GarmentPile: Point-Level Visual Affordance Guided Retrieval and Adaptation for Cluttered Garments Manipulation

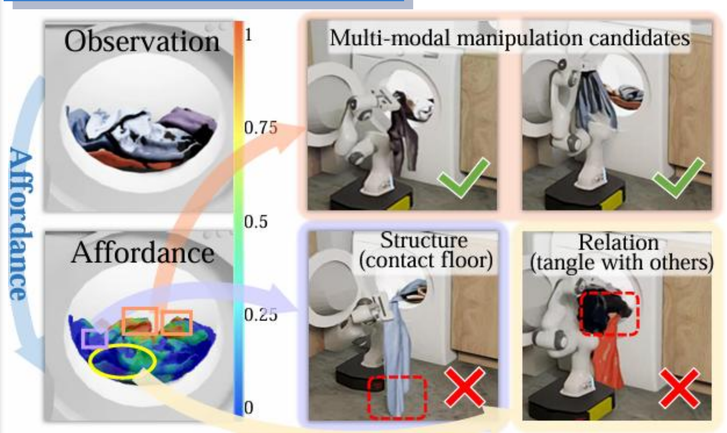
Ruihai Wu* Ziyu Zhu* Yuran Wang* Yue Chen Jiarui Wang Hao Dong

PEKING UNIVERSITY

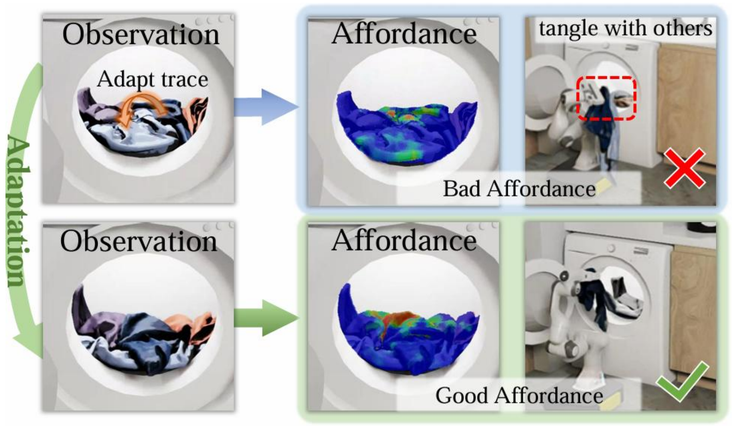
Project Page

Introduction

MOTIVATION: Cluttered garments manipulation requires suitable representation due to its complexity.



➢ Affordance can represent multi-modal output, garment structures and relations.

