

# Grasp, Slide, Roll: Comparative Analysis of Contact Modes for Tactile-Based Shape Reconstruction

Chung Hee Kim<sup>1,2</sup>, Shivani Kamtikar<sup>1,3</sup>, Tye Brady<sup>1</sup>, Taşkın Padır<sup>1</sup>, Joshua Migdal<sup>1</sup>

**Abstract**—Tactile sensing allows robots to gather detailed geometric information about objects through physical interaction, complementing vision-based approaches. However, efficiently acquiring useful tactile data remains challenging due to the time-consuming nature of physical contact and the need to strategically choose contact locations that maximize information gain while minimizing physical interactions. This paper studies how different contact modes affect object shape reconstruction using a tactile-enabled dexterous gripper. We compare three contact interaction modes: grasp-releasing, sliding induced by finger-grazing, and palm-rolling. These contact modes are combined with an information-theoretic exploration framework that guides subsequent sampling locations using a shape completion model. Our results show that the improved tactile sensing efficiency of finger-grazing and palm-rolling translates into faster convergence in shape reconstruction, requiring 34% fewer physical interactions while improving reconstruction accuracy by 55%. We validate our approach using a UR5e robot arm equipped with an Inspire-Robots Dexterous Hand, showing robust performance across primitive object geometries.

## I. INTRODUCTION

Robotic manipulation systems have traditionally relied on vision-based sensing to guide their interactions with objects. While effective in controlled environments, these systems often struggle when confronted with dynamic scenarios where objects shift during manipulation or in cluttered environments where visual occlusion becomes a significant challenge. These limitations stem from their dependence on pre-touch sensing and inability to adapt to real-time changes in the environment. Tactile sensing emerges as a promising complement to visual perception, offering robots the ability to gather precise geometric information through direct physical contact, much like the human sense of touch.

However, efficiently acquiring and utilizing tactile data presents several key challenges. Naive interaction strategies, such as simple grasp-and-release actions, often yield limited contact information, making accurate object reconstruction slow and impractical in real-world settings. Furthermore, the scarcity of large-scale tactile datasets limits the development of robust tactile perception algorithms. Together, these

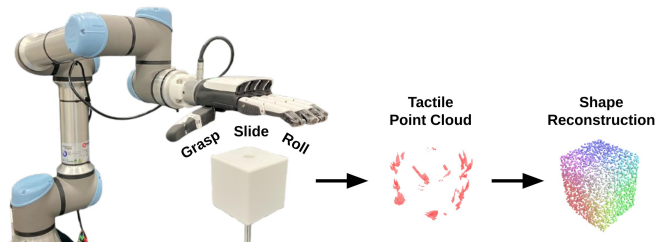


Fig. 1: Tactile information is promising to unlock contact-rich manipulation capabilities for robots. We investigate tactile data acquisition efficiency of various contact modes (sliding, rolling, and discrete contacts) in the context of 3D object reconstruction using tactile data only (no vision).

challenges highlight a central trade-off in tactile perception: richer object geometry estimates require more frequent contacts, yet practical systems must limit interactions to remain efficient. This motivates our central research question: *How can minimal tactile interaction be leveraged to enable reliable, efficient object identification and maximize geometric reconstruction accuracy?*

This paper presents a novel methodology to explore contact-rich object interactions for dexterous manipulators with tactile sensing through three main contributions. First, we employ a diffusion-based shape completion model that can reconstruct complete object geometry from sparse tactile measurements with zero-shot sim-to-real transfer. Second, we develop an information-theoretic exploration framework that strategically guides tactile sampling to maximize information gain. Thirdly, we present a comprehensive study of different contact interaction modes, comparing the efficiency of *grasp-releasing*, *finger-grazing*, and *palm-rolling* in tactile data acquisition for object reconstruction, characterized from contact kinematics of two surfaces [1].

We demonstrate our approach on a real robotic system comprising an anthropomorphic dexterous hand equipped with distributed tactile sensor arrays mounted on a 6-DOF robotic arm (see Fig. 1). Our experimental results show that dynamic contact modes (sliding and rolling) guided by our information-theoretic exploration policy achieve better performance compared to basic grasp-and-release interactions, requiring 34% fewer contacts while improving reconstruction accuracy by 55%. The system demonstrates robust shape completion across various object geometries, requiring an average of 8.4 interactions per object for accurate reconstruction. These findings highlight the critical role of diverse contact modes in robotic tactile exploration. Going beyond simple grasping can enhance tactile information acquisition, motivating exploration frameworks that strategically select

\*This work was done when Chung Hee Kim and Shivani Kamtikar was an intern at Amazon.

†Taşkın Padır holds concurrent appointments as a Professor of Electrical and Computer Engineering at Northeastern University and as an Amazon Scholar. This paper describes work performed at Amazon and is not associated with Northeastern University.

<sup>1</sup>Amazon Fulfillment Technologies & Robotics, Westborough MA, USA {bradytye, ptaskin, jmigdal}@amazon.com

<sup>2</sup>Robotics Institute at Carnegie Mellon University, Pittsburgh PA, USA chunghek@andrew.cmu.edu

<sup>3</sup>Siebel School of Computing and Data Science, University of Illinois at Urbana-Champaign, Champaign IL, USA skk7@illinois.edu

interaction modes for richer object understanding.

## II. RELATED WORK

### A. Tactile Sensing for Robotic Manipulation

Tactile sensing has been gaining attention as a rich feedback modality for enhancing robotic dexterity. Model-based approaches leverage tactile data to estimate object properties such as friction coefficients [2], slippage [3], material characteristics [4], as well as object pose estimation using tactile sensing integrated with vision [5], [6]. Grounded in physical principles, these methods offer valuable insights into tactile sensing; however, they often assume simplified gripper designs or rely on analytical models that may not generalize across diverse contact conditions and nonlinear interactions [7]. Recent trends shift toward deep learning-based approaches that directly utilize high-dimensional tactile data from vision-based tactile sensors for in-hand manipulation [8]–[11] and tactile servoing [12]. While these approaches remove the need for explicit modeling, they tend to struggle with out-of-distribution generalization and lack structured strategies for contact planning.

