DeepRob

Seminar 7
Explicit Scene-Level Representations
University of Michigan and University of Minnesota

This Week: Scene-Level Representations

- Seminar 7: Semantic Scene Graphs and Explicit Representations
  1. Image Retrieval using Scene Graphs, Johnson et al., 2015
  2. Semantic Robot Programming for Goal-Directed Manipulation in Cluttered Scenes, Zeng et al., 2018
  3. Semantic Linking Maps for Active Visual Object Search, Zeng et al., 2020
  4. Hydra: A Real-time Spatial Perception System for 3D Scene Graph Construction and Optimization, Hughes et al., 2022
- Seminar 8: Neural Radiance Fields and Implicit Representations
  1. NeRF: Representing Scenes as Neural Radiance Fields for View Synthesis, Mildenhall et al., 2020
  2. iMAP: Implicit Mapping and Positioning in Real-Time, Sucar et al., 2021
  3. NeRF-SLAM: Real-Time Dense Monocular SLAM with Neural Radiance Fields, Rosinol et al., 2022
  4. NeRF-Supervision: Learning Dense Object Descriptors from Neural Radiance Fields, Yen-Chen et al., 2022
  5. NARF22: Neural Articulated Radiance Fields for Configuration-Aware Rendering, Lewis et al., 2022

Today: Explicit Representation

- Seminar 7: Semantic Scene Graphs and Explicit Representations
    1. Image Retrieval using Scene Graphs, Johnson et al., 2015
    2. Semantic Robot Programming for Goal-Directed Manipulation in Cluttered Scenes, Zeng et al., 2018
    3. Semantic Linking Maps for Active Visual Object Search, Zeng et al., 2020
    4. Hydra: A Real-time Spatial Perception System for 3D Scene Graph Construction and Optimization, Hughes et al., 2022

- Seminar 8: Neural Radiance Fields and Implicit Representations
    1. NeRF: Representing Scenes as Neural Radiance Fields for View Synthesis, Mildenhall et al., 2020
    2. iMAP: Implicit Mapping and Positioning in Real-Time, Sucar et al., 2021
    3. NeRF-SLAM: Real-Time Dense Monocular SLAM with Neural Radiance Fields, Rosinol et al., 2022
    4. NeRF-Supervision: Learning Dense Object Descriptors from Neural Radiance Fields, Yen-Chen et al., 2022
    5. NARF22: Neural Articulated Radiance Fields for Configuration-Aware Rendering, Lewis et al., 2022

Semantic Robot Programming for Goal-Directed Manipulation of Cluttered Scenes

By: Zhen Zeng, Zheming Zhou, Zhiqiang Sui, Odest Chadwicke Jenkins

Presented by: Shaurya Gunderia, Sukruthi Chidananda

The Authors

- **Zhen Zeng**
  - Ph.D University of Michigan
    - Research Scientist @ JP Morgan AI research
- **Zheming Zhou**
  - Robotics Ph.D University of Michigan
    - Applied Scientist @ Amazon Lab 126
- **Zhiqiang Sui**
  - CS/Robotics Ph.D, University of Michigan
    - Senior Software Engineer @ Nuro
- **Odest Chadwicke Jenkins**
  - CS/Robotics Ph.D USC
    - Professor of EECS/Robotics at Michigan

Navigating Cluttered Environments

- Cluttered environments pose major challenges for robots
- Humans can navigate easily, not as easy for robots
- Goal is to teach robots to understand cluttered environments

Benefits of Semantic Robot Programming

- **Easier Programming:** Based on high-level, natural language, making it easier and faster to use than traditional programming
- **Faster Deployment:** Allows for quick robot programming, facilitating faster deployment in the industry
- **Better Flexibility:** Enables easy adaptation of robots to different tasks and environments, increasing flexibility.
- **Improved Human-Robot Interaction:** Enables robots to understand and respond to human language, enhancing human-robot interaction
- **Increased Safety:** Allows for programming of robots to perform risky taks, reducing the risk of harm to humans.
- **Better Performance:** Improves robot performance by enabling understanding of complex instructions and efficient task performance

