NEURAL RADIANCE FIELDS (NeRF) FOR OPTIMUM ROBOT NAVIGATION USING INFRARED SENSORS IN DARK REGIONS

Kanwar Mansoor Ali 21107103

MSME (Mechatronics)

Dr. Faraz Junejo

Dissertation/ Thesis

Faculty of (Mechatronics Engineering)
Shaheed Zulfikar Ali Bhutto Institute of Science and Technology

December / 2023

Acknowledgment

First and foremost, I extend my deepest appreciation to my academic advisor, Dr. Faraz Junejo, for their unwavering guidance, invaluable insights, and tireless commitment to excellence. His mentorship has been instrumental in shaping the direction and quality of this research.

I am also pleased to thank Sir Aneel Ahmed, Program Manager Mechatronics Department SZABIST, for his moral support, patience and steadfast encouragement throughout the Master's program.

A special thanks to all my family members.

Abstract

Primarily, this research focuses on investigating the impact of Neural Radiance Fields (NeRF) on differential drive robot's performance w.r.t its functions of Simultaneous Localization and Mapping (SLAM) leading to respective Autonomous Navigation, while traversing through low light intensity or indoor regions having different obstacles. Namely, how the NeRF based 3D representation of real-world environment, in which robot navigates, could impact the performance of said robot as compared to ideal and built-in representation of the simulated environment that is based on pixels or voxels. The entire research is performed utilizing state of the art software using ROS2 (Robot operating system- version2) and Gazebo ignition Simulator. It is pertinent to mention that ROS2 and Gazebo Simulator are employed for conducting the study because aforementioned open-source software has become the robotic industry standards and benchmarks for robotic research, design and testing of prototypes prior real time development and further production. It is proven scientific approach having remarkable financial and technical outcomes in terms of success. A well-known example is the adoption of said software by NASA for the development of humanoid Robot-Valkyrie for space missions. Secondly, this research targets the process of building a robot with different configuration. Namely, using LiDAR, depth-camera and infrared camera alternatively on differential drive robot. Then to compare the performance of robot with different sensors configuration in terms of the map quality generated during SLAM, time elapsed, synchronization of Autonomous navigation vs command instructions, object avoidance rate and velocities robot maintained. In addition, recent research efforts have unveiled the dominance of NeRF for scene regeneration over traditional method like photogrammetry and other implicit Neural representation methods. Firstly, due to its inborn characteristics including how it handles continuous differentiable functions via deep learning, which increases scene resolution. Additionally, it increases the system's computing efficiency because it doesn't deal with grids, voxels, or pixels. Until now few research studies based on utilization of NeRF with

visual cameras for robot navigation has been done. Despite, this newly evolving field still offer considerable void for investigation especially in relation with the field of robotics. Also, its implementation using ROS2 and Gazebo hasn't been done yet. Therefore, the contribution and innovative factor of this thesis is to evaluate the use of NeRF with infrared camera for performing SLAM and autonomous navigation of robot in environment, marked by low light intensity, using ROS2 and Gazebo. Furthermore, comparison of different robot configuration performance using ROS2 and gazebo is marked as secondary contribution of this research. In this regard, different simulated robotic configurations are built and tested using aforementioned robot development software. For instance, robot configurations individually mounted with Lidar, depth-camera and ultimately infrared camera, are tested one-by-one in terms of the quality that the said robots' configurations exhibit and time they elapsed w.r.t SLAM and Autonomous Navigation operations. Initially, all said configurations are tested in normal built-in Gazebo environment. Secondly, same robotic configurations are tested within NeRF based environment of real world by importing it into gazebo simulator. Conclusively, six experiment are conducted to analyse and compare the quality of SLAM and Autonomous Navigation operation along with respective times each configuration elapsed in different environmental representation of worlds.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

List of Abbreviation

Abbreviation	Meaning
NeRF	Neural Radiance Fields
ROS	Robot Operating System
CNN	Convolutional Neural Network
IR	Infrared
LiDAR	Light Detection and Ranging
IMU	Inertial Measurement Unit
MLP	Multi-Layer Perceptron
RGB	Red Green Blue

