

Feel Robot Feels: Tactile Feedback Array Glove for Dexterous Manipulation

Feiyu Jia^{2,1,*} Xiaojie Niu^{1,*} Sizhe Yang^{1,3,*} Qingwei Ben^{1,3} Tao Huang^{4,1}

Feng Zhao^{2,†} Jingbo Wang^{1,†} Jiangmiao Pang^{1,†}

¹ Shanghai AI Laboratory ² University of Science and Technology of China

³ The Chinese University of Hong Kong ⁴ Shanghai Jiao Tong University

* Equal Contributions † Corresponding Authors

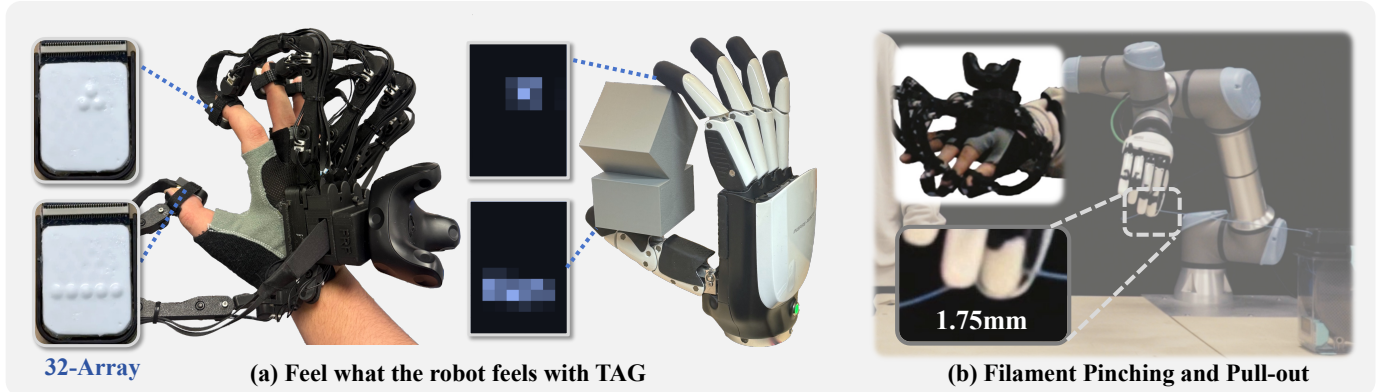


Fig. 1: (a) TAG is a Tactile Feedback Array Glove for dexterous manipulation and teleoperation. It achieves precise 21-DoF joint tracking and integrates a 32-point feedback array on each fingertip, bridging the gap between robot-side sensing and human-side perception. (b) Sequential filament pinching and pull-out.

Abstract—Teleoperation is a key approach for collecting high-quality, physically consistent demonstrations for robotic manipulation. However, teleoperation for dexterous manipulation remains constrained by: (i) inaccurate hand–robot motion mapping, which limits teleoperated dexterity, and (ii) limited tactile feedback that forces vision-dominated interaction and hinders perception of contact geometry and force variation. To address these challenges, we present TAG, a low-cost glove system that integrates precise hand motion capture with high-resolution tactile feedback, enabling effective tactile-in-the-loop dexterous teleoperation. For motion capture, TAG employs a non-contact magnetic sensing design that provides drift-free, electromagnetically robust 21-DoF joint tracking with joint angle estimation errors below 1° . Meanwhile, to restore tactile sensation, TAG equips each finger with a 32-actuator tactile array within a compact 2 cm^2 module, allowing operators to directly feel physical interactions at the robot end-effector through spatial activation patterns. Through real-world teleoperation experiments and user studies, we show that TAG enables reliable real-time perception of contact geometry and dynamic force, improves success rates in contact-rich teleoperation tasks, and increases the reliability of demonstration data collection for learning-based manipulation. Videos and code are available on our [project page](#).

I. INTRODUCTION

Dexterous manipulation is a core challenge in robotics and is increasingly addressed via learning from real-world demonstrations [32, 2, 5, 39, 41]. The success of these data-driven methods depends critically on the fidelity and coverage of the collected data [27]. This dependence has driven the development of recent teleoperation systems [36, 4, 1, 40], which offer practical and scalable interfaces for data collection.

Vision- and Virtual Reality (VR)-based teleoperation systems are widely used but are limited by visual blockage, pose estimation errors, and viewpoint misalignment [9, 4]. Glove-based approaches bypass these vision-dependent constraints through direct motion sensing, yet still face challenges in signal reliability: IMU-based systems suffer from drift over time, while contact-based mechanical sensors [40, 42] exhibit nonlinearity that increases with fatigue. Non-contact magnetic sensing [1] alleviates fatigue-related issues, but is typically limited to single-axis measurements, preventing precise angular estimation. Although commercial solutions [18] offer higher precision, they are both expensive and sensitive to electromagnetic interference. Consequently, reliable demonstration collection calls for a motion capture glove that provides both robust sensing and precise angular accuracy.

Beyond motion tracking, another critical challenge in teleoperation is the absence of haptic feedback. Without tactile cues, operators face a significant sensory gap, relying passively on vision and struggling to perceive object geometry or contact forces. While traditional pneumatic and mechanical systems offer strong feedback, their bulkiness severely restricts dexterity. Conversely, vibrotactile motors [6, 40] are lightweight but fail to render complex surface geometries. Recently, Electro-Osmotic Pumps (EEOs) have emerged as a promising solution, utilizing electro-hydrodynamics to generate high pressure in a silent, millimeter-scale form factor; however, current applications [22] are confined to tethered VR simulations. Therefore, realizing high-fidelity tactile feedback for real-

world robotic teleoperation remains an unsolved challenge.

In this paper, we introduce TAG, a low-cost glove system that features precise hand motion capture with high-resolution tactile feedback, enabling effective dexterous tactile-in-the-loop teleoperation. By deploying a non-contact magnetic tracking framework across a strategic 21-DoF configuration, our system guarantees drift-free joint estimation, enabling highly precise and stable robot manipulation.

To bridge the sensory gap, each fingertip is equipped with a compact 2 cm² module housing a 32-actuator EEOP tactile array. We optimize channel ratios and grid symmetry to ensure high fluid flux and consistent spatial resolution, while relaxing fabrication tolerances to enable handcrafted assembly. Through this high-resolution feedback module, we translate robotic data into human-perceivable cues via two distinct mapping strategies: a spatial mapping that encodes contact geometry into 2D actuator patterns for contour perception, and a pressure mapping that translates force magnitude into active contact area expansion for intuitive force feedback.

We validate TAG’s tactile-in-the-loop perceptual capabilities and its effectiveness for dexterous teleoperation through real-world experiments. Quantitative results show that TAG achieves sub-degree joint angle accuracy while remaining robust to electromagnetic interference, enabling reliable hand–robot motion retargeting. User studies further confirm that the rendered tactile feedback allows operators to consistently perceive contact geometry and dynamic force direction during manipulation. Teleoperation experiments on real robotic platforms demonstrate improved performance in contact-rich tasks that challenge vision-dominated control. Finally, imitation learning results indicate that demonstrations collected with TAG are reliable and physically consistent, making them suitable for learning-based manipulation.

The main contributions of this work are three-fold:

- We develop a robust 21-DoF magnetic tracking system that achieves drift-free, sub-degree precision for reliable demonstration collection.
- We construct high-resolution EEOP tactile modules and integrate them into real-world robotic loops, employing a dual-mode strategy to render geometry and force.
- We propose TAG and validate it on real-world challenging contact-rich tasks, demonstrating its superior capability in fine-grained teleoperation.

II. RELATED WORK

A. Data collection for Dexterous Hands

Large-scale, high-quality demonstrations are essential for learning autonomous dexterous manipulation policies [8, 12, 16]. One prevalent paradigm records human hand motions directly using vision-based pose estimation [26, 30], or wearable devices like exoskeletons and gloves [34, 7, 31]. However, vision-based methods are limited by pose estimation accuracy and are highly susceptible to occlusion, while exoskeletons are often tightly coupled to specific hand kinematics. More critically, transferring human motion to robot hands requires

complex kinematic mapping, which inevitably introduces an embodiment gap and action–vision mismatch. To ensure action–observation consistency, other works teleoperate the robot hand using vision-based pose estimation (VR [4, 17] or camera [9, 19]) or motion-capture gloves. Vision-based methods typically rely on inaccurate hand pose estimation and provide little physical contact feedback, limiting data quality in contact-rich tasks, while gloves suffer from sensor limitations. Specifically, resistive sensors [40, 42] exhibit mechanical fatigue and non-linearity; IMUs [15, 11] drift over time; and single-axis Hall sensors [1] lack precision. Commercial magnetic solutions [18] are drift-free but costly and sensitive to electromagnetic interference. TAG addresses these issues with robust, high-precision magnetic encoding and direct tactile feedback, enabling scalable, high-fidelity data collection.

B. Tactile Feedback Gloves

Tactile information is indispensable for dexterous manipulation [25]; yet, most teleoperation systems offer sparse feedback, rendering contact-rich tasks challenging despite available onboard sensing [24, 38]. To bridge this gap, prior works have introduced feedback via joint-level forces [3, 7, 20] or vibrotactile cues [40, 6]. However, these modalities suffer from inherent limitations: force feedback is predominantly kinesthetic rather than cutaneous, providing indirect cues while often restricting natural motion; vibrotactile feedback typically reduces complex contact geometries to binary signals. Similarly, electrotactile stimulation methods [29, 14] often induce unnatural sensations or suffer from instability due to varying skin impedance. While some commercial gloves achieve multi-point tactile feedback, they remain prohibitively expensive for widespread adoption. TAG embeds high-resolution electrochemical tactile modules [22] and integrates them into real-world robotic loops, employing a dual-mode strategy to render both geometry and force.

