

# Multistage Localization for High Precision Mobile Manipulation Tasks

Christopher James Mobley

Thesis submitted to the Faculty of the  
Virginia Polytechnic Institute and State University  
in partial fulfillment of the requirements for the degree of

Masters of Science  
in  
Mechanical Engineering

Tomonari Furukawa  
Brian Lattimer  
Kevin Kochersberger

Blacksburg, Virginia

Keywords: Autonomous Navigation, SLAM, Visual Servoing, Mobile Manipulation,  
Computer Vision, State Machine  
Copyright 2016, Christopher James Mobley

# Multistage Localization for High Precision Mobile Manipulation Tasks

Christopher James Mobley

## (ABSTRACT)

This paper will show the general framework necessary in order to solve two main problems preventing the development and widespread adoption of automation and robotics in construction (ARC). These problems include the fact that typical construction sites tend to be unstructured and are continuously evolving versus the highly controlled environments found in manufacturing. Also, the relationship between the part and manipulator has been reversed, causing increased complexity not seen in manufacturing environments where the part arrives at a fixed manipulator. The techniques that will be presented allow systems to create a 2-D map of their environment, localize themselves and complete the task(s) assigned. After localizing an augmented reality (AR) tag at the work site, the system is able to use a priori knowledge to localize points of interest (POIs) and complete several different types of operations achieving an accuracy of approximately  $\pm 2$  mm based on a multifaceted computer vision approach with only a USB webcam.

# Acknowledgments

Insert Acknowledgement

# Contents

<b>1</b>	<b>Introduction</b>	<b>1</b>
1.1	Background . . . . .	1
1.2	Objectives . . . . .	2
1.3	Summary of Original Contributions . . . . .	2
1.4	Outline . . . . .	2
<b>2</b>	<b>Literature Review</b>	<b>3</b>
2.1	Localization and Mapping For Autonomous Mobile Manipulators in Manufacturing and Construction . . . . .	3
2.2	Task Association and A Priori Knowledge for Mobile Manipulators . . . . .	3
2.3	Feature Localization Techniques for Mobile Manipulators . . . . .	3
2.4	Summary . . . . .	3
<b>3</b>	<b>Fundamentals of Autonomous Robotics</b>	<b>4</b>
3.1	ROS Concepts . . . . .	4
3.1.1	ROS Communication . . . . .	4
3.1.2	Rigid Body Transformations . . . . .	5
3.2	Simultaneous Localization and Mapping Concepts . . . . .	6
3.3	Localization and Path Planning Concepts . . . . .	8
3.4	Manipulation Concepts . . . . .	14
3.5	State Machines Concepts . . . . .	17
3.6	Camera Concepts . . . . .	17

3.7	Computer Vision Concepts . . . . .	21
3.7.1	Color Spaces . . . . .	21
3.7.2	Linear and Non-Linear Filters . . . . .	22
3.7.3	Binary Operations . . . . .	24
3.7.4	Hough Circle Detection . . . . .	25
3.7.5	Good Feature To Track . . . . .	27
3.7.6	Optical Flow . . . . .	30
3.7.7	Pose Estimation and Tracking Through Augmented Reality Tag Detection . . . . .	32
3.8	Summary . . . . .	33
<b>4</b>	<b>Multistage Localization for High Precision Mobile Manipulation Tasks</b>	<b>34</b>
4.1	Approach Overview . . . . .	34
4.2	Global Map Creation and Task Location Specification . . . . .	35
4.3	Autonomous Localization and Navigation . . . . .	35
4.4	Task Association and A Priori Knowledge . . . . .	36
4.5	Generic Framework for Multi-Stage Computer Vision Algorithm . . . . .	36
4.5.1	Initial Feature Location Prediction . . . . .	36
4.5.2	Corrected Feature Locations . . . . .	37
4.6	Approach Implementation . . . . .	39
4.6.1	System Overview . . . . .	39
4.6.2	Drilling Framework . . . . .	40
4.6.3	Sealant Application Framework . . . . .	41
4.7	Summary . . . . .	42
<b>5</b>	<b>Experiments and Results</b>	<b>43</b>
5.1	Hardware Architecture . . . . .	43
5.2	Software Architecture . . . . .	44
5.3	Experiments . . . . .	44

5.3.1	Camera Calibration Setup . . . . .	44
5.3.2	Navigation System Experimental Setup . . . . .	47
5.3.3	Drilling Operation Experimental Setup . . . . .	47
5.3.4	Sealant Application Experimental Setup . . . . .	48
5.4	Results . . . . .	49
5.4.1	Camera Calibration Accuracy Achieved . . . . .	49
5.4.2	Navigation System Accuracy Achieved . . . . .	49
5.4.3	Augmented Reality Tag Accuracy Achieved . . . . .	54
5.4.4	Drilling Operation Accuracy Achieved . . . . .	56
5.4.5	Sealant Application Accuracy Achieved . . . . .	57
5.5	Summary . . . . .	58
<b>6</b>	<b>Conclusion and Future Work</b>	<b>59</b>
<b>Bibliography</b>		<b>60</b>

# List of Figures

1.1	Vision for future transformable production systems . . . . .	2
3.1	Simple ROS Node Flowchart . . . . .	4
3.2	Conversion from Coordinate Frame A to B . . . . .	5
3.3	Robot Model with TFs . . . . .	6
3.4	Overview of SLAM Framework. [Riisgaard et al. 2005] . . . . .	6
3.5	Example of 2-D Occupancy Grid Map Produced. . . . .	7
3.6	Real World Performance Analysis of ROS Available SLAM Algorithms. [Santos et al. 2013] . . . . .	7
3.7	ROS Navigation Stack setup. . . . .	8
3.8	One-Dimensional Monte Carlo Localization Example. . . . .	10
3.9	Particle Filter Resampling Example . . . . .	12
3.10	Visual of AMCL in RVIZ . . . . .	13
3.11	Trajectory Rollout Path Planning Framework . . . . .	14
3.12	Moveit!'s System Architecture. . . . .	15
3.13	Forward and Inverse Kinematics Example . . . . .	16
3.14	Graphical View of State Machine using SMACH . . . . .	17
3.15	Photometric Image Formation. . . . .	18
3.16	Digital Camera Diagram. . . . .	19
3.17	Ground Sample Distance Effects on Image Quality. . . . .	20
3.18	Shutter Type Effects on Image Quality. . . . .	21
3.19	RGB and HSV color space models . . . . .	22

3.20	Effects of common linear filters.	23
3.21	Histogram equalization depiction.	24
3.22	Effect of histogram equalization on an image.	24
3.23	Effects of common binary operations.	25
3.24	Summary of preprocessing operation performed by Hough Circle detector.	26
3.25	Summary of Hough Circle detector with known radius.	26
3.26	Summary of Hough Circle detector with unknown radius.	27
3.27	Summary of Good Feature detector.	27
3.28	Image Gradient Example.	28
3.29	Difference Between Good Feature and Harris Corner Scoring Functions.	30
3.30	Example of optical flow.	31
3.31	Summary of Kanade-Lucas-Tomasi feature tracker.	32
3.32	Summary of ALVAR AR Tag Detection Framework.	33
4.1	Multistage Localization Approach Overview.	34
4.2	2-D map environment with specified start locations.	35
4.3	Robot Localization Framework.	36
4.4	Initial Feature Localization Framework.	36
4.5	Feature Detection and Tracking Pipeline.	38
4.6	Manipulator Correction Control Loop.	39
4.7	General system overview.	40
4.8	Drilling operation framework.	41
4.9	Sealant Application Framework.	42
5.1	Clearpath Robotics' Husky and Fetch Robotics' Fetch Mobile Manipulator	44
5.2	Camera Calibration Setup.	45
5.3	Extracted Corners of Calibration Pattern.	46
5.4	Extracted Corners and Laser Center	46
5.5	Navigation System Experimental Setup.	47

5.6	Drilling Operations Experimental Setup. . . . .	48
5.7	Sealant Application Experimental Setup. . . . .	49
5.8	Global X and Y Localization Accuracy Given Distance From Start - Clearpath Robotics' Husky. . . . .	50
5.9	Global $\psi$ Localization Accuracy Given Distance From Start - Clearpath Robotics' Husky. . . . .	51
5.10	Global X and Y Localization Accuracy Given Distance From Start - Fetch Robotics' Fetch. . . . .	52
5.11	Global $\psi$ Localization Accuracy Given Distance From Start - Fetch Robotics' Fetch. . . . .	53
5.12	Augmented Reality Accuracy Given Specific Start Conditions - Primesense Carmine 1.09. . . . .	55
5.13	Augmented Reality Accuracy Given Specific Start Conditions - Microsoft Kinect V2. . . . .	56
5.14	Drilling Operation Accuracy Given Distance From Feature Point. . . . .	57
5.15	Sealant Application Accuracy Given Distance From Feature Point. . . . .	58

# List of Tables

3.1	Real World Error Estimation for ROS Available SLAM Algorithm. [Santos et al. 2013] . . . . .	7
3.2	Comparison Between ROS Available Moveit! Inverse Kinematic Plugins. [Beeson et al. 2015] . . . . .	16
5.1	Global Localization Accuracy Given Distance From Start - Clearpath Robotics' Husky. . . . .	50
5.2	Global Localization Accuracy Given Distance From Start - Fetch Robotics' Fetch. . . . .	51
5.3	Augmented Reality Accuracy Given Specific Start Conditions - Primesense Carmine 1.09. . . . .	54
5.4	Augmented Reality Accuracy Given Specific Start Conditions - Microsoft Kinect V2. . . . .	55
5.5	Drilling Operation Accuracy Given Distance From Work Surface. . . . .	56
5.6	Sealant Application Accuracy Given Distance From Work Surface. . . . .	57

# Chapter 1

## Introduction

### 1.1 Background

Unlike the substantial benefits seen in the manufacturing industry through automation and robotics, automation and robotics in construction (ARC) has lagged far behind in adoption [1]. Consequently, when compared with other industries, construction has seen a significant decrease in productivity, as well as an increase in workplace injuries/fatalities over the last several decades [2]. While several technical complexities inherent in construction have hindered the development and adoption of field construction robots [4], through the capitalization of advances made by other industries, ARC can quickly close this gap. Thereby allowing dangerous and or mundane repetitive tasks to be accomplished autonomously. Thus, causing an increase in productivity and a decrease in workplace injuries/fatalities [1]. However, ARC faces two unique challenges when compared to other industries. Unlike manufacturing environments, which are tightly controlled, typical construction sites tend to lack structure and are continuously evolving. In addition, the reversal in relationship between the part and manipulator has dramatically increased the complexity of the problem to be solved [3]. Instead of the part appearing at a fixed manipulator, the manipulator must now move to and localize itself with respect to the part. The remainder of this paper is structured as follows: In Section 2, the author's technical approach is outlined, with particular focus on problem two, and experimental results are shown. Conclusions are then drawn and future work discussed in Section 3.

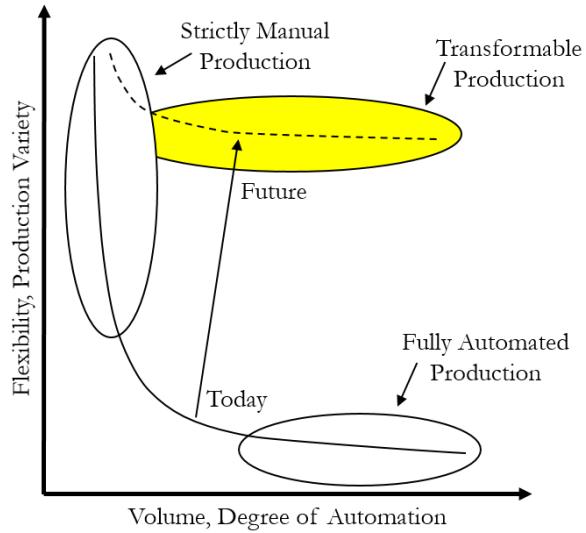


Figure 1.1: Vision for future transformable production systems.

## 1.2 Objectives

Insert Text Regarding the Objectives of This Work.

## 1.3 Summary of Original Contributions

Insert Text Regarding the Originals Contributions Presented in this Work.

## 1.4 Outline

Insert Text Outlining the Upcoming Chapters.

# **Chapter 2**

## **Literature Review**

### **2.1 Localization and Mapping For Autonomous Mobile Manipulators in Manufacturing and Construction**

Insert Text Summarizing Current SLAM and Localization Techniques Used.

### **2.2 Task Association and A Priori Knowledge for Mobile Manipulators**

Insert Text Summarizing Current Methods Used to Associate Mobile Manipulator to a Specific Task and How Prior Knowledge is Conveyed About the Task to be Performed.

### **2.3 Feature Localization Techniques for Mobile Manipulators**

Insert Text Summarizing Current Methods Used to Localize a Feature and Have a Mobile Manipulator Perform a Set Operation.

### **2.4 Summary**

Insert Text Summarizing This Chapter and Transiting to the Next.

# Chapter 3

## Fundamentals of Autonomous Robotics

Insert text outlining upcoming sections.

### 3.1 ROS Concepts

Insert Text Defining ROS.

#### 3.1.1 ROS Communication

Insert Text Regarding Node/Topic/Service Communication.

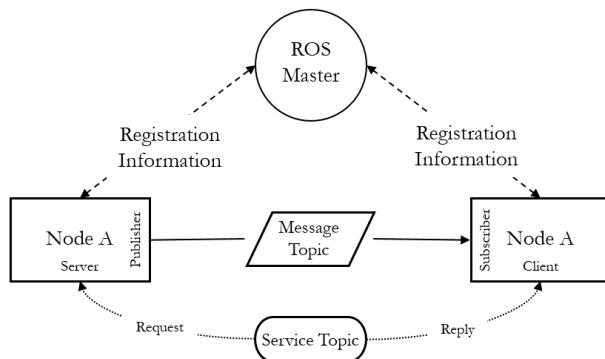


Figure 3.1: Simple ROS Node Flowchart

### 3.1.2 Rigid Body Transformations

Insert Text Explaining Rigid Body Transformations.

$$x_a = T_b^a x_b \quad (3.1)$$

where  $T_b^a$  is equal to

$$\begin{bmatrix} R_b^a & t_b^a \\ 0^T & 1 \end{bmatrix} \quad (3.2)$$

where  $R_b^a$  is the rotation matrix, which performs the rotation part of moving frame  $b$  into alignment with frame  $a$  and  $t_b^a$  is the translation matrix, which performs the translation part of moving frame  $b$  origin to frame  $a$ .

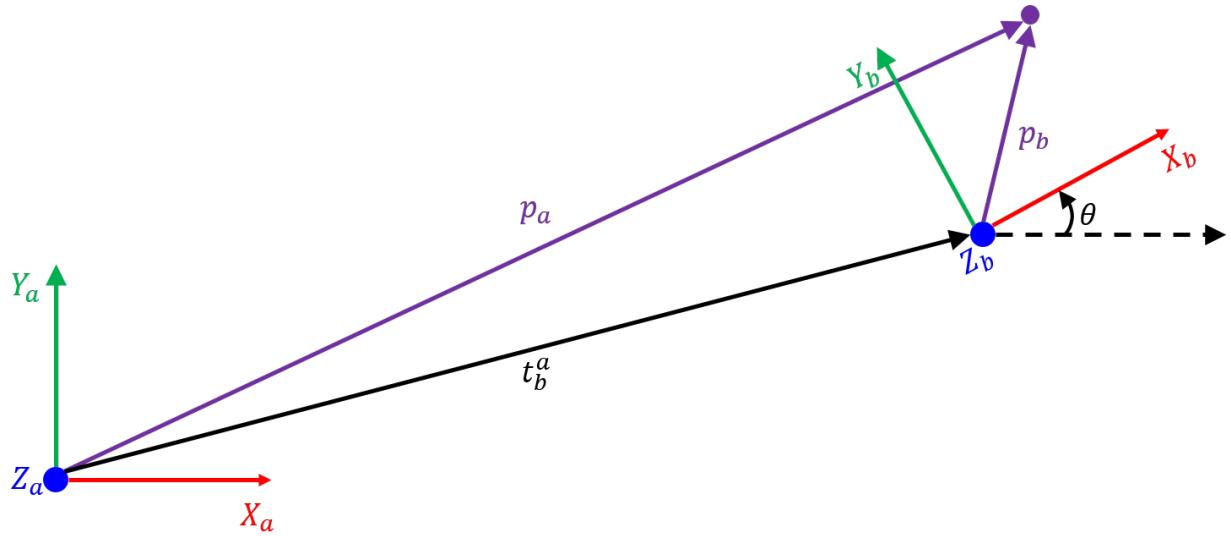


Figure 3.2: Conversion from Coordinate Frame A to B

Insert Text Regarding How a URDF is Setup in ROS so that Transformation Between Coordinates can be Managed with the Package TF.

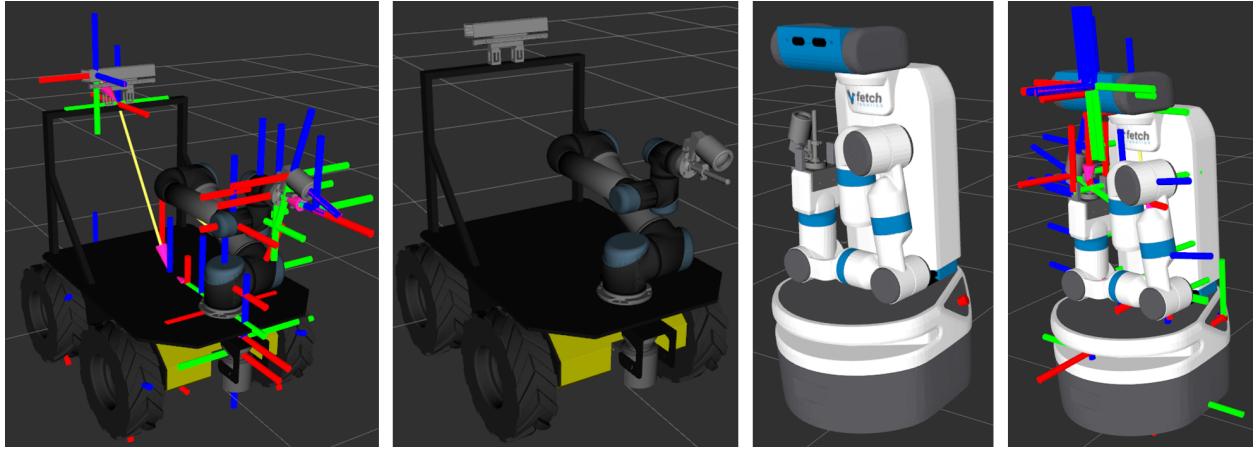


Figure 3.3: Robot Model with TFs

## 3.2 Simultaneous Localization and Mapping Concepts

Insert Text Outlining the Basic Intuition Behind SLAM.

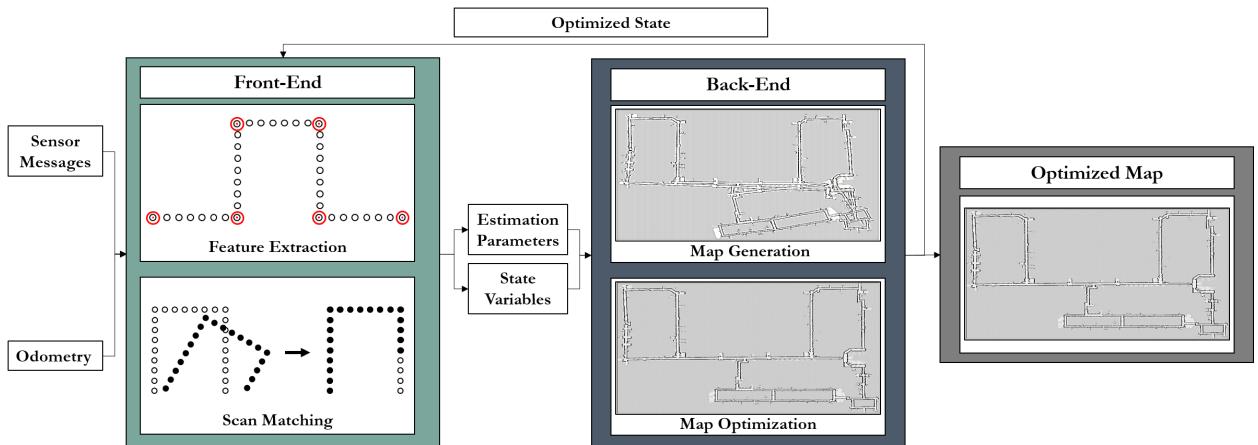


Figure 3.4: Overview of SLAM Framework. [Riisgaard et al. 2005]

Insert Text Explaining How ROS Uses SLAM to Create 2-D Occupancy Grid Maps and What They Are Used.

