Life-science Fetch Robot for Dexterous Manipulation of Centrifuge Tubes

Jiahui Zhu*, and Yizheng Zhang*

Abstract

In the field of robotics, service robots are increasingly becoming an industry hot-spot. Especially for mobile manipulation, many researchers have studied its applications and some teams are focused on the manipulator's pose estimation of grasping objects, and the other teams are working on the simultaneous localization and mapping, path planning and automatic obstacle avoidance of mobile vehicles. Very few of them explore the integrated system of mobile manipulation and only transport simple, regular objects. Since in life science experiments, many repetitive, simple tasks can be replaced by autonomous robots. This work is aiming to apply a mobile manipulation Fetch robot to execute the dexterous manipulation of centrifuge tubes in life science laboratories. Usually, those tubes have a screw-lid so that we need to unscrew and re-screw the lid. In this work: (a) to get an accurate centrifuge tube localization; (b) to grasp the tube from the side and place it in the designated position; (c) to unscrew the lid from the top of the centrifuge tube and place it on the test table; (d) to re-position the lid, grasp and re-screw it. Additionally, we will use AprilTag2 and DenseFusion two pose estimation methods to grasp the centrifuge tube and the lid.

I. Introduction

Thus life sciences is a research area that needs to do a lot of repetitive experiments. Usually, the only difference between experiments is the difference in parameters due to that there is a countless number of proteins, molecules and reagents that need to be explored by testing with them in real life science experiments. This is a work that requires manual labor that is quite often very repetitive and tiresome. Students in Universities or lab technicians in research institutes spend a lot of time on these experiments. Sometimes because of too tired, they make mistakes, and have to do it again. On the other hand, this kind of repetitive, simple experiment is perfect for robots to do it.

There is further literature regarding research on mobile manipulation for life science. [1] report on a robot called H20 that is concentrating on the hardware, control and planning aspects. Schmitt et al. [2] take robot dynamics and time-variant environments into consider and propose a new model for sequential manipulation tasks. It can interact with the object in the environment like human to avoid collision with them, which is pretty useful for assistant robot and life science robot. [3]describes a reinforcement learning (RL) strategy for manipulation and grasping of a mobile manipulator that reduces the complexity of the visual feedback and handle varying manipulation dynamics and uncertain external perturbations. [4] propose Neural Task Graph (NTG) Networks, which use conjugate task graph as the intermediate representation to modularize both the video demonstration and the derived policy. This network NTG improves data efficiency with visual input as well as achieve strong generalization without the need for dense hierarchical supervision

In the work of [5], they use self-supervision to learn a compact and multimodal representation of sensory inputs like force sensor, which can then be used to improve the sample efficiency of our policy learning. Srinivasa et al. [6] present practical techniques that improve performance in areas like the house by considering the complete system in the context of this specific domain. In an industrial scenario

*These authors have contributed equally to this work.

Stoyanov et al. [7] build a nearly market-ready robot to unload coffee sacks out of seafright containers in RobLog Project. They using feature-based object recognition, ROS and MoveIt! to solve those tasks successfully.

Fig. 1: A mobile manipulation Fetch robot in life science laboratories.

II. State of the Art

A. JIAHUI

1) Papers: In the last few years, there are a few works done on mobile manipulation as an integrated system and several mobile manipulators have been coming to the market, for example, Tiago by Pal Robotics, Mobile Baxter by Rethink Robotics, and Fetch and Freight by Fetch Robotics [8].

For labware transportation in life science laboratories, a mobile robot called H20 with dual arms capable of 6-Degree of Freedom (DOF) mobility, an indoor GPS navigation system and Kinect sensor developed by H. Liu. et al. [9], [10] can realize manipulation tasks for different experimental objects such as flat panels and labware containers. However, different from general wheeled mobile manipulation, [11] proposes a lab automation drone and a robotic limb is attached to the robotic rotor-craft which could manipulate and transfer lines rapidly between test work-spaces.

2) One further paper relevant to the project: In the paper [12], they describe a method for robustly detecting visual fiducials and propose a graph-based image segmentation algorithm based on local gradients that allows lines to be precisely estimated. They specify and provide results on a set of benchmarks which will allow better comparisons of fiducial systems in the future. Their system is composed of two major components: the tag detector and the coding system.

