Automatic precision robot assembly system with microscopic vision and force sensor

Yanqin Ma[1,2,3] ⬤, **Kai Du**[3], **Dongfeng Zhou**[4], **Juan Zhang**[3], **Xilong Liu**[1,2] and **De Xu**[1,2]

Abstract

An automatic precision robot assembly system is established. The robot assembly system mainly consists of an industrial robot, three cameras, a micro force sensor, and a specific gripper. The industrial robot is a six-axis serial manipulator, which is used to conduct grasping and assembly subtasks. Two microscopic cameras are fixed on two high accuracy translational platforms to provide visual information in aligning stage for assembly. While one conventional camera is installed on the robotic end effector to guide the gripper to grasp component. The micro force sensor is installed on the robotic end effector to perceive the contacted forces in inserting stage. According to the characteristics of components, an adsorptive gripper is designed to pick up components. In addition, a three-stage "aligning–approaching–grasping" control strategy for grasping subtask and a two-stage "aligning–inserting" control strategy for assembly subtask are proposed. Position offset compensation is computed and introduced into aligning stage for assembly to make the grasped component in the microscopic cameras' small field of view. Finally, based on the established robot assembly system and the proposed control strategies, the assembly tasks including grasping and assembly are carried out automatically. With 30 grasping experiments, the success rate is 100%. Besides, the position and orientation alignment errors of pose alignment for assembly are less than 20 μm and 0.1°.

Keywords

Grasping, aligning, visual servoing, precision assembly, robot system

Introduction

In recent years, with the wide application of microelectro-mechanism system in the fields of medicine, aerospace, and precision electronic engineering, precision assembly technology has received more and more attention.[1,2] Precision assembly is to package millimeter-size or less components with micrometer-level precision requirements.[3] Traditional precision assembly is conducted by human, and the manual assembly has the inherent limitations of low efficiency and precision. Besides, it causes high production costs. Automatic precision assembly as a promising assembly

[1] Institute of Automation, Chinese Academy of Sciences, Beijing, China
[2] University of Chinese Academy of Sciences, Beijing, China
[3] Research Center of Laser Fusion, China Academy of Engineering Physics, Sichuan, China
[4] China University of Mining and Technology, Beijing, China

Corresponding author:
De Xu, Research Center of Precision Sensing and Control, Institute of Automation, Chinese Academy of Sciences, Beijing 100190, China.
Email: de.xu@ia.ac.cn

technology has been investigated extensively, and numerous of precision assembly systems are established for different applications.[4,5] For example, Wason et al.[6] designed a vision-based micro-assembly system to assemble multiple probes. Xing et al.[7] presented a micro-assembly system with six manipulators to assemble irregular shaped components. Shen et al.[8] designed an automatic high precision assembly system to assemble microparts. In these assembly systems, a large number of assembly methods are proposed to achieve automatic precision assembly.

In general, assembly can be further classified into parallel assembly and serial assembly.[9] Parallel assembly divides the assembly task into several subtasks, and these subtasks are conducted simultaneously. Compared to parallel assembly, serial assembly has lower efficiency. But serial assembly is more suited to complex assembly manipulation for its ability of assembling varying types components with high flexibility.[10] Thus, serial assembly are commonly used in precision assembly. Grasping is one of the most basic assembly tasks.[8,9,11] Traditional grasping control strategies without vision sensors are prone to leading to grasping failure when uncertainties occur in pose estimation.[12] Visual servoing is an effective method to improve success rate of grasping for the promising ability to provide noncontact online information.[9,11] Series of visual servoing control algorithms have been employed in grasping manipulation, which dramatically increases the flexibility and reliability.[11,13–16] For example, the micro-grasping and micro-joining tasks are performed with visual servoing control.[9] With a simple 2-D gripper, the stable grasping in 3-D space is achieved with 2-D visual information.[14] Combining visual features and a camera on robotic end effector, a visual servoing grasping planning method in dynamic environment is proposed.[15] With an active zoom microscopic camera and nonlinear projective principle, the multiple scale visual information is used for robot micro-manipulation and micro-assembly.[16] According to the relation between hand and eye, the visual servoing-based grasping methods can be divided into two categories, that is, eye-to-hand grasping[13,14] and eye-in-hand grasping.[15,16] In eye-to-hand mode, the working space for grasping is limited. Besides, in order to control the end effector to approach object, it also needs to estimate the relative pose between the end effector and the object in real time.[13,14] With eye-in-hand grasping, the camera can move with the end effector, which expands the working space for grasping.[15,16]

The inserting assembly is one of classic assembly modes.[10,17–20] According to fit type between components, the inserting assembly can be categorized into clearance assembly, transition assembly, and interference assembly.[7] In precision assembly, the interference assembly is widely used.[17] In general, the interference inserting assembly can be divided into two stages, that is, aligning stage and inserting stage.[4,7] The aligning is usually completed based on visual features,[10] while the force-based control strategy is employed in inserting stage due to the visual features are blocked in inserting

process.[18–22] Liu et al.[19] proposed a high precision assembly control strategy. Firstly, the orientation alignment and position alignment are carried out separately. Then the inserting assembly is conducted based on force feedback. Xing et al.[4] proposed a hybrid control structure for precision contacted assembly, which includes a vision-based alignment controller and a force-based inserting controller. Liu et al.[21] presented a high efficient inserting assembly method, which models the inserting process as a stochastic state transition process and the uncertainties in inserting process are described by Gaussian function. Considering multiple points contacted assembly, Wiemer and Schimmels[22] designed a direct admittance selection method for force-guided assembly. Because most of the above precision assembly systems are equipped with multiple independent manipulators and the degree of freedom of each manipulator is limited, the orientation and position alignments are carried out sequentially.[4,7,18–21] Obviously, the efficiency of the decoupled pose alignment is low. In the works above, the grasping control is not investigated, and the components to be assembled are put on the manipulators by human.

With the characteristics of miniaturization, flexibility, high movement accuracy, and wide working space of industrial robots, the industrial robots are widely employed to automatically assemble objects.[23] For example, Koveos et al.[24] presented a task-based variable impedance method for assembly, which achieves assembly via robot-aided operation in a semistructured environment. Fang et al.[25] designed a robot assembly system for small components assembly, in which a dual-arm robot is used to improve the flexibility and collaboration in automatic assembly. To solve the problems among the precision, mobility, and global view in vision-based grasping, Muis and Ohnishi[26] used two robots achieve grasp positioning. The industrial robot assembly systems are mainly employed to assemble mesoscale or macroscale components. Because of high precision requirements in precision assembly,[27] it still be a challenge to achieve precision assembly with industrial robots.

Considering the shortages of existing assembly systems and assembly methods, an automatic robot precision assembly system used for assembling small components with high efficiency is established. The robot assembly system mainly consists of an industrial robot, a gripper, three cameras, a micro force sensor, and a host computer. The industrial robot is used to conduct the whole assembly manipulation. The gripper is a specific vacuum adsorptive device, which is used to grasp small components. One camera with large field of view is installed on the robotic end effector to guide robot to locate components in large working space. Two microscopic cameras with small field of view and high resolution are used to measure the pose errors of components in aligning stage for assembly. The micro force sensor provides force information in inserting stage. Based on the designed robot assembly system, the classic assembly manipulation: grasping and assembly are conducted. An "aligning–approaching–grasping" control strategy is designed for grasping

subtask. The proposed grasping method provides an efficient solution for eye-in-hand grasping with monocular vision, and it is suitable to be used in the engineering applications for its high success rate and low hardware cost. Besides, the Image Jacobian matrix needs to be calibrated only once, which is more flexible compared to existing grasping methods. An "aligning–inserting" control strategy for assembly subtask with microscopic vision and force perception is proposed. The assembly experiments of two millimeter-scale components are carried out to verify the practicability and efficiency of robot assembly system. The main contributions of this article are:

A novel automatic robot assembly system is established for precision assembly, which can automatically complete classic assembly tasks with high efficiency.

(i) A three-stage "aligning–approaching–grasping" control method for grasping is proposed, which improves the grasping success rate.
(ii) A two-stage "aligning–inserting" assembly control method is designed, which improves the efficiency of precision assembly.

The rest of this article is arranged as follows. The designed precision robot assembly system including the hardware system and the software system is described in the second section. The automatic precision assembly control, which includes the assembly control strategies and process, the grasping and assembly controllers, is presented in the third section. In the fourth section, the precision assembly experiments are conducted on the established robot assembly system. In the final section, the conclusions are presented.

Assembly task

The task of this article is to assemble two small cylindrical components that are shown in Figure 1. The whole assembly task consists of two subtask: grasping and assembly. The three-stage "aligning–approaching–grasping" grasping method is focused on aligning, approaching, and grasping components. And we take grasping component A as example to explain the grasping method. Of course, the component B can be picked up with the proposed grasping control method and put on the assembly platform automatically. It is not necessary to describe or discuss the grasping process for the components A and B repeatedly. The component B is assumed to be put on the assembly platform in advance in order to make this article more compact. The assembly subtask includes two substages, that is, aligning stage and inserting stage. The goals of aligning stage for assembly are to align the orientation of component A to component B and make the relative position between component A and component B reaches a desired status. While the goal of inserting stage is to complete interference inserting assembly without damaging the components.

