Sensor Integration and Fusion for Autonomous Screwing Task by Dual-Manipulator Hand Robot

R.L.A. Shauria,*, K. Saikib, S. Toritani and K. Nonamib

Faculty of Electrical Engineering, Universiti Teknologi MARA, 40450 Shah Alam, Selangor, Malaysia
Graduate School of Engineering, Division of Artificial Systems Science, Chiba University, 1-33 Yayoi-cho, Chiba-shi, Chiba, 263-8522, Japan

Abstract

Sensing and manipulation of small objects are among the most challenging, requiring a very high degree of coordinated actuation in producing assembly work robots. It is inevitable for the development of a robot that is designed to execute complex tasks to apply multiple sensors for determination of a target’s attributes. In this study, a seven-link dual-arm robot with three-fingered hands that exhibits human-like skills to manipulate small parts like nuts and bolts is proposed. The dynamical system of the robot requires attention in many aspects including control, sensor integration, image processing and trajectory planning strategies for small object manipulation. In this paper, the implementation of sensor integration between force and vision sensors for real-time autonomous grasping and screwing of an M10 nut and bolt by the robot is discussed. The measured data on the object or environment is used to create the robot’s trajectory in real-time operation and to ensure safety of the robot. The proposed method applies a visual servoing structure for the recognition by camera, and impedance control for the hand robot to produce softness to the fingers. The required kinematics adjustments to modify the trajectories of arms and the robot hand’s configuration are applied according to the measured data by the sensors. Experimental results indicate the performance of the proposed method in achieving the sensor-based planning and control for fully autonomous screwing task in real-time experiments. This shows the capability of the dual-manipulator hand robot for manipulating small objects and its applicability for more complex assembly tasks in the future.

* Corresponding author. Tel.:+6012-5002912; fax:+603-55435077.
E-mail address: ruhizan@salam.uitm.edu.my

1. Introduction

The application of anthropomorphic multi-arm robots with multiple sensors has attracted much attention. One of the essential features in building an autonomous robot is the selection of suitable sensors to fit the purpose of a robot for the assigned task in real-time operations. Hence, numerous researches have been conducted seeking ways to enhance the sensing of the environment or the object manipulated for robotic applications, which includes image processing and force control techniques. Among the studies are the dual-arm robots produced by Fanuc Ltd. [1] and Yaskawa Electric Corp [2], which use lasers and cameras to enable the robots perform autonomously in assembly lines. Yaskawa’s Motoman robots [3] also apply force sensor on their end effectors but the arms are only assigned for independent tasks and not manipulating an object cooperatively. Even though these robots are built with redundant 7-DOF arms, they use only simple 1-DOF gripper to handle parts.
Other example of using robots for industries includes work by Feddema and Mitchell [4], who showed a vision-guided servoing technique with feature-based trajectory generation on a PUMA robot for tracking the movement of a carburetor gasket. In addition, V. Lippiello et al. [5] proposed position-based visual servoing using two industrial robots named Comau SMART-3 S, with hybrid eye-in-hand (EIH) and eye-to-hand (ETH) multi-camera of a known model work piece. However, these robots manipulate objects that are bigger than assembly parts such as nuts and bolts, which require more accurate manipulation. Furthermore, a German firm, Pi4 Robotics has collaborated with Berlin’s Fraunhofer Institute for Production Systems and Design Technology, to create Pi4 WorkerBot [6], that could perform inspection tasks on a small chrome metal part with vision and force sensors using two arms on a mobile platform. However, despite of its ability to move with the manipulators, the robot has only gripper type of end effectors, which limit its capability for human-like skills manipulation.

The objective of this study is to build a 7-link dual-arm robot for autonomous and complex assembly tasks such as grasping and screwing of nuts and bolts. In this paper, the real-time implementation of the above tasks of an initial target of M10 nut using interaction between the robot and the environment/manipulated object in event-based approach is shown. The sensor integration which includes the visual servoing algorithm for target recognition and the impedance control algorithm for hand position control are discussed. Consequently, the trajectory planning module that contains the generation of dynamic trajectory path based on sensor integration in real-time and for unknown situations, is explained. Finally, experimental results on the recognition by vision, force measurements and positions of the end effector are presented.

2. Robot System Configuration and Structure

The latest version of the Dual-Arm Robot 3 with three-fingered hands, as shown in Fig 2(a), is low cost and is suitable for use on assembly lines. It is built in response to the needs of autonomous small-scale industries that manufacture specialized products. The robot is designed to possess human-like manipulative skills by making both arms similar to that of a human being in terms of shoulder size and the number of joints, except for the robot’s hands.

