Navigation Among Movable Obstacles with Mobile Manipulator using Learned Robot-Obstacle Interaction Model

Taegeun Yang[1], Minsung Yoon[1], Jeil Jeong[2] and Sungeui Yoon[1†]

***Abstract*— In this paper, we address the online Navigation Among Movable Obstacles (NAMO) problem by employing a mobile manipulator. Unlike mobile robots, mobile manipulators offer the advantage of effectively relocating obstacles out of the driving path while tracking a global path. However, the high degrees of freedom (DOF) of mobile manipulator complicates whole-body control. To address these challenges, we propose a Reinforcement Learning (RL) based Model Predictive Path Integral (MPPI) framework. This strategy includes identifying actions for stable pushing through RL, training robot-obstacle kinodynamic interaction model from policy-generated data, and applying this model in MPPI to maneuver obstacles while tracking the global path. In our experiment, we demonstrated that our method successfully pushes obstacles aside and maintains adherence to the global path when it is obstructed.**

Index Terms-Navigation Among Movable Obstacles (NAMO), Mobile Manipulator, RL based MPPI

I. INTRODUCTION

Robotic navigation in diverse real-world environments is crucial for task execution, such as warehouse management and delivery. Recent advances in mobile navigation field have enabled robots to navigate successfully in challenging environments including narrow passages and dynamic obstacles [1], [2]. Nonetheless, the predominant focus of navigation research has been on collision avoidance. While important, avoiding all obstacles can sometimes significantly increase the travel distance and time. In addition, this approach may lead to immobility problems, where the robot becomes stuck, surrounded by obstacles and unable to find the path to reach the goal [3]. In such scenarios, the ability of a robot to interact with obstacles to actively find a drivable path is important for navigation, and this problem is called Navigation Among Movable Obstacles (NAMO).

The NAMO problem is generally recognized as NP-hard because of the extensive search space involving the robot's movement, obstacle manipulation, and evolving configuration space over time [4]. Thus, many studies have relied on assumptions such as the robot aligning itself axially with obstacles to limit search space while pushing and simplifying the class of obstacles into binary categories of movable or immovable [5], [6]. However, such assumptions reduce the maneuverability of obstacles during pushing and do not capture their physical properties such as mass and friction coefficient. Therefore, our objective is to employ a mobile manipulator and robot-obstacle kinodynamic interaction model to mitigate these limitations. Utilizing a mobile manipulator, which enables the separation of arm pushing from base movement, can significantly enhances maneuverability (Fig. 1). Furthermore, we utilize estimated physical properties from interactions to predict robot and obstacles future trajectories with a kinodynamic interaction model.

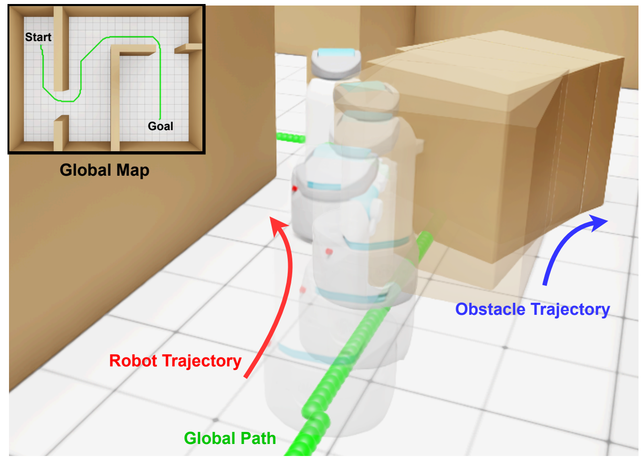

Fig. 1. Mobile manipulator whole-body control for NAMO problem: While tracking a global path (green), if the robot detects an obstacle with a potential for collision, it actively pushes the obstructing obstacles, enabling collision free navigation to the goal. The red arrow represents the movement of the robot as it pushes the obstacle, and the blue arrow indicates the movement of the obstacle. Left top is the global map and planned global path using A* algorithm.

There are several challenges in using a mobile manipulator to push obstacles. Maintaining contact is crucial for stable pushing [7], yet this becomes difficult when executing whole-body control with a mobile manipulator that possesses high degrees of freedom (DOF). Additionally, the reachability of the end-effector must be considered. To handle these challenges, we propose Reinforcement Learning (RL) based Model Predictive Path Integral (MPPI) approach. Through RL, we first identify a set of actions that maintains stable contact with obstacles while executing given command, and train a robot-obstacle kinodynamic interaction model using data obtained through a trained policy. Subsequently, we utilize the trained models within the MPPI framework and determine the finite-horizon optimal action sequence from the given set of actions derived from RL to generate drivable path when encountering obstacles.

