

# Contact-Based Pose Estimation of Workpieces for Robotic Setups

Yitaek Kim, Aljaz Kramberger, Anders Glent Buch, and Christoffer Sloth

**Abstract**—This paper presents a method for contact-based pose estimation of workpieces using a collaborative robot. The proposed pose estimation exploits positions and surface normal vectors along an arbitrary path on an object with known geometry, where surface normal vectors are estimated based on contact forces measured by the robot. When data is only available along a single path, it is difficult to find initial correspondences between source data (recorded points and normal vectors) and target data (CAD of an object); hence, a novel weighted incremental spatial search approach for generating correspondences based on point pair features is proposed. Subsequently, robust pose estimation is employed to reduce the effect of erroneous correspondences. The proposed pose estimation is verified in simulation on three paths on two objects and with different levels of noise on the source data to quantify the robustness of the algorithm. Finally, the method is experimentally validated to provide an average pose rotation and translation accuracy of  $0.55^\circ$  and  $0.51$  mm, respectively, when using the robust estimation cost function Geman-McClure.

## I. INTRODUCTION

High-mix low-volume robotic manufacturing has recently been getting attention since traditional robotic automation is hardly able to manage the high product variation in production processes [1]. High-volume production often relies on accurate fixtures to eliminate the kinematic uncertainty, which makes it possible to complete, e.g. assembly tasks in a fast and reliable way. Accurate fixtures are often too expensive for high-mix low-volume production; thus, workpiece calibration must be performed in a different manner. The workpiece pose estimation can be done with point set registration based on visual data (point cloud) acquired utilizing calibrated 3D scanning devices or with data obtained from contact-based calibration methods, which relate a point on an object with a point in a CAD file.

Point Set Registration (PSR) is the procedure of obtaining a rigid transformation between two sets of points, and has been used in various vision, and pattern recognition applications [2]. Correspondence-based point set registration methods follow the procedure

- 1) Collect source data  $\mathbf{p}_i \in \mathbb{R}^3$  for  $i = 1, \dots, N$ .
- 2) Given target data  $\mathbf{q}_j \in \mathbb{R}^3$  for  $j = 1, \dots, M$ . Find correspondence  $\gamma : \{1, \dots, N\} \rightarrow \{1, \dots, M\}$ .

All authors are with The Maersk Mc-Kinney Moller Institute, University of Southern Denmark, Denmark {yik, alk, anbu, chsl}@mmmi.sdu.dk

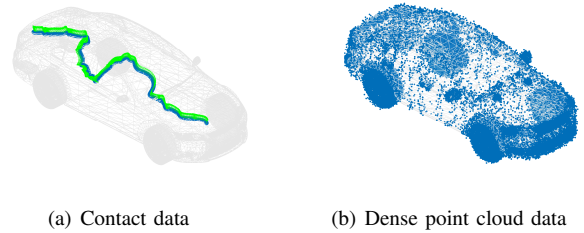


Fig. 1: Types of source data. (a) has contact points (blue) and surface normals (green), on the other hand, (b) has spatial points (blue).

- 3) Find rigid transformation  $(R, t) \in SE(3)$  by solving

$$\min_{(R,t) \in SE(3)} \sum_{i=1}^N d(\mathbf{p}_i, R\mathbf{q}_{\gamma(i)} + t) \quad (1)$$

where  $d$  is a distance function.

When the correspondence  $\gamma$  is known then (1) can be easily solved with  $d$  chosen as the Euclidean distance. This is the approach used in most contact-based calibration methods [3], [4]. However, when the correspondence  $\gamma$  has erroneous pairs, then outliers are introduced into the least squares problem, which greatly affects the accuracy of the pose estimation. To minimize the effect of outliers, heuristic estimation methods such as RANSAC [5] are being used, but the computation time of RANSAC grows exponentially with the outliers ratio [6]. Alternatively, optimization-based estimation methods such as ADAPT [7] can be used; ADAPT runs with an adaptive inlier threshold to reject the outliers. To decrease computational costs and increase robustness even further, Graduated Non-Convexity (GNC) [8] was proposed in [9] as a robust estimation method for spatial perception applications. The method enables robust estimation in case of a high outlier ratio and shows better performance than RANSAC and ADAPT.

All of the above methods rely on given correspondences between source and target data. One potential method to find the putative correspondences is the Point Pair Feature (PPF) which extracts local surface features based on the relative position and normal vector of given points [10]. Most feature matching methods based on PPF, such as Fast Point Feature Histogram (FPFH) [11] and Point Pair Feature Histogram (PPFH) [12] are likely to contain a high outlier ratio in the putative correspondences [6]; thus, the subsequent optimization problem to find the rigid transformation must be robust towards outliers. It should be noted that the available data in the considered pose estimation problem is considerably

different from the standard vision problem, as only partial data along a contact line is available. Fig. 1 illustrates a dense point cloud and a path including points and surface normal vectors which are considered in this paper. The addition of surface normals is expected to improve the pose estimation, as the combination with a point cloud and a surface normal generally performs better than using only a point cloud in the pose and camera calibration process [13][14][15]. Note that the surface normal cannot be computed based on only positional information along the path.

