294 de cámara, manteniendo de ese modo solo una imagen del objeto agarrado 231. El filtrado puede utilizar información de características visuales almacenada en la memoria 140 utilizada por la entidad 154 de extractor de objetos para eliminar las áreas que coinciden con la información de características visuales. Los vectores de movimiento inversos 224, 225 y 226 se utilizan para desplazar las imágenes 292, 293 y 294 de cámara para obtener imágenes 296, 297 y 298 de cámara desplazadas respectivas. En la secuencia de imágenes 295 - 298 de cámara desplazadas, otros objetos tales como el objeto 232 parecen estar en movimiento mientras que el objeto agarrado 231 parece estacionario. Debe indicarse que, en una realización de la invención, las imágenes 295 - 298 de cámara desplazadas pueden no formarse realmente en la memoria 140,

pero los vectores de movimiento inversos 224 - 226 pueden utilizarse, por ejemplo, en las etapas de calcular una imagen promedio y calcular una imagen de varianza como información de desplazamiento de imagen, cuando se leen valores de píxel de diferentes imágenes de cámara. En una realización de la invención, solo se puede mantener un área correspondiente a la intersección del área de las imágenes 295 - 298 de cámara. En una realización de la invención, las imágenes 295 - 298 de cámara se pueden modificar con píxeles de valor único, tales como píxeles de valor cero para las nuevas áreas de imagen que aparecen debido al desplazamiento. Las imágenes en cambio también pueden reordenarse. Por tanto, puede obtenerse un resultado equivalente al caso de una cámara montada en el elemento de agarre o el brazo robótico para los cálculos posteriores. A continuación, pueden seguirse las etapas de calcular una imagen promedio, calcular una imagen de varianza y la formación de una máscara de filtrado, tal como se describe en asociación con la figura 2A.

Las realizaciones de la invención descritas anteriormente en el presente documento con respecto a las figuras 2A y 2B pueden utilizarse en cualquier combinación entre sí. Varias de las realizaciones pueden combinarse entre sí para formar una realización adicional de la invención.

La figura 3 es un diagrama de flujo que ilustra un procedimiento para el filtrado de imágenes de objetos objetivo en un sistema robótico en una realización de la invención. El procedimiento puede aplicarse en un sistema robótico tal como se ilustra en la figura 1.

En la etapa 300, se agarra un objeto objetivo físico utilizando un elemento de agarre unido a un brazo robótico.

En la etapa 302, el objeto físico en el elemento de agarre se mueve con, como mínimo, uno del elemento de agarre y el brazo robótico.

En la etapa 304 se capturan una serie de imágenes utilizando una cámara conectada, por ejemplo, al brazo robótico del elemento de agarre. La cámara también puede estar colocada a una distancia del brazo robótico y el elemento de agarre. Las imágenes se capturan mientras el objeto objetivo está en movimiento. El objeto objetivo puede moverse sobre un fondo que comprende una serie de otros objetos que pueden someterse a manipulación posterior por el brazo robótico o que pueden ignorarse en la clasificación. El movimiento del brazo robótico puede comprender un movimiento inicial hacia un posible destino para el objeto objetivo, tal como varios compartimentos o estantes de ordenación. El fondo puede ser un escenario no estructurado. El número de imágenes puede ser, por ejemplo de 4, 5, 10 o un número natural arbitrario mayor de 1.

En la etapa 306 se calcula una imagen promedio de la serie de imágenes. La imagen promedio se puede calcular,

por ejemplo, utilizando la fórmula
$$P_{(x,y)} = \frac{\sum_{i=1}^n P_{i(x,y)}}{n}$$
, en la que i es un índice para las imágenes de cámara, n es el número de imágenes de cámara y x,y son las coordenadas del píxel.

En la etapa 308 se calcula una imagen de varianza de la serie de imágenes. La imagen de varianza se puede

calcular, por ejemplo, utilizando la fórmula
$$P_{(x,y)} = \frac{1}{n} \left(\sum_{i=1}^n P_{i(x,y)}^2 \right) - \left(\frac{1}{n} \left(\sum_{i=1}^n P_{i(x,y)} \right) \right)^2$$
, en la que i es un índice para las imágenes de cámara, n es el número de imágenes de cámara y x,y son las coordenadas del píxel.

En la etapa 310 se forma una imagen de máscara de filtración a partir de la imagen de varianza.

En una realización de la invención, la imagen de máscara de filtrado se obtiene ajustando el valor de píxel en un píxel dado $p_{(x,y)}$ a 1 en la imagen de máscara de filtrado, si el valor de ese píxel está por debajo de un umbral predefinido. En caso contrario, el valor en ese píxel se ajusta a 0.

En la etapa 312 se obtiene una imagen filtrada que comprende el objeto objetivo enmascarando la imagen promedio utilizando la máscara de filtrado.