Chapter 1: Introduction

1.1 Background of the research

SLAM and Navigation of robots through indoor regions having low light intensity has been an interesting area of research. Sensors like LiDAR, depth-cameras and Infrared cameras are frequently used to address this problem. These sensors offer useful information about the environment, allowing robots to move about in a safe and efficient manner while avoiding obstacles. However, at early stages, it isn't rational or economically feasible to design and construct a physical robot and test its navigational abilities. Due to the utilization of computationally powerful microcontrollers, industrial-grade sensors, precise actuators, and batteries, it would be impractical to spend on development costs at the preliminary research stage. Consequently, entire effort and investment could be lost in the event of failure. Furthermore, testing an actual robot in different environmental conditions is a difficult problem to address during preliminary stages. Because managing lightning conditions, outside noises, and system configuration could be extremely difficult, choosing the wrong or mismatched hardware and software for a certain working environment could have disastrous effects. Navigation and other functionalities of robot can be enhanced by ROS2 Framework. Complex robotic system development is made easier by the modularity and flexibility of ROS2 [1]. It offers a strong communication infrastructure, the ability to integrate sensors, simulation and visualisation tools. Developers can build dependable and effective navigation algorithms for robots operating in difficult settings by utilising ROS2 [2]. With Gazebo Ignition, a widely used simulator in the robotics community that is integrated with ROS2, researchers and developers can faithfully replicate real-world scenarios, including dark regions, and assess the effectiveness of navigation algorithms and sensor systems. On the other hand, Neural Radiance Fields (NeRF) is a novice method for representing the world in which a robot moves or Autonomously Navigates. NeRF is a state-of-the-art method that has attracted a lot of interest in the domain of computer vision. By simulating the volumetric radiance field, it is possible to reconstruct intricate 3D scenes from a collection of input photos [3]. NeRF portrays the scene as a continuous function, enabling high-fidelity reconstructions and the precise representation of fresh views. Robots can create thorough and accurate representations of their surroundings by utilising NeRF to train sensors' data, which will help with tasks like mapping, localization, and navigation. In comparing the NeRF based representation against normal simulation within Gazebo, following crucial factors are relevant for performing comparative Analysis and acquiring Results:

1. Obstacle Avoidance: The robot's capacity to recognise and steer clear of impediments in its path [5];
2. Safety: the ability of the navigation system to reliably and robustly avert collisions or mishaps;
3. Time Lag: The lag between sensor readings and robot responses has an impact on system performance overall and real-time decision-making;

4. Quality of Maps.

NeRF will shed light on the advantages and disadvantages of each approach for diverse sensor-based navigation within indoor regions. This research advances autonomous robotics, industrial automation, and search and rescue operations by improving the navigational capabilities of robots' indoor areas. The results will contribute to enhancing robot navigation in difficult and visually limited areas in terms of efficiency, safety, and effectiveness.

1.2 Research problem

This research addresses critical challenges in indoor lighting conditions for improving environmental estimation in SLAM and Autonomous Navigation operations particularly for differential drive robots. Until recently researcher have employed NeRF technology with only RGB camera and for aerial robots only for achieving better results for SLAM and Navigation. However, It aims to evaluate the integration of ROS2, Gazebo ignition simulator, and Neural Radiance Fields (NeRF) with LiDAR, depth cameras, and infrared sensors for accurate environment estimation in differential drive robots. Additionally, it compares NeRF with conventional simulation technologies to enhance navigation precision and reliability. The study also explores optimal sensor configurations to mitigate challenges in reduced visibility and GPS-assisted navigation absence.

1.3 Importance of the research

The importance of this study rests in its prospective contributions to the field of NeRF assisted robotic simulation using LiDAR, depth camera and infrared sensor-based robot navigation in low-light conditions. The research can have a number of significant towards following factors:

1. Better Navigation Capabilities: Improving robot navigation in low-light situations is essential for a variety of applications, such as search and rescue operations, exploration of dark spaces, and industrial automation. This work intends to enhance the precision, efficiency, and reliability of robot navigation, enabling them to successfully function in difficult environments.

2. Increased Safety: Robotics places a high priority on safety. The likelihood of collisions and mishaps while navigating increases. The project intends to increase the safety of robots operating in indoor situations by creating reliable navigation systems employing infrared sensors. Effective obstacle avoidance and trustworthy environment perception can reduce collision risk and guarantee safe robot operation.

3. Energy Efficiency: For autonomous robots with limited battery capacity, power consumption is a crucial factor. The work can aid in the creation of energy-efficient navigation systems by optimising navigation algorithms based on the training of infrared sensor data using NeRF. As a result, robots can operate for longer periods of time and require less frequent charging or battery replacement.

4. Real-time responsiveness: The time it takes for a robot to respond to a sensor input can affect how effective navigation is in dynamic situations. The study intends to measure and reduce the delay by analysing the time lag parameter, allowing robots to act quickly and adjust to changing circumstances in

real-time.

