In-Hand Pose Estimation with Optical Tactile Sensor for Robotic Electronics Assembly

Uksang Yoo and Wenzhen Yuan

The Robotics Institute, Carnegie Mellon University

***Abstract*—This report explores industrial electronics assembly application of tactile sensing with GelSight. A GelSight vision-based tactile sensor was attached to a parallel gripper and a 6 degree of freedom robot arm. Rudimentary in-hand pose estimation method of the grasped electronic components was proposed based on HSV thresholds. Then, a case study of electronics assembly tasks is presented. Based on the detected pose of the component, motion is planned to achieve desired rotation and translation with respect to the component frame. Then GelSight's marker displacement vectors are considered to detect component's fit with the PCB through-holes. Case study applications clearly showed that GelSight can provide useful tactile feedback for assembly tasks requiring precision and real-time pose estimation.**

I. Introduction

Despite numerous advancements in printed circuit board (PCB) assembly automation such as robotic surface mount technology, robotic insertion of odd-form electronic components presents a challenge. Primarily, automation of electronic component insertion into the PCB's through-holes still often requires expensive specialized equipment and suffer from low reliability in operation [1]. Much of these assembly tasks then have to be performed manually with human labor, adding to both assembly time and cost [2]. In such applications, articulated robotic arms and grippers have the potential to significantly reduce costs of equipment and operation due to their general applicability to different components without requiring specialized modifications. However, the integration of these robots in PCB assembly lines is hindered by their lack of ability to adapt to variances in grasping conditions, in-hand pose of the components and task environment. Then, improvements must be made to sensory feedback particularly pertaining to the component's in-hand pose during assembly tasks.

Pose estimation of objects is often done by visual perceptions in which position and orientation are estimated by various techniques such as template matching, iterative-closest-point (ICP), and various learning techniques [3]. However, using external vision for in-hand pose estimation is limited by occlusion, especially when the electronic components are relatively small compared to the gripper grasping it. Furthermore, visual perception lacks the ability to estimate applied force and torque measurement which is crucial during robotic manipulation to prevent slippage and to perceive interaction with the environment.

High-resolution tactile sensors such as GelSight can help to eliminate the hindrance and slippage problem. They can complement visual perception by measuring the direct contact force and torque distribution at the robotic gripper. Since such tactile sensors have only become available recently, how to best use this new modality for in-hand pose estimation, together or independent of visual perception, is still an open research problem. Performing in-hand pose estimation with tactile perception in the context of robotic electronics assembly is the focus of this investigation.

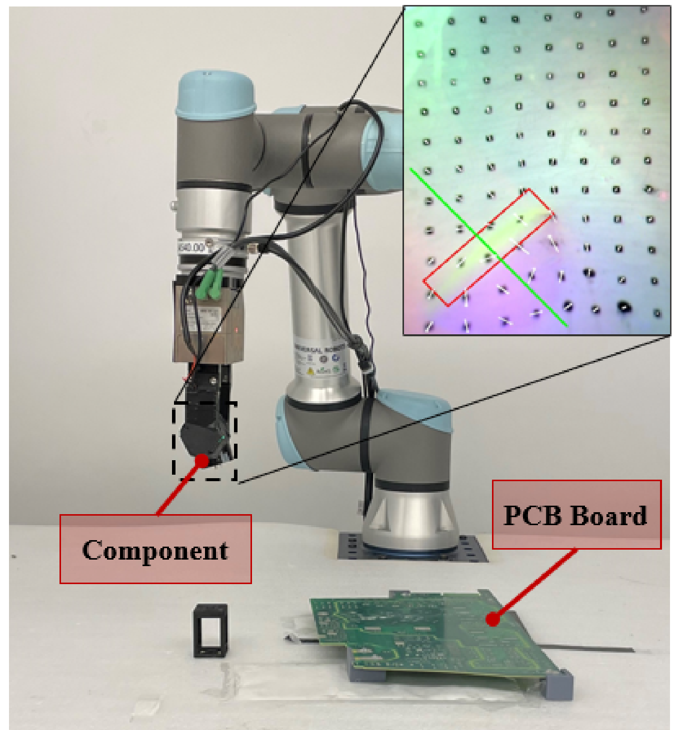

Fig. 1. UR5e 6-DoF robot arm and a WSG50 parallel gripper with the GelSight attachment that is gripping an electronic component. GelSight view is also illustrated with an automatically generated pose estimation where the red rectangle highlights the detected top edge of the component and the green line indicates the object's central axis.