### B. Tactile-Based 3D Reconstruction

A growing body of research investigates tactile-guided shape reconstruction, often in combination with vision [13]. Early methods fuse depth images with tactile contacts for voxel-based reconstruction [14], while more recent approaches employ neural fields for in-hand visuotactile SLAM [15] or actively select contact regions based on visual priors [16], [17]. Exploration strategies vary from learned tactile-based policies [18], [19] to random sampling [20] and heuristic approaches [21], [22]. With the emergence of 3D Gaussian Splatting, several studies have investigated vision-tactile fusion for high-fidelity surface reconstruction [23]–[25]. While these methods demonstrate improved surface modeling, they primarily treat tactile sensing as a complementary cue to vision. In contrast, our work focuses on diverse tactile interaction strategies to enable complete object reconstruction from purely tactile feedback, without reliance on visual input.

### C. Tactile-Sensing Anthropomorphic Grippers

Several anthropomorphic end-effectors have been developed with embedded tactile sensors [26]–[28], yet many are limited to hardware demonstrations without control or perception strategies. Some works integrate tactile sensing with Bayesian and RL-based exploratory behaviors for object identification [29] or implement force control loops for manipulation [30]. However, most of these efforts do not explicitly model or analyze different modes of contact. Reconstruction methods using tactile-equipped anthropomorphic hands have shown promising results [15], [16], but none investigate the role of distinct contact modes in shaping object understanding. Our work is the first to study the impact of contact modes on tactile shape exploration using an anthropomorphic, tactile-enabled hand, offering new insights into how diverse exploratory behaviors can be harnessed for efficient reconstruction.

## III. SPARSE TACTILE SENSING TO DENSE GEOMETRY

This section presents our approach to reconstructing complete object geometry from sparse tactile measurements, detailing the shape completion model, key improvements for real-world deployment, and the synthetic data generation pipeline.

### A. Diffusion-Based 3D Shape Completion

For 3D shape completion, we build upon and enhance the Sparse Point-Voxel Diffusion (SPVD) model [31], with key improvements enabling position, scale, and rotation invariance (detailed in Sec. III-B). Specifically designed to diffuse point clouds, SPVD’s dual-space architecture combines point-space processing for local feature extraction with voxel-space encoding for global shape understanding through 3D convolutions. The model implements a diffusion process that iteratively denoises  $N$  randomly sampled points  $P_{pred} = \{p_i \in \mathbb{R}^3 \mid i = 1, \dots, N\}$ , conditioned on tactile measurements represented by a partial point cloud  $P_{meas} = \{p_j \in \mathbb{R}^3 \mid j = 1, \dots, M\}$ . Starting from Gaussian noise, the model progressively refines  $P_{pred}$  through  $T$  denoising steps to generate a dense point cloud representing the complete object geometry (see Fig. 7).

### B. Improvements for Real-World Deployment

The original SPVD model was designed for clean 3D assets, making it unsuitable for direct use with tactile data. Real-world tactile sensing presents several challenges absent in simulations. Physical sensor readings contain measurement noise and drift, objects are encountered in random poses, and their scale can vary arbitrarily. Additionally, the reference frame for tactile observations depend on the robot’s configuration, unlike standard 3D shape assets where canonical object poses are readily available. To enable real-world tactile shape completion, we introduced critical enhancements and data augmentations that were necessary for the model to process tactile measurements:

1) *Model Enhancements*: We applied three key modifications to enable the model’s real-world applicability:

- **Point Normal Input Features**: We enrich the input representation by incorporating surface normal vectors at each point. These normal vectors, which can readily be approximated as perpendicular to the contact surface, provide directional guidance for shape generation. Fig. 2(a) demonstrates how the absence of contact normal vectors leads to incorrect shape completion (left), while incorporating normal information enables reconstruction aligned with the true surface orientation (right).
- **Novel Point Generation**: We modified the model to generate entirely new output points, departing from SPVD’s default approach where measured points are retained in the final prediction ( $P_{meas} \subseteq P_{pred}$ ). While the assumption that input points belong to the true shape works for clean data, tactile measurements often contain noise. Our modification ensures the predicted point cloud is distinct from the input ( $P_{meas} \cap P_{pred} = \emptyset$ ),

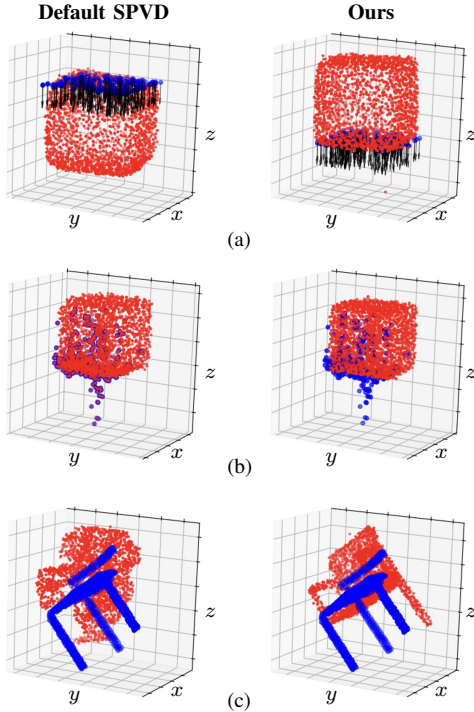


Fig. 2: Comparative examples of original SPVD results and our improved approach. Partial input points and shape completed output points are colored blue and red, respectively. (a) Shape completion with (right) and without (left) **normal vectors** (shown as black arrows) illustrates how normal guidance disambiguates shape generation direction. (b) **Novel point generation** filters noisy measurements (right), while the original approach preserves noisy inputs (left). (c) **Rotation-invariant** shape completion achieved through random rotation augmentation, demonstrated with a chair point cloud.

enabling the model to effectively filter noisy inputs and handle input point sets of unconstrained size. Fig. 2(b) illustrates this noise-filtering capability.

- **Position and Scale Invariance:** We applied per-sample normalization to achieve position and scale invariance. For each input point cloud  $P = P_{meas}$ , we apply:

$$P' = (P - \bar{P}) \quad (1)$$

$$P_{normalized} = P' / (\lambda \sigma_{\|P'\|_2}) \quad (2)$$

where  $\bar{P}$  represents the mean,  $\lambda$  is a scaling factor, and  $\sigma_{\|P'\|_2}$  denotes the standard deviation of L2-norms of the centered points  $P'$ . This normalization enables origin-independent and scale invariant predictions during inference.