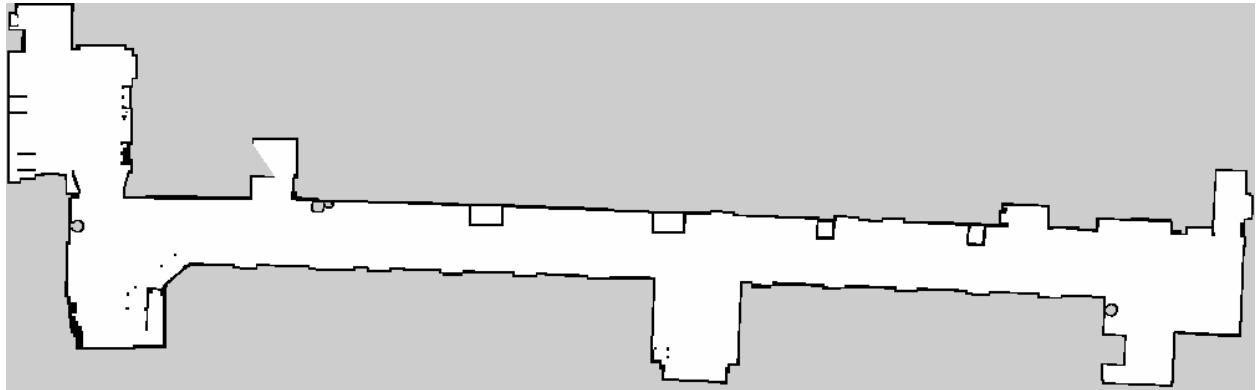


Figure 3.5: Example of 2-D Occupancy Grid Map Produced.

Insert Text Specifying Available 2-D SLAM Techniques and Why KartoSLAM was Chosen.

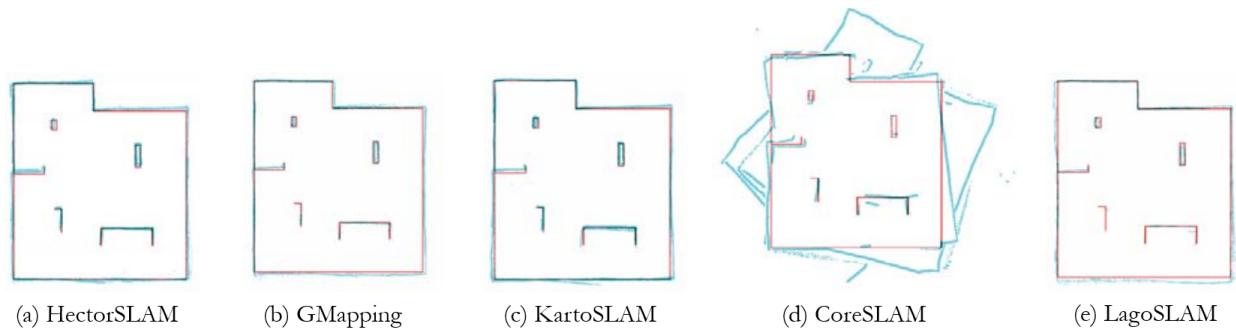


Figure 3.6: Real World Performance Analysis of ROS Available SLAM Algorithms. [Santos et al. 2013]

Table 3.1: Real World Error Estimation for ROS Available SLAM Algorithm. [Santos et al. 2013]

Real World Experiments				
HectorSLAM	GMapping	KartoSLAM	CoreSLAM	LagoSLAM
1.1972	2.1716	1.0318	14.75333	3.0264
0.5094	0.6945	0.3742	7.9463	0.8181
1.0656	1.6354	0.9080	7.5824	2.5236

### 3.3 Localization and Path Planning Concepts

Insert Text Outlining The Navigation Stack's Framework.

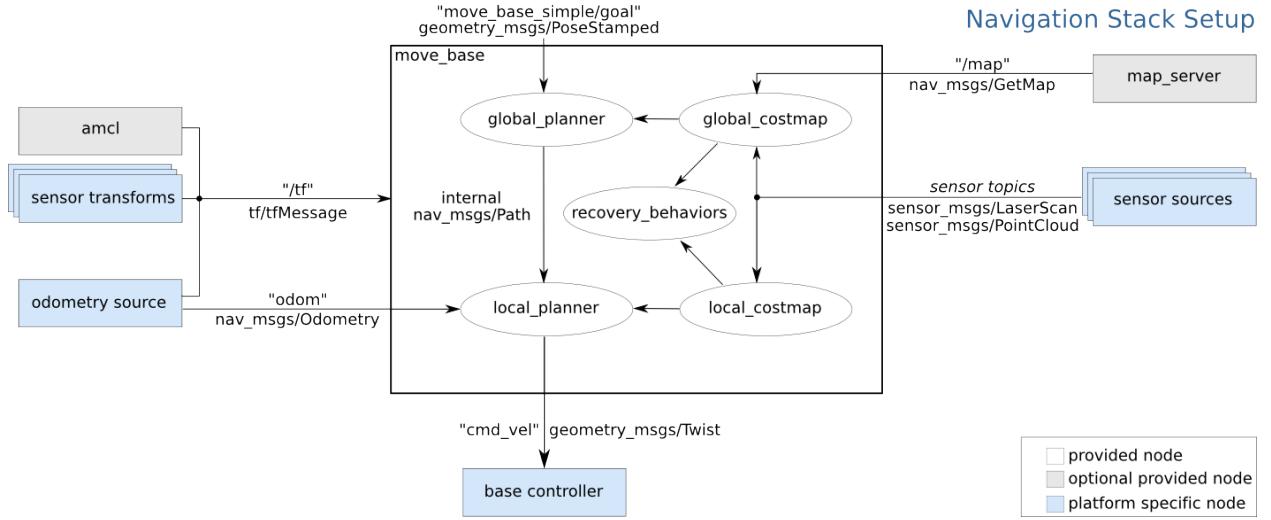


Figure 3.7: ROS Navigation Stack setup.

After building a two-dimensional occupancy grid map, shown in Figure 3.5, using SLAM, it becomes crucial to accurately localize the robot within this predefined map; so that, the robot can both plan and execute appropriate trajectories to its next destination. Localization involves estimating the position and orientation of the robot, known collectively as pose, while the robot moves throughout its environment. One routine localization technique used involves tracking the robot from an initial known starting pose. Through the measurement of wheel rotation and the integration of accelerations provided by an inertial measurement unit (IMU), the distance traveled by the robot from the initial position can be calculated and the robot's pose in the map estimated with some certainty. However, these methods do not account for wheel slippage or measurement error. As a result, the accuracy of the pose estimate will degrade over time. Consequently, a solution which can compensate for the accumulated odometry error and inaccuracies in the initial starting pose is needed. One accepted solution to this problem is Monte Carlo Localization (MCL), which utilizes a particle filters to keep track of the robot's pose. However, additional options include Kalman Filters and Markov Localization, which employ Gaussian distributions and histograms respectively.

Figure 3.8 depicts MCL using a one-dimensional corridor with a few doors. The robot initially has no information about where it is in this corridor. As a result, the graph of the robot's belief states, which defines the probability of the robot being at a particular position, is drawn from a uniform distribution of discretely sampled positions along the

corridor. A measurement update is performed at each time step. A measurement update involves convolving the measurement model, the probability of receiving a specific sensor measurement in the corridor, with the belief states to get an updated belief state. The updated belief state is the same as the previous belief state; however, the weight of each particle have been updated based on the sensor reading. At step k=1 the robot senses a door; so, the weight of particles at the three door are increased. At the next step, a motion model update is performed. The odometry indicated that the robot moved forward a specified distance d. As a result, the belief state is updated by moving the particles forward that specific distance with noise added to account for the aforementioned odometry errors. It should be noted that the particles at this stage in Figure 3.8 were also resampled, which will be covered in the following paragraphs. The motion model update is followed by a measurement update. The robot again senses a door. As a result, the measurement model is the same as the previous time step. After convolving the current measurement model with the current belief state, the cumulative probability mass is centered at door two indicating that the robot is likely at this location [4].

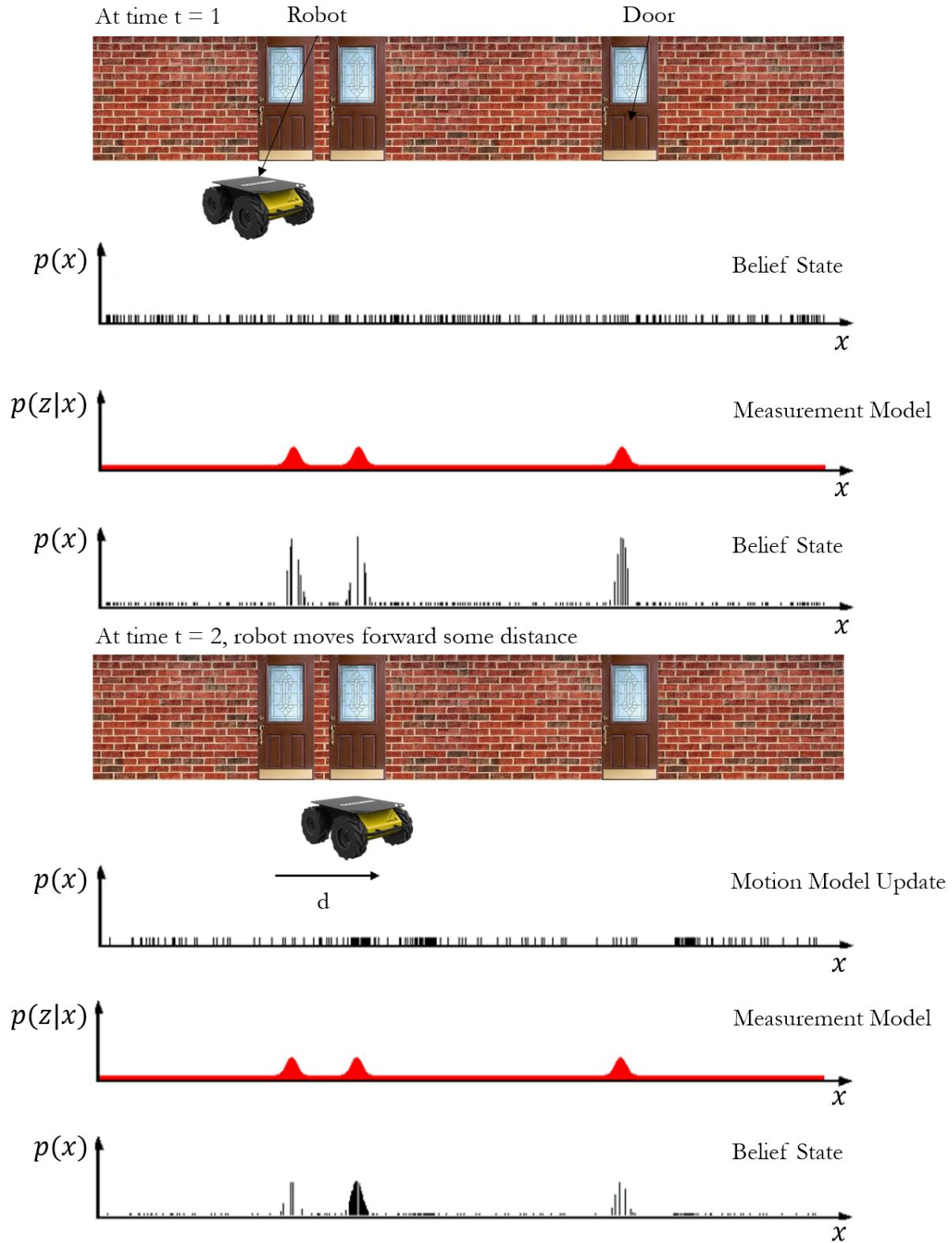


Figure 3.8: One-Dimensional Monte Carlo Localization Example.

Two-dimensional MCL follows the same format of a motion model and then measurement update. The initial particles are drawn from the current odometry with added noise. At each step, the particles are updated via the odometry and then corrected via a measurement update. For each particle, the correlation between the two-dimensional occupancy grid map, seen in Figure 3.5, and laser scan is calculated using Equation 3.3, where  $A$  is the predefined map,  $B$  is the map created by the current laser scan, and  $\bar{A}$  and  $\bar{B}$  are the mean values of the pixels of both maps respectively. Note that while obstacle pixels are black and have a value of 1, free space pixels are grey and have a value of 0. While  $m$  and  $n$  are the x and y pixel values. The particle (pose) with the highest correlations score is chosen as the pose for the current step. The weight of each particles at step  $k$  is found by multiplying the particles weight at step  $k - 1$  by its normalized correlation score at time step  $k$ , as seen in Equation 3.4.

$$s = \frac{\sum_m \sum_n (A_{mn} - \bar{A})(B_{mn} - \bar{B})}{\sqrt{(\sum_m \sum_n (A_{mn} - \bar{A})^2)(\sum_m \sum_n (B_{mn} - \bar{B})^2)}} \quad (3.3)$$

$$W_k \leftarrow W_{k-1}s \quad (3.4)$$

As weights are multiplied over steps, the particles with consecutive small correlation scores are reduced to very small values. As a result, the particles filter eventually has every few effective particles to ensure that good results are produced. Consequently, re-sampling, show in Figure 3.9, is performed when the number of effective particles become too small. This is done by drawing samples close to the particles that have a higher weights, indicated by their size. Thus the new sample have a higher density near the position where the particle with higher weight existed. Each particle after resampling has same weight. As a result, the particle filter begins from scratch.

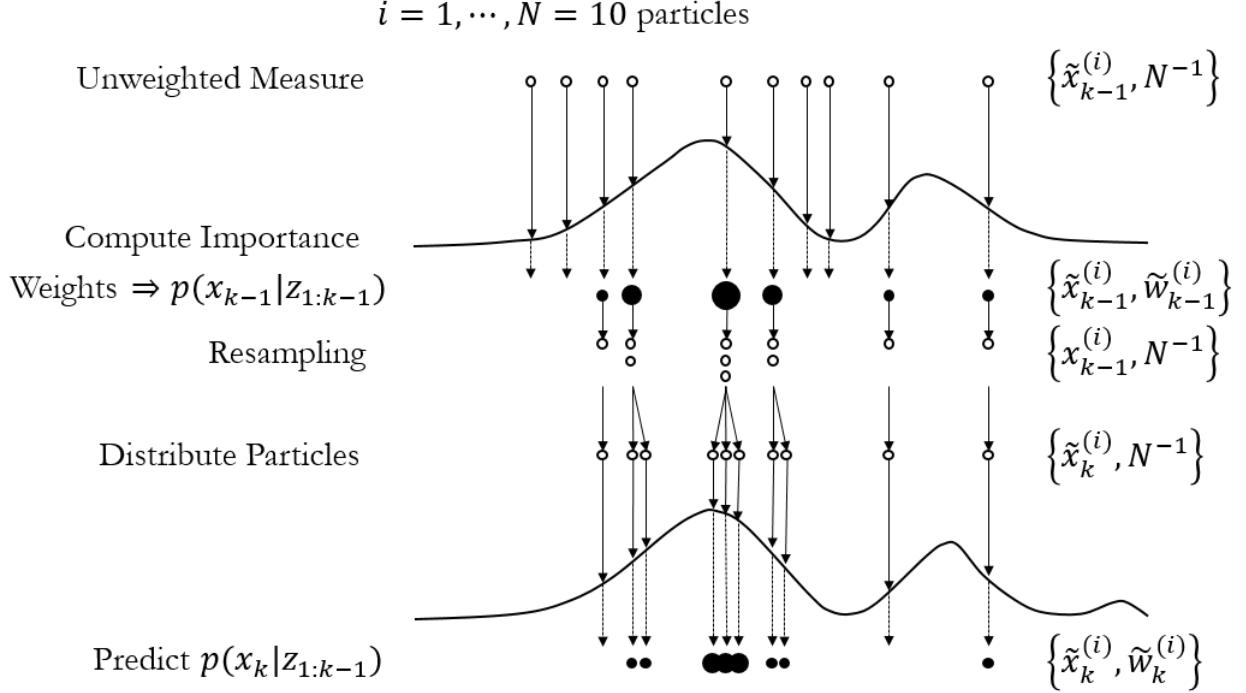


Figure 3.9: Particle Filter Resampling Example

Due to the computational complexity inherent in iteratively having to calculate the correlation score for each particle, an optimization technique known as *KLD-sampling* is used. *KLD-sampling*, derived from *Kullback-Leibler divergence*, is a technique that determine the number of particles needed such that the error between the sample and true posterior is less than  $\epsilon$  [4]. KLD-sampling basically control the number of particles based on the difference in odometry and particle base location.

Figure 3.10 shows how KLD-sampling effectively work. Initially when the position is unknown, the particle cloud is large due to the uncertainty in the position and orientation of the robot. However, as the robot moves, the particle converges and the particle cloud size reduces as KLD-sampling removes the redundant particles and improves computational performance.

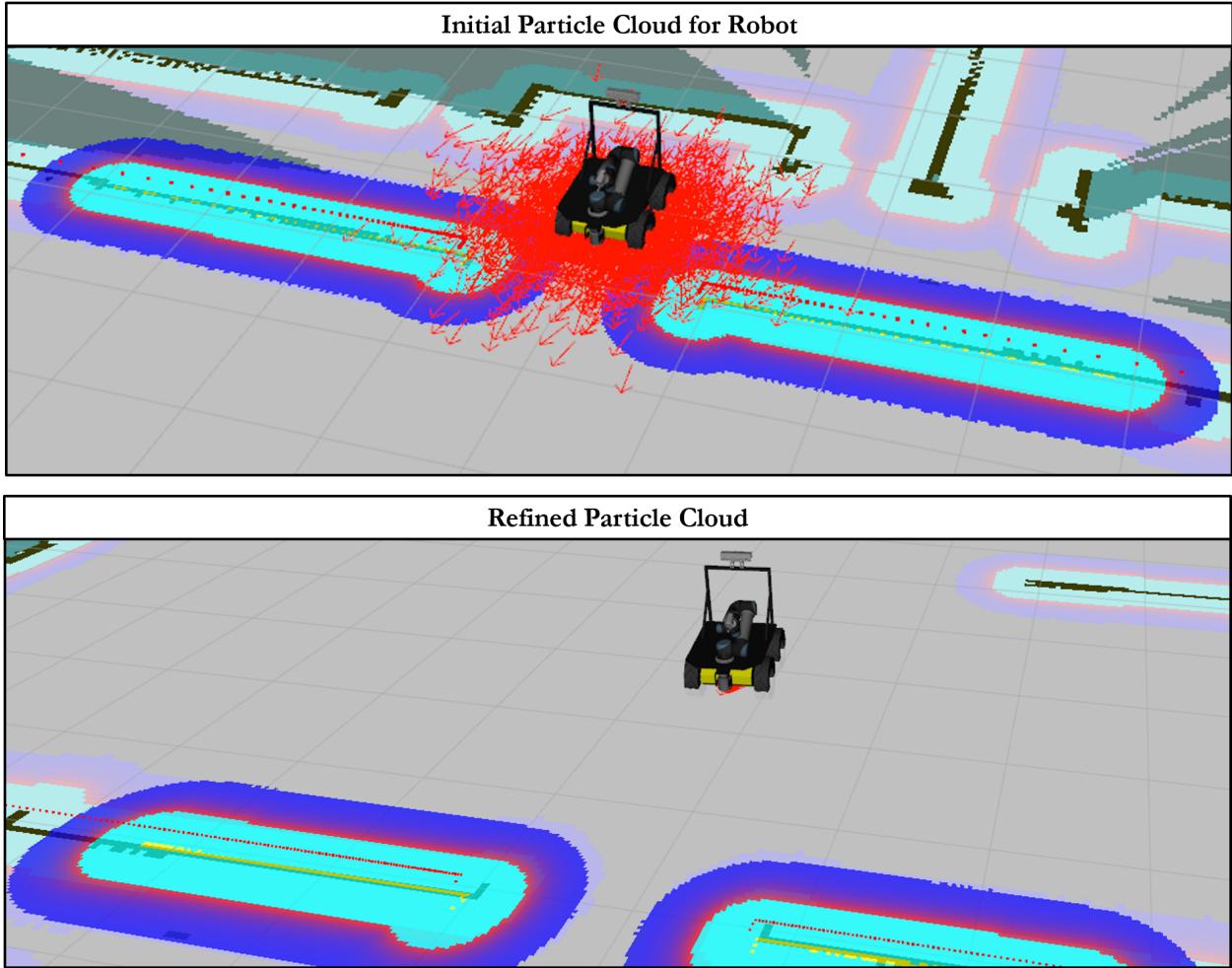


Figure 3.10: Visual of AMCL in RVIZ

Insert Text Explaining How Trajectory Rollout Works for Path Planning.

$$C(k) = \alpha Obs + \beta Gdist + \gamma Pdist + \delta \frac{1}{\dot{x}^2} \quad (3.5)$$

Where  $Obs$  is the sum of grid cell cost through which the trajectory passes (taking account of the robot's actual footprint in the grid);  $Gdist$  and  $Pdist$  are the estimated shortest distance from the endpoint of the trajectory to the goal and the optimal path, respectively; and  $\dot{x}$  is the translation component of the velocity command that produces the trajectory.