We first present a HSV thresholding and arbitrarily aligned bounding method for contact edge detection using GelSight. We use the contact edge to estimate the in-hand pose of the components. Then we introduce two case study electronics assembly tasks to further explore GelSight's utility in industrial applications. In each task, we let the controller adjust either the component's orientation or its position to achieve proper fit with the PCB. The report then concludes with potential future works to expand on this exploratory study toward reliable implementation in industrial settings.

II. Related Work

In-hand pose estimation of small objects has been a subject of study in recent works [4]–[7]. Specifically, researchers have extracted USB connector pose from GelSight images through BRISK feature-based matching [4]. Later works have looked at matching simulated GelSight images from object's accurate 3D geometric models to estimate the object's current pose probability distribution [6]. A recent work studied using the in-hand pose estimation of a USB cable extracted as principal component axes on the GelSight depth image to guide the cable connector end to the USB port [7]. In our study, we focus a largely object-agnostic top edge detection based on GelSight image color thresholds that can be used with known geometric dimensions of the component to get the object pose estimation.

Researchers have also studied tactile sensing to detect grasp quality and slip. Particularly, researchers took advantage of GelSight's marker displacement information to detect slip based on the relationship between applied shear forces and marker displacements [8]. In another work, the researchers set empirically derived thresholds on ratio between the maximum marker displacement and peripheral marker displacement around the contact boundary to detect slip [9]. For this study, we will explore the application of these marker displacements to guide insertion of components into under-defined holes.

III. Top Edge Detection

Since it is likely that the grasped component will not lie entirely within the sensorized region of GelSight, it is necessary to rely on the contact edges that are visible to estimate the pose of the component. For simplicity, we assume here that the component's top edge is always visible. Otherwise, a more sophisticated shape mapping method will have to be utilized. In this section, we discuss the methods used to detect contact with the component and estimate the component's position and orientation in the GelSight view. The following assumptions are made to simplify the problem of in-hand pose estimation:

1) Component's in-plane movement is considered fully constrained by the gripper.
2) The component is assumed to be nearly symmetric with continuous edges.
3) The top edge of the component is sufficiently wide (> 5 mm).

The raw image from GelSight without markers is illustrated in Fig. 2A. The Red-Green-Blue (RGB) LED lights embedded in the sensor, highlight gelpad's indentation from contact [10]. Specifically, the top edge of the contact region noticeably stands out because it generates the steepest depth change in the surface of the gelpad.

To isolate the component's top edge based on the assumptions above, we subtracted the initial GelSight frame without contact from the current frame (Fig. 2B). Then, the top edge contact region was isolated based on HSV threshold to generate a binary image of the contact edge (Fig. 2C). Then a boundary box of the contact edge region with minimum

