# Real Time Appearance Based Mapping SLAM Overview & Simulation in ROS

Karim Hassanieh

**Abstract**—In this paper we present a theoretical overview of generic Grid Based SLAM with particular focus and attention on Real Time Appearance Based SLAM - RTAB SLAM. Moreover we will present the advantages and contributions of RTAB SLAM to the traditional Grid Based SLAM family of algorithms. Simulations of RTAB SLAM are performed and presented in this paper using ROS. Each simulation will be done in a separate world environment. The first environment is a kitchen environment while the second environment will be an office hallway-like environment. We will present the results obtained from both environments to evaluate RTAB SLAM and determine potential weakness , if any, in the algorithm.

**Index Terms**—Robot, IEEEtran, RTAB SLAM,Graph-Based SLAM,ROS.

## 1 Introduction

Simultaneous localization and mapping has been a hot topic within the robotics community over the past years. Its solution is seen as the "holy grail" to both researchers and the robotics community as it would provide the means to obtain a fully autonomous and independent system in environments which are unknown and have never been visited before.

SLAM recovers two important aspects for robot navigation in a simultaneous manner:

1) Creating the map needed for good localization.
2) Estimating the pose needed to produce an accurate map.

Due to the fact that the mapping and localization/pose estimation aspects are interlinked with one another, SLAM has been described as a chicken and egg problem. Thus, SLAM is somehow considered a hard and challenging problem.

In the following sections we will present the theory behind Graph Based SLAM with attention to Real Time Based Appearance Mapping SLAM ( a variation of Graph Based SLAM developed by Labbe et al. [1] ). We will also show and discuss the simulation results preformed in ROS for RTAB SLAM under two different environments. The first environment is a simulated kitchen environment. The second environment is an indoor office hallway like environment. We will also discuss the potential shortcomings,if any, of RTAB SLAM noticed in the simulation results.

## 2 Simultaneous Localization And Mapping

The genesis of SLAM goes as back to 1986 when probabilistic methods were introduced in both Robotics and AI. A number of researchers were trying estimation methods for the localization and mapping problem [2].

The formulation of the SLAM problem can be seen in Figure 1. The motion model is first used to estimate the pose of the robot based on the input control. The Observation is then made via sensor measurements to detect/estimate the distance of landmark or objects in the environment. The estimated robot pose and landmark position are used to update to a more accurate robot position and map containing the location of the features.

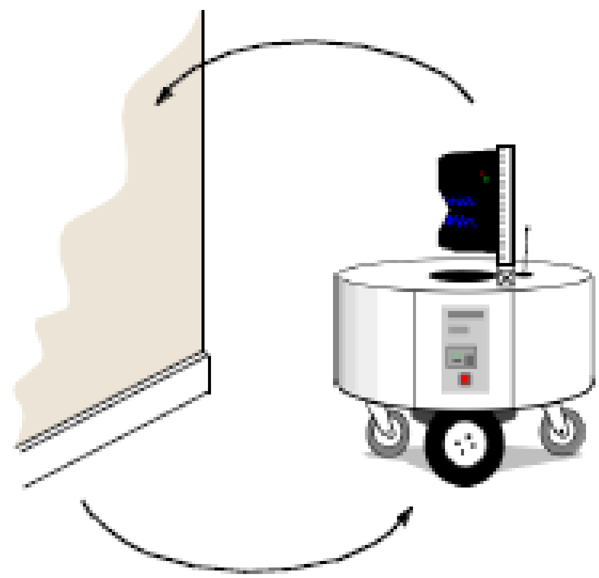

Fig. 1. SLAM is considered a hard problem since both map and pose estimation are preformed simultaneously with no prior knowledge of the environment.

