GraphMimic: Graph-to-Graphs Generative Modeling from Videos for Policy Learning

Guangyan Chen[1], Te Cui[1], Meiling Wang[1], Chengcai Yang[1], Mengxiao Hu[1], Haoyang Lu[1], Yao Mu[2], Zicai Peng[1], Tianxing Zhou[1], Xinran Jiang[1], Yi Yang[1], Yufeng Yue[1]

[1]Beijing Institute of Technology, Beijing, P. R. China

[2]The University of Hong Kong, Hong Kong, P. R. China

Motivation

Current methods:

- Model representations from pixel space directly, neglecting the modeling of important structures

- Graph representations inherently encode spatial relational bias, enabling effective structural modeling.

- Visual action vertices can be incorporated into graphs, capturing the relationships between objects and effectors.

Can we formulate video pre-training as graph-to-graphs generative modeling, enabling policy learning on limited action-labeled data?

Motivation

Three critical challenges

(1) Embodiment discrepancy:
The source video and target scene often exhibit substantial embodiment discrepancy, hindering the transfer of learned action knowledge.

We propose the action-informed transferable graph representation, which extracts visual action vertices by abstracting key interaction points of the embodiment.

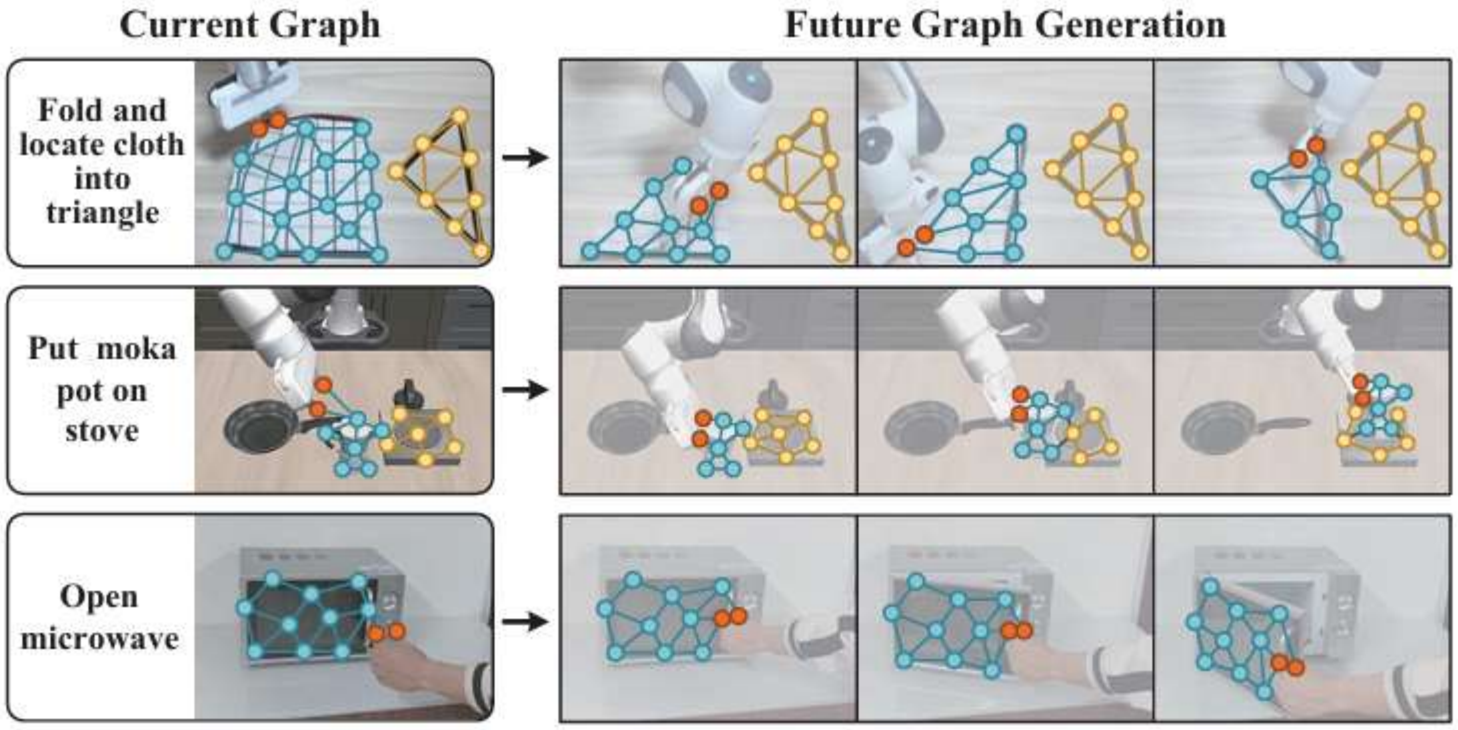
(2) Object property diversity:
The properties of distinct objects exhibit significant variability, such as stiffness, inhibiting the generative modeling network from capturing diverse object behaviors.