III. HARDWARE DESIGN

This section presents the hardware architecture of TAG, describing how its design principles (Sec. III-A) guide the kinematic structure (Sec. III-B) and the tight integration of high-precision joint tracking (Sec. III-C) with high-density tactile feedback (Sec. IV-B) for dexterous teleoperation.

A. Design Principles

1) **Efficient:** TAG achieves sub-degree joint tracking accuracy ($<0.8^\circ$) with strong electromagnetic interference (EMI) resilience using non-contact magnetic encoders. Combined with a 32-point tactile feedback array, the system enables low-latency, high-fidelity teleoperation.

2) **Cross-platform:** TAG interfaces seamlessly with diverse robotic hands, tactile sensing schemes, and end-effector pose tracking methods, enabled by its high kinematic freedom and tactile sensing fidelity.

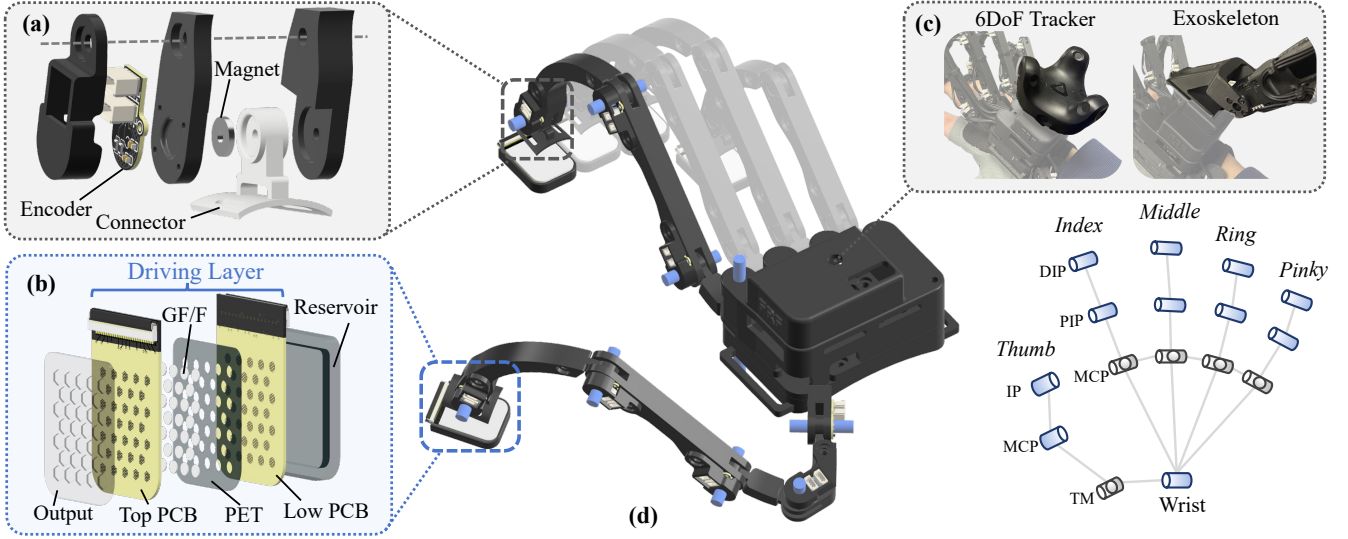


Fig. 2: Hardware architecture of TAG. (a) Exploded view of the joint encoder, utilizing a magnetometer and ring magnet for angle sensing. (b) Exploded view of the tactile feedback module, a multi-layer EOP stack sandwiching a membrane between PCBs. (c) TAG is compatible with diverse arm tracking solutions, such as 6DoF trackers or exoskeletons. (d) Kinematic configuration distributing 21 encoders for full-hand motion capture.

3) **Low-cost:** Existing commercial haptic gloves are typically expensive (over \$5,000) for widespread research [10, 18]. In contrast, TAG lowers this barrier by offering accurate full-hand motion capture and high-resolution tactile feedback at a total cost below \$500.

4) **Open-source:** To foster community-driven research, TAG is fully open-source. We provide comprehensive CAD models, PCB designs, and fabrication guides, ensuring that the system can be easily reproduced using standard materials.

B. Kinematic Design

To achieve anthropomorphic and high-fidelity hand tracking, the kinematic structure of TAG must closely replicate the biological degrees of freedom (DoF) of a human hand. This high-fidelity capture is essential for acquiring precise demonstrations for dexterous manipulation tasks. Based on this principle, we implement a 21-DoF configuration by strategically distributing 21 magnetic encoder modules across all key anatomical joints, as illustrated in Fig. 2(d).

For the four fingers (index to pinky), each is equipped with four encoders to monitor the Distal Interphalangeal (DIP), Proximal Interphalangeal (PIP), and Metacarpophalangeal (MCP) joints. Specifically, the MCP tracking is decoupled into flexion/extension and adduction/abduction to capture the complex spreading and grasping motions of the palm.

The thumb, as the most critical component for dexterous manipulation, requires an expansive range of motion to facilitate opposition, circumduction, and fine-grained pinching. Consequently, we assign 5 DoF to the thumb by deploying one encoder at the Interphalangeal (IP) joint, one at the Metacarpophalangeal (MCP) joint, and three encoders at the complex Trapeziometacarpal (TM) joint to precisely track its three-dimensional rotation. This 21-DoF design ensures a precise mapping between the operator and diverse robotic end-effectors, effectively minimizing information loss during

complex teleoperation tasks.

C. Joint Angle Encoders

Precise joint tracking is fundamental to effective teleoperation. Unlike conventional designs that rely on potentiometers [40] or single-axis Hall sensors [1], which are susceptible to mechanical wear and installation misalignment, TAG adopts a robust non-contact sensing approach utilizing 21 MLX90393 magnetometers. Selected for its high-performance Triaxis technology, this sensor resolves the magnetic flux density vector with a built-in 16-bit ADC, offering high linearity (typically $\pm 0.1\%$) and direct digital output that eliminates complex external circuitry.

As illustrated in Fig. 2(a), a miniature, diametrically magnetized Neodymium Iron Boron (NdFeB) ring magnet is positioned directly above the sensor, aligned parallel to the sensor’s XY plane. As the joint rotates, the sensor captures the orthogonal components of the magnetic flux density vector B . This relationship is modeled as:

$$\begin{cases} B_x = B_0 \cos(\theta) + O_x \\ B_y = B_0 \sin(\theta) + O_y \end{cases}, \quad (1)$$

where B_0 is the magnetic flux amplitude, θ is the rotation angle, and $O_{x,y}$ represent magnetic offsets induced by environmental bias. In the TAG encoder, these offsets are calibrated by rotating the joint through its full range:

$$O_i = (B_{i,\max} + B_{i,\min})/2, \quad i \in \{x, y\}. \quad (2)$$

The angular position is then analytically derived using the four-quadrant arctangent function:

$$\theta = \text{atan2}(B_y - O_y, B_x - O_x). \quad (3)$$

Crucially, since this calculation relies on the ratio of field components, common-mode errors, such as those induced by temperature changes, are naturally cancelled out.

Since the mechanical structure of the glove is fixed, we can accurately calculate the 3D pose of each fingertip and retarget these to various dexterous hands. Detailed implementation of the retargeting algorithms is provided in the [Appendix](#).

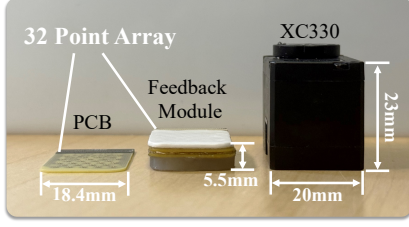


Fig. 3: Physical dimensions and form factor of the TAG feedback module. The entire assembly measures only $29 \times 18.4 \times 5.5$ mm.

D. Tactile Feedback Module

1) **Module Design:** To deliver high-fidelity, multi-point tactile feedback within a compact 2 cm^2 fingertip area, we develop a tactile feedback module driven by Electro-Osmotic Pumps (EOPs). The module’s architecture is conceptually similar to prior works [23, 22, 37] and consists of three primary layers: an output layer, a driving layer, and a reservoir layer. The output and reservoir components are cast from Ecoflex™ 00-30 using 3D-printed precision molds. The whole module is illustrated in [Fig. 3](#).

The core driving mechanism, illustrated in [Fig. 2\(b\)](#), utilizes a glass fiber (GF/F) membrane sandwiched between dual-layer PCBs. The TAG tactile feedback module is powered by a 12V source (either a battery or external supply), which is stepped up to the required high voltage 200 V via a boost converter. By modulating the potential difference applied to each independent channel, we enable Tri-state control for each taxel: active *protrusion* (+), *neutral* (0), and *retraction* (-).