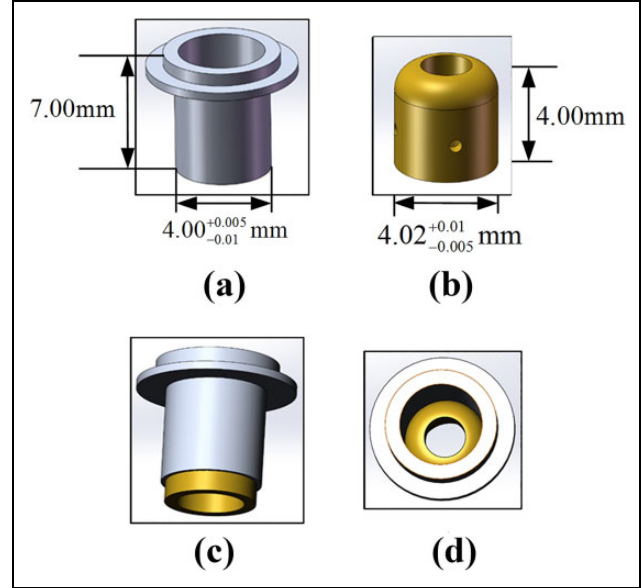The main difficulties of this assembly task can be summarized as follows.

Figure 1. Components to be assembled. (a) Component A. (b) Component B. (c) Components in assembly process. (d) Assembled component.

(i) Design of the hardware system: To ensure the high success rate of grasping subtask and improve the efficiency of assembly subtask, the configuration of the robot assembly system is one of the key points. The hardware system configuration consists of the model selection and the installation of hardware devices. In particular, when designing the gripper used for grasping components, it needs to ensure the reliable grasping and avoid blocking component's image features.

(ii) Control strategy for grasping subtask: According to the configuration of robot assembly system, the central axis of the gripper is approximately parallel to the optical axis of camera installed on robotic end effector. The camera cannot directly observe the end of the gripper, which makes the grasping is an open-loop control based on vision guidance. In order to improve the success rate of grasping, it is necessary to design a reasonable and applicable grasping control strategy.

(iii) Control strategy for assembly subtask: To improve the efficiency of aligning stage for assembly, the orientation and position alignments are expected to be conducted simultaneously. However, how to keep components in the microscopic cameras' small field of view in aligning stage is still a problem. In addition, due to the two components are thin-walled parts and the inserting assembly is interference assembly, it is necessary to consider how to design inserting control strategy to protect thin-walled components from damage.
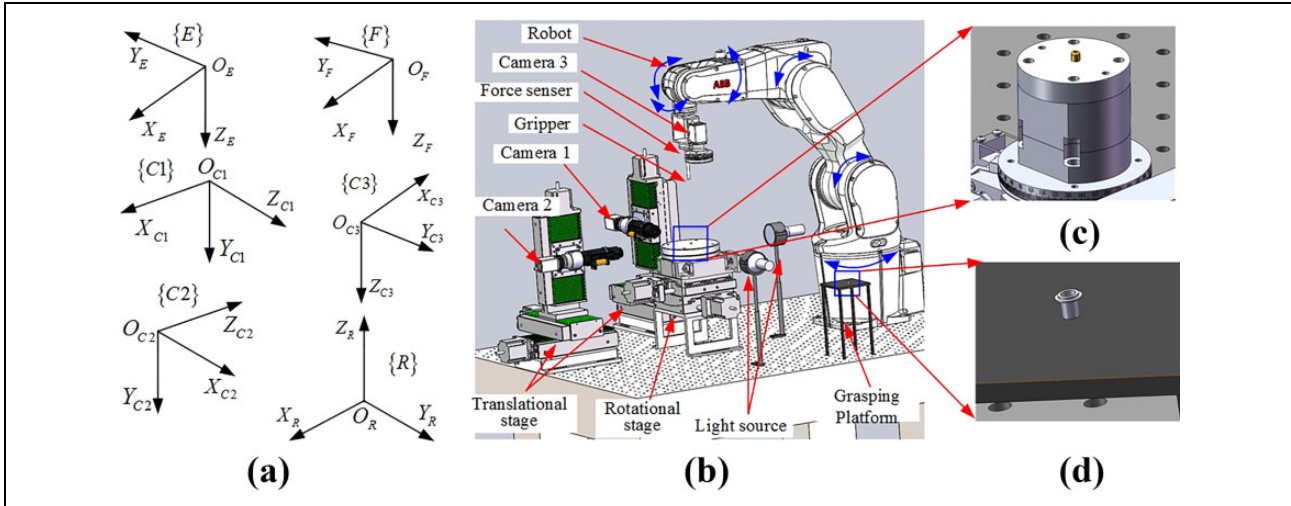
Figure 2. The configuration of the robot assembly system. (a) The established frames. (b) The robot assembly system. (c) Component B on the assembly platform. (d) Component A on the grasping platform.

Precision robot assembly system

Hardware system

A robot assembly system is established as shown in Figure 2. The hardware system mainly consists of a robot system, a vision system, a force sensor, a specific gripper, two translational platforms, a rotational platform, and a host computer. The robot system includes an industrial robot and a robot controller. The vision system includes three cameras: camera 1, camera 2, and camera 3. Camera 1 and camera 2 are microscopic vision. While camera 3 is conventional vision. The gripper is adsorptive cylindrical structure.

Robot system. The ABB IRB1200 industrial robot is a six-axis serial manipulator. The resolution of each axis is approximately $0.01°$. The reach and position accuracy of the robot are 703 mm and 0.02 mm, which fulfills the requirements of grasping in large working space and high precision requirements of assembly. In addition, the robot system is equipped with an IRC5 controller. With the IRC5 controller, we can control the robot to conduct assembly task.

Vision system. The resolution of the three cameras are 2448 × 2050 pixels. The microscopic camera 1 and camera 2 are fixed on high accuracy translational platforms. The optical axes of two microscopic cameras are approximately perpendicular. The two microscopic cameras are equipped with ring light source, coaxial light source, and collimated light emitting diode (LED) backlight source. The camera 3 is mounted on robotic end effector. It is equipped with a ring light source. The ring light source is installed in front of the camera's lens to make camera observe the components' features clearly. The coaxial light source is installed
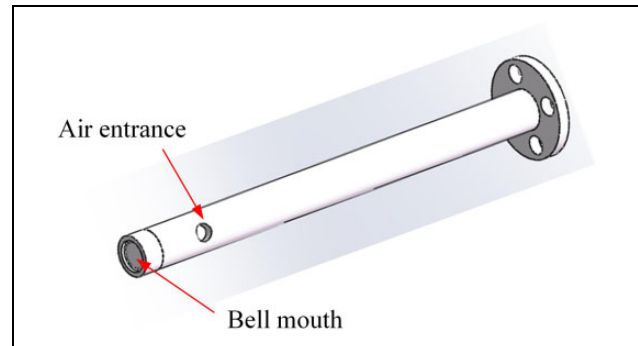

Figure 3. The specific gripper.

inside the camera's lens to improve the imaging of the planar reflective components. The backlight source is installed on the extension line between the camera's optical center and the components. It can make the contour features of components clearly and easily be extracted. The microscopic cameras are equipped with NAVITAR 6000 motorized focusing lens, while the conventional camera 3 is equipped with a Computer M2518-MPV lens whose focal length is 25 mm.

Force sensor. The fit style between component A and component B is interference fit. To measure the contacted forces in inserting stage, the ATI six-axis micro force and torque sensor Nano43 is employed. The force sensor can measure force and torque simultaneously. The force sensor is mounted between the robotic end effector and the gripper. The axes X_F, Y_F, and Z_F of force sensor are approximately parallel to the axes X_E, Y_E, and Z_E of robotic end effector. The range of force sensor is within $[-18, 18]$ N via configuring the calibration file, and the resolution of force sensor is 4 mN. In this article, only the contacted forces are used in inserting assembly.
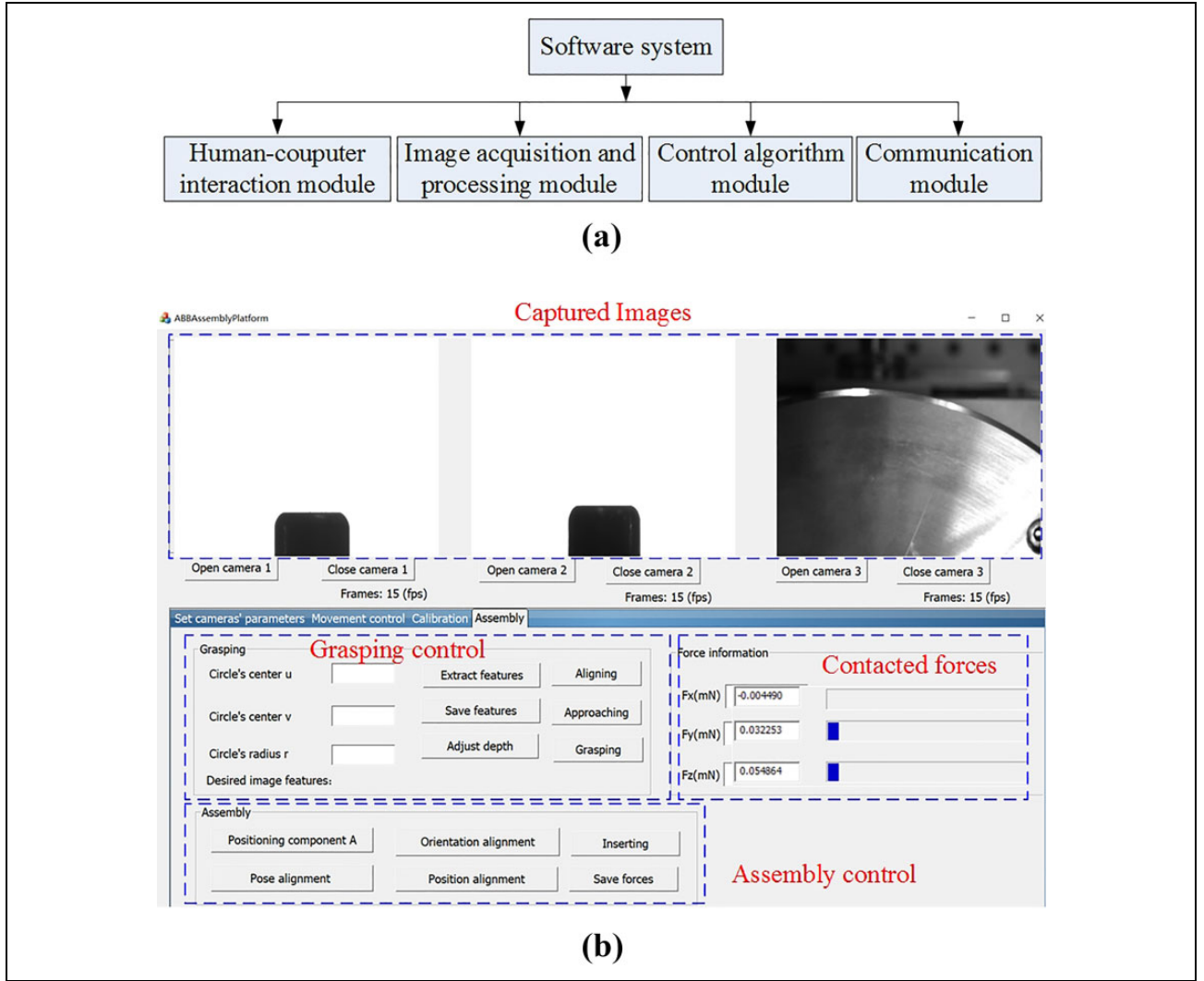
Figure 4. Software system of robot assembly system. (a) The modules of software system. (b) The operation interface of assembly system.