In this paper, we propose a novel approach for contact-based pose estimation, where initial correspondences are not known and source data is represented as points and normal vectors along a path on the object. Since the data corresponds to a partial contact line, we show that it is not possible to use conventional feature matching methods such as FPFH and PPFH for determining the correspondences. Therefore, a new PPF-based weighted incremental spatial search method is proposed for finding correspondences along a given path. In addition, the robust pose estimation problem in [16] is slightly modified to use the surface normal correspondences. We show that the proposed pose estimation applies to the considered problem, and is robust towards noise in the source data.

The paper is structured as follows. Section II formulates the problem addressed in the paper. The proposed method is presented in Section III and explains how the correspondences are obtained and subsequently how the robust pose estimation is solved. The method is verified in simulation in Section IV and experimental results are provided in Section V. Lastly, conclusions are provided in Section VI.

## II. PROBLEM FORMULATION

To estimate the 6D pose of a given workpiece with a robot, it is generally important to know multiple points on the surface of the workpiece and preferably the surface normals at each point. To this end, we need to estimate the points and the surface normals directly while a robot moves smoothly on the surface of the workpiece, keeping contact as shown in Fig. 2. However, the estimated contact information as source data is much more partial than the dense point cloud provided by the vision-based approach, which leads to difficulty extracting features and matching them between source data and target data to find putative correspondences.

In the subsequent step the optimal rigid transformation w.r.t., a robot base on the putative correspondences, has to be obtained, which can be accomplished by using the existing non-minimal solvers [16]. However, since the performance of the non-minimal solvers depends on how many noises and outliers are inherent in the data, we need to take advantage of the robust pose estimation method. The following section proposes a solution to the presented problem and provides an accurate and robust method for pose estimation when given only contact data and a CAD model.

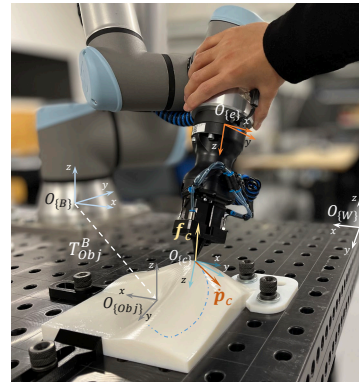


Fig. 2: Data collection with a robot.  $f_c$  is a contact force, and  $\dot{p}_c$  is a contact velocity, and  $T_{obj}^B$  is the transformation between robot base and an object. Five frames are defined:  $O_{\{B\}}$ - robot base,  $O_{\{e\}}$ - end-effector,  $O_{\{W\}}$ - world,  $O_{\{c\}}$ - contact, and  $O_{\{obj\}}$ - object.

## III. PROPOSED METHOD

The proposed pose estimation method consists of the following three steps

- 1) **Data Collection:** The collected data is a sequence of pairs  $(p_i, n_i)$ , where  $p_i \in \mathbb{R}^3$  is the measured point, and  $n_i \in \mathbb{R}^3$  is the estimated surface normal vector for  $i = 1, \dots, N$ , where  $N$  is the number of samples. The noise on the data is assumed to be given by a zero mean Gaussian process with known covariance.
- 2) **Correspondence Generation:** Find correspondences between source data  $(p_i, n_i)$  and target data  $(q_j, m_j)$ , i.e. a function  $\gamma : \{1, \dots, N\} \rightarrow \{1, \dots, M\}$ . The correspondences are found using a novel method explained in Section III-B that uses point pair feature [10].
- 3) **Pose Estimation:** Find a homogeneous transformation  $(R, t) \in SE(3)$  from target to source data. The transformation is found using the robust pose estimation method proposed by [9], where the surface normal is included in the distance measure as explained in Section III-C.

### A. Data Collection

As shown in Fig. 2, a pointy tool attached to a robot is guided along an object (blue dashed line). Recorded tool position, velocity, and contact wrench are used for estimating surface normal vectors along the path. The surface normal is estimated using the integral adaptive law presented in [17].

### B. Correspondence Generation

The Point Pair Feature (PPF) method describes the object features formulated with all pairs that include the relative information of points and their normal vectors [10]. A PPF is defined as:

$$\mathbf{F}(p_1, p_2) = (v_d, \angle(n_1, v_d), \angle(n_2, v_d), \angle(n_1, n_2)) \quad (2)$$

where  $p_1, p_2 \in \mathbb{R}^3$  are points on the surface,  $n_1, n_2 \in \mathbb{R}^3$  are surface normals at each point,  $v_d = \|p_2 - p_1\|_2$  is the

