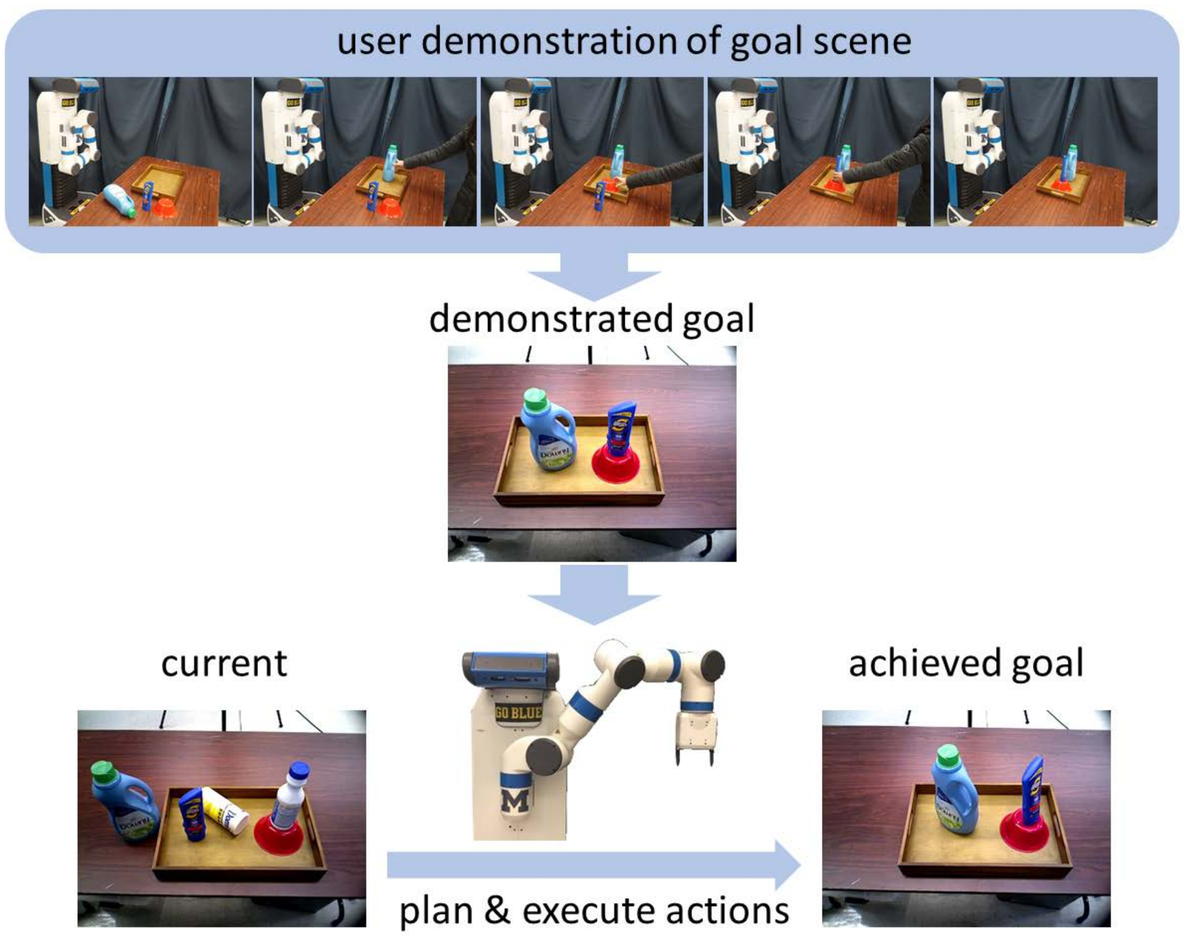
Semantic Robot Programming for Goal-Directed Manipulation in Cluttered Scenes