5. Compare and contrast: The work sheds light on the benefits and drawbacks of these training methods for infrared sensor-based navigation by contrasting the performance of NeRF and normal simulation environment.

Researchers and practitioners can choose the best strategy depending on the unique requirements of their applications with the aid of this comparative analysis, which can also help them comprehend the trade-offs.

Overall, this research has the potential to enhance the field of robot navigation in low-light conditions, enhancing the effectiveness, safety, and dependability of robots operating in these difficult circumstances. The research could have applications in a number of fields and businesses, such as robotics, automation, surveillance, and disaster response, ultimately resulting in the creation of more powerful and flexible robotic systems.

1.4 Organisation of the Thesis

The following chapters make up the remainder of the work, and the summaries of each chapter are provided below:

- Chapter 2: A survey of the literature to help you comprehend the thesis
- Chapter 3: Research Methodology
- Chapter 4: Tools and Techniques, Key Ideas, including ROS2, Gazebo, NeRF, SLAM, Vision-Based In-depth discussions are held on navigation.
- Chapter 5: Results and Comparative Analysis
- Chapter 6: Conclusion, Recommendation and Future Direction

Chapter 2: Literature Review

2.1 Introduction

Journals, conferences, and technical reports are the sources for research articles. The relevant sources include, but are not limited to, IEEE Xplore, Google Scholar, and the websites of Cornell University. The domain of Vision based Navigation is presented in the paragraphs that follow, followed by research on NeRF and its applications. Finally, research focused on the use of Gazebo and ROS2 for research and development purposes is discussed.