The main solutions of SLAM can be categorized as follows :

1) ***EKF-SLAM*** which is one of the most popular SLAM algorithms. In EKF-SLAM the noise for the motion and measurement models are represented in the form of Gaussian distributions. The main weakness in EKF-SLAM is that it employs a linearized model of non-linear motion and observation models which leads to inconsistencies in the solution. The solution will only converge under a linear case. In addition

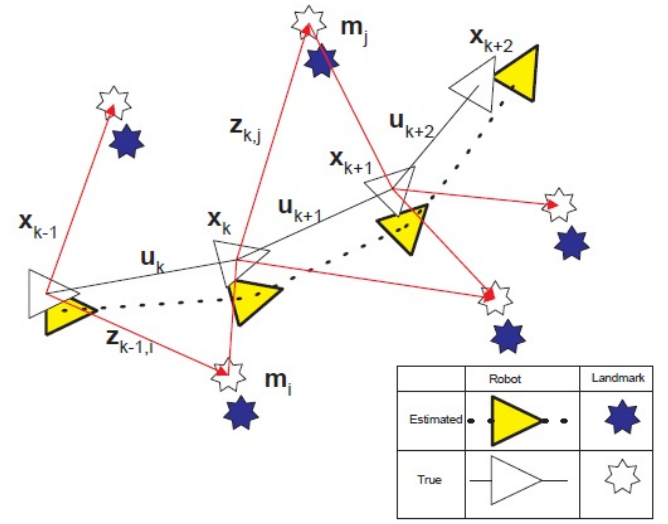


Fig. 2. The essential SLAM problem. A simultaneous estimate of both robot and landmark locations is required. The true locations are never known or measured directly. Observations are made between true robot and landmark locations [1].

EKF-SLAM represents the map in the form of landmark representation.

2) ***Particle Filter-SLAM a.k.a FAST SLAM*** solves the SLAM problem using particles. Each Particle would estimate the path, mean and covariance of the landmarks in the map. The robot posterior is solved by a Rao-Blackwellized filter, where each landmark is represented by a 2x2 Extended Kalman Filter. Also due to the factorization of the Roa-Blackwellized filter each particles has landmark estimates which are independent from other particles.
3) ***Graph Based SLAM*** was born from the intuition that SLAM can be represented as a sparse graph of node with constraints between the different nodes. Figure 3 shows such representation. It could be considered that Graph Based SLAM represents the full solution as an elastic net. The full SLAM solution would be considered to be found by computing the minimal energy of this net.

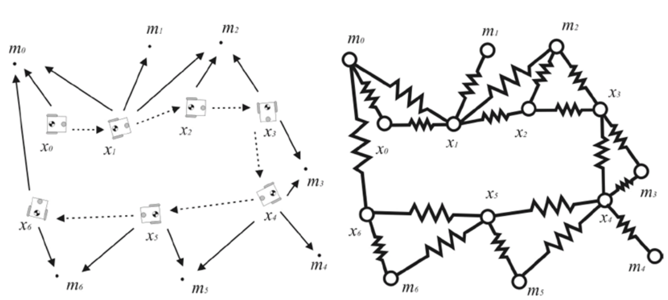

Fig. 3. Graph Based SLAM Representation. The nodes in the graph are the robot position and landmarks. The constraints are the relative distances between consecutive poses of the robot and the distance of the landmark from the robot.

## 2.1 Real Time Appearance Based Mapping SLAM

Real Time Appearance Based Mapping SLAM (also known as RTAB SLAM) is a graph based type of 3D SLAM. RTAB SLAM mainly uses the input from a depth sensor such as RGB-D Depth Sensor or Stereo Based Camera Vision Sensor for 3D perception. Localization data is also collected from sensors such as Wheel encoders or Laser Range Finders. Wheel encoders can provide direct pose readings, whereas through laser range finder we can use Iterative Closest Point - ICP between two different laser scans to estimate the Rotation and Translation between consecutive laser readings.

RTAB SLAM can be split into two steps:

1) ***FRONT END*** : In the front end step, images are collected and visual features are detected.Image features are matched between consecutive images. Figure 4 shows an example of features being detected and matched between two images. Once the matching features are identified there 3D location are extracted from the depth information received. With the 3D information, rotation and translation are estimated using RANSAC rigid transformation.

