Reactive Whole-Body Control for Grasping Moving Objects with Mobile Manipulators

Ahmet Tekden[1*] Francesco Gigante[1*] Shijie Zhou[1*] Dongxiang Li[1*] Yasemin Bekiroglu[1,2]

***Abstract*— Grasping moving objects remains a challenging task for fixed manipulators due to their limited workspace, which restricts their ability to respond to object motion. This limitation becomes particularly important in human-robot collaborative scenarios. In this work, we propose a framework for grasping moving objects with a mobile manipulator that dynamically repositions its base to expand its reachable workspace and improve grasp feasibility. To coordinate the system motion, we adopt a reactive whole-body control strategy that treats the mobile base and manipulator as a unified kinematic system. As the object moves, the target pre-grasp pose is updated based on the estimated object pose, allowing the robot to track and approach a grasp-relevant target. The proposed method is validated in both simulation and real-world experiments, demonstrating the effectiveness of the unified control strategy for coordinated tracking and grasp-oriented motion generation when interacting with moving objects.**

I. INTRODUCTION

Stationary robot manipulators have succeeded in tasks such as grasp planning [1]–[3], throwing objects into bins [4], and pick-and-place [5]; however, their effectiveness depends on being well positioned relative to objects. As a result, tasks must be designed so the robot can reach and manipulate targets from its initial pose. In contrast, mobile manipulators can perform similar tasks with fewer placement constraints.

The manipulation task becomes even more challenging when the target object moves or its pose changes over time. In such scenarios, the robot must continuously estimate the object state and adapt its motion accordingly. This requires real-time integration of perception, tracking, and control, where inaccuracies in any one of these components can lead to failure. Dynamic grasping approaches [6]–[10] address this problem by tracking the object pose, generating grasp or pre-grasp targets relative to the object, and controlling the manipulator to follow these poses during motion. Although these methods have shown promising results, they are primarily designed for fixed manipulators operating within a constrained workspace.

When the object moves outside the reachable region of the arm, the task may become infeasible regardless of the quality of the perception or control strategy. Mobile manipulators provide a natural way to overcome this limitation by enabling the robot to reposition its base and expand the feasible grasping region. However, in many mobile manipulation systems, robot placement and manipulation are treated as separate stages, where the robot first plans and moves to a suitable base configuration, for example via reachability analysis [3], [11]–[13], before executing the manipulation task. While effective for many static manipulation scenarios, this separation can limit the robot's ability to reactively adapt its motion when interacting with moving objects.

This work was supported by by Chalmers AI Research Center (CHAIR), Gender Initiative for Excellence (Genie), and the Wallenberg AI, Autonomous Systems and Software Program (WASP). *Equal contribution, [1]Chalmers University of Technology, [2] University College London.

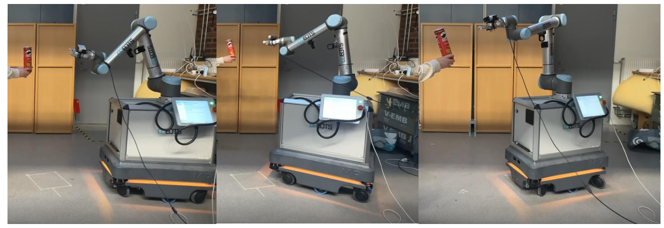


Fig. 1: Closed-loop tracking of an object: perception updates the object pose and the robot reacts with whole-body motion.

In this work, we address these challenges by enabling reactive grasp pose tracking with a mobile manipulator through unified control of the mobile base and manipulator. Instead of treating base and arm motion as separate stages, the controller coordinates both within a unified control framework, enabling the robot to continuously follow grasp-related target poses generated from the moving object, and avoids obstacles during motion.

II. METHOD

A high-level overview of the proposed four-stage system is depicted in Fig. 2. The system processes acquired scene depth data into a filtered point cloud, which is used to segment the target, estimate its pose, and determine a corresponding grasp pose. Based on this, a pre-grasp pose is selected to serve as the target pose for the holistic controller, which then guides the robot's motion. The pipeline operates iteratively, enabling target pose updates as the object moves.