2.2 Research on Vision Navigation

The study by Arun Narenthiran Sivakumar et al. [6] scrutinizes challenges faced by agricultural robots navigating under canopies. While their proposed deep learning and computer vision solution seems promising, its practical implementation and limitations require thorough examination. The reliance on visual data raises concerns about resilience to environmental changes and the system's effectiveness under dynamic conditions. Additionally, the study lacks in-depth discussion on potential drawbacks and trade-offs. The reported experiments could benefit from a more comprehensive evaluation across various conditions. Furthermore, broader socio-economic implications and ethical considerations merit attention. In essence, while the study presents a novel approach, a critical examination of its implementation, performance, and broader implications is crucial for ensuring its efficacy and sustainability. Saurabh Gupta et al. [7] introduce a new technique in 2017 aiming to enhance visual navigation capabilities by merging map-based and landmark-based representations. Though proposing leveraging both approaches' strengths, critical examination is needed regarding practical implementation and effectiveness. The integration of map-based and landmark-based representations seems promising in theory. However, the paper lacks thorough exploration of potential challenges and limitations in merging these approaches. Questions arise about compatibility of different data representations and potential trade-offs in computational complexity or memory requirements. Furthermore, while discussing technical aspects of integration like feature extraction and fusion strategies, potential pitfalls or uncertainties in methodologies are overlooked. Robustness to environmental variations and scalability to different terrains require scrutiny. Additionally, the paper could benefit from comprehensive evaluation of the framework's performance metrics. While mentioning navigation accuracy, deeper analysis of robustness to environmental changes and adaptability to dynamic surroundings is necessary. Moreover, absence of comparative analysis with existing techniques raises doubts about framework's superiority. Benchmarking against alternative methods is crucial to assess benefits against implementation costs. Kaichun Mo et al. [8] present the Adobeindoornav dataset and its application for deep reinforcement learning in indoor robot vision navigation. While the paper discusses the challenges of visual navigation in intricate indoor environments, a critical examination is necessary regarding the dataset's utility and the effectiveness of deep reinforcement learning methods. The introduction of the Adobe Indoor Navigation dataset appears promising, offering a substantial collection of RGB-D photos with depth maps and robot trajectories. However, the paper lacks a thorough discussion of potential limitations or biases inherent in the dataset. Moreover, while highlighting the potential of deep reinforcement learning for indoor visual navigation, the paper overlooks the complexities and uncertainties associated with training reinforcement learning agents in real-world environments. Factors like environmental variability, occlusions, and lighting conditions are briefly mentioned but not deeply explored. The technical aspects of the experimental design, including network topologies and training procedures, are discussed, but the paper lacks detailed analysis of the algorithm's performance metrics and robustness. Furthermore, while emphasizing the dataset's value in advancing research on visual navigation algorithms for indoor robots, the paper fails to acknowledge potential biases or limitations that may affect algorithm performance in real-world applications. Without addressing these concerns, the dataset's practical utility may be overestimated. Yiding Qiu et al. [9] propose an innovative method for target-driven visual navigation that leverages object relationships, aiming to enhance the efficacy and efficiency of visual navigation systems. While the research ambitiously combines object connection reasoning and target-driven navigation, critical scrutiny is needed to assess the practical implementation and effectiveness of this approach. While the paper outlines the methodology's theoretical framework, it lacks comprehensive discussion on potential limitations or challenges in integrating object relationship reasoning and target-driven navigation. Moreover, while the technical details of the framework are briefly mentioned, a deeper analysis of the models, algorithms, and methods utilized for object connection reasoning and target-driven navigation is warranted. The paper may benefit from providing more insight into how object relationships are modeled, visual cues extracted, and integrated into the navigation process. Furthermore, while experiments and assessments are likely conducted to evaluate the approach, a thorough analysis of performance metrics and comparisons with baseline techniques is necessary to validate the framework's effectiveness. Without addressing these concerns, the practical utility and generalizability of the proposed method may be questionable. Khanh Nguyen and Hal Daumé III [10] introduce an approach for visual navigation that integrates natural multimodal support through retrospective curiosity-encouraging imitation learning, aiming to enhance visual navigation systems by combining human demonstrations and natural language instructions. While the paper proposes an ambitious framework combining curiosity-driven inquiry with imitation learning, critical examination is necessary to assess its practical implementation and effectiveness. While the technical discussion outlines the framework's design and procedure for gathering data, it lacks in-depth analysis of potential limitations or challenges in integrating human demonstrations and natural language instructions. Moreover, while the framework promotes curiosity-driven exploration, the paper overlooks the complexities and uncertainties associated with training the policy network under self-supervision. The paper may benefit from providing more insight into the methodology's robustness and generalizability across different environments and situations. Furthermore, while experiments and assessments are likely conducted, a more comprehensive evaluation of performance metrics and comparisons with benchmark techniques is needed to validate the framework's efficacy. Without addressing these concerns, the practical utility and reliability of the proposed method may be questioned. Fengda Zhu et al. [11], provide a comprehensive assessment of deep learning methods in embodied vision navigation, aiming to review developments, difficulties, and trends in the field. While the paper offers an extensive overview of deep learning's role in navigation tasks, critical scrutiny is needed to evaluate the practical implications and limitations of the discussed methods. Although the authors emphasize the importance of deep learning in this context, the paper lacks in-depth analysis of the challenges and constraints faced by current methods, such as generalization to unseen environments and robustness to sensory noise. Furthermore, while the study covers various deep learning techniques and architectures for navigation components, it overlooks potential biases or limitations inherent in these approaches. Additionally, while exploring benchmarks and datasets, the paper fails to provide sufficient insight into the reliability and generalizability of these evaluation metrics. Moreover, while anticipating future trends, the paper briefly mentions potential research directions without elaborating on their feasibility or practical implications. Without addressing these concerns, the paper's contribution to advancing the field of embodied vision navigation may be limited. Haiyang Wang et al.'s presentation of a new method for group visual navigation [12] appears to address the challenges of navigation in complex environments through collaboration among multiple agents. While the paper emphasizes the benefits of cooperative navigation, critical examination is warranted to assess its practical implications and limitations. Although the framework proposes leveraging the collective intelligence of multiple agents, it lacks comprehensive discussion on potential constraints or challenges in coordinating agent actions and sharing information effectively. Furthermore, while technical aspects of the collaborative visual navigation framework are briefly mentioned, the paper overlooks potential biases or limitations inherent in the communication protocols and coordination algorithms employed. Additionally, while experiments may have been conducted to evaluate the viability of the cooperative navigation strategy, the paper lacks detailed analysis of performance metrics and comparisons with alternative methodologies. Without addressing these concerns, the practical utility and reliability of the proposed collaborative approach may be questionable, especially in dynamically changing environments. Kevin Chen et al. [13] introduce a method of visual navigation by combining graph localization networks, emphasizing the utilization of behavioral cues and graph-based representations to address challenges in navigation. While the paper presents an innovative approach, critical scrutiny is essential to assess its practical implications and limitations. Although the framework integrates behavioral cloning with graph neural networks, the paper lacks in-depth discussion on potential constraints or challenges in effectively blending these components. Furthermore, while technical details of the proposed framework are briefly mentioned, the paper overlooks potential biases or limitations inherent in the architecture and training procedures employed. Additionally, while experiments may have been conducted to evaluate the strategy's efficacy, the paper lacks detailed analysis of performance metrics and comparisons with alternative methodologies. Without addressing these concerns, the practical utility and reliability of the proposed approach may be questionable, especially in dynamically changing environments.