Absorptive gripper. The designed adsorptive gripper is shown in Figure 3. The gripper is internal hollow cylindrical structure. The air entrance of the gripper is connected with a suction device, and the end of the gripper has lots of holes. When the gripper approaches component A, component A can be picked up due to the adsorption force. The designed gripper can grasp component without blocking the visual features of component A. In addition, the end of gripper is designed into a bell mouth. The diameter of the bell mouth is larger than the outer diameter of component A by 0.01 mm, which can improve the success rate of grasping.

Various frames are established for convenience, which are shown in Figure 2. The robot base frame $\{R\}$ is established on the robot base. The axis X_R points to the front of the robot when the robot is in the default initial pose. The axis Z_R points to the second joint from the base. The axis Y_R is set according to the right-hand rule. The robot end frame $\{E\}$ is established on the robotic end effector with origin in its default tool center. The frame $\{E\}$ moves as the robotic end effector is

adjusted. The force frame $\{F\}$ is established on the force sensor. The axes X_F, Y_F, and Z_F are approximately parallel to the axes X_E, Y_E, and Z_E. The frames $\{C1\}$, $\{C2\}$, and $\{C3\}$ are established on the camera 1, camera 2, and camera 3's optical centers. The axes Z_C are parallel to their optical axes and point from origins to scenes. The axes X_C and Y_C are parallel to axes U and V of their images.

Software system

As shown in Figure 4(a), the software system mainly consists of four modules: human–computer interaction module, communication module, control algorithm module, and image acquisition and processing module. The human–computer interaction module includes the input and output devices and the data display interface. The operation interface of the robot assembly system is shown in Figure 4(b). The operator can control the assembly task and obtain the status of assembly system through the human–computer

interaction module. The communication module includes the communication programs with cameras,[28] the communication programs with translational and rotational platforms, the communication programs with force sensor, and the communication programs with robot. The control algorithm module mainly includes assembly control programs, and it is the core of assembly software system. The image acquisition and processing module mainly includes image acquisition and image processing. The basic image processing programs includes image filter, edge detection, and image feature extraction.

In addition, the assembly software in the computer is written with C++ language based on Microsoft Foundation Classes (MFC) library. While the programming language of ABB robot is RAPID. The RAPID programming language is designed for ABB robots and can be used in all ABB robotic products. The socket communication is employed between the host computer and the robot.

Automatic assembly control

Assembly control strategies and process

In this article, we mainly focus on the grasping and assembly subtasks.

Grasping subtask. In the grasping subtask, the "aligning–approaching–grasping" control strategy is proposed. In the aligning stage for grasping, the image-based visual servoing is employed to make the image features of component A coincide with the desired image features. In the approaching stage, the robotic end effector moves $-P$ to make the gripper approach component A. In the grasping stage, the gripper picks up component A based on absorption force.

The image point feature is sensitive to the translational movement perpendicular to the optical axis of camera, and the image area feature is sensitive to the translational movement parallel to the optical axis of camera.[28] In the aligning stage of grasping subtask, the center point feature and contour area feature of component A are manually selected as the image feature for image-based visual servoing. The desired and current image features of component A are denoted as (u_d, v_d, s_d) and (u, v, s). The desired image feature and position change for approaching are obtained as follows. Firstly, the robotic end effector moves to make the gripper can grasp component A and the pose of robotic end effector is recorded. Secondly, the robotic end effector moves to make component A within the clear field of view of camera 3. The robotic end effector's position change P for approaching is recorded, and the center point feature and contour area feature are saved as the desired image features (u_d, v_d, s_d).

The automatic grasping process is shown in Figure 5. In each control cycle, the camera 3 captures the image of component A, and component A's edge points are extracted based on the grayscale of image automatically. Then the edge of component A is determined via Random Sample
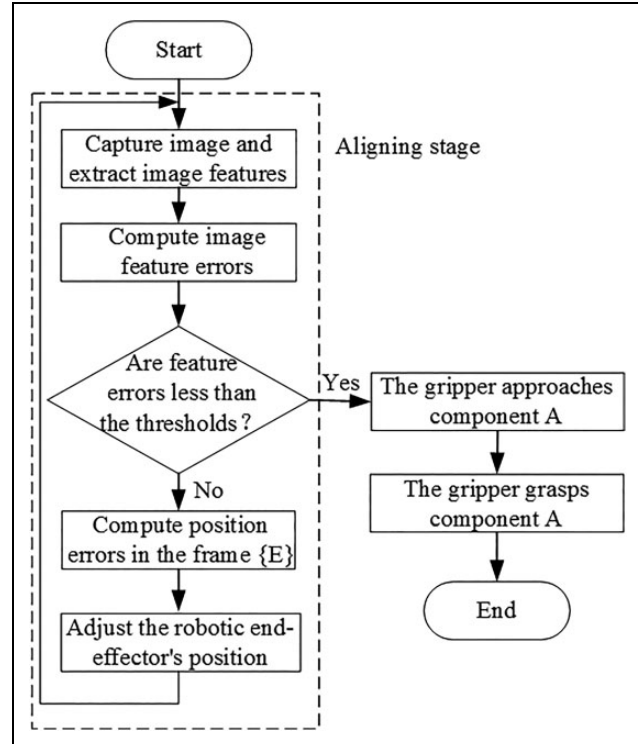

Figure 5. The grasping process.

Consensus (RANSAC) method. The center (u, v) and radius r of component A are obtained. The current image features $(u, v, s = \pi r^2)$ are computed. Then the image feature errors are calculated by comparing with the desired image features (u_d, v_d, s_d). If the image feature errors are larger than the designed thresholds, the corresponding position errors in the frame $\{E\}$ are computed based on the image Jacobian matrix, and the robotic end effector moves according to the designed grasping controller (1). Until the image feature errors are less than the thresholds, the aligning stage for grasping finishes. The robotic end effector moves $-P$ and the gripper grasps component A.

Assembly subtask. In the assembly subtask, the "aligning–inserting" control strategy is employed. In the aligning stage for assembly, the image-based visual servoing is used. And in the inserting stage, the force-based feedback control is used.

In aligning stage for assembly, the point feature and line feature are selected as the image features for visual servoing control. The point feature is used for position alignment, while the line feature is used for orientation alignment. The center of component A's lower edge line in the images of microscopic cameras is selected as the point feature of component A. The center of component B's upper edge line in the images of microscopic cameras is selected as the point feature of component B. The component A's central axis and component B's central axis are selected as the line features. And the angle of the central axis is used to represent the line feature. In the visual
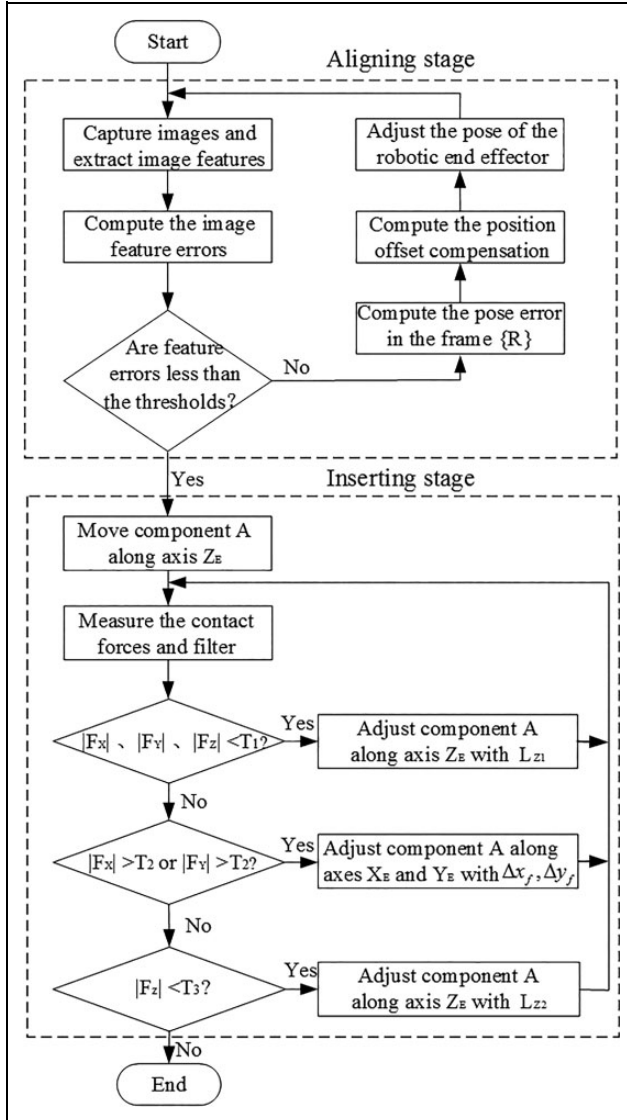
Figure 6. The assembly process.

inserting stage starts automatically. During inserting process, the force-based feedback control strategy (15) is used. If the forces $F_X, F_Y,$ and F_Z are less than T_1, it indicates that the two components do not contact. Then component A is moved along axis Z_E to conduct inserting. When the contacted force F_X or F_Y is larger than T_2, it indicates that component A needs to be adjusted along axis X_E and Y_E. Then component A is moved along axis Z_E until the contacted force F_Z is larger than T_3, the inserting stage finishes.

