IROS 2026 CONTACT-RICH MANIPULATION · FEW-SHOT IMITATION LEARNING

PHASE: Compliance-Enabled Tactile Phase Retrieval for Few-Shot Insertion Learning

PROJECT PAGE

SINIC X

Jeremy Siburian[1,2] Cristian C. Beltran-Hernandez[1] Tatsuya Matsushima[2] Yusuke Iwasawa[2] Masashi Hamaya[1†] Mai Nishimura[1†]

[1]OMRON SINIC X Corporation, Tokyo, Japan [2]The University of Tokyo †Joint last authors. Contact information: masashi.hamaya@sinicx.com

01 WHAT PROBLEM?

Learn **peg-in-hole insertion from a handful of demonstrations** by reusing prior insertion experience.

CHALLENGES

- **Phase identification:** Search-to-insert transitions appear in force and tactile signals but may be missed by kinematics.
- **Variable duration:** Geometry and start pose change phase duration; fixed windows can split contact phases.

KEY INSIGHT

A compliant wrist sustains contact, producing **tactile and force signals** that reveal insertion phases and guide phase-aligned retrieval.

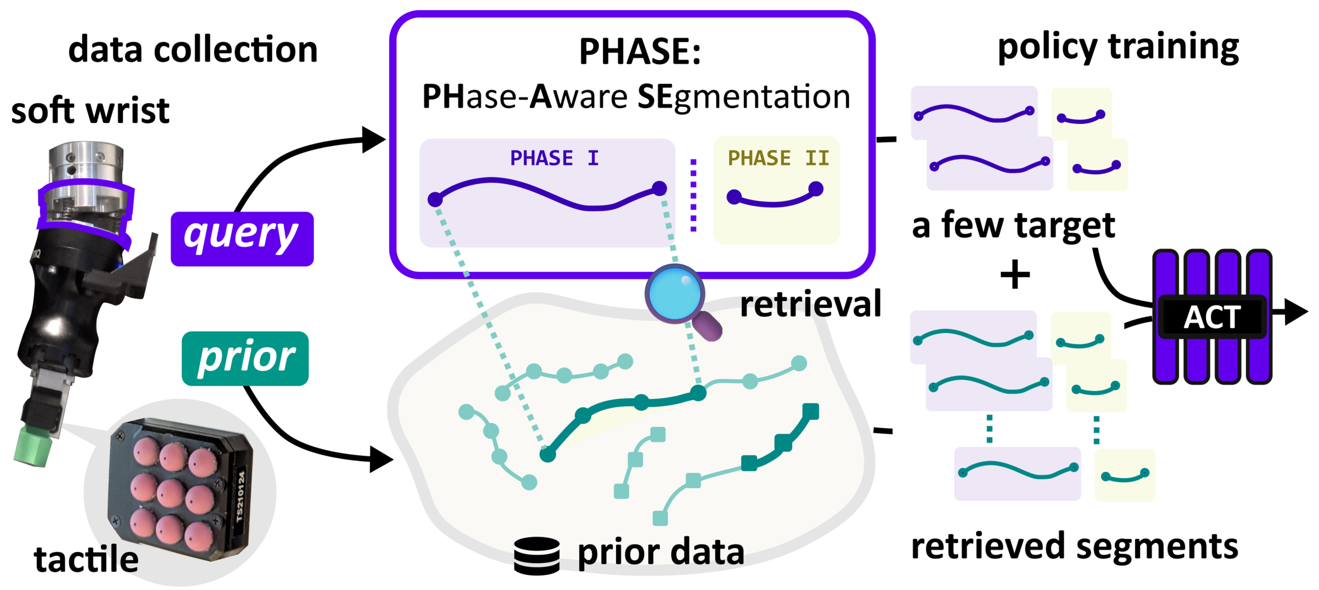
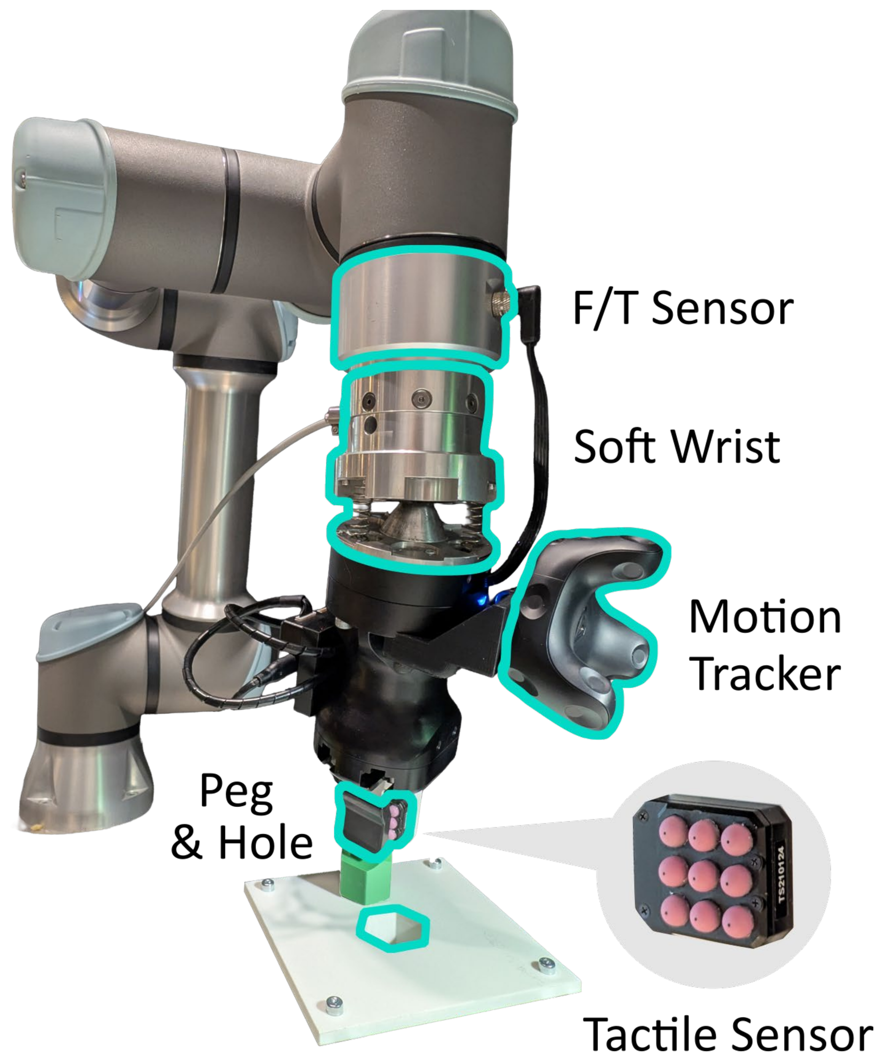