2) *Data Augmentation:* To improve sim-to-real transfer and enhance model robustness, we applied the following data augmentations (illustrated in Fig. 3):

- **Contact Point Truncation:** Number of input points were varied by truncating  $P_{meas}$  in reverse chronological order, simulating different contact sequences and partial object interactions.
- **Sensor Noise Simulation:** Gaussian noise was injected into  $P_{meas}$  to emulate physical sensor noise.

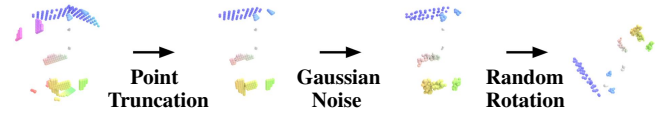


Fig. 3: Data augmentation pipeline on a simulated tactile point cloud including contact point truncation, Gaussian noise injection, and random rotation, illustrated using contact data from a sphere. Points are colored to show correspondence before and after applying the augmentation.

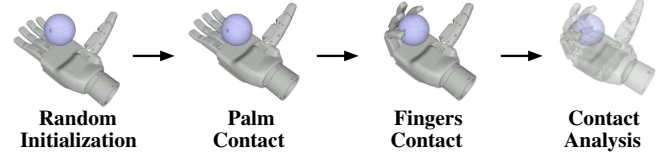


Fig. 4: Contact motion sequence of the tactile contact simulator. The target object is a sphere shaded in purple.

- **Random Rotation:** Input/target point clouds were rotated arbitrarily during training, ensuring the model learns rotation-invariant representations of object geometry as illustrated in Fig. 2(c).

### C. Synthetic Tactile Data Generation

To train the shape completion model, we developed a custom tactile contact simulator that generates synthetic training data using gripper forward kinematics and collision meshes (see Sec. V-B for gripper specifications). The simulator emulates a four-step heuristic contact motion sequence (illustrated in Fig. 4):

- 1) The gripper is randomly initialized at a preset distance from the target object, with the palm facing the object.
- 2) The gripper approaches the target object along the palm-normal axis until initial contact is detected.
- 3) Each finger closes independently until object-finger contact is detected.
- 4) The positions of tactile contact points are computed through gripper-object mesh intersection analysis.

This process is repeated multiple times per object, resulting in a synthetic training dataset of paired tactile contact point clouds ( $P_{meas}$ ) and ground truth object meshes, enabling efficient scaling of our training data.

## IV. TACTILE EXPLORATION FRAMEWORK

This section presents our approach to address two critical challenges in tactile-driven perception: autonomous exploration and efficient data acquisition. We propose an information-theoretic framework for guiding optimal sampling locations, while investigating alternative contact interaction modes to improve tactile sensing efficiency.

### A. Information-Theoretic Exploration

We present an exploration strategy that optimizes tactile sampling for 3D object reconstruction. Our framework iteratively estimates object geometry with a pretrained shape completion network based on tactile measurements. By quantifying uncertainties in these estimates, we compute

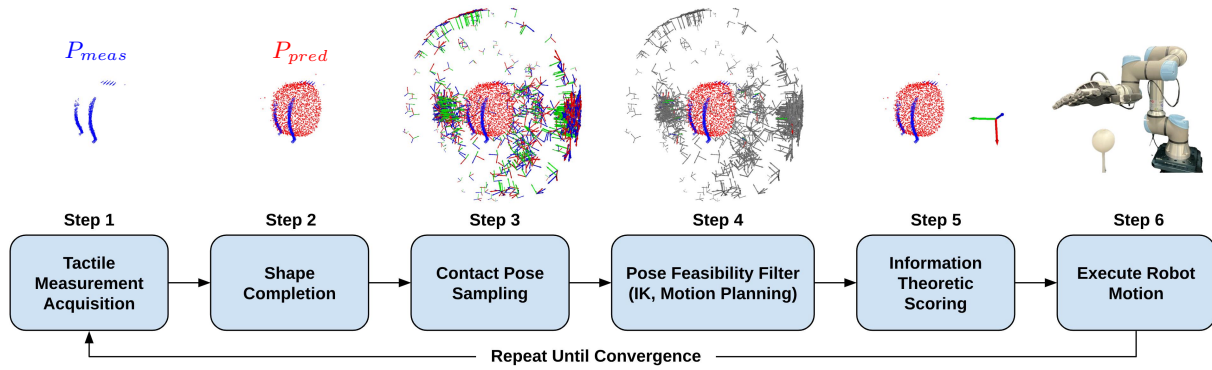


Fig. 5: Process diagram of the information-theoretic tactile exploration. Measured tactile points  $P_{meas}$  are colored blue (Step 1), while the shape completed point cloud  $P_{pred}$  is colored red (Step 2). The coordinate frames in Step 3 represent candidate contact poses, while gray coordinate frames in Step 4 represents filtered candidates. The contact pose with the highest score is selected (Step 5) and executed (Step 6).

information gain metrics to identify the most informative regions for subsequent sampling. The robot targets high-uncertainty areas, ensuring new tactile contacts improve our shape understanding. The framework illustrated in Fig. 5 enables systematic object exploration with the objective of shape reconstruction with minimal tactile interactions.

1) *Tactile Measurement Acquisition*: We obtain 3D tactile measurements through forward kinematics of the robot arm and hand. The tactile sensor locations are transformed from the hand frame to the base frame:

$$P^{base} = T_{TCP}^{base} T_{hand}^{TCP} P^{hand} \quad (3)$$

where  $P^{base}$  and  $P^{hand}$  represent the tactile point clouds in their respective frames, and the transformation matrices map between the robot base, tool center point (TCP), and hand frames.

2) *Shape Completion*: A trained shape completion model (discussed in Sec. III) processes the tactile measurements  $P_{meas}$  to predict complete object geometry, establishing our current belief state  $P_{pred}$  of the object shape.

