Robot Autonomy Final Report

Michael Beck **Akshay Bhagat** **Matthew Lauer**

Che-Yen Lu **Jin Zhu**

Project Overview: Warehouse Storage Robot

The purpose of this project was to develop a robotic system for the task of automated item stowing/storage. The robot was to pick items from a jumbled tote and place those items onto a shelf. The item locations on the shelf were be known after being stowed, with the system creating a listing of the anticipated item locations after it had completed stowing all of the items. This is a legacy project at CMU and is part of the development for one of the primary portions of this year's Amazon Picking Challenge.

1 Amazon Picking Challenge Background: Stowage Portion

Contestants for the challenge have been given a set of 40 known items as well as rules and constraints around building a robotic system that can pick or stow those items to and from a shelf. The shelf will be fabricated by contestants and must contain between 2-10 distinct storage bins, as well as fit certain volumetric constraints. In addition, this year's competition will feature an additional set of items which will be provided on the day of the competition, and competitor's will have less than 30 minutes to train their system to recognize and pick these items in addition to the 40 which have already been provided. The stowage portion of the competition will consist of attempting to store 20 items which will be selected at the time of the competition by Amazon, with 10 items being from the additional set that is provided on competition day. These items will be stored together in a single jumbled tote, and contestants will be given 15 minutes to stow as many of the items as possible. The competition is highly focused on accurate item location reporting, with most points available to be gained or lost being focused around a system's ability to report the final bin locations of the items within the shelf storage system. Michael Beck, Akshay Bhagat, Matthew Lauer, Che-Yen Lu, and Jin Zhu are all members of CMU's team for this year's competition and will be participating at RoboCup in Japan this summer.

Additional information about the competition can be found at the Amazon Robotics web page:

https://www.amazonrobotics.com/#/roboticschallenge

2 Problem Definition

The goal for this project was to complete the foundational work to autonomously stow items from a tote to a shelf in a highly occluded, cluttered environment and generate an item report listing the inventory of the shelf and tote. The scope of this work was determined in line with necessary time frames for participating in this year's Amazon Picking Challenge. To that end a baseline stowage run consisted of stowing items from a tote populated with 20 jumbled items as seen in Figure 1, with 3 of those items being non-suction pickable. These items were to be picked by the robotic arm and placed into one of the 8 system bins as seen in Figure 2. The arm was expected to pick 10 items with suction within a 15 minute time frame, and to accurately report the bin location within the shelf for all items at the end of the run. The presence of any item outside of the stowage tote, an amnesty tote, or the shelf bins at the end of the run was considered a failure case, as this represents hefty penalties at the competition. This project did not include any goals pertaining to the additional item set which Amazon will be providing on the competition day, and only involved the known 40 items which have already been provided (addressing the additional items is a future task for our group which will be completed in early summer).

Figure 1: Sample tote populated with 20 jumbled competition items.

Figure 2: Robotic arm system with 4 drawer shelf. Each shelf separates into 2 bins, for a total of 8 separately identifiable bins.

2.1 Stowage vs. Picking

The competition features two primary components: a timed picking run, where the robot attempts to pick items from bins within the shelf and place them in totes, and a timed stowage run, where the robot attempts to stow items from a tote to the shelf system.The stowage portion of the competition is the primary focus of this report, with the picking portion being beyond this scope of work. Both competition components have different challenging aspects to them. For the picking portion the system has to have a much more dynamic planning range as it is a necessity to reach all bins (although this may be desirable with stowage as well for item isolation purposes to aid vision and location reporting) as well as the ability to quickly identify desired items within the shelf system from a given order list. Item locations are much more structured for the picking run than the stowage run and much less likely to be occluded, with only 4 items expected to be in a given bin at any time. The Stowage challenge on the other hand presents greater challenges for both vision and grasping, as there are 20 items which are all heavily occluded within a single tote to grasp from. This makes accidental grasps and item misreporting much more likely, which are both detrimental to competition performance. The time allotted for the run is also more demanding, with a perfect run being stowage of all 20 items within a 15 minute time window (this is quite fast based off of past performances of successful teams in previous competitions, see the Related Works section below). Another challenging aspect of this process is having the robot work around objects inside the tote which are not graspable, while still managing to grip and stow as many of the other nearby occluded items as possible.