Fig. 1 —PHASE retrieves contact-phase segments from prior demonstrations for few-shot insertion learning.

02 ROBOT SETUP



- UR5e + integrated F/T sensor
- **Soft wrist [1]** – three parallel coil springs, passive 6D compliance
- **3×3 taxel** tactile sensor on the gripper
- VIVE tracker for gripper pose; teleoperated demos at 50 Hz

03 METHOD

1. **TACTILE REPRESENTATION LEARNING**

   The MAT^3 [2] masked tactile-proprioceptive encoder learns per-frame embeddings $z_t \in \mathbb{R}^{256}$ from taxels, F/T and proprioception — self-supervised, no phase labels. Trained on D_{prior}, then frozen.

2. **PHASE SEGMENTATION & RETRIEVAL**

   Tactile-estimated torque $\hat{\tau} = \Sigma\, \mathrm{r_i} \times f_i$. The boundary is the first post-peak frame where $CV = {}^{\sigma}/_{\mu} < \eta$ — the onset of stable engagement. Variable-length **search** / **insert** segments are matched within the same phase by FastDTW.

3. **POLICY TRAINING**

   An **ACT** policy is trained on $D_{\text{train}} = D_{\text{target}} \cup D_{\text{retrieval}}$. Retrieved segments are independent trajectories; the 4 target demos stay complete to preserve the search-to-insert transition.