3) *Contact Pose Sampling*: We generate candidate contact poses based on both measured tactile points  $P_{meas}$  and the predicted shape point cloud  $P_{pred}$ . For each point  $p_i \in P_{pred}$ , we compute a coverage distance:

$$d_{min}(p_i) = \min_j \|p_i - p_j\|_2 \text{ where } p_j \in P_{meas} \quad (4)$$

which is the Euclidean distance to its nearest measured tactile point. Points exceeding a 5mm coverage threshold are importance-sampled ( $k = 500$ ) to identify underexplored regions. To convert these points into actionable 6-DoF end-effector poses  $\mathbf{p}$ , we align the hand's palm with the surface normal at each sampled point, oriented toward the object, with a uniformly sampled yaw angle as depicted in Fig. 5.

4) *Feasibility Filtering*: Candidate poses undergo two-stage filtering: (1) inverse kinematics and joint limit checks, and (2) collision-free motion planning via cuRobo [32]. We retain poses with valid robot trajectories  $\mathcal{T}_{\mathbf{p}}$ .

5) *Information-Theoretic Scoring*: We evaluate candidate poses using an information-gain versus motion-cost metric:

$$Score(\mathbf{p}) = IG_{coverage}(\mathbf{p}) - C_{motion}(\mathcal{T}_{\mathbf{p}}) \quad (5)$$

The information gain  $IG_{coverage}$  measures the potential value of unexplored regions. We apply a Gaussian kernel mapping to convert the coverage distance  $d_{min}(p_i)$  to a heuristic probability  $P(p_i)$ :

$$P(p_i) = \exp\left(-\frac{d_{min}(p_i)^2}{2\sigma^2}\right) \quad (6)$$

Applying Shannon's definition of self-information [33]:

$$IG_{coverage}(p_i) = -\log P(p_i) = \frac{d_{min}(p_i)^2}{2\sigma^2} \quad (7)$$

where  $\sigma$  is a bandwidth parameter controlling the sensitivity. This maps large (small) distances to greater (smaller) information gain, consistent with the intuition that predicted surfaces further from measured points represent regions of higher uncertainty.

The motion cost  $C_{motion}$  is calculated as a weighted sum of joint-space movements along the planned trajectory  $\mathcal{T}_{\mathbf{p}}$ :

$$C_{motion}(\mathcal{T}_{\mathbf{p}}) = \sum_{t=1}^T \sum_{j=1}^6 w_j |q_{t,j} - q_{t-1,j}| \quad (8)$$

where  $q_{t,j}$  represents the position of joint  $j$  at timestep  $t$ , and  $w_j$  denotes joint-specific weights that allow for differential penalization of joint movements. The equation computes the total motion cost by summing weighted joint displacements across all timesteps.

6) *Execution*: The robot executes the highest-scoring feasible trajectory and interacts with the object. The process repeats until shape prediction converges, identified when the Chamfer distance between  $K$  consecutive beliefs remain below a threshold.

## B. Contact Interaction Modes

Naive binary grasp-and-release approaches yield limited contact points per interaction. To address this limitation, we consider two alternative contact interaction modes, namely *finger-grazing* and *palm-rolling*, which meaningfully improve tactile sensing efficiency. Our contact strategy implements these interactions in two phases. First, the robot moves its palm-oriented hand toward the target object along a linear

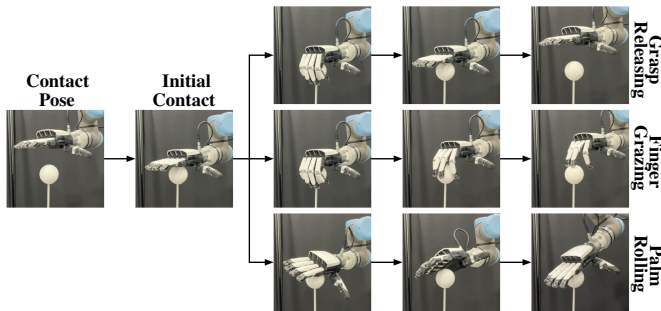


Fig. 6: Three contact interaction modes including grasp-releasing (top), finger-grazing (middle), and palm-rolling (bottom).

path. Upon detecting contact, the robot interacts with the object in one of three interaction modes (see Fig. 6):

- *Grasp-Releasing*: A sequence of power-grasp and releasing motion.
- *Finger-Grazing*: Fingers curl inward under torque control while the arm retracts under force control, inducing sliding contacts.
- *Palm-Rolling*: Fingers maintain open positions while the arm rotates under force control inducing rolling contacts: (1)  $\pm 30$  degree bidirectional rotation about the gripper axis, followed by (2) a 30 degree forward pitch about the palm’s lateral axis.

## V. EXPERIMENTS

### A. Shape Completion Results on Synthetic Data

We evaluated our shape completion model on a synthetic dataset comprising three primitive shapes: balls, boxes, and cylinders [34]. For each shape category, we generated a dataset that consists of 100 training, 15 validation, and 15 test meshes with randomized aspect ratios. Each mesh underwent 20 simulated contact-grasp interactions to create input-output data sample pairs. To enhance real-world robustness of the model, we augmented the dataset by applying point truncation ranging from 0% to 90% in 10% increments, resulting in 3,000 training samples. During training, we dynamically applied Gaussian noise and random rotations. The model was trained for 2,000 epochs with a batch size of 64 and learning rate of  $5e-4$ . We used a DDIM scheduler with 1,000 training steps and 100 inference steps to balance computational efficiency and reconstruction quality, resulting in an inference time of 2.5 seconds per sample on an NVIDIA GeForce RTX 2080 Super GPU. Fig. 7 shows qualitative results on the synthetic test dataset, demonstrating the model’s ability to accurately reconstruct complete object shapes from partial observations.

### B. Real-World Experiment Setup

We conducted real-world experiments using the anthropomorphic Inspire-Robots Dexterous Hand mounted on a UR5e robot arm (Fig. 8(a)). The hand exhibits five fingers with six degrees of freedom (extra DOF for the thumb) and is equipped with tactile sensing capabilities, featuring pressure

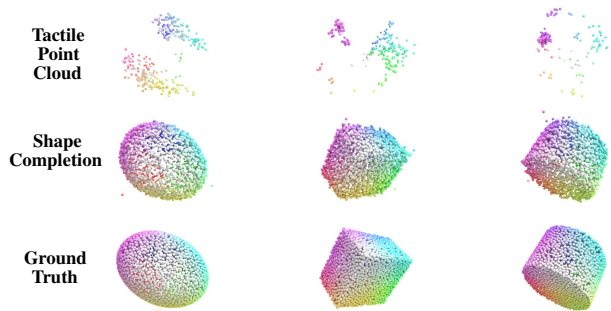


Fig. 7: Shape completion results on simulated tactile point clouds.