- Detection process
  The detector attempts to find four-sided regions that have a darker interior than their exterior. Tags are composed of black and white blocks similar to the QR code and different tags contain different ID information. First, paste some tags in the scene. Secondly, the gradient at every pixel is obtained,

and then compute their magnitudes and direction. Using a graph-based method, pixels with similar gradient directions and magnitude are clustered into components. Using weighted least squares, a line segment is then fit to the pixels in each component. Once four lines have been found, a candidate quad detection is created. Finally, they compute the 3×3 homography matrix that projects 2D points in homogeneous coordinates from the tag's coordinate system to the 2D image coordinate system. The homography is computed using the Direct Linear Transform (DLT) algorithm.

- Payload decoding
  The final task is to read the bits from the payload field. They do this by computing the tag-relative coordinates of each bit field, transforming them into image coordinates using the homography, and then thresholding the resulting pixels. In order to be robust to lighting (which can vary not only from tag to tag, but also within a tag), they use a spatially-varying threshold.

3) ROS package: AprilTag is a free and open source visual fiducial system developed by the APRIL Robotics Laboratory of the University of Michigan, useful for a wide variety of tasks including augmented reality, robotics, and camera calibration. Targets can be created from an ordinary printer, and the AprilTag detection software computes the precise 3D position, orientation, and identity of the tags relative to the camera. The fiducial design and coding system are based on a near-optimal lexicographic coding system, and the detection software is robust to lighting conditions and view angles. It uses a 2D bar code style "tag" similar to the QR code, allowing full 6-DOF localization of features from a single image.

In this work, we used an open source ROS package apriltag_ros for objects' pose estimation. The apriltag_ros package is a ROS wrapper of the AprilTag 3 visual fiducial detection algorithm. Provides full access to the core AprilTag 3 algorithm's customizations and makes the tag detection image and detected tags' poses available over ROS topics (including tf). The core AprilTag 3 algorithm is extended to allow the detection of tag bundles and a bundle calibration script is provided (bundle detection is more accurate than single tag detection). Continuous (camera image stream) and single image detector nodes are available. The package works as shown in Fig. 2.

Fig. 2: Overview of the apriltag_ros package.

B. YIZHENG

1) Papers: Schmitt et al. [2] propose a new model for sequential manipulation tasks that also considers robot dynamics and time-variant environments. It can automatically derive constraint-based controllers and use them as steering functions in a kinodynamic manipulation planner using this model. Thakar et al. [13] propose an active learning based approach that can pick-up parts while the mobile base and the gripper are moving to minimize the operation time.

Wang et al. [14] propose a generic framework for estimating 6D pose of a set of known objects from RGB-D images and deploy the proposed method to a real robot to grasp and manipulate objects based on the estimated pose.

2) One further paper relevant to the project: In this work, the author present DenseFusion, a generic framework for estimating 6D pose of a set of known objects from RGB- D images. DenseFusion is a heterogeneous architecture that processes the two data sources individually and uses a novel dense fusion network to extract pixel-wise dense feature embedding, from which the pose is estimated. Furthermore, they integrate an end-to-end iterative pose refinement procedure that further improves the pose estimation while achieving near real-time inference. The experiments show that this method outperforms state-of-the-art approaches in two datasets, YCB-Video and LineMOD. The method's architecture contains two main stages. The first stage takes the color image as input and performs semantic segmentation for each known object category. Then, for each segmented object, feed the masked depth pixels (converted to 3D point cloud) as well as an image patch cropped by the bounding box of the mask to the second stage. The object pose is obtained by the second stage processes. It comprises four components:

- a fully convolutional network that processes.
- a PointNet-based network that processes each point in the masked 3D point cloud to a geometric feature embedding.
- a pixel-wise fusion network that combines both embeddings and outputs the estimation of the 6D pose of the object based on an unsupervised confidence scoring.
- an iterative self-refinement methodology to train the network in a curriculum learning manner and refine the estimation result iteratively.

We will try to use this algorithm to estimate the centrifuge tube's pose and then grasp it.