2.3 Research papers on Neural Radiance Fields (NeRF)

Ben Mildenhall et al. [14] introduce Neural Radiance Fields (NeRF) as a novel technique for view synthesis, aiming to generate realistic photographs of complex scenes from fresh perspectives. While the paper highlights NeRF's potential in creating accurate and visually coherent views, critical examination is necessary to assess its practical implications and limitations. Although NeRF represents scenes as continuous 3D functions, the paper lacks in-depth discussion on potential challenges or limitations in handling complex scene geometry, occlusions, and challenging lighting conditions. Furthermore, while technical details of the NeRF framework are briefly mentioned, the paper overlooks potential biases or limitations inherent in the design and training methods of the multi-layer perceptrons (MLPs) used in NeRF. Additionally, while experiments may have been conducted to evaluate NeRF's performance, the paper lacks detailed analysis of performance metrics and comparisons with alternative techniques. Without addressing these concerns, the practical utility and reliability of NeRF in real-world applications may be questionable, especially in scenarios with sparse data or dynamic environments. Michal Adamkiewicz et al. [15] propose a method for vision-only robot navigation using neural radiation fields (NeRF), aiming to address the challenges of robot navigation without relying on explicit geometric models or sensor fusion techniques. While the paper presents an innovative approach, critical examination is necessary to assess its practical implications and limitations. Although the framework leverages NeRF to model the environment as a continuous 3D function, the paper lacks in-depth discussion on potential challenges or limitations in incorporating visual observations into the NeRF representation and extracting relevant features for navigation. Furthermore, while technical aspects of the vision-only navigation framework are briefly mentioned, the paper overlooks potential biases or limitations inherent in the perception pipeline and localization algorithms used. Additionally, while experiments may have been conducted to evaluate the strategy's efficacy, the paper lacks detailed analysis of performance metrics and comparisons with alternative techniques. Without addressing these concerns, the practical utility and reliability of the vision-only navigation system may be questionable, especially in real-world scenarios with dynamic environments or limited visual cues. A thorough overview of perception and navigation techniques in autonomous systems is provided by Yang Tang et al. [16], which was published in IEEE Transactions on Neural Networks and Learning Systems in 2022. A special emphasis is placed on the improvements made by learning-based approaches. The research examines how machine learning and neural networks can be used to help autonomous systems overcome difficulties in sensing and navigating their environment. Beginning with perception, localization, mapping, and navigation, the authors outline the basic elements of autonomous systems. They talk about the conventional techniques employed in these fields before delving into the nascent topic of learning-based approaches. The paper gives a general introduction of the most important learning methods, including deep learning, reinforcement learning, and unsupervised learning, along with examples of how they are applied to perception and navigation problems. A wide range of subjects pertaining to perception and navigation in autonomous systems are covered in the survey. It looks at how learning approaches have been used to enhance the resilience and accuracy of several sensory modalities, such as vision, LiDAR, and radar. The authors look at localization and mapping techniques that use learning algorithms to increase efficiency and accuracy. The paper also explores navigation tactics such as motion planning, obstacle avoidance, and exploration and emphasises how learning-based approaches have improved these fields. It examines the advantages and difficulties of employing learning techniques for navigation, including the requirement for data efficiency and safety considerations, as well as the capacity to adjust to dynamic surroundings. The authors will probably analyse the benchmarks, assessment measures, and datasets already available in the area of autonomous perception and navigation. They might also talk about unresolved issues like generalisation to new settings, the interpretability of learned models, and the merging of many modalities. The research's importance can be found in its thorough analysis of the improvements in perception and navigation brought about by learning-based techniques in autonomous systems. The survey is a useful tool for scholars and practitioners in the field because it summarises the body of literature and highlights the most important strategies and trends. It offers information on how learning algorithms may enhance the vision and navigational capabilities of autonomous systems, allowing them to function more efficiently and securely in real-world situations. In 2022, Janaki Sun et al. [17] introduced a technique called NeRF-Loc that provides precise object localization within neural radiation fields (NeRF). The work addresses the problem of object localization in NeRF representations of complicated scenes that generally lack explicit object-level annotations. The authors provide a transformer-based method for localising objects in NeRF's continuous 3D representation by fusing visual information and 3D positional encoding. To capture interactions between various regions of the picture, the approach makes use of the inherent spatial information encoded within the light fields and