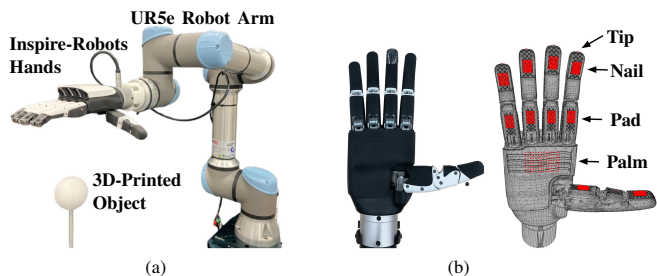


Fig. 8: (a) Our system features a UR5e robot arm and the Inspire-Robots Dexterous Hands embedded with tactile sensors. (b) Tactile sensor arrays on the palmar side of the anthropomorphic end-effector is visualized in red.

sensor arrays distributed across multiple contact surfaces (see Fig. 8(b)):  $3 \times 3$  arrays on fingertips,  $12 \times 8$  on nails,  $10 \times 8$  on finger pads, and  $8 \times 14$  on the palm. While our implementation utilizes this specific hardware configuration, our approach is designed to be compatible with any tactile gripper capable of generating contact point cloud data.

We tested our approach on nine 3D-printed objects comprised of three primitive shapes. For each shape category, we included three variations: a standard geometric form (perfect sphere, cube and cylinder), a scaled variant 1.5 times larger, and a deformed variant with modified aspect ratios and an orientation offset of 45-degrees (see Fig. 9 and Table I). Each object was fixed relative to the robot base and was tested across 5 different initial robot poses per contact mode. To ensure consistent initialization, we manually positioned the robot such that linear movement along the palm-axis would result in object contact during the first interaction. Episodes were limited to a maximum of 20 contact interactions or until convergence criteria was met (discussed in Sec. IV-A). This experimental design yielded 135 total trials (3 contact modes  $\times$  9 objects  $\times$  5 trials). The shape completion model, trained exclusively on synthetic data as described in Sec. V-A, was applied directly without additional fine-tuning, with each exploration episode conditioned only on the guaranteed contact initialization and the pre-trained model.

### C. Shape Completion on Real-World Tactile Data

Fig. 9 presents qualitative examples of object shape reconstruction through tactile exploration using the three interaction modes (grasp-releasing, finger-grazing, and palm-rolling). Our results demonstrate successful sim-to-real transfer, with the shape completion model effectively generalizing to real-world data. This transferability can be attributed

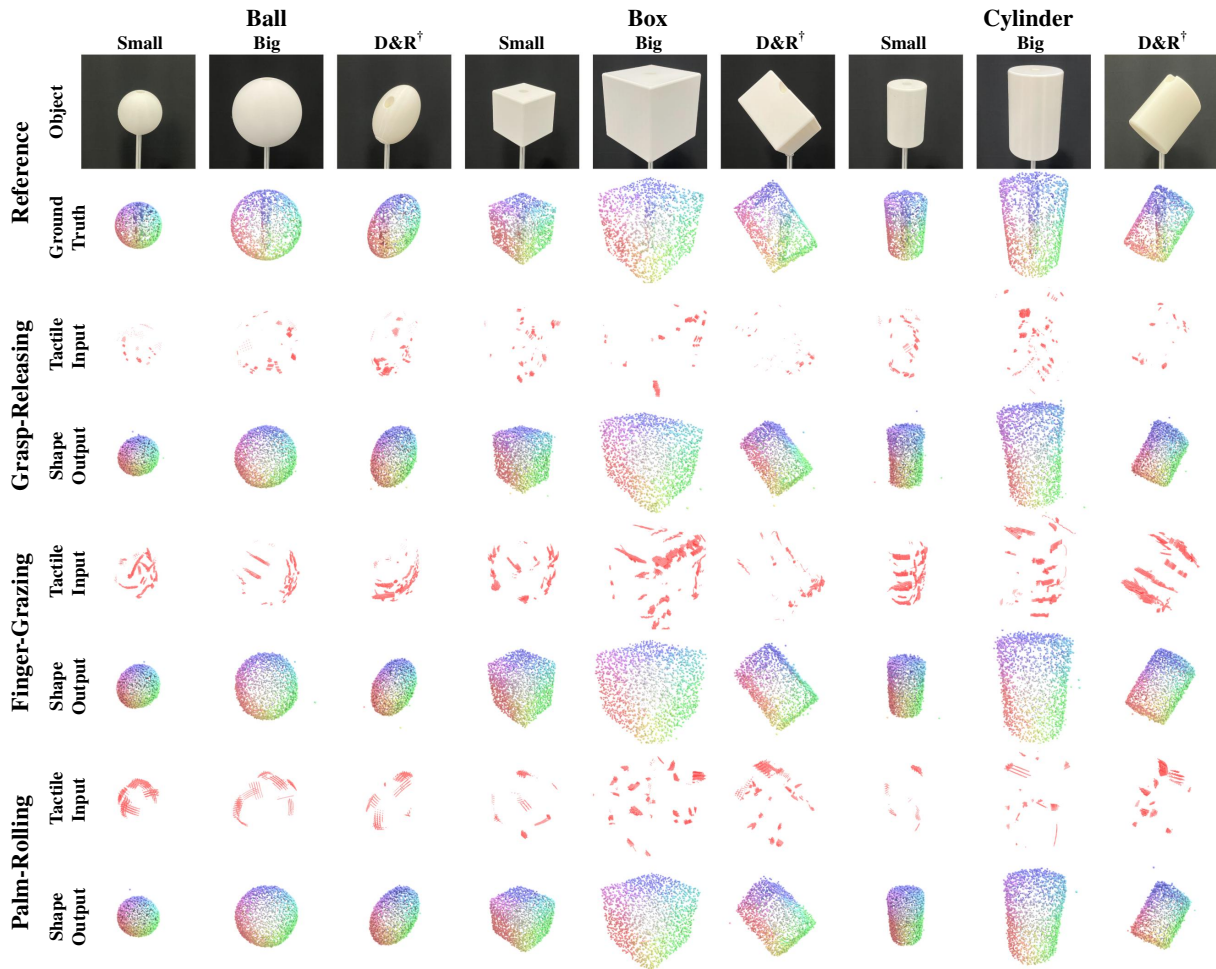


Fig. 9: Qualitative results of shape reconstruction on real-world tactile data. Tactile measurements are obtained through information-theoretic exploration using one of three contact interaction modes. We tested our framework on nine objects fixed relative to the robot base. <sup>†</sup>D&R: Deformed & Rotated.