3) ROS package: The move_base package provides an implementation of an action that, given a goal in the world, will attempt to reach it with a mobile base. The move_base node links together a global and local planner to accomplish its global navigation task. It supports any global planner adhering to the nav_core::BaseGlobalPlanner interface specified in the nav_core package and any local planner adhering to the nav_core::BaseLocalPlanner interface specified in the nav_core package. The move_base node also maintains two costmaps, one for the global planner, and one for a local planner that are used to accomplish navigation tasks. The pip line of the package is shown in Fig. 3

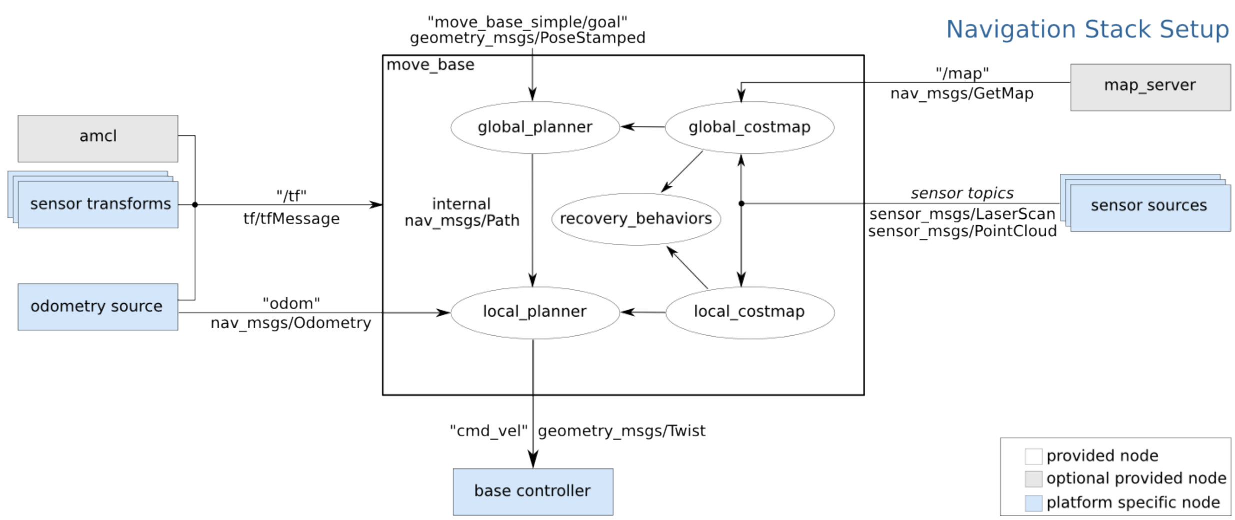

Fig. 3: The framework of the move base node.

III. System Description

The system contains 6 components:

- MoveIt framework that is used to control the manipulator to grasp the object.
- Cartographer algorithm to build a map.
- Move base ros package that is used to navigate the robot to the target place where the object should be grasped.

- AprilTag 3 visual fiducial detection algorithm is used to obtain the object pose.
- We will try to implement DenseFusion framework to estimate the object pose to replace the AprilTag 3 algorithm since this algorithm does not need to add the QR code on the object we want to grasp.
- We will design a device that can clamp for test tubes on the robot.

Fig. 4: Centrifuge tubes with screw lid.

Firstly, we put the AprilTag on the tubes. Once the camera sees the AprilTag, the pose of the tube is obtained and we set the coordinate as the goal of the inverse kinematics. The manipulator reaches directly above the object, then drops vertically to grab it. After the tube is grasped, the manipulator places the tube on the device mounted on the robot. and then unscrew the lid from the top of the centrifuge tube (Fig. 4)and place it on the test table. Now we can add something into the tube if needed. The manipulator re-position the lid, grasp and re-screw it. Also, we use cartographer, a graph-based SLAM algorithm to build the map of the lab. And then use this map to navigate our robot to the test table through move base ros package.

IV. System Evaluation

This work aims to use the Fetch robot (Fig. 1) for grasping and placing the centrifuge tube where the desired labware is positioned. Due to the centrifuge tube has a lid, we need to use the Fetch robot's parallel grippers to unscrew and re-screw the lid. Thus, we will conduct these experiments and choose four criteria to evaluate the performance.