The robot hands were developed for both the right and left arms having the same structures and designs as shown in Fig 2(b). An additional rotational feature is introduced to increase the DOF of the robot hand and, consequently, enhances the hand’s capability to autonomously grasp objects of various shapes without having to reallocate the finger segments. A stereo camera from Point Grey Research (PGR), as shown in Fig 3(a), acts as an ETH camera and fixed on top of the Dual-Arm Robot 3 to capture a wider view of a worktable placed in front of the robot. Meanwhile, a monocular camera from the same manufacturer, as shown in Fig 3(b), acts as an EIH camera and fixed on each of the end effectors of the robot to provide a proximity view of the targets when the robot arms are moved closer towards the targets. For the sense of touch, 6-axis force-torque sensors, as shown in Fig 3(b), are fixed on each of the fingertips of the left hand.

The system structure of the robot mainly consists of the robot, sensors, actuators and the control mechanism. Vision PCs calculate the vision data from EIH and ETH cameras, while the Host PCs contain the control algorithms. The control input is produced by the Target PCs and executed on the robot via xPC Target on a MATLAB/Simulink environment. Data socket functions are used for receiving and transmitting data between the Target PCs, Host PCs and the vision systems.

3. Sensor Integration and Fusion

The development of planning and control systems that can integrate various types of sensory information and human knowledge is important to produce a robot for real-time operation in unconditioned environment. In this study, sensory information, i.e. from vision of several cameras and force measurement of the robot fingers are integrated to provide the required position for the robot end effector. Even though vision sensor can be a powerful single source of sensory information, its integration with force measurement is very useful to compensate the difficulty encountered especially when assembling small object which requires robot’s sensitivity to be as similar as possible to human senses. As shown in Fig 4, two separate modules from both vision and force are responsible for generating the robot’s trajectory. The position and orientation of the target from vision is used to calculate the pose of the end effector, while the force data of hands are used in the impedance control algorithm to prevent collision between the robot arm and the environment, and to reposition nut during the screwing task. The current and desired postures of the end effector, denoted by \( r_{AWS} \) and \( r_{AHD} \), respectively, are then used to generate the reference trajectory, \( r_{ref} \). Finally, the reference angles for all links, \( \Theta_{ref} \), is calculated using inverse kinematics calculations of the robot arm.

3.1. Hierarchical Phase-Template Paradigm

The multisensory integration approach that is applied in this study falls into the hierarchical phase-template paradigm, which has been discussed by Luo and Kay in [7]. The four phases which are “far away”, “near to”, “touching” and
“manipulate”, show how the sensors are being utilized for the required purposes to assist the robot for the completion of the task. For the first phase, “far away”, an overview of the worktable by vision camera provides the detection and location while excluding the unrelated target for further data processing in the following phase. In the “near to” phase, an EIH camera mounted on the robot arm is used to provide the proximity views for detailed recognition process. The integrated information from previous phases then is used to implement the “touching” phase, in which the force sensor is used to absorb the collision between fingers and the worktable. Finally, the “manipulate” phase integrates both the vision and force sensor information to match the position between both nut and bolt for screwing.

Fig. 1. Hexagonal nuts and bolts

Fig. 2. (a) Dual-Manipulator Hand Robot 3 (b) Robot hand (right)

Fig. 3. External sensors: (a) stereo camera (b) monocular camera (c) 6-axis force sensor

Fig. 4. Control Scheme of Robot Arm
3.2. Three-Level Hierarchy

The application of multiple types of sensors can increase the accuracy or the determination of a measured parameter by improving the weaknesses of each other on the system. This will then result to a reduced error of the measured parameter or failure in achieving the assigned task. Data fusion has been classified by [8, 9] into several levels including data-level fusion, feature-level fusion and decision-level fusion. In this study, data-level and feature-level fusions are both applied at the image processing stage, where the former is used in the triangulation process of two 2-D images from ETH camera, while the latter is used for the target recognition and pose estimation based on the EIH camera views. Finally, the decision-level fusion is applied by integrating the vision and force data to yield the final inference for achieving the task.