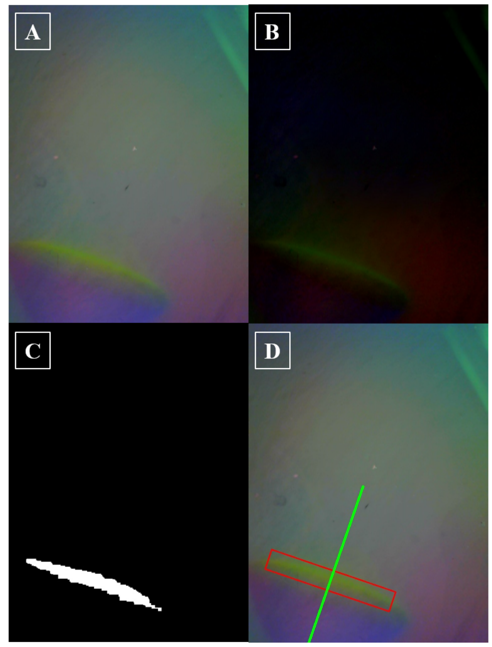

Fig. 2. GelSight view of the contacted capacitor slanted at 30 degrees. Fig. 2A shows the raw image view. Fig. 2B shows the image subtracted by the initial GelSight frame. Fig. 2C shows the HSV thresholded binary mask. Based on the mask, Fig. 2D shows the boundary rectangle and orientation of the component. Fitting a ellipse was also attempted; however, the angle estimations were found to be noisier.

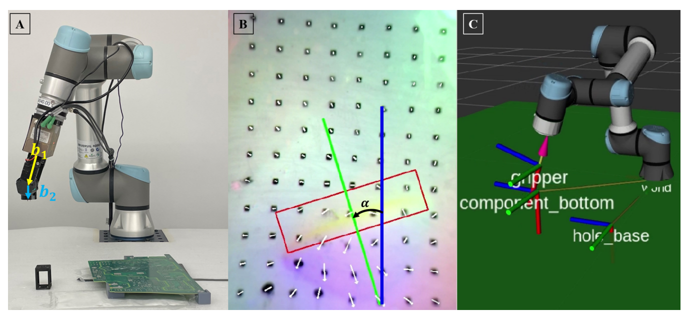

Fig. 3. A: Front view of the robot grasping a capacitor. b_1^G represents the translational vector from gripper frame to the detectected top frame of the component in gripper coordinate frame. b_2^C is the translational vector from the top of the component to the component's bottom. B: GelSight view of the component contact. C: RViz visualization of the frames. Red segment represents the x-axis, green segment represents the y-axis and blue segment represents the z-axis.

area was computed based on rotating calipers algorithm [11]. Based on assumption (2), we are able to find the orientation and position of the component's top edge in GelSight frame with the rectangle's axes of symmetry. Based on assumption (3), we can use the rectangle's secondary axis through the rectangle's width to define the component's central axis.

IV. Reference Transformation

This section will briefly discuss the use of the pose estimation in assisting component manipulation. We first note that it is convenient to perform motion planning with respect to the component frame. This is especially evident when the task requires isolated rotation around the object's central axis even when it is misaligned with the gripper's frame. Given the inverse kinematics model of the robot, it is sufficient to define the spatial transformation matrix from the gripper frame to the component frame.

In the tasks presented in this report, the component's central/x-axis (as defined in Fig. 3C) orientation is either the target of adjustment through interaction with the PCB or is assumed to be known by mechanical constraint. And since the component's z-axis rotation and the its position with respect to the gripper's y-axis is constrained by the grippers based on assumption (1), we only need to define the component's rotation with respect to gripper's y-axis and its position with respect to gripper's x and z-axes to fully define the component's pose.

We can define a gripper-fixed vector b_1 from the gripper frame to the center of GelSight's gelpad. Since the gripper and GelSight are aligned, we can define it as b_1 in gripper frame as $b_1^G = [d_c, 0, 0]^T$ where d_c is the distance from the gripper frame to the center of the gelpad. Then we define the vector b_2 from the center of gelpad to the component's top edge center found in the previous section in gripper frame as $b_2^G = [d_x, d_y, 0]^T$ where d_x and d_y are distances in gripper's x and y-axes from the center of the gelpad to the component's top edge center respectively. Note that to find d_x and d_y, a conversion from pixel distance to metric unit distance can be computed using the black dot markers' known distance in the images (2 mm in this case). The sum of the two vectors then yield translation from the gripper frame origin to the component's top edge frame origin. Based on the methods presented in the previous section, we can find the y-axis rotation angle α as defined in Fig. 3B. Based on these parameters, we can define the rotation matrix from gripper to component as such:

$$R_C^G(\alpha) = \begin{bmatrix} cos(\alpha) & 0 & sin(\alpha) \\ 0 & 1 & 0 \\ -sin(\alpha) & 0 & cos(\alpha) \end{bmatrix}$$

V. Case Study 1: Capacitor Assembly

We first approached the capacitor assembly task. The capacitor used in this task has three pins that must be matched to the three PCB through-holes (Fig. 3 A). Then, GelSight is tasked with two responsibilities: 1. in-hand pose estimation to update the planning reference frame and allow for isolated capacitor rotations without translation 2. recognition of task completion where capacitor is in correct orientaiton to fit into the holes. As noted previously, both subtasks are difficult to achieve reliably solely with external vision due to occlusion.