Building on this architecture, our implementation introduces targeted design optimizations to enhance hydraulic efficiency, spatial resolution, and manufacturability:

- **Enhanced hydraulic efficiency:** We enlarge the dimensions of the *vias* to 0.25/0.4 mm (inner/outer). With 12 vias arranged within each 1.6 mm-diameter feedback taxel, this design yields a conductive-to-surface area ratio of $\sim 29.3\%$. This optimization significantly enhances hydraulic efficiency, enabling us to lower the driving voltage from 250 V to 200 V while maintaining high-quality fluid flux, thereby effectively reducing power consumption.
- **Optimized spatial perception:** Prioritizing alignment with robotic sensing profiles over anatomical mimicry, we adopt a centrally symmetric layout. Unlike human fingertips which taper distally, robotic end-effectors typically feature uniform sensing arrays; our symmetric design ensures spatial mapping and consistent tactile resolution across the active area.
- **Simplified manufacturing:** We intentionally increase the inter-node spacing to relax fabrication tolerances. This geometric adjustment eliminates the dependency on high-precision industrial tools like laser cutters, ensuring the entire assembly can be handcrafted.

Comprehensive details regarding the fabrication process and electronic control circuitry are provided in the [Appendix](#).

2) **Tactile Feedback Mapping:** To bridge the gap between robot-side sensing and human-side perception, we develop two distinct mapping strategies for the 32-point tactile array, tailored to different task requirements.

Shape Mapping Mode: Intended for tasks demanding high geometric discrimination, this mode preserves the spatial topology of the contact. Raw tactile data from the end-effector is processed via bilinear interpolation and normalized to align with the staggered layout of the 32-point tactile feedback array. The mapping function maintains the spatial distribution of the input—distinguishing, for instance, between the diffuse patch of a broad contact and the narrow profile of a sharp edge. As illustrated in [Fig. 7](#), the activation state of each array cell $k \in \{0, \dots, 31\}$ is driven by the normalized intensity at its corresponding mapped coordinate, ensuring a direct spatial correspondence for perceiving object shapes.

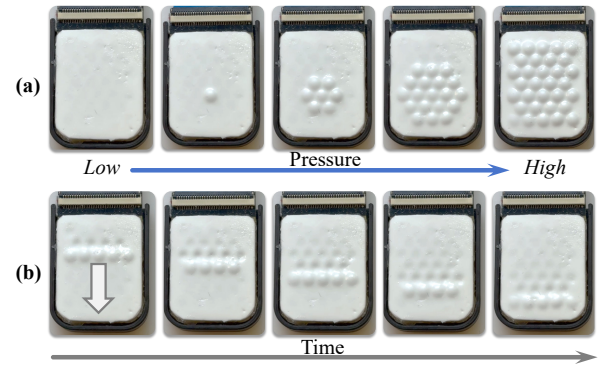


Fig. 4: Dynamic demonstration of the TAG feedback module. (a) Active area expansion in Pressure Mode under varying pressure levels. (b) Dynamic force profiles simulating forward-to-backward motion.

Pressure Mapping Mode: This mode prioritizes the perception of force magnitude by converting intensity into contact area. Rather than a direct spatial map, the system monitors the normalized peak pressure $P_{max} \in [0, 1]$ from the sensor array. This value triggers a discrete, multi-level area expansion based on configurable thresholds (e.g., $T = \{0.1, 0.35, 0.7\}$):

$$\mathcal{I}_{active} = \{k \mid k \in \text{Indices}(T_n) \text{ if } P_{max} \geq T_n\}. \quad (4)$$

As P_{max} surpasses each threshold, the active area expands radially from a single central point to 7 or 19 points. As shown in [Fig. 4\(a\)](#), this stepwise expansion creates an intuitive *tactile scale*, allowing the operator to gauge pressure levels through changes in the perceived contact area.

E. Arm Pose Acquisition

TAG is designed to be highly compatible with diverse teleoperation approaches for robot arms by simply adjusting its mechanical mounting adapters. In our experiments, we implement two primary methods for capturing wrist pose: utilizing the Forward Kinematics (FK) of the HOMIE exoskeleton [1] and obtaining pose with HTC VIVE Trackers [28]. This modular interfacing ensures that TAG can be readily integrated

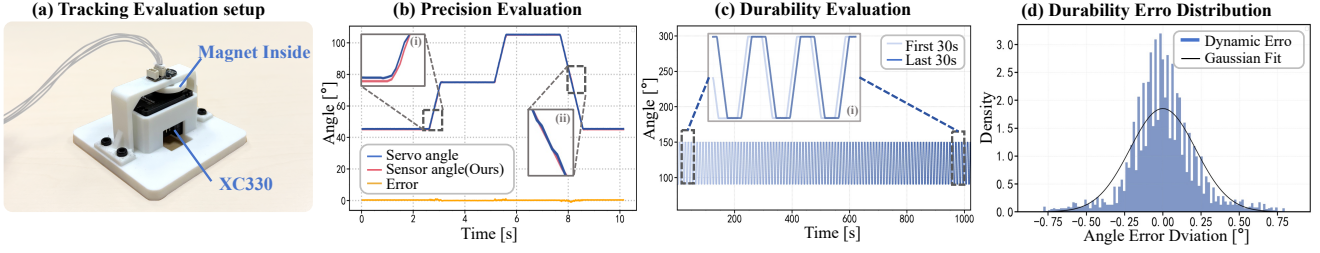


Fig. 5: Exoskeleton glove tracking performance. (a) Experimental setup for joint-tracking. (b) Real-time tracking results for discrete and continuous trajectories. (c) Long-term stability test over 1000s of operation. (d) Statistical distribution of instantaneous tracking errors with Gaussian fit.

into various frameworks, including vision-based or VR-based systems, to meet specific user requirements.

IV. HARDWARE EVALUATION

In this section, we first evaluate the exoskeleton glove by characterizing its motion-tracking precision and robustness to EMI in Sec. IV-A. With reliable glove-based teleoperation established, we then characterize the tactile sensing performance of diverse robotic platforms in Sec. IV-B to inform hardware selection and validate our tactile mapping strategies.

A. Glove Exoskeleton Evaluation

1) **High-Precision Joint Tracking:** To quantify the joint-tracking accuracy of the TAG encoder, we use a Dynamixel XC330-M288-T servo (0.088° resolution) as a high-precision ground-truth reference, following established benchmarking protocols [21, 43]. As illustrated in Fig. 5(a), a magnetic encoder module is mounted via a custom adapter to replicate its integration within the glove. The servo is commanded to execute a predefined angular trajectory ($45^\circ \rightarrow 75^\circ \rightarrow 105^\circ \rightarrow 45^\circ$), spanning a typical range of finger joint motion.

The tracking results in Fig. 5(b) demonstrate that TAG achieves rapid response and high-fidelity tracking during both (i) discrete angle transitions and (ii) continuous motion. Across the entire sequence, the maximum tracking error remains within $\pm 0.35^\circ$, demonstrating sub-degree joint-tracking precision suitable for high-fidelity hand motion capture.

2) **Long-Term Tracking Stability:** Beyond instantaneous accuracy, sustained teleoperation also requires that such accuracy be maintained over extended periods. To evaluate long-term stability, the encoder is driven by the XC330 servo under prolonged reciprocating motion for 1000s. As illustrated in Fig. 5(c), a comparison between the initial and final 30s intervals shows highly consistent angular estimates, with an average discrepancy of only $\sim 0.02^\circ$, indicating negligible drift over extended operation.

Beyond evaluating start-end drift, we aim to characterize the distribution and temporal consistency of instantaneous tracking errors over the entire duration. We analyze the statistical characteristics of these errors, and the corresponding Gaussian fit is shown in Fig. 5(d). The distribution is tightly centered at 0° , with a standard deviation of $\sigma = 0.215^\circ$, and the vast majority of samples fall within $\pm 0.8^\circ$. These results demonstrate stable and robust joint tracking during long-duration operation.

3) **Strong EMI Resilience:** While tracking accuracy is critical for sustained operation, teleoperation systems also need to remain reliable in electromagnetically complex environments, making robustness to electromagnetic interference (EMI) essential. To this end, we evaluate the EMI resilience of TAG by measuring motion-tracking stability while approaching an active PC chassis, a representative source of electromagnetic interference, and compare its performance against the commercial Manus glove [18]. During the experiment, the operator gradually moves the index finger from a distance of 50 cm to 5 cm relative to the interference source.

As shown in Fig. 6(b), the Manus glove deviates noticeably during the approach phase (ii), reaching up to 5.69° . In contrast, TAG maintains stable tracking during approaching, with a maximum deviation of only 0.24° . These results highlight the superior EMI resilience of TAG, enabling reliable motion capture in electronics-intensive teleoperation scenarios.

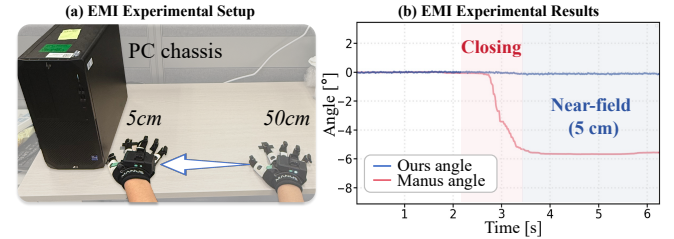


Fig. 6: EMI resilience evaluation. (a) Setup for near-field interference testing. (b) Angular stability comparison between the Manus glove and TAG.

B. Tactile Sensing Characterization

Following the tactile mapping methodology established in Sec. III-D, practical implementation requires a characterization of the end-effector sensing capabilities. This evaluation is crucial not only for calibrating mapping parameters (e.g., gain and threshold) to ensure consistent feedback but also for selecting the most appropriate hardware platform for specific validation tasks. We characterize four distinct platforms: the custom resistive *LeapTac* array [13] (integrated into *LeapHand* [21]), the commercial resistive *Inspire Hand*, the electromagnetic *XHand*, and the vision-based *GelSight Mini*.