We propose to infer properties from historical observations and integrate them into graphs, facilitating predictions for objects with distinct properties.

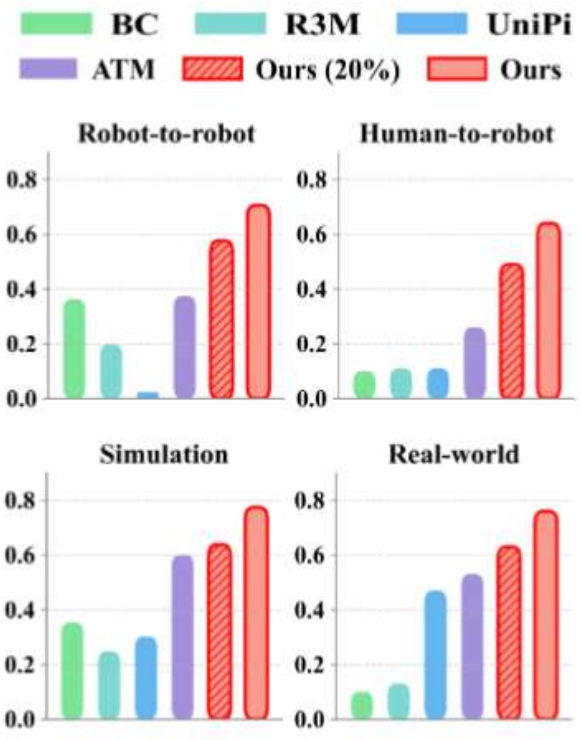
(3) Contradiction between internal and global modeling:
Modeling long-range dependencies is crucial for capturing spatial relationships, but directly expanding edge construction range potentially overwhelms important internal structures

We present a hierarchical architecture, integrating global vertices for long-range information propagation.