Based on the stated focus of this study, some assumptions were made. First, we assumed we know the position but not the orientation of the PCB through-holes. Next, we assumed we approximately know the position of the capacitor. Both of these assumptions were made based on the expectation that external vision can sufficiently define these parameters in realistic industrial operations.

In the task, we first let the robot arm and gripper to grasp the capacitor from the table. Then, we perturbed both the position and the orientation of the capacitor within the gripper. Based on the in-hand pose estimnation of the capacitor with GelSight, the robot arm adjusts the capacitor orientation and position in the world frame to match those of the through-holes (Fig. 4B). The capacitor was pushed up against the PCB until small displacement of the markers was detected. Then, the capacitor was rotated about its central axis until we can detect that the holes matched with the capacitor pins.

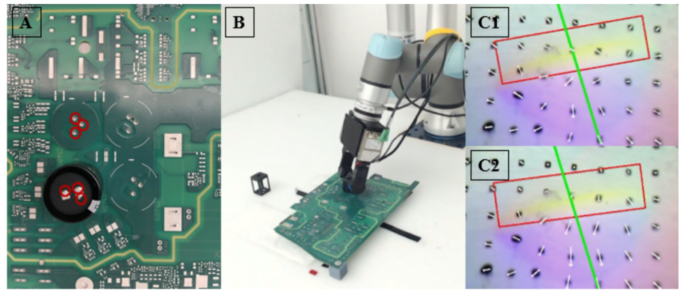

Fig. 4. A: three PCB through-holes and three capacitor pins that must be matched up for assembly through rotation B: robot arm rotating the capacitor against the PCB C1: GelSight view of the capacitor with hole mismatch C2: GelSight view of the capacitor with hole match

Primarily, we looked at GelSight's markers that provide insight into the applied distributed forces on the object. White 2D vectors in Fig. 4C1 provide both direction and magnitude of the marker movements on the gelpad. When the through-holes of the PCB match with the capacitor's pins, the small applied force on the component will push the pins in to the holes. Then, with any further rotation, the capacitor will resist the applied torque and GelSight markers will be significantly displaced. Unweighted mean average of all of the marker displacement vector magnitudes were taken. We heuristically set the threshold to 20% increase in average magnitude for the robot arm to stop rotating and drop the capacitor into the hole. Out of 10 trials performed, 6 trials were successful with arbitrary perturbation. All failures occured when the change in the marker displacement was not sufficiently large to indicate completion of the task. Much of the variances among trials that resulted in these failures can be attributed to nonexistent feedback on the location of PCB in between trials. As noted earlier, we assume that the PCB location is known and it was not updated in between trials. However, because of the interaction with the component, the PCB position may shift slightly after a trial, leading to misalignment and failures. Another factor may be the method which we used find the threshold. We took an unweighted mean average of all displacement vectors' magnitudes. More robust methods using weighted averages are discussed in Future Works section. We expect that with the inclusion of external vision to adjust for changing location of the PCB and a more robust weighted average threshold task completion criteria, the success rate will be drastically improved.

VI. Case Study 2: PCB Connector Assembly

We also explored translational adjustments for electronics assembly using GelSight with PCB connector components. The connector is a cuboid with two guiding plastic pins that correspond to two guiding through-holes on the PCB. In this case, the orientation is not a concern as the rectangular cross-sections allow us to fully define the orientation of the component when it is grasped. However, we do not assume we know the exact position of the through-holes on the PCB.