En una realización de la invención, la imagen de máscara se utiliza para eliminar de la imagen promedio aquellos píxeles $p_{(x,y)}$ que tenían un valor 0 en x,y en la imagen de máscara. Esos píxeles $p_{(x,y)}$ se ajustan a cero. El resultado de la operación de enmascaramiento se almacena como la imagen de resultado, es decir, la imagen filtrada. Dicho de otro modo, la máscara de filtrado es una máscara bidimensional de bits que se utiliza junto con la imagen promedio en una operación que devuelve un valor de píxel a partir de la imagen promedio si el valor del píxel correspondiente en la máscara de bits es 1. Esto puede formularse de la siguiente manera formal:

$$\left\{ \begin{array}{l} \text{si } \text{máscara}_{(x,y)} = 1 \Rightarrow \text{resultado}_{(x,y)} := \text{prom.}_{(x,y)} \\ \text{si } \text{máscara}_{(x,y)} = 0 \Rightarrow \text{resultado}_{(x,y)} := 0 \end{array} \right. ,$$

5 en la que máscara(x,y) representa un píxel en la imagen de máscara, resultado(x,y) representa un píxel en la imagen de resultado y prom(x, y) representa un píxel en la imagen promedio. El símbolo “:=” representa la operación de asignación de valor.

10 En una realización de la invención, en la etapa 314, se eliminan las áreas de imagen que tienen la textura del elemento de agarre o el brazo robótico de la imagen filtrada para facilitar de ese modo un mejor reconocimiento, por ejemplo, de la forma del objeto objetivo. A continuación, termina el procedimiento.

15 En una realización de la invención, los datos de entrada recibidos desde la cámara son una imagen digital que consiste en una matriz de píxeles bidimensional, teniendo cada píxel un valor numérico para la componente de color rojo, verde y azul, designadas a continuación en el presente documento como los valores de R, G y B, respectivamente. El número de píxeles en los datos corresponde a la resolución de la cámara.

20 En una realización de la invención, para disminuir la cantidad de cálculos, se realiza un submuestreo de los datos de imagen recibidos desde la cámara a una resolución determinada adecuada para el análisis.

25 En una realización de la invención, se normaliza entonces la imagen submuestreada resultante para tener en cuenta los cambios en las condiciones de iluminación. Entonces se puede realizar la normalización de manera independiente para cada píxel en la imagen submuestreada.

30 En una realización de la invención, el aparato está configurado para reconocer una imagen del elemento de agarre en las, como mínimo, dos imágenes de origen. El aparato calcula como mínimo un desplazamiento para una imagen del elemento de agarre entre una primera imagen de origen y una segunda imagen de origen y determina una colocación mutua de la primera imagen de origen y la segunda imagen de origen como mínimo para las etapas de calcular una imagen promedio y calcular una imagen de varianza en base al desplazamiento. El desplazamiento puede utilizarse para desplazar la segunda imagen de origen para superponer de manera precisa las imágenes del objeto agarrado en las imágenes de origen primera y segunda. La imagen real del objeto agarrado puede eliminarse. El desplazamiento puede ser solo lógico y utilizarse solo como un valor o índice de desplazamiento en el cálculo de las imágenes promedio y de varianza.

35 Las realizaciones de la invención descritas anteriormente en el presente documento con respecto a la figura 3 pueden utilizarse en cualquier combinación entre sí. Varias de las realizaciones pueden combinarse entre sí para formar una realización adicional de la invención. Un procedimiento, un sistema, un aparato, un programa informático o un producto de programa informático al que se refiere la invención puede comprender como mínimo una de las realizaciones de la invención descritas anteriormente en el presente documento.

40 La figura 4 es un diagrama de flujo que ilustra un procedimiento para un procedimiento de reconocimiento de objetos basado en el movimiento de objetos en un sistema robótico en una realización de la invención. El movimiento del objeto utilizado se determina utilizando un movimiento conocido de un elemento de agarre o un brazo robótico que sujeta el objeto que se va a reconocer.

45 En la etapa 400, el objeto que se va a reconocer se agarra utilizando un elemento de agarre unido a un brazo robótico.

50 En la etapa 402, el objeto que se va a reconocer se mueve con, como mínimo, uno del elemento de agarre y el brazo robótico.

55 En la etapa 404 se capturan una serie de imágenes utilizando una cámara conectada, por ejemplo, al brazo robótico o al elemento de agarre. La cámara también puede ser estacionaria y visualizar el movimiento del elemento de agarre desde una distancia del brazo robótico. Las imágenes se capturan mientras el objeto objetivo está en movimiento. El objeto que se va a reconocer puede moverse en relación con un fondo que comprende una serie de otros objetos que pueden someterse a manipulación posterior por el brazo robótico o que pueden ignorarse en la clasificación. El movimiento del brazo robótico puede comprender un movimiento inicial hacia un posible destino para el objeto que se va a reconocer tal como varios compartimentos o estantes de ordenación. El fondo puede ser un escenario no estructurado. El número de imágenes puede ser, por ejemplo de 4, 5, 10 o un número natural arbitrario mayor de 1.