Motivation

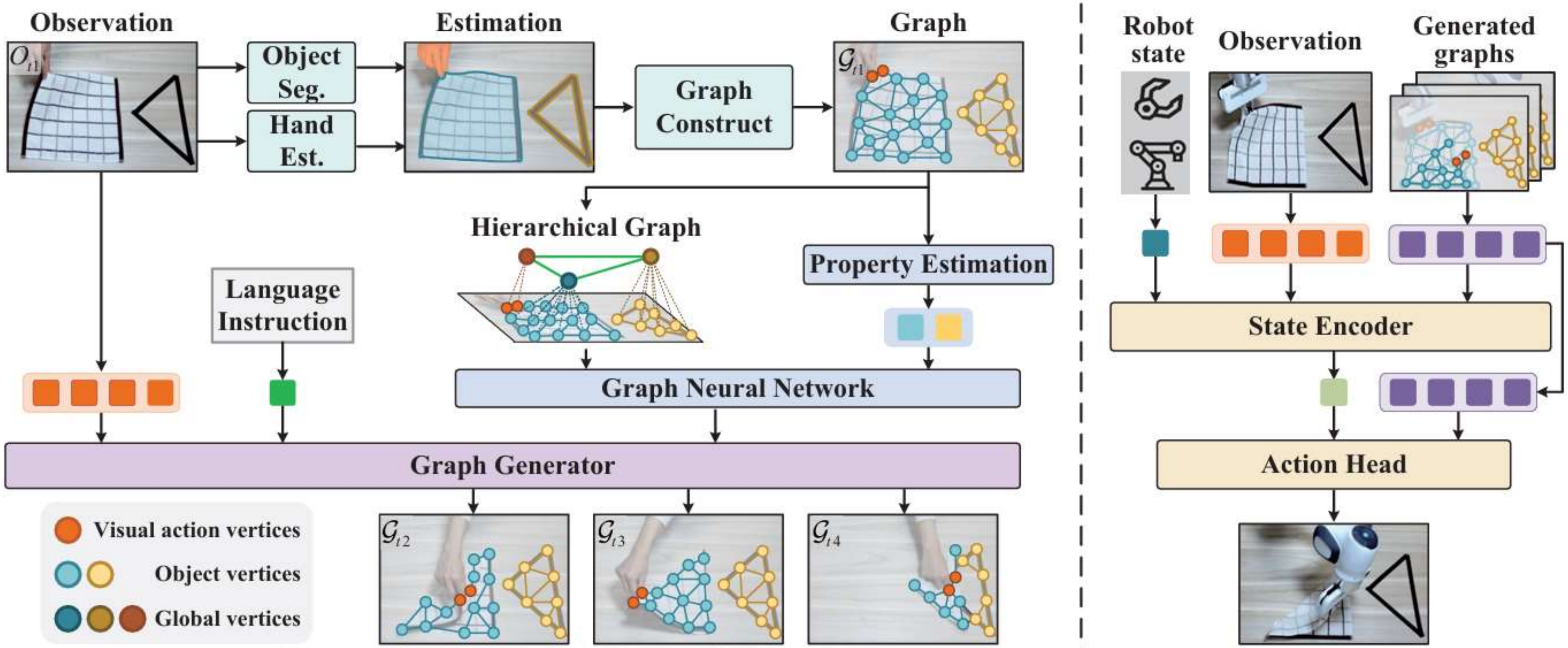


(a) Graph-to-Graphs Generative Modeling



(b) Comparisons with Baselines

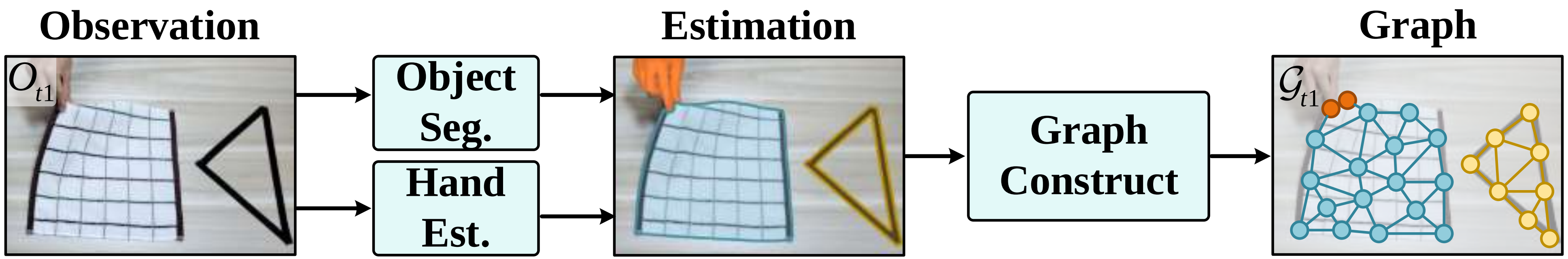
Methods



(a) Property-aware Graph Generative Modeling

(b) Graph-guided Policy

Action-informed Transferable Graph

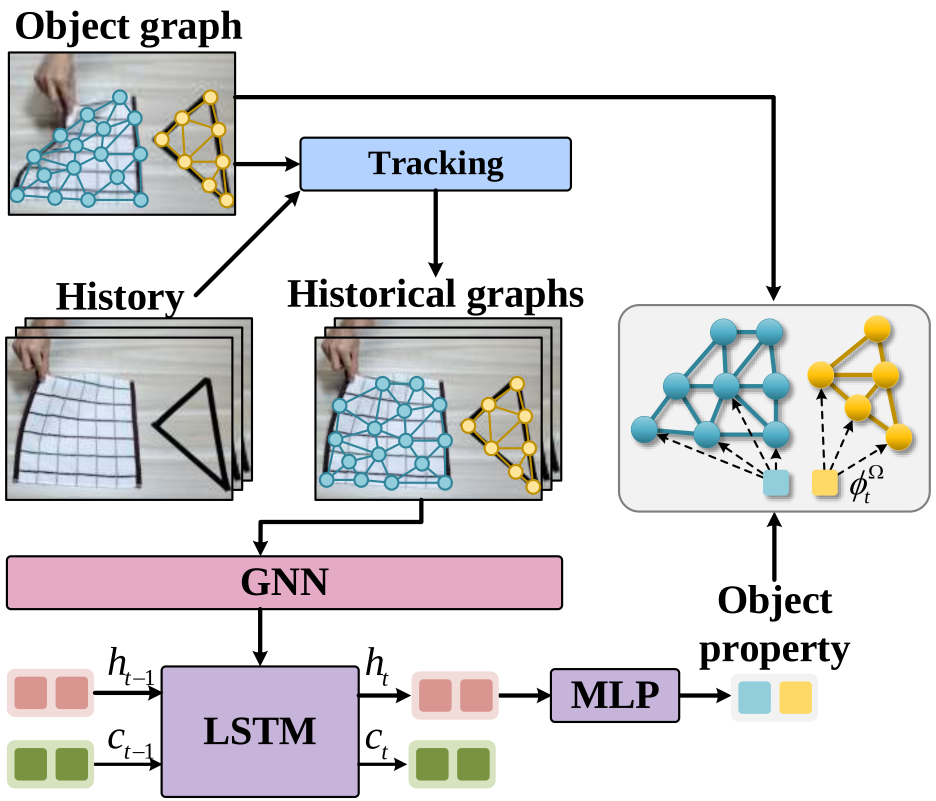