Controller design for grasping

In the aligning stage for grasping, the image-based visual servoing is used, and the incremental proportional-integral (PI) controller is designed as follows

$$
\begin{bmatrix} \Delta x_{Ek} \\ \Delta y_{Ek} \\ \Delta z_{Ek} \end{bmatrix} = K_{1i} J_h^{-1} \begin{bmatrix} \Delta u_k \\ \Delta v_k \\ \Delta s_k \end{bmatrix} \tag{1}
$$

$$
+ K_{1p} J_h^{-1} \left(\begin{bmatrix} \Delta u_k \\ \Delta v_k \\ \Delta s_k \end{bmatrix} - \begin{bmatrix} \Delta u_{k-1} \\ \Delta v_{k-1} \\ \Delta s_{k-1} \end{bmatrix} \right)
$$

where lower mark k and $k-1$ represent sampling time. K_{1p} and K_{1i} are proportional and integral coefficients of incremental PI controller. $\Delta u_k = u_d - u_k$, $\Delta v_k = v_d - v_k$, and $\Delta s_k = s_d - s_k$ are image feature errors. Δx_{Ek}, Δy_{Ek}, and Δz_{Ek} are position adjustments of robotic end effector along axes X_E, Y_E, and Z_E in sampling time k. J_h is image Jacobian matrix, which represents the relation between robotic end effector translational movement in the frame {E} and change of image feature in image captured by camera 3.

Controller design for assembly

Position offset compensation. In aligning stage for assembly, the orientation adjustment of robotic end effector with respect to frame {R} will lead to the position offset of component A. The position offset may cause component A moving out of the microscopic cameras' small field of view. Based on the differential transformation between frame {R} and frame {E}, the position offset of component A resulting from orientation adjustment is computed.

Given the pose of robotic end effector with respect to frame {R} as T with $T = \begin{bmatrix} R & p \\ 0 & 1 \end{bmatrix}$, $R = [n, o, a]$ and p represent the orientation matrix and position vector. A differential orientation adjustment of robotic end effector in the frame {R} will lead to robotic end effector's differential translation and rotation with respect to frame {E}. As is shown in Figure 7, suppose the orientation adjustment in the frame {R} is $\delta = [\delta_x, \delta_y, \delta_z]^T$, the pose of robotic end effector in the frame {R} after orientation adjustment yields

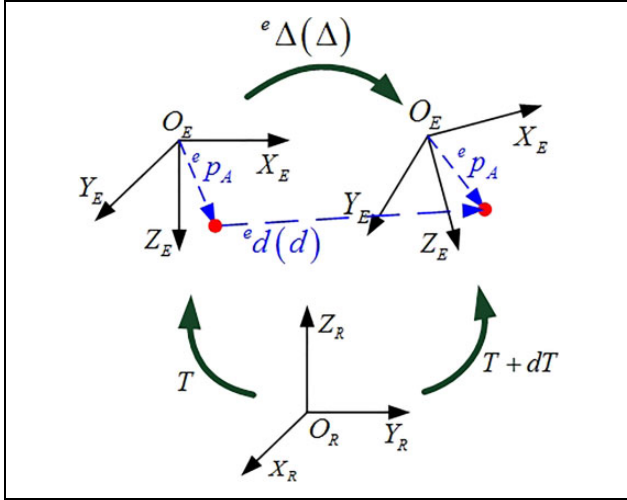servoing control, the image features are automatically extracted and tracked by designed region of interest.

The automatic assembly process is shown in Figure 6. In each control cycle, the microscopic cameras capture the images of component A and B. The edge points of two components are extracted based on the grayscale of images. Then the two side edge lines and lower (upper) edge line of component A (B) are determined via RANSAC method. The average line of the two side edge lines is set as component A (B)'s central axis. The point features and line features of two components are extracted and feature errors are calculated. If the feature errors are larger than the designed thresholds, the corresponding pose errors in the frame {R} are computed with image Jacobian matrix. The position offset compensation of component A due to robotic end effector's orientation adjustment is computed from (11) and (14). The pose of the robotic end effector is adjusted. Until the feature errors are less than the thresholds, the aligning stage for assembly finishes. The

Figure 7. The differential transformation for position offset compensation.

$$T + \mathrm{d}T = \mathrm{Rot}(x, \delta_x)\mathrm{Rot}(y, \delta_y)\mathrm{Rot}(z, \delta_z)T \qquad (2)$$

where $\mathrm{Rot}(\cdot, \cdot)$ represents homogeneous rotation matrix. δ_x, δ_y, and δ_z are the orientation adjustments around axes X_R, Y_R, and Z_R, respectively.

The differential transformation in the frame {R} can be expressed as

$$\Delta = \begin{bmatrix} 0 & -\delta_z & \delta_y & 0 \\ \delta_z & 0 & -\delta_x & 0 \\ -\delta_y & \delta_x & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \qquad (3)$$

The corresponding differential transformation in the frame {E} yields

$$^e\Delta = \begin{bmatrix} 0 & -^e\delta_z & ^e\delta_y & ^ed_x \\ ^e\delta_z & 0 & -^e\delta_x & ^ed_y \\ -^e\delta_y & ^e\delta_x & 0 & ^ed_z \\ 0 & 0 & 0 & 0 \end{bmatrix} \qquad (4)$$

where $^e\delta_x$, $^e\delta_y$, and $^e\delta_z$ are differential rotations around axes X_E, Y_E, and Z_E. ^ed_x, ^ed_y, and ^ed_z are differential translations around axes X_E, Y_E, and Z_E.

With equivalent differential transformation, equations (3) and (4) satisfy

$$^e\Delta = T^{-1}\Delta T \qquad (5)$$

Combining (3) and (5), the differential transformation (4) can be rewritten as

$$^e\Delta = \begin{bmatrix} 0 & -\delta \cdot a & \delta \cdot o & \delta \cdot (p \times n) \\ \delta \cdot a & 0 & -\delta \cdot n & \delta \cdot (p \times o) \\ -\delta \cdot o & \delta \cdot n & 0 & \delta \cdot (p \times a) \\ 0 & 0 & 0 & 0 \end{bmatrix} \qquad (6)$$

Since component A is grasped by robotic end effector, the position of component A with respect to frame {E}, that is, $^ep_A = [l_x, l_y, l_z]^\mathrm{T}$ can be obtained from the CAD models and layouts of the gripper and component A. The homogeneous differential translation of component A resulting from differential orientation of robotic end effector can be computed.

$$^ed' = {}^e\Delta \cdot \begin{bmatrix} ^ep_A \\ 1 \end{bmatrix} = \begin{bmatrix} -l_y \cdot \delta \cdot a + l_z \cdot \delta \cdot o + \delta \cdot (p \times n) \\ l_x \cdot \delta \cdot a - l_z \cdot \delta \cdot n + \delta \cdot (p \times o) \\ -l_x \cdot \delta \cdot o + l_y \cdot \delta \cdot n + \delta \cdot (p \times a) \\ 0 \end{bmatrix}$$
$$(7)$$

The differential translation of component A with respect to frame {E} can be determined from (7)

$$^ed = R^\mathrm{T} \cdot (\delta \times p) + \delta \cdot \begin{bmatrix} -l_y \cdot a + l_z \cdot o \\ l_x \cdot a - l_z \cdot n \\ -l_x \cdot o + l_y \cdot n \end{bmatrix} \qquad (8)$$

According to equivalent differential transformation, the differential translation of component A in the frame {R} yields

$$d = R \cdot {}^ed - \delta \times p \qquad (9)$$

Combining (8) and (9), the differential translation of component A with respect to frame {R} can be further written as

$$d = R \cdot \delta \cdot \begin{bmatrix} -l_y \cdot a + l_z \cdot o \\ l_x \cdot a - l_z \cdot n \\ -l_x \cdot o + l_y \cdot n \end{bmatrix} \qquad (10)$$

The position offset compensation in the frame {R} is $-d$.

Aligning control design. The image-based incremental PI controller used for orientation alignment is designed.

$$\Delta\theta_{Rk} = K_{2p}J_r^{-1}\left(\Delta\theta_{Ik} - \Delta\theta_{I(k-1)}\right) + K_{2i}J_r^{-1}\Delta\theta_{Ik} \quad (11)$$

where lower mark k and $k-1$ represent sampling time. K_{2p} and K_{2i} are proportional and integral coefficients of the incremental PI controller. $\Delta\theta_{Rk} = [\Delta\theta_{xk}, \Delta\theta_{yk}]^\mathrm{T}$ is the orientation adjustment of robotic end effector on sampling time k. $\Delta\theta_{Ik} = [\Delta\theta_{1k}, \Delta\theta_{2k}]^\mathrm{T}$ consists of the angle errors of component B's central axis and component A's central axis in images of microscopic cameras. Image Jacobian matrix J_r represents the relation between image line angle's change and rotational increment of robotic end effector with respect to frame {R}.