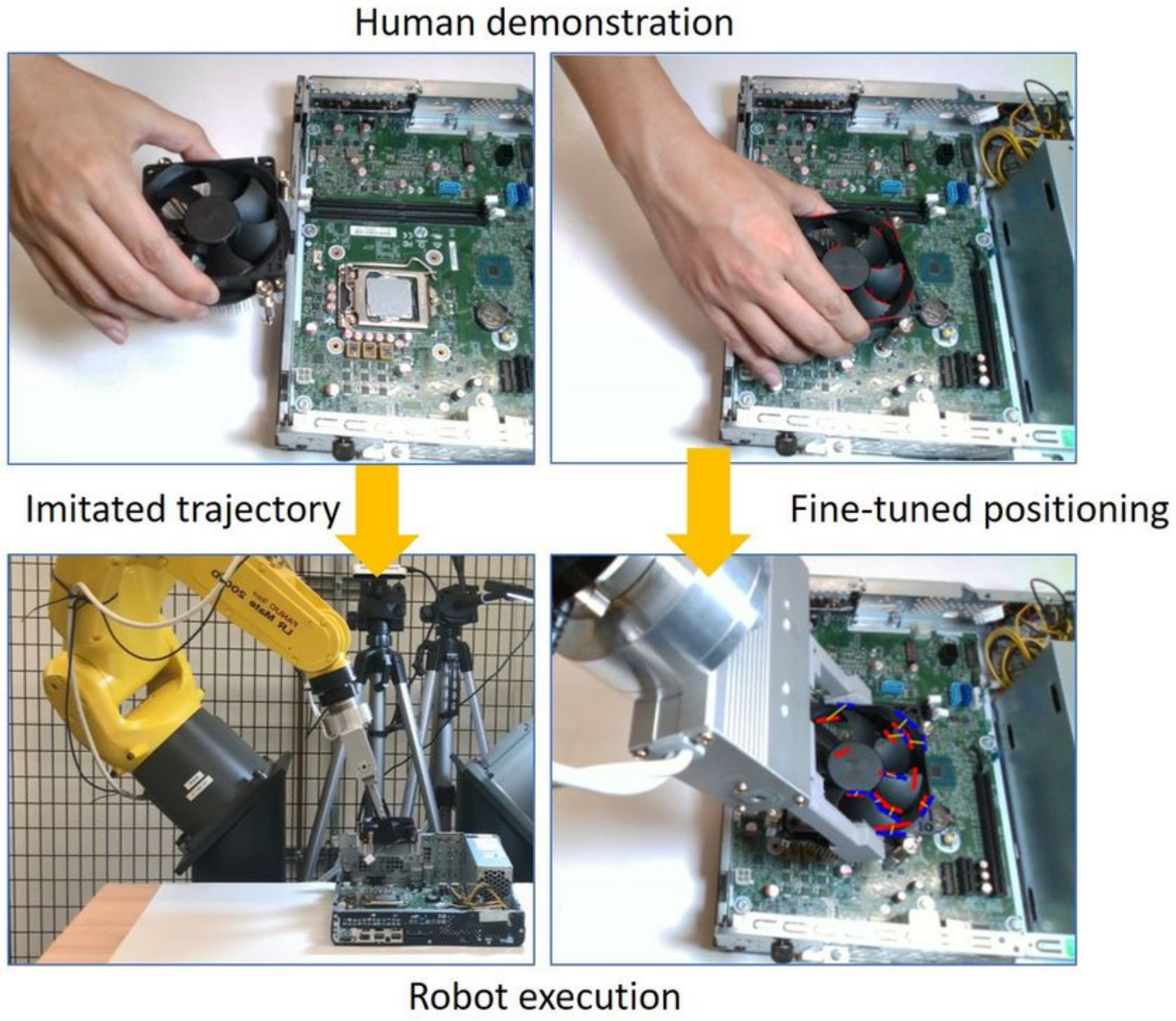
Contributions

1. SRP enables robots to perform goal-directed manipulation in cluttered environments
2. Uses semantic representation of environment and tasks
3. Understands user input and formulates high-level goals.

- Robot programming in cluttered environments is hard, SRP addresses some of these challenges.