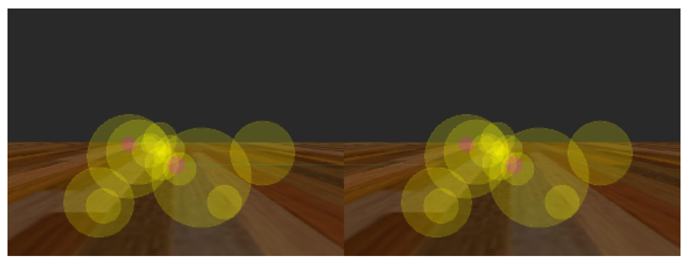

Fig. 4. Feature Matching During SLAM. The yellow circles indicate detected image features.The circle in pink indicate to matched features.

2) ***BACK END*** : In the back end step the graph is created. Nodes are created containing the depth image, RGB image data, and pose estimate on that particular node. The link is the created being the transformation between the two consecutive nodes. In the back end loop closure detection is also preformed. New images acquired are being compared to previous images obtained. Once the old and new images are found to be same, loop closure is preformed and the graph is optimized. The 2D occupancy grid map and 3D Point Cloud map are then generated based on the optimized graph.

Figure 5 outlines the algorithm pipeline.

# 3 Scene and Robot Configuration

## 3.1 Robot Configuration

We will discuss the robot configuration used in this section. The robot is similar to a vacuum cleaning robot with a circular chassis. It carries an RGB-D camera along with a Hokuyu 2D laser range finder. Transform tree of the robot is shown in Figure 6

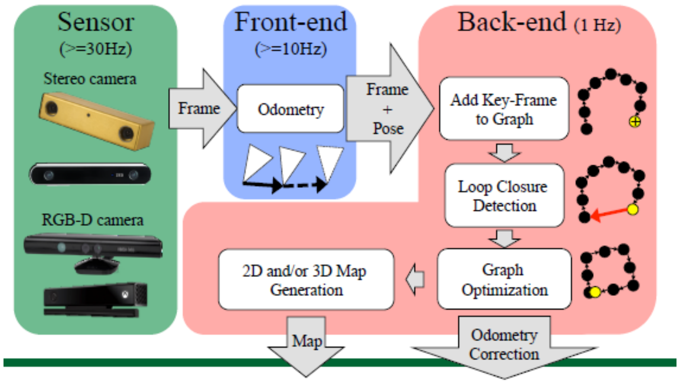


Fig. 5. RTAB-SLAM Pipeline

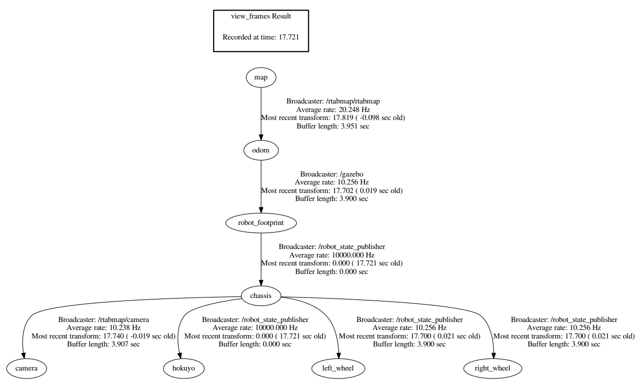


Fig. 6. Transform Tree of the Robot

### 3.2 Scene Configuration

Two scenes were configured. The first kitchen like environment is representative of a normal indoor kitchen environment provided for the project. The second environment is similar to that of an office hallway. Objects were placed in the environment and included barrels,bowls and benches to introduce clutter and objects to the scene.

## 4 Simulations

Simulations were performed in two separate and different environments. The first environment used was the Kitchen world Figure 7. The second environment was the office hallway world Figure 8. Both environments contain cluttered objects in the scene such as barrels,tables,chairs, which the robot has to navigate around.

The worlds were launched in gazebo in addition the "teleop script" was launched which allowed to navigate the robot in the 3D environment. Figures 9 and 10 show the robot in simulation scanning the environment.

## 5 Results

Results were obtained after the simulation was completed. They are presented in below sections.