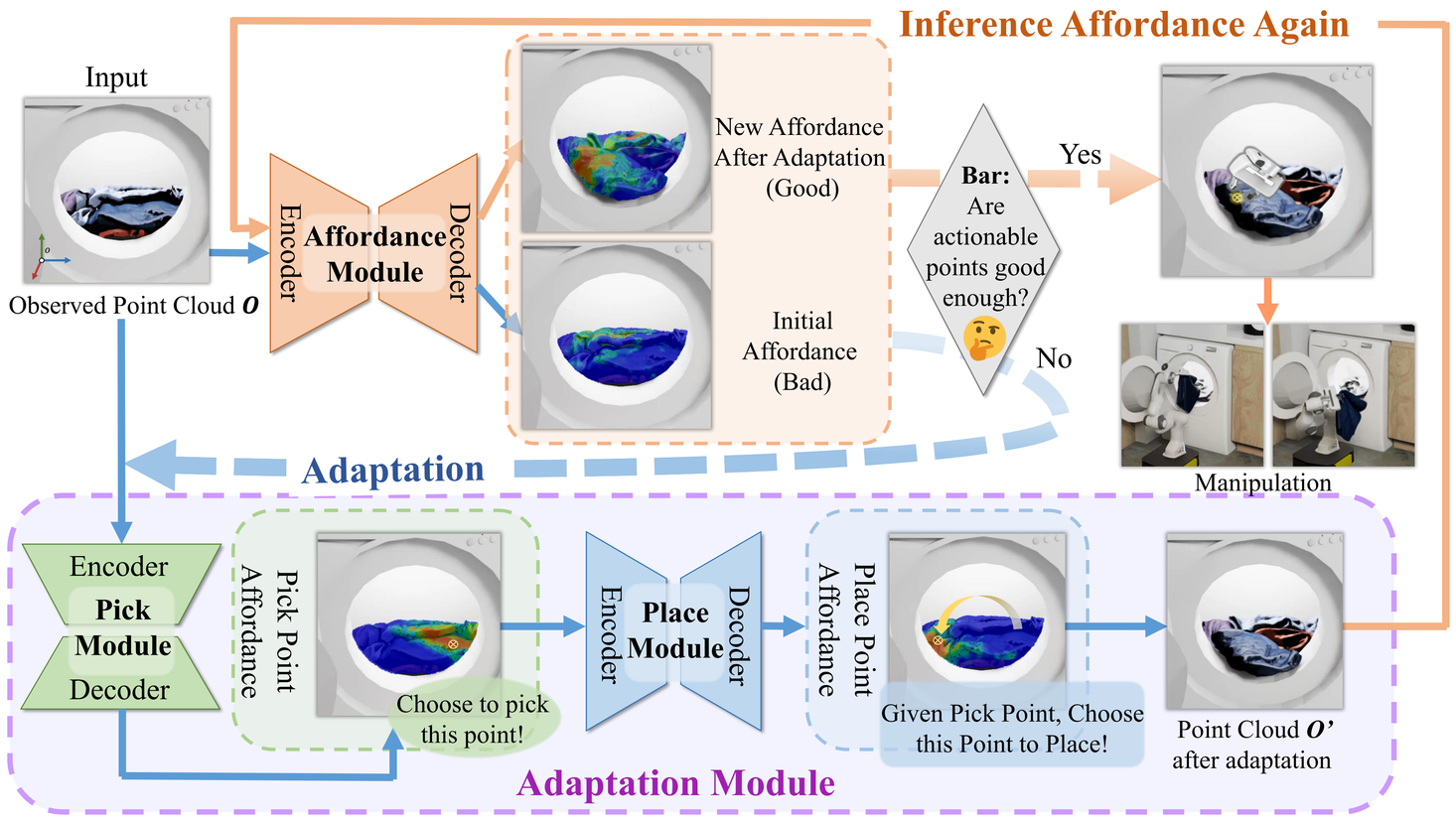


➢ Affordance can guide reorganizing the scene into a plausible state.

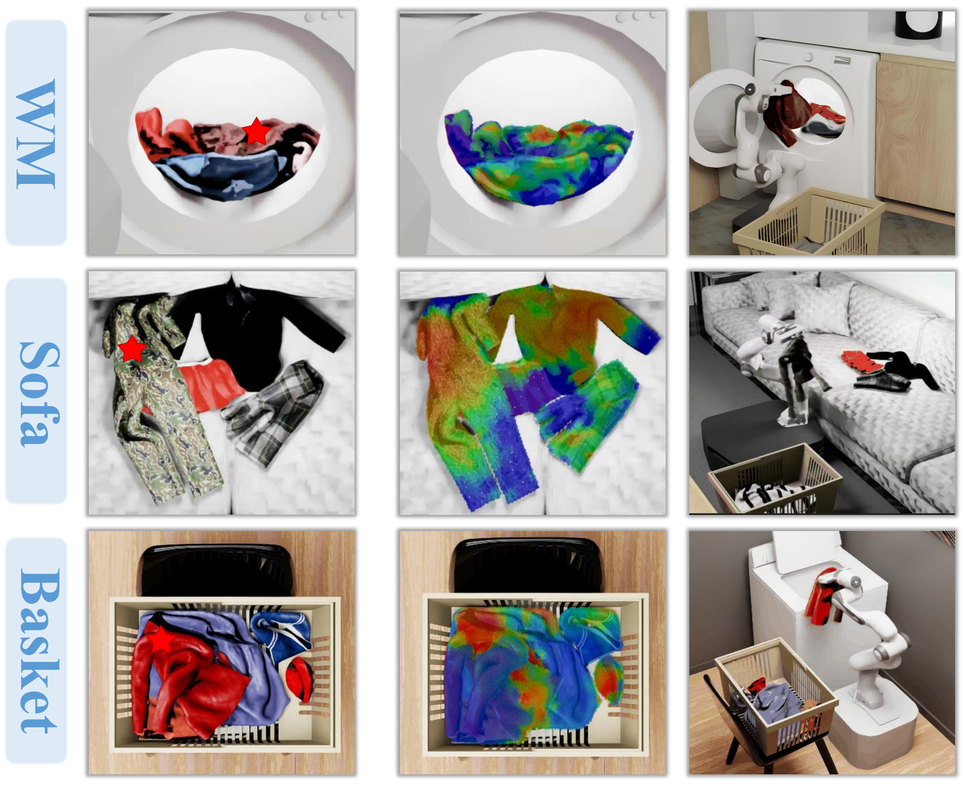
- We introduce **point-level affordance** learning for cluttered garments manipulation for efficient representation of highly complex state and action spaces, and **multi-modal policy outputs**.
- We further develop **the adaptation module** to efficiently adapt the cluttered garments to states easy to successfully manipulate.