Prior Work and Background

- SRP builds on PbD and scene perception for manipulation
- PbD learns low-level skills from users
- Scene perception enables manipulation in real world scenarios
- DIGEST estimates scene graph for goal-directed manipulation in initial world states
- Prior methods struggle in cluttered environments - rely on known object geometry, colors, etc.



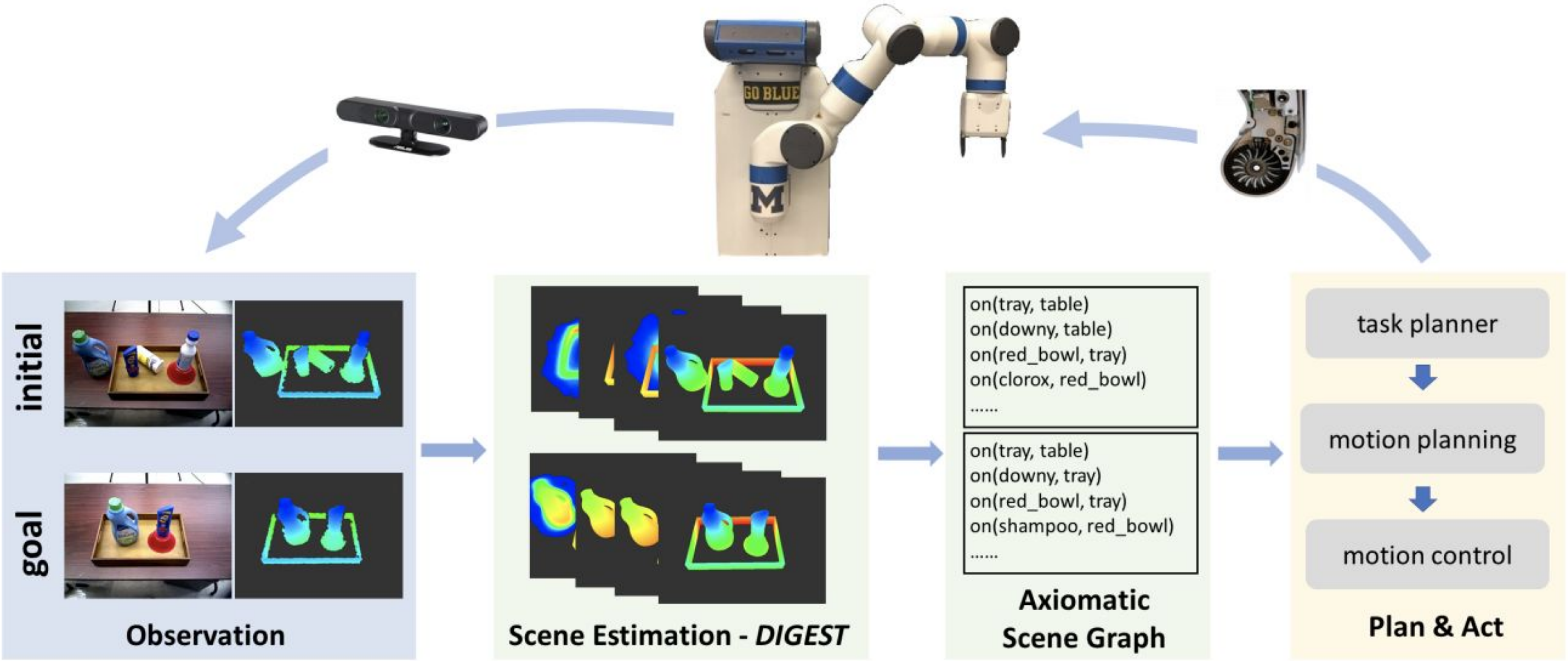
Semantic Robot Programming

- Traditionally → manually specify the robot's movements for each task
- SRP uses high level descriptions of task to allow robots to autonomously generate movements
- The approach
  - objects are represented as semantic entities with attrs. such as location, size, shape.
  - tasks represented as semantic descriptions.
  - planner generates a sequence of movements that satisfy task description.