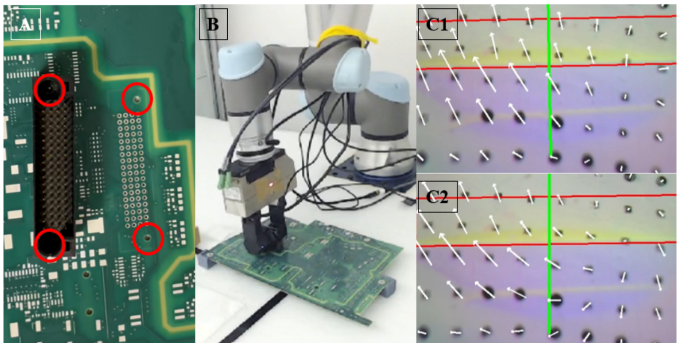

Fig. 5. A: two protruding pins on the connectors that must be matched up with the two guiding through-holes in the PCB through translation. B: robot arm performing raster-pattern search for the hole. C1: GelSight view of the connector pushed up against the PCB with hole mismatch. C2: GelSight view of the connector pushed up against the PCB with hole match. We can note changes in both the direction and the magnitude of the white 2D marker displacement vectors.

We performed similar early procedures as the previous case study with grasping and pose perturbation. The connector was then tilted about its y-axis (axis that is in-plane to the GelSight view) by 10 degrees. This was to force only one of component's pins to be in contact with the PCB and simplify the hole search. The pin was then pushed against the PCB board to achieve significant gelpad marker displacement. We may note in Fig. 5C1 that GelSight is able to clearly highlight the moment being applied on the connector where to the left of the central axis (green line), the the marker displacement vectors form a circular pattern. The pin was then dragged across the region of interest in a raster-search pattern until the pin fell into the hole on the PCB hole.

Similar to the previous case study application with the capacitor, we look at the displacement vectors of the markers through the search. We take an unweighted mean average of the displacement vector magnitudes and similar to before, we heuristically set the conditional threshold to 12% change in magnitude. In this task, the robot arm applied greater force on the component initially and the controller looked at the decrease in magnitude to indicate that the pin has fallen into the hole. We may note that in Fig. 5C2, there is a significant overall decrease in the magnitude of the displacement vectors. Once a significant decrease in the magnitude was detected, the arm applied rotation and translation to the connector to match the orientation of the PCB board and push the other pin into the hole. Out of 9 trials performed, 4 was successful. Similar to the previous task, we expect the success rates to be significantly improved with the addition of external vision and adjustment to the methods we use to set threshold for task completion.

VII. Future Works

In the presented case study tasks, thresholding to isolate the top edge sufficiently defined the pose of the components and heuristically set change threshold detected completion of tasks. For reliable integration of GelSight in industrial applications, however, significant improvements can be made in the generalizability of the presented methods. First, in-hand pose estimation of the assembly components can benefit from

TABLE I
Task Descriptions and Success Rate

Task	Adjustment Mode	Description	Success Rate
Capacitor Assembly	Pure Rotation	Rotate the capacitor until it falls into through-holes	6/10
Connector Assembly	Pure Translation	Raster-pattern search for corresponding hole	4/9

incorporating the full GelSight representation of contact. A convolutional residual network trained on high-fidelity GelSight simulations has been implemented for the application of estimating GelSight contact region and height map [12]. Similar approaches may enable a more robust pose estimation. Particularly, relaxing assumption (1) by also estimating in-plane movement of the component may provide valuable insight in both in-hand pose-estimation and feedback-based manipulation.

Modifications can be made to the presented methods for tactile feedback during assembly tasks to significantly improve task success rates. For example, in assessing the thresholds of the marker displacements, we simply took the unweighted average mean of the magnitudes in this work, ignoring localized marker displacements and the vector directions. However, as it can be seen with Fig. 5C, both direction and magnitude of the displacement vectors vary significantly based on their location on the gelpad with respect to component contact region. Therefore, incorporating such information by dynamically varying the vectors' weights based on their location may increase the reliability.