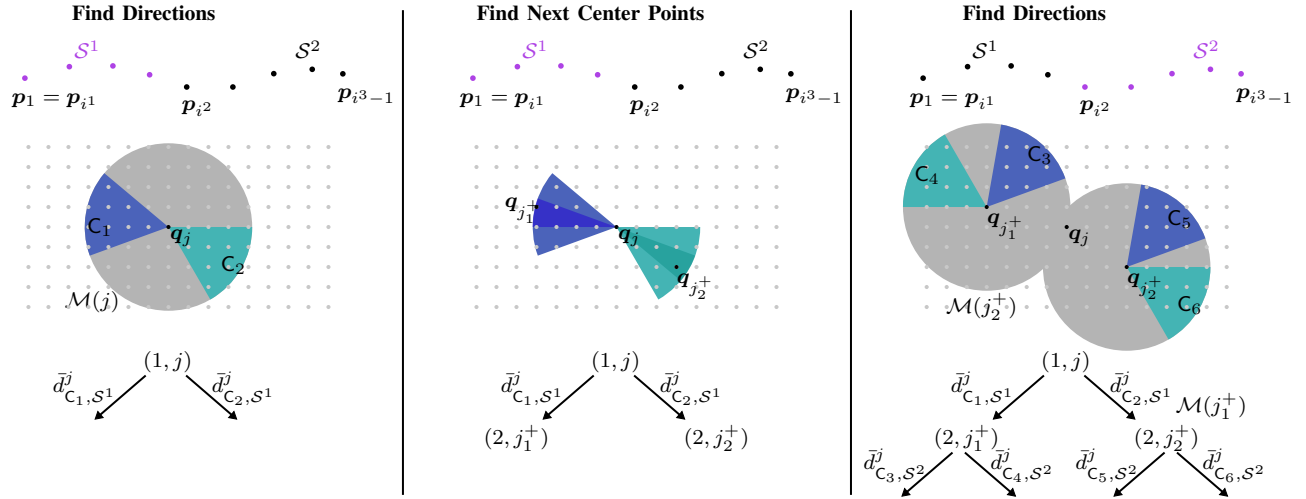


Fig. 3: Illustration of the algorithm for the three steps of the spatial search method. First (left) the two best clusters are identified based on average PPF distance, and a weighted directed graph is generated relating the first segment of source data to the two clusters. Second, the next center points are identified within the two clusters. Third, the procedure is repeated by identifying the two best clusters for the two center points.

distance, and  $\angle(\mathbf{a}, \mathbf{b})$  is the unsigned angle between vectors  $\mathbf{a}$  and  $\mathbf{b}$ . Each element of a PPF should be normalized to be in the same order of magnitude. For calculation efficiency, we define an  $r$ -neighborhood around each point in the target data as:

$$\mathcal{M}(j) = \{k \in \{1, \dots, M\} \setminus \{j\} \mid \|\mathbf{q}_j - \mathbf{q}_k\|_2 \leq r\} \quad (3)$$

where  $\mathbf{q}_j$  is said to be the center point of the neighborhood  $\mathcal{M}(j)$ . In addition, we partition the source data into  $S$  segments; we assume that the source data is ordered, i.e.  $\mathbf{p}_1$  is the start of the path and  $\mathbf{p}_N$  is the end of the path. Let index be  $i^1 = 1$  and for  $s = 2, \dots, S$

$$i^s = \min_{k \in \{i^{s-1}, \dots, N\}} k \text{ s.t. } \|\mathbf{p}_k - \mathbf{p}_{i^{s-1}}\|_2 > r. \quad (4)$$

The source data segments are defined as

$$\mathcal{S}^s = \{i^s, \dots, i^{s+1} - 1\}, s = 1, \dots, S \quad (5)$$

where  $i^{S+1} = N + 1$ .

For each neighborhood of the target data, we introduce a cluster  $\mathcal{C}_h^o(j)$  defined from a neighboring point  $\mathbf{q}_h$  as:

$$\mathcal{C}_h^o(j) = \{k \in \mathcal{M}(j) \mid \angle(\mathbf{q}_j - \mathbf{q}_h, \mathbf{q}_j - \mathbf{q}_k) < \theta\}$$

where  $\theta > 0$  is a threshold on the angle. A feature map with respect to  $\mathbf{q}_j$  is defined for a cluster  $\mathcal{C}$  as

$$\mathcal{F}_{\mathcal{C}}(j) = \bigcup_{k \in \mathcal{C}} \mathbf{F}(\mathbf{q}_j, \mathbf{q}_k). \quad (6)$$

Finally, we define the distance between a point pair feature  $F$  and a feature map  $\mathcal{F}$  as

$$d(F, \mathcal{F}) = \min_{\tilde{F} \in \mathcal{F}} \|\mathbf{F} - \tilde{\mathbf{F}}\|_2. \quad (7)$$

In our case, the source data only consists of points and surface normals on a contact line; hence, the PPF method is more sensitive to noise than when having a dense point cloud.