In position alignment control, the desired position of component A is computed based on the position and orientation of component B. The orientation of component B in
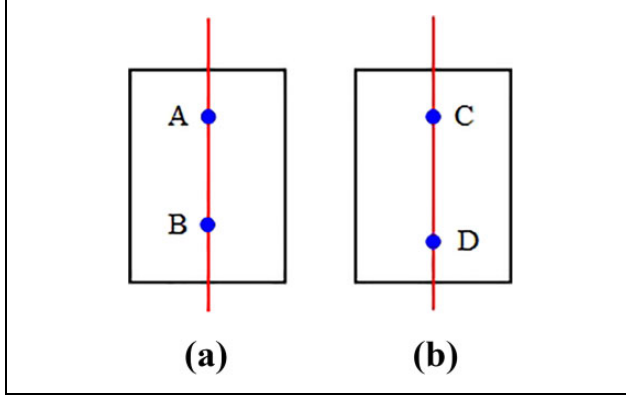
Figure 8. Orientations of component B in images. (a) Image captured by camera 1. (b) Image captured by camera 2.

the frame {R} is estimated firstly. As shown in Figure 8, in the image of microscopic camera 1, we extract two feature points A and B on the central axis of component B. In the image of microscopic camera 2, two feature points C and D on the central axis of component B are extracted. Then the orientations of component B in images of microscopic cameras are $d_1 = [du_1, dv_1]^T = [u_A - u_B, v_A - v_B]^T$ and $d_2 = [du_2, dv_2]^T = [u_C - u_D, v_C - v_D]^T$. According to the definition of image Jacobian matrix, the orientation of component B in the frame {R} can be expressed as

$$\begin{bmatrix} d_x \\ d_y \\ d_z \end{bmatrix} = \begin{bmatrix} J_{t11} & J_{t12} & J_{t13} \\ J_{t21} & J_{t22} & J_{t23} \\ J'_{t31} & J'_{t32} & J'_{t33} \end{bmatrix}^{-1} \begin{bmatrix} du_1 \\ dv_1 \\ 0 \end{bmatrix} \quad (12)$$

where $J'_{t31} = du_2 \cdot J_{t41} - dv_2 \cdot J_{t31}$, $J'_{t32} = du_2 \cdot J_{t42} - dv_2 \cdot J_{t32}$, and $J'_{t33} = du_2 \cdot J_{t43} - dv_2 \cdot J_{t33}$. J_{t11} to J_{t43} are elements of image Jacobian matrix J_t. J_t describes the relation between image point feature's variation and the translational movement of robotic end effector in the frame {R}, which is a constant matrix and can be calibrated with active motions.

As shown in Figure 9, the desired alignment position for component A yields

$$\begin{bmatrix} u_{d1} \\ v_{d1} \\ u_{d2} \\ v_{d2} \end{bmatrix} = l \cdot J_t \cdot \begin{bmatrix} d_{x1} \\ d_{y1} \\ d_{z1} \end{bmatrix} + \begin{bmatrix} u_{b1} \\ v_{b1} \\ u_{b2} \\ v_{b2} \end{bmatrix} \quad (13)$$

where $[d_{x1}, d_{y1}, d_{z1}]^T$ is the normalized orientation vector of component B. (u_{b1}, v_{b1}) and (u_{b2}, v_{b2}) are the centers of component B's upper edge line in images of microscopic camera 1 and camera 2. l is the distance between the desired alignment position and the upper edge's center of component B in the frame {R}, which is given according to the requirement of position alignment.
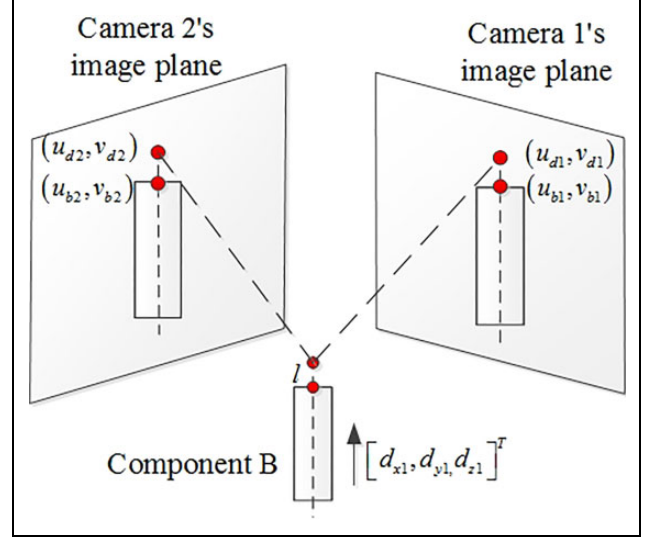

Figure 9. The desired image features in aligning stage for assembly.

The position adjustment for pose alignment is

$$\begin{cases} \Delta p_{pk} = K_{3p} J_t^+ \left(\Delta P_{Ik} - \Delta P_{I(k-1)} \right) + K_{3i} J_t^+ \Delta P_{Ik} \\ \Delta p_k = \Delta p_{pk} - d \end{cases} \quad (14)$$

where lower mark k and $k-1$ represent sampling time. K_{3p} and K_{3i} are proportional and integral coefficients of the incremental PI controller for position alignment. $\Delta p_{pk} = [\Delta x_{pk}, \Delta y_{pk}, \Delta z_{pk}]^T$ is position adjustment for position alignment. $\Delta p_k = [\Delta x_k, \Delta y_k, \Delta z_k]^T$ is total position adjustment in pose alignment, which consists of position adjustment Δp_{pk} for position alignment and position offset compensation $-d$. $\Delta P_{Ik} = [\Delta u_{1k}, \Delta v_{1k}, \Delta u_{2k}, \Delta v_{2k}]^T$ represents position alignment errors in images of microscopic cameras. J_t^+ is the pseudo-inverse of J_t, which can be computed based on least squares.

Inserting control design. In the inserting stage, the force-based feedback control law is given

$$\begin{cases} \begin{bmatrix} \Delta x_f \\ \Delta y_f \end{bmatrix} = J_f^{-1} \begin{bmatrix} F_X \\ F_Y \end{bmatrix} & \text{if } |F_X| > T_2 \text{ or } |F_Y| > T_2 \\ \Delta z_f = \begin{cases} L_{Z1} & \text{if } |F_X|, |F_Y| < T_1, |F_Z| < T_1 \\ L_{Z2} & \text{if } |F_X|, |F_Y| < T_2, |F_Z| < T_3 \\ 0 & \text{others} \end{cases} \end{cases} \quad (15)$$

where Δx_f, Δy_f, and Δz_f are the movements of robotic end effector along axes X_E, Y_E, and Z_E. L_{Z1} and L_{Z2} are step length during inserting. T_1, T_2, and T_3 are the designed
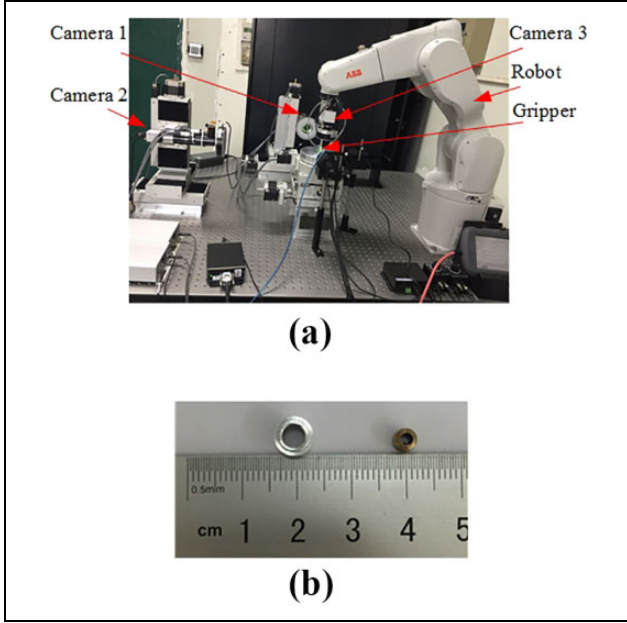
Figure 10. The robot assembly system and components. (a) Robot assembly system. (b) Components to be assembled.

force thresholds according to assembly requirements. Jacobian matrix J_f represents the relation between the movement of robotic end effector and the change of contacted forces, which can be calibrated in advance.

Experiments

Robot assembly system and Jacobian matrices calibration

The established robot assembly system is shown in Figure 10(a). The components to be assembled are shown in Figure 10(b), and both components are small parts with cylindrical symmetry structure.

