Reinforcement Learning-based Adaptive Probabilistic Movement Primitives in Hybrid Scenarios (Extended Abstract)

Shibao Yang, Pengcheng Liu, and Nick Pears

University of York, York, UK

1 Introduction

Research in the field of robotic manipulation has traditionally focused on either static or dynamic environments, largely overlooking the intricacies associated with hybrid scenarios where both immovable and movable objects coexist. These hybrid environments present unique challenges, such as distinguishing between static and movable objects, and dealing with the unpredictable shifts in these settings, as well as the activities of other agents [2]. Furthermore, enabling a robotic arm to adapt to the myriad situations that may emerge within these hybrid scenarios presents considerable difficulties. Such scenarios could encompass a range of combinations, including interactions with both static and moveable objects, and a blend of both within a single scene.

We propose an adaptive probabilistic motion primitive (ProMP) [4] based approach to solve the hybrid scene problem, where we combine non-prehensile actions [3] and vision techniques [5] to enhance the learning of the properties of the hybrid scene, enabling the robot arm to accurately identify object states and generate appropriate trajectories for the manipulation task. The main components of our approach are an exploration module and an adaptive ProMPs framework for trajectory generation.

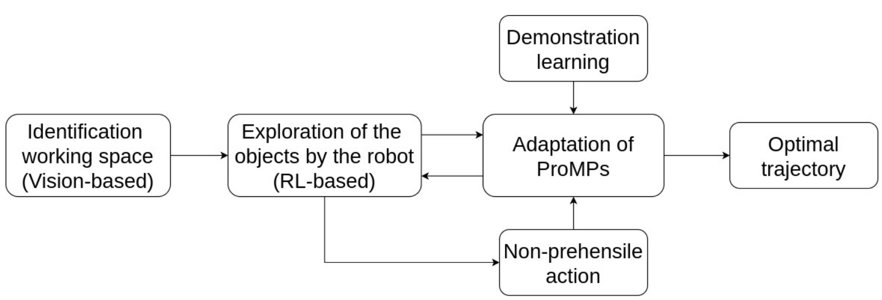

Fig. 1: RL-based Adaptive ProMPs.

The primary objectives of the research are to:

– Develop an exploration module to efficiently explore and understand hybrid scenarios. This should accurately perceive the environment and build a representation that captures the location, properties, and types of objects (static and movable).
– Extend the existing ProMPs framework to autonomously adapt to the problems of motion planning in hybrid scenarios, such as moveable obstacles interfering with predicted motion trajectories.
– Evaluate the performance of the proposed pipeline through simulations and real-world experiments, focusing on accuracy, robustness, and adaptability.

1.1 Method

The Fig. 1 shows the structure of the RL-based adaptive ProMPs.

– **Vision Technique**: A fully convolutional network (FCN) [7] can be used to facilitate a more nuanced segmentation process. This implementation permits the robotic system to achieve a more comprehensive understanding of the spatial relationships between different objects. When integrated with the Robot Operating System (ROS) within the Gazebo simulator, the resulting system is able to perform an accurate object segmentation within the environment.
– **Non-prehensile Actions**: In the exploration module, we utilize reinforcement learning (RL) algorithms, specifically Proximal Policy Optimization (PPO) [6], to facilitate the discovery of feasible trajectories within hybrid-scenario environments by the Franka robot arm. The PPO algorithm has been chosen due to its recognized sample efficiency, stability, and reliability - attributes that render it suitable for the complexities of hybrid environments. Furthermore, PPO's robustness to hyperparameter choices, ease of implementation, compatibility with prevalent RL libraries, as well as its scalability for large neural networks and parallelized training, make it well-suited for tasks involving high-dimensional state and action spaces that characterize hybrid scenarios.
– **Adaptive ProMPs for Trajectory Generation**: We extend the existing Probabilistic Movement Primitives (ProMPs) framework [1] to handle hybrid scenarios. To do this, we incorporate the property of the scenario into the ProMPs model. We propose adding virtual via-points around obstacles, calculated as the closest points on the obstacle's surface moved away in the direction of the normal vector. Conditioning the ProMP distribution on these via-points allows the generated trajectory to circumvent different types of obstacles, thereby optimizing the trajectory.

In this investigation, we devise a hybrid-scenario experimental framework, encompassing four types of items: static boxes, spheres, cylinders, and movable boxes. These were arranged into three distinct experimental settings. The initial setting solely consisted of static items such as boxes, spheres, and cylinders. The subsequent setting incorporated two movable boxes, while the final setting integrated a diverse assortment of all item types.

The experimental process is initiated with the robotic arm identifying and negotiating obstacles within its operational pathway. Utilizing a combination of reinforcement learning and sensory inputs, including vision, the arm interacted with its environment, thereby enabling the detection and subsequent classification of obstacles into stationary or movable categories.