The quantitative evaluation results are summarized in Tab. I. Based on these performance metrics, we assign specific platforms to corresponding validation experiments to maximize system efficacy. As indicated in the table, the *Inspire Hand* possesses a contiguous sensing area with sufficient spatial resolution (12×8) to reliably distinguish basic geometric

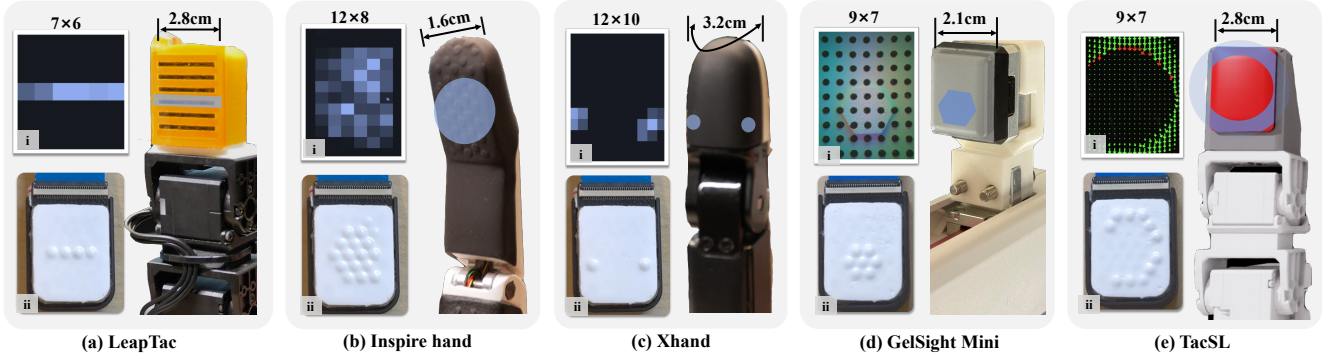


Fig. 7: Tactile sensing and shape mapping performance across various end-effectors. Comparison of raw sensing inputs (i) and physical rendering outputs on TAG (ii) across various robot end-effectors (right). TAG effectively translates diverse sensor geometries into coherent tactile feedback: (a) LeapTac: linear ridge; (b) Inspire Hand: circular area; (c) XHand: dual contact points; (d) GelSight Mini and (e) TacSL: complex shapes.

TABLE I: Comparative Analysis of Tactile Sensing Performance. Res: Spatial Resolution. **Thres:** Minimum Force Threshold. **Var.** indicates sensitivity is tunable via gel stiffness. **EM:** Electromagnetic.

End-effector	Sensing Principle	Spatial Res.	Sensing Thres.(N)	Basic Shape	Complex Shape
LeapTac	Resistive	7×6	0.14	✓	✗
Inspire	Resistive	12×8	0.38	✓	✗
XHand	EM-based	12×10	0.08	✗	✗
GelSight Mini	Vision	9×7	Var.	✓	✓

primitives (point, line, plane). Consequently, we select this platform to conduct the user studies detailed in Sec. V-A, as these tasks require stable spatial pattern reconstruction.

Conversely, the XHand features a fingertip-wrapping sensor. While this geometry limits the frontal contact area for planar differentiation (✗), it provides multi-directional sensing and exceptional sensitivity. Therefore, the XHand is assigned to the delicate Teleoperation Tasks in Sec. V-B, ensuring that even minimal interaction forces are captured and rendered.

Finally, while the GelSight Mini captures precise geometric shapes with resultant forces mapped to a 9×7 grid, and the custom LeapTac provides a cost-effective baseline, the integration of these diverse sensors validates the universality of our mapping framework. As illustrated in Fig. 7, by applying the proposed shape mapping mode, raw data from these varying modalities is successfully normalized into coherent tactile stimuli on the TAG glove.

V. TELEOPERATION EXPERIMENTS

In this section, we conduct real-world teleoperation experiments using TAG on two different platforms: the Unitree G1 equipped with Inspire Hand, and the UR5e integrated with XHand. These experiments aim to demonstrate the capabilities and efficiency of TAG in contact-rich teleoperation tasks.

A. User Study

To demonstrate how TAG empowers operators to “feel what the robot feels” and accurately discriminate the contact conditions, we conduct a user study with five trained participants. The study comprises two parts: contact shape discrimination

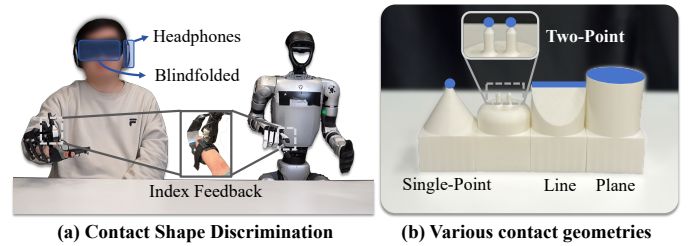


Fig. 8: User Study. (a): The operator, blindfolded and wearing headphones, identifies the shape of the object pressed against the index finger of G1 solely through tactile feedback. (b): Four objects with different contact geometries: single-point, two-point, line, and plane.

and dynamic force direction discrimination, both conducted on the G1 robot with an Inspire Hand.

1) Contact Shape Discrimination:

Setup: As illustrated in Fig. 8, the operator wears the TAG device and teleoperates G1 (equipped with the Inspire Hand), holding the arm raised with the index finger extended. An experimenter presses on the Inspire Hand’s index finger using four objects with various geometries (specifically: single-point, two-point, line, and plane) for 5s. The operator, while blindfolded and wearing noise-canceling headphones, is required to identify the perceived objects based solely on the tactile feedback provided by the TAG device to their index finger. The typical tactile feedback patterns are shown in Fig. 7. A total of five participants are recruited, each performing 20 trials. The presentation order of the four objects is fully randomized, with each object being presented five times to each participant.

Results: The recognition performance for contact shape discrimination is visualized in Fig. 9. The system demonstrates high fidelity in rendering spatial contact geometries particularly for **single-point** contacts, which are identified with a 100% accuracy. Minor misidentifications primarily occur between geometrically similar categories; for instance, **line** contacts are occasionally perceived as **two-point** ($n = 2$) or **plane** ($n = 2$) contacts. Similarly, **two-point** and **plane** geometries achieve high recognition rates, with 23 correct identifications out of 25 trials each. The strong alignment between the actual and perceived geometries validates the effectiveness of the TAG device in conveying distinct contact

patterns to the operator.

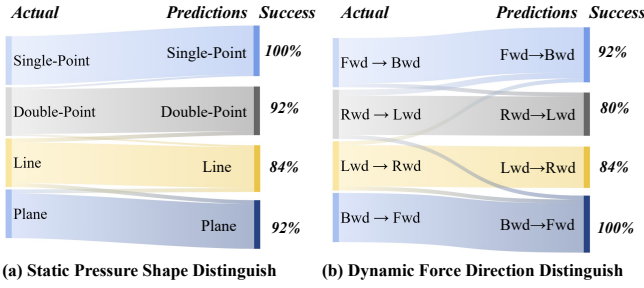


Fig. 9: Classification results of the user study. The Sankey diagrams visualize the mapping between actual stimuli and user predictions. The width of the flows represents the number of trials.

2) Dynamic Force Direction Discrimination:

Setup: The experimental setup follows the same protocol as the contact shape experiment. In this session, the experimenter applies a dynamic load by rolling a cylinder across the fingertip of the Inspire Hand for approximately 2s per trial. Four distinct motion directions are evaluated: **forward-to-backward**, **backward-to-forward**, **left-to-right**, and **right-to-left**. The dynamic evolution of these tactile patterns is illustrated in Fig. 4(b). Participants are required to determine the direction of the rolling force based solely on the dynamic tactile cues rendered by the TAG device. All other experimental conditions remain consistent with the contact shape discrimination study.

Results: As summarized in Fig. 9, superior discriminability is observed for longitudinal motions, with participants achieving 100% accuracy for **Bwd** → **Fwd** and 92% (23/25) for **Fwd** → **Bwd**. Lateral recognition is slightly lower, with **R** → **L** and **L** → **R** correctly identified in 20 and 21 trials, respectively. This disparity stems from the physical aspect ratio of the TAG module; the narrower lateral dimension accommodates fewer sensing and actuation points compared to the longitudinal axis, resulting in reduced resolution for side-to-side movements. Nonetheless, the overall high recognition rates confirm that TAG effectively renders continuous directional cues, providing operators with intuitive motion perception.

B. Teleoperation Performance

To evaluate TAG in complex contact-rich scenarios, we conduct three real-world teleoperation tasks using a UR5e arm with an XHand. These fine-motor tasks primarily involve the coordinated control of the index finger and thumb.

1) Multimodal Control of a Smart Desk Lamp:

Tasks: As shown in Fig. 10(a), this task requires hierarchical control of a smart lamp via a single rotary knob. To ensure precise force regulation, TAG is configured in **Press Mode**, mapping detected pressure to the feedback contact area. The task is decomposed into three stages:

- **Stage 1: Power Toggling (Short Press)** – A short press (< 2s) toggles the power. Precise temporal control is essential as exceeding 2s triggers mode switching.