University of Michigan, Ann Arbor USA

Zhen Zeng Zheming Zhou Zhiqiang Sui Odest Chadwicke Jenkins

Introduction



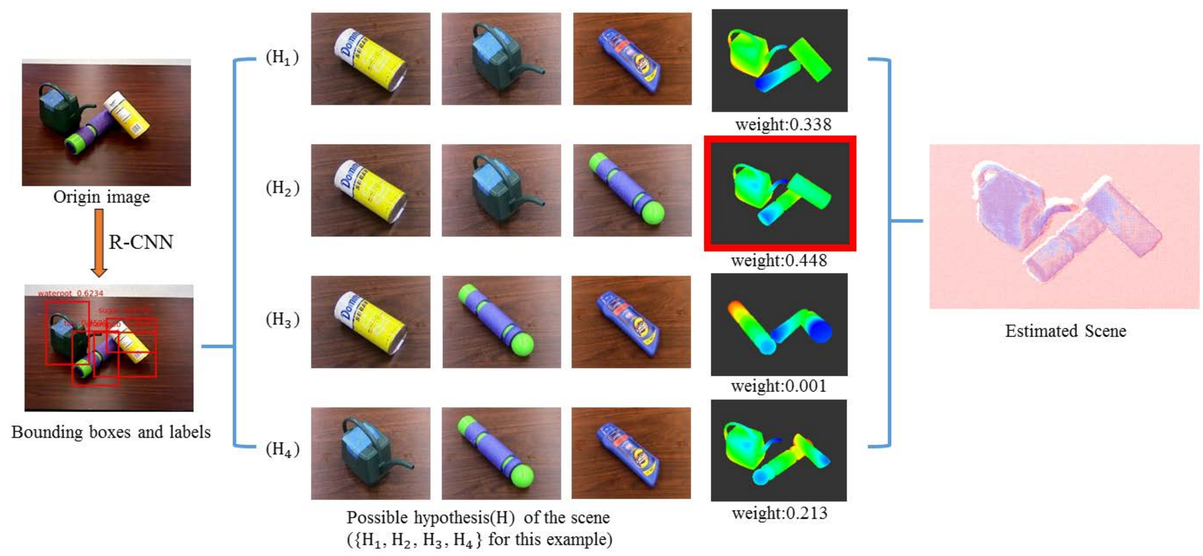
We present the **Semantic Robot Programming** (*SRP*) paradigm as a convergence of robot programming by demonstration and semantic mapping. In *SRP*, a user can declaratively program a robot manipulator by demonstrating a snapshot of their intended goal scene in workspace. The efficacy of *SRP* is demonstrated for the task of tray-setting with a Michigan Progress Fetch robot. Scene perception is evaluated on a public household occlusion dataset and our cluttered scene dataset.