4. Sensor Fusion in Planning and Control

4.1. Position-based Visual Servoing Structure for Target Recognition

In the most related studies, robots are dependent on the direct information of the target posture obtained by image processing methods prior to manipulations performed by the robot. A vision-based closed-loop motion control or visual servoing is needed to control the pose of the robot’s end effector relative to the target. In this study, a robotic visual servoing (RVS) structure that provides the position estimation based on the constraints of small-sized nuts to control the arm posture is applied. This method executes iterative steps that include recognition, confirmation, estimation and rotation of robot joint which is abbreviated as CER (Confirm-Estimate-Rotate), as detailed in [10]. The advantage of this method is that, it can ensure a safer generation of the complex trajectory of the seven-link robot. Besides, within the constraints of the target, this method is capable of increasing the accuracy of the calculations that are difficult to solve using other pose estimation methods such as by Zhang [11] (closed-form method) and DeMenthon and Davis POSIT (iterative method) [12]. Thus, in this study, a static position-based look-and-move approach as illustrated in previous Fig 4, is proposed. The control architecture is hierarchical, in that the vision is used to calculate the reference for the arm’s posture in task coordinates, and a feedback loop is applied internally to control the position of the joints. The vision and the robot control routines are performed in a sequential manner, in which the robot needs to be in a complete rest before the camera could acquire a new image. The detail configuration of this module is shown in Fig 5 which can be divided into three parts; target recognition by the EIH camera, target pose estimation before grasping, and target position matching before screwing.

![Fig. 5. Divisions of vision processing module](image-url)
4.2. Force Control for Hand robot

The force control introduced by K. Saiki in [13] uses impedance control algorithm to calculate the finger position based on external force, \( f_{ext} \), measured by the force sensor. This provides the softness to hand when colliding with the worktable and also solves the object slip problem in the screwing task. According to the hand block diagram of the force control module shown in previous Fig 4, the reference force, \( f_{ref} \) and \( f_{ext} \) produce the input force values for the impedance model to generate the displacement \( \Delta P_k \) that changes the previous reference position, \( P_P \), to produce the new reference position \( P_{ref} \) of the fingers. The Eq. (1) below represents the dynamic equation used to implement the impedance control of all three fingers. \( M_d \), \( D_d \) and \( K_d \) are the virtual mass, damping and stiffness coefficients, respectively, determined through trial and error in experiments and can be illustrated as in Fig 6.

\[
f_{ext} - f_{ref} = M_d(\ddot{P} - \ddot{P}_d) + D_d(\dot{P} - \dot{P}_d) + K_d(P - P_d)
\]  

(1)

Besides, for the screwing, after the first matched position is given by vision as illustrated in the “Target Position Matching” in Fig 5, the force measurement is applied to correct the contact position of both targets in order to make them fit for screwing. The moment equations (2) ~ (4) are used to estimate the contact point, \((x_t, y_t)\) when nut make contact with the bolt. Here, \((x_1, y_1) \), \((x_2, y_2) \) and \((x_3, y_3) \) are the positions of each fingertips, and \((f_{z_1}, f_{z_2}, f_{z_3}) \) is the external force measured by each fingers in z-axis.

\[
f_{z_1}x_1 + f_{z_2}x_2 + f_{z_3}x_3 = (f_{z_1} + f_{z_2} + f_{z_3})x_t
\]  

(2)

\[
f_{z_1}y_1 + f_{z_2}y_2 + f_{z_3}y_3 = (f_{z_1} + f_{z_2} + f_{z_3})y_t
\]  

(3)

\[
x_t = \frac{f_{z_1}x_1 + f_{z_2}x_2 + f_{z_3}x_3}{(f_{z_1} + f_{z_2} + f_{z_3})}, \quad y_t = \frac{f_{z_1}y_1 + f_{z_2}y_2 + f_{z_3}y_3}{(f_{z_1} + f_{z_2} + f_{z_3})}
\]  

(4)

4.3. Grasping and Screwing Trajectory Planning

In this sub-section, trajectory for screwing task is explained. The fingers of each arm will grasp the nut and bolt individually before the arms changed to the screwing pose. Then, the arms’ position is corrected based on the recognition of the bolt held using the left camera on the left hand of robot. Consequently, with respect to the screwing strategy, after the nut and bolt make their first contact, the robot trajectory is designed such that it can automatically recheck the force measurement and correct the position of the nut. The correction is based on the moment of force calculation applied by the bolt onto the surface of nut and when it finds the best fit, it will begin to rotate the end effector for screwing both targets together. The correction by measuring the moment of force between the two targets has been detailed in previous section. The integration between vision and force control for both tasks in continuous steps, can be described by the flowchart in Fig 7. The force thresholds \( \alpha \) and \( \gamma \) are predetermined values by trial and error.

