

TacLoc: Global Tactile Localization on Objects from a Registration Perspective

Zirui Zhang, Fumin Zhang and Huan Yin

Abstract—Existing tactile-based methods generally rely on rendering data or pre-trained models, which limits generalizability and efficiency. In this study, we propose TacLoc, a novel tactile localization framework that addresses these challenges by formulating the problem as a one-shot point cloud registration task. TacLoc introduces a graph-theoretic partial-to-full registration method, leveraging tactile sensor-generated dense point clouds and surface normals for efficient and accurate pose estimation. The approach eliminates the need for extensive training data. It achieves robust performance through normal-guided graph pruning and a hypothesis-and-verification pipeline. Evaluations on the YCB-based dataset demonstrate the superiority of TacLoc in terms of accuracy, efficiency, and generalization. We also perform the real-world tests using Gelsight Mini sensor.

I. INTRODUCTION

Tactile sensing presents a promising strategy for contact-based sensing in robotic manipulation, particularly for tasks subject to visual occlusion. A crucial initial step for successful manipulation is accurately determining the robotic manipulator’s initial pose relative to an object. Current tactile localization methods often employ Monte Carlo Localization (MCL) [1], [2] or brute-force search within the $se(3)$ space [3], [4]. These techniques estimate pose by matching real sensor data against a pre-rendered or simulated model, a process that can be computationally demanding and may necessitate extensive training. In this study, we introduce TacLoc, a novel tactile localization framework that surpasses point cloud registration methods in accuracy. Utilizing a hypothesis-and-verification scheme, TacLoc obviates the need for sequential Bayesian inference. Instead, TacLoc identifies the optimal transformation by minimizing point-to-plane residuals in a single, one-shot localization. It achieves a mean rotation error (RE) of 22.38° , translation error (TE) of 5.0mm (for a total movement of 10mm) in TACTO simulator [5] when tested on 10 YCB objects [6]. We also demonstrate its application on a real 3D-printed object (Figure 1).

II. METHOD

From Raw Data to Initial Correspondence. As depicted in Figure 2, the process begins with the conversion of raw tactile images into points and normals. A point cloud submap can then be constructed from a sequence of these measurements, leveraging end-effector poses. Alternatively, a single-shot measurement can be used for one-shot pose estimation.

To ensure computational efficiency during pose estimation, voxel grid downsampling is applied to the point clouds. Key-points are then detected using the Intrinsic Shape Signatures

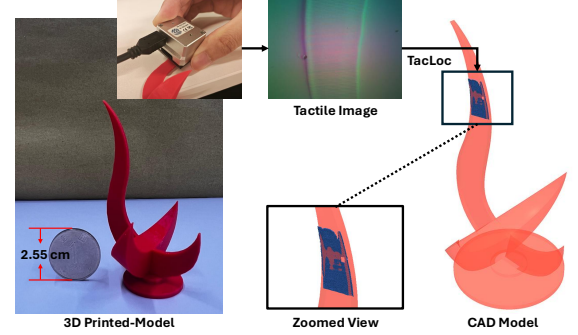


Fig. 1. Real-world demonstration. TacLoc performs fast, accurate, and robust 6DoF object localization based on only touch sensing.

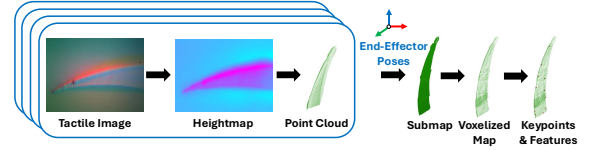


Fig. 2. For each tactile frame, we convert it to a point cloud with normals and build a submap using end-effector poses. This submap is then processed by our front-end pipeline.

(ISS) and are encoded FPFH for each keypoint. Finally, initial correspondences are established by performing Manhattan distance matching in the feature space.

Multiple Pose Hypotheses Generation. A compatibility graph is constructed based on pairwise geometric consistency check, including distance, normal alignment, and injective correspondence checks. In this graph, nodes represent individual correspondences, while edges connect pairs of correspondences deemed mutually compatible. Maximal cliques within this graph are subsequently extracted and ranked by size using a modified Bron-Kerbosch algorithm [7]. The top K cliques are selected to estimate the pose.