The Control Approach: We adopt the whole-body control framework of [14], where the mobile base and manipulator are treated as a unified kinematic system. To incorporate base motion into the differential kinematics, the base is modeled as virtual degrees of freedom, allowing the robot to be represented as an equivalent kinematic chain. Base and joint velocities are then computed through a velocity-based whole-body control formulation expressed as a constrained optimization problem. To prevent the target object from leaving the camera field of view, we introduce a field-of-view constraint based on the estimated object pose in the camera frame, which is used to monitor the object position in the image plane. As the object approaches the image boundaries, the allowable camera velocity in the corresponding

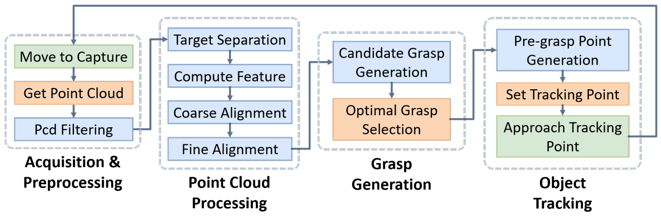


Fig. 2: Vision-guided tracking, grasping and control

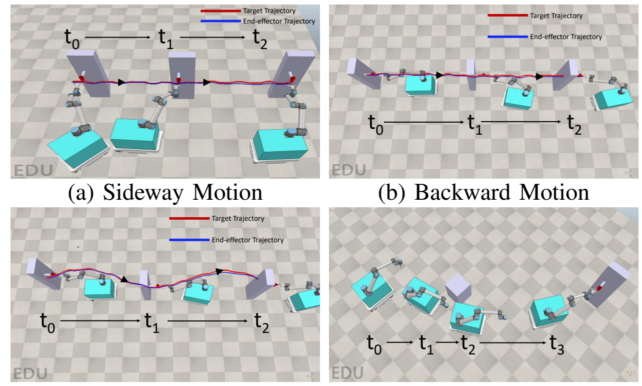
direction is progressively reduced to keep the object within the image bounds. This constraint is incorporated into the whole-body controller as an additional inequality constraint. Additionally, we define safety zones around obstacles on the map. When the robot is far from an obstacle, navigation mode is activated using ROS Navigation [15]). In close proximity to obstacles, we incorporate distance-dependent velocity inequality constraints into the control framework to guarantee the tracking performance. These constraints ensure that the robot's velocity asymptotically attenuates to zero as it nears an obstacle, thereby providing a formal guarantee for collision-free motion.

Target Pose Generation for Grasping Moving Objects: The system processes depth observations to obtain a point cloud representation of the scene, which is filtered and segmented to isolate the target object using the Density-Based Spatial Clustering of Applications with Noise (DBSCAN) algorithm [16]. To maintain and update the estimate of the object pose over time, standard registration techniques are employed, including coarse alignment using Random Sample Consensus (RANSAC) [17] and fine alignment using Point-to-Plane Iterative Closest Point (ICP) [18], together with point cloud completion. Based on the estimated object pose, a grasp pose is chosen from a curated set according to task-relevant constraints, defining the corresponding pre-grasp target used as the reference for control. As the object moves, the target pose is continuously updated and followed by the controller in a closed-loop manner.

III. Results

We validate the proposed method in both simulation and real-world settings. The experimental platform consists of a UR10 robotic arm equipped with an RG2-FT gripper, mounted on a MIR200 mobile base, with a wrist-mounted RGB-D camera (Femto Bolt [19] in the real-world setup). The experiments evaluate the robot's ability to continuously follow object-relative pre-grasp target poses as the object moves.

In simulation, we evaluate whether the proposed controller can follow target poses under object motion and handle obstacle avoidance. To this end, we consider three representative target motion patterns with noise: sideways, backwards, and nonlinear trajectories, capturing both motion dynamics and directional changes. As shown in Fig. 3a–3c, the robot is able to reach an initial pre-grasp configuration and continuously track the updated target poses



(a) Sideway Motion (b) Backward Motion (c) Nonlinear Motion (d) Obstacle Avoidance

Fig. 3: Example trajectories and obstacle avoidance. Red: target noisy trajectory; Blue: end-effector trajectory.

throughout the object motion. Quantitatively, the average pre-grasp tracking errors are 6.21 cm for the sideways motion, 3.20 cm for the backward motion, and 3.34 cm for the nonlinear motion. We further evaluate the controller in an obstacle avoidance scenario, illustrated in Fig. 3d, where the robot safely avoids the obstacle while still reaching the target pre-grasp pose. Overall, these results demonstrate the effectiveness of the control strategy for coordinated and reactive motion generation for moving targets with obstacle avoidance.

In real-world experiments, we verify the framework using a hand-held target object. As shown in Fig. 1, the system can detect the target object and estimate its pose from the observed point cloud. As the object is moved by a human, the robot maintains tracking behavior and follows the updated target pose. These results indicate that the proposed framework can operate in real-world settings and support reactive tracking of grasp-related target poses for moving objects.

IV. Conclusion

In this work, we present a framework for grasping of moving objects using a mobile manipulator with whole-body control. By coordinating the motion of the mobile base and manipulator within a unified control formulation, the system can adapt its motion in real time to maintain feasible grasp configurations. Future work will focus on more extensive evaluation of the proposed method, improving robustness and extending the system to operate in more complex environments. In particular, we aim to incorporate active camera viewpoint selection for improved object reconstruction and tracking, integrate predictive models of object motion to better anticipate target movement, and develop more adaptive base motion strategies that account for task and environmental constraints.