Therefore, it is required to look at the trends of the data. To achieve this, we propose a PPF-based weighted incremental spatial search method for finding putative correspondences. The proposed method searches for the most similar sets of points in the target and source data by leveraging both sequential information of the contact line and its features. Thus, we generate  $M$  directed weighted graphs, where each graph relates source and target data from one initial point in the target data  $\mathbf{q}_j$ ; the shortest path through all  $M$  graphs gives a correspondence between segments of source data and clusters of target data. The generation of the weighted graphs is computationally demanding; thus, the beam search method [18][19] is used when generating the graph. The beam search method only adds the best  $k$  candidates to the graph at each iteration. An overview of the algorithm for finding correspondences is given in the following and supported by Fig. 3.

- 1) **Find Directions:** The direction of the source data in segment  $\mathcal{S}^s$  with respect to the neighborhood  $\mathcal{M}(\bar{j})$  in the target data is found by a) filtering to minimize noise, and b) identifying the cluster with the best fit of PPF. The filtering is accomplished by finding the average distance between point pair features for all  $\mathbf{p} \in \mathcal{S}^s$  with respect to  $\mathbf{p}_1$  and the feature map  $\mathcal{F}_{\mathcal{C}}(j)$ , i.e.,

$$\bar{d}_{\mathcal{C}, \mathcal{S}^s}^j = \frac{1}{\#\mathcal{S}^s} \sum_{k \in \mathcal{S}^s} d(F(\mathbf{p}_1, \mathbf{p}_k), \mathcal{F}_{\mathcal{C}}(j)) \quad (8)$$

where  $\#\mathcal{S}^s$  denotes the cardinality of  $\mathcal{S}^s$ . The best direction is found from the above average distance as

$$h^*(s, \bar{j}) = \arg \min_{h \in \mathcal{M}(\bar{j})} \bar{d}_{\mathcal{C}_h^o(\bar{j}), \mathcal{S}^s}^j. \quad (9)$$

This is illustrated in Fig. 3 (left) for the first segment and in Fig. 3 (right) for the second segment. In the illustration, we use the two best clusters to complete a

beam search. The distance (8) is used as a weight in the weighted graph, see bottom of Fig. 3.

- 2) **Find Next Center Points:** The next center point  $\mathbf{q}_{j^+}$  in the target data is chosen to have the maximal distance to the previous center point  $\mathbf{q}_{\bar{j}}$  within cluster  $C_{h^*(s,\bar{j})}^\epsilon(j)$  with  $\epsilon < \theta$ , i.e.,

$$j^+ = \arg \max_{k \in C_{h^*(s,\bar{j})}^\epsilon(j)} \|\mathbf{q}_{\bar{j}} - \mathbf{q}_k\|_2.$$

The center point  $\mathbf{q}_{j^+}$  is associated to the next segment of the source data  $\mathcal{S}^{s+1}$ . This is illustrated in Fig. 3 (middle) where the center points are identified within the darker cones with a smaller angle.

- 3) **Find Correspondences:** After iterating between 1) and 2) until the last segment of the source data, a weighted graph starting at each point in the target data is available. We identify the best starting point  $\mathbf{q}_{j^*}$  by finding the graph with the shortest path. The weighted graph relates segments and clusters, i.e. it provides a sequence of pairs  $(\mathcal{S}^s, \mathcal{C}_s)_{s=1,\dots,N}$ . The correspondence  $\gamma$  is generated from the sequence of segments and clusters as:

$$\gamma(i) = \arg \min_{k \in \mathcal{C}_s} d(F(p_1, p_i), F(p_{j^*}, p_k)) \quad (10)$$

where  $i \in \mathcal{S}^s$  and  $s = 1, \dots, S$ .

### C. Robust Pose Estimation

To obtain an optimal transformation, we need to formulate an optimization problem based on the correspondences given by  $\gamma$ . However, the correspondences are likely to have outliers which negatively influence the accuracy of the pose estimation. Therefore, we eliminate outliers in  $\gamma$  by robust optimization. According to [9], one way to eliminate the effect of outliers is to apply robust cost function  $\rho(\cdot)$  such as Geman-McClure (GM) and Truncated Least Squares (TLS) to the optimization problem (1) as

$$\min_{(R,t) \in SE(3)} \sum_{i=1}^N \rho(\|r(\mathbf{p}_i, \mathbf{q}_{\gamma(i)})\|_2) \quad (11)$$

where  $r(\cdot)$  is the residual error between source data and transformed target data. We use the method proposed in [9] for robust pose estimation that combines Black-Rangarajan Duality [20] and Graduated Non-Convexity (GNC) [21]. In addition, we reformulate the least square optimization problem proposed by [16] in order to exploit our estimated surface normal vectors.

1) *Black-Rangarajan Duality:* This theory is utilized to reconcile between robust estimation and outlier process. According to [20], the problem (11) can be replaced with:

$$\min_{(R,t) \in SE(3)} \sum_{i=1}^N \left[ W_i \|r(\mathbf{p}_i, \mathbf{q}_{\gamma(i)})\|_2^2 + \Phi_\rho(W_i) \right] \quad (12)$$

where  $W_i$  is a set of weights and  $\Phi_\rho(W_i)$  is an outlier process depending on  $W_i$  and  $\rho(\cdot)$ .