2.2 Challenging Aspects

The challenge of this competition is attempting to pick as many items as possible from a heavily occluded environment within an alloted time. The environment itself presents challenges in terms of planning and lighting, as the stowage tote has small curves and nooks which are hard to plan in and out of when grasping items, and which also cast shadows on objects that can confuse classifiers. Additionally, some items present greater challenges with identification or grasping than others that make them much more difficult to pick, including having uneven and/or deformable surfaces and/or spectral qualities.

3 Related Works

This competition resembles considerable challenges for robotics in the areas of vision, planning, and grasping, all of which are pertaining to a current unsolved problem for industry applications. As such the competition has attracted a diverse set of participants in both the commercial and academic sectors. The following recorded seminars are representative of academic work in this area:

Amazon Picking Challenge 2016 - Team MIT-Princeton - Summary

Lessons from the 1st Amazon Picking Challenge and Rutgers' Participation

Motion Planning for Industrial Robots and Warehouse Automation

Further works can be found as a collection as part of MIT's Workshop on Automation for Warehouse Logistics.

4 Approach

4.1 Foundational Work

This project is in its second year as a legacy project at CMU. The previous year's team laid the foundation for this year's system design, as well as for the rudimentary planning interface and system implementation. This semester features a new robotic arm (a switch from the Universal Robots UR5 to the UR10), a new 1-DOF slider for the robot which allows reach for all bins, a new end effector, new grasping mechanisms (1-DOF suction and electromagnets vs. stationary suction), and all rebuilt vision and grasping algorithms. The planner also underwent a substantial overhaul to accommodate the new arm, slider, and 1-DOF suction gripper. This is the first year Amazon is allowing participants to design and use their own shelf system, which are aspects of the project which warranted large system re-designs. The entire system setup can be seen in Figure 3.

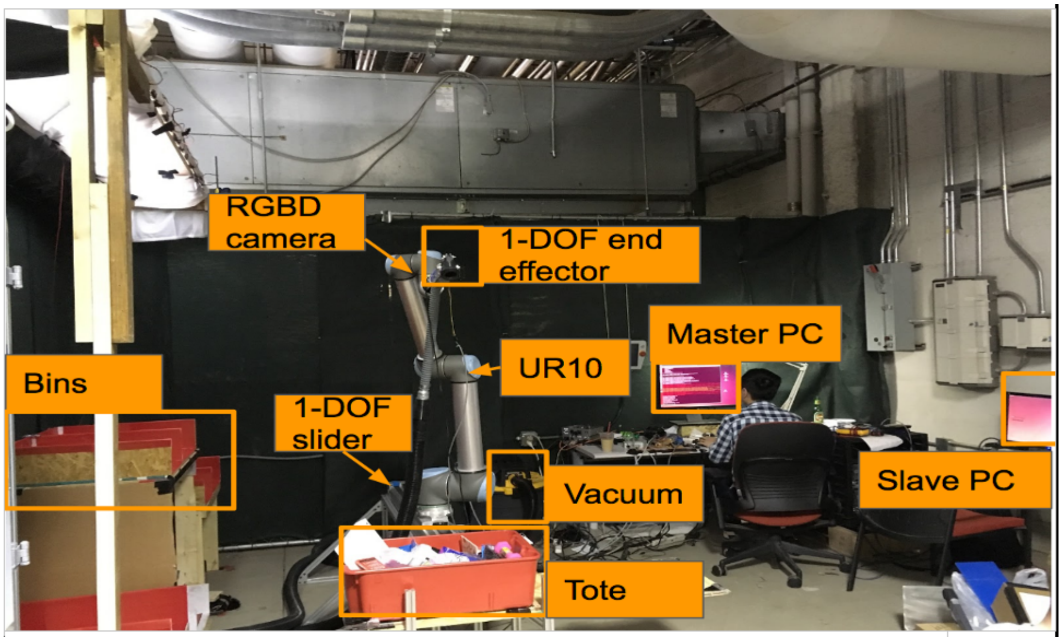

Figure 3: System hardware setup.