60 En la etapa 406, se registra el movimiento del elemento de agarre durante la captura de la serie de imágenes. Naturalmente, esta etapa se realiza en paralelo con la etapa 404. El movimiento se registra, por ejemplo, utilizando datos de sensor obtenidos del brazo robótico. Los datos de sensor pueden corresponder al cambio de posición del elemento de agarre en las coordenadas del mundo real. El cambio de posición del elemento de agarre puede

traducirse en una función de correlación de coordenadas del mundo real a coordenadas de imagen, que representan el movimiento del elemento de agarre dentro de las imágenes capturadas con la cámara, tal como la serie de las imágenes. La función de correlación puede construirse utilizando el aprendizaje de la máquina.

5 En la etapa 408, se determina como mínimo un primer vector de movimiento para el movimiento entre la serie de imágenes en base al movimiento registrado del elemento de agarre. El, como mínimo, un primer vector de movimiento puede representar el movimiento del elemento de agarre en coordenadas de imagen, por ejemplo, en píxeles entre cada imagen capturada subsiguiente.

10 En la etapa 410, como mínimo una imagen en la serie de imágenes se divide en una serie, por ejemplo, como mínimo de cuatro áreas. Las áreas pueden ser píxeles, bloques de píxeles o áreas de forma variable.

En la etapa 412, se determina como mínimo un segundo vector de movimiento para como mínimo un área en base a una comparación de datos de imagen en imágenes subsiguientes dentro de la serie de imágenes. La comparación puede comprender la búsqueda de las características visuales de un área de una primera imagen a partir de una segunda imagen posterior.

En la etapa 414, el, como mínimo, un segundo vector de movimiento se hace coincidir con el, como mínimo, un primer vector de movimiento. La comparación puede referirse a la dirección y la longitud, es decir, a la magnitud. Si la comparación revela que las direcciones y las longitudes de los vectores coinciden dentro de una tolerancia de error o que los vectores son vectores inversos dentro de una tolerancia de error, se selecciona como mínimo un área con un vector de movimiento correspondiente al, como mínimo, un primer vector de movimiento. La tolerancia de error comprende la tolerancia de error con respecto a los vectores de dirección y longitud.

25 En una realización de la invención, se elimina un área seleccionada del proceso adicional o se elige para el proceso adicional. Las áreas para el proceso adicional se someten a etapas de reconocimiento de objetos adicionales.

En una realización de la invención, las áreas seleccionadas se eliminan de la serie de imágenes. A continuación, pueden seguir las etapas de calcular una imagen promedio, calcular una imagen de varianza y la formación de la imagen de filtrado de la manera explicada con respecto a estas etapas en la figura 3, indicadas con los números de referencia 306, 308 y 310, respectivamente. La eliminación puede comprender el ajuste de los valores de píxel a un valor predefinido tal como cero o uno.

En una realización de la invención, como mínimo se utiliza un área seleccionada de una imagen entre la serie de imágenes para obtener directamente una imagen de resultado. La imagen de resultado se puede utilizar en la clasificación del objeto que se va a reconocer. Puede haber etapas de reconocimiento adicionales antes de la clasificación, tal como la eliminación de las características visuales del elemento de agarre de la imagen de resultado.

40 Las realizaciones de la invención descritas anteriormente en el presente documento con respecto a la figura 4 pueden utilizarse en cualquier combinación entre sí. Varias de las realizaciones pueden combinarse entre sí para formar una realización adicional de la invención.

Las realizaciones a modo de ejemplo de la invención se pueden incluir dentro de cualquier dispositivo adecuado, por ejemplo, incluyendo cualquier servidor, estación de trabajo, PC, ordenador portátil, PDA, aplicación de Internet, dispositivos de mano, teléfono móvil, dispositivo inalámbrico, otros dispositivos y similares adecuados, que puedan realizar los procesos de las realizaciones a modo de ejemplo, y que puedan comunicarse a través de uno o varios mecanismos de interfaz, incluyendo, por ejemplo, acceso a Internet, telecomunicaciones en cualquier forma adecuada (por ejemplo, voz, módem y similar), medios de comunicaciones inalámbricas, una o varias redes de comunicaciones inalámbricas, redes de comunicaciones móviles, redes de comunicaciones 3G, redes de comunicaciones 4G, red telefónica pública conmutada (PSTN), redes de datos en paquetes (PDN), Internet, intranet, una combinación de los mismos y similares.

Ha de entenderse que las realizaciones a modo de ejemplo son para fines a modo de ejemplo, ya que son posibles muchas variaciones del hardware específico utilizado para implementar las realizaciones a modo de ejemplo, tal como apreciarán los expertos en la materia de hardware. Por ejemplo, la funcionalidad de uno o varios de los componentes de las realizaciones a modo de ejemplo puede implementarse a través de uno o varios dispositivos de hardware.