Our contributions are summarized as follows:

- We propose an RL based MPPI approach to address the complexities of whole-body control in mobile ma-

[1]T. Yang, M. Yoon, and S. Yoon are with the School of Computing in Korea Advanced Institute of Science and Technology (KAIST), Daejeon, 34141, Republic of Korea.

[2]J. Jeong is with the Robotics Program at the same institute, KAIST.

[†]S. Yoon is a corresponding author; sungeui@kaist.edu.

nipulators. It facilitates stable and effective obstacle manipulation in the context of NAMO settings.

- We propose a physical properties informed robot-obstacle kinodynamic interaction model for predicting the future trajectory of robot and obstacles, which is then utilized in planning.

II. RELATED WORK

A. Navigation Among Movable Obstacles (NAMO)

The Navigation Among Movable Obstacles (NAMO) problem presents a major challenge in robotics, addressing the complex task of reasoning through environments where robots can strategically move obstacles to create free space for driving. Wilfong initially demonstrated the NP-hard of the NAMO problem, attributing its complexity to dynamic configuration spaces and extensive search space, even under simplified conditions [4]. Further research has aimed at enhancing algorithmic efficiency, either by limiting the action space or by developing composite planners that consist of several task-specific sub-planners [8].

Aforementioned offline approaches traditionally presuppose complete prior environmental knowledge, prompting recent research interest in online NAMO solutions that begin with only partial knowledge. [9] first tackled the online NAMO challenge within unknown environments characterized by both static and movable obstacles, where the robot proactively learns obstacle mobility through interaction. Meanwhile, [5] focused on real-world NAMO challenges, leveraging a photorealistic simulator for efficient training data generation, and facilitating sim2real transfer through direct visual application on real robots.

However, these studies merely categorize obstacles mobility into binary states—movable or immobile—and confine interactions to axis-aligned pushes. Such constraints notably limit the robot's maneuverability with obstacles and obstacles' physical properties. In contrast, our approach discerns physical properties through active interactions, employing a mobile manipulator to enhances obstacle maneuverability through arm utilization, allowing the base to adhere to the global path while displacing obstacles from the driving path.

B. RL based MPPI

Model Predictive Path Integral Control (MPPI) is a sampling-based MPC approach where an agent determines the optimal control sequence at each time step using online rollouts. By leveraging the Feynman-Kac lemma [10], the MPPI approach addresses non-linear Stochastic Optimal Control (SOC) problems by replacing the Hamilton-Jacobi-Bellman equation with evaluations of expected future trajectories through Monte-Carlo sampling. Consequently, this method enables the identification of the optimal control sequence based on these predictive assessments of trajectories.

Due to MPPI's suitability for non-linear, non-convex dynamic systems, there has been significant interest in its application for robotic control. [11] demonstrates the real-time application on high-dimensional robots like manipulators across various tasks. Meanwhile, [12] introduces a variant of

TABLE I
VARIABLE REPRESENTATION AND REWARD FUNCTIONS

Notation		Components	Dim	Total
state $\mathbf{s}_t$	observation $\mathbf{o}_t$	q_{arm}	7	23
		v_{base}	2	
		$\{pose_{obj}\}^{robot}$	4	
		$\{pose_{ee}\}^{robot}$	4	
		$\{vel_{obj}\}^{robot}$	3	
		$\{vel_{ee}\}^{robot}$	3	
	privileged $\mathbf{p}_t$	m_{obj}	1	7
		COM_{obj}	3	
		$friction\ coeff_{obj}$	2	
		$restitution_{obj}$	1	
command $\mathbf{c}_t$		v_{base}^{target}	2	3
		$push\ point$	1	

Reward	Expression
r_1^{cmd}	follow target vel
r_2^{cmd}	follow push point
r_3^{constr}	continue contact
r_4^{constr}	perpenticular contact
r_5^{constr}	maintain ee height
r_6^{smooth}	smooth arm action
$\mathbf{r}_{tot} = k_1 r_1 + k_2 r_2 + k_3 r_3 + k_4 r_4 + k_5 r_5 + k_6 r_6$	

MPPI designed to react safely to static and dynamic obstacles while executing tasks, leveraging a signed distance function for enhanced safety and efficiency.