References

[1] A. Tekden, M. P. Deisenroth, and Y. Bekiroglu, "Grasp transfer based on self-aligning implicit representations of local surfaces," *IEEE Robotics and Automation Letters*, vol. 8, no. 10, pp. 6315–6322, 2023.

[2] A. Ten Pas, M. Gualtieri, K. Saenko, and R. Platt, "Grasp pose detection in point clouds," *The International Journal of Robotics Research*, vol. 36, no. 13-14, pp. 1455–1473, 2017.

[3] F. Burget and M. Bennewitz, "Stance selection for humanoid grasping tasks by inverse reachability maps," in *2015 IEEE International conference on robotics and automation (ICRA)*. IEEE, 2015, pp. 5669–5674.

[4] A. Zeng, S. Song, J. Lee, A. Rodriguez, and T. Funkhouser, "Tossingbot: Learning to throw arbitrary objects with residual physics," *IEEE Transactions on Robotics*, vol. 36, no. 4, pp. 1307–1319, 2020.

[5] M. Shridhar, L. Manuelli, and D. Fox, "Perceiver-actor: A multi-task transformer for robotic manipulation," in *Conference on Robot Learning*. PMLR, 2023, pp. 785–799.

[6] F. Husain, A. Colomé, B. Dellen, G. Alenya, and C. Torras, "Realtime tracking and grasping of a moving object from range video," in *2014 IEEE International Conference on Robotics and Automation (ICRA)*. IEEE, 2014, pp. 2617–2622.

[7] N. Marturi, M. Kopicki, A. Rastegarpanah, V. Rajasekaran, M. Adjigble, R. Stolkin, A. Leonardis, and Y. Bekiroglu, "Dynamic grasp and trajectory planning for moving objects," *Autonomous Robots*, vol. 43, no. 5, pp. 1241–1256, 2019.

[8] M. Tuscher, J. Hörz, D. Driess, and M. Toussaint, "Deep 6-dof tracking of unknown objects for reactive grasping," in *2021 IEEE International Conference on Robotics and Automation (ICRA)*. IEEE, 2021, pp. 14 185–14 191.

[9] C.-C. Wong, M.-Y. Chien, R.-J. Chen, H. Aoyama, and K.-Y. Wong, "Moving object prediction and grasping system of robot manipulator," *IEEE Access*, vol. 10, pp. 20 159–20 172, 2022.

[10] E. Maranci, C. A. Avizzano, and S. D'Avella, "Reactive grasp controller for dynamic grasping of moving objects," *Robotics and Autonomous Systems*, p. 105401, 2026.

[11] F. Zacharias, W. Sepp, C. Borst, and G. Hirzinger, "Using a model of the reachable workspace to position mobile manipulators for 3-d trajectories," in *2009 9th IEEE-RAS International Conference on Humanoid Robots*. IEEE, 2009, pp. 55–61.

[12] N. Vahrenkamp, T. Asfour, and R. Dillmann, "Robot placement based on reachability inversion," in *2013 IEEE International Conference on Robotics and Automation*. IEEE, 2013, pp. 1970–1975.

[13] A. Makhal and A. K. Goins, "Reuleaux: Robot base placement by reachability analysis," in *2018 Second IEEE International Conference on Robotic Computing (IRC)*. IEEE, 2018, pp. 137–142.

[14] J. Haviland, N. Sünderhauf, and P. Corke, "A Holistic Approach to Reactive Mobile Manipulation," *IEEE Robotics and Automation Letters*, vol. 7, no. 2, pp. 3122–3129, Apr. 2022, arXiv:2109.04749 [cs]. [Online]. Available: http://arxiv.org/abs/2109.04749

[15] Open Robotics, "Ros navigation," https://wiki.ros.org/navigation, rOS Wiki, accessed 2026-04-04.

[16] M. Ester, H.-P. Kriegel, J. Sander, and X. Xu, "A density-based algorithm for discovering clusters in large spatial databases with noise," in *Proceedings of the Second International Conference on Knowledge Discovery and Data Mining*, ser. KDD'96. AAAI Press, 1996, p. 226–231.

[17] M. A. Fischler and R. C. Bolles, "Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography," *Communications of the ACM*, vol. 24, no. 6, pp. 381–395, 1981.

[18] P. J. Besl and N. D. McKay, "Method for registration of 3-D shapes," in *Sensor Fusion IV: Control Paradigms and Data Structures*, P. S. Schenker, Ed., vol. 1611, International Society for Optics and Photonics. SPIE, 1992, pp. 586 – 606. [Online]. Available: https://doi.org/10.1117/12.57955

[19] Orbbec, "Femto bolt," https://www.orbbec.com/products/tof-camera/femto-bolt/, accessed: 2026-04-04.