60 Las realizaciones a modo de ejemplo pueden almacenar información en relación con diversos procesos descritos en el presente documento. Esta información puede almacenarse en una o varias memorias, tales como un disco duro, disco óptico, disco magnetoóptico, RAM y similares. Una o varias bases de datos pueden almacenar la información utilizada para implementar las realizaciones a modo de ejemplo de las presentes invenciones. Las bases de datos se pueden organizar utilizando estructuras de datos (por ejemplo, registros, tablas, matrices, campos, gráficos, árboles, listas y similares) incluidos en una o varias memorias o dispositivos de almacenamiento enumerados en el presente documento. Los procesos descritos con respecto a las realizaciones a modo de ejemplo pueden incluir estructuras

de datos apropiadas para almacenar los datos recogidos y/o generados por los procesos de los dispositivos y subsistemas de las realizaciones a modo de ejemplo en una o varias bases de datos.

5 La totalidad o una parte de las realizaciones a modo de ejemplo puede implementarse mediante la preparación de circuitos integrados específicos de la aplicación o mediante la interconexión de una red apropiada de circuitos de componentes convencionales, tal como apreciarán los expertos en la materia eléctrica.

10 Tal como se indicó anteriormente, los componentes de las realizaciones a modo de ejemplo pueden incluir memorias o un medio legible por ordenador según lo dado a conocer en las presentes invenciones y para mantener estructuras de datos, tablas, registros y/u otros datos descritos en el presente documento. Un medio legible por ordenador puede incluir cualquier medio adecuado que participe en proporcionar instrucciones a un procesador para su ejecución. Un medio de este tipo puede adoptar muchas formas incluyendo, pero sin limitarse a, medios no volátiles, medios volátiles, medios de transmisión y similares. Los medios no volátiles pueden incluir, por ejemplo, discos ópticos o magnéticos, discos magnetoópticos y similares. Los medios volátiles pueden incluir memorias dinámicas y similares. Los medios de transmisión pueden incluir cables coaxiales, hilo de cobre, fibra óptica y similares. Los medios de transmisión también pueden adoptar la forma de ondas acústicas, ondas ópticas, ondas electromagnéticas y similares, tal como las generadas durante las comunicaciones por radiofrecuencia (RF), comunicaciones de datos por infrarrojos (IR) y similares. Las formas comunes de medios legibles por ordenador pueden incluir, por ejemplo, un disquete, un disco flexible, disco duro, cinta magnética, cualquier otro tipo de medio magnético adecuado, un CD-ROM, CDRW, DVD, cualquier otro medio óptico adecuado, tarjetas perforadas, cinta de papel, hojas de marcas ópticas, cualquier otro medio físico adecuado con patrones de orificios u otros indicios reconocibles ópticamente, una RAM, una PROM, una EPROM, una FLASH-EPROM, cualquier otro chip o cartucho de memoria adecuado, una onda portadora o cualquier otro medio adecuado que pueda leer un ordenador.

25 Es obvio para un experto en la materia que con el avance de la tecnología, la idea básica de la invención puede implementarse de diversos modos. La invención y sus realizaciones no se limitan, por tanto, a los ejemplos descritos anteriormente; en cambio pueden variar dentro del alcance de las reivindicaciones.

REIVINDICACIONES

1. Procedimiento, que comprende:

5 agarrar un objeto con un elemento de agarre unido a un brazo robótico;
capturar (304) como mínimo dos imágenes de origen de un área que comprende el objeto con un sensor de imagen
mientras el objeto se está moviendo con el brazo robótico;
calcular (306) una imagen promedio de las, como mínimo, dos imágenes de origen;
caracterizado por
10 calcular (308) una imagen de varianza de las, como mínimo, dos imágenes de origen;
formar (310) una imagen de filtrado a partir de la imagen de varianza; y
obtener (312) una imagen de resultado enmascarando la imagen promedio utilizando la imagen de filtrado como una
máscara de bits.

15 2. Procedimiento, según la reivindicación 1, comprendiendo además el procedimiento:

detectar (300, 302) un agarre satisfactorio del objeto en un aparato conectado al elemento de agarre y el brazo
robótico;
en respuesta a la detección, indicar al sensor de imagen que capture las, como mínimo, dos imágenes de origen
20 mediante un aparato; y
obtener (304) las, como mínimo, dos imágenes de origen capturadas para el aparato.

3. Procedimiento, según la reivindicación 1, comprendiendo además el procedimiento:

25 reconocer una imagen del elemento de agarre en las, como mínimo, dos imágenes de origen;
calcular (308) como mínimo un desplazamiento para una imagen del elemento de agarre entre una primera imagen
de origen y una segunda imagen de origen; y
determinar una colocación mutua de la primera imagen de origen y la segunda imagen de origen como mínimo para
las etapas de calcular una imagen promedio y calcular una imagen de varianza en base al desplazamiento.