MPPI has advanced significantly in the domain of robotic control, yet it encounters a challenge where the quality of rollouts critically impacts its optimality [13]. Specifically, in robots with a high degrees of freedom, such as mobile manipulators, identifying the optimal action sequence through random sampling within the large configuration space is difficult. [14] suggests using offline reinforcement learning for generating initial rollouts, leveraging learned stochastic action distribution as starting parameters. This approach has outperformed traditional MPPI methods in UAV control tasks. We further develop this approach by integrating commands into reinforcement learning, ensuring compliance with certain constraints, and applying MPPI in the command space. This adjustment not only allows for a wider range of movements than traditional approaches but also boosts the likelihood of identifying an optimal action sequence within predefined constraints. It simultaneously enhances the robot's adaptability to dynamic environments and unforeseen situations through a broader exploration of potential scenarios.

III. METHOD

Our task focuses on driving a mobile manipulator through obstacles, utilizing whole-body control to displace obstructions encountered while navigating the global path. We assume that obstacles are uniform in size and shape but differ in physical properties without any prior knowledge. We employ a hierarchical strategy that first utilizes RL for both stable pushing of obstacles without base collision and action generation to execute given commands, and then applies

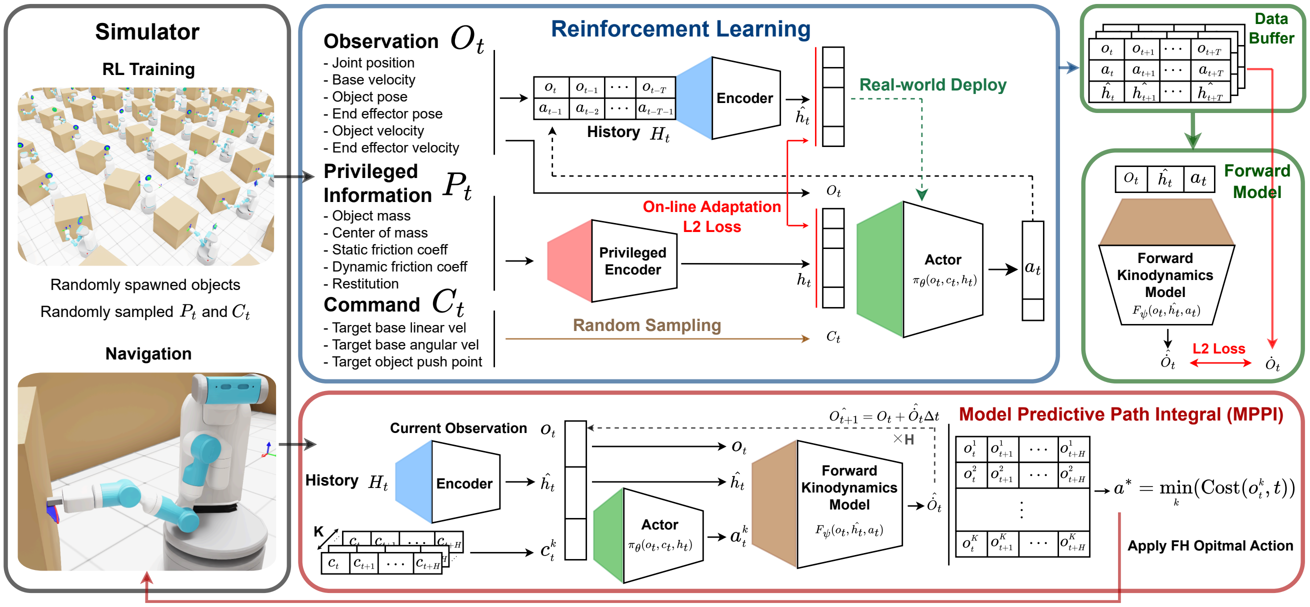

Fig. 2. Overall framework: We first train a policy π_θ via reinforcement learning that enables the robot to stably push obstacles while executing given commands. Subsequently, data collected using this policy is employed to train a robot-obstacle kinodynamics interaction model F_ϕ, which is then utilized as the model for MPPI control.

MPPI control within the command space. Our system is illustrated in Fig. 2 and we provide a detailed description of each component in the following sections.

A. Training Whole-Body Stable Pushing Policy

We first utilize reinforcement learning to train the mobile manipulator to follow given commands while stably pushing obstacles during its movement. The definitions of commands and variables, as well as the reward function, are detailed in Table. I. The observation o_t encompasses both proprioception and exteroception, obtainable from the onboard sensors. Privileged information is derived from randomly sampling the physical properties of obstacles within predetermined ranges. Regarding commands, the target object push point, essential for robot-obstacle maneuverability, indicates the point on the obstacle's surface to be pushed, chosen randomly within a specified range. To simplify the end-effector's movement, we facilitate its maintenance of a consistent height. Stable pushing of an obstacle necessitates maintaining contact, with perpendicular alignment enhancing this stability. The reward function is therefore tailored to ensure these conditions are satisfied while executing the commands. To reduce the sim2real gap, online adaptation approach is employed [15].