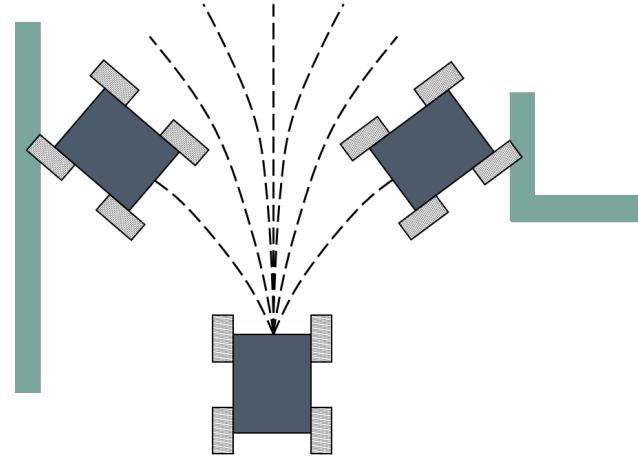


Figure 3.11: Trajectory Rollout Path Planning Framework

### 3.4 Manipulation Concepts

Insert Text Outlining Moveit!'s System Architecture.

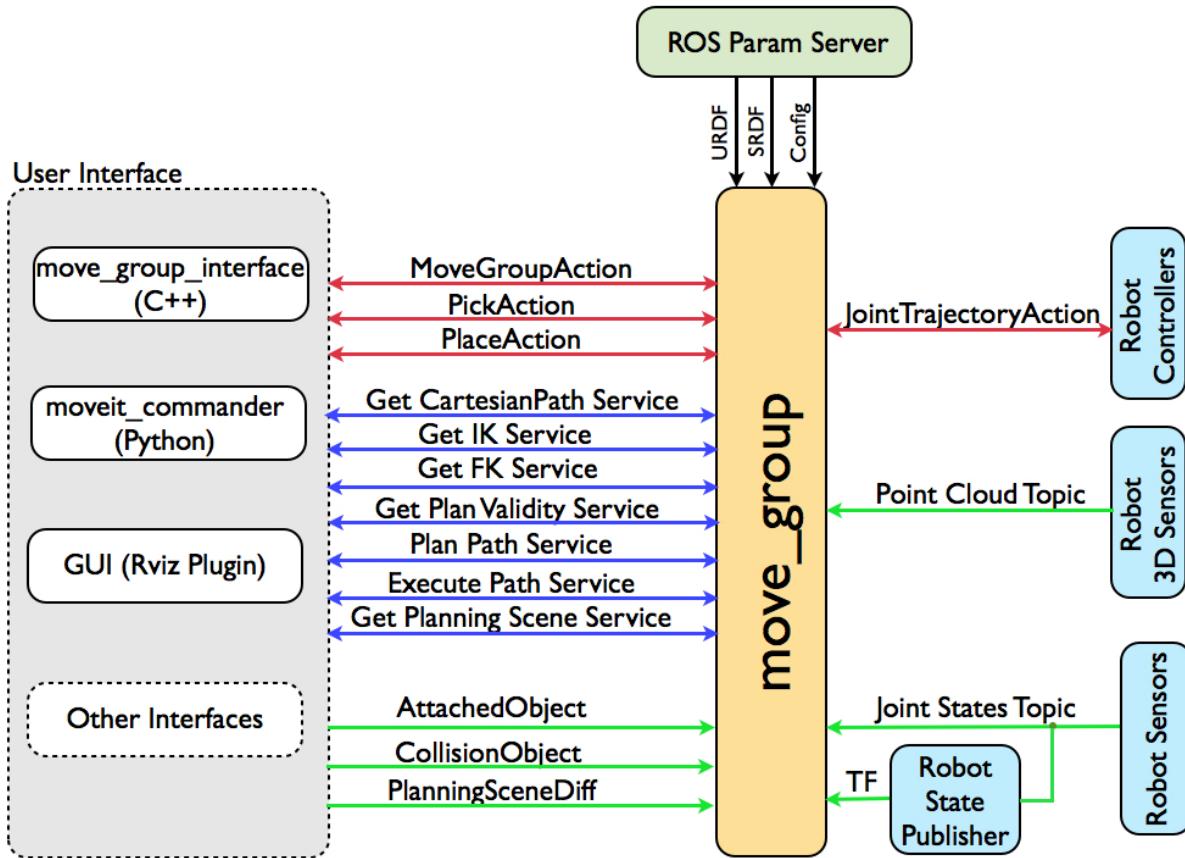


Figure 3.12: Moveit!'s System Architecture.

Insert Text Explaining the Basic Intuition Behind Forward and Inverse Kinematics.

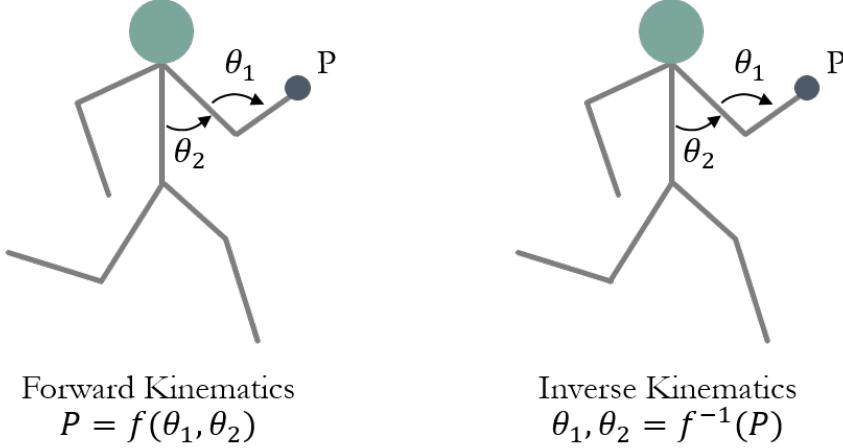


Figure 3.13: Forward and Inverse Kinematics Example

Insert Test Explaining the Difference Between The Numerical Solvers available in ROS (KDL, KDL-RR and Trac\_IK) and Why Trac\_IK was chosen to be used.

Table 3.2: Comparison Between ROS Available Moveit! Inverse Kinematic Plugins. [Beeson et al. 2015]

Robot	DOFs	IK Technique					
		Orcos' KDL		KDL-RR		TRAC-IK	
		Solve Rate (%)	Avg Time (ms)	Solve Rate (%)	Avg Time (ms)	Solve Rate (%)	Avg Time (ms)
Atlas 2013 Arm	6	75.54	1.35	97.13	0.39	99.97	0.33
Atlas 2015 Arm	7	75.71	1.50	93.13	0.81	99.18	0.48
Baxter Arm	7	61.07	2.21	89.52	1.02	99.17	0.60
Denso VS-068	6	27.92	3.69	98.13	0.42	99.78	0.38
Fanuc M-430iA/2F	5	21.07	3.99	88.34	0.92	99.16	0.58
Fetch Arm	7	92.49	0.73	93.82	0.72	99.96	0.44
Jaco2	6	26.23	3.79	97.66	0.58	99.51	0.58
LBR iiWA 14 R820	7	37.71	3.37	94.02	0.73	99.63	0.56
KUKA LWR 4+	7	67.80	1.88	95.40	0.62	99.95	0.38
PR2 Arm	7	83.14	1.37	86.96	1.27	99.84	0.59
NASA Robonaut2 'Grasping Leg'	7	61.27	2.29	87.57	1.10	99.31	0.67
NASA Robonaut2 'Leg + Waist + Arm'	15	97.99	0.80	98.00	0.84	99.86	0.79
NASA Robonaut2 Arm	7	86.28	1.02	94.26	0.73	99.25	0.50
NASA Robosimian Arm	7	61.74	2.44	99.87	0.36	99.93	0.44
TRACLabs Modular Arm	7	79.11	1.35	95.12	0.63	99.80	0.53
UR10	6	36.16	3.29	88.05	0.82	99.47	0.49
UR5	6	35.88	3.30	88.69	0.78	99.55	0.42
NASA Valkyrie Arm	7	45.18	3.01	90.05	1.29	99.63	0.61

### 3.5 State Machines Concepts

Insert Text Explaining the Basics of State Machine and How They are Implemented in SMACH.

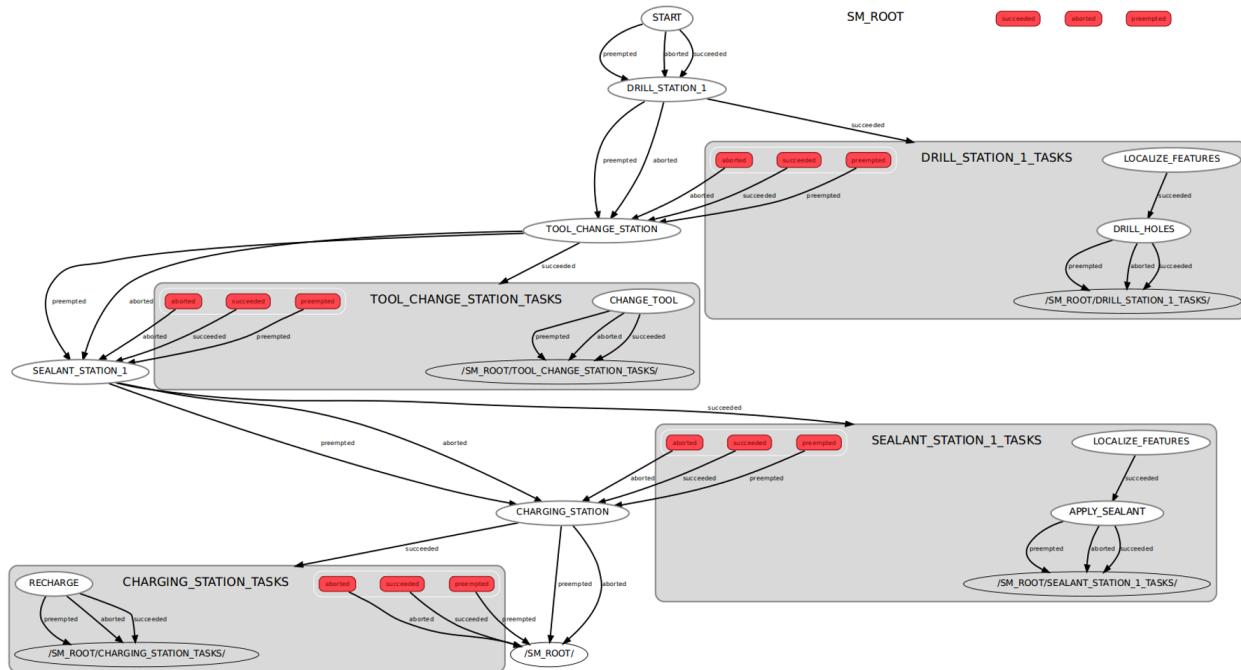


Figure 3.14: Graphical View of State Machine using SMACH

### 3.6 Camera Concepts

Insert Text Explaining How Digital Image Are Formed and Stored.

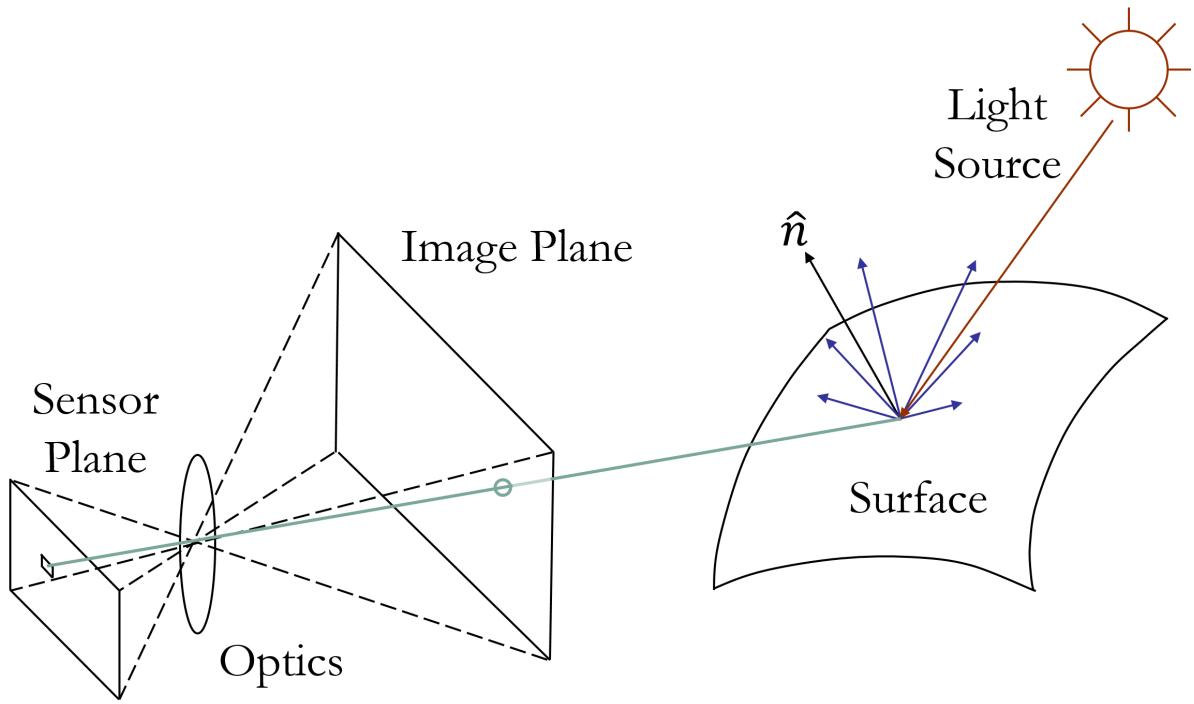


Figure 3.15: Photometric Image Formation.

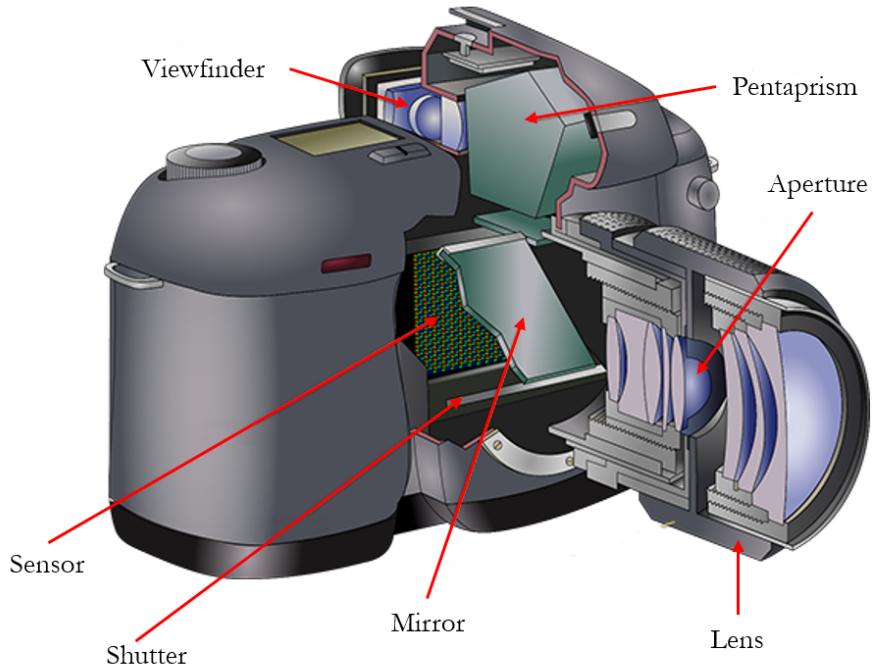


Figure 3.16: Digital Camera Diagram.

Insert Text Explaining Ground Sample Distance and It's Effects on Image Quality Thereby Effecting Ones Camera Choice.

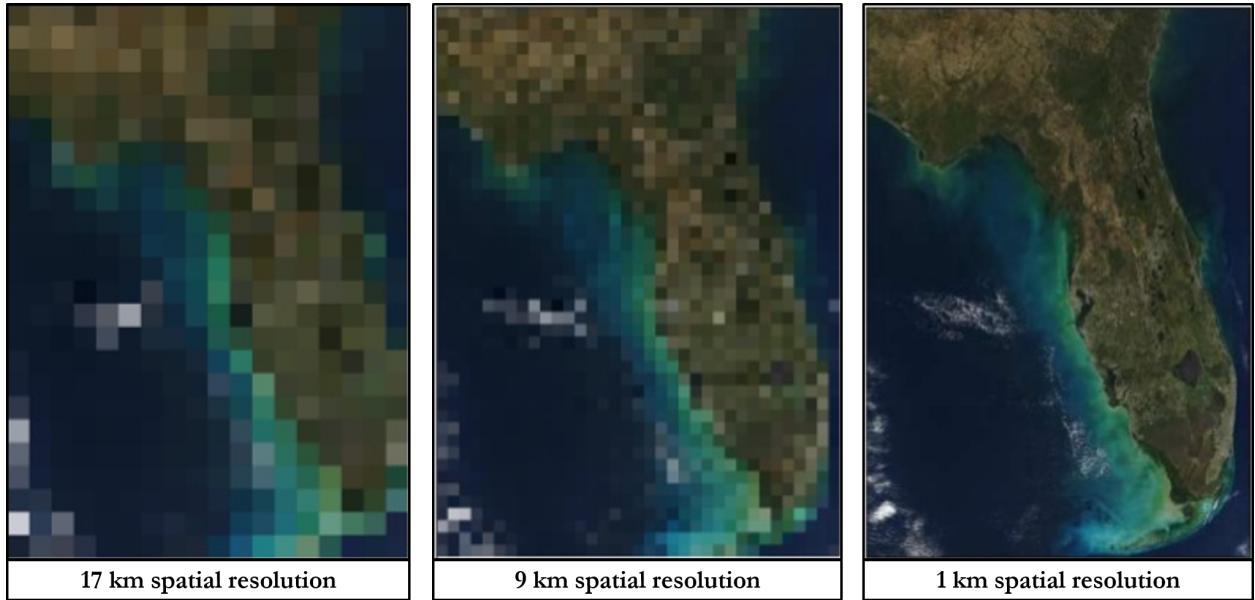


Figure 3.17: Ground Sample Distance Effects on Image Quality.

Insert Text Explaining The Effects of Different Shutter Type on Image Quality Thereby Effecting Ones Camera Choice.