30 4. Procedimiento, según la reivindicación 1, comprendiendo además el procedimiento:
determinar una colocación mutua de las, como mínimo, dos imágenes de origen, cuyo movimiento corresponde a un
movimiento registrado del elemento de agarre, como mínimo para las etapas de calcular una imagen promedio y
calcular una imagen de varianza en base al desplazamiento.

35 5. Procedimiento, según la reivindicación 1, comprendiendo además el procedimiento:
eliminar como mínimo un área de imagen de la imagen de resultado que comprende textura visual de como mínimo
uno del elemento de agarre y el brazo robótico.

40 6. Procedimiento, según la reivindicación 1, comprendiendo además el procedimiento:

clasificar el objeto en la imagen de resultado en base a como mínimo una característica visual en la imagen de
resultado; y
mover el objeto a una ubicación correspondiente a la clasificación.

45 7. Procedimiento, según la reivindicación 1, en el que la etapa de calcular una imagen de varianza comprende:

calcular una varianza de cada píxel respectivo en las, como mínimo, dos imágenes de origen; y
ajustar la varianza calculada como el valor del píxel respectivo en la imagen de varianza.

50 8. Procedimiento, según la reivindicación 1, en el que la etapa de formar una imagen de filtrado comprende:
ajustar cada píxel respectivo a 1 en la imagen de filtrado para la que el píxel respectivo en la imagen de varianza
tiene un valor por debajo de un umbral predefinido.

55 9. Procedimiento, según la reivindicación 1, en el que la etapa de obtener una imagen de resultado comprende:
seleccionar cada píxel respectivo para la imagen de resultado a partir de la imagen promedio solo si el valor del píxel
respectivo en la imagen de filtrado es 1.

60 10. Procedimiento, según la reivindicación 1, en el que las, como mínimo, dos imágenes se toman contra un fondo
que comprende, además de los objetos de interés, objetos de características desconocidas.

11. Procedimiento, según la reivindicación 1, en el que el sensor de imagen es como mínimo uno de una cámara,
una cámara de infrarrojos y un escáner láser.

65 12. Procedimiento, según la reivindicación 1, en el que las etapas de calcular una imagen promedio, calcular una
imagen de varianza y formar una imagen de filtrado se realizan por separado para cada canal de color de píxel.

13. Procedimiento, según la reivindicación 1, en el que el sensor de imagen está unido al elemento de agarre y está configurado para moverse junto con el elemento de agarre.

5 14. Aparato, que comprende:

medios (120) para controlar un elemento de agarre (112) y un brazo robótico (116) para agarrar un objeto;
medios (152) para obtener como mínimo dos imágenes de origen de un área que comprende el objeto con un sensor
de imagen mientras el objeto se está moviendo con el brazo robótico;

10 medios (142) para calcular una imagen promedio de las, como mínimo, dos imágenes de origen;

caracterizado por

medios (142) para calcular una imagen de varianza de las, como mínimo, dos imágenes de origen;

medios (142, 154) para formar una imagen de filtrado a partir de la imagen de varianza; y

15 medios (142, 154) para obtener una imagen de resultado enmascarando la imagen promedio utilizando la imagen de filtrado como una máscara de bits.

15. Programa informático implementado en un medio legible por ordenador, comprendiendo el programa informático código para controlar un procesador para ejecutar un procedimiento que comprende:

20 controlar un elemento de agarre y un brazo robótico para agarrar un objeto;

obtener como mínimo dos imágenes de origen de un área que comprende el objeto con un sensor de imagen
mientras el objeto se está moviendo con el brazo robótico;

calcular una imagen promedio de las, como mínimo, dos imágenes de origen;