B. MPPI Control using Learned Kinodynamics Model

An policy network π_θ, trained via reinforcement learning, is employed to collect data for training a robot-obstacle kinodynamics interaction model. This model F is designed to predict the temporal derivative of observations $\dot{o}_t$ based on current observations o_t, actions a_t, and estimated physical properties of obstacles $\hat{h}_t$. Subsequently, the trained model is utilized as the forward model in the MPPI framework. It generates random command sequences within the command space and performs rollout for future observations corresponding to each sequence. It then determines the optimal command sequence based on the cost associated with each rollout, subsequently executing the corresponding action. The cost function utilized is specified as follows:

$$C(o_t) = w_1 Speed(o_t) + w_2 Track(o_t) + w_3 Push(o_t) \quad (1)$$

The $Speed$ and $Track$ terms encourage the robot to tracking the global path at an fast speed, while the $Push$ term aims to push obstacles away from the robot's driving path. Utilizing these terms facilitates the generation of a path that allows the robot to navigate to the next waypoint without collisions.

C. Navigation Strategy

We address an online NAMO task with the initial assumption of possessing a global map for static obstacles, such as walls. Utilizing this map, a global path is determined via the A* algorithm. Upon encountering obstacles that could potentially lead to collisions while following the global path, the RL based MPPI is deployed to dynamically generate and navigate a collision-free path to the next waypoint. In this context, the next waypoint is set as the nearest position within a fixed distance from the robot. This process is iterated until the goal is reached.

IV. EXPERIMENT AND RESULT

As this research is currently on process, only qualitative results will be discussed. The global map used for navigation can be viewed in Fig. 1. This section addresses a scenario where obstacles are present in the corridor through which the robot must travel to reach its goal, known as the key-hole problem. The robot navigates along the global path, and upon detecting obstacles that may lead to collisions, it actively generates a collision free path by displacing the obstacles.

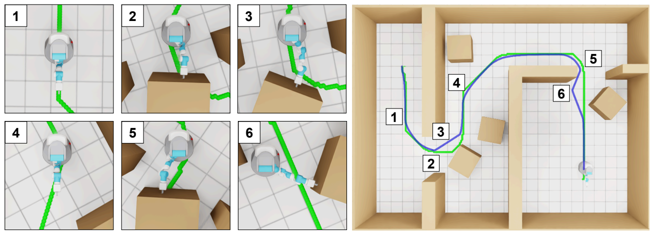

Fig. 3. Navigation Result: Left images depict the robot pushing obstacles while following the global path. Right image illustrates the result of navigation, with the green path representing the precomputed global path and the blue path is the robot's traveled path.

A. Environment Setting

To demonstrate the robot's capability to create collision free paths by pushing obstacles, they were strategically placed within passages that the robot must traverse. In such scenarios, methods solely focused on collision avoidance are insufficient for reaching the goal. Additionally, to assess the robot's ability to evaluate collision risks, obstacles that do not present a risk were also interspersed. As it conducted in simulation, we assumed that the pose of boxes can be precisely detected, although the physical properties of the obstacles remain unknown.

B. Navigation Result

The navigation result is visualized in Fig. 3, where the robot starts from the top left of the map and moves towards the goal located at the bottom right. The images on the left are indexed in chronological order, with images 1 and 4 depicting following the global path, while images 2, 3, 5 and 6 illustrates the pushing action. To navigate without base collisions, the robot must push aside two obstacles, while the remaining obstacles pose no collision risk. The green path represents the global path derived solely from the global map, whereas the blue path illustrates the robot's traversed path. Significantly, the robot exhibits behavior aimed at adhering to the global path, actively displacing obstacles identified as collision risks.

V. CONCLUSION AND FUTURE WORKS

We proposed a method to solve the online NAMO problem in a partially known environment using a mobile manipulator with RL based MPPI. The proposed method mitigates the limited obstacle maneuverability caused by the axis-aligned constraint in previous studies and enables safe driving without collision of the base. In addition, the problem of extensive search space when performing whole-body control on a mobile manipulator was solved through reinforcement learning. Using simulation, we show that the proposed method can successfully navigate in the NAMO setting. However, in practice, when the robot pushes the obstacles, it may not be moved by a wall or other obstacles. Also, it may be more efficient to bypass the obstacle if it is not easily pushed. Addressing these issues as future work will make online NAMO task works further in real world scenarios.