- **Stage 2: CCT Adjustment (Press & Turn)** – The operator must maintain a steady downward force exceeding 3N while rotating the knob to adjust the color temperature.
- **Stage 3: Brightness Adjustment (Turn Only)** – Rotation requires a subtle downward force within a narrow window (0.5N–3N). The force must be sufficient for friction but light enough to avoid accidentally triggering the “press” mode in Stage 2.

Results: Results are summarized in Tab. II. In **Stage 1**, although feedback slightly increases time (3.6s vs. 3.2s), the tactile “click” confirms actuation, ensuring a 100% success rate. In **Stage 2**, array feedback allows operators to maintain stable pressure (>3N) by perceiving the full contact area. The benefit is most evident in **Stage 3**. Regulating force (0.5–3N) is difficult with single-point haptics, often leading to over-pressing (>3N). In contrast, TAG renders the *incremental expansion* of the contact patch, providing intuitive confirmation at minimal force levels to safely engage the knob.

TABLE II: Results of multimodal desk lamp control.

Teleoperation System	Success Rate (↑)	Average Time (↓)
Stage I: Power Toggling (Short Press)		
HOMIE [1]	12/15	6.1s
TAG w/o feedback	14/15	3.2s
TAG w/ feedback	15/15	3.6s
Stage II: CCT Adjustment (Press & Turn)		
HOMIE [1]	9/15	23.5s
TAG w/o feedback	12/15	14.8s
TAG w/ feedback	14/15	8.4s
Stage III: Brightness Adjustment (Turn Only)		
HOMIE [1]	1/15	31.3s
TAG w/o feedback	4/15	26.2s
TAG w/ feedback	12/15	15.7s

2) Filament Pinching and Pull-out:

Tasks: As shown in Fig. 10(b), this task requires the operator to pinch and extract a 1.75mm diameter 3D printing filament in two consecutive trials. Each trial requires a stable pull exceeding 20cm in length. With TAG in **Shape Mode**, the operator must distinguish between an *empty pinch* (diffuse tactile feedback area) and a *successful capture* (narrow tactile feedback area), confirming the filament is securely centered before initiating the long pull.

Results: The results are summarized in Tab. III. In the absence of feedback, operators frequently fail either midway through the pulling motion or during a second attempt due to tactile ambiguity, often initiating the pull before the filament is securely centered. In contrast, TAG provides distinct geometric differentiation. The transition to a sharp, narrow tactile feedback area serves as an intuitive “trigger signal,” ensuring secure engagement for the long-distance pull. The precision feedback from TAG allows for an 86.7% (13/15) success rate and reduces completion time to 22.4s.

3) Multi-level Spring Compression:

Tasks: As shown in Fig. 10(c), this task evaluates the precise control of force magnitude using a spring-loaded

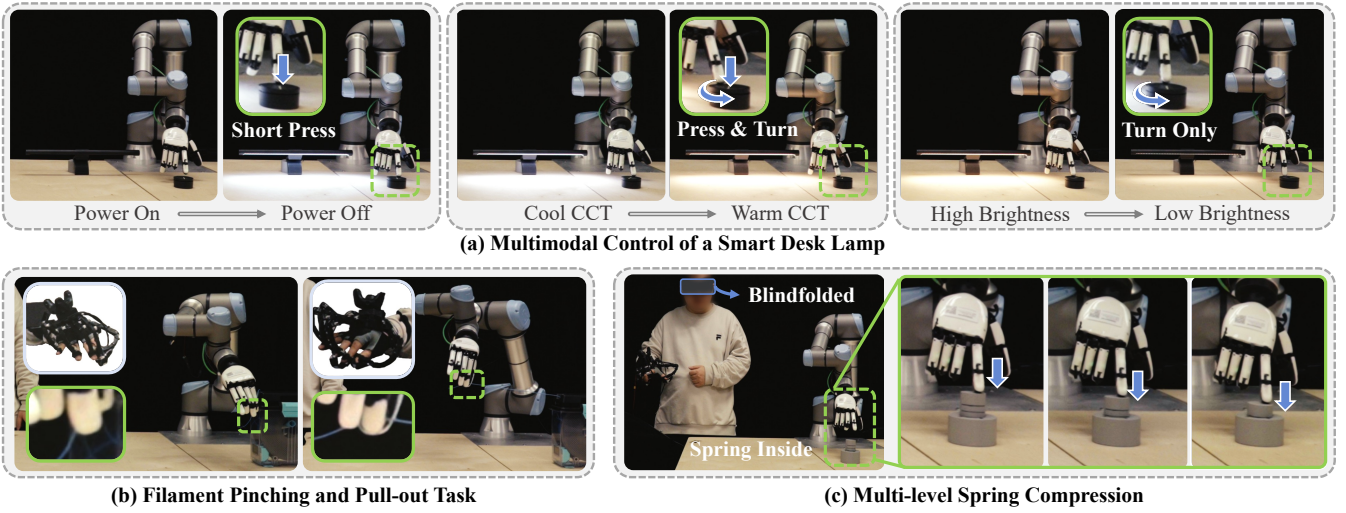


Fig. 10: Real-world teleoperation tasks using TAG. (a) **Multimodal lamp control:** Hierarchical interaction via short press, press-and-turn, and subtle rotation. (b) **Sequential filament pinching:** Distinguishing successful capture of a 1.75mm filament from empty pinches via localized tactile feedback. (c) **Multi-level spring compression:** Maintaining three distinct pressure levels (0.5N, 3.5N, 7.5N) guided by rendered tactile feedback.

TABLE III: Results of sequential filament pinching (2 consecutive trials).

Teleoperation System	Success Rate ($\uparrow$)	Average Time ($\downarrow$)
HOMIE [1]	5/15	42.8s
TAG w/o feedback	7/15	35.6s
TAG w/ feedback	13/15	22.4s

button. TAG is configured in **Press Mode**. Operators must sequentially reach and maintain three pressure levels for $>2s$: *Initial Contact* (0.5N), *Half-compressed* (3.5N), and *Fully-compressed* (7.5N). Tests are conducted under two conditions, blindfolded and with visual feedback, to assess the system’s ability to substitute visual cues with tactile feedback.

Results: As shown in Tab. IV, baseline systems under blindfolded conditions yield success rates $\leq 20\%$, with operators noting these few successes are purely stochastic and lack repeatability. Without tactile feedback, operators struggle to distinguish pressing force, resulting in unreliable proprioceptive cues. In contrast, TAG enables the operators to achieve an 86.7% success rate by rendering the incremental expansion of the contact patch as intuitive feedback. This allows operators to “lock” onto specific force levels with high confidence. Supplementing visual feedback with tactile feedback via TAG increases the success rate from 12/15 to 15/15, confirming that TAG provides a more immediate and precise signal for force maintenance than visual estimation of spring displacement.

TABLE IV: Success rate of multi-level spring compression.

Teleoperation System	Blindfolded ($\uparrow$)	With Visual ($\uparrow$)
HOMIE [1]	2/15	10/15
TAG w/o feedback	3/15	12/15
TAG w/ feedback	13/15	15/15

C. Imitation Learning

We train a manipulation policy via imitation learning to verify the usability of data collected through TAG teleoperation.

ManiFlow [35] is employed as the policy. ManiFlow takes point clouds and the robot state as input and outputs actions. The actions are defined as the end-effector pose of the robot arm and the joint positions of the robot hand. As illustrated in Fig. 11(a), we use a RealSense D455 camera to capture scene point clouds. Tactile information is included as part of the robot state input to the policy. We evaluate the learned policy on two representative tasks: CCT Adjustment and Full Spring Compression, as shown in Fig. 11(b). The objects are randomly placed within a 0.1×0.1 m area. The policy trained on data collected via TAG achieves success rates of up to 15/15 and 13/15 on the two tasks, demonstrating the effectiveness of TAG data collection.

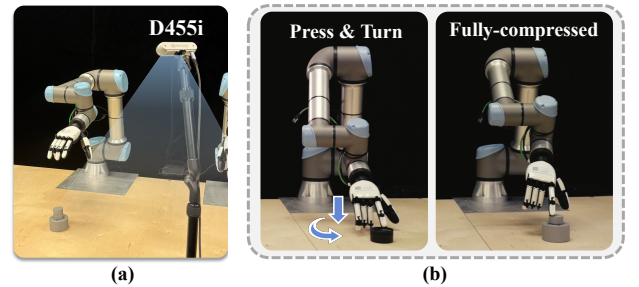


Fig. 11: Imitation learning. (a) Experimental setup for real-world tasks. (b) Task illustration: CCT Adjustment and Fully Spring Compression.

VI. CONCLUSION AND LIMITATIONS

In this paper, we introduce TAG, a low-cost glove system integrating precise motion capture with high-resolution tactile feedback. By leveraging 21-DoF non-contact 3-axis magnetic sensors, our system achieves sub-degree precision and interference-resistant joint tracking. To bridge the sensory gap, we incorporate compact EEOP arrays that render tactile cues via dual mapping strategies, effectively encoding contact shapes and pressure magnitudes. Real-world experiments validate that TAG empowers operators to perform

delicate, contact-rich tasks with superior situational awareness compared to baselines. Ultimately, this approach significantly increases the reliability of demonstration data collection for learning-based manipulation.

Limitations Despite its capabilities, TAG has limitations. First, while our EEOP modules provide high-resolution pressure and geometry cues, they lack kinesthetic feedback to physically restrict finger motion. Furthermore, although dynamic shear can be approximated through temporal modulation, the system cannot render static shear forces necessary for detecting steady-state slip. Second, the manual assembly of fluidic circuits introduces variability and potential durability concerns. Future work will focus on hybrid actuation and standardized micro-fabrication to address these issues.