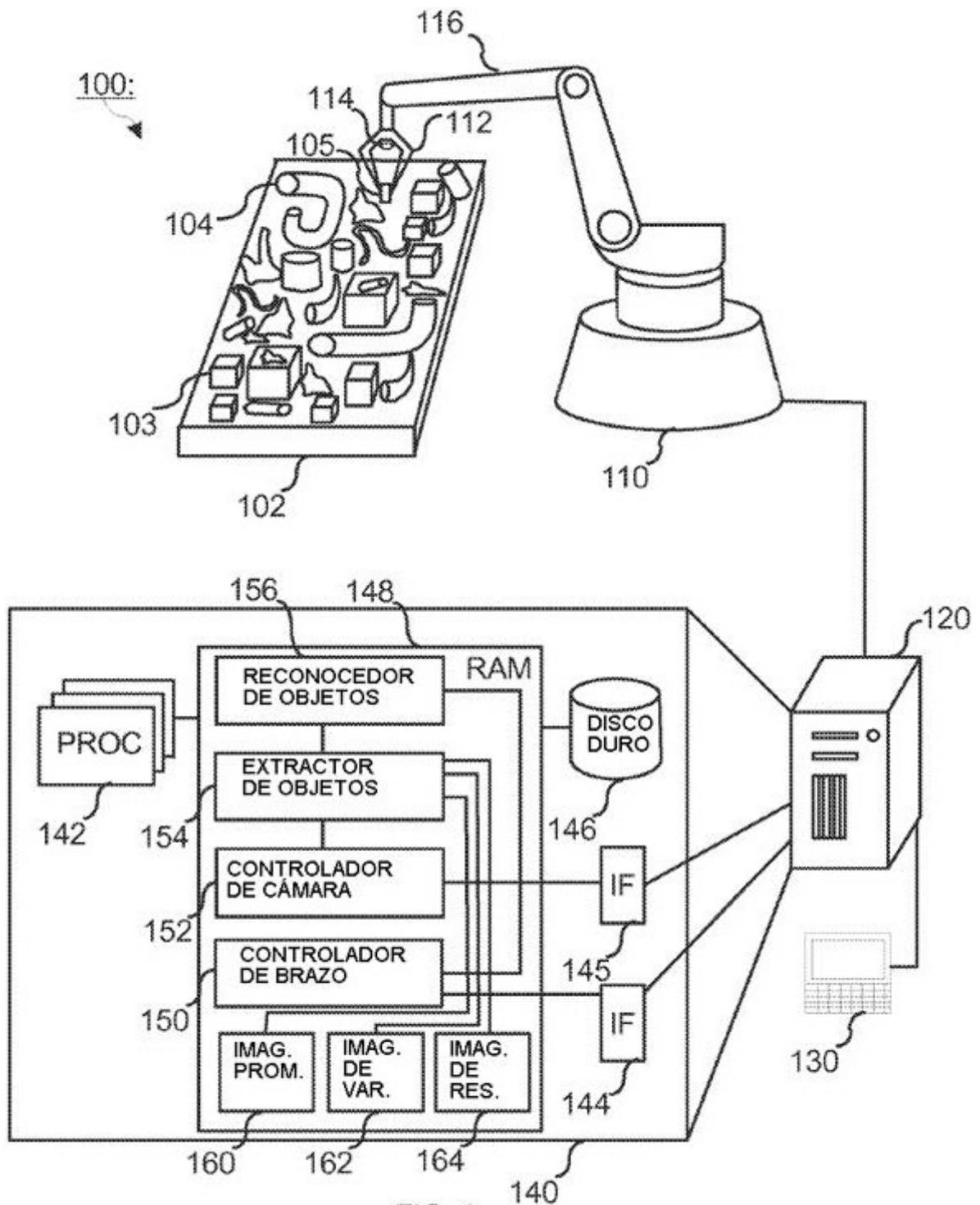
caracterizado por

25 calcular una imagen de varianza de las, como mínimo, dos imágenes de origen;

formar una imagen de filtrado a partir de la imagen de varianza; y

obtener una imagen de resultado enmascarando la imagen promedio utilizando la imagen de filtrado como una
máscara de bits.

30 16. Programa informático, según la reivindicación 15, en el que dicho medio legible por ordenador es una tarjeta de memoria extraíble, una memoria holográfica, un disco magnético o un disco óptico.