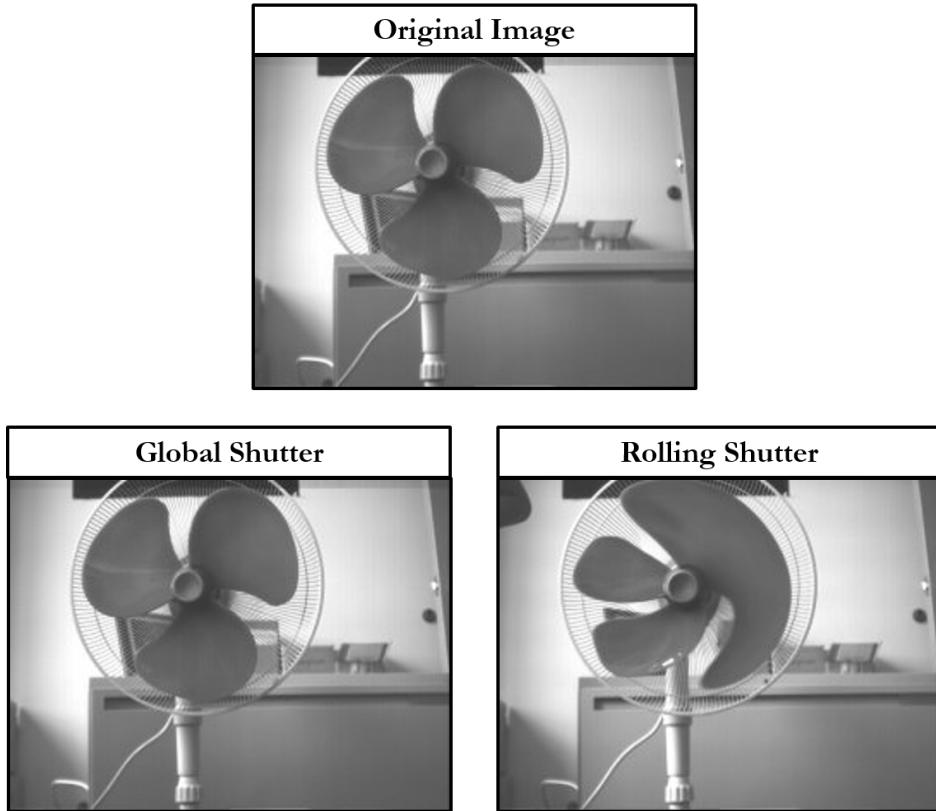


Figure 3.18: Shutter Type Effects on Image Quality.

## 3.7 Computer Vision Concepts

Insert Text Regarding Upcoming Sections.

### 3.7.1 Color Spaces

Insert Text Explaining Both RGB Color Space and HSV/HSL Color Space and Why One Would Be Chosen Over The Other.

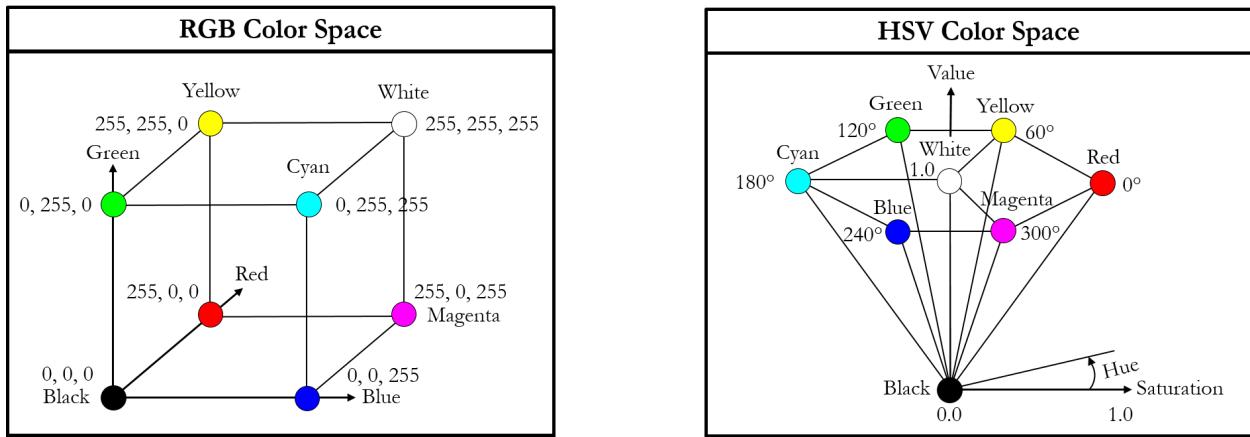


Figure 3.19: RGB and HSV color space models

### 3.7.2 Linear and Non-Linear Filters

Insert Text Explaining the Basics behind Linear Filters, Which Ones Were Chosen and Why.

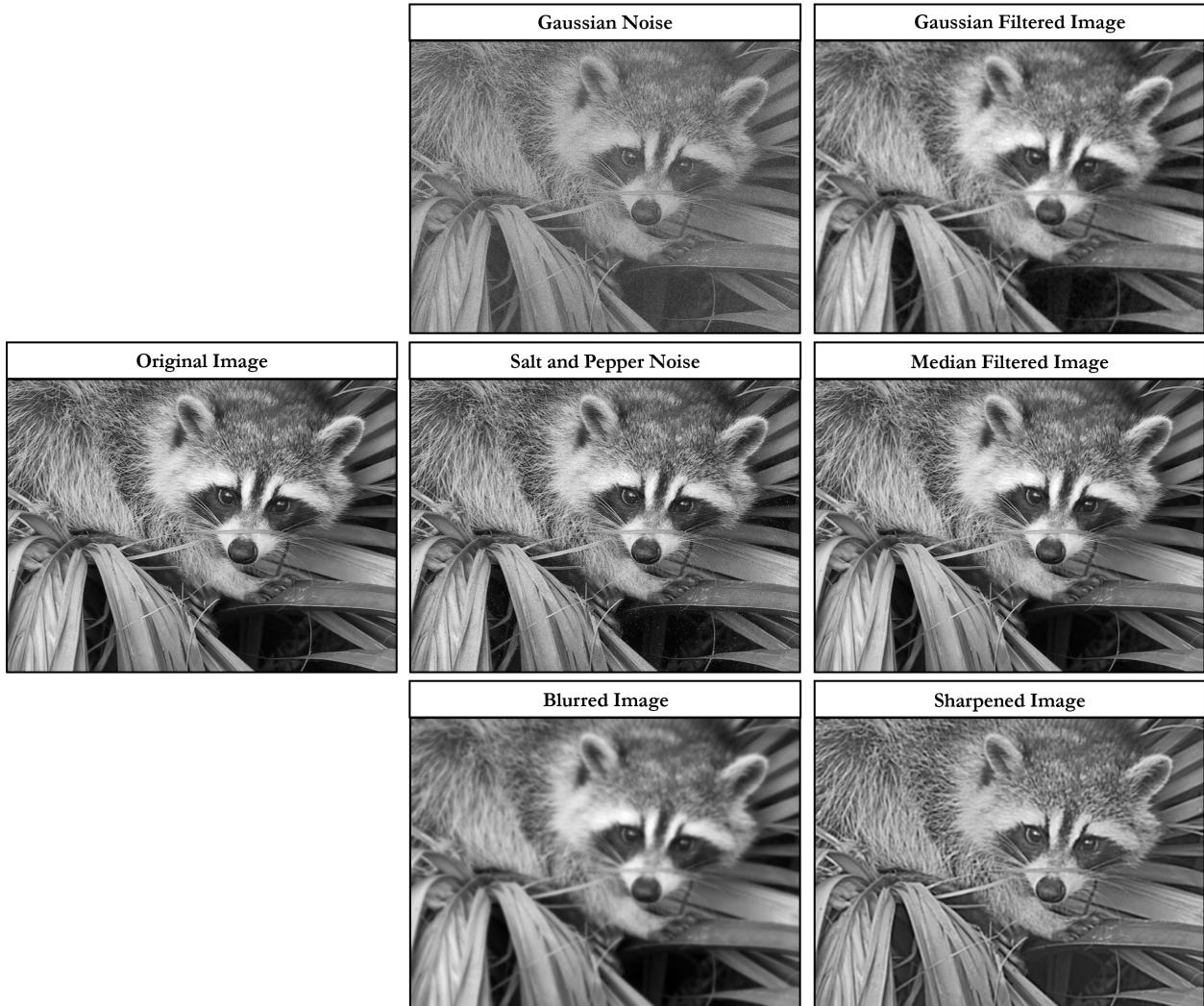


Figure 3.20: Effects of common linear filters.

Insert Text Explaining the Basics behind Histogram Equalization.

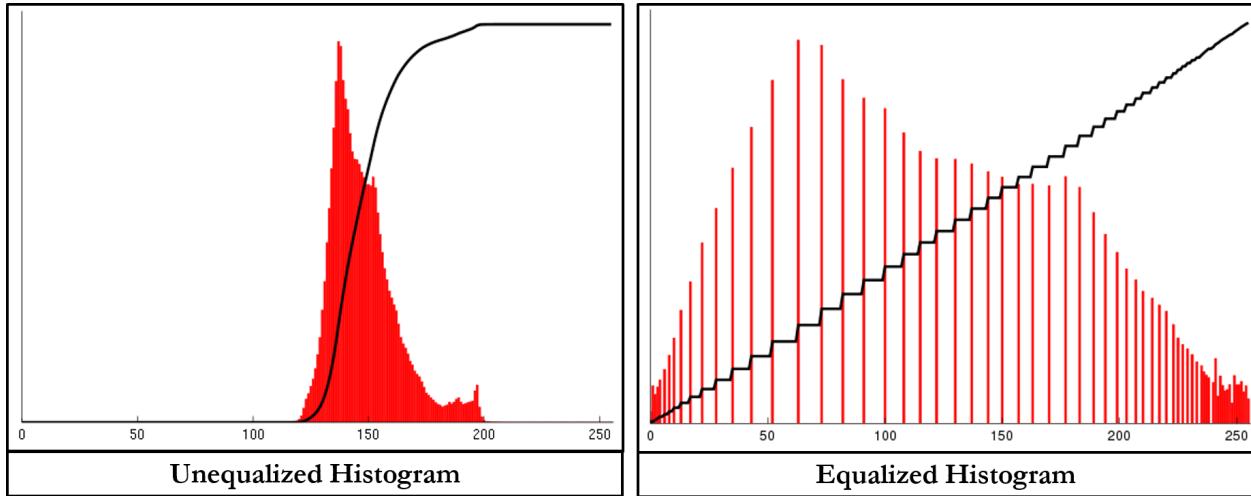


Figure 3.21: Histogram equalization depiction.

Insert Text Explaining The Effect of Histogram Equalization on Images.



Figure 3.22: Effect of histogram equalization on an image.

### 3.7.3 Binary Operations

Insert Text Explaining the Basics behind Binary Operations, Which Ones Were Chosen and Why.

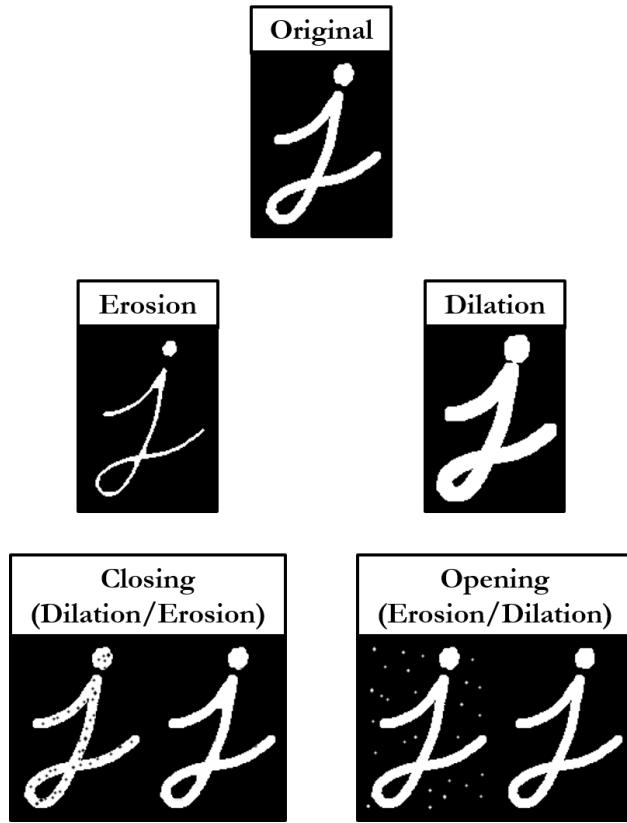


Figure 3.23: Effects of common binary operations.

### 3.7.4 Hough Circle Detection

Insert Text Explaining The Basic Preprocessing Operation Performed by Hough Circle Transform In Order to Get Edges.

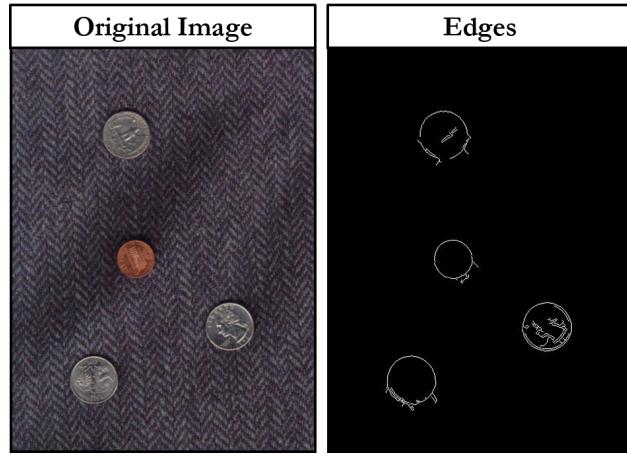


Figure 3.24: Summary of preprocessing operation performed by Hough Circle detector.

Insert Text Explaining the Basic Intuition Behind Hough Circle with Known Radius.

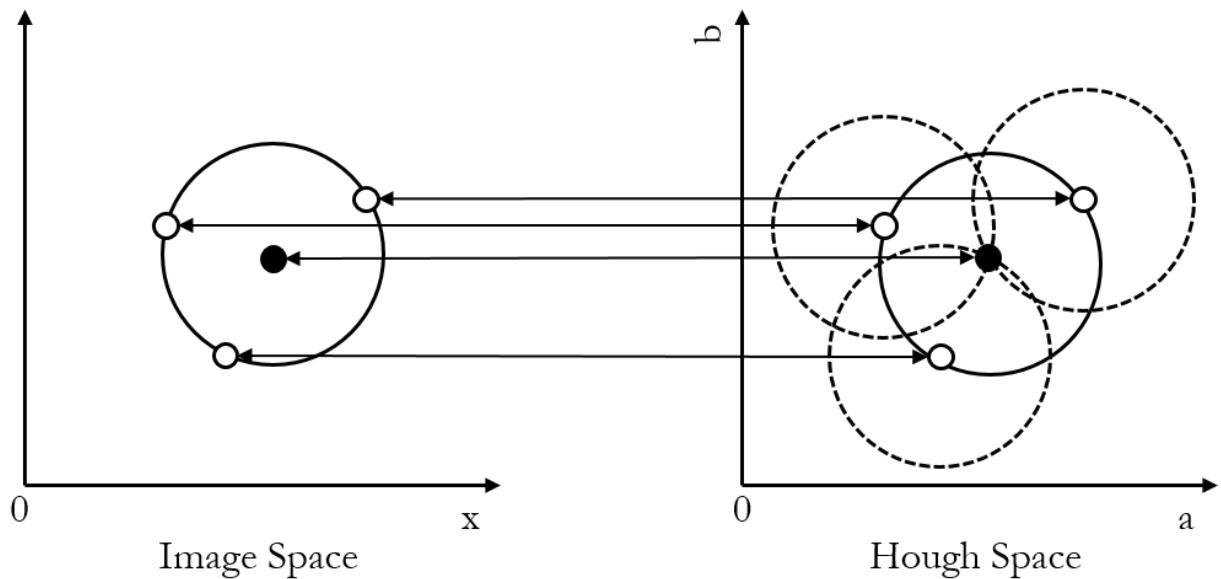


Figure 3.25: Summary of Hough Circle detector with known radius.

Insert Text Expanding Basic Intuition Behind Hough Circle with Unknown Radius.

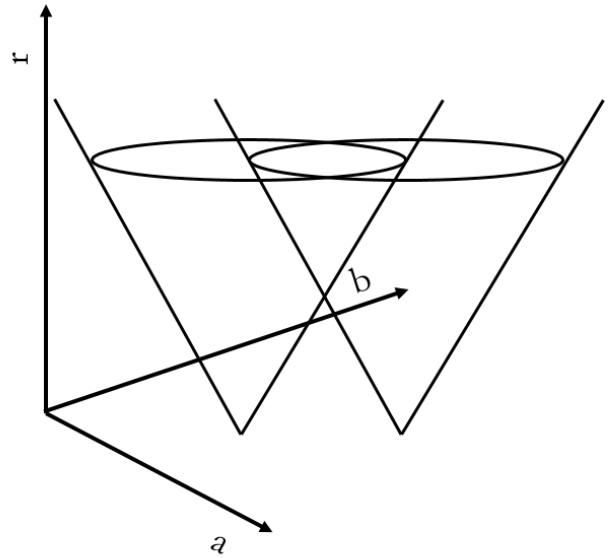


Figure 3.26: Summary of Hough Circle detector with unknown radius.

### 3.7.5 Good Feature To Track

Insert Text Explaining the Basic Intuition Behind Good Feature.

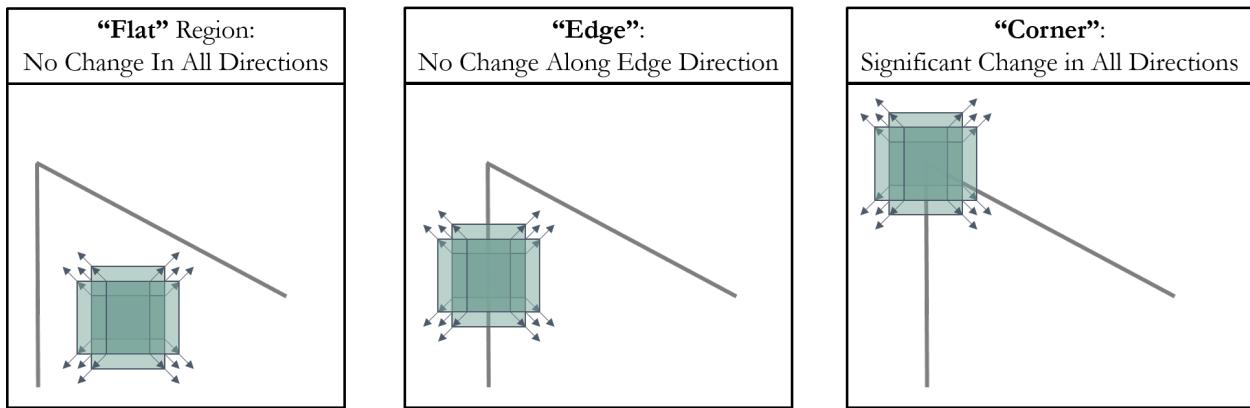


Figure 3.27: Summary of Good Feature detector.

Insert Text Explaining an Image Gradient/Derivative and Why We Look for Them.