The Jacobian matrices J_r, J_t, J_h, and J_f are calibrated with active motions.[29] In calibrating J_h, the depth of robotic end effector changes little around the depth of alignment position, which can guarantee the stability of the control system when the image Jacobian matrix of the desired position is used in image-based visual servoing.[30] While in calibrating J_r and J_t, the robotic end effector rotates and moves in the microscopic cameras' clear imaging planes, respectively. Then the Jacobian matrices can be computed according to least square method.

$$J_r = \begin{bmatrix} -0.02 & 0.89 \\ -0.94 & -0.05 \end{bmatrix}$$

$$J_t = \begin{bmatrix} 94.13 & -1.12 & -1.32 \\ -2.38 & -0.19 & -93.07 \\ 4.00 & 93.40 & -2.17 \\ 1.26 & -2.34 & -94.36 \end{bmatrix}$$
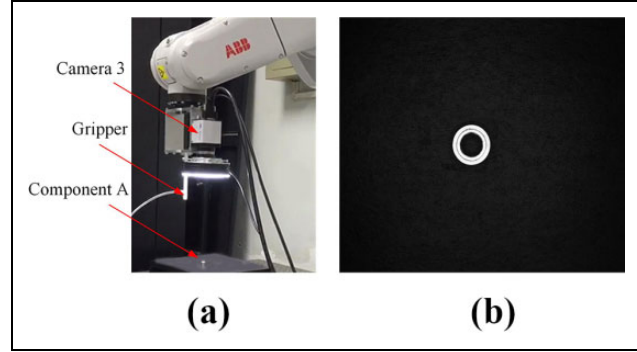

Figure 11. Status of camera 3 and component A's desired image after aligning stage for grasping. (a) Status of camera 3. (b) Component A's desired image.

$$J_h = \begin{bmatrix} -42.93 & 0.27 & 0.43 \\ -0.21 & -42.32 & -0.18 \\ 5.56 & -1.66 & 725.61 \end{bmatrix}$$

$$J_f = \begin{bmatrix} 3.85 & -0.73 \\ -0.22 & 3.40 \end{bmatrix}$$

Grasping experiments

The desired image features of component A and the position change P in grasping subtask are recorded as $(u_d, v_d, s_d) = (1125 \text{ pixels}, 1025 \text{ pixels}, 69{,}369.85 \text{ pixel}^2)$, $P = [0, 60, -93.35]^T \text{mm}$. The parameters of the controller (1) are $K_{1i} = 0.4$ and $K_{1p} = 0.15$. The position error threshold in aligning stage is set 20 μm via experiments. After aligning for grasping, the status of camera 3 and the desired image of component A are shown in Figure 11. In one grasping experiment, the initial position error is $[\Delta x, \Delta y, \Delta z]^T = [3125.4, -5519.5, -4384.1]^T$ μm in aligning stage for grasping.

The position error in aligning stage for grasping is shown in Figure 12. It can be seen that the position feature error reduces quickly and steadily. The aligning stage for grasping finishes after eight steps, and the time-consuming is 10 s.

Once the aligning stage for grasping finishes, the gripper moves $-P$ to make the gripper approach component A. Then the gripper grasps component A with absorption force. The total time-consuming for grasping subtask is 22 s.

In comparison experiments, the position-based visual servoing method in Yacine and Rosmiwati[31] is used in aligning stage for grasping. Thirty grasping experiments are conducted using different batches but the same kind of component A. With the 30 grasping experiments, the success rate of the proposed method is 100%. While the success rate of the method in Yacine and Rosmiwati[31] is 90%. The standard deviation, absolute average, and Root Mean Square (RMS) of position errors in the frame {E} are listed in Table 1. Besides, the position alignment errors and
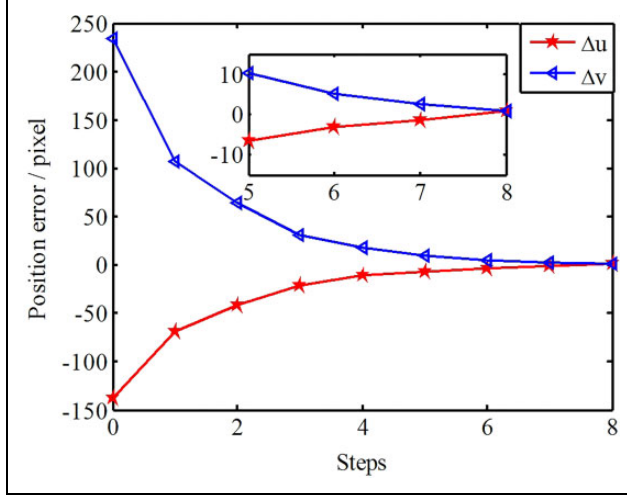
Figure 12. Position error in aligning stage for grasping.

Table 1. Position errors after aligning stage for grasping with the proposed method and comparative method[31] in 30 experiments.

Error type	Proposed method			Comparative method[31]		
	Standard deviation	Absolute average	RMS	Standard deviation	Absolute average	RMS
Δx (μm)	21.1	13.4	16.2	39.2	23.4	36.3
Δy (μm)	21.0	10.3	17.0	33.7	16.8	34.5
Δz (μm)	14.9	11.1	12.1	27.4	19.3	22.1

RMS: Root Mean Square.

the standard deviations of the 30 comparative experiments are given in Figure 13. The red circle represents position alignment error with the proposed method, and the red line passing through the red circle represents corresponding standard deviation. The blue asterisk represents position alignment error with comparative method in Yacine and Rosmiwati,[31] and the blue line passing through the blue asterisk represents corresponding standard deviation. From the experiments, we can see that the proposed grasping strategy has high precision and grasping success rate. With the position-based visual servoing, the component A's 3-D position in the frame $\{C3\}$ is computed with the geometrical algorithm in Chen and Huang.[32] The accuracy of position measurement is low. Besides, the position-based visual servoing needs to calibrate the intrinsic and extrinsic parameters of camera 3. Due to the small depth of field of camera 3, the defocused calibration method in Ding et al.[33] is used, and the calibration process is complicated. Generally speaking, the image Jacobian matrix in image-based visual servoing methods is the function of depth, so it is necessary to estimate the depth. Different from the existing image-based visual servoing methods, the image area of component A is selected to indicate the gripper's depth in our method. The proposed method only needs to calibrate the image Jacobian matrix in the condition that the end

effector is at the alignment height over the component A, it is flexible and convenient.

Assembly experiments

Aligning stage experiments. In aligning stage for assembly, the desired distance for position alignment is set $l = 1.5$ mm. In the image-based microscopic visual servoing with the constant Jacobian matrix, its stability can be ensured with the adequate parameters of PI controller, which can be seen in Liu et al.[19] The parameters of the controllers (2) and (14) are regulated as $K_{2i} = 0.5$, $K_{2p} = 0.15$, $K_{3i} = 0.4$, and $K_{3p} = 0.15$ with the Ziegler–Nichols method. The position of component A in the frame $\{E\}$ when component A is grasped by the gripper is $^e p_A = [0, 0, 171.15]^T$ mm. In one aligning experiment for assembly, the initial position and orientation alignment errors are $[\Delta x, \Delta y, \Delta z]^T = [-760.5, -1433.8, -6367.5]^T$ μm and $[\Delta \theta_x, \Delta \theta_y]^T = [-0.56°, -0.38°]^T$.

In one aligning experiment for assembly, the captured images of two components by microscopic cameras before and after alignment are shown in Figure 14. The position alignment error and orientation alignment error are shown in Figure 15. Figure 15(a) shows the position alignment error, while Figure 15(b) shows the orientation alignment error. From Figure 15, it can be seen that the position and orientation alignment errors reduce stably with only nine steps. After the aligning stage for assembly, the position and orientation alignment errors reach $[\Delta x, \Delta y, \Delta z]^T = [-9.6, -7.5, -13.7]^T$ μm and $[\Delta \theta_x, \Delta \theta_y]^T = [-0.02°, -0.02°]^T$.

Series of comparison experiments are conducted with the decoupled orientation and position alignment method in Liu et al.[19] for assembly. The parameters of the controller in the comparative method are the same with the proposed algorithm. The position offset compensation is introduced into the comparative method in Liu et al.[19] to avoid alignment failure. After aligning for assembly, the standard deviation of position and orientation errors with the proposed method are $x_s = 10.36$ μm, $y_s = 9.36$ μm, $z_s = 14.07$ μm, $\theta_{xs} = 0.03°$, $\theta_{ys} = 0.02°$. And the standard deviation of position and orientation errors with the comparative method in Liu et al.[19] are $x_s = 12.78$ μm, $y_s = 10.67$ μm, $z_s = 13.85$ μm, $\theta_{xs} = 0.04°$, $\theta_{ys} = 0.03°$. From the results, we can see the high precision of the proposed alignment method for assembly.

The time-consuming of aligning experiments for assembly with different initial pose errors are listed in Table 2. From Table 2, we can see that the method in Liu et al.[19] requires more time than the proposed control strategy. Based on the robot assembly platform and proposed alignment method, the position and orientation alignments can be achieved simultaneously. Therefore, the proposed pose alignment method has high efficiency.
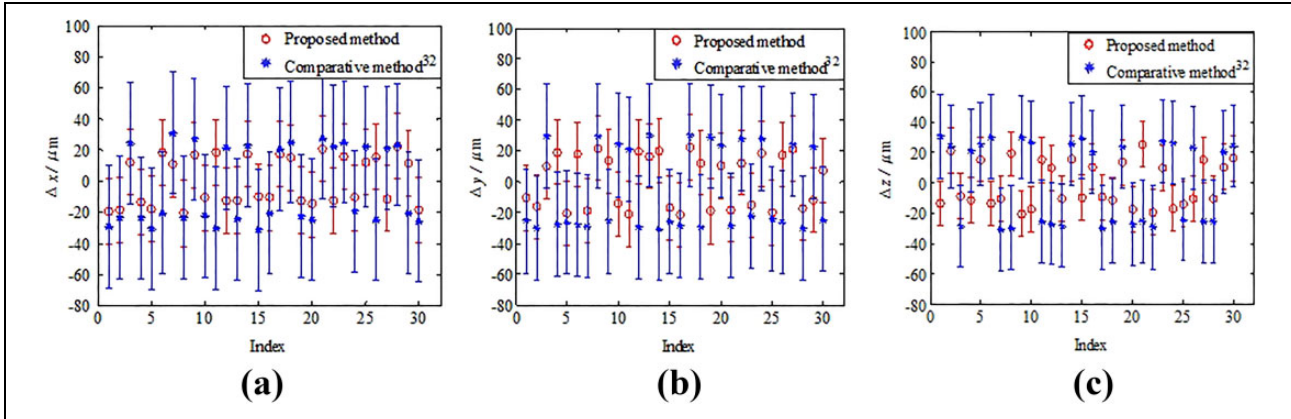
Figure 13. Position alignment errors and the standard deviations of the 30 grasping comparative experiments. (a) Position alignment errors and the standard deviations along axis X_E. (b) Position alignment errors and the standard deviations along axis Y_E. (c) Position alignment errors and the standard deviations along axis Z_E.
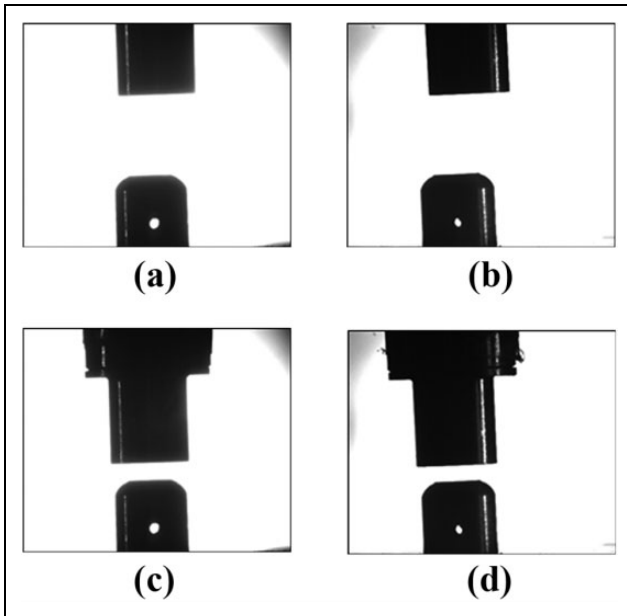