ACKNOWLEDGMENTS

This work is supported by the Shanghai Artificial Intelligence Laboratory. We thank Zhongwei Xiang, Jingping Wang, Huayi Wang, and Xiao Chen for their helpful assistance and valuable discussions. We are grateful to Shenghan Zhang and Quanli Xuan for their support in hardware development, and Yanming Shao for coordinating the shared use of the Xhand.

REFERENCES

- [1] Qingwei Ben, Feiyu Jia, Jia Zeng, Juntong Dong, Dahua Lin, and Jiangmiao Pang. Homie: Humanoid loco-manipulation with isomorphic exoskeleton cockpit. In *Robotics Science and Systems (RSS)*, 2025.
- [2] Kevin Black, Noah Brown, Danny Driess, Adnan Esmail, Michael Equi, Chelsea Finn, Niccolo Fusai, Lachy Groom, Karol Hausman, Brian Ichter, et al. π_0 : A vision-language-action flow model for general robot control. *arXiv preprint arXiv:2410.24164*, 2024.
- [3] Dapeng Chen, Haojun Ni, Juncheng Lou, Lianshun Shen, Zhong Wei, Jia Liu, and Aiguo Song. Design and evaluation of a compact force feedback glove for virtual reality applications. *Sensors and Actuators A: Physical*, 2026.
- [4] Xuxin Cheng, Jialong Li, Shiqi Yang, Ge Yang, and Xiaolong Wang. Open-television: Teleoperation with immersive active visual feedback. In *Conference on Robot Learning (CoRL)*, 2024.
- [5] Cheng Chi, Zhenjia Xu, Siyuan Feng, Eric Cousineau, Yilun Du, Benjamin Burchfiel, Russ Tedrake, and Shuran Song. Diffusion policy: Visuomotor policy learning via action diffusion. *The International Journal of Robotics Research (IJRR)*, 2025.
- [6] Runyu Ding, Yuzhe Qin, Jiyue Zhu, Chengzhe Jia, Shiqi Yang, Ruihan Yang, Xiaojuan Qi, and Xiaolong Wang. Bunny-visionpro: Real-time bimanual dexterous teleoperation for imitation learning. In *International Conference on Intelligent Robots and Systems (IROS)*, 2025.
- [7] Hao-Shu Fang, Branden Romero, Yichen Xie, Arthur Hu, Bo-Ruei Huang, Juan Alvarez, Matthew Kim, Gabriel Margolis, Kavya Anbarasu, Masayoshi Tomizuka, et al. Dexop: A device for robotic transfer of dexterous human manipulation. *arXiv preprint arXiv:2509.04441*, 2025.
- [8] Ying Feng, Hongjie Fang, Yinong He, Jingjing Chen, Chenxi Wang, Zihao He, Ruonan Liu, and Cewu Lu. Learning dexterous manipulation with quantized hand state. *arXiv preprint arXiv:2509.17450*, 2025.
- [9] Ankur Handa, Karl Van Wyk, Wei Yang, Jacky Liang, Yu-Wei Chao, Qian Wan, Stan Birchfield, Nathan Ratliff, and Dieter Fox. Dexipilot: Vision-based teleoperation of dexterous robotic hand-arm system. In *International Conference on Robotics and Automation (ICRA)*, 2020.
- [10] HaptX. Haptx official website, 2026. URL <https://haptx.com/>.
- [11] A. Harrison, A. Jester, S. Mouli, A. Fratini, and A. Jabran. Systematic evaluation of IMU sensors for application in smart glove system for remote monitoring of hand differences. *Sensors*, 2024.
- [12] Liang Heng, Haoran Geng, Kaifeng Zhang, Pieter Abbeel, and Jitendra Malik. Vitacformer: Learning cross-modal representation for visuo-tactile dexterous manipulation. *arXiv preprint arXiv:2506.15953*, 2025.
- [13] Binghao Huang, Yixuan Wang, Xinyi Yang, Yiyue Luo, and Yunzhu Li. 3d-vitac: Learning fine-grained manipulation with visuo-tactile sensing. In *Conference on Robot Learning (CoRL)*, 2024.
- [14] Chutian Jiang, Yinan Fan, Junan Xie, Emily Kuang, Kaihao Zhang, and Mingming Fan. Designing unobtrusive modulated electrotactile feedback on fingertip edge to assist blind and low vision (blv) people in comprehending charts. In *Conference on Human Factors in Computing Systems*, 2024.
- [15] Shuang Li, Jiaxi Jiang, Philipp Ruppel, Hongzhuo Liang, Xiaojian Ma, Norman Hendrich, Fuchun Sun, and Jianwei Zhang. A mobile robot hand-arm teleoperation system by vision and imu. In *International Conference on Intelligent Robots and Systems (IROS)*, 2020.
- [16] Toru Lin, Yu Zhang, Qiyang Li, Haozhi Qi, Brent Yi, Sergey Levine, and Jitendra Malik. Learning visuotactile skills with two multifingered hands. In *International Conference on Robotics and Automation (ICRA)*, 2025.
- [17] Chenhao Lu, Xuxin Cheng, Jialong Li, Shiqi Yang, Mazeyu Ji, Chengjing Yuan, Ge Yang, Sha Yi, and Xiaolong Wang. Mobile-television: Predictive motion priors for humanoid whole-body control. In *International Conference on Robotics and Automation (ICRA)*, 2025.
- [18] MetaGloves. Metagloves official website, 2026. URL <https://manus-meta.com/>.
- [19] Yuzhe Qin, Wei Yang, Binghao Huang, Karl Van Wyk, Hao Su, Xiaolong Wang, Yu-Wei Chao, and Dieter Fox. Anyteleop: A general vision-based dexterous robot arm-hand teleoperation system. In *Robotics Science and Systems (RSS)*, 2024.
- [20] Branden Romero, Hao-Shu Fang, Pulkit Agrawal, and Edward Adelson. Eyesight hand: Design of a fully-actuated dexterous robot hand with integrated vision-based tactile sensors and compliant actuation. In *Inter-*

national Conference on Intelligent Robots and Systems (IROS), 2024.

- [21] Kenneth Shaw, Ananye Agarwal, and Deepak Pathak. Leap hand: Low-cost, efficient, and anthropomorphic hand for robot learning. In *Robotics Science and Systems (RSS)*, 2023.
- [22] Vivian Shen, Tucker Rae-Grant, Joe Mullenbach, Chris Harrison, and Craig Shultz. Fluid reality: High-resolution, untethered haptic gloves using electroosmotic pump arrays. In *ACM Symposium on User Interface Software and Technology*, 2023.
- [23] Craig Shultz and Chris Harrison. Flat panel haptics: Embedded electroosmotic pumps for scalable shape displays. In *Conference on Human Factors in Computing Systems*, 2023.
- [24] Zilin Si and Wenzhen Yuan. Taxim: An example-based simulation model for gelsight tactile sensors. *Robotics and Automation Letters (RA-L)*, 2022.
- [25] Zilin Si, Gu Zhang, Qingwei Ben, Branden Romero, Zhou Xian, Chao Liu, and Chuang Gan. Diff tactile: A physics-based differentiable tactile simulator for contact-rich robotic manipulation. In *International Conference on Learning Representations (ICLR)*, 2024.
- [26] Aravind Sivakumar, Kenneth Shaw, and Deepak Pathak. Robotic telekinesis: Learning a robotic hand imitator by watching humans on youtube. In *Robotics Science and Systems (RSS)*, 2022.
- [27] Generalist AI Team. Gen-0: Embodied foundation models that scale with physical interaction. *Generalist AI Blog*, 2025. <https://generalistai.com/blog/nov-04-2025-GEN-0>.
- [28] VIVE Tracker. Vive official website, 2026. URL <https://vive.com/>.
- [29] Sebastian Vizcay, Panagiotis Kourtesis, Ferran Arge-laguet, Claudio Pacchierotti, and Maud Marchal. Electrotactile feedback for enhancing contact information in virtual reality. *International Conference on Artificial Reality and Telexistence and Eurographics Symposium on Virtual Environments*, 2021.
- [30] Chen Wang, Linxi Fan, Jiankai Sun, Ruohan Zhang, Li Fei-Fei, Danfei Xu, Yuke Zhu, and Anima Anand-kumar. Mimicplay: Long-horizon imitation learning by watching human play. In *Conference on Robot Learning (CoRL)*, 2023.
- [31] Chen Wang, Haochen Shi, Weizhuo Wang, Ruohan Zhang, Li Fei-Fei, and C Karen Liu. Dexcap: Scalable and portable mocap data collection system for dexterous manipulation. In *Robotics Science and Systems (RSS)*, 2024.
- [32] Ruoshi Wen, Guangzeng Chen, Zhongren Cui, Min Du, Yang Gou, Zhigang Han, Liqun Huang, Mingyu Lei, Yunfei Li, Zhuohang Li, et al. Gr-dexter technical report. *arXiv preprint arXiv:2512.24210*, 2025.
- [33] Philipp Wu, Yide Shentu, Zhongke Yi, Xingyu Lin, and Pieter Abbeel. Gello: A general, low-cost, and intuitive teleoperation framework for robot manipulators. In *International Conference on Intelligent Robots and Systems (IROS)*, 2024.
- [34] Mengda Xu, Han Zhang, Yifan Hou, Zhenjia Xu, Linxi Fan, Manuela Veloso, and Shuran Song. Dexumi: Using human hand as the universal manipulation interface for dexterous manipulation. In *Conference on Robot Learning (CoRL)*, 2025.
- [35] Ge Yan, Jiyue Zhu, Yuquan Deng, Shiqi Yang, Ri-Zhao Qiu, Xuxin Cheng, Marius Memmel, Ranjay Krishna, Ankit Goyal, Xiaolong Wang, et al. Manifold: A general robot manipulation policy via consistency flow training. In *Conference on Robot Learning (CoRL)*, 2025.
- [36] Shiqi Yang. Ace: A cross-platform visual-exoskeleton system for low-cost dexterous teleoperation. In *Conference on Robot Learning (CoRL)*, 2024.
- [37] Tianyu Yu, Peisheng He, Bob Tianqi Wei, Chenyuheng Wang, Xueqing Li, Xuezhong Wang, Yao Lu, Wei Yue, Megan Teng, Zihan Wang, Liwei Lin, Haipeng Mi, Qi Lu, and Lining Yao. Morphingskin: A skin-like platform that integrates multimodal hydraulic actuators based on flexible electroosmotic pumps. In *ACM Symposium on User Interface Software and Technology*, 2025.
- [38] Wenzhen Yuan, Siyuan Dong, and Edward H Adelson. Gelsight: High-resolution robot tactile sensors for estimating geometry and force. *Sensors*, 2017.
- [39] Yanjie Ze, Gu Zhang, Kangning Zhang, Chenyuan Hu, Muhan Wang, and Huazhe Xu. 3d diffusion policy: Generalizable visuomotor policy learning via simple 3d representations. In *Robotics Science and Systems (RSS)*, 2024.
- [40] Han Zhang, Songbo Hu, Zhecheng Yuan, and Huazhe Xu. Doglove: Dexterous manipulation with a low-cost open-source haptic force feedback glove. In *Robotics Science and Systems (RSS)*, 2025.
- [41] Tony Z Zhao, Vikash Kumar, Sergey Levine, and Chelsea Finn. Learning fine-grained bimanual manipulation with low-cost hardware. In *Robotics Science and Systems (RSS)*, 2023.
- [42] Shuo Zhu, Angus Stuttaford-Fowler, Ashraf Fahmy, Chunxu Li, and Johann Sienn. Development of a low-cost data glove using flex sensors for the robot hand teleoperation. In *Intelligent Manufacturing Technology (ISRIMT)*, 2021.
- [43] Anya Zorin, Irmak Guzey, Billy Yan, Aadithya Iyer, Lisa Kondrich, Nikhil X. Bhattasali, and Lerrel Pinto. Ruka: Rethinking the design of humanoid hands with learning. In *Robotics Science and Systems (RSS)*, 2025.