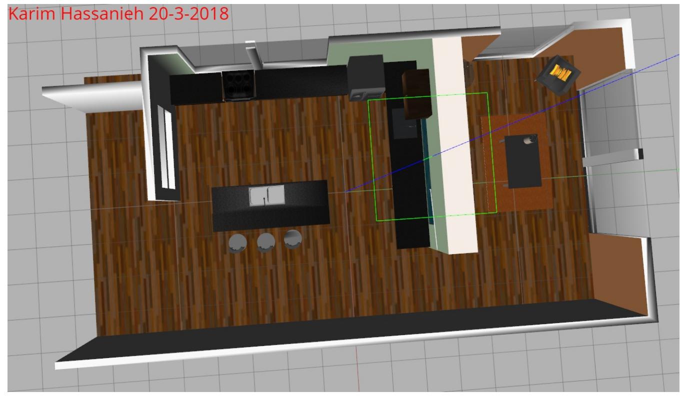


Fig. 7. The Kitchen World Simulated in Gazebo

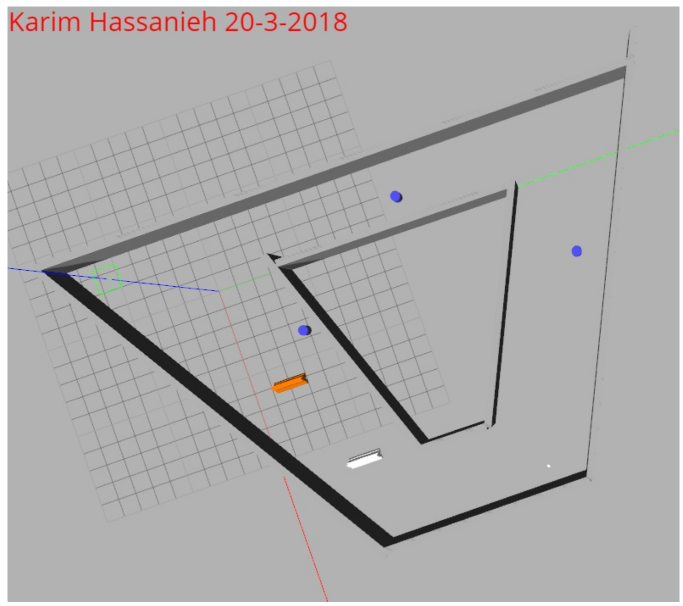


Fig. 8. Office Hallway World Simulated in Gazebo

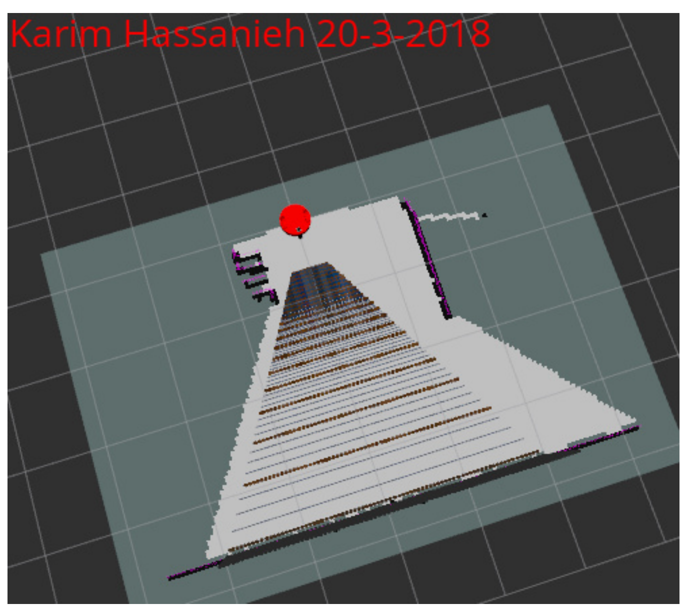


Fig. 9. Robot Scanning Environment in Rviz

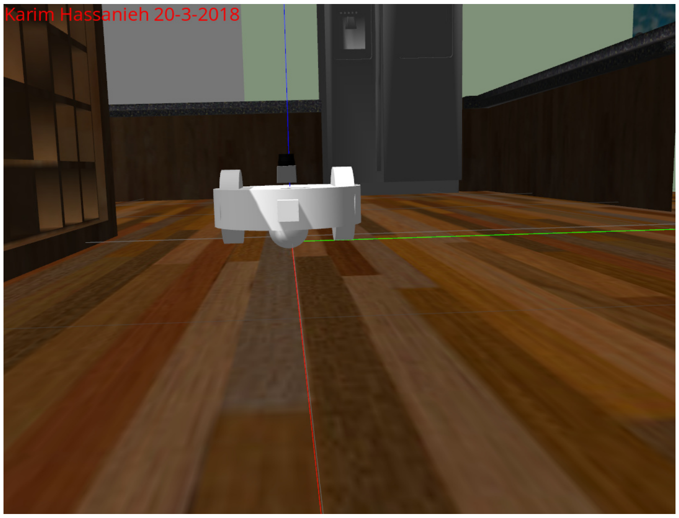


Fig. 10. Robot Scanning Environment in Gazebo

### 5.1 Kitchen World

Figure 11 shows the map using laser scans. Figure 12 shows the occupancy grid map obtained. Figure 13 , 14, 15 show the 3D maps and views obtained at the end of the mapping simulation.


Fig. 11. 2D Laser Map Kitchen World

### 5.2 Office Hallway World

Figure 16 shows the map using laser scans. Figure 17 shows the occupancy grid map obtained. Figure 18 , 19 show the 3D maps and views obtained at the end of the mapping simulation.

## 6 Discussion

3D Maps were obtained for both environments, however it was noticed that performance of RTAB SLAM obtained for the Kitchen environment was much more accurate and representative of the actual environment then the performance of RTAB SLAM in the office hallway environment. Also, the number of loop closures in "Kitchen World" was