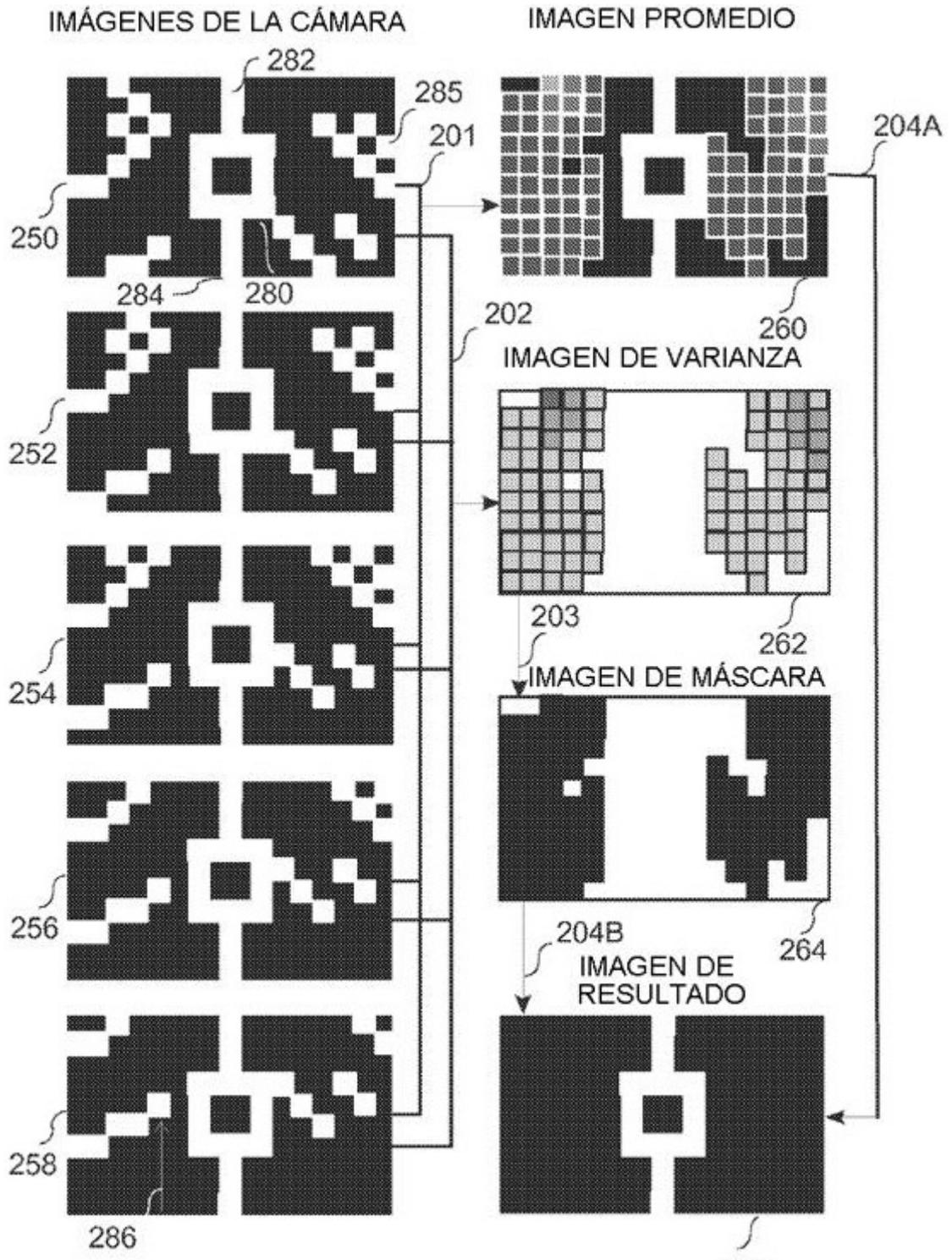


FIG. 2A

266

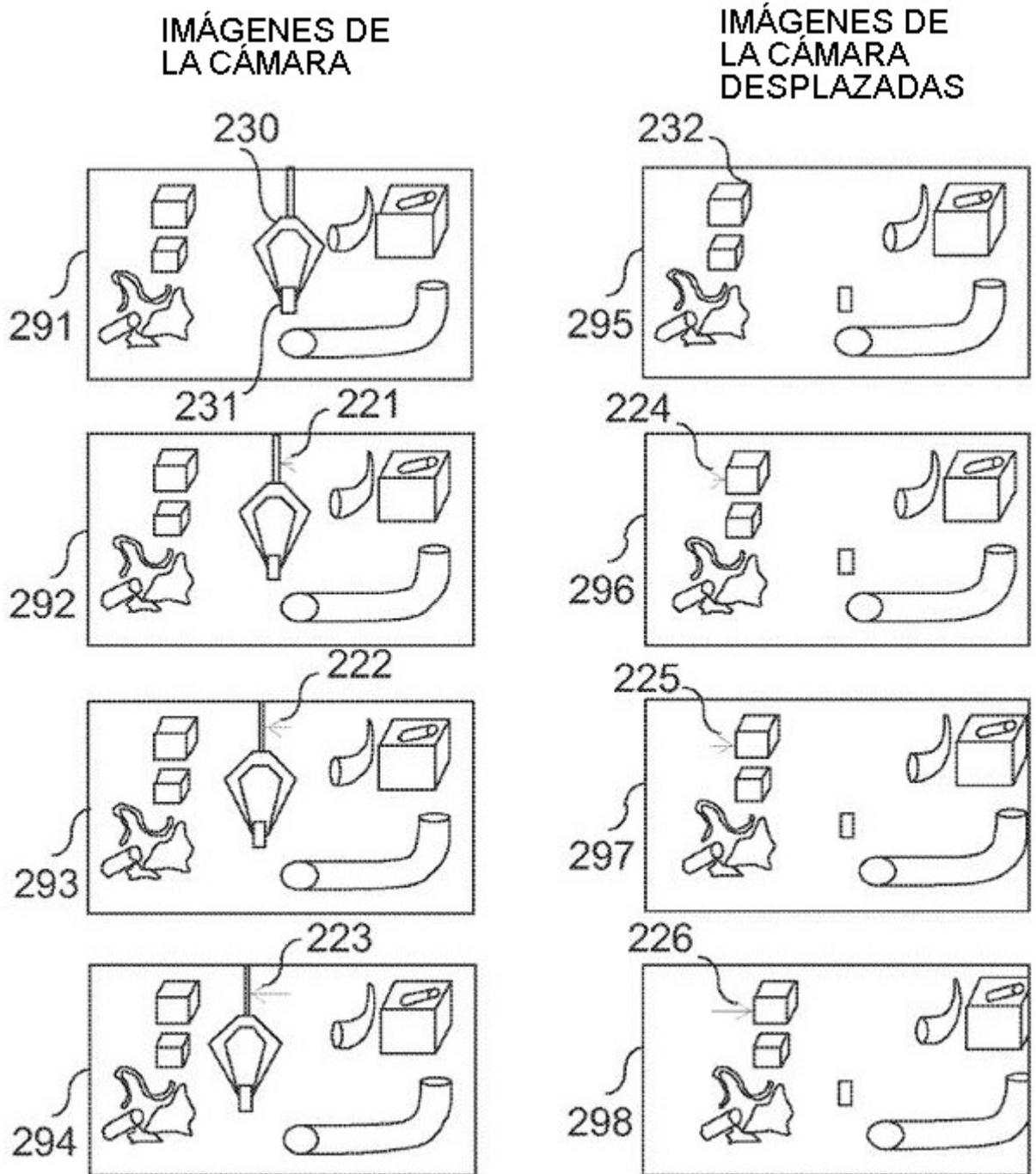