ACKNOWLEDGMENT

This work was supported by the National Research Foundation of Korea (NRF) and the Institute of Information & communications Technology Planning & Evaluation (IITP), grant funded by the Korea government (MSIT) (No. RS-2023-00208506(2024) and RS-2023-00237965)

REFERENCES

[1] Ankit A Ravankar, Abhijeet Ravankar, Takanori Emaru, and Yukinori Kobayashi, "Hpprm: hybrid potential based probabilistic roadmap algorithm for improved dynamic path planning of mobile robots", *Ieee Access*, vol. 8, pp. 221743–221766, 2020.

[2] Utsav Patel, Nithish K Sanjeev Kumar, Adarsh Jagan Sathyamoorthy, and Dinesh Manocha, "Dwa-rl: Dynamically feasible deep reinforcement learning policy for robot navigation among mobile obstacles", in *2021 IEEE International Conference on Robotics and Automation (ICRA)*. IEEE, 2021, pp. 6057–6063.

[3] Kai Zhang, Eric Lucet, Julien Alexandre dit Sandretto, and David Filliat, "Navigation among movable obstacles using machine learning based total time cost optimization", in *2023 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. IEEE, 2023, pp. 11321–11327.

[4] Gordon Wilfong, "Motion planning in the presence of movable obstacles", in *Proceedings of the fourth annual symposium on Computational geometry*, 1988, pp. 279–288.

[5] Kirsty Ellis, Henry Zhang, Danail Stoyanov, and Dimitrios Kanoulas, "Navigation among movable obstacles with object localization using photorealistic simulation", in *2022 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. IEEE, 2022, pp. 1711–1716.

[6] Kirsty Ellis, Denis Hadjivelichkov, Valerio Modugno, Danail Stoyanov, and Dimitrios Kanoulas, "Navigation among movable obstacles via multi-object pushing into storage zones", *IEEE Access*, vol. 11, pp. 3174–3183, 2023.

[7] Kevin M Lynch and Matthew T Mason, "Stable pushing: Mechanics, controllability, and planning", *The international journal of robotics research*, vol. 15, no. 6, pp. 533–556, 1996.

[8] Kei Okada, Atsushi Haneda, Hiroyuki Nakai, Masayuki Inaba, and Hirochika Inoue, "Environment manipulation planner for humanoid robots using task graph that generates action sequence", in *2004 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)(IEEE Cat. No. 04CH37566)*. IEEE, 2004, vol. 2, pp. 1174–1179.

[9] Hai-ning Wu, Martin Levihn, and Mike Stilman, "Navigation among movable obstacles in unknown environments", in *2010 IEEE/RSJ International Conference on Intelligent Robots and Systems*. IEEE, 2010, pp. 1433–1438.

[10] Gregory F Lawler, *Introduction to stochastic processes*, Chapman and Hall/CRC, 2018.

[11] Mohak Bhardwaj, Balakumar Sundaralingam, Arsalan Mousavian, Nathan D Ratliff, Dieter Fox, Fabio Ramos, and Byron Boots, "Storm: An integrated framework for fast joint-space model-predictive control for reactive manipulation", in *Conference on Robot Learning*. PMLR, 2022, pp. 750–759.

[12] Vasileios Vasilopoulos, Suveer Garg, Pedro Piacenza, Jinwook Huh, and Volkan Isler, "Ramp: Hierarchical reactive motion planning for manipulation tasks using implicit signed distance functions", in *2023 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. IEEE, 2023, pp. 10551–10558.

[13] Muhammad Kazim, JunGee Hong, Min-Gyeom Kim, and Kwang-Ki K Kim, "Recent advances in path integral control for trajectory optimization: An overview in theoretical and algorithmic perspectives", *Annual Reviews in Control*, vol. 57, pp. 100931, 2024.

[14] Yue Qu, Hongqing Chu, Shuhua Gao, Jun Guan, Haoqi Yan, Liming Xiao, Shengbo Eben Li, and Jingliang Duan, "Rl-driven mppi: Accelerating online control laws calculation with offline policy", *IEEE Transactions on Intelligent Vehicles*, 2023.

[15] Ashish Kumar, Zipeng Fu, Deepak Pathak, and Jitendra Malik, "Rma: Rapid motor adaptation for legged robots", *arXiv preprint arXiv:2107.04034*, 2021.