For each selected clique, we estimate its transformation $T_k = (R_k, t_k)$ by minimizing both point-to-point and normal-to-normal residuals. The rotation component is estimated as:

$$R_k = \arg \min_{R_k} \sum_{(i,j) \in C_k} (\|q'_j - R_k p'_i\|^2 + \alpha \cdot \arccos^2 \langle m_j, R_k n_i \rangle) \quad (1)$$

where p'_i and q'_j represent centered points from the source and target clouds, respectively; n_i and m_j denote their corresponding normals. The term α balances the weight between distance and normal differences. The translation component is then estimated following:

$$t_k = \arg \min_{t_k} \sum_{(i,j) \in C_k} \|t_k - (q_j - R_k p_i)\|^2 \quad (2)$$

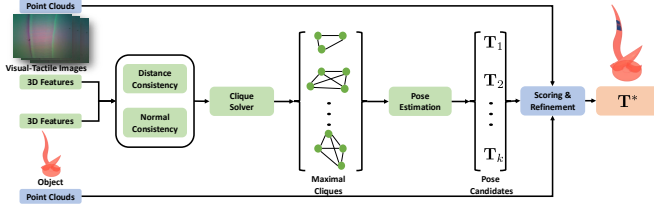


Fig. 3. TacLoc aligns a tactile-sensed point cloud to a CAD model by finding the transformation that best fits the points to the model’s surfaces, minimizing point-to-plane error.

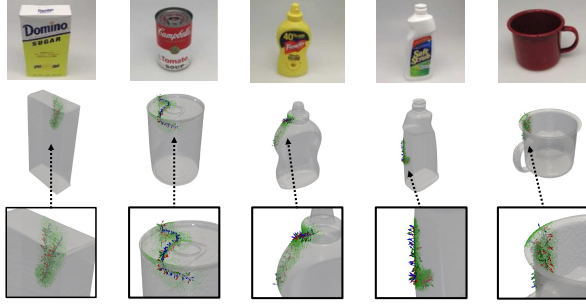


Fig. 4. Selected samples from the YCB-Reg benchmark.

Finally, we generate K pose candidates based on the pruned correspondences, which are obtained from the maximal cliques used for multiple hypotheses generation.

Pose Verification and Refinement. We utilize a point-to-plane loss function for geometric verification and refinement. This function, which is based on spatial proximity rather than feature descriptors, is expressed as follows:

$$\mathcal{L}(\delta \mathbf{R}_k, \delta \mathbf{t}_k) = \sum_i \langle \delta \mathbf{R}_k \hat{\mathbf{p}}_i^{\text{src}} - \tilde{\mathbf{q}}_i + \delta \mathbf{t}_k, \tilde{\mathbf{m}}_i \rangle^2 \quad (3)$$

where $\hat{\mathbf{p}}_i^{\text{src}} = \mathbf{R}_k \mathbf{p}_i^{\text{src}} + \mathbf{t}_k$, $\tilde{\mathbf{q}}_i$ denotes the closest point to the transformed source point $\hat{\mathbf{p}}_i^{\text{src}}$ in the downsampled target cloud, and $\tilde{\mathbf{m}}_i$ represents the associated normal of $\tilde{\mathbf{q}}_i$. The refined solution $\mathbf{T}_k^* = (\delta \mathbf{R}_k^* \mathbf{R}_k, \delta \mathbf{R}_k^* \mathbf{t}_k + \delta \mathbf{t}_k^*)$ is regarded as the hypothesis θ_k for clique C_k . And the weight is granted by $w_k = \exp[-\mathcal{L}(\delta \mathbf{R}_k^*, \delta \mathbf{t}_k^*)]$. The transformations achieving lower residual errors receive higher weights, with the highest-scoring estimate selected as the final pose \mathbf{T}^* .

III. EXPERIMENTS AND RESULTS

YCB-Reg Benchmark Comparison. We validate our method on global localization tasks using the YCB-Reg tactile dataset (Figure 4), consisting of ten YCB objects [6]. Each object includes ten 10cm sliding with TACTO simulator [5].

