CS 6301 Introduction to Robot Manipulation and Navigation Project Proposal Description

Professor Yu Xiang

September 9, 2024

1 Introduction

For the robotics course project, students can choose a topic related to robotics and explore the topic in one of the three different ways:

- **Research-oriented.** In this direction, students are going to propose a new idea that has not been explored before in the literature, then implement the new idea and conduct experiments to verify it.
- **Application-oriented.** In this direction, students can apply an existing robotics algorithm or method to a new problem or a new application. For example, if a method is proposed for domain A, the project can explore applying the method to domain B where different data are collected.
- **Implementation-oriented.** In this direction, students can select an existing robotics algorithm or method, and then implement it and perform experiments to verify the implementation. Since most robotics methods are open source these days, for implementation-oriented projects, students cannot just use an open-source code and run experiments with it.

For project evaluation, all three categories will be considered equally. A project will be evaluated according to its quality in terms of implementation, experiments, presentation, and writing, regardless of its category. However, students are encouraged to consider research-oriented projects and application-oriented projects. Even if the novelty introduced is incremental, it is still exploring new things researchers have not tried before or applying an approach to new applications. Moreover, if you use the Fetch mobile manipulator for experiments, it is possible to try your methods on the real robot in my lab.

2 Proposal Format

The project proposal should be prepared using the ICRA double column latex format. A useful online LaTex tool is Overleaf `https://www.overleaf.com/`. We have the ICRA latex template accessible here through the overleaf: `https://www.overleaf.com/read/rwmhwnwjkrmc`. You can download a copy of the template or make a copy on the overleaf for your own project, and then edit it. Please make sure that you use this latex template for your project. Otherwise, some point will be deducted.

The project proposal should be 1-page PDF using the latex template with the following items:

- **Title**. Let us give a name to your project.
- **Team Members**. List the names of the team members as authors in the proposal. We expect you to work in groups of 2-3 students for course projects.
- **Problem Statement**. Describe the problem you are trying to solve in this project.
- **Approach**. Describe your idea to solve the problem. It is fine if some details have not been determined in the project proposal. But students should have rough ideas of how to proceed.

  First, explicitly state which category the project is in: research-oriented, application-oriented, or implementation-oriented. Second, for research-oriented projects, describe the proposed idea and the novelty of the idea. For application-oriented projects, describe which approach is going to be used and how to apply this approach to a new application. For implementation-oriented projects, describe which approach is to be implemented and the plan for implementation.
- **Simulation Environments**. Describe the simulation environment that will be used in the project for robot experiments.
- **Evaluation**. Describe how to evaluate the success of the project. For example, what evaluation metrics will be used to evaluate the performance of the method?
- **References**. Cite related works in the proposal.

There are two mandatory requirements for the course project:

- The project needs to have a robot.
- The project needs to have robot manipulation.

3 Suggested Topics

Based on the materials we cover in the lectures, we suggest the following topics for the course project. However, the scope of the project is not limited to the topics mentioned below. Students can explore other topics in robotics as well. Also, the references in the suggested topics are recent representative works. Students can explore methods beyond these references and propose new ideas for different topics.

- **Model-based Grasping**. Using 3D models of objects, we can first estimate the 6D object pose, i.e., 3D rotation and the 3D translation of objects [35, 27, 29, 5]. Then model-based grasp planning [20, 6] and motion planning (e.g., MoveIt [2]) approaches can be used for robot grasping.

- **Model-free Grasping Planning**. If we cannot get 3D models of objects but we can obtain images or 3D point clouds of objects using object segmentation methods [36, 18, 14], grasps can be planned using images or point clouds [19, 4, 21, 26].
- **Reinforcement Learning for Grasping**. Reinforcement learning approaches have been widely studied for robot grasping. Most approaches focus on the use of RL for top-down grasping [22, 12, 10, 28, 41]. A few works apply RL to 6D grasping [24, 31, 30].
- **Manipulation of Articulated or Deformable Objects**. In addition to manipulation of rigid objects, active research interests focus on manipulation of articulated objects [13, 34, 17, 37] and deformable objects [40, 23, 15, 16].
- **Mobile Manipulation.** In this case, a robot needs to navigate and manipulate objects [11, 25, 7, 9, 33]. Mobile manipulation is suitable for robots working in large environments.
- **Language-guided Manipulation.** Humans can use natural language to instruct robots to perform manipulation tasks [1, 8, 32].
- **Human-Robot Handover.** Robots and humans can work together to perform tasks. Human-robot handover is a simple example in which humans hand objects to robots or vice versa [38, 31, 39, 3].