APPENDIX

A. Tactile Feedback Module Details

1) **Feedback Module Fabrication:** The general fabrication workflow aligns with the methods established in [23]. However, a key distinction in our approach is the elimination of laser cutting. As discussed in Sec. III-D1, our design optimizations—specifically the increased inter-node spacing—allow the entire assembly to be handcrafted without relying on high-precision industrial machining.

The driving layer consists of two PCB layers bonded together using Pressure Sensitive Adhesive (PSA, 3M 468MP) and a PET spacer. Top-down views of the PCB layout and the assembled feedback module are illustrated in Fig. 12(a). The fabrication process is as follows:

- **Spacer Preparation:** PSA sheets are applied to both sides of a PET film. Using a custom 3D-printed mold and a circular punch, through-holes are cut into the PET film to match the feedback via patterns on the PCBs, as shown in the PET layer of Fig. 2(b).
- **Membrane Integration:** The prepared PET spacer is bonded to the bottom PCB layer. Glass fiber (GF/F) membranes (1.6 mm diameter), which are similarly cut using a circular punch, are then carefully placed into the 32 designated voids within the spacer.
- **Sealing:** The top PCB layer is aligned and bonded onto the assembly. Constant pressure is applied during curing to ensure a hermetic seal and effective isolation between adjacent fluidic channels.

To integrate the soft structure with the rigid driving layer, we use Sil-Poxy adhesive to bond the Ecoflex™ 00-30 output and reservoir components to the PCB. Once the module is fully bonded and cured, Propylene Carbonate (PC) is introduced as the dielectric working fluid. Using a syringe, PC is injected into the reservoir; simultaneously, trapped air is evacuated from the chamber to ensure a bubble-free hydraulic system.

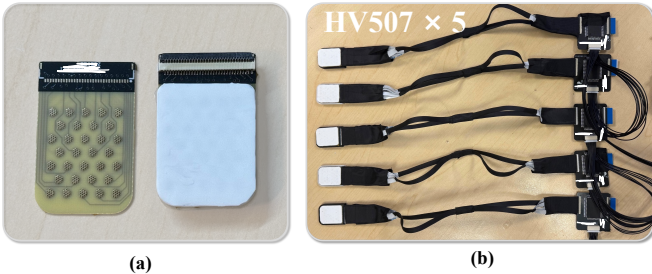


Fig. 12: Tactile Feedback Module and Control Circuit. (a) Top view of the Tactile Feedback Module and its single-layer PCB; (b) Five modules correspond to the five fingers, each independently driven by an HV507 chip.

2) **Control Circuitry:** Each feedback module features a 32-taxel tactile array requiring independent voltage regulation for both the top and bottom electrodes, totaling 64 control channels. This configuration enables tri-state control for each taxel: active protrusion (+), neutral (0), and retraction (-). As illustrated in Fig. 12(b), each module is driven by a single HV507 (Microchip Technology), a 64-channel high-voltage

serial-to-parallel converter with push-pull outputs. The HV507 supports a daisy-chain configuration and features a Latch Enable (LE) function, allowing the states of multiple HV507 chips to be synchronized and controlled using a minimal number of digital I/O pins.

B. Retargeting Details

1) **Hand Retargeting:** The tactile glove provides its joint angle information θ_g at a frequency of 260 Hz. Through forward kinematics, we can obtain the poses (p_g, q_g) of the glove's end-links relative to the wrist frame. To match the size of the dexterous hand, we scale p_g by a factor λ_f , defined as the ratio between the dexterous hand's finger length and the operator's corresponding finger length, yielding the target fingertip positions p_{gt} , while keeping the target orientations unchanged ($q_{gt} = q_g$).

Preserving thumb–finger opposition is crucial for successful manipulation during retargeting. Therefore, we incorporate an opposition constraint to encourage the dexterous hand to reproduce the operator's finger-closing behavior. Specifically, we formulate the following optimization problem:

$$\begin{aligned} \min_{\theta} \quad & w_1 \sum_i \|p_i(\theta) - p_{gt,i}\|_2^2 + w_2 \sum_i \|R_i(\theta) - R_{gt,i}\|_F^2 \\ & + \alpha \|\theta - \theta_{\text{last}}\|_2^2 \\ & + \sum_j \text{weight}_j \left(d_{\text{thumb},j}(\theta) - d_{\text{thumb},j}^{gt} \right)^2 \\ \text{s.t.} \quad & \theta_k^{\min} \leq \theta_k \leq \theta_k^{\max}, \quad \forall k. \end{aligned} \quad (5)$$

where $p_i(\theta)$ and $R_i(\theta)$ denote the position and orientation (as a rotation matrix) of the i -th fingertip link under joint angles θ ; $d_{\text{thumb},j}(\theta)$ is the distance between the thumb and the j -th finger; and weight_j is a piecewise function of the current thumb–finger distance $d_{\text{thumb},j}$: it equals 0 when $d_{\text{thumb},j}$ exceeds an upper threshold d_{max} , increases from 0 to w_{max} as $d_{\text{thumb},j}$ decreases from d_{max} to a lower threshold d_{min} , and remains clamped at w_{max} when $d_{\text{thumb},j} \leq d_{\text{min}}$. The bounds $\theta_k^{\min}$ and $\theta_k^{\max}$ denote the joint-angle limits for each component of θ . Solving Eq. (5) yields the dexterous-hand joint command θ_t at time step t , which is then sent to the hand controller to track the retargeted motion. We can solve this optimization problem at approximately 3000 Hz on an i7-12700 CPU, which is sufficient for real-time control. The qualitative results of the proposed retargeting framework are illustrated in Fig. 13.

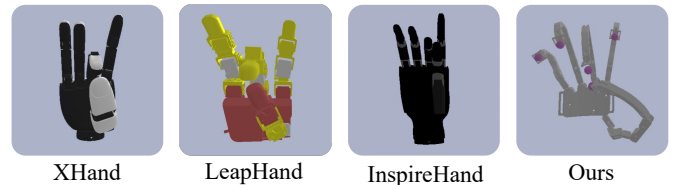


Fig. 13: Hand retargeting performance. The optimization-based method is applied to three distinct robotic hands: XHand, LeapHand and InspireHand. Despite differences in kinematics and dimensions, all hands successfully reproduce the target gesture captured from TAG.

2) **Arm Retargeting:** We implement arm retargeting by mapping the operator’s wrist pose directly to the robot’s end-effector. As detailed in [Sec. III-E](#), we employ two distinct interfaces for pose acquisition: the HOMIE exoskeleton and the HTC VIVE trackers. Below, we describe the implementation details for each setup.

HOMIE Exoskeleton: We adopt the exoskeleton interface design from HOMIE [1] and integrate it into our TAG system. By utilizing Forward Kinematics (FK), we acquire both the arm joint poses and the wrist pose.