The DIGEST Framework

- Framework for goal-directed manipulation in cluttered scenes
- Has three modules
  - Scene understanding: CV to identify objects
  - Goal Formulation: formulates high-level goals using knowledge base and user input
  - Task planning: generates sequence of movements to achieve goal using task planning

Approach Visualized



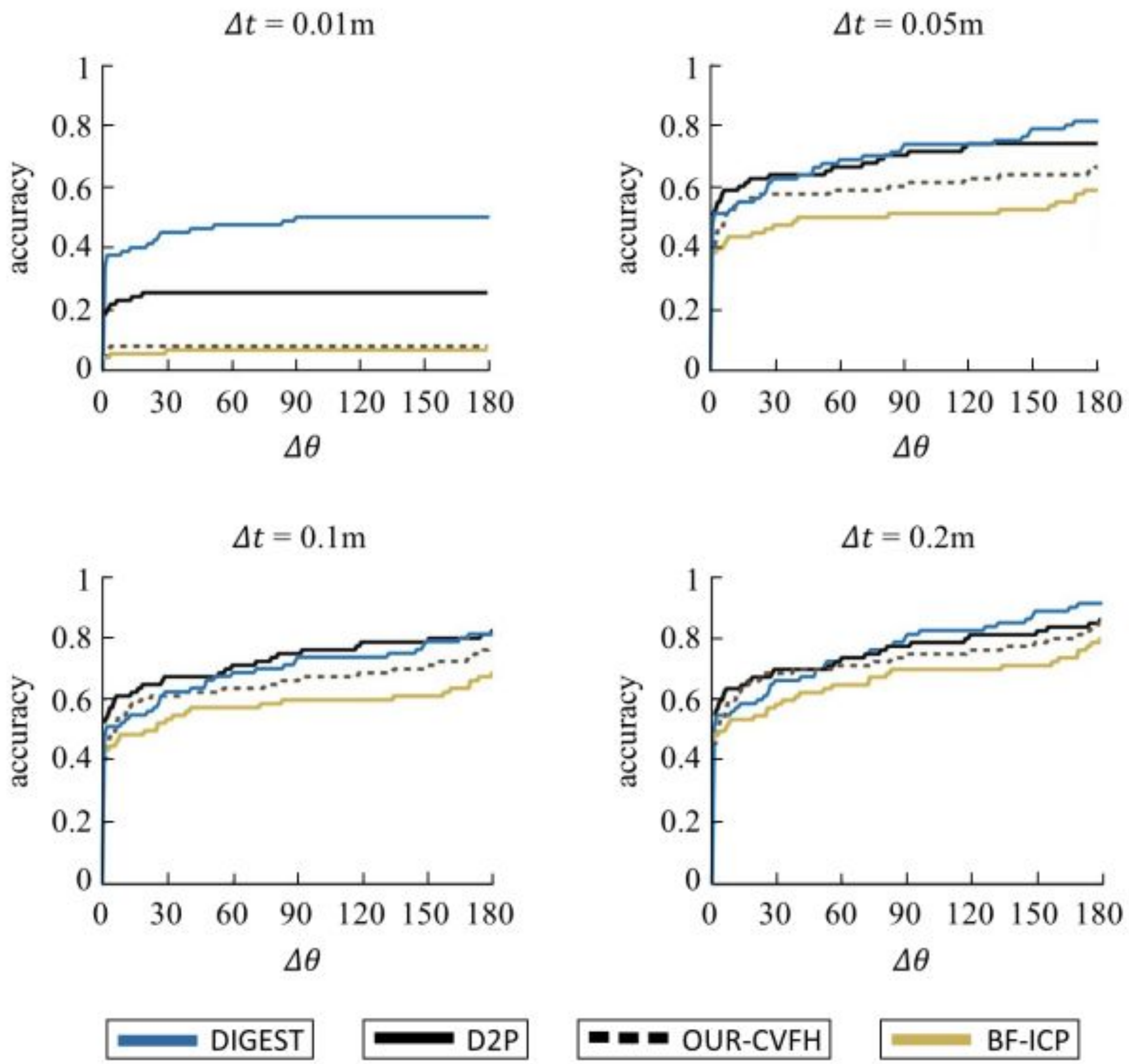
Insights for SRP