applies self-attention mechanisms. NeRF-Loc can successfully deduce the exact positions and extents of objects within the scene by introducing transformers into the localization process. The technical aspects of the NeRF-Loc framework, such as the architecture of the transformer model, the incorporation of visual features and positional encoding, and the training process to improve localization performance, are probably presented in the paper. The transformer model learns to focus on pertinent traits and spatial correlations, enabling precise and effective object localization, as the authors might explain. The authors most likely performed trials using benchmark datasets or artificial situations to gauge NeRF-Loc's effectiveness. Metrics like recall, precision, and localization accuracy may be included in the evaluation. For the purpose of showcasing NeRF-Loc's advantages in terms of precision and robustness, the authors may compare its performance with that of other object localization techniques. The discovery is significant because NeRF-Loc can localise objects within NeRF representations, which is useful for a number of applications, such as augmented reality, virtual reality, and scene comprehension. The suggested solution expands NeRF's capabilities by providing precise item localization, which opens the door for better scene comprehension and interaction in virtual and augmented environments. Neural Radiance Fields (NeRF) are used in Arunkumar Byravan et al.'s [18] technique, which they name Nerf2real, to enable the transfer of vision-guided bipedal motion abilities from virtual environments to the actual world. The research addresses the problem of skill transfer from simulation to the real world, which frequently suffers from domain shift and sensor data restrictions. The authors provide a system that combines reinforcement learning and NeRF-based sensing methods to let bipedal robots move around and interact with their surroundings. NeRF is used to depict the simulated environment because it offers a continuous and accurate 3D scene representation. Utilising the visual data offered by NeRF, the robots learn to comprehend their surroundings and develop motion skills through reinforcement learning. The document probably includes the technical specifics of the Nerf2real framework, such as the creation of the reinforcement learning algorithms, the mechanisms for sim2real transfer, and the integration of NeRF-based perception into the learning pipeline. In order to bridge the gap between simulated and real-world environments and guarantee that learned abilities may be properly transferred, the authors may explore strategies like domain adaptation and finetuning. The authors probably carried out tests using actual robot platforms and simulated surroundings to gauge Nerf2real's efficacy. Metrics like task completion rate, locomotor efficiency, and generalisation capacity may be included in the evaluation. In order to show Nerf2Real's advantages in terms of skill transfer, adaptability, and resilience, the authors may compare its performance with that of other Sim2Real transfer methods. The research is significant because it demonstrates how Nerf2Real may use NeRF-based perception to close the Sim2Real gap in vision-guided bipedal motion skills. The proposed system enables robots to learn and apply abilities from simulation to the real world by utilising NeRF's capability to deliver realistic scene representations and combining it with reinforcement learning. This has potential applications in a variety of fields, including interactive jobs, robotic manipulation, and mobility, where vision-based abilities are essential. In 2022, Kyle Gao et al. [19] presented their thorough analysis of neural radiation fields (NeRF) in the context of 3D vision as an arXiv preprint. The goal of the study is to showcase NeRF and its derivatives' contributions to the field of 3D scene interpretation by summarising their important ideas, developments, and applications. The notion of NeRF, which describes a scene as a continuous 3D function that converts 3D coordinates to radiance values, is introduced by the authors first. The volumetric representation, the usage of neural networks to model the radiance function, and the rendering procedure to produce fresh views of the scene are some of the fundamental ideas of NeRF that are covered in their discussion. The document most likely gives a summary of the various NeRF architectural variations, along with each one's advantages and disadvantages. The review examines the enhancements and extensions to the initial framework made in NeRF-related research. Multi-object NeRF, dynamic scene modelling, scale-aware NeRF, and adding temporal information into NeRF representations are some of the subjects covered by the writers. They probably explore the consequences of each variation's primary contributions for 3D vision tasks including reconstruction, rendering, and scene comprehension. In addition, the article looks at NeRF's uses in a number of fields, including augmented reality, virtual reality, robotics, and autonomous driving. The writers will probably go over how NeRF has been used for things like object manipulation, scene reconstruction, new view synthesis, and realistic rendering. They can also draw attention to the difficulties and potential avenues for further development and NeRF implementation in practical applications. The review's importance can be attributed to its thorough examination and synthesis of the 3D vision research that has been done in the field of NeRF. The publication offers a useful guide for scholars and professionals in the field by condensing the main ideas, variants, and applications of NeRF. It opens the door for more improvements and developments in this interesting field