2) *Rigid Transform Optimization with GNC:* To solve problem (12), we use Graduated Non-Convexity (GNC) [21], where an adaptive parameter  $\mu$  is gradually increased or decreased to recover the robust cost functions  $\rho(\cdot)$  from an initially convex surrogate cost function  $\rho_\mu(\cdot)$ . The parameter  $\mu$  is changed in an outer loop while a transformation  $(R, t)$  is found in an inner loop for the fixed  $\mu$  by alternating between the following two optimization problems [9]:

- 1) **Transformation optimization**

$$(R, t)^{(m)} = \arg \min_{(R,t) \in SE(3)} \sum_{i=1}^N W_i^{(m-1)} \|r(\mathbf{p}_i, \mathbf{q}_{\gamma(i)})\|_2^2 \quad (13)$$

- 2) **Weight optimization**

$$W_i^{(m)} = \arg \min_{W_i \in [0,1]} \sum_{i=1}^N \left[ W_i \|r(\mathbf{p}_i, \mathbf{q}_{\gamma(i)})\|_2^2 + \Phi_{\rho_\mu}(W_i) \right] \quad (14)$$

where  $m \in \mathbb{N}$  is iteration number. According to [9], Weight optimization (14) is solved using a closed-form solution that depends on the used robust function  $\rho_\mu$ . In this paper, we use GM and TLS. In the optimization problem (13), we use the solver proposed by [16] to find  $R$  and  $t$ .

The residual error in (13) and (14) is defined using *point-to-plane* and *normal-to-normal* distances:

- 1) *point-to-plane* (used in e.g. [16], [22])

$$D_{\text{pl}}^i := \|\mathbf{m}_{\gamma(i)}^\top (\mathbf{p}_i - (R\mathbf{q}_{\gamma(i)} + t))\|_2^2 \quad (15)$$

- 2) *normal-to-normal* (introduced in this paper)

$$D_{\text{n}}^i := \|(\mathbf{n}_i - R\mathbf{m}_{\gamma(i)})\|_2^2 \quad (16)$$

When using the distances defined in (15) and (16) in the optimization problems (13) and (14), we replace  $\|r(\mathbf{p}_i, \mathbf{q}_{\gamma(i)})\|_2^2$  with  $\alpha D_{\text{pl}}^i + \bar{\alpha} D_{\text{n}}^i$ , where the two weights  $\alpha, \bar{\alpha} \geq 0$  should be scaled according to the noise covariance on the source data.

It is seen that optimization problem (13) is quadratic in the unknown variables; thus, we follow [16] and reformulate (13) as the following constrained quadratic program

$$(R^*, t^*) = \arg \min_{(R,t) \in SE(3)} \tilde{\tau}^\top \tilde{U} \tilde{\tau}, \quad \tilde{\tau} = \begin{bmatrix} \text{vec}(R) \\ t \\ 1 \end{bmatrix} \quad (17)$$

where  $\tilde{U}$  is generated from the source data, target data, and weights  $\alpha, \bar{\alpha}$ .

To solve (17), we follow [16] and approximate the constraint  $R \in SO(3)$  by quadratic constraints, which leads to a Quadratically Constrained Quadratic Program (QCQP) for finding the rotation matrix  $R$  (the translation  $t$  is subsequently computed based on  $R$ ). The dual problem to the QCQP is then solved to find  $R$  using CVX [23]. The solution to the dual problem is empirically shown to obtain the global optimal solution of the primal problem, see [16] for details.

#### IV. SIMULATION

In this section, we verified the proposed method in simulation. For verification purposes, we used two types of mesh objects; an L-shaped membrane and a car object. The L-shaped object is defined with a mathematical equation, on the other hand, the car mesh has complex shapes provided by the PASCAL+ dataset [24]. For each mesh model, we extract a point cloud as target data from the model and provide an arbitrary contact path including surface normals as source data. We firstly show that existing methods cannot find the putative correspondences properly. Subsequently, we conduct several simulations to verify the proposed method in terms of sensitivity toward noise, paths, and geometry.

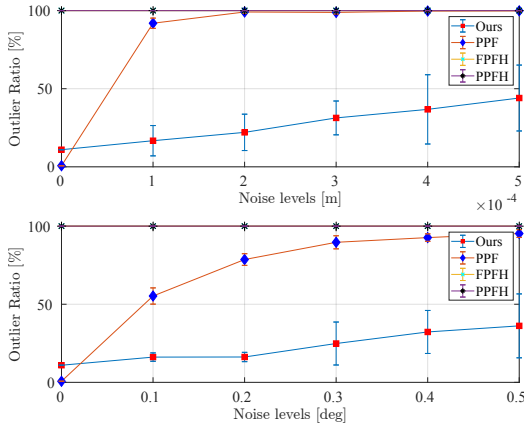


Fig. 4: Comparison with the existing methods in terms of outlier ratio as a function of noise variance.