Object vertices: The object graph dynamics describe the change in object state, irrelevant to the agent embodiment, enhancing the cross-embodiment learning ability

Visual action vertices: The incorporation of visual action vertices enables our approach to capture object-embodiment interactions

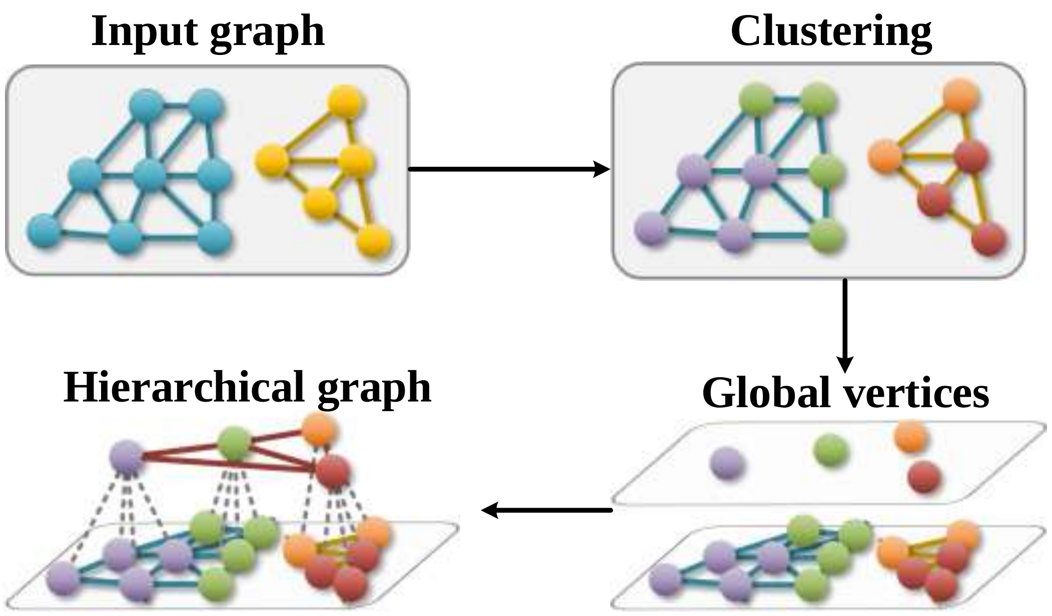
Property-aware Graph Generative Modeling