- Localization accuracy
  In the beginning, the Fetch robot needs to move to the place where it grasps the centrifuge tube. There are two ways to evaluate the localization accuracy. 1. We can use the data from the tracing system as the ground truth and compare it with the data from amcl. 2. We add another QR code on the table, and obtain the transform from table to robot, and use this data compare with the data from amcl.

- Success rate of grasping and placing
  After the robot moves to the designated position, it needs to grasp the tube and place it on its fixed platform. We can do it in 50 attempts and see how many times it successfully grasps and places the tube.
- Alignment error
  Because the centrifugal tube has a lid, and in this experiment, it is necessary to screw the lid and place it on the table, then grasp the lid and re-screw it back on the tube. At this point, the alignment error between the lid and the tube is involved. We can evaluate the performance of the manipulator by measuring the eccentricity distance between the lid and the tube.
- Task time
  In this experiment, we use different IK methods to test the time required for the robot to complete the whole tasks and a shorter time means better performance.

V. Work Plan

We make a plan for this project and we divide the entire experiment into seven stages including algorithm, implementation, testing, and evaluation, as shown in Table I below. In order to show the procedure clearly, we draw the Gantter (Fig. 5) of each stage. The mid-term report will be finished at the end of November and the final demo, the final report and the project website will be performed at the end of December.

TABLE I: The stage of the whole experiment.

Stage	**Move base**	**Object place**	**Fixed tube**	**Unscrew**	**Place the lid**	**Grasp**	**Re-screw**
1	No	Fixed	Manual	Yes	No	No	No
2	No	Fixed	Manual	Yes	No	No	Yes
3	No	Fixed	Automatic	Yes	No	No	Yes
4	No	Random	Automatic	Yes	No	No	Yes
5	No	Random	Automatic	Yes	Yes	Yes	Yes
6	Yes	Random	Automatic	Yes	Yes	Yes	Yes(AprilTag)
7	Yes	Random	Automatic	Yes	Yes	Yes	Yes(DenseFusion)

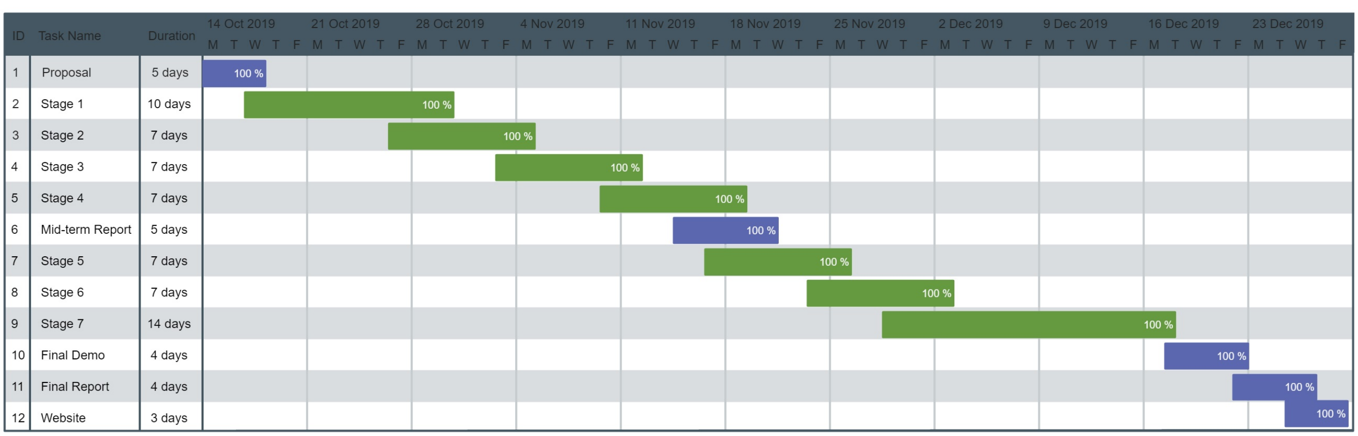

Fig. 5: The Gantter of the work plan.

VI. Conclusions

In this project, we have proposed a method for labware transportation and dexterous manipulation of centrifuge tubes in life science laboratories using a Fetch robot. Also, we adopt two pose estimation methods, AprilTag and DenseFusion for objects recognition and both methods work in the experiment while DenseFusion performs better than AprilTag because of the unrestricted shape of labware containers. In the future, the development of repeatable and simple experiments in life science laboratories, very often plagued with manual transportation, could profit from this.