Fig. 6. Virtual damping and stiffness coefficients for impedance control
5. Experimental Results

5.1. Processed Image for Target Recognition

ETH camera through its recognition algorithm has identified a small object on the worktable, marked with a red dot in Fig 8(a). The position of the object is sent to the robot arm, and as the arm approaches the object, the recognition by EIH algorithm begins. The results obtained by the EIH camera, as depicted in Fig 8(b), shows that the proposed vision has performed well when recognizing multiple objects consisting of M8, M4 and M2 nuts, and M8 and M10 bolts for grasping. Consequently, Fig 9 proves that the applied CER method is capable of identifying nuts placed at different orientations.

![Diagram](image)

Fig. 7. Integration of sensors for grasping and screwing tasks

![Images](image)

Fig. 8. (a) Detection using ETH camera (b) Target recognition by EIH camera

![Images](image)

Fig. 9. Estimating the inclination direction of nuts at different orientations
5.2. Vision and Force Data for Screwing

The robot depends on real-time vision and force data in a sequence of steps as depicted in Fig 10 for the autonomous screwing task. First, the position of the end effector (EE) in y-z direction is corrected based on measurement by vision as shown in Fig 11. It can be observed that the position by vision, denoted by \(\{Y_{EE}, Z_{EE}\}\), at 805 s, is used to align the y-z position of the left end effector \(\{Y_{Exp}, Z_{Exp}\}\) at 807 s. The efficiency of the vision can be determined by measuring the force at each finger of the left hand that holds the nut when it makes contact with the bolt. It is assumed that if the nut makes full contact with the surface of the bolt’s tail, which is not a desirable result, then the contact force in the z-direction of the fingers could exceed 1 N (positive and negative signs). It can be observed from the force measurements depicted in Fig 12 at around 834 s, where the measured forces from all fingers are small, thus shows that vision is reliable to guide the hand to ensure the contact of targets. Consequently, position correction steps based on force calculation are implemented, which consists of judgement 1 to confirm the alignment, estimation of contact point, and judgement 2 to confirm the fit. As shown in Fig 12, the estimation of the contact point and the arm’s position corrections are made repeatedly at 837 s, 854 s and 866 s before the bolt is judged fit into the nut to proceed screwing at 870 s. The execution of each step in Fig 10 is represented with the same line styles in Fig 12. Consequently, the position correction of the end effector in y-z direction for screwing is shown in Fig 13, similarly with the steps involved in Fig 10 outlined accordingly. The results showed how both of the sensors data are used effectively for the robot to achieve the complex screwing task.

From ten trials that have been conducted, without the integration of the force sensor with vision, the success rate was only 15%. However, in another ten trials when force sensor is integrated with vision, the autonomous check and re-correct trajectory strategies worked successfully, and the success rate has increased to 60%. Finally, as depicted by the snapshot sequence in Fig 14, the robot grasped the nut and bolt in pictures 2 and 3, respectively, before directing the arms to the screwing pose as shown in picture 4. For screwing, the bolt recognition by vision is done in picture 5, and is used to move the left arm closer to the bolt, as in picture 6. After contact between both targets is made, the position corrections of the left hand are implemented, resulted to the fit contact between targets as shown in picture 7, followed by the screwing actions in picture 8, before fingers release the nut in picture 9 to complete the task.

![Fig. 10. Matching nut and bolt position in y-z direction by force for screwing](image1)

![Fig. 11. Matching targets’ position in y-z direction by vision for screwing](image2)
Fig. 12. Estimation of contact point based on external force

Fig. 13. End effector’s position correction (a) y-axis  (b) z-axis

Fig. 14. Grasping and screwing in autonomous steps
6. Conclusion

A combination of vision from multiple cameras, such as the ETH and EIH cameras, has been used to tackle the constraints of small and indistinguishable coloured nuts and bolts as the targets. The integration of sensors has been applied in which the vision and the force sensors operate sequentially, such that the vision could provide the initial measurements while the force sensors start measuring only after the fingers make contact with the environment or target. Furthermore, the trajectory planning, which utilizes the sensor outputs, is designed with the manipulation strategies to successfully control the robot movement. Though the accuracy of manipulation has not been fully achieved in the experiments, the experimental results have proven that the system could be put into practice in real-time operations.

Several real-time experiments to measure the performance of the proposed methods on actual targets have been conducted. A 60% success rate has been achieved in the screwing tests involving M10 nut and bolt. However, the method used must be further improved before the robot is able to successfully screw M2 nut and bolt. Noting from the experimental results, it can be concluded that the dual-manipulator hand robot is capable of executing the screwing task autonomously without human assistance.

References