4 Robot Simulator Resources

Here are a few robot simulators and simulation environments that you can use for the course project.

- Gazebo `https://gazebosim.org/home`. Gazebo is integrated with ROS. If your code runs in Gazebo, it can be easily transferred to a real robot.
- PyBullet `https://pybullet.org/wordpress/`. PyBullet is an easy-to-use simulator with Python interfaces.
- NVIDIA Isaac Gym `https://developer.nvidia.com/isaac-gym`. Isaac Gym can use GPU acceleration and parallel runs of thousands of environments. It is useful for RL.
- iGibson `https://svl.stanford.edu/igibson/`. iGibson is a simulation environment based on PyBullet. It can be used for robot manipulation and navigation.
- AI2-THOR `https://ai2thor.allenai.org/`, a simulation environment for navigation.
- Habitat `https://aihabitat.org/`, a simulation environment for embodied AI.
- SAPIEN `https://sapien.ucsd.edu/`. SAPIEN contains asserts of articulated objects. It can be useful for manipulating articulated objects.
- BulletArm `https://github.com/ColinKohler/BulletArm`, a PyBullet-based simulation environment for benchmarking of several robot manipualtion tasks.

5 Deep Learning Resources

The most recent vision methods leverage deep learning to train neural networks to tackle various problems in robotics. If your project requires deep neural network training, you may need GPUs for training. **Google Colab** is a great free resource for small amounts of GPU resources: `https://colab.research.google.com/`. Two widely-used deep learning frameworks:

- PyTorch `https://pytorch.org/`
- TensorFlow `https://www.tensorflow.org/`

6 Resources from IRVL at UT Dallas

Our lab, Intelligent Robotics and Vision Lab (IRVL) at UT Dallas, provides several software that can be useful for your projects:

- The SceneReplica benchmark for robot grasping `https://irvlutd.github.io/SceneReplica/`. This project contains links to several object perception, grasp planning, and motion planning software.
- Grasping trajectory optimization `https://github.com/IRVLUTD/GraspTrajOpt`
- Isaac Sim for grasping `https://irvlutd.github.io/MultiGripperGrasp/`
- GraspIt! `https://github.com/IRVLUTD/neuralgrasps-dataset-generation/tree/multigrippergrasp`
- Reinforcement learning for 6D grasping `https://sites.google.com/view/gaddpg`

You can find more in `https://labs.utdallas.edu/irvl/resources/`.

References

[1] Anthony Brohan, Yevgen Chebotar, Chelsea Finn, Karol Hausman, Alexander Herzog, Daniel Ho, Julian Ibarz, Alex Irpan, Eric Jang, Ryan Julian, et al. Do as i can, not as i say: Grounding language in robotic affordances. In *Conference on Robot Learning*, pages 287–318. PMLR, 2023.

[2] Sachin Chitta, Ioan Sucan, and Steve Cousins. Moveit![ros topics]. *IEEE Robotics & Automation Magazine*, 19(1):18–19, 2012.

[3] Sammy Christen, Wei Yang, Claudia Pérez-D'Arpino, Otmar Hilliges, Dieter Fox, and Yu-Wei Chao. Learning human-to-robot handovers from point clouds. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 9654–9664, 2023.

[4] Fu-Jen Chu, Ruinian Xu, and Patricio A Vela. Real-world multiobject, multigrasp detection. *IEEE Robotics and Automation Letters*, 3(4):3355–3362, 2018.

[5] Xinke Deng, Arsalan Mousavian, Yu Xiang, Fei Xia, Timothy Bretl, and Dieter Fox. Poserbpf: A rao–blackwellized particle filter for 6-d object pose tracking. *IEEE Transactions on Robotics*, 37(5):1328–1342, 2021.

[6] Clemens Eppner, Arsalan Mousavian, and Dieter Fox. Acronym: A large-scale grasp dataset based on simulation. In *2021 IEEE International Conference on Robotics and Automation (ICRA)*, pages 6222–6227. IEEE, 2021.

[7] Stephen Hart, Ana Huamán Quispe, Michael W Lanighan, and Seth Gee. Generalized affordance templates for mobile manipulation. In *2022 International Conference on Robotics and Automation (ICRA)*, pages 6240–6246. IEEE, 2022.

[8] Siyuan Huang, Zhengkai Jiang, Hao Dong, Yu Qiao, Peng Gao, and Hongsheng Li. Instruct2act: Mapping multi-modality instructions to robotic actions with large language model. *arXiv preprint arXiv:2305.11176*, 2023.

[9] Xiaoyu Huang, Dhruv Batra, Akshara Rai, and Andrew Szot. Skill transformer: A monolithic policy for mobile manipulation. *arXiv preprint arXiv:2308.09873*, 2023.