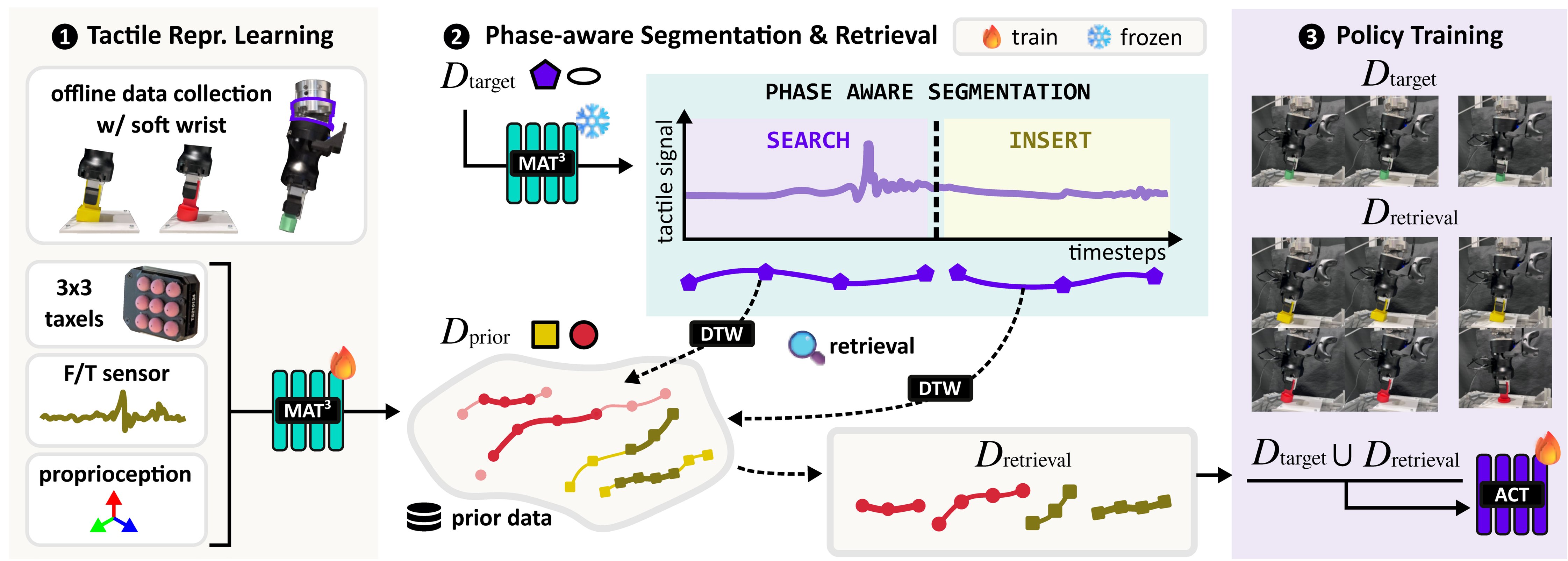


Fig. 2 – The three stages of PHASE: tactile representation learning, phase-aware segmentation & retrieval, and policy training.

WHAT IS A PHASE?

A phase is defined by contact interaction: search locates the hole, while insert advances the aligned peg.