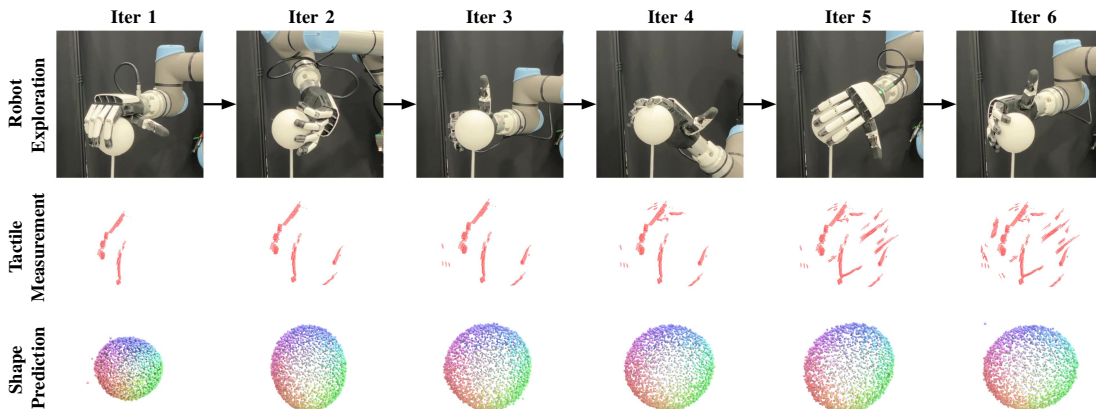


Fig. 10: Sequential robotic tactile exploration of a 3D-printed ball using finger-grazing motions. (Top row) Contact locations from each exploration step. (Middle row) Accumulated tactile point cloud data. (Bottom row) Progressive shape reconstruction results. See supplementary video for detailed visualizations.

to both the relatively low sim-to-real gap of point cloud representations (compared to RGB modalities) and the use of comprehensive augmentation and normalization strategies. The results further show that all three modes were capable of reconstructing the object when guided by the information-theoretic exploration policy described in Sec. IV-A and

illustrated in Fig. 10 (see supplementary video).

#### D. Quantitative Results & Discussion

Despite this shared capability, the reconstruction accuracy and the number of interactions required for convergence varied across contact modes. Detailed quantitative results

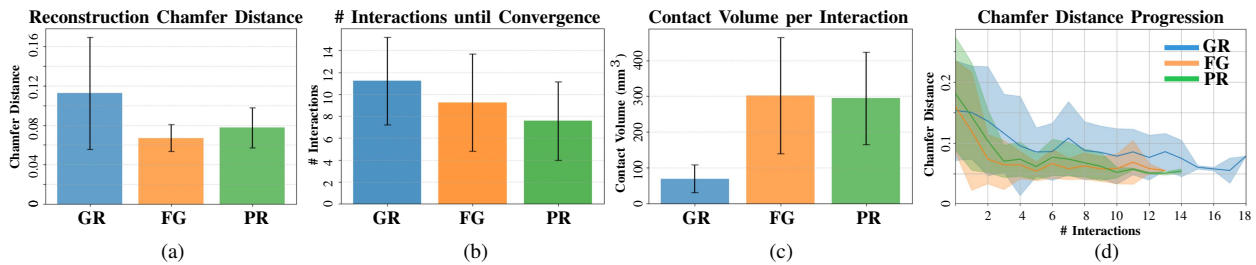


Fig. 11: Key metrics across different contact modes – **GR**: Grasp-Releasing / **FG**: Finger-Grazing / **PR**: Palm-Rolling. (a) **Reconstruction Chamfer Distance**: Average Chamfer distance for shape reconstruction at convergence. (b) **Number of Interactions Until Convergence**: Average number of interactions required to achieve shape reconstruction convergence. (c) **Contact Volume per Interaction**: Average volume of tactile data collected during each interaction. (d) **Chamfer Distance Progression**: Illustrates how reconstruction Chamfer distance decreases over successive interactions.

TABLE I: Quantitative Results for Real-World Tactile Shape Reconstruction

Shape Category	Variation	Dimension (mm)	Chamfer Distance			# Interactions to Converge			Contact Volume (mm <sup>3</sup> )		
			GR	FG	PR	GR	FG	PR	GR	FG	PR
Ball	Small	$r_x = r_y = r_z = 30$	0.079	<b>0.046</b>	0.048	6.8	5.2	<b>3.2</b>	52	263	<b>472</b>
	Big	$r_x = r_y = r_z = 45$	0.095	<b>0.073</b>	0.088	10.0	<b>5.4</b>	7.0	114	174	<b>404</b>
	D&R <sup>†</sup>	$r_x = r_y = 30, r_z = 45$	<b>0.047</b>	0.048	0.060	11.0	6.2	<b>4.8</b>	89	<b>377</b>	319
Box	Small	$w = h = l = 60$	0.102	0.060	<b>0.042</b>	13.4	11.6	<b>7.2</b>	65	<b>417</b>	137
	Big	$w = h = l = 90$	0.099	<b>0.090</b>	0.095	11.8	<b>10.6</b>	11.4	77	238	<b>292</b>
	D&R <sup>†</sup>	$w = h = 60, l = 90$	0.243	<b>0.062</b>	0.100	13.6	16.6	<b>10.0</b>	39	234	<b>362</b>
Cylinder	Small	$r_x = r_y = 25, h = 80$	0.084	<b>0.050</b>	0.074	9.6	<b>4.8</b>	<b>4.8</b>	50	<b>384</b>	189
	Big	$r_x = r_y = 37.5, h = 120$	0.191	<b>0.114</b>	0.125	12.6	12.4	<b>8.4</b>	80	<b>242</b>	200
	D&R <sup>†</sup>	$r_x = 25, r_y = 37.5, h = 80$	0.070	<b>0.051</b>	0.059	12.6	<b>10.4</b>	11.4	65	<b>426</b>	283

\*All entries are averaged over 5 trials, initialized from different end-effector poses. **GR**: Grasp-Releasing / **FG**: Finger-Grazing / **PR**: Palm-Rolling

<sup>†</sup>**D&R**: Deformed & Rotated (45°)

are reported in Table I. The mean Chamfer distances at convergence were 0.114, 0.067, and 0.077 (Fig. 11(a)), with corresponding mean interaction counts of 11.2, 9.2, and 7.6 (Fig. 11(b)) for grasp-releasing, finger-grazing, and palm-rolling, respectively. We further quantified the tactile information acquired per interaction by voxelizing the accumulated tactile point clouds at a resolution of 1 mm. The resulting mean contact volumes were 70, 302, and 294 mm<sup>3</sup> per interaction for the three modes, respectively (Fig. 11(c)).