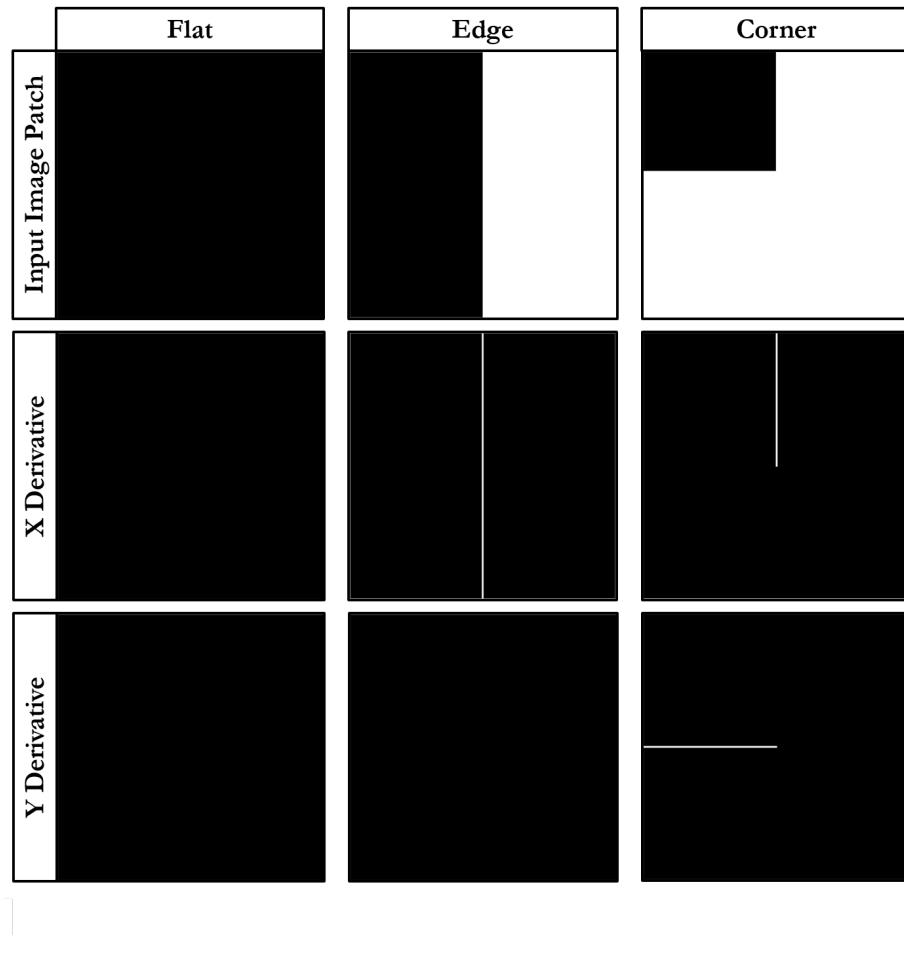


Figure 3.28: Image Gradient Example.

Insert Text Explaining Math.

$$E(u, v) = \sum_{x,y} w(x, y)[I(x + u, y + v) - I(x, y)]^2 \quad (3.6)$$

$$\sum_{x,y} [I(x + u, y + v) - I(x, y)]^2 \quad (3.7)$$

$$E(u, v) \approx \sum_{x,y} [I(x, y) + uI_x + vI_y - I(x, y)]^2 \quad (3.8)$$

$$E(u, v) \approx \sum_{x,y} u^2 I_x^2 + 2uv I_x I_y + v^2 I_y^2 \quad (3.9)$$

$$E(u, v) \approx [u \ v] \left( \sum_{x,y} w(x, y) \begin{bmatrix} I_x^2 & I_x I_y \\ I_x I_y & I_y^2 \end{bmatrix} \right) \begin{bmatrix} u \\ v \end{bmatrix} \quad (3.10)$$

$$E(u, v) \cong [u \ v] M \begin{bmatrix} u \\ v \end{bmatrix} \quad (3.11)$$

Where

$$M = \sum_{x,y} w(x, y) \begin{bmatrix} I_x^2 & I_x I_y \\ I_x I_y & I_y^2 \end{bmatrix}$$

Insert Text Explaining How Both Harris Corner and Good Feature Scoring Functions Work and Why Good Feature Performs Slightly Higher.

$$R = \det(M) - k(\text{trace}(M))^2 \quad (3.12)$$

Where  $\det(M) = \lambda_1 \lambda_2$  and  $\text{trace}(M) = \lambda_1 + \lambda_2$ . So,

$$R = \lambda_1 \lambda_2 - k(\lambda_1 + \lambda_2)^2 \quad (3.13)$$

$$R = \min(\lambda_1, \lambda_2) \quad (3.14)$$

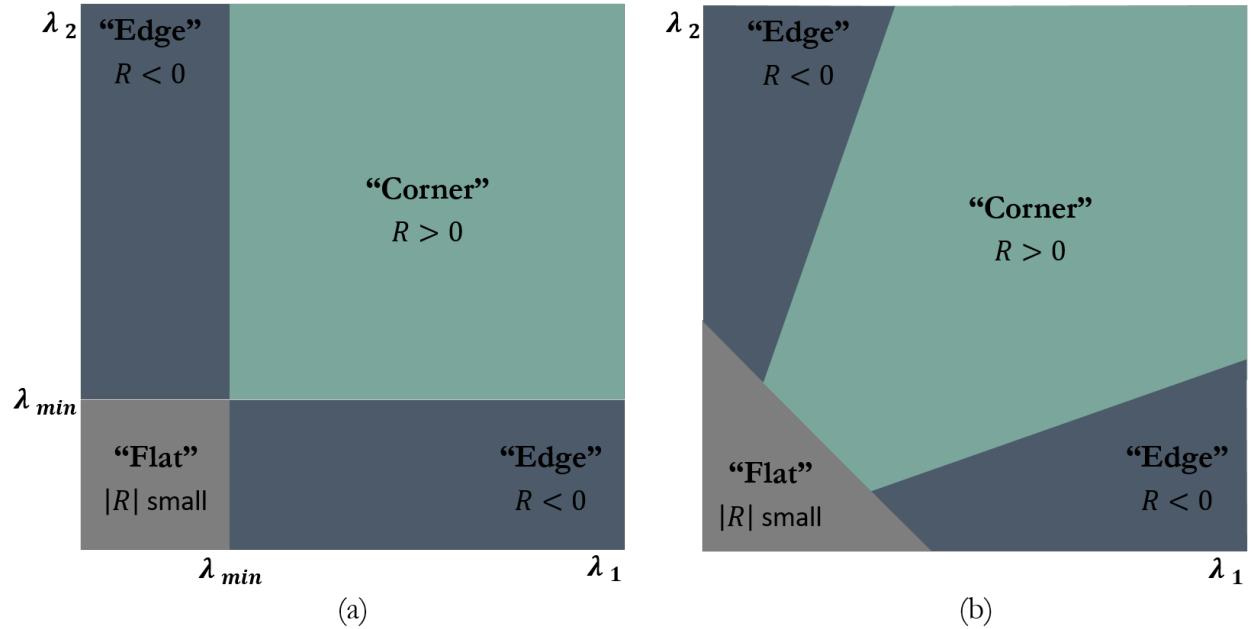


Figure 3.29: Difference Between Good Feature and Harris Corner Scoring Functions.

### 3.7.6 Optical Flow

Insert Text Explaining the Basic Intuition Behind Optical Flow.



Figure 3.30: Example of optical flow.

Insert Text Explaining the Basic Intuition Behind Kanade-Lucas-Tomasi Feature Tracker.

$$\sum_x [T(\mathbf{W}(\mathbf{x}; \Delta \mathbf{p})) - I(\mathbf{W}(\mathbf{x}; \mathbf{p}))]^2 \quad (3.15)$$

$$\sum_x \left[ T(\mathbf{W}(\mathbf{x}; \mathbf{0})) + \nabla T \frac{\partial \mathbf{W}}{\partial \mathbf{p}} \Delta \mathbf{p} - I(\mathbf{W}(\mathbf{x}; \mathbf{p})) \right]^2 \quad (3.16)$$

$$\Delta \mathbf{p} = H^{-1} \sum_x \left[ \nabla T \frac{\partial \mathbf{W}}{\partial \mathbf{p}} \right]^T [I(\mathbf{W}(\mathbf{x}; \mathbf{p})) - T(\mathbf{x})] \quad (3.17)$$

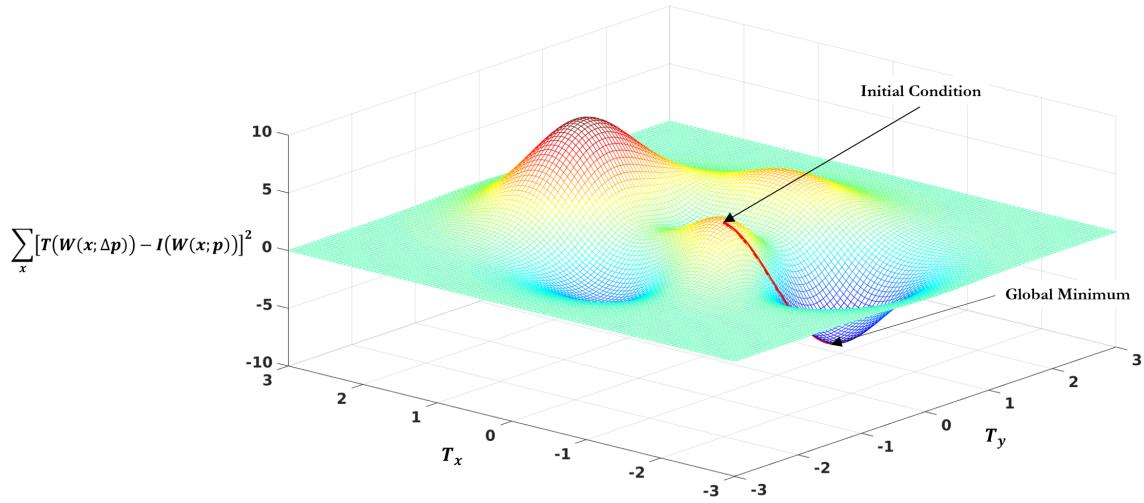


Figure 3.31: Summary of Kanade-Lucas-Tomasi feature tracker.

### 3.7.7 Pose Estimation and Tracking Through Augmented Reality Tag Detection

Insert Text Describing How AR Tag Detection, Pose Estimation, and Tracking Works with AR\_Track\_Alvar.

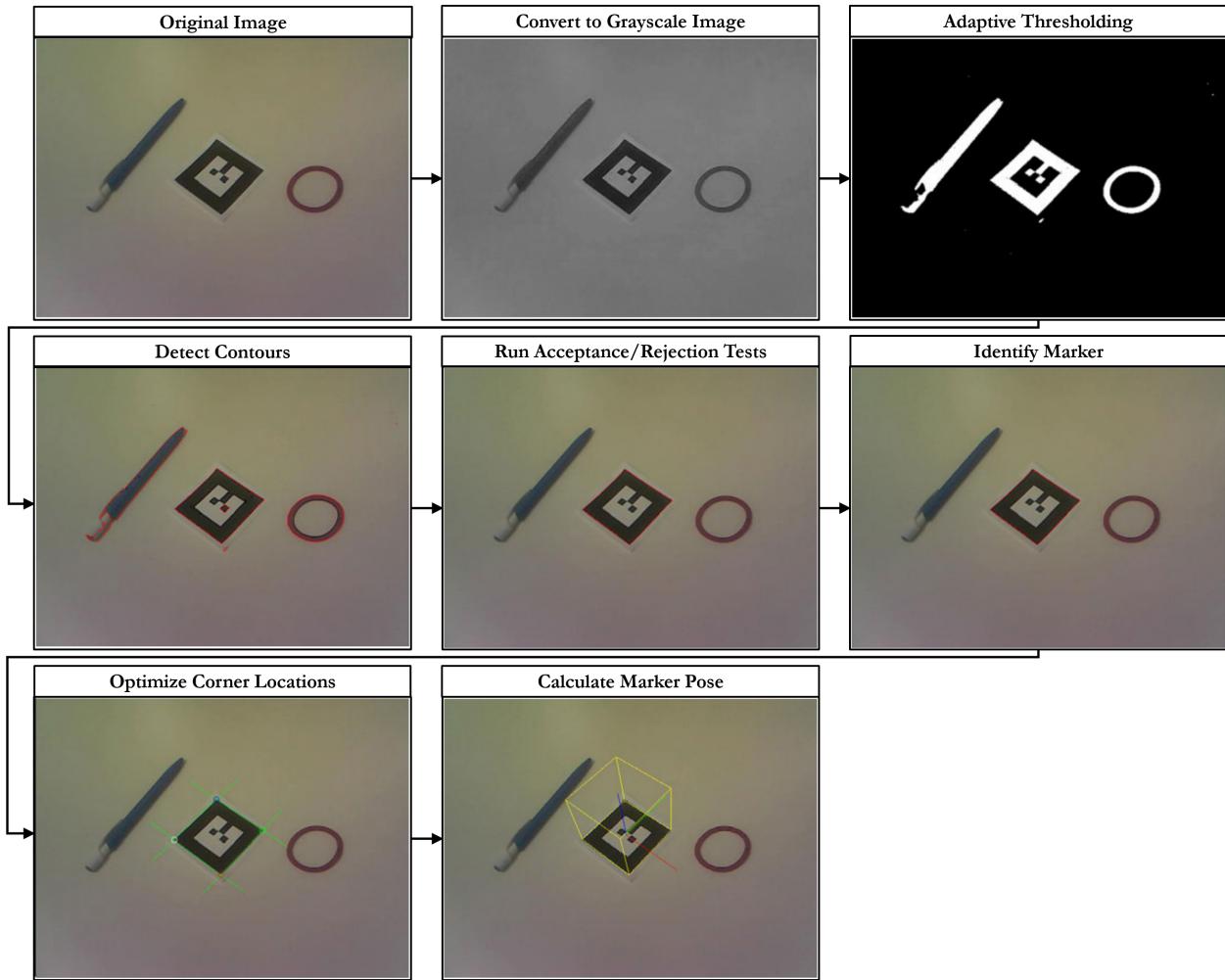


Figure 3.32: Summary of ALVAR AR Tag Detection Framework.

## 3.8 Summary

Insert Text Summarizing This Chapter and Transiting to the Next.

# Chapter 4

## Multistage Localization for High Precision Mobile Manipulation Tasks

Insert Text Outlining the Upcoming Sections.

### 4.1 Approach Overview

Insert Text Outlining the Steps of the Approach.

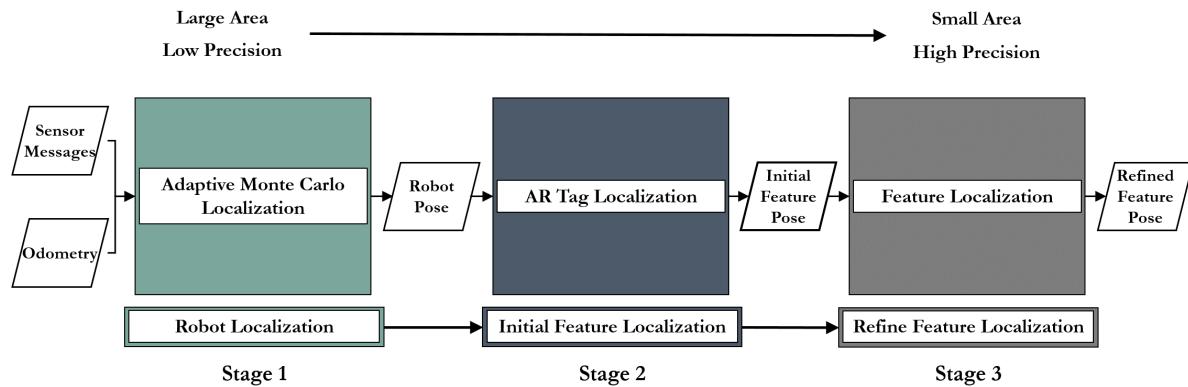


Figure 4.1: Multistage Localization Approach Overview.

## 4.2 Global Map Creation and Task Location Specification

The purpose of this section is reiterate why KartoSLAM was chosen, as well as describe how maps are created, the map files themselves, and how global locations are specified.

Before the manipulator can localize itself with respect to the part, the system must first navigate to the general vicinity in which the work will take place. In order to achieve this, the system utilizes odometry data, given by wheel encoders and an onboard inertial measurement unit (IMU), as well as sensor data, such as laser scans from a LIDAR or point clouds from an RGB-D sensor, to output safe velocity commands that will be sent to the mobile base of the system.

First, a Simultaneous Localization and Mapping (SLAM) technique named KartoSLAM uses the system's odometry and laser scan or point cloud data, to create a 2-D map of the environment in which the operation(s) will take place. After which, the global location of specific operation(s) are defined, as shown in Figure 4.2.

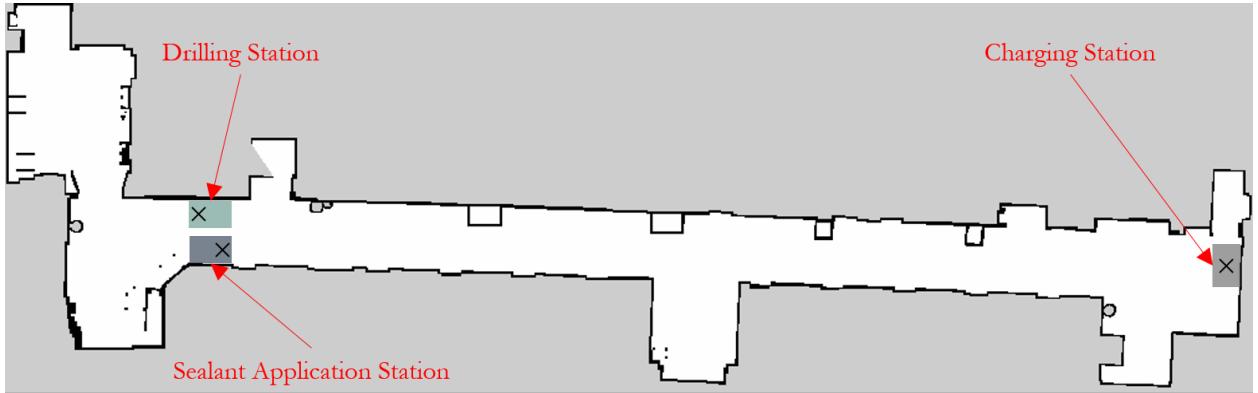


Figure 4.2: 2-D map environment with specified start locations.

## 4.3 Autonomous Localization and Navigation

Adaptive Monte Carlo localization (amcl) is used to localize the system within the map. Subsequently, odometry data is combined with a global and local cost map, in which obstacles and a specific distance around them represent a cost. These maps are used to plan optimal and obstacle free paths through the environment. The global path is computed before the system begins moving and takes into account all known obstacles, while the local path monitors incoming sensor data to compute suitable linear and angular velocities for the system to complete the current section of the global path. The local path is typically

computed at a rate of 20 Hz; however, this parameter is adjustable given the needs of the system.

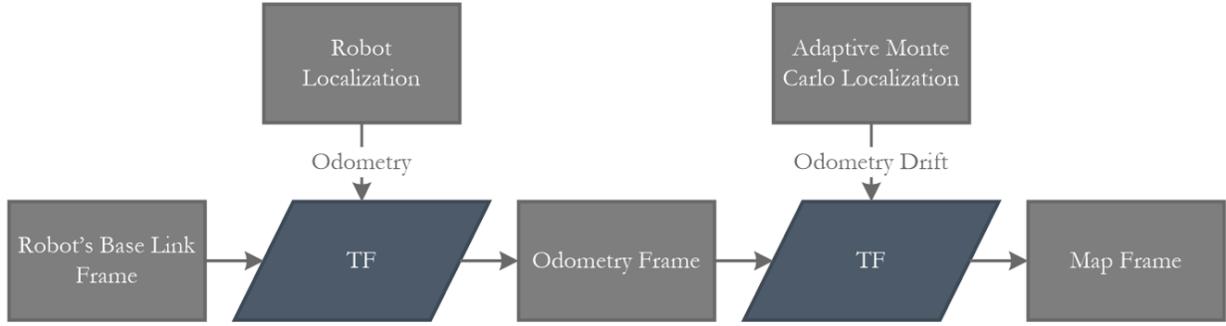


Figure 4.3: Robot Localization Framework.

## 4.4 Task Association and A Priori Knowledge

Once arriving in the general vicinity of the task to be accomplished, the system then locates an augmented reality (AR) tag [5], which allows it to localize and transform points of interest (POIs) associated with the AR tag into the system's frame of reference. This general localization framework is presented in Figure ??.



Figure 4.4: Initial Feature Localization Framework.

## 4.5 Generic Framework for Multi-Stage Computer Vision Algorithm

Insert Text Outlining the Upcoming Sections.

### 4.5.1 Initial Feature Location Prediction

This subsection will go through AR tags and how they are used to populate point.

#### 4.5.2 Corrected Feature Locations

Currently the sponsor of this work uses a hole template on the object to be drilled to ensure accuracy within  $\pm 0.3$  mm. Using the predicted hole locations, given by the AR tag, an inverse kinematic solver is used to move the manipulator to the specified Cartesian location. A camera mounted on the manipulator is then used to further correct the position of the end-effector. Canny Edge Detection, Hough Transforms, as well as the camera's intrinsic characteristics and a priori knowledge of each hole's size is used to output an adjusted Cartesian location of the circle on the templet closest to the predicted position [6]. If a hole to be drilled is not found or is outside the range of the manipulator, that hole will be added to a list and the customer notified of all such holes after the operation is completed. In addition to the high precision achieved by the above technique, it can also provide a video log of all work done for inspection.