[10] Stephen James, Paul Wohlhart, Mrinal Kalakrishnan, Dmitry Kalashnikov, Alex Irpan, Julian Ibarz, Sergey Levine, Raia Hadsell, and Konstantinos Bousmalis. Sim-to-real via sim-to-sim: Data-efficient robotic grasping via randomized-to-canonical adaptation networks. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 12627–12637, 2019.

[11] Snehal Jauhri, Jan Peters, and Georgia Chalvatzaki. Robot learning of mobile manipulation with reachability behavior priors. *IEEE Robotics and Automation Letters*, 7(3):8399–8406, 2022.

[12] Dmitry Kalashnikov, Alex Irpan, Peter Pastor, Julian Ibarz, Alexander Herzog, Eric Jang, Deirdre Quillen, Ethan Holly, Mrinal Kalakrishnan, Vincent Vanhoucke, et al. Scalable deep reinforcement learning for vision-based robotic manipulation. In *Conference on Robot Learning*, pages 651–673. PMLR, 2018.

[13] Dov Katz and Oliver Brock. Manipulating articulated objects with interactive perception. In *2008 IEEE International Conference on Robotics and Automation*, pages 272–277. IEEE, 2008.

[14] Alexander Kirillov, Eric Mintun, Nikhila Ravi, Hanzi Mao, Chloe Rolland, Laura Gustafson, Tete Xiao, Spencer Whitehead, Alexander C Berg, Wan-Yen Lo, et al. Segment anything. *arXiv preprint arXiv:2304.02643*, 2023.

[15] Rita Laezza, Robert Gieselmann, Florian T Pokorny, and Yiannis Karayiannidis. Reform: A robot learning sandbox for deformable linear object manipulation. In *2021 IEEE International Conference on Robotics and Automation (ICRA)*, pages 4717–4723. IEEE, 2021.

[16] Xingyu Lin, Zhiao Huang, Yunzhu Li, Joshua B Tenenbaum, David Held, and Chuang Gan. Diffskill: Skill abstraction from differentiable physics for deformable object manipulations with tools. *arXiv preprint arXiv:2203.17275*, 2022.

[17] Liu Liu, Wenqiang Xu, Haoyuan Fu, Sucheng Qian, Qiaojun Yu, Yang Han, and Cewu Lu. Akb-48: A real-world articulated object knowledge base. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 14809–14818, 2022.

[18] Yangxiao Lu, Yuqiao Chen, Nicholas Ruozzi, and Yu Xiang. Mean shift mask transformer for unseen object instance segmentation. *arXiv preprint arXiv:2211.11679*, 2022.

[19] Jeffrey Mahler, Jacky Liang, Sherdil Niyaz, Michael Laskey, Richard Doan, Xinyu Liu, Juan Aparicio Ojea, and Ken Goldberg. Dex-net 2.0: Deep learning to plan robust grasps with synthetic point clouds and analytic grasp metrics. *arXiv preprint arXiv:1703.09312*, 2017.

[20] Andrew T Miller and Peter K Allen. Graspit! a versatile simulator for robotic grasping. *IEEE Robotics & Automation Magazine*, 11(4):110–122, 2004.

[21] Arsalan Mousavian, Clemens Eppner, and Dieter Fox. 6-dof graspnet: Variational grasp generation for object manipulation. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*, pages 2901–2910, 2019.

[22] Deirdre Quillen, Eric Jang, Ofir Nachum, Chelsea Finn, Julian Ibarz, and Sergey Levine. Deep reinforcement learning for vision-based robotic grasping: A simulated comparative evaluation of off-policy methods. In *2018 IEEE International Conference on Robotics and Automation (ICRA)*, pages 6284–6291. IEEE, 2018.

[23] Daniel Seita, Pete Florence, Jonathan Tompson, Erwin Coumans, Vikas Sindhwani, Ken Goldberg, and Andy Zeng. Learning to rearrange deformable cables, fabrics, and bags with goal-conditioned transporter networks. In *2021 IEEE International Conference on Robotics and Automation (ICRA)*, pages 4568–4575. IEEE, 2021.

[24] Shuran Song, Andy Zeng, Johnny Lee, and Thomas Funkhouser. Grasping in the wild: Learning 6dof closed-loop grasping from low-cost demonstrations. *IEEE Robotics and Automation Letters*, 5(3):4978–4985, 2020.

