Project Update: X-Ray for Lateral Access Mechanical Search

Michael Danielczuk
UC Berkeley
mdanielczuk@berkeley.edu

Anelia Angelova
Google Brain
anelia@google.com

Ken Goldberg
UC Berkeley
goldberg@berkeley.edu

1 Project Summary

Our efforts on mechanical search, or trying to efficiently extract a target object from a scene with occluding objects, started with our work in ICRA 2019 [1], where we formalized this class of problems and studied an instance where occluding objects are heaped over the target object in a bin. We presented a set of 5 baseline policies where the robot uses an RGBD perception system and control policies to iteratively select among, parameterize, and perform 3 actions – push, suction, grasp – until the target object is successfully retrieved. We realized from this effort that intermediate representations of the scene, such as a geometric prior on the occluded target object's location based on the geometries and poses of the other objects in the scene, could significantly boost the policy performance.

Figure 1: Mechanical search with a fully occluded target object (top row) and a partially occluded target object (bottom row). X-RAY predicts the target object occupancy distribution, which depends on the target object's visibility and the heap (second column). Each pixel value in the distribution image corresponds to the likelihood of that pixel containing part of the target object. X-Ray plans a grasp on the object that minimizes the estimated support of the resulting occupancy distribution to minimize the number of actions to extract the target object.

This insight led us to our IROS 2020 submission of X-Ray [2], a dataset generation pipeline and trained neural network that learns the *target occupancy distribution* using a synthetic dataset of RGBD heap images labeled for a set of standard bounding box targets with varying aspect ratios. We found that this intermediate representation could induce a greedy policy (grasp and remove the object

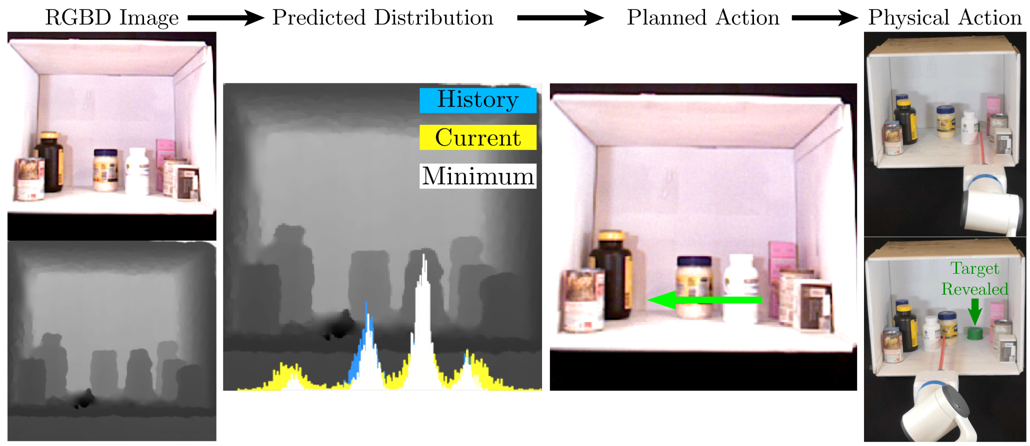

Figure 2: LAX-RAY's perception pipeline predicts an occupancy distribution over target object locations at the current step (yellow), the previous time step (blue), and the minimum of the two (white). LAX-RAY's mechanical-search policy computes a push action indicated by the green arrow, and the robot executes it.

that overlaps most with the predicted occupancy distribution at each timestep) which outperformed the policies introduced in the previous work. Figure 1 shows examples of the predicted target object occupancy distribution and action taken by the policy for both fully-occluded and partially-occluded target objects.

Recently, we have extended X-Ray to lateral-access environments such as shelves and cabinets in a paper that will be presented at IROS 2021 [3]. LAX-RAY (Lateral-Access X-Ray) addresses several new challenges and constraints introduced in the lateral-access setting. For example, objects on a shelf must lie in one of their stable poses, as opposed to being in an arbitrary 6D pose when lying in a heap. The action space is also restricted due to the presence of the shelf walls, which can make pushing actions more accessible and efficient than grasping actions in this setting. In LAX-RAY, we adapt the occupancy distribution training process from X-Ray to account for these constraints, as well as shelf depth and camera perspective effects. We propose two pushing polices that encode a history of occupancy distribution predictions and lookahead to avoid repeated actions. Figure 2 shows an example of lateral-access mechanical search using the LAX-RAY policy on a shelf using a physical Fetch robot.

2 Takeaways and Current Work

Intermediate representations such as a target occupancy distribution can greatly increase the efficiency of the mechanical search problem. However, in the lateral-access problem, novel challenges and constraints require further perception and policy improvements. We are currently exploring how to mitigate some of the assumptions in LAX-RAY, such as pushing multiple objects, stacked objects, or rotating objects through pushing. We also are investigating expanding the action set by using pneumatically-activated suction cups to lift and pull occluding objects from the shelf.

References

[1] M. Danielczuk, A. Kurenkov, A. Balakrishna, M. Matl, D. Wang, R. Martín-Martín, A. Garg, S. Savarese, and K. Goldberg, "Mechanical search: Multi-step retrieval of a target object occluded by clutter," in *Proc. IEEE Int. Conf. Robotics and Automation (ICRA)*, 2019, pp. 1614–1621.

[2] M. Danielczuk, A. Angelova, V. Vanhoucke, and K. Goldberg, "X-ray: Mechanical search for an occluded object by minimizing support of learned occupancy distributions," in *Proc. IEEE/RSJ Int. Conf. on Intelligent Robots and Systems (IROS)*, 2020.

[3] H. Huang, M. Dominguez-Kuhne, J. Ichnowski, V. Satish, M. Danielczuk, K. Sanders, A. Lee, A. Angelova, V. Vanhoucke, and K. Goldberg, "Mechanical search on shelves using lateral access x-ray," in *Proc. IEEE/RSJ Int. Conf. on Intelligent Robots and Systems (IROS)*, 2021.