Method

➢Our inference pipeline guided by affordance.



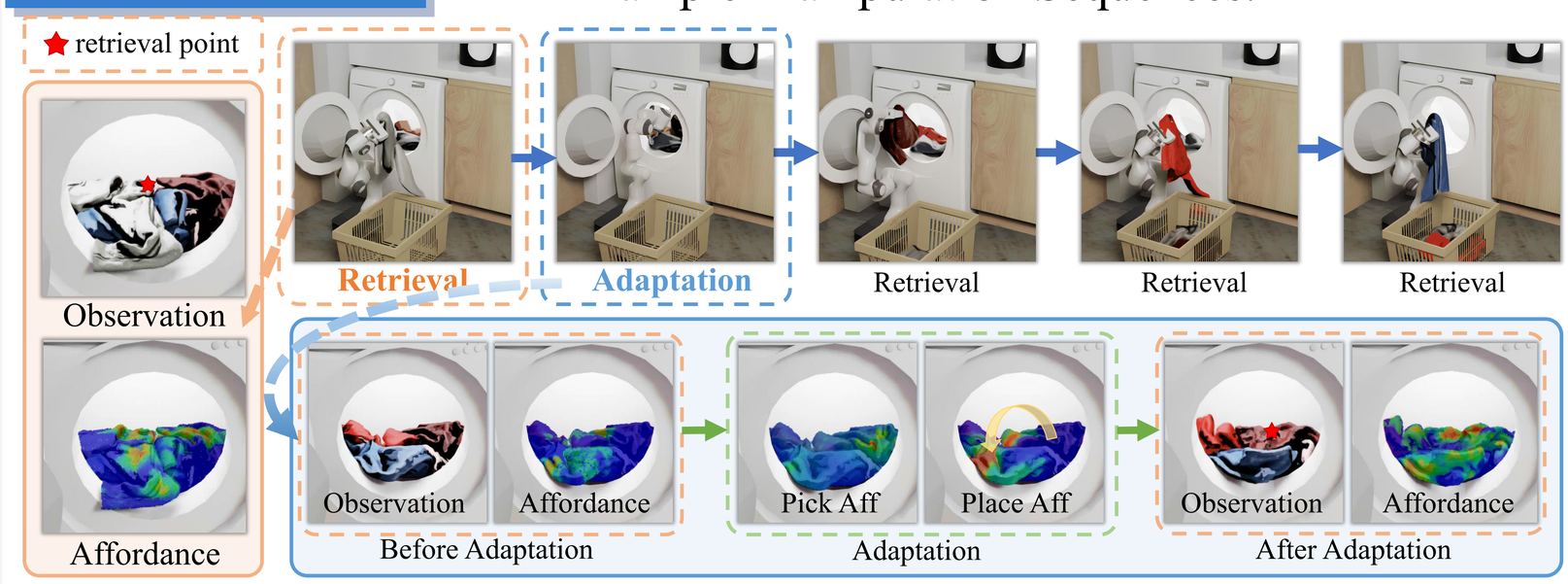
Results in Simulation



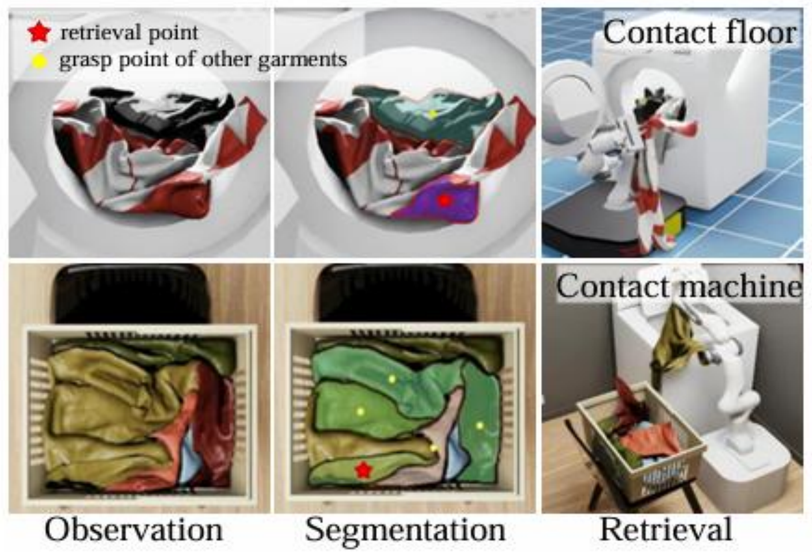
➢Retrieval results in different scenarios.

Whole Process

➢Example Manipulation Sequences.



Baseline



➢Guided by Segmentation and Support Relation.

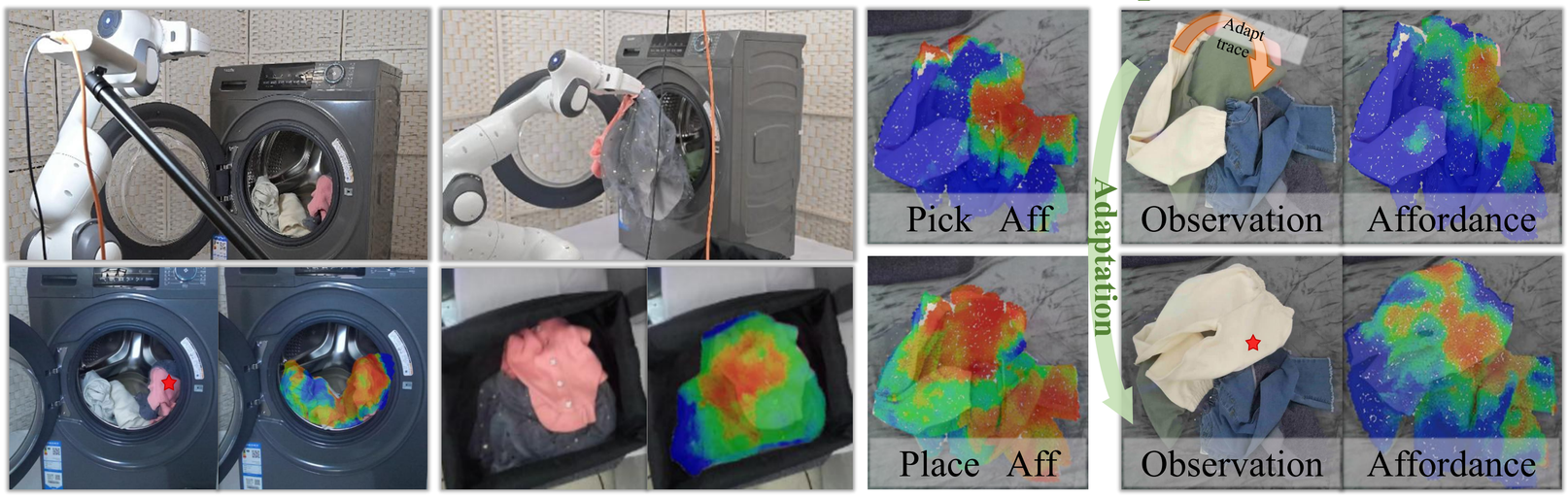
➢Guided by GPT-4o.

Method	WashingMachine	Sofa	Basket
Where2Act [31]	9 / 15	10 / 15	8 / 15
Support-M [18]	8 / 15	12 / 15	9 / 15
GPT-Fabric [36]	6 / 15	7 / 15	6 / 15
Ours	**12 / 15**	**13 / 15**	**12 / 15**

➢Real-World Success Rate in different scenarios.

Results in the Real World

➢Retrieval in WM ➢Retrieval in Basket ➢Adaptation in Sofa



➢Retrieval and adaptation results in different scenarios.