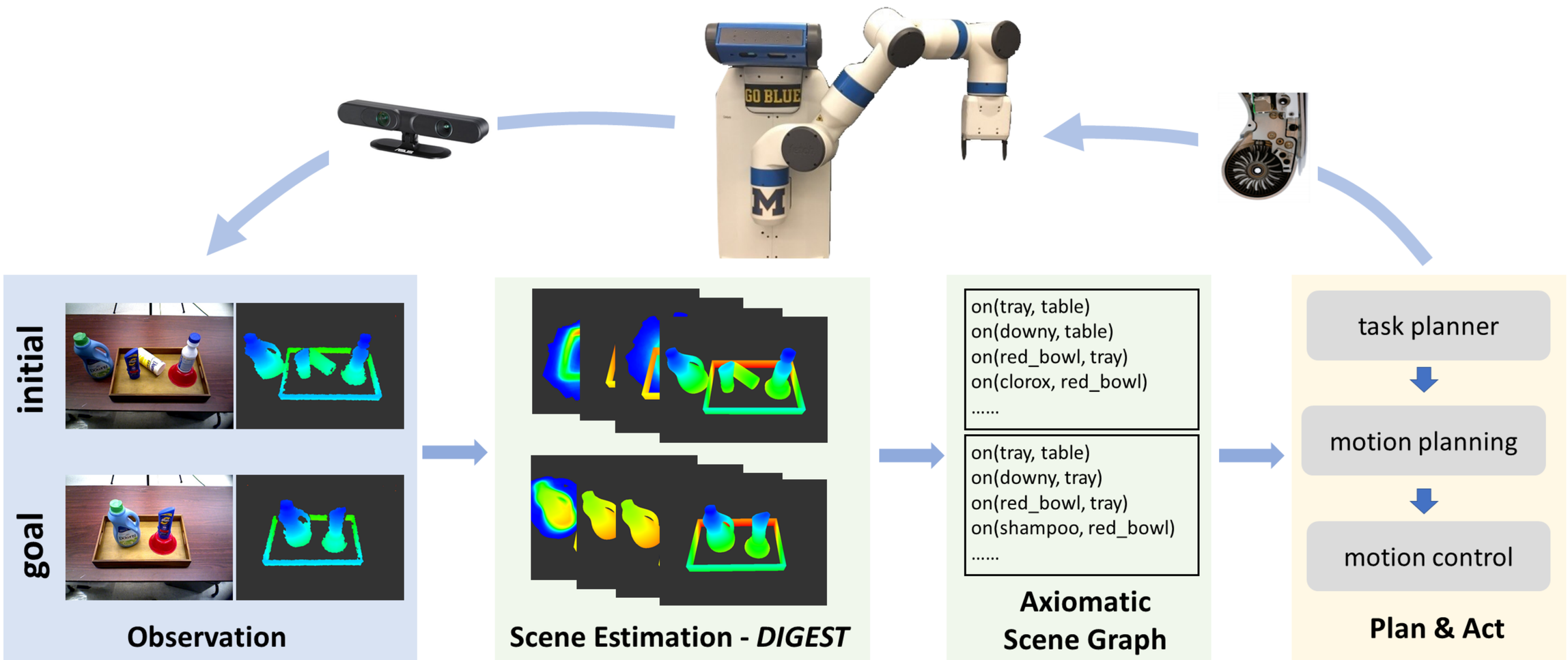
SRP Overview

Given an RGB-D observation of the previously demonstrated goal scene and current scene, the robot parses each as an axiomatic scene graph, comprised of object poses and inter-object relations. Then the robot plans and executes a sequence of manipulation actions to rearrange objects to a satisfying goal scene. *SRP* assumes the number of objects present, and 3D mesh models for all objects.

Scene Estimation

Within *SRP*, we propose **Discriminatively-Informed Generative Estimation of Scenes and Transforms** (*DIGEST*) to estimate scenes from RGB-D observations. DIGEST refines object detections from R-CNN through generative 6 DOF pose estimation. Axioms representing the estimated scene graph are heuristically derived from the inferred object poses.