Our experimental results demonstrate distinct advantages and trade-offs among the three contact interaction modes. The dynamic nature of sliding and rolling interactions proved particularly effective at gathering rich tactile information. While grasp-releasing provides discrete contact points, both finger-grazing and palm-rolling enable continuous surface sampling through sweeping motions, resulting in approximately four times more tactile data acquisition per interaction in terms of contact volume. This enhanced data density translated to improved reconstruction accuracy, with finger-grazing and palm-rolling achieving 70% and 48% improved accuracy respectively, while requiring 22% and 47% fewer contact interactions compared to grasp-releasing. This efficiency is evident in Fig. 11(d), which illustrates the faster convergence rates of finger-grazing and palm-rolling in terms of Chamfer distance reduction over episode progression.

However, each interaction mode presents distinct operational considerations influenced by both hardware constraints and control complexity. Grasp-releasing, while yielding the

least tactile data per interaction, offers the advantages of minimal execution time and reduced sensor noise due to the absence of motion-induced latency issues. Finger-grazing achieves efficient data acquisition but requires precise coordination between arm and hand motions, introducing potential latency challenges in real-world implementation. Palm-rolling, despite requiring longer execution times due to extended arm rotations, minimizes finger motion-related noise and proves particularly effective for curved surfaces. However, its performance is notably constrained when exploring flat surfaces (in box-like objects), primarily due to the limited tactile surface area and the concave geometry of the palmar side of the hand. These findings highlight the importance of considering both the geometric properties of the tactile sensor surfaces and the target object characteristics when selecting an appropriate interaction strategy.

## VI. CONCLUSION

In this work, we investigated the effectiveness of different contact modes for tactile data acquisition for the task of object reconstruction using an anthropomorphic hand equipped with tactile sensors. A diffusion-based shape completion model, trained on simulated data, was employed to reconstruct complete object geometry from partial tactile point clouds. Real-world experiments, guided by an information-theoretic exploration framework, compare three interaction modes: grasp-releasing, finger-grazing, and palm-rolling. The results show that dynamic contact modes achieve a reconstruction accuracy improvement of 55% while requiring

34% fewer interactions than basic grasping, underscoring the value of diverse contact strategies for efficient tactile exploration. Future directions include adaptive contact mode selection informed by local geometry inference and real-time tactile feedback.