Fig. 4 shows that the existing methods: PPF [10], FPFH [11], and PPFH [12] cannot provide the putative correspondences when only data along a contact line is provided; therefore, they cannot be used for our problem. This is because we use only partial data in the matching, i.e. data from a line of the object is available; thus, other methods cannot complete the matching. On the other hand, our proposed method provides putative correspondences even if noise levels are increased. The performance of PPF is better than the proposed method when there is no noise, since PPF visits all points in a neighborhood, whereas our proposed method uses threshold parameters to define the best cluster in each neighborhood. However, it is clearly shown that the outlier ratio of PPF is highly increased when noise is added to the source data, which indicates that it cannot be used in the real environment.

To statistically analyze the robustness to noise, we conduct 50 simulations for 20 noise levels in translation and surface normal. The standard deviations of noise on translation and surface normal are increased in steps of 0.1 mm and  $0.1^\circ$ , respectively. As shown in Fig. 5, the proposed method is robust toward noise on translation and surface normal and provides correspondences with a low outlier ratio compared to other methods in the case of noisy source data.

We generate various paths on the objects as shown in Fig. 6 to assess if the performance of the method, depends on

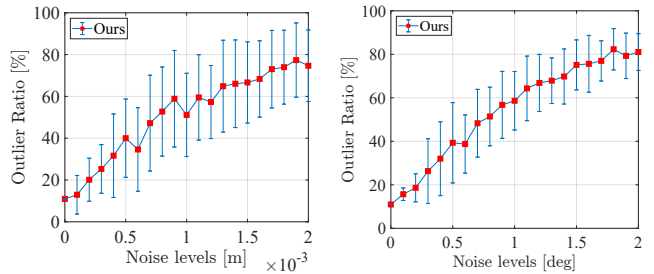


Fig. 5: Performance of finding correspondences depending on the gradual increase of noise levels.

the specific path or the specific object geometry. As shown in TABLE I, the results for both models with three paths indicate that GNC-TLS and GNC-GM have overall good performances in all evaluations. The accuracy of ADAPT is similar to the GNC methods, but it has a longer computation time. The results of the rotation error in the second path show that RANSAC obtains the best performance, but the translation error is very large compared to other methods. In conclusion, we verified that the proposed method estimates the pose of a given object when provided with various contact paths. Note that we normalize all errors since the two mesh models have different mesh resolutions and sizes.

#### V. EXPERIMENTAL EVALUATION

We evaluated the proposed method on a real robotic setup including a Universal Robot UR10e. In order to perform the experiments, we designed a pointy tip attachment that enables the human operator to record exact in-contact motion, via kinesthetic teaching. Furthermore, in order to validate the performance of our proposed approach, we carried out a table-to-object calibration following the procedures proposed in [25]. Based on the calibration, we defined the transformation  $T_{obj}^B$  as the ground truth pose of the test L-shaped object shown in Fig. 2.

In the first phase of the experiment, we recorded the contact path with kinesthetic guidance. The operator, physically grabbed the robot by the tool and with the help of a compliance controller, demonstrated the positional path while maintaining contact between the tool and the object. Afterward, we replayed the demonstrated path in order to eliminate the human effect on the robot and recorded the data such as the contact point, wrench, and velocity. Based on the recorded contact data, we estimated the surface normals, which represent the source data with contact points. The recorded source data and the target data extracted from the CAD file were used to find correspondences and estimate the transformation,  $T_{obj}^B$ . In total, we performed three experiments to validate our proposed method. TABLE II shows the performance of our proposed method. It can be seen that the proposed method provides putative correspondences in the real environment with an average outlier ratio of 70.92%. The outlier ratio is influenced by various factors such as the error of the surface normal estimation and tool calibration. However, our outlier ratio is small compared to

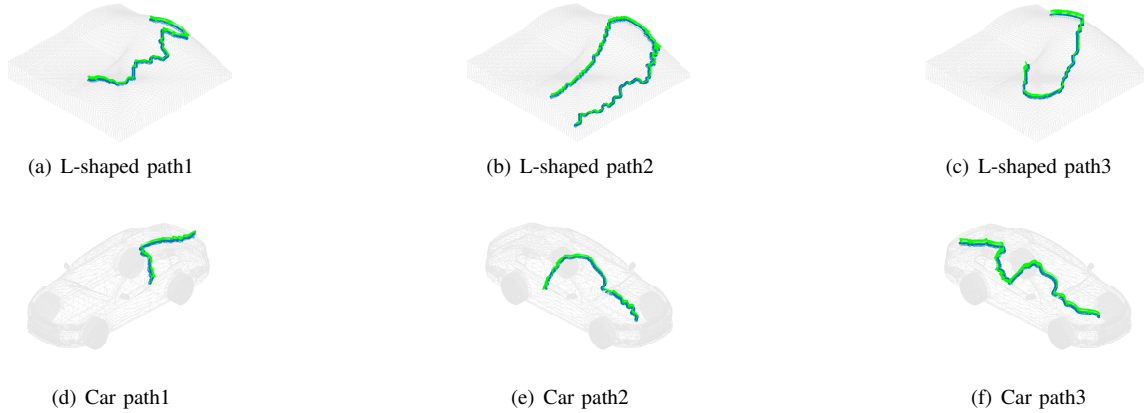


Fig. 6: Different contact paths in simulation. Green vectors are surface normal vectors, and blue dots are contact points.