of research and sheds light on NeRF's potential as a potent tool for 3D scene interpretation. Neural Radiance Fields (NeRF) are used in a technique dubbed Loc-NeRF by Dominic Maggio et al. [20], which is used for Monte Carlo localization. The goal of this study is to develop a method for accurately determining a mobile agent's pose and location in a given environment from visual data. The authors provide a framework that integrates Monte Carlo localization, a probabilistic method frequently used in robotics for posture prediction, with NeRF, which offers a continuous 3D scene representation. In Loc-NeRF, synthetic data are simulated using NeRF as a generative model and compared to actual observations to update the agent's posture estimate. Loc-NeRF seeks to increase the precision and robustness of localization by making use of the extensive 3D data included in NeRF. The Loc-NeRF framework's technical specifics, including the incorporation of NeRF into the Monte Carlo localization pipeline, the creation of synthetic observations, and the update of the agent's pose estimate via particle filtering or comparable methods, are most likely presented in the paper. The authors might go through how NeRF makes it possible to create plausible synthetic observations and how the combination of artificial and actual data enhances localization performance. The authors probably ran tests on benchmark datasets or in simulated settings to gauge Loc-NeRF's effectiveness. Metrics including localization accuracy, convergence rate, and resistance to noise and occlusions may be included in the evaluation. For the purpose of showcasing Loc-NeRF's advantages in terms of precision and effectiveness, the authors may compare its performance with that of other localization techniques. The research is significant because it combines NeRF with Monte Carlo localization to tackle the difficult issue of precise localization in intricate situations. Loc-NeRF presents a potential method for reliable and accurate localization in a variety of applications, such as robotics, autonomous navigation, and augmented reality. It does this by combining the advantages of NeRF's rich scene representation and Monte Carlo localization's probabilistic estimation. In order to enforce safety in vision-based controllers, Mukun Tong, Charles Dawson, and Chuchu Fan [21], released as an arXiv preprint in 2022, describe a technique that combines control barrier functions (CBFs) and neural radiation fields (NeRF). The problem of maintaining safety in control systems that rely on visual input is discussed in the study, especially in situations when standard safety limits are difficult to describe. The authors suggest a system that combines NeRF, a potent instrument for scene representation and perception, with CBFs, a control theory term for safety enforcement. The combination of these methods makes it possible to assure safety while utilising the thorough scene comprehension offered by NeRF. The document most likely outlines the technical specifics of how safety limitations developed using CBFs and based on visual data from NeRF are implemented into the controller design. The authors probably ran simulations or trials in situations requiring safety enforcement, including collision avoidance or obstacle navigation, to gauge the viability of the suggested approach. Metrics including safety violation rates, controller performance, and robustness to various environmental conditions may be included in the evaluation. The authors may contrast their strategy with alternative approaches to illustrate its benefits in terms of performance and safety enforcement. The discovery is significant because it integrates CBFs with NeRF to address the fundamental problem of safety in vision-based control systems. The suggested framework offers a systematic mechanism to impose safety limitations based on visual information by combining the advantages of both approaches. This has significant ramifications for many applications where safety is a top priority, such as robots, autonomous driving, and human-machine interaction. E-nerf is a technique introduced by Simon Klenk et al. [22] in IEEE Robotics and Automation Letters in 2023 that permits the creation of neural radiation fields (NeRF) using a moving event camera. The challenge of integrating event cameras—novel sensors that record asynchronous changes in a scene—into the NeRF framework for scene representation and rendering is covered in this study. The authors suggest a paradigm for producing precise and dynamic NeRF representations by combining event camera data with conventional intensity images. They develop a neural network architecture that can successfully process both event and intensity data in order to meet the special properties of event cameras, such as their high temporal resolution and sparse data format. The technical aspects of the E-nerf framework, such as the network design, the training process, and the merging of event and intensity data, are probably presented in the paper. The authors probably ran tests utilising real-world event camera datasets to gauge E-NeRF's efficacy. Metrics like reconstruction accuracy, rendering quality, and generalisation ability may be included in the evaluation. To show E-NeRF's advantages in handling event camera data and capturing dynamic scenes, the authors may compare its performance with conventional NeRF techniques. The research is significant because it combines event cameras with NeRF to capture and portray dynamic situations with asynchronous and sparse data. E-nerf creates new opportunities for scene analysis, reconstruction, and rendering in applications like robotics, augmented reality, and autonomous vehicles by utilising the special qualities of event cameras. The study adds to the