4.2 Gripper Choice

The system features a high-flow vacuum attached to a pivoting 1-DOF suction head, which has a range from 0 to 90 degrees. This vacuum is capable of picking 34 out of 40 of the known competition items for this year (although one of the 34 items is very challenging), and has been the main gripper type used in past competitions due to its reliability. The 1-DOF functionality allows for grasping in tight spacing and corners beyond what a stationary gripper can accommodate. There is also an intention to install an electromagnet into the system early this summer which should let pick an additional 5 known ferrous items. The 40th item has currently been blacklisted, meaning it has been deemed to be non-pickable with the current system design. Ideally the system would also incorporate a two-finger gripper mechanism, which would allow picking for all of the challenge items including any additional items provided on the day of the competition. A two-finger gripper has not currently been implemented due to time and budget constraints.

4.3 Software

4.3.1 Planning

For planning, the system uses EGWA (experience graph weighted A*). Motions with the arm involve executing pre-trained plans from a built experience graph in order to have reliable motions into poses for image capturing, item transportation and drop-off, and pre-grasp points. Path constraints have also been built into the planner in order to disallow any motions which would cause the arm to tangle itself with its vacuum hosing or electrical wiring. Planning from pre-grasp points to item grasp points is executed without an experience graph using weighted A*. Grasp points for this project were determined by using a combinational weight between point cloud centroids and point cloud heights. The reasoning was that the higher a point within a point cloud was, the less likely that section of the item was to be occluded by another object. Centroid grasping allows for greater likelihood of a solid grasp on an item without causing the item to have a large cantilever weight on the suction head that could cause a grasp failure.

4.3.2 Vision

The system uses FCN for item identification. FCN proved itself to have a high accuracy in item identification, and provides the benefit of pixel-wise labeling which grants valuable information when dealing with occluded cases. The network was trained using 470 images which contained between

1-40 instances of the known competition items, such as the image in Figure 4. The images were hand-labeled using the free online service LabelMe.

Figure 4: Example training data after being hand-labeled.

4.4 Runtime Logic

Before beginning a run confidence scores for the 3 non-suction pickable items were set to 0 in the state machine, assuring that the arm would not try to pick them from the stowage tote. The system was programmed to first move the arm into a camera pose and identify as many items as possible, and to segment each item identification as a separate point cloud. Those identifications were then checked against the item confidence scores, and point clouds corresponding to items with confidence scores above 0 were queried for their highest point cloud value in the upward direction. This item was then given the highest priority for grasping, based off of the assumption that it was less likely to be occluded due to its height in the stowage tote. The arm then moved to a pre-determined pre-grasp pose, and the grasp pose calculator found a suitable pose based on point cloud centroid locations and height. Once the arm had planned to the new pose the suction system was engaged, and a pressure sensor was checked to ensure a suction seal had been made. After verifying a good seal the item was moved from the stowage tote in a pre-defined motion to the shelf and dropped at one of 10 hard-coded locations within the shelf bins, executed in a sequential order from back to front within each bin. This process was repeated as many times as possible within the 15 minute time frame.

5 Results

5.1 Vision

The FCN net operated very well, identifying between 56% to 96% of pixels for all items, even in heavily occluded environments. Results for the network in a heavily occluded case can be see in Figure 5.