[25] Charles Sun, Jdrzej Orbik, Coline Manon Devin, Brian H Yang, Abhishek Gupta, Glen Berseth, and Sergey Levine. Fully autonomous real-world reinforcement learning with applications to mobile manipulation. In *Conference on Robot Learning*, pages 308–319. PMLR, 2022.

[26] Martin Sundermeyer, Arsalan Mousavian, Rudolph Triebel, and Dieter Fox. Contact-graspnet: Efficient 6-dof grasp generation in cluttered scenes. In *2021 IEEE International Conference on Robotics and Automation (ICRA)*, pages 13438–13444. IEEE, 2021.

[27] Jonathan Tremblay, Thang To, Balakumar Sundaralingam, Yu Xiang, Dieter Fox, and Stan Birchfield. Deep object pose estimation for semantic robotic grasping of household objects. *arXiv preprint arXiv:1809.10790*, 2018.

[28] Dian Wang, Robin Walters, Xupeng Zhu, and Robert Platt. Equivariant q learning in spatial action spaces. In *Conference on Robot Learning*, pages 1713–1723. PMLR, 2022.

[29] Gu Wang, Fabian Manhardt, Federico Tombari, and Xiangyang Ji. GDR-Net: Geometry-guided direct regression network for monocular 6d object pose estimation. In *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 16611–16621, June 2021.

[30] Lirui Wang, Xiangyun Meng, Yu Xiang, and Dieter Fox. Hierarchical policies for cluttered-scene grasping with latent plans. *IEEE Robotics and Automation Letters*, 7(2):2883–2890, 2022.

[31] Lirui Wang, Yu Xiang, Wei Yang, Arsalan Mousavian, and Dieter Fox. Goal-auxiliary actor-critic for 6d robotic grasping with point clouds. In *Conference on Robot Learning*, pages 70–80. PMLR, 2022.

[32] Tianyu Wang, Yifan Li, Haitao Lin, Xiangyang Xue, and Yanwei Fu. Wall-e: Embodied robotic waiter load lifting with large language model. *arXiv preprint arXiv:2308.15962*, 2023.

[33] Jimmy Wu, Rika Antonova, Adam Kan, Marion Lepert, Andy Zeng, Shuran Song, Jeannette Bohg, Szymon Rusinkiewicz, and Thomas Funkhouser. Tidybot: Personalized robot assistance with large language models. *arXiv preprint arXiv:2305.05658*, 2023.

[34] Ruihai Wu, Yan Zhao, Kaichun Mo, Zizheng Guo, Yian Wang, Tianhao Wu, Qingnan Fan, Xuelin Chen, Leonidas Guibas, and Hao Dong. Vat-mart: Learning visual action trajectory proposals for manipulating 3d articulated objects. *arXiv preprint arXiv:2106.14440*, 2021.

[35] Yu Xiang, Tanner Schmidt, Venkatraman Narayanan, and Dieter Fox. Posecnn: A convolutional neural network for 6d object pose estimation in cluttered scenes. *arXiv preprint arXiv:1711.00199*, 2017.

[36] Yu Xiang, Christopher Xie, Arsalan Mousavian, and Dieter Fox. Learning rgb-d feature embeddings for unseen object instance segmentation. *arXiv preprint arXiv:2007.15157*, 2020.

[37] Zhenjia Xu, Zhanpeng He, and Shuran Song. Universal manipulation policy network for articulated objects. *IEEE Robotics and Automation Letters*, 7(2):2447–2454, 2022.

[38] Wei Yang, Chris Paxton, Arsalan Mousavian, Yu-Wei Chao, Maya Cakmak, and Dieter Fox. Reactive human-to-robot handovers of arbitrary objects. In *2021 IEEE International Conference on Robotics and Automation (ICRA)*, pages 3118–3124. IEEE, 2021.

[39] Wei Yang, Balakumar Sundaralingam, Chris Paxton, Iretiayo Akinola, Yu-Wei Chao, Maya Cakmak, and Dieter Fox. Model predictive control for fluid human-to-robot handovers. In *2022 International Conference on Robotics and Automation (ICRA)*, pages 6956–6962. IEEE, 2022.

[40] Jihong Zhu, Andrea Cherubini, Claire Dune, David Navarro-Alarcon, Farshid Alambeigi, Dmitry Berenson, Fanny Ficuciello, Kensuke Harada, Xiang Li, Jia Pan, et al. Challenges and outlook in robotic manipulation of deformable objects. *arXiv preprint arXiv:2105.01767*, 2021.

[41] Xupeng Zhu, Dian Wang, Ondrej Biza, Guanang Su, Robin Walters, and Robert Platt. Sample efficient grasp learning using equivariant models. *arXiv preprint arXiv:2202.09468*, 2022.