Object property estimation



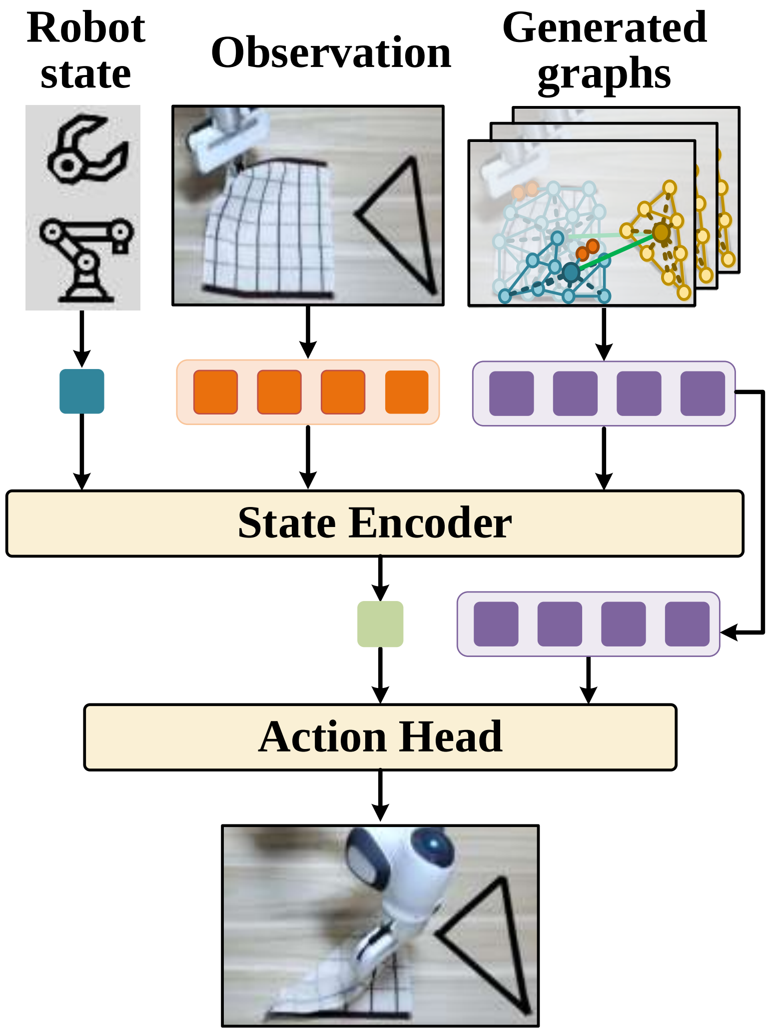
We retrieve object-centric graphs and track vertex set across historical observations. The historical information is encoded into object properties.

Hierarchical graph modeling



Vertices are clustered with the root global vertex added to each cluster. Edges are then constructed to connect the hierarchical structure.

Graph Guided Policy Learning



State encoder: The state encoder encodes all modalities to generate the state representation.

Action Head: The MLP action head generates action sequences.

Experiments - Video Pre-training for Imitation Learning

Table 1. Average success rate on LIBERO benchmark. 50 videos are collected for pre-training, 'Num of L-demos' indicates the number of action-labeled demonstrations utilized for training. Results are averaged over three seeds.