	Methods	Rotation Error				Translation Error				Execution Time [s]			
		Path 1	Path 2	Path 3	Average	Path 1	Path 2	Path 3	Average	Path 1	Path 2	Path 3	Average
L-Shaped	GNC-TLS	<b>0.383</b>	2.209	0.486	0.631	0.052	<b>0.014</b>	0.045	0.037	24.82	40.7	31.19	32.24
	GNC-GM	0.682	2.241	<b>0.459</b>	0.704	<b>0.039</b>	0.027	<b>0.02</b>	<b>0.028</b>	<b>16.26</b>	<b>25.68</b>	<b>18.11</b>	<b>20</b>
	ADAPT	0.403	2.178	0.481	<b>0.63</b>	0.047	0.022	0.041	0.037	219.8	478.2	276.6	324.9
	RANSAC	1.0	<b>1.0</b>	1.0	1.0	1.0	1.0	1.0	1.0	35.78	67.39	37.93	47.03
Car	GNC-TLS	0.022	<b>0.021</b>	<b>0.103</b>	<b>0.032</b>	0.064	0.053	0.036	0.054	34.98	19.11	34.78	29.64
	GNC-GM	<b>0.019</b>	0.038	0.123	0.041	<b>0.05</b>	<b>0.033</b>	<b>0.028</b>	<b>0.039</b>	<b>25.44</b>	<b>15.08</b>	<b>26.46</b>	<b>22.33</b>
	ADAPT	0.022	0.022	0.146	0.038	0.078	0.059	0.037	0.062	211.8	151.5	309.9	224.4
	RANSAC	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	28.52	29.38	36.86	31.59

TABLE I: Simulation performance of contact-based pose estimation with robust optimization. All results except for execution time are normalized relative to RANSAC. All executions have the same noise levels; standard deviations of rotation and translation are  $2^\circ$  and 1mm, respectively. The average outlier ratio is 65.49% for all paths.

	Methods	Rotation Error [deg]				Translation Error [mm]				Execution Time [s]			
		Path 1	Path 2	Path 3	Average	Path 1	Path 2	Path 3	Average	Path 1	Path 2	Path 3	Average
L-Shaped	GNC-TLS	0.979	0.729	0.391	0.699	1.22	<b>0.54</b>	0.925	0.895	18.79	26.06	24.71	23.14
	GNC-GM	<b>0.635</b>	<b>0.25</b>	0.781	<b>0.555</b>	<b>0.487</b>	0.64	<b>0.404</b>	<b>0.51</b>	<b>13.61</b>	<b>19.62</b>	<b>12.71</b>	<b>15.44</b>
	ADAPT	1.428	2.109	<b>0.316</b>	1.284	1.585	5.571	0.864	2.673	158.8	190.6	192.69	184.4
	RANSAC	3.849	0.981	3.069	2.633	122.3	131.3	128.2	127.2	44.35	36.11	46.73	43.13
Outlier ratio [%]									79.49	60.11	73.17	<b>70.92</b>	

TABLE II: Experimental validation with the performance of contact-based pose estimation.

other works that report outlier ratios of more than 95% with real data [6]. Although the putative correspondences have outliers, the pose estimation shows rotation and translation average accuracy of  $0.55^\circ$  and 0.51 mm, respectively, when using GNC-GM. As shown in the previous section, GNC approaches have the best performance overall, having rotation and translation errors under  $0.7^\circ$  and 1 mm, respectively. We, therefore, conclude that the results of the real experiments are comparable with the simulations and that the proposed method provides a high pose estimation accuracy despite the noise in the source data.

## VI. CONCLUSIONS

In this paper, we presented a contact-based pose estimation method for localizing a workpiece in the workspace of a collaborative robot. We compared the proposed method with other state-of-the-art methods and concluded that the proposed method for finding correspondences outperforms the other methods when the source data is only given along

a single path. Furthermore, we reformulated the distance function in the pose estimation method to use surface normal-to-normal information and demonstrated that we can perform robust pose estimation of the workpiece. A simulation-based comparison between existing PPF-based methods for finding correspondences and our method showed that the proposed method is significantly more robust towards noise. Finally, the proposed contact-based pose estimation was experimentally evaluated on a real robotic setup. The outcome of the experiments confirmed the simulation results by validating the method's high pose accuracy.

## ACKNOWLEDGMENT

This work was supported by MADE FAST. The authors thank MIT-SPARK and Jesus Briaes for sharing the implementation codes of the following references [5], [7], [9], [16].

