Where2Act: From Pixels to Actions for Articulated 3D Objects

Kaichun Mo, Leonidas J. Guibas, Mustafa Mukadam, Abhinav Gupta, Shubham Tulsiani

Task-aware Visual Affordance
What actions one can take?

- Pull the drawer handle
- Push the door closed
- Pull the door by edge
- ......
What actions one can take?

Pull the drawer handle outward

What?  Where?  How?
Problem Formulation

What?

push

Where?

pull

How?

(a) Where?

(b) How?

(c) 0.81
0.94
0.92
0.35
0.95
0.92
0.85
0.29
Networks

2D-UNet
(a)

3D-PN++
(b)

or

Actionability score:
Is this action-type good for this point?

Action Proposals:
What actions can be taken?

Action Score:
Is this specific action likely to succeed?
Dataset and Simulation Environment


Xiang et al., “SAPIEN: A SimulAted Part-based Interactive ENvironment”, CVPR 2020
Learn from Simulated Interaction (Trial-and-Error)

SAPIEN

No Human-Annotation or Demonstration Needed

Short-term Hard-coded Trajectory Parameterized by <Contact Point, Gripper Orientation>

Xiang et al., “SAPIEN: A Simulated Part-based Interactive Environment”, CVPR 2020
- Task-aware Localized Geometric Features
- Generalize to novel test object categories
Results: Action Score Predictions

- Localized Geometric Features
- Generalize to novel test object categories
Results: Action Proposals

Diverse proposals
Results: Real Data Transfer Testing

Reasonable Results
Where2Act: From Pixels to Actions for Articulated 3D Objects

With Prof. Leonidas J. Guibas, Mustafa Mukadam, Prof. Abhinav Gupta, Prof. Shubham Tulsiani

Thank you for Watching!

(ask in the chat box or attend live sessions for questions)

Email: kaichun@cs.stanford.edu
Web: https://cs.stanford.edu/~kaichun/where2act/