Methods	Num of L-demos	Libero-Spatial	Libero-Object	Libero-Goal	Libero-Long	Libero-90
BC-Full-Trainset	50	0.71 ± 0.03	0.71 ± 0.08	0.76 ± 0.01	0.24 ± 0.02	0.38 ± 0.01
BC	10	0.39 ± 0.08	0.51 ± 0.13	0.42 ± 0.04	0.16 ± 0.03	0.29 ± 0.01
R3M [29]	10	0.49 ± 0.04	0.52 ± 0.02	0.05 ± 0.01	0.09 ± 0.03	0.09 ± 0.00
UniPi [18]	10	0.69 ± 0.04	0.59 ± 0.03	0.11 ± 0.02	0.05 ± 0.02	0.07 ± 0.01
ATM [21]	10	0.68 ± 0.02	0.68 ± 0.06	0.77 ± 0.01	0.39 ± 0.15	0.48 ± 0.02
Ours (20%)	**2**	$\mathbf{0.75 \pm 0.02}$	$\mathbf{0.75 \pm 0.02}$	$\mathbf{0.74 \pm 0.03}$	$\mathbf{0.43 \pm 0.05}$	$\mathbf{0.52 \pm 0.02}$
Ours	**10**	$\mathbf{0.88 \pm 0.02}$	$\mathbf{0.89 \pm 0.02}$	$\mathbf{0.87 \pm 0.03}$	$\mathbf{0.56 \pm 0.03}$	$\mathbf{0.67 \pm 0.01}$

Experiments - Video Pre-training for Imitation Learning

Table 2. Success rates on real-world manipulation experiments. 50 videos are collected for pre-training, 'Num of L-demos' indicates th number of action-labeled demonstrations utilized for training.

Methods	BC-Full-Trainset	BC	R3M	UniPi	ATM	**Ours (20%)**	**Ours**
Overall	0.44	0.10	0.13	0.48	0.53	**0.63**	**0.76**

Methods	Num of L-demos	Open drawer	Stack block	Open oven	Put fruit on plate	Press button	Pour from cup to cup
BC-Full-Trainset	50	0.7	0.7	0.4	0.7	0.6	0.4
BC	10	0.3	0.2	0.2	0.2	0.1	0.0
R3M [29]	10	0.4	0.3	0.2	0.3	0.0	0.0
UniPi [18]	10	0.6	0.6	0.5	0.7	0.6	0.5
ATM [21]	10	0.8	0.6	0.7	0.7	0.7	0.5
Ours (20%)	**2**	**0.7**	**0.7**	**0.7**	**0.8**	**0.7**	**0.6**
Ours	**10**	**0.8**	**0.8**	**0.8**	**0.9**	**0.8**	**0.8**

Methods	Num of L-demos	Sweep table	Insert box	Push box	Fold cloth	Fold and put cloth	Straighten rope
BC-Full-Trainset	50	0.6	0.5	0.2	0.3	0.0	0.2
BC	10	0.1	0.0	0.0	0.1	0.0	0.0
R3M [29]	10	0.1	0.1	0.0	0.1	0.0	0.1
UniPi [18]	10	0.7	0.5	0.3	0.3	0.1	0.3
ATM [21]	10	0.7	0.4	0.5	0.4	0.1	0.3
Ours (20%)	**2**	**0.6**	**0.6**	**0.5**	**0.7**	**0.4**	**0.5**
Ours	**10**	**0.8**	**0.6**	**0.7**	**0.8**	**0.6**	**0.7**

Experiments - Cross-embodiment Transfer

Table 3. Robot-to-robot experiment results on LIBERO benchmark. We collect 50 videos from a Franka arm and 10 action-labeled demonstrations from a UR arm for each task. 'Num of L-demos' indicates the number of action-labeled demonstrations utilized for training. Results are averaged over three seeds.