Insert Text Regarding Computer Vision Detection and Tracking System.

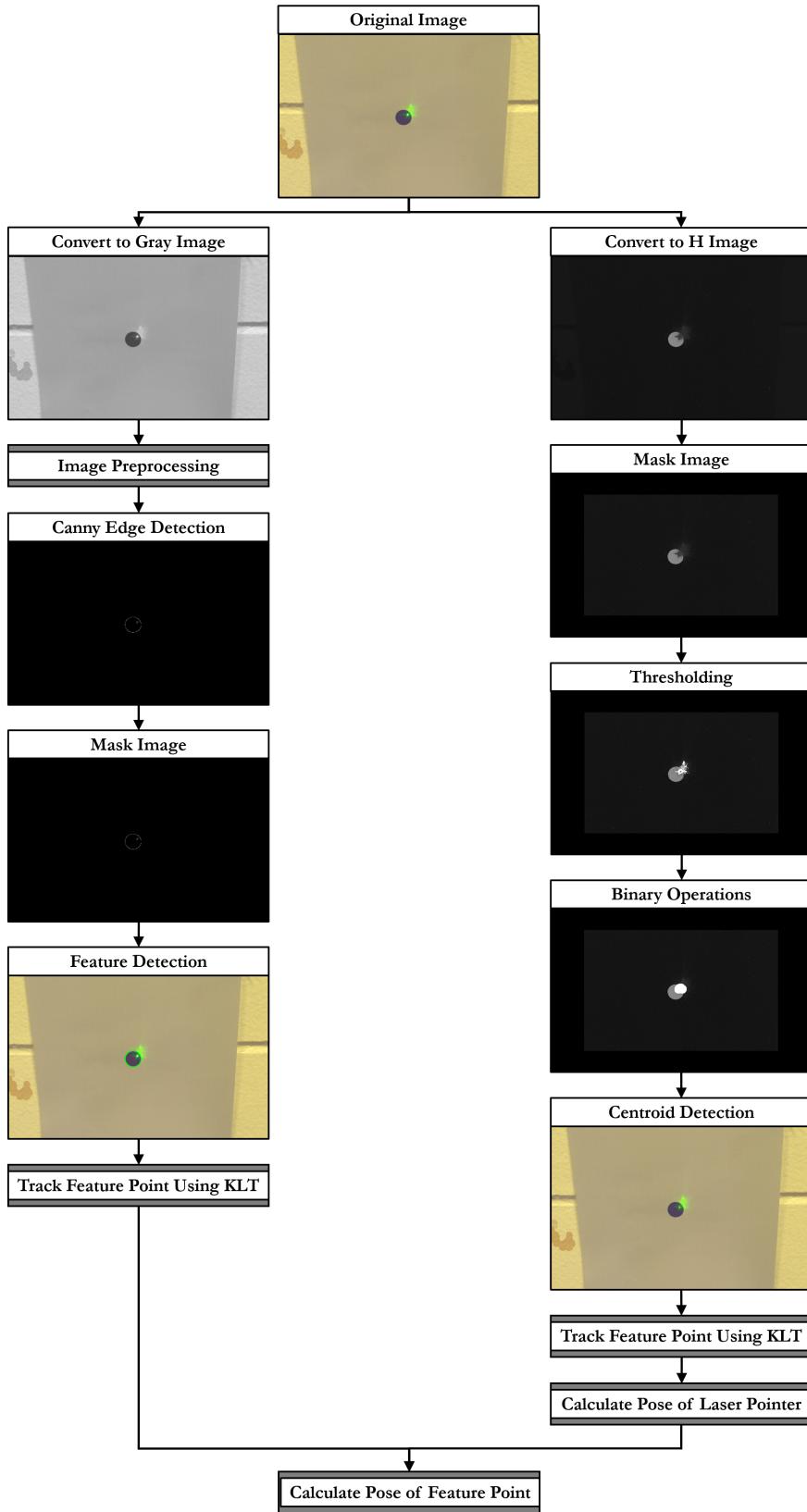


Figure 4.5: Feature Detection and Tracking Pipeline.

Insert Text Regarding Control Loop.

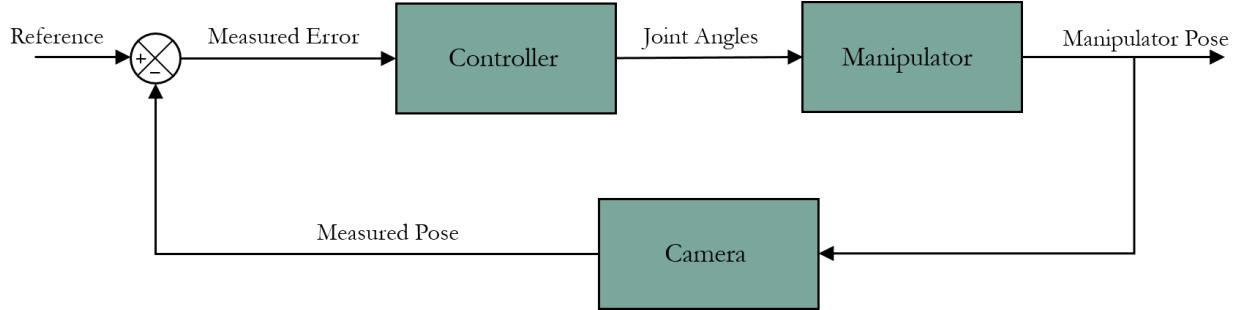


Figure 4.6: Manipulator Correction Control Loop.

## 4.6 Approach Implementation

Outline the upcoming sections.

### 4.6.1 System Overview

The generic framework of the system is depicted in Figure ???. The system allows the user to input predefined tasks. Given a priori knowledge of each tasks and their global start locations, the mobile system navigates to and performs the requested operation(s) while monitoring its battery level to ensure mission completion. The framework shown was implemented using a hierarchical state machine and was written in such a way as to make it compatible with the Robot Operating System (ROS) to ensure ease of use across differing robotic platforms.

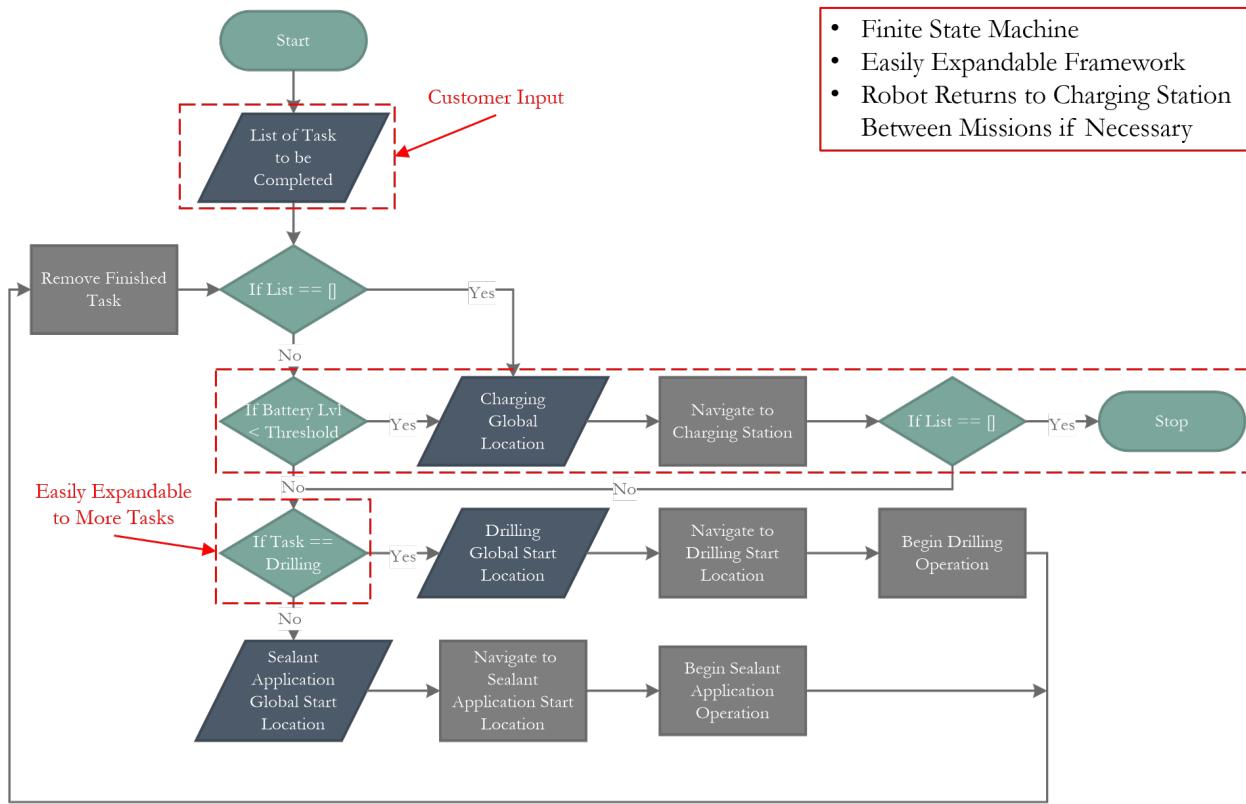


Figure 4.7: General system overview.

#### 4.6.2 Drilling Framework

This section will go in depth into how the robot simulates drilling.

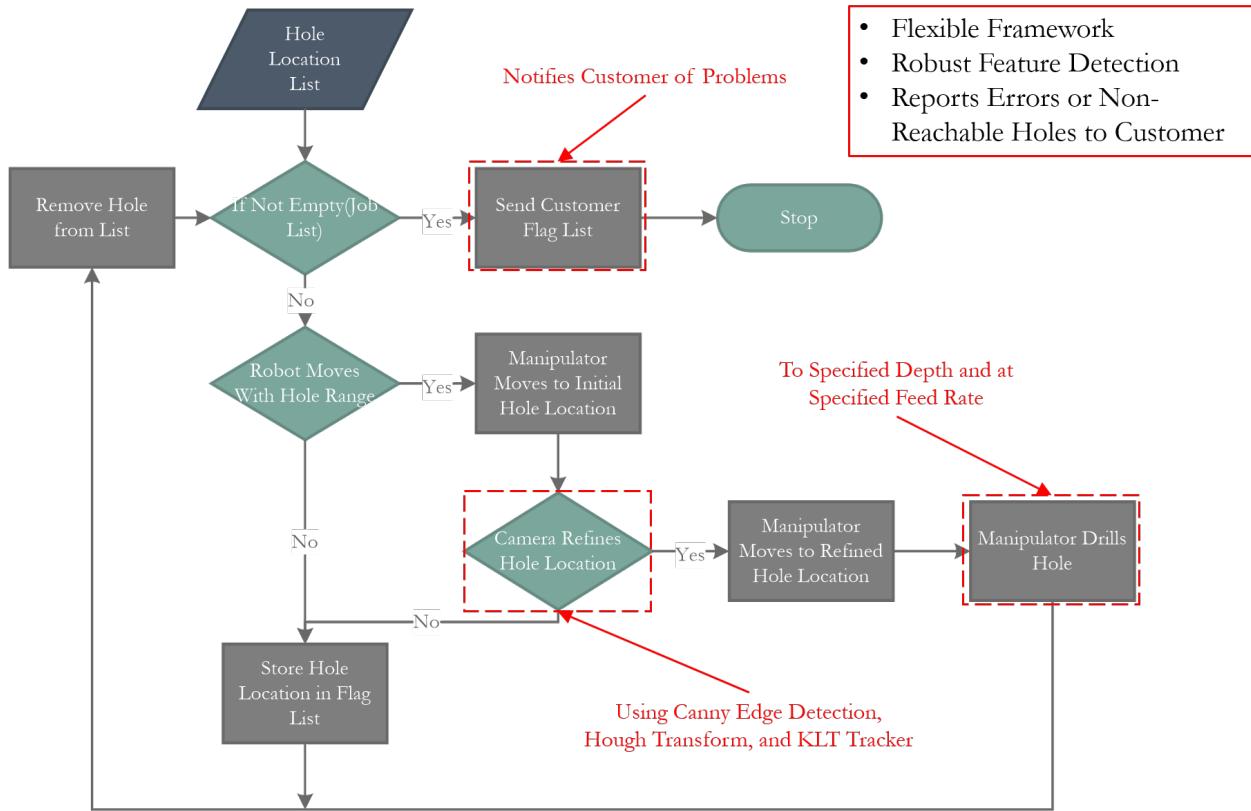


Figure 4.8: Drilling operation framework.

#### 4.6.3 Sealant Application Framework

This section will go in depth into how the robot simulates sealant application.

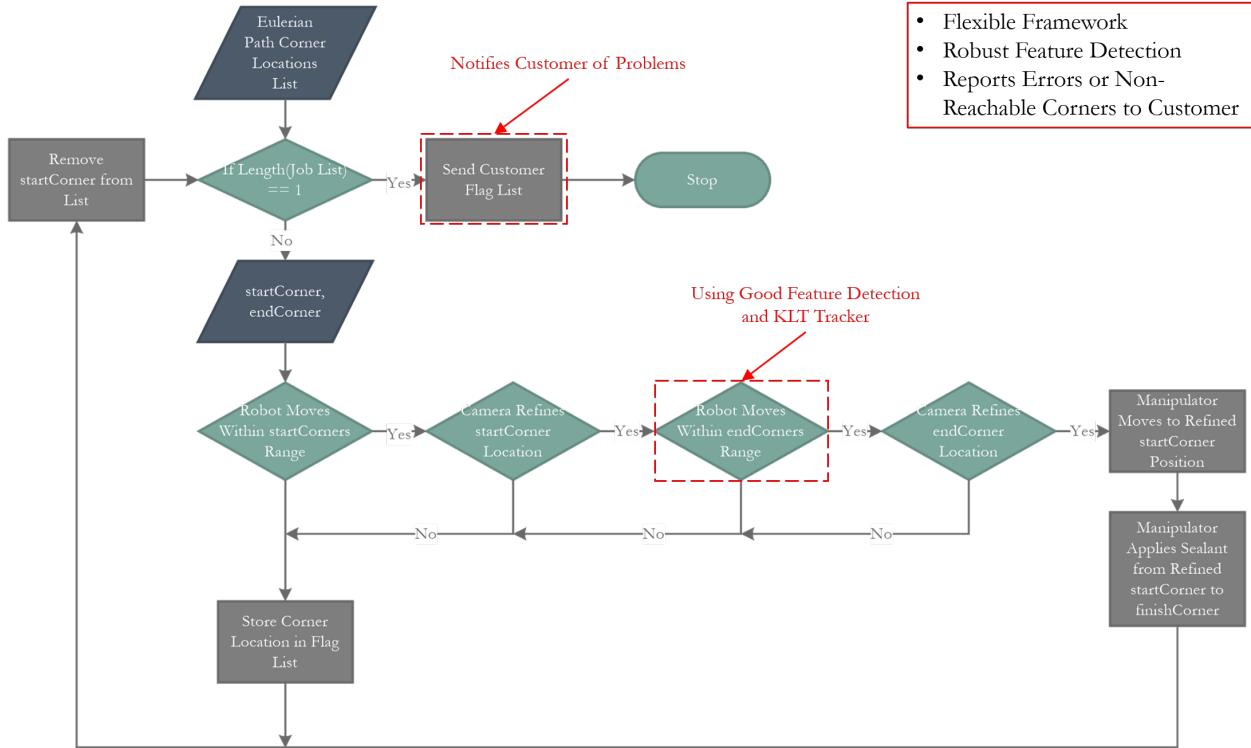


Figure 4.9: Sealant Application Framework.

## 4.7 Summary

Insert Text Summarizing This Chapter and Transiting to the Next.

# **Chapter 5**

## **Experiments and Results**

Outline the upcoming sections.

### **5.1 Hardware Architecture**

Insert Text Regarding Robot, Sensors, etc.

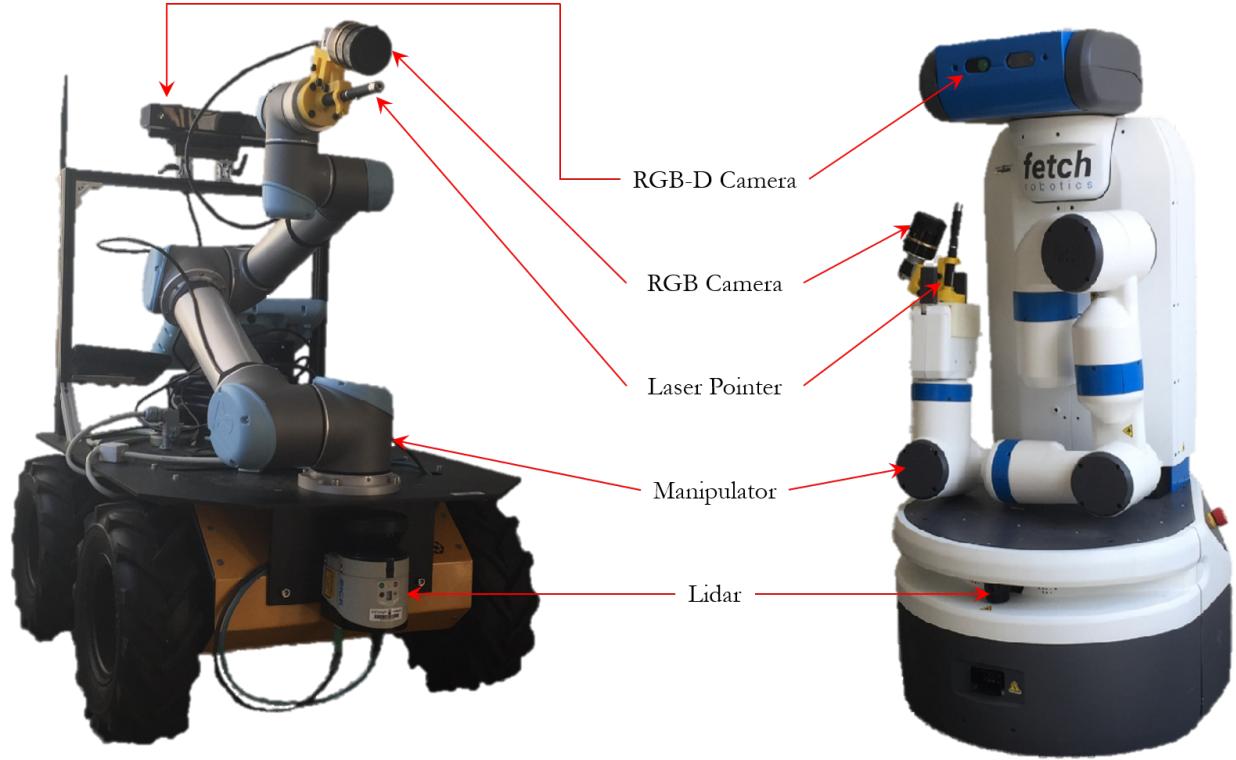


Figure 5.1: Clearpath Robotics' Husky and Fetch Robotics' Fetch Mobile Manipulator

## 5.2 Software Architecture

Insert Text Regarding The Software Architecture.

## 5.3 Experiments

Insert Text Regarding Upcoming Sections.

### 5.3.1 Camera Calibration Setup

Insert text setup for camera calibration.



Figure 5.2: Camera Calibration Setup.