Figure 14. The images captured by microscopic cameras before and after pose alignment. (a) Image captured by camera 1 before alignment. (b) Image captured by camera 2 before alignment. (c) Image captured by camera 1 after alignment. (d) Image captured by camera 2 after alignment.

Inserting stage experiments. In the inserting stage, the parameters of the force-based control law (15) are set $L_{Z1} = 50\ \mu m$, $L_{Z2} = 30\ \mu m$, $T_1 = 20\ mN$, $T_2 = 100\ mN$, and $T_3 = 1000\ mN$ via real inserting assembly experiments. The sampling frequency of the micro force sensor is 1000 Hz. The Butterworth low-pass filter is designed to filter out high frequency noise.

In one inserting experiment, the adjusting curve of robotic end effector and the measured forces are shown in Figure 16. Figure 16(a) shows the adjustment of robotic end effector, while Figure 16(b) shows the contacted forces. From Figure 16, it can be seen that the radial
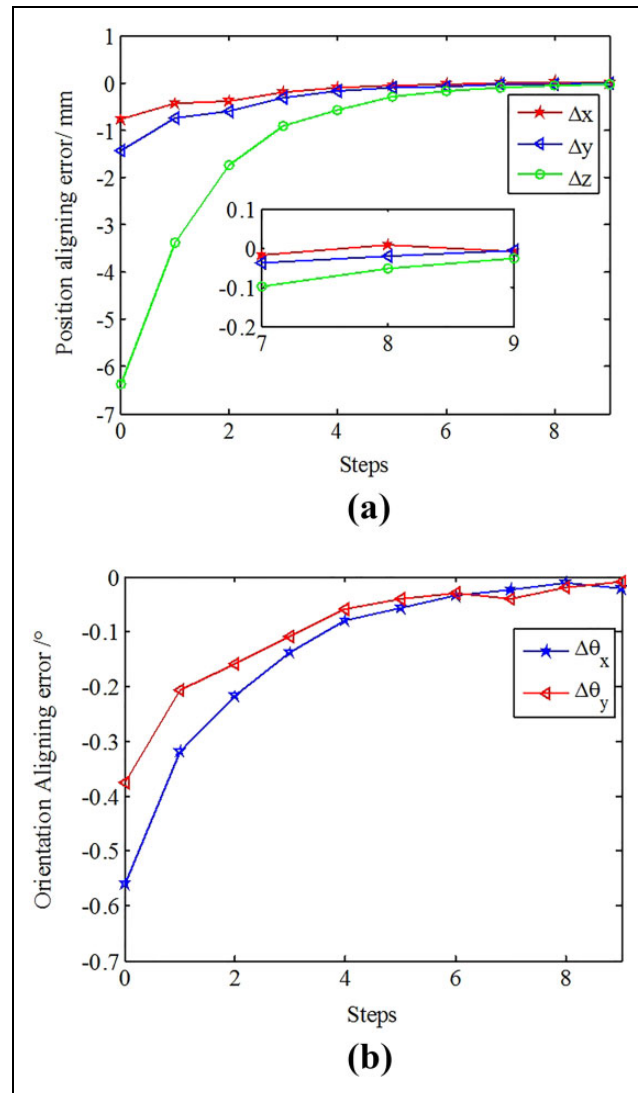

Figure 15. Position and orientation alignment errors in aligning stage for assembly. (a) The position alignment error. (b) The orientation alignment error.

Table 2. Time-consuming of aligning stage for assembly with the proposed method and the comparative method[19] in 7 times experiments.

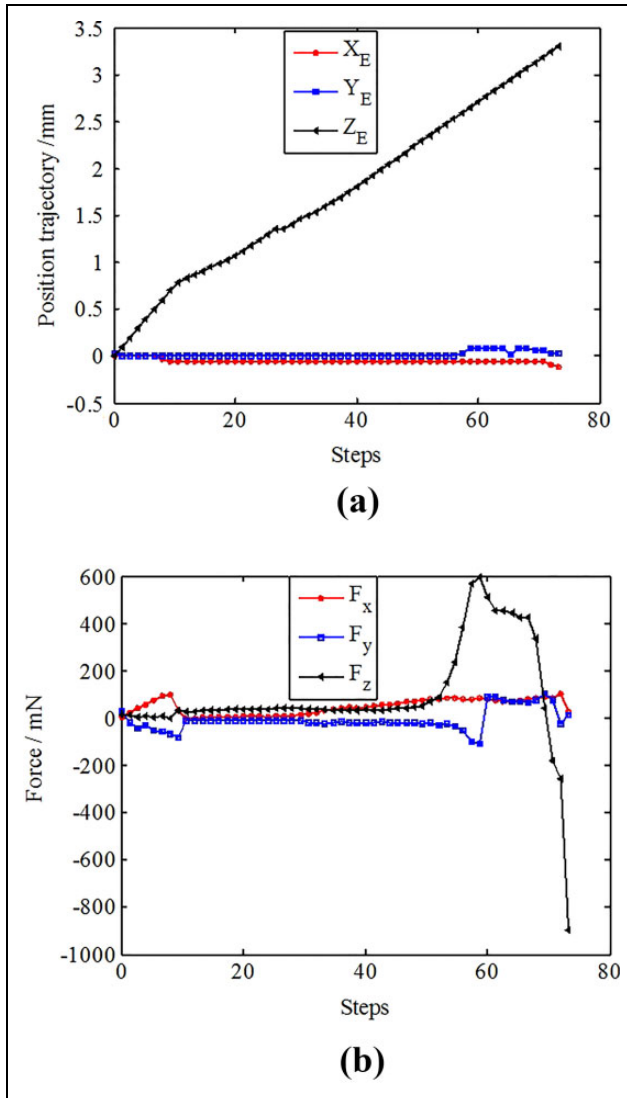| No.	Proposed method Time(s)	Comparative method[19] Time(s)
1	63	86
2	57	78
3	74	90
4	65	87
5	86	103
6	91	115
7	76	95

Figure 16. Position adjustment curve and contacted forces in inserting stage. (a) Adjustment curve of robotic end effector. (b) The contacted forces.

Figure 17. Position adjustment curve and contacted forces in inserting stage with method.[19] (a) Adjustment curve of robotic end effector. (b) The contacted forces.

contacted forces F_X and F_Y are controlled within acceptable range, which avoids the damage to components. In addition, the adjustments of robotic end effector are smooth, which ensures the stability of the inserting process.
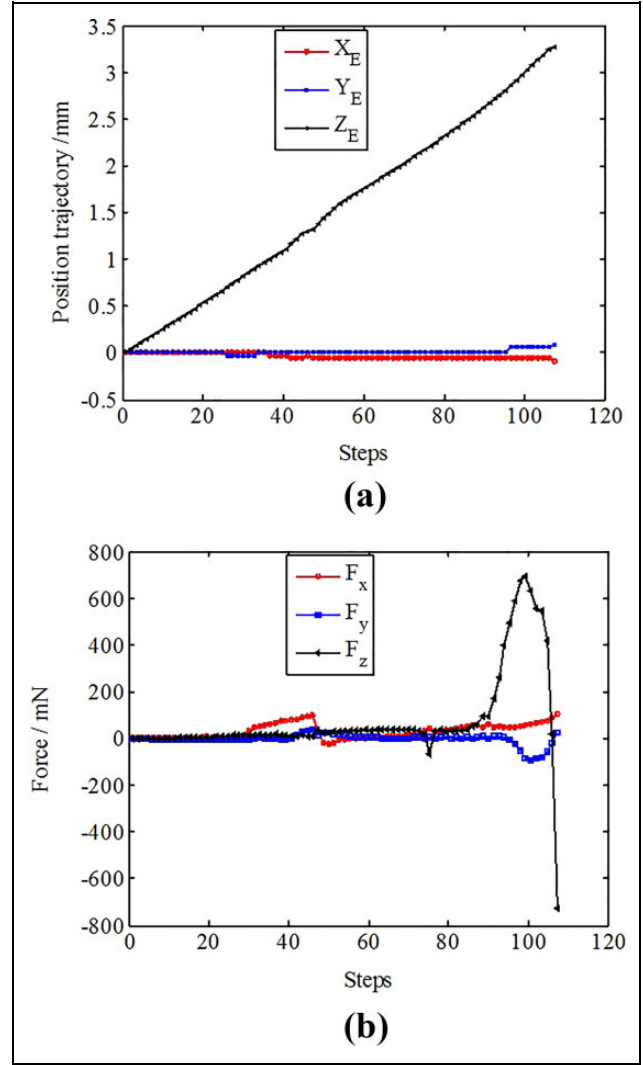
The inserting assembly comparison experiments are conducted with the method in Liu et al.[19] The adjusting curve of robotic end effector and the measured forces are shown in Figure 17(a) and (b). From Figures 16 and 17, it can be seen that the adjusting steps in inserting stage for assembly with the proposed method are 72, and the adjusting steps in inserting stage with the comparative method in Liu et al.[19] are 108. The adjusting steps of the proposed inserting method are less than that of the method in Liu et al.[19] The reason is that the adjusting step length along the axes X_E and Y_E are constants in method in Liu et al.[19] While the adjusting step length of the proposed method along the axes X_E and Y_E are computed based on the forces F_X and F_Y. Compared to the constant step length, the proposed assembly method is more efficient.