For the Unitree G1, our setup remains consistent with the original HOMIE implementation. We employ a 7-DoF isomorphic exoskeleton for the right arm, which is directly interfaced with TAG. We do not extend the FK-based exoskeleton adaptation of TAG to the UR5e. For exoskeleton-based teleoperation of the UR5e, we refer readers to the implementation in GELLO [33].

HTC VIVE Trackers: To acquire precise wrist pose information, we utilize the Vive tracking system, which consists of two base stations (infrared (IR) emitters) and a tracker. The system employs infrared laser scanning to achieve high-precision localization of the tracker. Each tracker provides 6-DoF pose data relative to the base station coordinate system, including 3D position $p \in \mathbb{R}^3$ and orientation represented as a quaternion q , at a sampling frequency of 250 Hz.

In our setup, we attach a Vive tracker to the operator’s glove to capture accurate wrist pose. We first transform the pose information into a coordinate system aligned with the robot’s base frame and record the initial pose p_0, q_0 . At each time step t , we compute the incremental change in pose relative to the initial pose as $\Delta p_t, \Delta q_t$. Since the operator’s arm movement range typically exceeds that of the robot, we introduce a scaling factor λ to adjust the positional increment Δp_t .

For the Unitree G1, we apply this scaled positional increment and the rotational change to the initial pose of the robot’s end-effector p_{r0}, q_{r0} to obtain the target pose p_{rt}, q_{rt} . Finally, we map the target pose to the robot’s joint space using inverse kinematics. Specifically, we formulate an optimization problem:

$$\begin{aligned} \min_{\theta} \quad & w_1 \|p(\theta) - p_{rt}\|^2 + w_2 \|R(\theta) - R_{rt}\|_F \\ & + \alpha \|\theta - \theta_{last}\|^2 \\ \text{s.t.} \quad & \theta_k^{\min} \leq \theta_k \leq \theta_k^{\max}, \quad \forall k. \end{aligned} \quad (6)$$

where $p(\theta)$ denotes the end-effector position of the robot at joint angles θ , and α is the weight of the regularization term. The bounds $\theta_k^{\min}$ and $\theta_k^{\max}$ denote the joint-angle limits for each component of θ . By solving this optimization problem, we obtain the robot’s joint angles θ_t at the current time step t , enabling pose tracking of the robot’s end-effector through the robot’s control interface.

For the UR5e, we similarly add the increments $\Delta p_t, \Delta q_t$ to the initial pose of the arm’s end-effector p_{r0}, q_{r0} to obtain the target pose p_{rt}, q_{rt} , and then utilize the UR5e’s built-in interface for target pose tracking.

3) **Key Note:** A critical implementation detail concerns signal interference. The RealSense D455 camera utilizes an active IR projector for depth estimation. When its projection field overlaps with the VIVE base stations, the IR emission causes crosstalk, resulting in tracking jitter or severe localization errors. To avoid this, we strategically adjust the spatial positioning of the base stations to eliminate signal overlap, successfully ensuring robust tracking performance.

C. Experimental Details

1) **User Study:** We use the **Contact Shape Discrimination** task to illustrate the protocol. The experimental workflow follows the procedures described in [Sec. V-A1](#). Participants are blindfolded and wear noise-canceling headphones to eliminate visual and auditory cues. Since they cannot actively align the robotic finger with the target shapes, a second experimenter positions the geometries beneath the Inspire Hand’s index finger.

In each trial, the second experimenter shuffles and randomly selects one geometry, pressing it against the sensor with sufficient force to generate clear tactile features. The second experimenter also monitors signal quality in real-time to ensure proper data capture, maintaining contact for 5 seconds. Upon removal, the participant immediately reports the perceived shape relying solely on tactile feedback. This process continues until all four geometries are tested, completing one experimental block. Each participant performs 5 blocks (20 trials total). With 5 participants, a total of 100 trials are conducted. The results are presented in [Fig. 9](#).

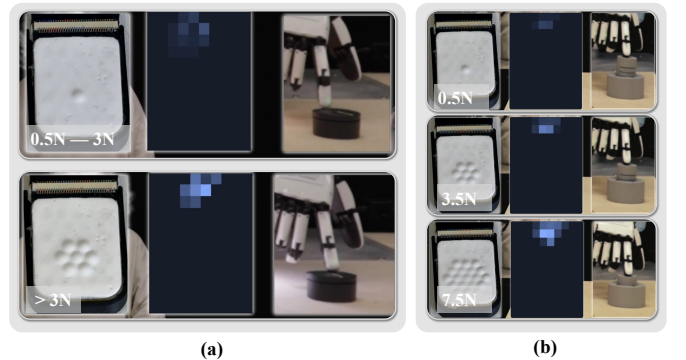


Fig. 14: Illustration of teleoperation tasks in Press Mode. (a) Multimodal Control of a Smart Desk Lamp. (b) Multi-level Spring Compression.

2) **Teleoperation Experiments Details:** The general experimental setup follows the description in [Sec. V-B](#). For comparative analysis, we select the HOMIE system [1] as the baseline—specifically, the HOMIE Glove integrated with a VIVE Tracker. This baseline device is a 15-DoF exoskeleton glove that employs single-axis Hall sensors for joint encoding but lacks tactile feedback capabilities.

- **Multimodal Control of a Smart Desk Lamp** – Supplementary to the protocol, Stages 2 and 3 require the operator to perform multiple rotational clutching motions (typically 3–5 repetitions) to traverse the full control range. The trial ends immediately upon reaching the target state, defined

as the warmest CCT or minimum brightness. A trial is classified as a failure if any incorrect regulation occurs, particularly in Stage 3: if the downward force exceeds 3N during rotation, the system unintentionally switches to the CCT mode of Stage 2, resulting in an immediate failure. Since TAG is configured in **Press Mode**, the tactile protrusion is modulated based on real-time pressure; specifically, TAG protrudes by 1 point (0.5N-3N) and by 7 points ($> 3N$), as illustrated in Fig. 14(a).

- **Filament Pinching and Pull-out** – This task entails two consecutive pinch-and-pull actions. The trial fails if either attempt is unsuccessful. Failure modes include missing the grasp initially (grasping air) or the filament slipping from the fingertips during pull-out. The schematic of the thumb and index finger pinching the filament is shown in Fig. 15.
- **Multi-level Spring Compression** – Whether blindfolded or sighted, the operator must verbally report the perceived force level upon engagement and maintain pressure for 2 seconds. A trial fails if the actual compression depth mismatches the reported level—specifically, if the visual markers on the spring-loaded button fall short of or exceed the target lines. The three pressure levels and corresponding feedback are detailed in Fig. 14(b).



Fig. 15: Illustration of Filament Pinching and Pull-out Task. The schematic shows the thumb and index finger pinching the filament.

D. LeapTac Design

The LeapHand is extensively employed in robotic manipulation research; however, it notably lacks integrated tactile sensing capabilities. To address this limitation, as detailed in Sec. IV-B, we present *LeapTac*, a custom-designed resistive tactile array tailored for the LeapHand fingertip. The fundamental sensing principle of LeapTac aligns with the piezoresistive approach described in [13].

Structurally, the sensor comprises a piezoresistive Velostat sheet sandwiched between two flexible printed circuit (FPC) layers containing orthogonal conductive traces. The top layer features transverse traces while the bottom layer features longitudinal traces, forming a grid architecture where each intersection point functions as an individual sensing element, or *taxel*, as illustrated in Fig. 16(a). This sandwich structure ensures that pressure applied to the surface compresses the Velostat locally, reducing its electrical resistance at the contact point. To seamlessly accommodate the curved geometry of the

LeapHand fingertip, the FPC electrodes are designed with a trace width of 2 mm and a center-to-center spacing of 4 mm. This configuration yields a spatial resolution of 7×6 taxels per fingertip, providing sufficient density for object localization and force estimation.

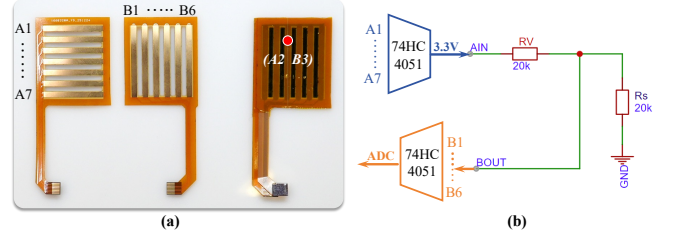


Fig. 16: Physical prototype and schematic of the LeapTac sensor. (a) The two-layer Flexible Printed Circuit (FPC) electrodes and the fully assembled LeapTac module. (b) The circuit schematic illustrating the readout mechanism for individual sensing points (taxels).

The electronic interface is designed to minimize wiring complexity while maximizing sampling speed. Data acquisition is managed by two 8-channel analog multiplexers (74HC4051) operating in a row-column scanning mode. As shown in the schematic in Fig. 16(b), one multiplexer controls the voltage driving lines (rows), while the other handles the signal readout lines (columns). To query a specific taxel (R_i, C_j), the driving multiplexer selectively applies a reference voltage V_{cc} to row R_i . Current flows through the resistive film at the intersection to column C_j , which is connected via the readout multiplexer to a microcontroller's ADC through a voltage divider circuit. By measuring the output voltage, we can calculate the instantaneous resistance of the Velostat using Ohm's law. This resistance change is inversely proportional to the normal force applied at the specific coordinate (R_i, C_j), allowing the system to reconstruct a real-time pressure map of the contact area.