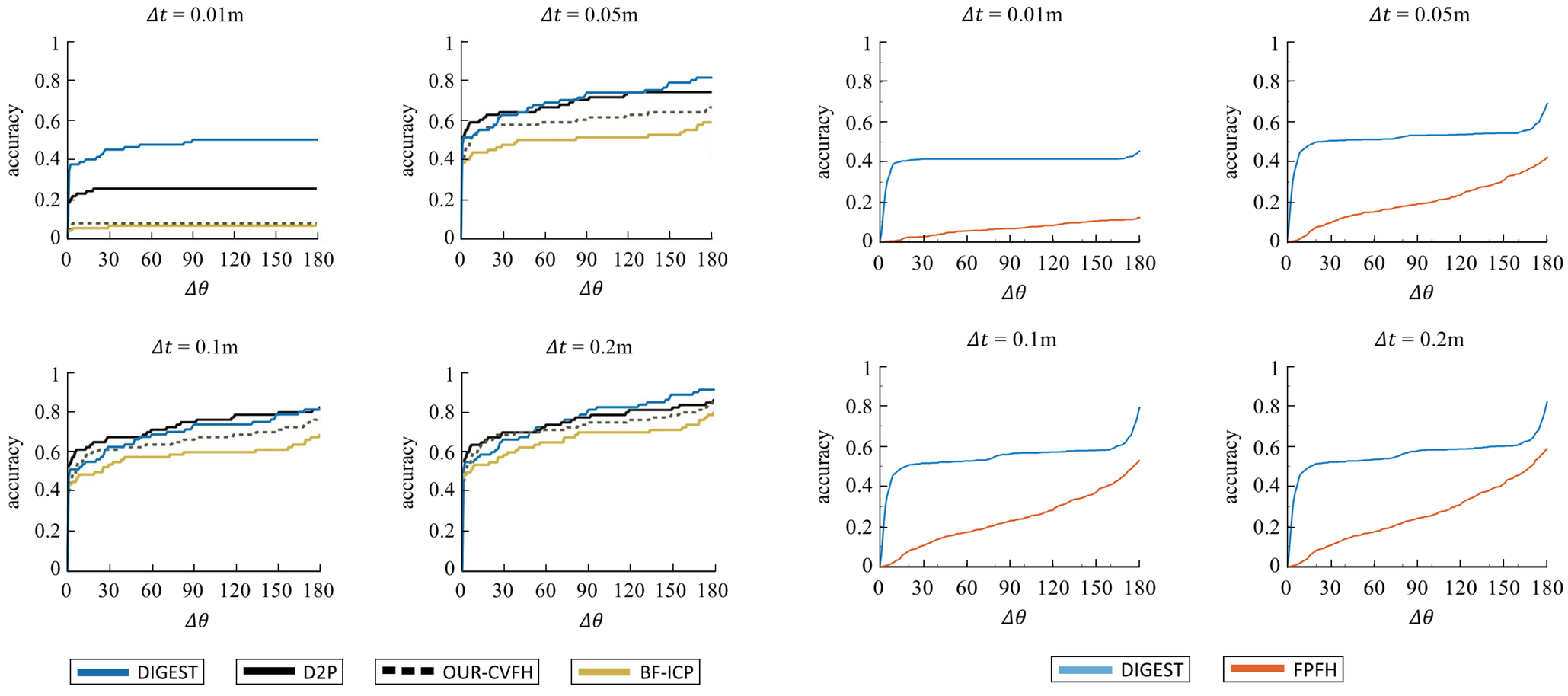
Semantic Robot Programming for Goal-Directed Manipulation in Cluttered Scenes