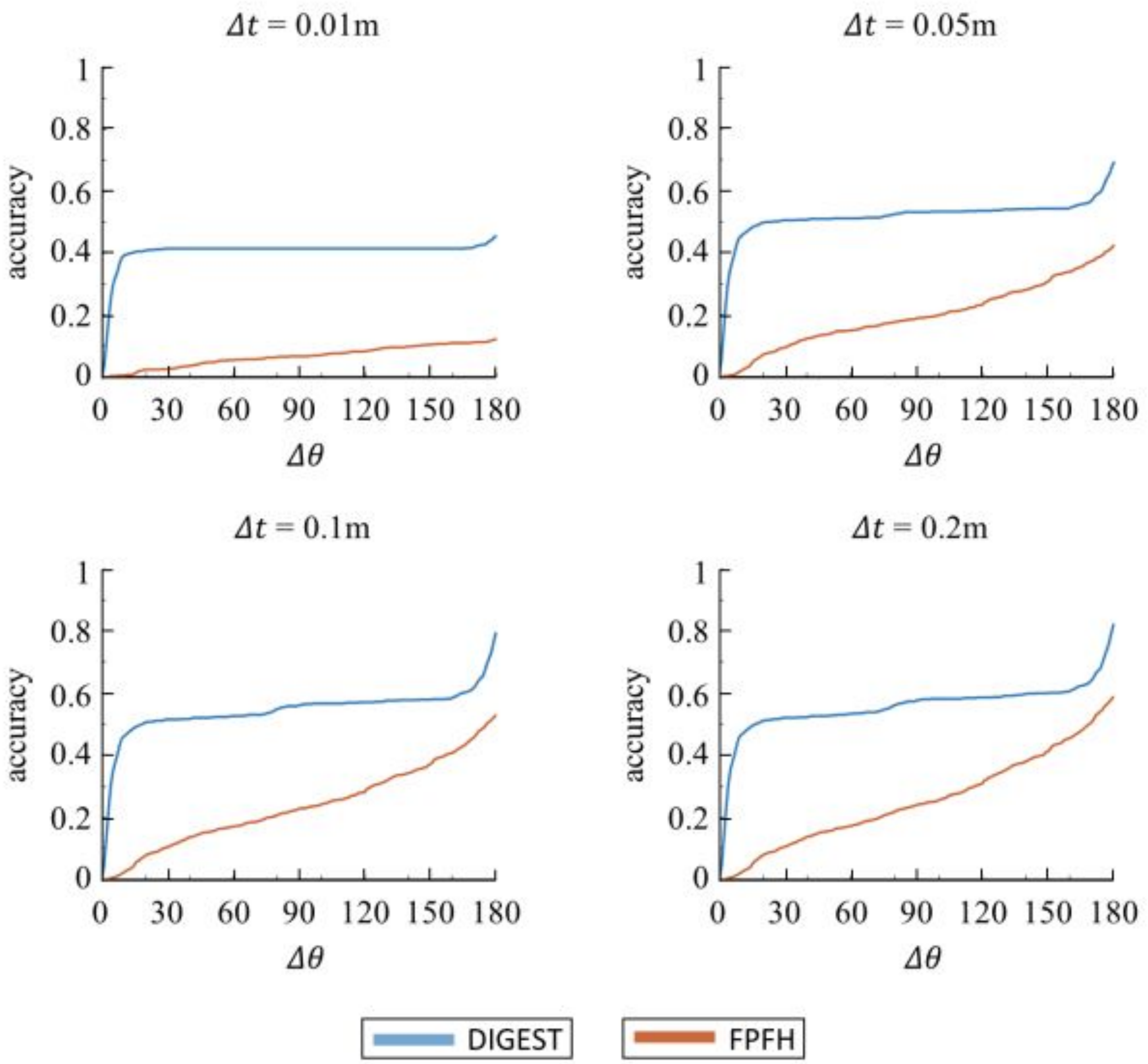
- **Scene Segmentation:** Segment the scene into objects with semantic attributes
- **Object Affordances:** Associate objects with actions that can be performed on them (grasping, pushing, etc.)
- **Task Representation:** Develop a method for representing tasks as semantic descriptions
- **Planning:** Develop a planner than can generate a sequence of movements to satisfy a task description