TacLoc is compared against three point cloud registration baselines: RANSAC [8], TEASER++ [9], and 3DMAC [10]. Average rotation error (RE), average translation error (TE), recall rate (RR) under 10mm and 30°, and average execution time are reported in Table I. Additionally, Figure 5 illustrates three representative failure cases, which were attributed to repetitive features that introduced ambiguity during the registration process.

Real-world Demonstration. The proposed TacLoc is further demonstrated using a GelSight Mini sensor on a 3D-printed model. The target point cloud is generated by uni-



Fig. 5. Three failure case studies. In each subfigure, the left shows the ground truth while the right displays the misalignment.

TABLE I
6 DOF LOCALIZATION RESULTS

| Method | RE (°) | TE (mm) | RR (%) | Time (s) |
|---------------|--------------|-------------|-------------|-------------|
| RANSAC [8] | 127.89 | 99.57 | 1.0 | 1.74 |
| TEASER++ [9] | 65.92 | 12.24 | 46.0 | 12.24 |
| 3DMAC [10] | 58.43 | 31.65 | 46.0 | 23.92 |
| TacLoc (Ours) | 22.38 | 5.00 | 85.0 | 3.16 |

formly sampling the original CAD model, shown in Figure 1. For the measurements, the GelSight SDK provides the heightmap, gradient map, and contact mask. Subsequently, the point cloud and surface normals are derived from the reconstructed geometry.

To ensure thorough validation, the demonstration was repeated ten times on different regions of the model. Three key challenges were identified: a *single, non-sliding touch* was used, which is more difficult than sliding-based methods; the 3D-printed model was highly symmetrical and lacked texture; and minor manufacturing defects caused discrepancies between the CAD and physical models. Under these challenging conditions, we report the success rate for the real-world TacLoc was 6/10.

REFERENCES

- [1] A. Petrovskaya and O. Khatib, “Global localization of objects via touch,” *IEEE Transactions on Robotics*, vol. 27, no. 3, pp. 569–585, 2011.
- [2] S. Suresh, Z. Si, S. Anderson, M. Kaess, and M. Mukadam, “Midas-Touch: Monte-Carlo inference over distributions across sliding touch,” in *Proc. Conf. on Robot Learning, CoRL*, Auckland, NZ, Dec. 2022.
- [3] M. Bauza, A. Bronars, and A. Rodriguez, “Tac2pose: Tactile object pose estimation from the first touch,” *The International Journal of Robotics Research*, vol. 42, no. 13, pp. 1185–1209, 2023.
- [4] Y. Dou, F. Yang, Y. Liu, A. Loquercio, and A. Owens, “Tactile-augmented radiance fields,” in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2024, pp. 26 529–26 539.
- [5] S. Wang, M. Lambeta, P.-W. Chou, and R. Calandra, “Tacto: A fast, flexible, and open-source simulator for high-resolution vision-based tactile sensors,” *IEEE Robotics and Automation Letters*, vol. 7, no. 2, pp. 3930–3937, 2022.
- [6] B. Calli, A. Singh, J. Bruce, A. Walsman, K. Konolige, S. Srinivasa, P. Abbeel, and A. M. Dollar, “Yale-cmu-berkeley dataset for robotic manipulation research,” *The International Journal of Robotics Research*, vol. 36, no. 3, pp. 261–268, 2017.
- [7] C. Bron and J. Kerbosch, “Algorithm 457: finding all cliques of an undirected graph,” *Communications of the ACM*, vol. 16, no. 9, pp. 575–577, 1973.
- [8] M. A. Fischler and R. C. Bolles, “Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography,” *Communications of the ACM*, vol. 24, no. 6, pp. 381–395, 1981.
- [9] H. Yang, J. Shi, and L. Carlone, “Teaser: Fast and certifiable point cloud registration,” *IEEE Transactions on Robotics*, vol. 37, no. 2, pp. 314–333, 2020.
- [10] J. Yang, X. Zhang, P. Wang, Y. Guo, K. Sun, Q. Wu, S. Zhang, and Y. Zhang, “Mac: Maximal cliques for 3d registration,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 46, no. 12, pp. 10 645–10 662, 2024.