## REFERENCES

- [1] L. Malin, A. Peter, J. Caroline, W. Boel, and W. Magnus, "Evaluation of flexible automation for small batch production," *Procedia Manufacturing*, vol. 25, pp. 177–184, 2018.
- [2] B. Maiseli, Y. Gu, and H. Gao, "Recent developments and trends in point set registration methods," *Journal of Visual Communication and Image Representation*, vol. 46, pp. 95–106, 2017.
- [3] X. Li, T. A. Fuhlbrigge, S. Choi, and B. Zhang, "Automatic offline program calibration in robotic cells," in *The 4th Annual IEEE International Conference on Cyber Technology in Automation, Control and Intelligent*, 2014, pp. 585–590.
- [4] W. Zhang, X. Ma, L. Cui, and Q. Chen, "3 points calibration method of part coordinates for arc welding robot," in *International Conference on Intelligent Robotics and Applications*, 10 2008, pp. 216–224.
- [5] M. A. Fischler and R. C. Bolles, "Random sample consensus: A paradigm for model fitting with applications to image analysis and automated cartography," *Commun. ACM*, vol. 24, no. 6, p. 381–395, jun 1981.
- [6] A. P. Bustos and T.-J. Chin, "Guaranteed outlier removal for point cloud registration with correspondences," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 40, no. 12, pp. 2868–2882, 2018.
- [7] V. Tzoumas, P. Antonante, and L. Carlone, "Outlier-robust spatial perception: Hardness, general-purpose algorithms, and guarantees," in *2019 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, 2019, pp. 5383–5390.
- [8] S. Gold and A. Rangarajan, "A graduated assignment algorithm for graph matching," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 18, no. 4, pp. 377–388, 1996.
- [9] H. Yang, P. Antonante, V. Tzoumas, and L. Carlone, "Graduated non-convexity for robust spatial perception: From non-minimal solvers to global outlier rejection," *IEEE Robotics and Automation Letters*, vol. 5, no. 2, pp. 1127–1134, 2020.
- [10] B. Drost, M. Ulrich, N. Navab, and S. Ilic, "Model globally, match locally: Efficient and robust 3D object recognition," in *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 07 2010, pp. 998–1005.
- [11] R. B. Rusu, N. Blodow, and M. Beetz, "Fast point feature histograms (FPFH) for 3D registration," in *2009 IEEE International Conference on Robotics and Automation*, 2009, pp. 3212–3217.
- [12] A. G. Buch and D. Kraft, "Local point pair feature histogram for accurate 3D matching," in *BMVC*, 2018.
- [13] S. Tang, X. Wang, X. Lv, T. X. Han, J. Keller, Z. He, M. Skubic, and S. Lao, "Histogram of oriented normal vectors for object recognition with a depth sensor," in *Asian conference on computer vision*. Springer Berlin Heidelberg, 2012, Conference Proceedings, pp. 525–538.
- [14] D. Nehab, S. Rusinkiewicz, J. Davis, and R. Ramamoorthi, "Efficiently combining positions and normals for precise 3D geometry," *ACM Transactions on Graphics (Proc. of ACM SIGGRAPH 2005)*, vol. 24, no. 3, Aug. 2005.
- [15] I. Eichhardt and L. Hajder, "Improvement of camera calibration using surface normals," in *2016 23rd International Conference on Pattern Recognition (ICPR)*, 2016, pp. 3745–3750.
- [16] J. Briales and J. Gonzalez-Jimenez, "Convex global 3D registration with lagrangian duality," in *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2017, pp. 5612–5621.
- [17] C. Sloth and I. Iturrate, "Simultaneous contact point and surface normal estimation during soft finger contact," in *Proceedings of the 20th International Conference on Advanced Robotics*, 2021, pp. 19–25.
- [18] D. Furcy, "Limited discrepancy beam search," in *In Proceedings of the International Joint Conference on Artificial Intelligence (IJCAI)*, 2005, pp. 125–131.
- [19] R. Zhou and E. A. Hansen, "Beam-stack search: Integrating backtracking with beam search," in *ICAPS*, 2005.
- [20] M. J. Black and A. Rangarajan, "On the unification of line processes, outlier rejection, and robust statistics with applications in early vision," *International Journal of Computer Vision*, vol. 19, no. 1, pp. 57–91, 1996.
- [21] A. Blake and A. Zisserman, *Visual Reconstruction*. The MIT Press, 1987.
- [22] C. Olsson, F. Kahl, and M. Oskarsson, "Branch-and-bound methods for Euclidean registration problems," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 31, no. 5, pp. 783–794, 2009.
- [23] M. Grant and S. Boyd, "CVX: Matlab software for disciplined convex programming, version 2.1," <http://cvxr.com/cvx>, Mar. 2014.
- [24] Y. Xiang, R. Mottaghi, and S. Savarese, "Beyond PASCAL: A benchmark for 3D object detection in the wild," in *IEEE Winter Conference on Applications of Computer Vision*, 2014, pp. 75–82.
- [25] K. S. Arun, T. S. Huang, and S. D. Blostein, "Least-squares fitting of two 3-D point sets," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. PAMI-9, no. 5, pp. 698–700, 1987.