The FCN can be used to process an image of the scene as input and deduce the 3D location of each object at that particular pixel in the scene. For clarity in differentiation, we can mark the obstacles in the scene, for example boxes are in blue, cylinders in red and spheres in yellow.

The research also incorporated learning from demonstration (LfD). We provided demonstrations to train Probabilistic Movement Primitives (ProMPs). Subsequently, these trained ProMPs were updated or "conditioned" to reflect the current state of the environment. This encompassed the positions and properties of any remaining stationary or movable objects, along with any task-specific constraints or goals.

In hybrid scenarios, we designate via-points around static and movable obstacles at known locations to condition the ProMPs for adaptation. This conditioning ensured the generation of a trajectory that was finely attuned to the current environment, aiding in the prevention of collisions with any residual objects.

Notably, when a movable obstacle appeared in the predicted trajectory, the robot arm executed a non-prehensile action, a push action, to displace the obstacle. This displacement allowed the arm to acquire a new trajectory, thereby maintaining the fluidity of the process.

Finally, an optimal trajectory is sampled from the conditioned ProMPs. This trajectory signifies a smooth and efficient movement of the robot arm that accomplishes the intended task, taking into account the current environment and any remaining stationary or movable objects.

1.2 Evaluation

We will evaluate the proposed pipeline through a series of simulation-based and real-world experiments. The performance metrics will include:

- **Accuracy**: The ability of the pipeline to generate appropriate trajectories based on the detected object states.
- **Efficiency**: The time required for the adaptive ProMPs to generate and execute trajectories.
- **Adaptability**: The capability of the pipeline to handle changes in the environment and object states, demonstrating its robustness in dynamic situations.

Accuracy verification starts with the success rate of the motion planning, the ability to successfully avoid dynamic objects and the ability to ensure safe movement of the robot arm.

We record the time taken by the adaptive ProMPs to generate a trajectory after conditioning on the current environment. We perform this measurement across multiple trials and compute the average trajectory generation time. This will help assess the computational efficiency of the ProMPs. In addition to evaluating efficiency in terms of time, consider other performance metrics such as trajectory smoothness, task success rate, and safety (e.g., collision avoidance). This will provide a more comprehensive view of the adaptive ProMPs' performance in the hybrid-scenario environment.

We will evaluate the ability of ProMPs to adapt to external perturbations, such as sudden changes in the environment or the presence of obstacles. Design experiments where the learned ProMPs need to adapt in real-time to these perturbations and measure their performance in terms of task success rate, response time, and other relevant metrics. Compare the results with other methods to demonstrate the robustness of ProMPs in adapting to perturbations.

2 Conclusion

This research aims to extend the ProMPs framework for robotic manipulation in hybrid scenarios with static and movable objects. By developing a reinforcement learning-based pipeline that includes exploration, object state detection, and adaptive trajectory generation, we seek to improve the accuracy, efficiency, and adaptability of robot arm motion planning and manipulation. This method has the potential to significantly enhance the performance of robotic systems in real-world applications, where environments are often dynamic and unpredictable.

References

1. Gomez-Gonzalez, S., Neumann, G., Schölkopf, B., Peters, J.: Adaptation and robust learning of probabilistic movement primitives. IEEE Transactions on Robotics **36**(2), 366–379 (2020)
2. Mohanan, M., Salgoankar, A.: A survey of robotic motion planning in dynamic environments. Robotics and Autonomous Systems **100**, 171–185 (2018)
3. Murali, P.K., Dutta, A., Gentner, M., Burdet, E., Dahiya, R., Kaboli, M.: Active visuo-tactile interactive robotic perception for accurate object pose estimation in dense clutter. IEEE Robotics and Automation Letters **7**(2), 4686–4693 (2022)
4. Paraschos, A., Daniel, C., Peters, J., Neumann, G.: Using probabilistic movement primitives in robotics. Autonomous Robots **42**(3), 529–551 (2018)
5. Riansyah, M.I., Nugraha, Y.P., Ridlwan, H.M., Trilaksono, B.R.: 3d mapping hexacopter simulation using gazebo and robot operating sytem (ros). In: Proceedings of the 9th International Conference on Machine Learning and Computing. pp. 507–510 (2017)
6. Schulman, J., Wolski, F., Dhariwal, P., Radford, A., Klimov, O.: Proximal policy optimization algorithms. arXiv preprint arXiv:1707.06347 (2017)
7. Zeng, A., Song, S., Welker, S., Lee, J., Rodriguez, A., Funkhouser, T.: Learning synergies between pushing and grasping with self-supervised deep reinforcement learning. In: 2018 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS). pp. 4238–4245. IEEE (2018)