Results

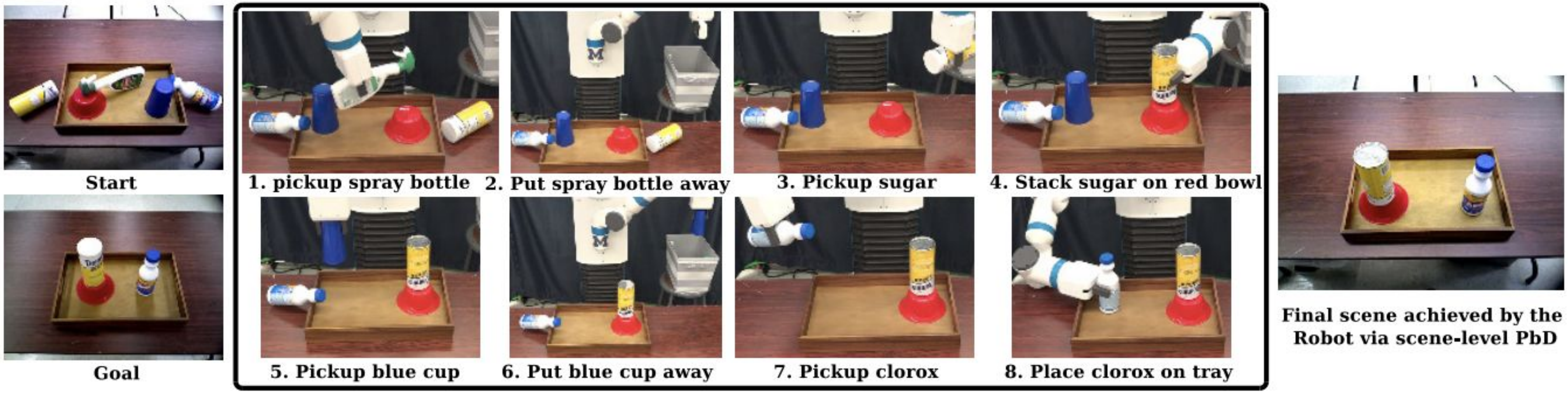
- Evaluated on household occlusion dataset and cluttered scene dataset.
  - Accuracy based on % of correctly localized objects
  - Object is correctly localized if pose error within position/rotation threshold
- SRP outperformed traditional programming methods, based on success rate and time taken for completion
- Number of actions also reduced

PH/CS Dataset Performance





SRP in Action



Conclusions

- SRP framework for goal-directed manipulation in cluttered scenes
- Enables easier/faster programming and improved deployment, flexibility, human-robot interaction, safety, and performance
- DIGEST achieve high accuracy in object localization and task planning metrics
- Potential in various industries such as manufacturing, healthcare

Limitations and Directions for Future Work

- Expand range of tasks that can be accomplished with semantic robot programming
- Better CV algorithms for scene segmentation
- Use ML to incorporate prior knowledge/experience

Thank you

Next Time: Implicit Representations

- Seminar 7: Semantic Scene Graphs and Explicit Representations
    1. Image Retrieval using Scene Graphs, Johnson et al., 2015
    2. Semantic Robot Programming for Goal-Directed Manipulation in Cluttered Scenes, Zeng et al., 2018
    3. Semantic Linking Maps for Active Visual Object Search, Zeng et al., 2020
    4. Hydra: A Real-time Spatial Perception System for 3D Scene Graph Construction and Optimization, Hughes et al., 2022
- Seminar 8: Neural Radiance Fields and Implicit Representations
    1. NeRF: Representing Scenes as Neural Radiance Fields for View Synthesis, Mildenhall et al., 2020
    2. iMAP: Implicit Mapping and Positioning in Real-Time, Sucar et al., 2021
    3. NeRF-SLAM: Real-Time Dense Monocular SLAM with Neural Radiance Fields, Rosinol et al., 2022
    4. NeRF-Supervision: Learning Dense Object Descriptors from Neural Radiance Fields, Yen-Chen et al., 2022
    5. NARF22: Neural Articulated Radiance Fields for Configuration-Aware Rendering, Lewis et al., 2022

DeepRob

Seminar 7
Explicit Scene-Level Representations
University of Michigan and University of Minnesota