The case studies underscored the types of useful information GelSight can provide in assembly tasks. However, how this high-dimensional tactile information should be processed and utilized for optimal performance are still open questions. Real-time pose perturbation rejection, fusion with external vision, and robust component fit search are particularly promising directions for future work to enable reliable implementation in industrial settings.

VIII. Conclusion

In this work, in-hand pose estimation was explored for the application of automated electronics assembly. By thresholding, contacted component's top edge was isolated from which the pose of the component was estimated. Furthermore, case studies with two geometrically different components were performed to showcase the distinct contributions that GelSight can make in assembly tasks. Particularly, GelSight was found to be useful in sensing of PCB-component interactions for pose adjustment and component insertion into the appropriate holes. Overall, the experiments and the results revealed multiple promising directions of future work that can build on relaxing the constraints imposed in this work for reliable applications in industrial settings.

References

[1] K. Zhang, Y. Li, Y. Guan, L. He, Y. Gan, and Z. Dong, "A Vision Detection System for Odd-Form Components," in *2018 IEEE International*

Conference on Robotics and Biomimetics (ROBIO), pp. 120–125, Dec. 2018. Journal Abbreviation: 2018 IEEE International Conference on Robotics and Biomimetics (ROBIO).

[2] S. P. Chan, “A disturbance observer for robot manipulators with application to electronic components assembly,” *IEEE Transactions on Industrial Electronics*, vol. 42, pp. 487–493, Oct. 1995.

[3] G. Du, K. Wang, S. Lian, and K. Zhao, “Vision-based robotic grasping from object localization, object pose estimation to grasp estimation for parallel grippers: a review,” *Artificial Intelligence Review*, vol. 54, pp. 1677–1734, Mar. 2021.

[4] R. Li, R. Platt, W. Yuan, A. ten Pas, N. Roscup, M. A. Srinivasan, and E. Adelson, “Localization and manipulation of small parts using GelSight tactile sensing,” in *2014 IEEE/RSJ International Conference on Intelligent Robots and Systems*, pp. 3988–3993, Sept. 2014. Journal Abbreviation: 2014 IEEE/RSJ International Conference on Intelligent Robots and Systems.

[5] J. Bimbo, S. Luo, K. Althoefer, and H. Liu, “In-Hand Object Pose Estimation Using Covariance-Based Tactile To Geometry Matching,” *IEEE Robotics and Automation Letters*, vol. 1, pp. 570–577, Jan. 2016.

[6] M. Bauza, E. Valls, B. Lim, T. Sechopoulos, and A. Rodriguez, “Tactile Object Pose Estimation from the First Touch with Geometric Contact Rendering,” 2020. _eprint: 2012.05205.

[7] Y. She, S. Wang, S. Dong, N. Sunil, A. Rodriguez, and E. Adelson, “Cable manipulation with a tactile-reactive gripper,” *The International Journal of Robotics Research*, p. 02783649211027233, Aug. 2021. Publisher: SAGE Publications Ltd STM.

[8] W. Yuan, R. Li, M. A. Srinivasan, and E. H. Adelson, “Measurement of shear and slip with a GelSight tactile sensor,” in *2015 IEEE International Conference on Robotics and Automation (ICRA)*, pp. 304–311, May 2015. Journal Abbreviation: 2015 IEEE International Conference on Robotics and Automation (ICRA).

[9] S. Dong, W. Yuan, and E. Adelson, “Improved GelSight Tactile Sensor for Measuring Geometry and Slip,” *arXiv e-prints*, p. arXiv:1708.00922, Aug. 2017. _eprint: 1708.00922.

[10] W. Yuan, S. Dong, and E. H. Adelson, “GelSight: High-Resolution Robot Tactile Sensors for Estimating Geometry and Force,” *Sensors*, vol. 17, no. 12, 2017.

[11] D. Eberly, “Minimum-area rectangle containing a set of points,” *Geometric Tools, LLC*, 2015.

[12] S. Suresh, Z. Si, J. G. Mangelson, W. Yuan, and M. Kaess, “Efficient shape mapping through dense touch and vision,” 2021. arXiv.