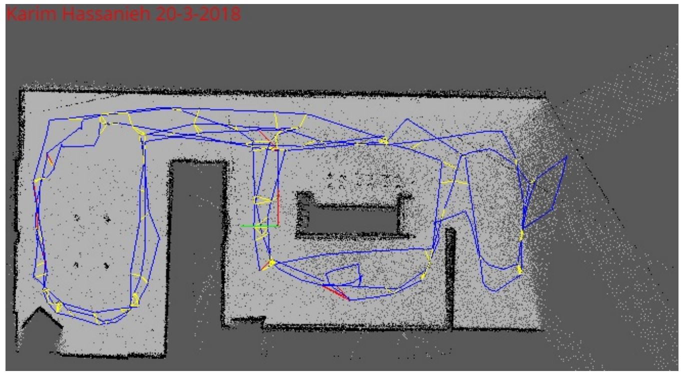


Fig. 12. Occupancy Gird Map Kitchen World


Fig. 13. 3D Map Kitchen World

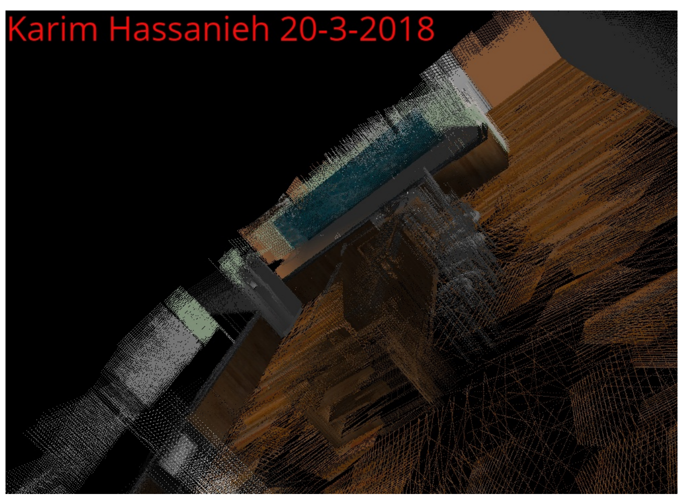


Fig. 14. 3D First Person View Kitchen World


Fig. 15. 3D First Person View Kitchen World

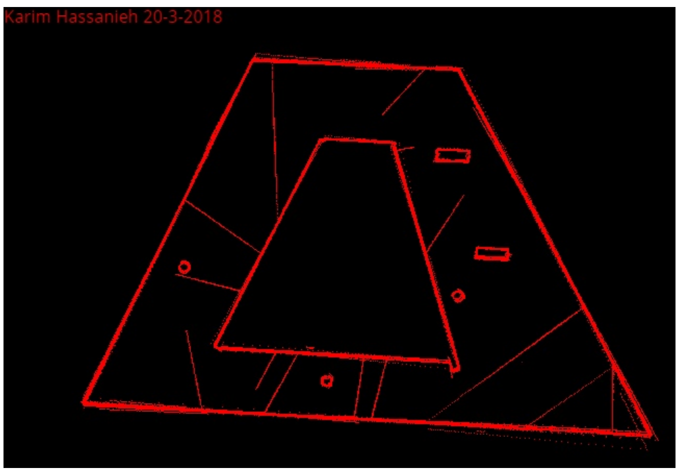


Fig. 16. 2D Laser Map Office World

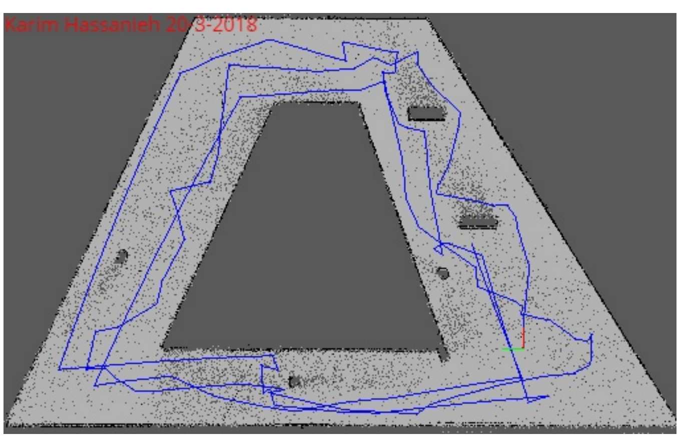


Fig. 17. Occupancy Gird Map Office World


Fig. 18. 3D Map Office World

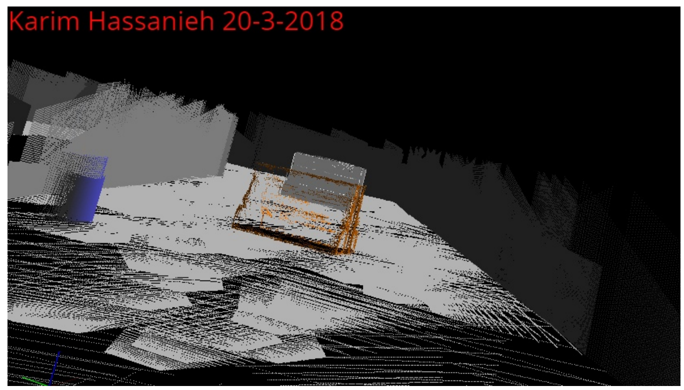


Fig. 19. 3D First Person View Office World

significantly greater. In the Kitchen world the number of loop closures reached up to 16. This is due to the fact that the kitchen environment contained more distinctive texture which made it easier to detect image features. Image features were crucial for the detection of loop closures.

It was also noticed that the mapping of the walls for both environments in Figure 14 and 19 could be further improved in the sense that only partial areas of the wall or floor have been mapped. This can be further improved by assuming that the wall and floor are perpendicular planes. Therefore missing depth information can be assumed to be on the same plane as wall and floor. Of-course this will only hold true for indoor environments.

## 7 Future work

RTAB-MAP SLAM algorithm has shown to have good potential in successful and accurate mapping of environments. For future work such tool can then be used in various number of applications and places where humans are unable to reach. One can think of the places which are hazardous and toxic places for people such as toxic gases. Another application could be perhaps inside places where people usually don't visit such as tunnels or pipes.

## REFERENCES

[1] M. Labb and F. Michaud, "Memory management for real-time appearance-based loop closure detection.," in *IROS*, pp. 1271–1276, IEEE, 2011.

[2] H. Durrant-Whyte and T. Bailey, "Simultaneous localisation and mapping (slam): Part i the essential algorithms," *IEEE ROBOTICS AND AUTOMATION MAGAZINE*, vol. 2, p. 2006, 2006.