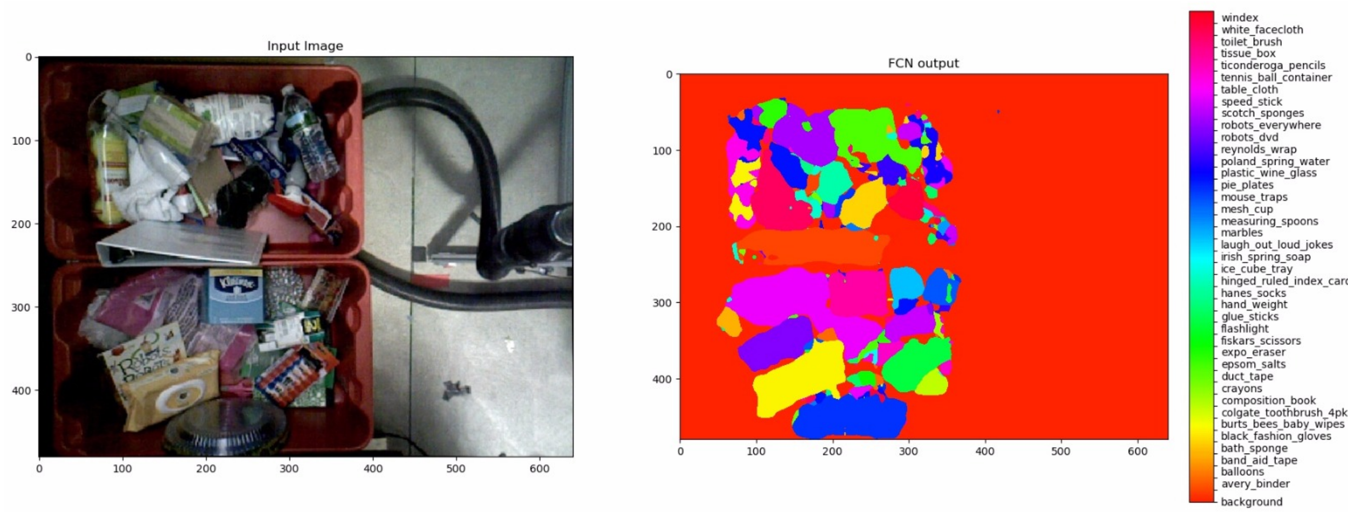

Figure 5: FCN results for a heavily occluded stowage environment.

5.2 Planning

EGWA plan times were on the order of .2-.3 seconds per plan, which were well within time requirements. Path constraints operated as desired, with the arm rejecting any plans that would cause tangling of any of its hosing or wiring. Planning execution was slower than anticipated, at around 3 seconds per plan. This delay in execution time has been identified as being caused by a combination of redundant collision check parameters and too high of detail in the collision modeling for the planning environment. A memory leak was also discovered in the planning module which only allows the arm to execute 32 plans before causing the driver to crash. Both of these issues will be resolved as part of continuing work this summer.

5.3 Grasping

Grasp metrics worked as expected and generated appropriate poses for each item. A new issue was discovered during runtime however, in which the 1-DOF suction head would quickly move from 0-90 degrees or vice versa, and either sheer an item off on the edge of the stowage tote or simply fling the item away. This issue can be solved by creating motion constraints for the 1-DOF gripper that keep it in a downward orientation while it is grasping an item. These constraints will be implemented as part of continuing work this summer.

5.4 Overall

Overall the system was able to pick 7-8 items at maximum per run, out of the initially proposed 10. This was primarily due to the memory leak within the system planner, as 32 motions was simply not enough planning executions to stow more than 7-8 items. The average stowage time for each item was 1 minute and 10 seconds. Extrapolating this over a 15 minute run represents a theoretical 12 items stowed within the allotted time, which would have been within the desired metric. Additionally the system did occasionally show failure cases during stowage runs, either by sheering or dropping an item during transportation, or by placing an item too close to a shelf edge and causing it to fall into an unintended location within the shelf. Stowage runs which demonstrated these failure cases accounted for about 1 out of 3 runs. With the exception of these failure cases all item locations were reported by the system correctly at the end of each run.

6 Work Division

This work division for this project reflected primary ownership for various tasks, and these roles are continuing into further work this summer. All team members are and have been expected to provide support for all aspects of the project in addition to their primary responsibilities.

6.1 Michael Beck

Michael is the project manager. He is responsible for managing deadlines for all the team's tasks and making sure any unforeseen obstacles get resolved in a timely manner. He is also in charge of hardware purchasing and fabrication for the system.

6.2 Akshay Bhagat

Akshay is in charge of system calibration, grasping, and assisting Jin with the vision system. This includes both intrinsic and extrinsic calibration of the system sensors, all algorithms pertaining to generation of grasp points for items to be passed to the arm planner, and providing aid in training and troubleshooting the vision system.

6.3 Matthew Lauer

Matthew is in charge of arm planning. This includes maintaining the planning scene including all obstacles within the ROS MoveIt! environment, managing the SBPL planner, generating arm poses and training the planner's experience graph, and executing arm movements for tests and demonstrations.

6.4 Che-Yen Lu, "Leo"

Leo is in charge of the system's software architecture, as well as bin localization. As the software architect he is routinely updating and managing the system code, as well as providing instruction to other teammates as to their respective algorithm designs. For bin localization he is responsible for writing the code to allow the system sensor to locate an accurate pose of the system bins and storage tote through the use of April tags.

6.5 Jin Zhu

Jin is in charge of training the vision systems, and assisting Michael with project management. Vision system training includes the collection of data, the labeling of data for ground truth, and the training of convolutional neural networks from that data.

7 Video Demonstration

A video demonstration of the system can be found here. This video shows the system picking a total of nine items as part of two separate runs, with the first run demonstrating some failure cases.