Conclusion

An automatic robot assembly system with an industrial robot is established to assemble small components. The designed robot assembly system solves the problem of limited working space for grasping and reduces the complexity of assembly manipulation in the semi-automatic precision assembly with multiple manipulators. The assembly system uses an industrial robot, which improves the efficiency and practicality of assembly as well as lays the foundation for batch assembly in industrial. The proposed "aligning–approaching–grasping" grasping control method can be adapted to grasp other irregular components by replacing the gripper, which is very helpful to eye-in-hand grasping. The proposed "aligning–inserting" assembly control scheme can be used to assemble macroscale/mesoscale or microscale parts with high efficiency. In addition, in aligning stage for assembly, the proposed position offset compensation strategy avoids the grasped component moving out of the microscopic cameras' field of view. Experiments verify the feasibility and high efficiency of the robot assembly system.

The efficiency of inserting stage for assembly is still lower due to the designed step length along axis Z_E, and the alignment accuracy of assembly system need be improved. In the future, we will focus on optimal control methodology to further improve the efficiency of assembly system. Besides, by installing high precision motion stage on robotic end effector to improve the accuracy of assembly system.

ORCID iD

Yanqin Ma 🄳 https://orcid.org/0000-0002-7799-466X

References

1. Zhang J, Xu D, Zhang Z, et al. Position/force hybrid control system for high precision alignment of small gripper to ring object. *Int J Autom Comp* 2013; 10(4): 360–367.
2. Zou Y, Rong W, Sun L, et al. Tele-assembly system for final assembly of the fusion ignition target. In: *Proceedings of the 2014 IEEE international conference on robotics and biomimetics*, Bali, Indonesia, 5–10 December 2014, pp. 1823–1827. Bali, Indonesia: IEEE.
3. Xing D, Xu D, Liu F, et al. Precision assembly among multiple thin objects with various fit types. *IEEE/ASME Trans Mech* 2016; 21(1): 364–378.
4. Zhang P, Xu D, Zou W, et al. Positioning cylindrical target based on three-microscope vision system. *IEEE/ASME Trans Mech* 2014; 19(5): 1612–1624.
5. Cappelleri DJ, Cheng P, Fink J, et al. Automated assembly for mesoscale parts. *IEEE Trans Autom Sci Eng* 2011; 8(3): 598–613.
6. Wason JD, Wen JT, Gorman JJ, et al. Automated multiprobe microassembly using vision feedback. *IEEE Trans Robot* 2012; 28(5): 1090–1103.
7. Xing D, Xu D, and Li H. A sequence of micro-assembly for irregular objects based on a multiple manipulator platform. In: *2014 IEEE/RSJ international conference on intelligent robots and systems*, Chicago, IL, USA, 14–18 September 2014, pp. 761–766. Chicago, IL, USA: IEEE.
8. Shen F, Wu W, Yu D, et al. High-precision automated 3-D assembly with attitude adjustment performed by LMTI and vision-based control. *IEEE/ASME Trans Mech* 2015; 20(4): 1777–1789.
9. Wang L, Mills JK, and Cleghorn WL. Automatic microassembly using visual servo control. *IEEE Trans Elec Pack Manu* 2008; 31(4): 316–325.
10. Liu S, Xu D, Liu F, et al. Relative pose estimation for alignment of long cylindrical components based on microscopic vision. *IEEE/ASME Trans Mech* 2016; 21(3): 1388–1398.
11. Tamadazte B, Piat NL, and Dembele S. Robotic micromanipulation and microassembly using monoview and multiscale visual servoing. *IEEE/ASME Trans Mech* 2011; 16(2): 277–287.
12. Lu Q, and Hermans T. Modeling grasp type improves learning-based grasp planning. *IEEE Robot Autom Lett* 2019; 4(2): 784–791.
13. Wang L, Ren L, Mills JK, et al. Automated 3-D micrograsping tasks performed by vision-based control. *IEEE Trans Autom Sci Eng* 2010; 7(3): 417–426.
14. Liu C, Qiao H, Su J, et al. Vision-based 3-D grasping of 3-D objects with a simple 2-D gripper. *IEEE Trans Syst Man Cybern Syst* 2014; 44(5): 605–620.
15. Sharifi FJ, and Wilson WJ. Automatic grasp planning for visual-servo controlled robotic manipulators. *IEEE Trans Syst Man Cybern B Cybern* 1998; 28(5): 693–711.
16. Recatala G, Carloni R, Melchiorri C, et al. Vision-based grasp tracking for planar objects. *IEEE Trans Syst Man Cybern C Appl Rev* 2008; 38(6): 844–849.
17. Song HC, Kim YL, and Song JB. Automated guidance of peg-in-hole assembly tasks for complex-shaped parts. In: *2014 IEEE/RSJ international conference on intelligent robots and systems, Chicago*, IL, USA, 14–18 September 2014, pp. 4517–4522. Chicago, IL, USA: IEEE.
18. Wyk KV, Culleton M, Falco J, et al. Comparative peg-in-hole testing of a force-based manipulation controlled robotic hand. *IEEE Trans Robot* 2018; (99): 1–8.

19. Liu S, Xu D, Zhang D, et al. High precision automatic assembly based on microscopic vision and force information. *IEEE Trans Autom Sci Eng* 2016; 13(1): 382–393.

20. Zheng Y, Zhang X, Chen Y, et al. Peg-in-hole assembly based on hybrid vision/force guidance and dual-arm coordination. In: *Proceedings of the 2017 IEEE international conference on robotics and biomimetics*, Macau SAR, China, 5–8 December 2017, pp. 418–423. Macau SAR, China: IEEE.

21. Liu S, Li Y, Xing D, et al. An efficient insertion control method for precision assembly of cylindrical components. *IEEE/ASME Trans Mech* 2018; 65(10): 8062–8072.

22. Wiemer SC, and Schimmels JM. Optimal admittance characteristics for planar force-assembly of convex polygonal parts. In: *Proceedings of IEEE international conference on robotics and automation*, Saint Paul, MN, USA, 14–18 May 2012, pp. 2578–2583. Saint Paul, MN, USA: IEEE.

23. Chang WC, Weng YH, Tsai YH, et al. Automatic robot assembly with eye-in-hand stereo vision. In: *Proceedings of the 8th World Congress on intelligent control and automation*, Taipei, China, 21–25 June, 2011, pp. 914–919. Taipei, China: IEEE.

24. Koveos Y, Papageorgiou D, Doltsinis S, et al. A fast robot deployment strategy for successful snap assembly. In: *2016 IEEE international symposium on robotics and intelligent sensors*, Tokyo, Japan, 17–20 December 2016, pp. 80–85. Tokyo, Japan: IEEE.

25. Fang S, Huan X, Chen H, et al. Dual-arm robot assembly system for 3C product based on vision guidance. In: *2016 IEEE international conference on robotics and biomimetics*, Qingdao, China, 3–7 December 2016, pp. 807–812. Qingdao, China: IEEE.

26. Muis A, and Ohnishi K. Eye-to-hand approach on eye-in-hand configuration within real-time visual servoing. *IEEE/ASME Trans Mech* 2005; 10(4): 404–410.

27. Xing D, Liu F, Qin F, et al. Coordinated insertion control for inclined precision assembly. *IEEE Trans Ind Electron* 2016; 63(5): 2990–2999.

28. Xu D, Lu J, Wang P, et al. Partially decoupled image-based visual servoing using different sensitive features. *IEEE Trans Syst Man Cybern Syst* 2017; 47(8): 2233–2243.

29. Xing D, Xu D, Li H, et al. Active calibration and its applications on micro-operating platform with multiple manipulators. In: *2014 IEEE international conference on robotics and automation*, Hong Kong, China, 31 May–7 June 2014, pp. 5455–5460. Hong Kong, China: IEEE.

30. Bruno S, and Oussama K.*Springer handbook of robotics*. Berlin: Springer-Verlag, 2008, pp.565–570.

31. Yacine B, and Rosmiwati MM. Position-based visual servoing through Cartesian path-planning for a grasping task. In: *2012 IEEE international conference on control system, computing and engineering*, Penang, Malaysia, 23–25 November 2012, pp.410–415. Penang, Malaysia: IEEE.

32. Chen Z, and Huang JB. A vision-based method for the circle pose determination with a direct geometric interpretation. *IEEE Trans Robot Automat* 1999; 15(6): 1135–1141.

33. Ding W, Liu X, Xu D, et al. A robust detection method of control points for calibration and measurement with defocused images. *IEEE Trans Instrum Meas* 2017; 66(10): 2725–2735.