FIG. 2B



FIG. 3

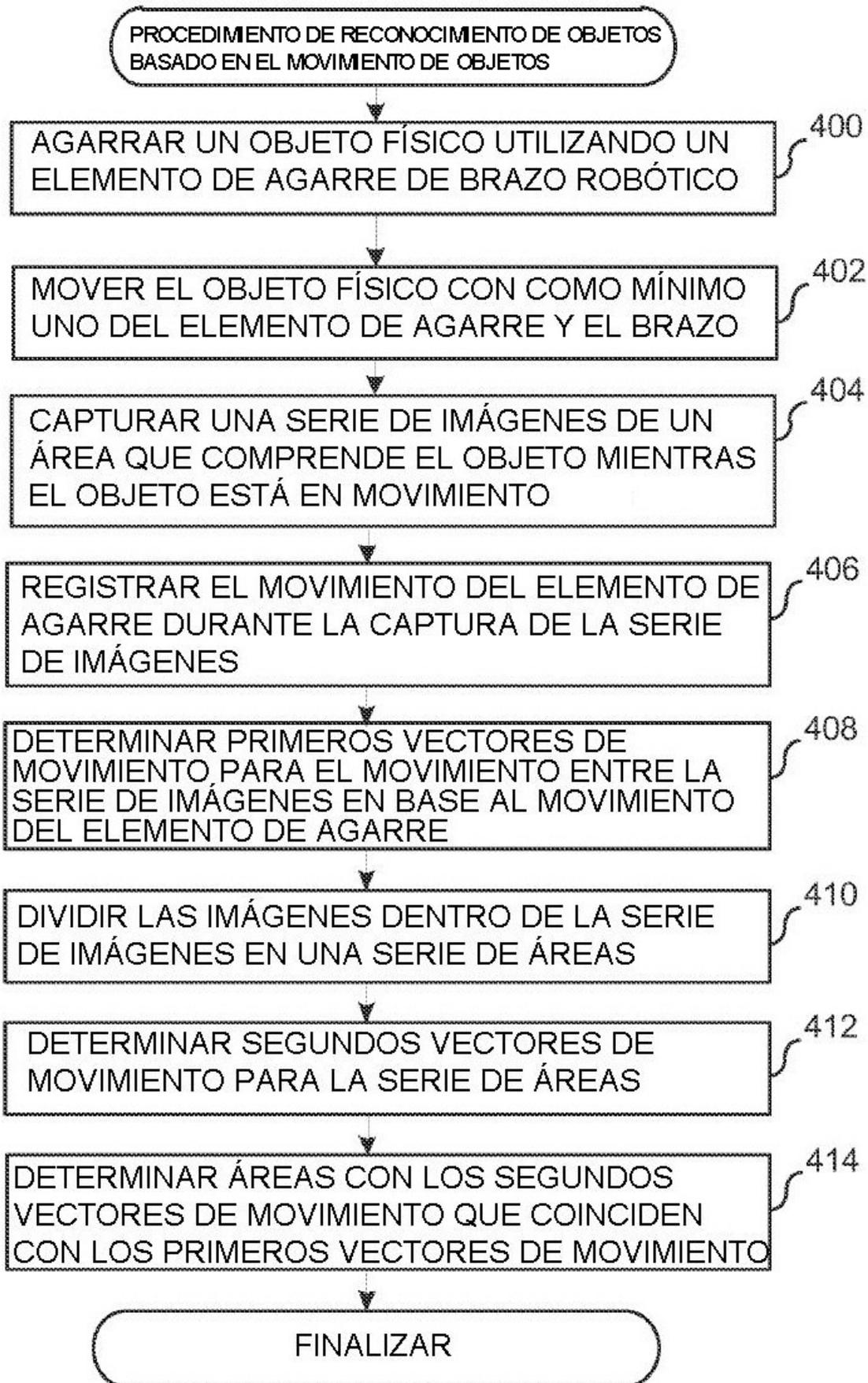


FIG. 4