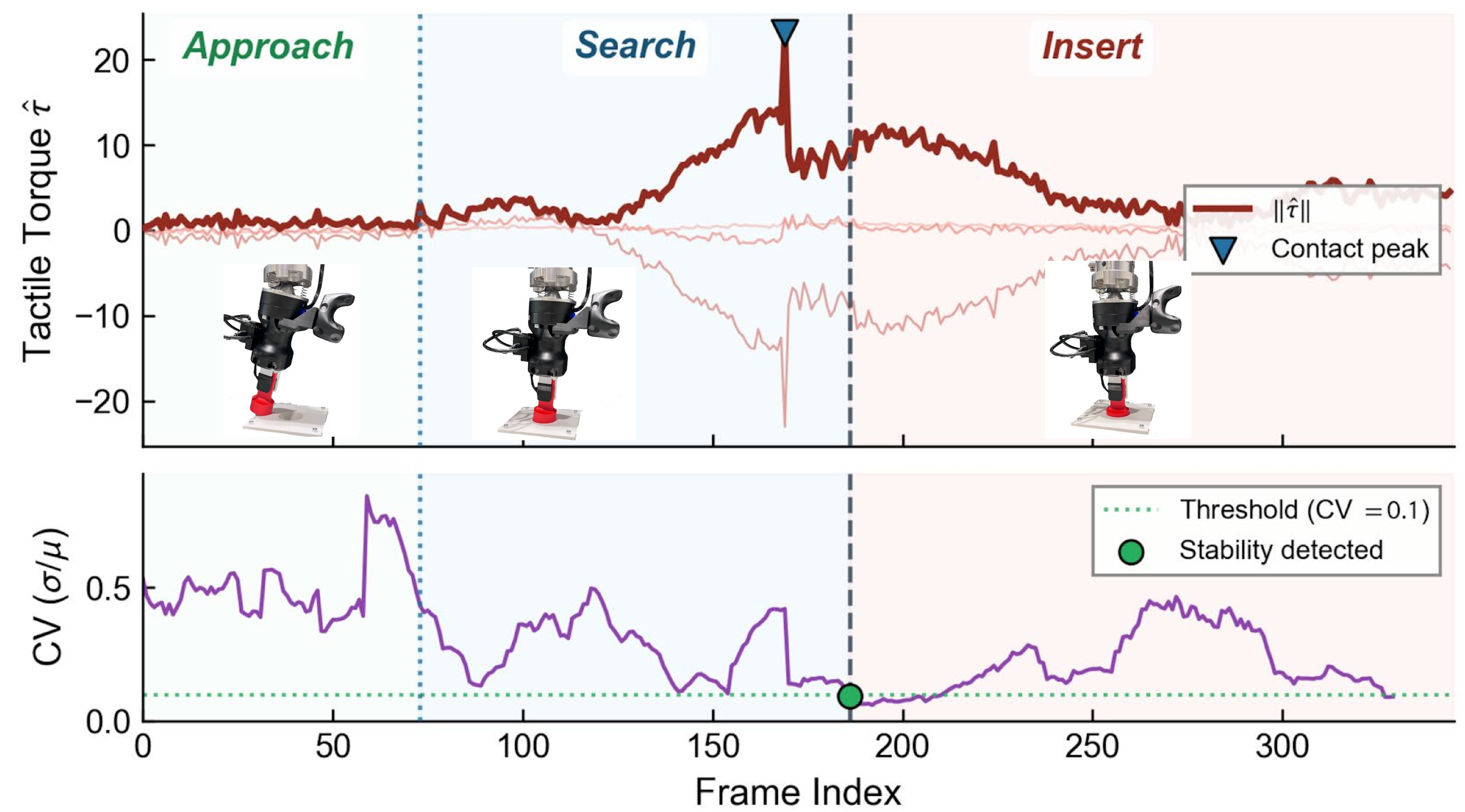


Fig. 3 – The boundary is detected from the tactile torque magnitude and its coefficient of variation.

04 RESULTS

OVERVIEW

77%

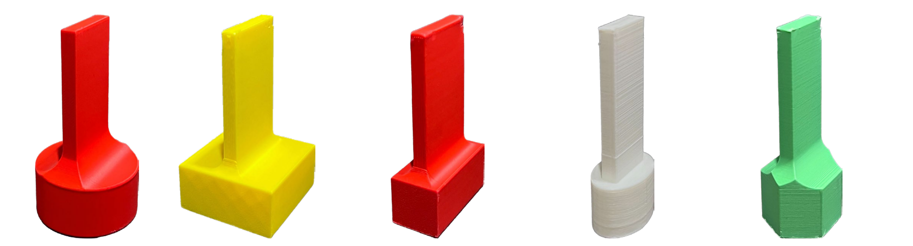
overall success — **+13 pt** over the strongest non-phase-aware baseline

+30 pt

under **unseen starting positions** (47% vs 17%)

SEEN & UNSEEN INSERTION

Success rate, 20 trials per shape · 100 trials per method



Method	Circle Seen	Square Seen	Rect. Unseen	Oval Unseen	Hex. Unseen	Total
Target only	13	6	9	15	7	50%
BC-Prior	19	10	4	17	14	64%
Retrieval-Single	13	7	7	14	15	56%
Retrieval-Full	16	9	7	19	13	64%
Retrieval-Window	16	10	6	16	7	55%
PHASE (Ours)	18	12	13	20	14	77%

UNSEEN START POSES

Success rate from start positions perturbed by ±2 cm (100 trials)

Method	Circle	Square	Rect.	Oval	Hex.	Total
Target only	0	0	0	0	0	0%
BC-Prior	5	2	3	5	2	17%
Retrieval-Single	0	1	1	0	0	2%
Retrieval-Full	4	4	4	2	0	14%
Retrieval-Window	6	5	5	0	0	16%
PHASE (Ours)	12	8	12	9	6	47%

05 LIMITATIONS & FUTURE WORK

Current segmentation uses insertion-specific torque heuristics. Learned contact-phase representations could extend PHASE to other contact-rich tasks.

References

[1] Kazutoshi Tanaka, Felix von Drigalski, Masashi Hamaya, Robert Lee, Chisato Nakashima, Yoshiya Shibata, Yoshihisa Ijiri, "A Compact, Cable-driven, Activatable Soft Wrist with Six Degrees of Freedom for Assembly Tasks," IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS 2020), pp 8752-8757, 2020

[2] Tatsuya Kamijo, Mai Nishimura, Nodoka Shibasaki, Jeremy Siburian, Cristian C. Beltran-Hernandez, Masashi Hamaya, "Tactile Memory with Soft Robot: Robust Object Insertion via Masked Encoding and Soft Wrist," IEEE Robotics and Automation Letters, 2026