Insert text regarding methods used for camera calibration (citing paper)

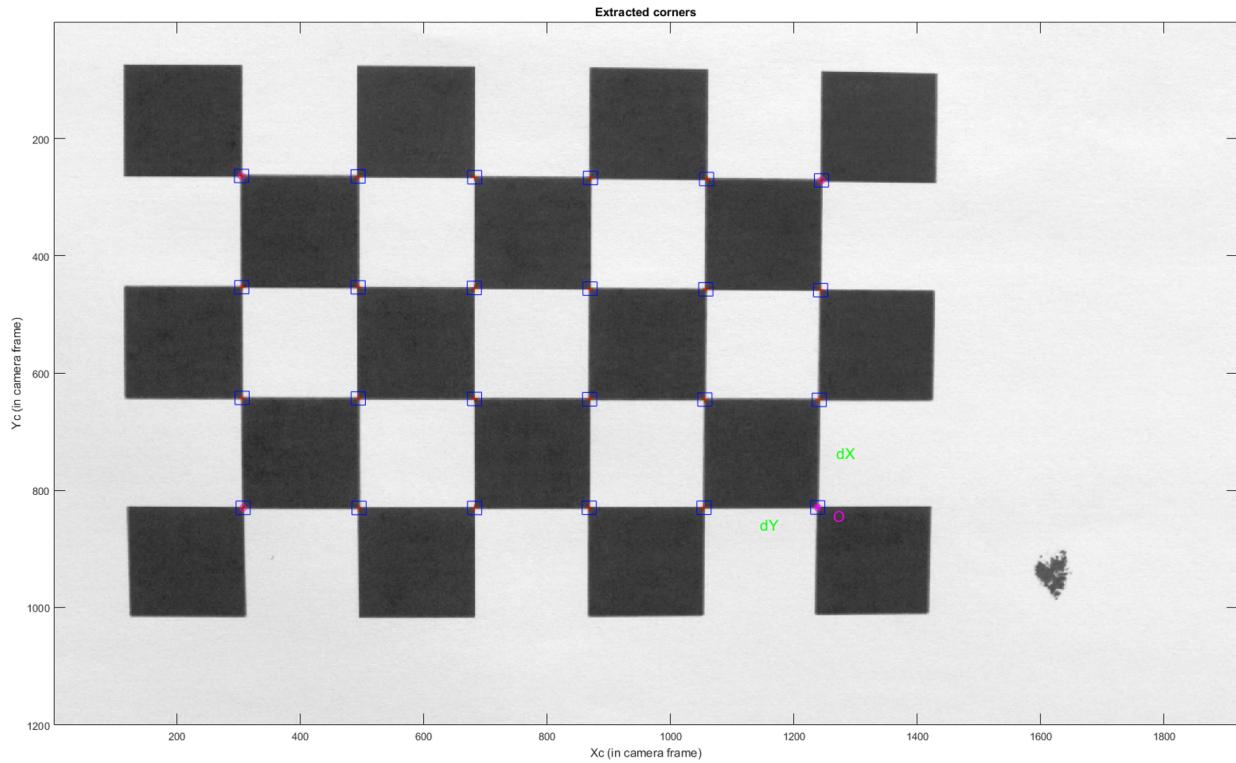


Figure 5.3: Extracted Corners of Calibration Pattern.

Insert Text and Math Regarding Calibration of Laser.

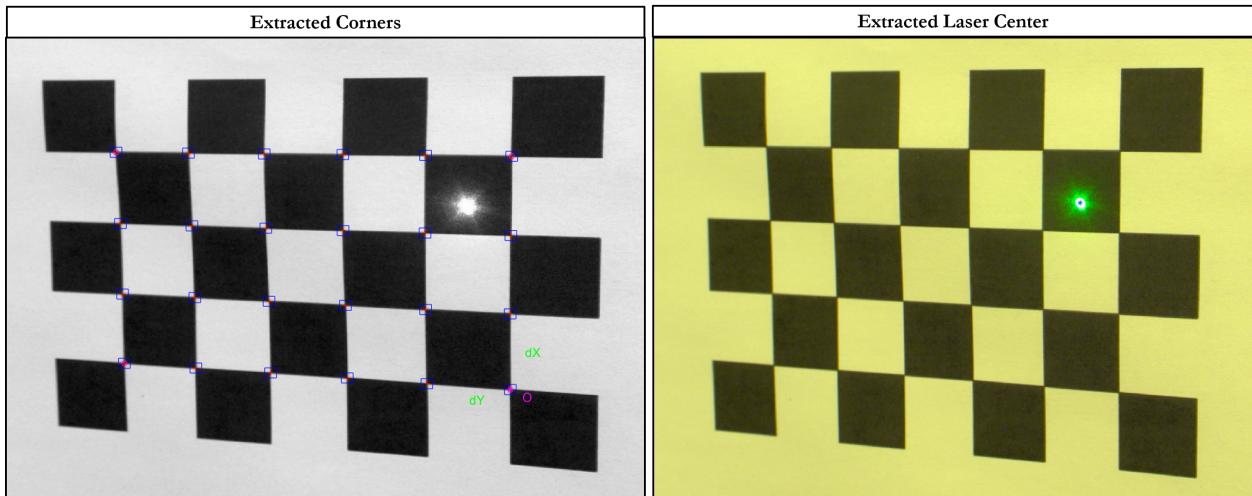


Figure 5.4: Extracted Corners and Laser Center

### 5.3.2 Navigation System Experimental Setup

Insert Text Regarding the Navigation System Experimental Setup.

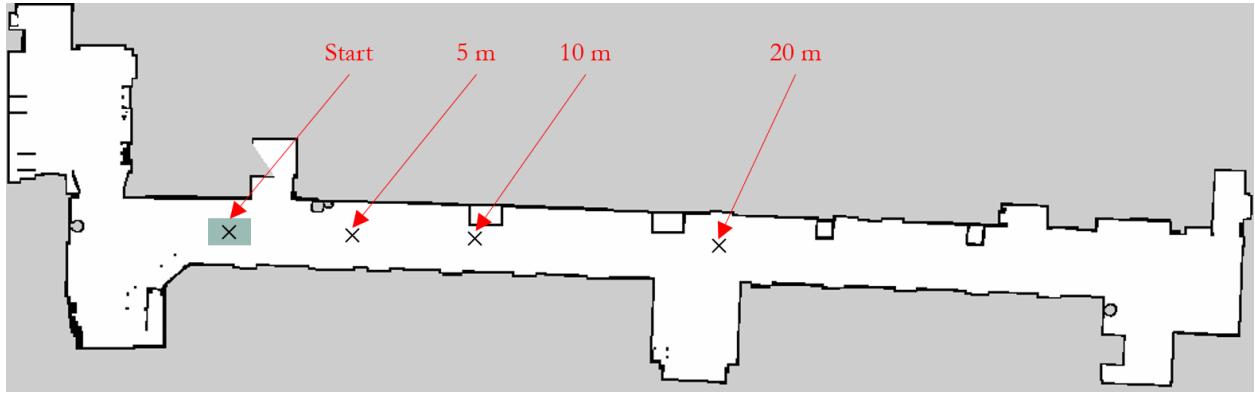


Figure 5.5: Navigation System Experimental Setup.

### 5.3.3 Drilling Operation Experimental Setup

Insert Text Regarding the Drilling Operations Experimental Setup.

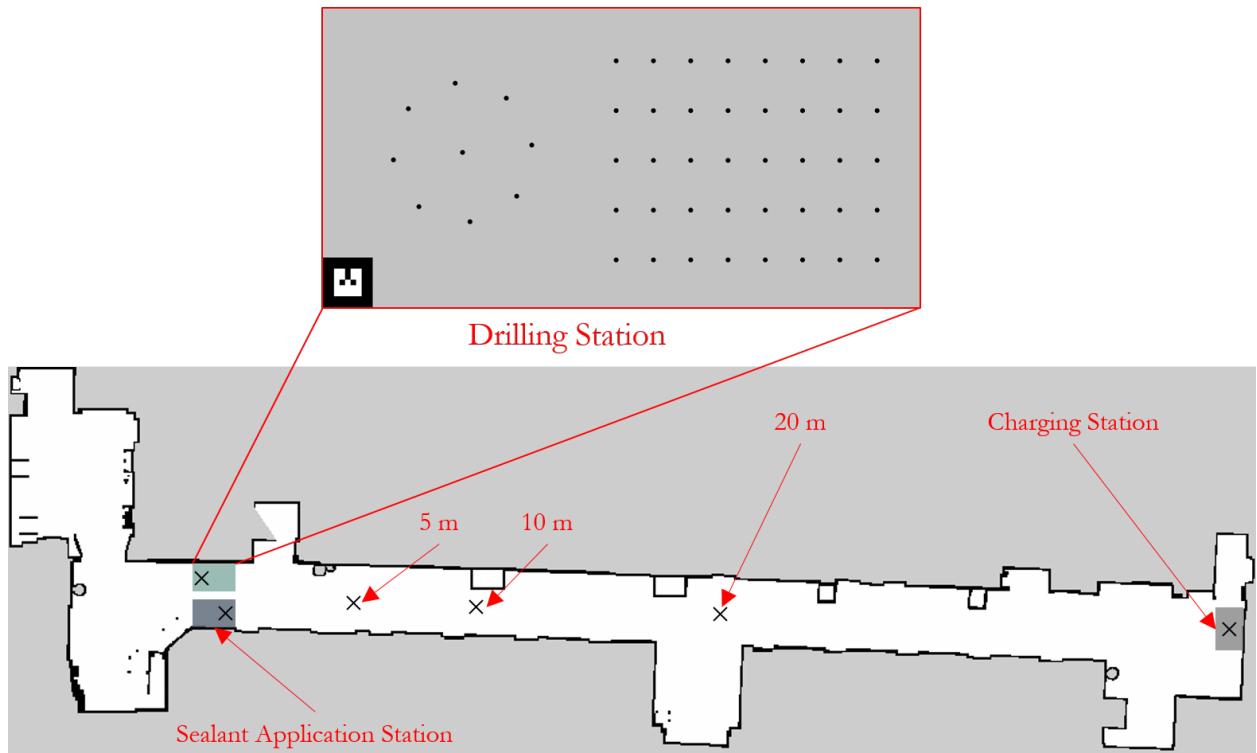


Figure 5.6: Drilling Operations Experimental Setup.

### 5.3.4 Sealant Application Experimental Setup

Insert Text Regarding the Sealant Application Experimental Setup.

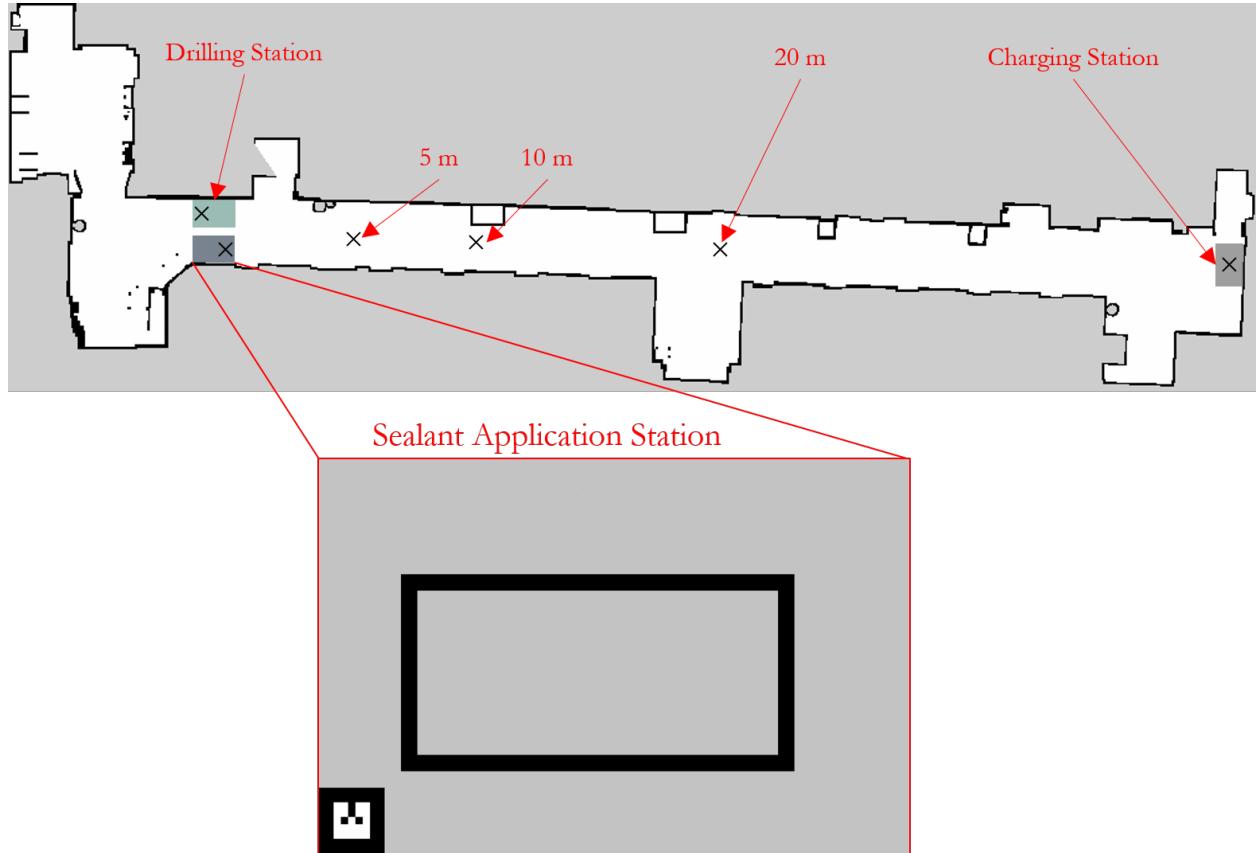


Figure 5.7: Sealant Application Experimental Setup.

## 5.4 Results

### 5.4.1 Camera Calibration Accuracy Achieved

Insert Text Regarding Pixel Error of Kinect V2.

Insert Text Regarding Pixel Error of Basler ACA150-UC.

Insert Text Regarding Pixel Error of Laser Calibration.

### 5.4.2 Navigation System Accuracy Achieved

Insert Text Regarding Accuracy of Navigation System and Ways To Improve Accuracy.

Table 5.1: Global Localization Accuracy Given Distance From Start - Clearpath Robotics' Husky.

	x (cm)	x std (cm)	y (cm)	y std (cm)	$\psi$ (rad)	$\psi$ std (rad)
5 m	6.29	4.20	4.50	3.58	0.03	0.02
10 m	4.16	3.16	2.88	2.97	0.03	0.03
20 m	6.35	4.55	2.37	1.89	0.02	0.01

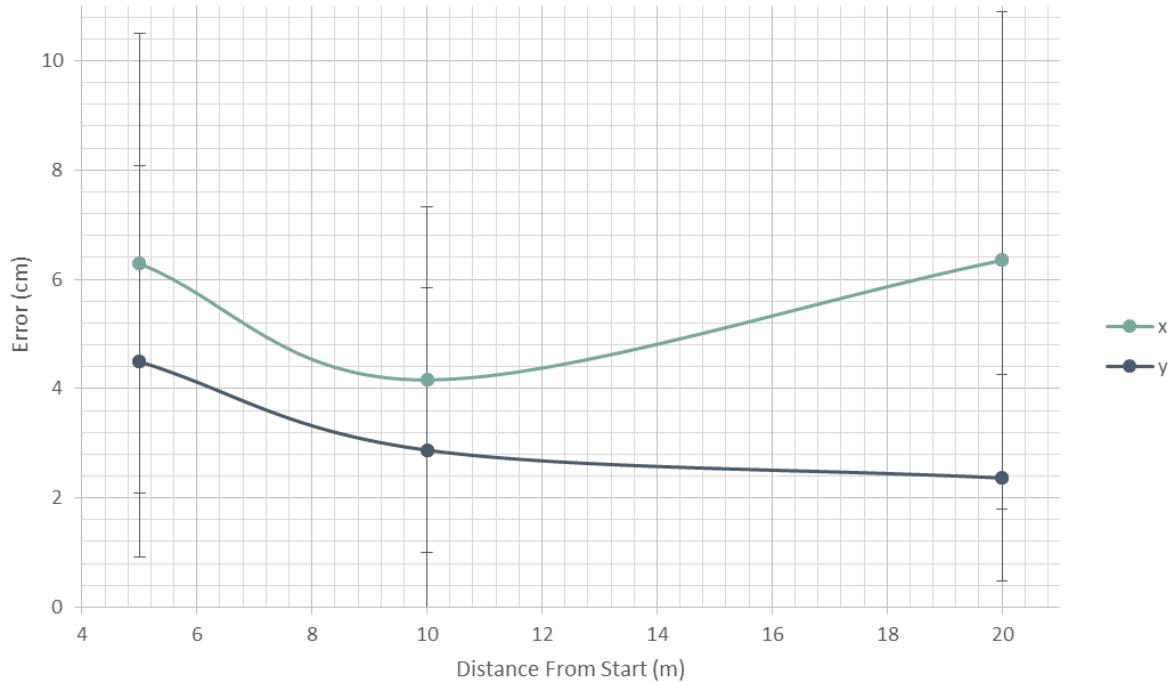


Figure 5.8: Global X and Y Localization Accuracy Given Distance From Start - Clearpath Robotics' Husky.

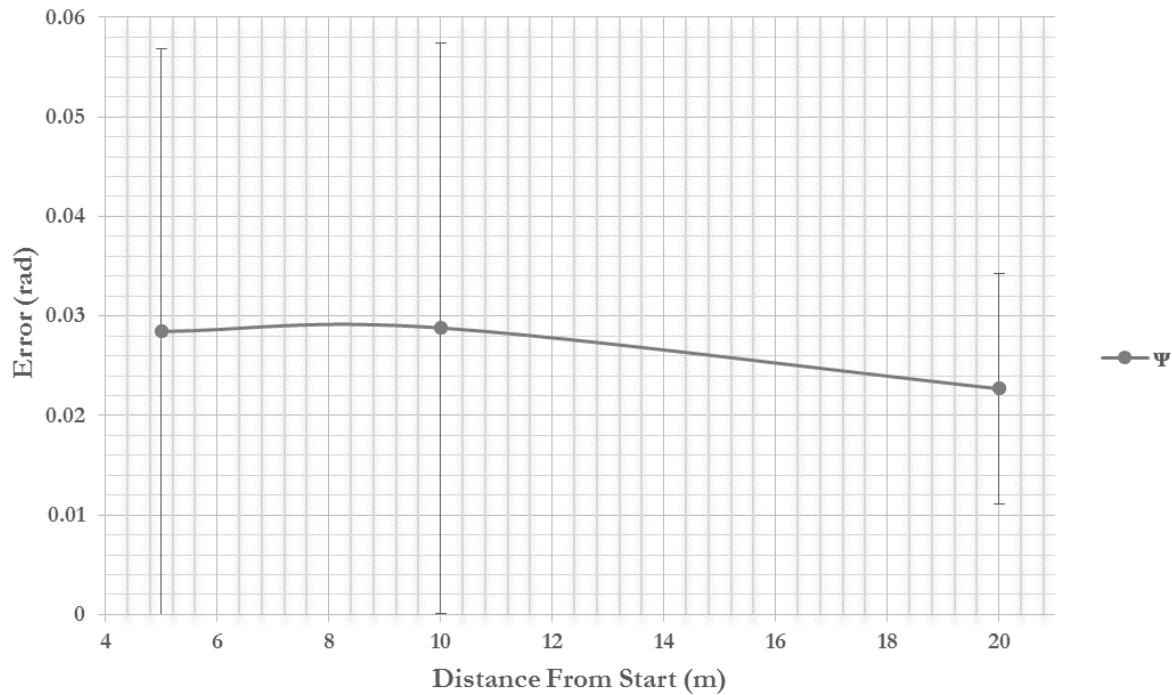


Figure 5.9: Global  $\psi$  Localization Accuracy Given Distance From Start - Clearpath Robotics' Husky.