## REFERENCES

- [1] C. Cai and B. Roth, "On the spatial motion of a rigid body with point contact," in *Proceedings. 1987 IEEE International Conference on Robotics and Automation*, vol. 4, pp. 686–695, 1987.
- [2] W. Chen, H. Khamis, I. Birznieks, N. F. Lepora, and S. J. Redmond, "Tactile sensors for friction estimation and incipient slip detection—toward dexterous robotic manipulation: A review," *IEEE Sensors Journal*, vol. 18, no. 22, pp. 9049–9064, 2018.
- [3] V. A. Ho and S. Hirai, "Two-dimensional dynamic modeling of a sliding motion of a soft fingertip focusing on stick-to-slip transition," in *2010 IEEE International Conference on Robotics and Automation*, pp. 4315–4321, 2010.
- [4] K. Dai, X. Wang, A. M. Rojas, E. Harber, Y. Tian, N. Paiva, J. Gnehm, E. Schindewolf, H. Choset, V. A. Webster-Wood, and L. Li, "Design of a biomimetic tactile sensor for material classification," in *2022 International Conference on Robotics and Automation (ICRA)*, pp. 10774–10780, 2022.
- [5] P. K. Murali, A. Dutta, M. Gentner, E. Burdet, R. Dahiya, and M. Kaboli, "Active visuo-tactile interactive robotic perception for accurate object pose estimation in dense clutter," *IEEE Robotics and Automation Letters*, vol. 7, no. 2, pp. 4686–4693, 2022.
- [6] H. Li, J. Akl, S. Sridhar, T. Brady, and T. Padir, "Vita-zero: Zero-shot visuotactile object 6d pose estimation," in *2025 IEEE International Conference on Robotics and Automation (ICRA)*, 2025.
- [7] Z. Kappassov, J. A. Corrales Ramon, and V. Perdereau, "Tactile sensing in dexterous robot hands - Review," *Robotics and Autonomous Systems*, vol. 74, Part A, pp. 195–220, 2015.
- [8] M. Lambeta, P.-W. Chou, S. Tian, B. Yang, B. Maloon, V. R. Most, D. Stroud, R. Santos, A. Byagowi, G. Kammerer, D. Jayaraman, and R. Calandra, "Digit: A novel design for a low-cost compact high-resolution tactile sensor with application to in-hand manipulation," *IEEE Robotics and Automation Letters*, vol. 5, no. 3, pp. 3838–3845, 2020.
- [9] I. Guzey, B. Evans, S. Chintala, and L. Pinto, "Dexterity from touch: Self-supervised pre-training of tactile representations with robotic play," in *7th Annual Conference on Robot Learning*, 2023.
- [10] H. van Hoof, T. Hermans, G. Neumann, and J. Peters, "Learning robot in-hand manipulation with tactile features," in *2015 IEEE-RAS 15th International Conference on Humanoid Robots (Humanoids)*, pp. 121–127, 2015.
- [11] G. Khandate, S. Shang, E. T. Chang, T. L. Saidi, Y. Liu, S. M. Dennis, J. Adams, and M. Ciocarlie, "Sampling-based exploration for reinforcement learning of dexterous manipulation," *arXiv preprint arXiv:2303.03486*, 2023.
- [12] S. Tian, F. Ebert, D. Jayaraman, M. Mudigonda, C. Finn, R. Calandra, and S. Levine, "Manipulation by feel: Touch-based control with deep predictive models," in *2019 International Conference on Robotics and Automation (ICRA)*, pp. 818–824, 2019.
- [13] A. Dutta, E. Burdet, and M. Kaboli, "Vitract: Robust object shape perception via active visuo-tactile interaction," *IEEE Robotics and Automation Letters*, vol. 9, no. 12, pp. 11250–11257, 2024.
- [14] S. Wang, J. Wu, X. Sun, W. Yuan, W. T. Freeman, J. B. Tenenbaum, and E. H. Adelson, "3d shape perception from monocular vision, touch, and shape priors," in *2018 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pp. 1606–1613, 2018.
- [15] S. Suresh, H. Qi, T. Wu, T. Fan, L. Pineda, M. Lambeta, J. Malik, M. Kalakrishnan, R. Calandra, M. Kaess, J. Ortiz, and M. Mukadam, "Neuralfeels with neural fields: Visuotactile perception for in-hand manipulation," *Science Robotics*, vol. 9, no. 96, p. ead10628, 2024.
- [16] E. J. Smith, D. Meger, L. Pineda, R. Calandra, J. Malik, A. Romero, and M. Drozdal, "Active 3d shape reconstruction from vision and touch," in *Advances in Neural Information Processing Systems* (A. Beygelzimer, Y. Dauphin, P. Liang, and J. W. Vaughan, eds.), 2021.
- [17] P. Oikonomou, G. Retsinas, P. Maragos, and C. S. Tzafestas, "Proactive tactile exploration for object-agnostic shape reconstruction from minimal visual priors," *arXiv preprint arXiv:2505.11975*, 2025.
- [18] Y. Wang, Z. Zhang, J. Qiu, D. Sun, Z. Meng, X. Wei, and X. Yang, "Touch2shape: Touch-conditioned 3d diffusion for shape exploration and reconstruction," 2025.
- [19] A.-H. Shahidzadeh, S. J. Yoo, P. Mantripragada, C. D. Singh, C. Fermüller, and Y. Aloimonos, "Actexplore: Active tactile exploration on unknown objects," in *2024 IEEE International Conference on Robotics and Automation (ICRA)*, pp. 3411–3418, 2024.
- [20] M. Comi, Y. Lin, A. Church, A. Tonioni, L. Aitchison, and N. F. Lepora, "Touchsdf: A deepsf approach for 3d shape reconstruction using vision-based tactile sensing," *IEEE Robotics and Automation Letters*, vol. 9, no. 6, pp. 5719–5726, 2024.
- [21] F. Luan, C. Wang, Z. Wang, J. Yue, Y. Zhou, and B. He, "A practical tactile exploration with enhanced sampling for fast shape estimation," *Measurement*, vol. 252, p. 117350, 2025.
- [22] H.-J. Huang, M. Kaess, and W. Yuan, "Normalflow: Fast, robust, and accurate contact-based object 6dof pose tracking with vision-based tactile sensors," *IEEE Robotics and Automation Letters*, vol. 10, no. 1, pp. 452–459, 2025.
- [23] A. Swann, M. Strong, W. K. Do, G. S. Camps, M. Schwager, and M. Kennedy, "Touch-gs: Visual-tactile supervised 3d gaussian splatting," in *2024 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pp. 10511–10518, 2024.
- [24] M. Comi, A. Tonioni, J. Tremblay, M. Yang, V. Blukis, Y. Lin, N. F. Lepora, and L. Aitchison, "Snap-it, tap-it, splat-it: Tactile-informed 3d gaussian splatting for reconstructing challenging surfaces," in *International Conference on 3D Vision 2025*, 2025.
- [25] I. Fang, K. Shi, X. He, S. Tan, Y. Wang, H. Zhao, H.-J. Huang, W. Yuan, C. Feng, and J. Zhang, "Fusionsense: Bridging common sense, vision, and touch for robust sparse-view reconstruction," in *2025 IEEE International Conference on Robotics and Automation (ICRA)*, 2025.
- [26] D. Goger, N. Gorges, and H. Worn, "Tactile sensing for an anthropomorphic robotic hand: Hardware and signal processing," in *2009 IEEE International Conference on Robotics and Automation*, pp. 895–901, 2009.
- [27] A. Schmitz, P. Maiolino, M. Maggiali, L. Natale, G. Cannata, and G. Metta, "Methods and technologies for the implementation of large-scale robot tactile sensors," *IEEE Transactions on Robotics*, vol. 27, no. 3, pp. 389–400, 2011.
- [28] B. Choi, S. Lee, H. R. Choi, and S. Kang, "Development of anthropomorphic robot hand with tactile sensor : Skku hand ii," in *2006 IEEE/RSJ International Conference on Intelligent Robots and Systems*, pp. 3779–3784, 2006.
- [29] D. Xu, G. E. Loeb, and J. A. Fishel, "Tactile identification of objects using bayesian exploration," in *2013 IEEE International Conference on Robotics and Automation*, pp. 3056–3061, 2013.
- [30] E. Chelly, A. Cherubini, P. Fraise, F. B. Amar, and M. Khoramshahi, "Tactile-based force estimation for interaction control with robot fingers," 2025.
- [31] I. Romanelis, V. Fotis, A. Kalogeras, C. Alexakos, K. Moustakas, and A. Munteanu, "Efficient and scalable point cloud generation with sparse point-voxel diffusion models," *arXiv preprint arXiv:2408.06145*, 2024.
- [32] B. Sundaralingam, S. K. S. Hari, A. Fishman, C. Garrett, K. Van Wyk, V. Blukis, A. Millane, H. Oleynikova, A. Handa, F. Ramos, et al., "Curobo: Parallelized collision-free robot motion generation," in *2023 IEEE International Conference on Robotics and Automation (ICRA)*, pp. 8112–8119, IEEE, 2023.
- [33] C. E. Shannon, "A mathematical theory of communication," *The Bell System Technical Journal*, vol. 27, no. 3, pp. 379–423, 1948.
- [34] S. A. Papert, "The summer vision project," 1966.