Methods	Num of L-demos	Libero-Spatial	Libero-Object	Libero-Goal	Libero-Long	Libero-90
BC	10	0.42 ± 0.05	0.54 ± 0.09	0.41 ± 0.04	0.17 ± 0.04	0.27 ± 0.01
R3M [29]	10	0.39 ± 0.03	0.44 ± 0.04	0.03 ± 0.03	0.06 ± 0.05	0.07 ± 0.02
UniPi [18]	10	0.03 ± 0.02	0.04 ± 0.01	0.03 ± 0.02	0.01 ± 0.01	0.01 ± 0.01
ATM [21]	10	0.45 ± 0.04	0.56 ± 0.09	0.41 ± 0.04	0.17 ± 0.09	0.28 ± 0.04
Ours (20%)	**2**	$\mathbf{0.69 \pm 0.05}$	$\mathbf{0.71 \pm 0.03}$	$\mathbf{0.67 \pm 0.05}$	$\mathbf{0.34 \pm 0.06}$	$\mathbf{0.47 \pm 0.02}$
Ours	**10**	$\mathbf{0.81 \pm 0.04}$	$\mathbf{0.85 \pm 0.02}$	$\mathbf{0.81 \pm 0.05}$	$\mathbf{0.46 \pm 0.03}$	$\mathbf{0.60 \pm 0.02}$

Experiments - Cross-embodiment Transfer

Table 4. Human-to-robot transfer experiment results on real-world. We collect 50 videos of a human performing the task and 10 action-labeled robot demonstrations for each task. 'Num of L-demos' indicates the number of action-labeled demonstrations utilized for training.

Methods	BC	R3M	UniPi	ATM	**Ours (20%)**	**Ours**
Overall	0.10	0.11	0.11	0.27	**0.49**	**0.64**

Methods	Num of L-demos	Open drawer	Stack block	Open oven	Put fruit on plate	Press button	Pour from cup to cup
BC	10	0.3	0.2	0.2	0.2	0.1	0.0
R3M [29]	10	0.3	0.2	0.2	0.2	0.0	0.0
UniPi [18]	10	0.2	0.1	0.1	0.1	0.2	0.0
ATM [21]	10	0.4	0.4	0.4	0.5	0.2	0.2
Ours (20%)	**2**	**0.7**	**0.4**	**0.6**	**0.7**	**0.5**	**0.4**
Ours	**10**	**0.7**	**0.6**	**0.7**	**0.8**	**0.7**	**0.6**

Methods	Num of L-demos	Sweep table	Insert box	Push box	Fold cloth	Fold and put cloth	Straighten rope
BC	10	0.1	0.0	0.0	0.1	0.0	0.0
R3M [29]	10	0.2	0.1	0.0	0.1	0.0	0.0
UniPi [18]	10	0.1	0.1	0.1	0.2	0.0	0.1
ATM [21]	10	0.3	0.2	0.2	0.3	0.1	0.0
Ours (20%)	**2**	**0.5**	**0.4**	**0.4**	**0.6**	**0.3**	**0.4**
Ours	**10**	**0.6**	**0.6**	**0.6**	**0.7**	**0.5**	**0.6**

Experiments – Ablation Analysis

Table 5. Ablation experiments with GraphMimic on robot-to-robot transfer experiments. Default settings are marked in gray.

(a) State representations.

Variants	Spatial	Object	Goal	Long
Graph with grid vertices	0.51	0.60	0.43	0.25
Flow with related points	0.66	0.65	0.53	0.23
Graph with related vertices	0.81	0.85	0.81	0.46

(b) Visual action representations.

Variants	Spatial	Object	Goal	Long
No action	0.41	0.61	0.42	0.18
End-effector	0.56	0.59	0.51	0.19
Interaction points	0.81	0.85	0.81	0.46

(c) Object representations.

Variants	Spatial	Object	Goal	Long
Manipulated objects	0.71	0.76	0.67	0.31
All objects	0.55	0.66	0.52	0.27
Task-related objects	0.81	0.85	0.81	0.46

(d) Hierarchical modeling.

Variants	Spatial	Object	Goal	Long
Local connection	0.73	0.79	0.71	0.38
Global connection	0.75	0.72	0.73	0.37
Hierarchical architecture	0.81	0.85	0.81	0.46

Table 6. Ablation study on the object property estimation module. Default settings are marked in gray.

Variants	Real-world	Human-to-robot
No estimation	0.66	0.57
No LSTM	0.73	0.60
Graph+LSTM	0.76	0.64

Thank you!