Table 5.2: Global Localization Accuracy Given Distance From Start - Fetch Robotics' Fetch.

	x (cm)	x std (cm)	y (cm)	y std (cm)	$\psi$ (rad)	$\psi$ std (rad)
5 m	4.25	2.50	1.48	1.23	0.03	0.02
10 m	3.93	2.62	2.57	1.66	0.03	0.02
20 m	5.82	3.43	1.52	1.11	0.03	0.02

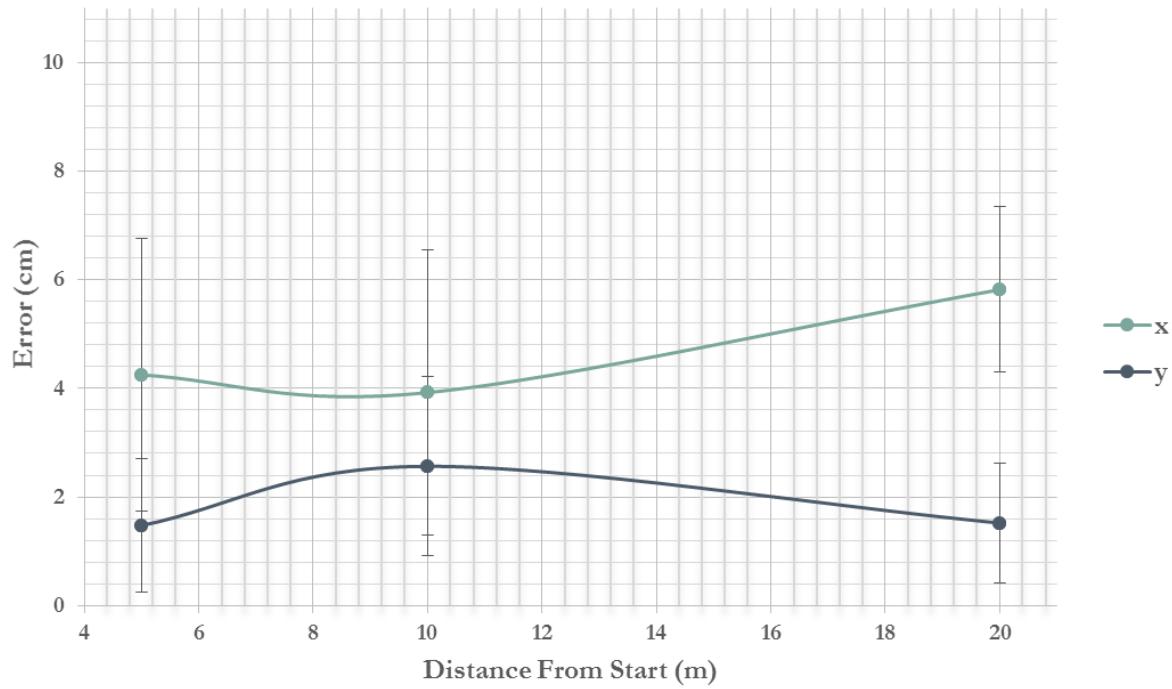


Figure 5.10: Global X and Y Localization Accuracy Given Distance From Start - Fetch Robotics' Fetch.

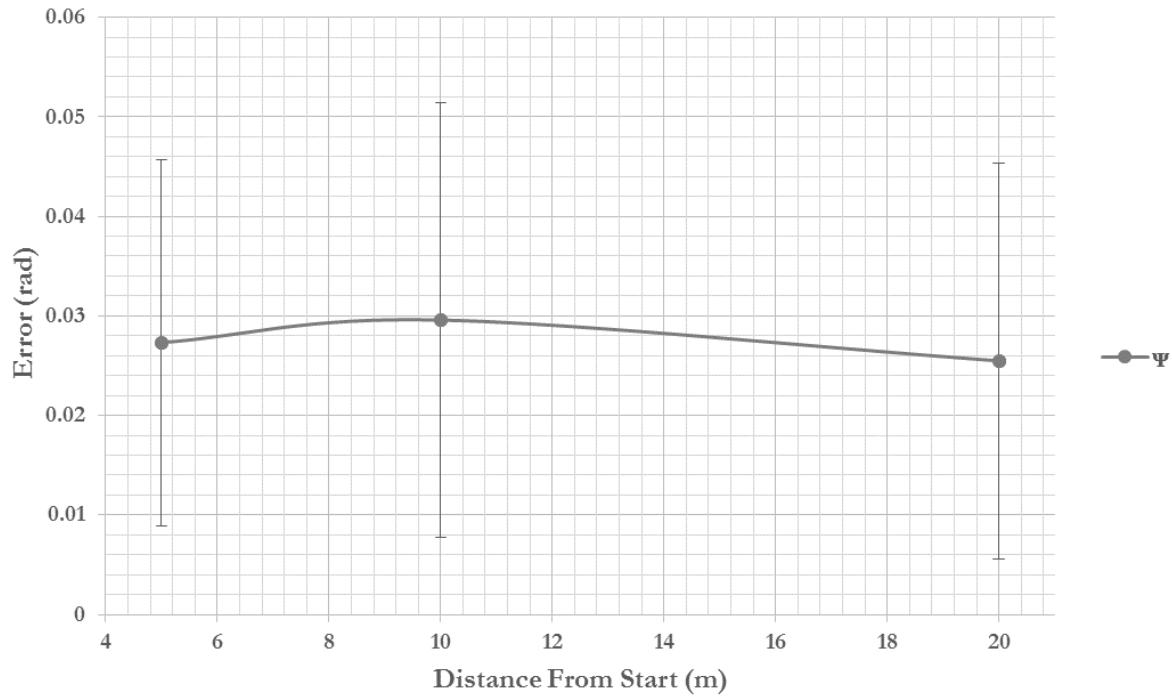


Figure 5.11: Global  $\psi$  Localization Accuracy Given Distance From Start - Fetch Robotics' Fetch.

### 5.4.3 Augmented Reality Tag Accuracy Achieved

Table 5.3: Augmented Reality Accuracy Given Specific Start Conditions - Primesense Carmine 1.09.

	x	x std (mm)	y	y std (mm)	z	z std (mm)
1	50.41	22.81	148.87	12.33	150.96	12.61
2	23.71	16.06	151.82	6.86	117.57	12.81
3	23.06	14.85	128.39	11.48	173.24	23.24
4	37.19	14.85	183.21	7.30	115.85	16.27
5	24.36	17.47	138.66	6.77	42.80	16.37
6	39.16	23.52	7.27	3.70	105.24	9.16

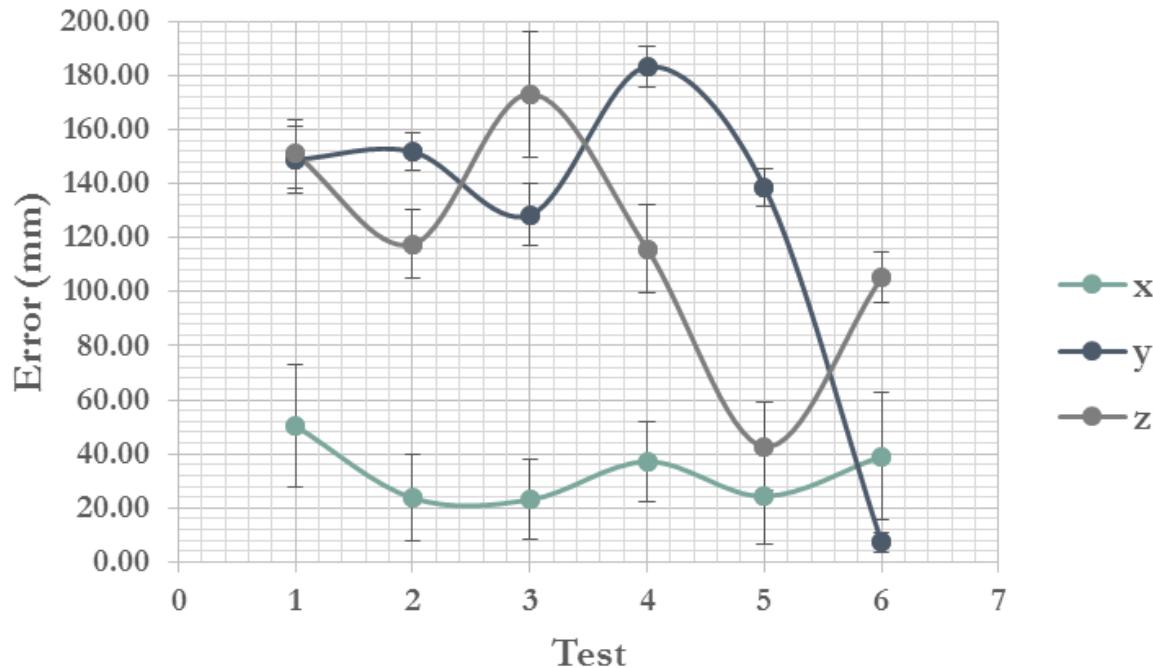


Figure 5.12: Augmented Reality Accuracy Given Specific Start Conditions - Primesense Carmine 1.09.

Table 5.4: Augmented Reality Accuracy Given Specific Start Conditions - Microsoft Kinect V2.

	x (mm)	x std (mm)	y (mm)	y std (mm)	z (mm)	z std (mm)
Test 1	14.32	9.75	4.04	1.65	84.77	2.50
Test 2	163.85	57.58	36.32	9.65	25.31	12.46
Test 3	84.33	44.37	50.31	5.48	55.24	8.57

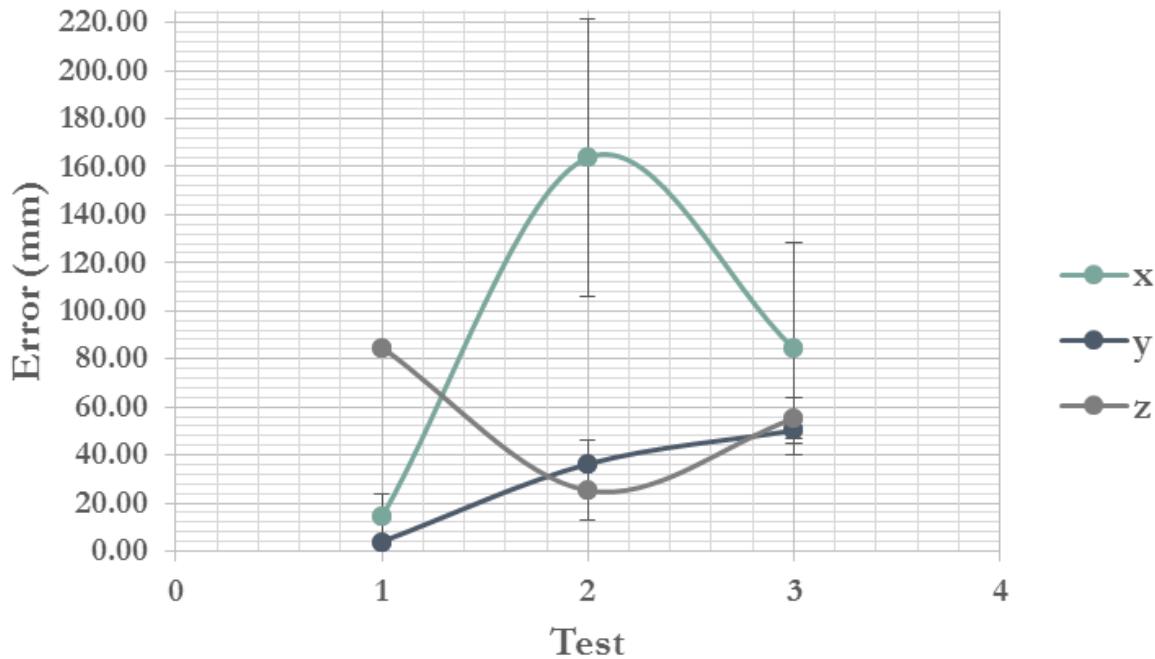


Figure 5.13: Augmented Reality Accuracy Given Specific Start Conditions - Microsoft Kinect V2.

#### 5.4.4 Drilling Operation Accuracy Achieved

Insert Text Regarding Accuracy of Drilling Operations and Ways To Improve Accuracy.

Table 5.5: Drilling Operation Accuracy Given Distance From Work Surface.

	x (mm)	x std (mm)	y (mm)	y std (mm)	z (mm)	z std (mm)
Distance Less Than 5 mm from Hole Center	13.25	1.43	0.70	0.43	0.82	0.56
Distance Between 5 and 10 mm from Hole Center	13.62	2.71	0.66	0.39	1.89	0.46
Distance Between 10 and 20 mm from Hole Center	10.78	2.55	0.53	0.30	4.06	0.70
Distance Between 20 and 30 mm from Hole Center	10.55	2.48	0.31	0.21	10.25	1.38
Mean	12.05	2.55	0.55	0.38	4.25	3.75

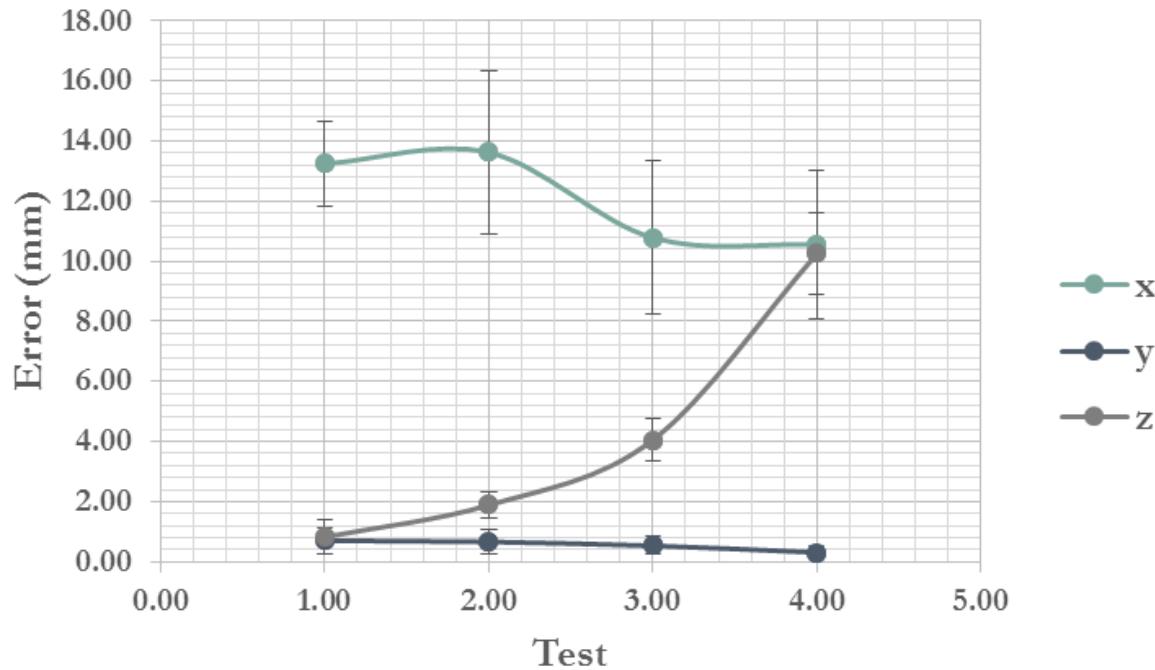


Figure 5.14: Drilling Operation Accuracy Given Distance From Feature Point.

#### 5.4.5 Sealant Application Accuracy Achieved

Insert Text Regarding Accuracy of Sealant Application and Ways To Improve Accuracy.

Table 5.6: Sealant Application Accuracy Given Distance From Work Surface.

	x (mm)	x std (mm)	y (mm)	y std (mm)	z (mm)	z std (mm)
Distance Less Than 5 mm from Corner	10.90	2.44	1.46	0.21	1.17	0.71
Distance Between 5 and 10 mm from Corner	11.95	3.21	1.27	0.22	1.48	0.91
Distance Between 10 and 20 mm from Corner	16.32	2.56	0.78	0.34	3.81	0.32
Distance Between 20 and 30 mm from Corner	13.92	3.34	0.79	0.60	8.21	0.40
Mean	13.27	3.33	1.08	0.48	3.67	2.88

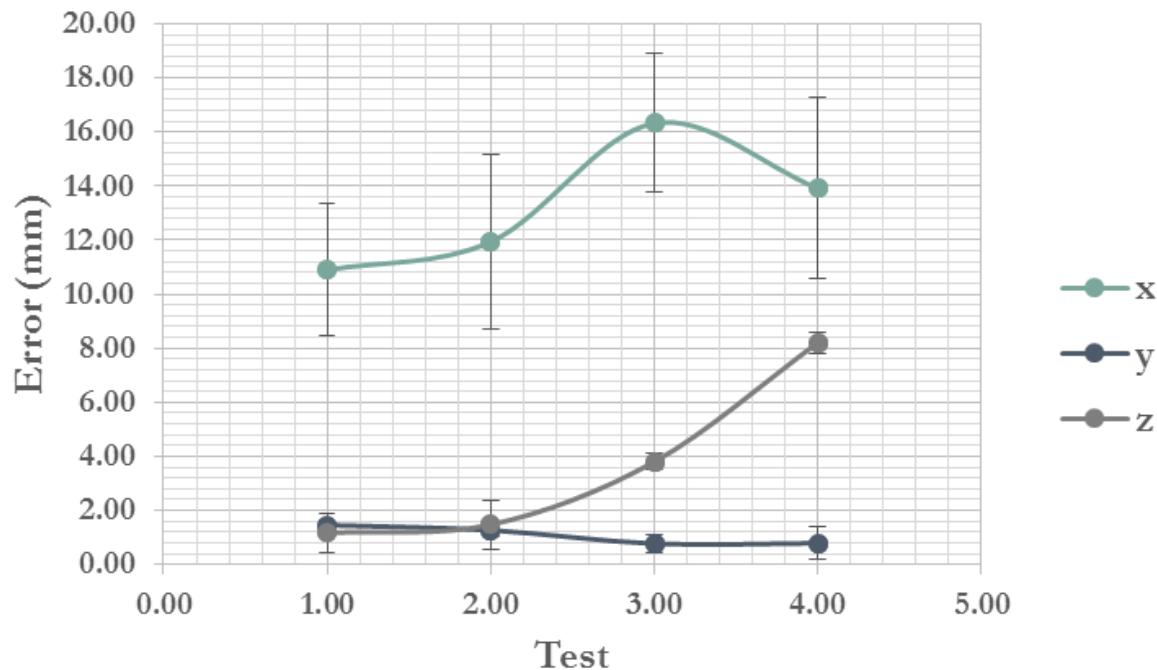


Figure 5.15: Sealant Application Accuracy Given Distance From Feature Point.

## 5.5 Summary

Insert Text Summarizing This Chapter and Transiting to the Next.

# Chapter 6

## Conclusion and Future Work

This paper showed the general framework necessary in order to solve the two main problems preventing the development and widespread adoption of ARC. These problems have caused a decrease in productivity and increase in workplace injuries/fatalities over the past several decades in comparison to other industries [2]. These problems include the fact that typical construction sites tend to be unstructured and are continuously evolving versus the highly controlled environments found in manufacturing. Also, the relationship between the part and manipulator has been reversed, causing increased complexity not seen in manufacturing environments where the part arrives at a fixed manipulator [3]. The techniques presented allow systems to create a 2-D map of their environment, localize themselves and complete the task(s) assigned. After localizing an AR tag at the work site, the system is able to use a priori knowledge to localize POIs and complete a plethora of operations achieving an accuracy of approximately  $\pm 2$  mm based on a multifaceted computer vision approach with only a USB webcam.

Future work to be explored includes increasing the accuracy of the computer vision system to the sub-millimeter levels through the use of a machine vision camera, as well as a relatively new calibration technique developed by Feng et al [3]. In addition, automated approaches to create 3-D maps are being looked into in order to provide updated data about the robot's surroundings and task(s) automatically without human intervention. Also, human and robot collaboration over a distributed network is being explored.

# Bibliography

- [1] C Balaguer. "NOWADAYS TRENDS IN ROBOTICS AND AUTOMATION IN CONSTRUCTION INDUSTRY: TRANSITION FROM HARD TO SOFT ROBOTICS". In: ().
- [2] Eddy M Rojas, Peerapong Aramvareekul, and Chris Hendrickson. "DISCUSSION AND CLOSURES Discussion of " Is Construction Labor Productivity Really Declining? """. In: 129129.1 (2003). doi: 10.1061/().
- [3] Chen Feng et al. "Vision guided autonomous robotic assembly and as-built scanning on unstructured construction sites". In: *Automation in Construction* (2015). ISSN: 09265805. doi: 10.1016/j.autcon.2015.06.002.
- [4] Sebastian Thrun, Wolfram Burgard, and Dieter Fox. *Probabilistic robotics*. Intelligent robotics and autonomous agents. Cambridge (Mass.) (London): the MIT Press, 2005. ISBN: 0-262-20162-3. URL: <http://opac.inria.fr/record=b1102986>.
- [5] Sanni Siltanen. "Theory and applications of marker-based augmented reality". In: ().
- [6] T D Alter. "3D Pose from 3 Corresponding Points under Weak-Perspective Projection". In: (1992).