References

[1] M. M. Ali, H. Liu, N. Stoll, and K. Thurow, "Intelligent arm manipulation system in life science labs using h20 mobile robot and kinect sensor," in *2016 IEEE 8th International Conference on Intelligent Systems (IS)*, pp. 382–387, IEEE, 2016.

[2] P. S. Schmitt, F. Wirnshofer, K. M. Wurm, G. v. Wichert, and W. Burgard, "Modeling and planning manipulation in dynamic environments," in *2019 International Conference on Robotics and Automation (ICRA)*, pp. 176–182, May 2019.

[3] Z. Li, T. Zhao, F. Chen, Y. Hu, C. Su, and T. Fukuda, "Reinforcement learning of manipulation and grasping using dynamical movement primitives for a humanoidlike mobile manipulator," *IEEE/ASME Transactions on Mechatronics*, vol. 23, pp. 121–131, Feb 2018.

[4] D.-A. Huang, S. Nair, D. Xu, Y. Zhu, A. Garg, L. Fei-Fei, S. Savarese, and J. C. Niebles, "Neural task graphs: Generalizing to unseen tasks from a single video demonstration," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 8565–8574, 2019.

[5] M. A. Lee, Y. Zhu, K. Srinivasan, P. Shah, S. Savarese, L. Fei-Fei, A. Garg, and J. Bohg, "Making sense of vision and touch: Self-supervised learning of multimodal representations for contact-rich tasks," *arXiv preprint arXiv:1810.10191*, 2018.

[6] S. S. Srinivasa, A. M. Johnson, G. Lee, M. C. Koval, S. Choudhury, J. E. King, C. M. Dellin, M. Harding, D. T. Butterworth, P. Velagapudi, *et al.*, "A system for multi-step mobile manipulation: Architecture, algorithms, and experiments," in *International Symposium on Experimental Robotics*, pp. 254–265, Springer, 2016.

[7] T. Stoyanov, N. Vaskevicius, C. A. Mueller, T. Fromm, R. Krug, V. Tincani, R. Mojtahedzadeh, S. Kunaschk, R. M. Ernits, D. R. Canelhas, *et al.*, "No more heavy lifting: Robotic solutions to the container unloading problem," *IEEE Robotics & Automation Magazine*, vol. 23, no. 4, pp. 94–106, 2016.

[8] M. A. Roa, D. Berenson, and W. Huang, "Mobile manipulation: Toward smart manufacturing [tc spotlight]," *IEEE Robotics & Automation Magazine*, vol. 22, no. 4, pp. 14–15, 2015.

[9] H. Liu, N. Stoll, S. Junginger, and K. Thurow, "Mobile robot for life science automation," *International Journal of Advanced Robotic Systems*, vol. 10, no. 7, p. 288, 2013.

[10] M. M. Ali, H. Liu, N. Stoll, and K. Thurow, "Kinematic analysis of 6-dof arms for h20 mobile robots and labware manipulation for transportation in life science labs," *Journal of Automation Mobile Robotics and Intelligent Systems*, vol. 10, 2016.

[11] D. Kim and P. Y. Oh, "Lab automation drones for mobile manipulation in high throughput systems," in *2018 IEEE International Conference on Consumer Electronics (ICCE)*, pp. 1–5, IEEE, 2018.

[12] E. Olson, "Apriltag: A robust and flexible visual fiducial system," in *2011 IEEE International Conference on Robotics and Automation*, pp. 3400–3407, IEEE, 2011.

[13] S. Thakar, P. Rajendran, V. Annem, A. Kabir, and S. Gupta, "Accounting for part pose estimation uncertainties during trajectory generation for part pick-up using mobile manipulators," in *2019 International Conference on Robotics and Automation (ICRA)*, pp. 1329–1336, May 2019.

[14] C. Wang, D. Xu, Y. Zhu, R. Martin-Martin, C. Lu, L. Fei-Fei, and S. Savarese, "Densefusion: 6d object pose estimation by iterative dense fusion," in *The IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, June 2019.