University of Michigan, Ann Arbor USA

Zhen Zeng Zheming Zhou Zhiqiang Sui Odest Chadwicke Jenkins

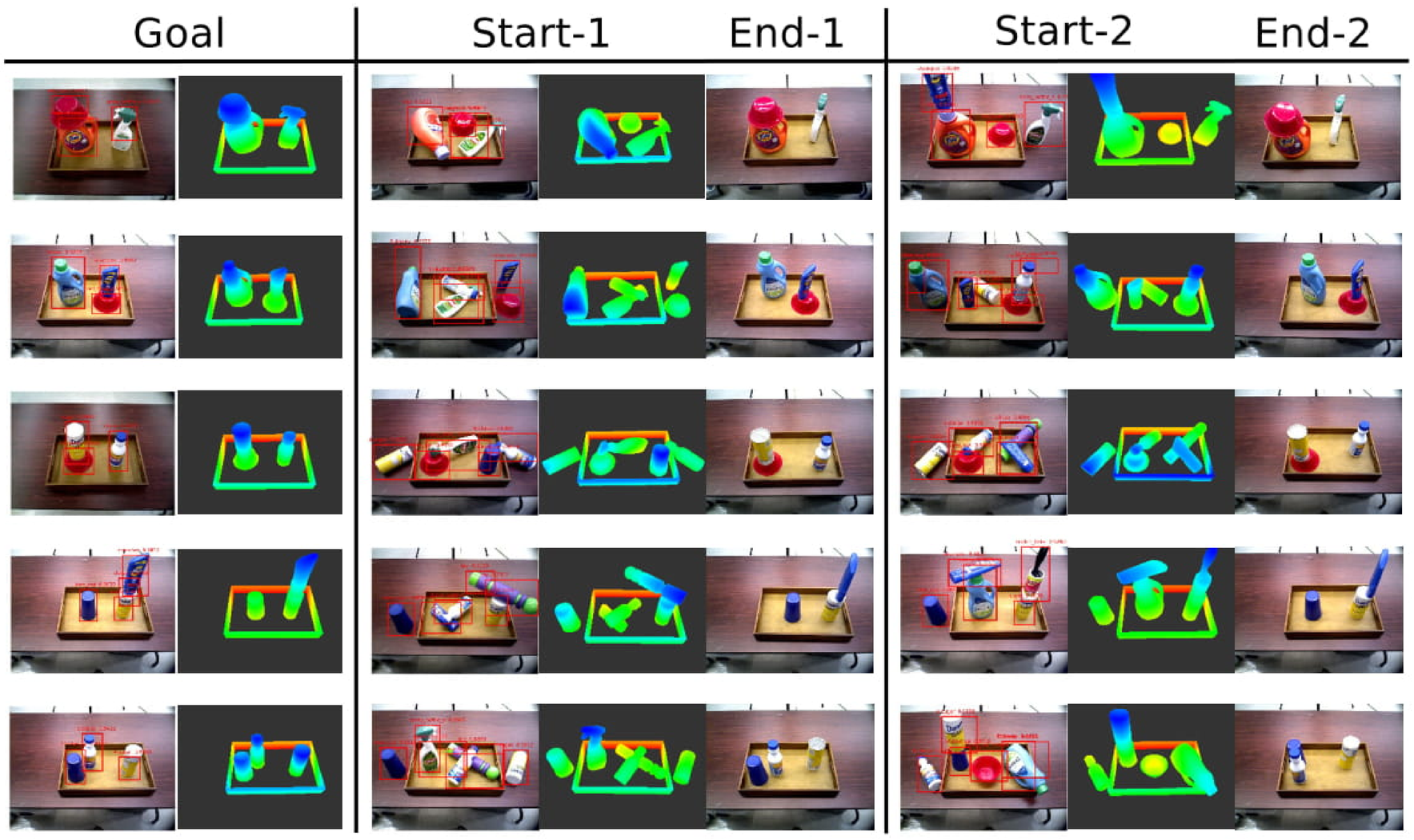
Experiments

We evaluate the proposed *DIGEST* scene estimation method on a public household occlusion dataset[1] and our cluttered scene dataset. *DIGEST* outperforms the state of the art method D2P [2] on the household occlusion dataset, and outperforms FPFH[3] on our cluttered scene dataset.



y-axis: pose estimation accuracy; $\Delta\theta$: rotation error threshold; Δt: translation error threshold

Our *SRP* experiments are designed around a service robot scenario, where the robot needs to prepare a tray as demonstrated by the user in the goal scene. *SRP* is tested on scenes of 4 to 6 objects including the tray, with different inter-object relations. When faced with different initial scene configurations, *SRP* enables the robot to seamlessly adapt to reach the demonstrated goal.



References

[1] Aitor Aldoma, Zoltan-Csaba Marton, Federico Tombari, Walter Wohlkinger, Christian Potthast, Bernhard Zeisl, Radu Bogdan Rusu, Suat Gedikli, and Markus Vincze. Point Cloud Library. IEEE Robotics & Automation Magazine, 1070(9932/12), 2012.

[2] Narayanan, Venkatraman, and Maxim Likhachev. "Discriminatively-guided Deliberative Perception for Pose Estimation of Multiple 3D Object Instances." *RSS*. 2016.

[3] R. B. Rusu, N. Blodow, and M. Beetz. Fast point feature histograms (fpfh) for 3d registration. In Robotics and Automation. *ICRA*. 2009.