expanding field of neural network-based scene representation in conjunction with cutting-edge sensor technology. A method dubbed CATNIPS for collision avoidance utilising neural implicit probabilistic scenarios is presented by Timothy Chen, Preston Culbertson, and Mac Schwager [23], and it will be available as an arXiv preprint in 2023. The research uses neural networks to simulate the implicit probabilistic representation of the world in order to tackle the problem of collision avoidance in dynamic situations. The authors suggest a system to calculate collision probability and produce collision-free trajectories that blends neural networks with probabilistic modelling. A neural implicit representation, used by CATNIPS to represent the scene, captures both the geometry and the uncertainty of the surrounding environment. The neural network design, the training process, and the computation of collision probability based on the probabilistic scene representation are expected to be covered in detail in the paper. The authors probably ran tests with dynamic obstacles in simulative or real-world environments to gauge CATNIPS' efficacy. Metrics including collision rates, trajectory variation, and computing efficiency may be included in the evaluation. To show CATNIPS' advantages in terms of precision, safety, and real-time operation, the authors may evaluate its performance in comparison to that of other collision avoidance techniques. The research is significant because it offers a collision avoidance method that makes use of implicit probabilistic neural scenes. In order to estimate collision risks and produce safe paths in dynamic situations, CATNIPS combines neural networks and probabilistic modelling. By tackling the significant problem of collision avoidance, which is essential for the safe and dependable functioning of robots in a variety of applications, the study makes a contribution to the field of robotics and autonomous systems. At the IEEE/CVF Conference, Matthew Tancik et al.'s [24] research was presented. Block-nerf, a technique for scalable and effective neural view synthesis in large scenes, is introduced in this study. In order to create new perspectives of a scene from a small number of input views, neural vision synthesis is used. Existing neural radiation fields (NeRF)-based techniques, however, can be computationally expensive and difficult to scale to huge settings. By dividing the view into chunks and training a different neural network for each block, Block-nerf overcomes these difficulties. When compared to a single network for the full picture, this enables parallel computation and lowers memory requirements. The picture is to be divided into blocks of varied sizes, with smaller blocks capturing finer details and larger blocks depicting global structure, according to the authors' proposed hierarchical spatial partitioning system. Technical information on partitioning and how networks are trained to provide high-quality new perspectives is probably included in the study. The authors probably ran tests on huge datasets or intricate scenes to gauge Block-NeRF's performance. Metrics including rendering quality, view synthesis accuracy, and processing efficiency may be included in the evaluation. To show Block-NeRF's benefits in terms of scalability, memory effectiveness, and visual fidelity, the authors may compare its performance with other cutting-edge techniques. The research is significant because it addresses the difficulties with the scalability of neural vision synthesis in large situations. Block-nerf makes it possible to quickly generate fresh views in settings that were previously difficult to manage with existing techniques by introducing the idea of block-wise neural networks. This has significant ramifications for computer graphics, virtual reality, and gaming applications, where the capacity to synthesise realistic images of intricate scenes is essential. Kevin Lin and Brent Yi's presentation at the Robotics Science and Systems conference in 2022 introduces a technique for active view planning within the context of Radiance Fields, neural network-based representations of 3D scenes for view synthesis and rendering. While the research addresses a pertinent issue in view planning, critical examination is essential to assess its practical implications and limitations. Although the authors propose a framework combining reinforcement learning with uncertainty estimation for view selection, the paper lacks in-depth discussion on potential challenges or limitations in effectively integrating these components. Furthermore, while technical details of the reinforcement learning agent and uncertainty estimation process are briefly mentioned, the paper overlooks potential biases or limitations inherent in the evaluation metrics used. Additionally, while experiments may have been conducted to evaluate the strategy's efficacy, the paper lacks detailed analysis of performance metrics and comparisons with alternative view planning systems. Without addressing these concerns, the practical utility and reliability of the proposed active view planning technique may be questionable, especially in real-world scenarios with dynamic scenes or limited view information. Raia Hadsell et al.'s paper [32] in the Journal of Field Robotics in 2009 presents a vision-based system for autonomous off-road driving, aiming to address challenges in navigating unstructured outdoor environments. While the research proposes an innovative approach, critical examination is necessary to assess its practical implications and limitations. Although the authors highlight the drawbacks of conventional sensor-based methods and the advantages of vision-based systems, the paper lacks in-depth discussion on potential challenges or limitations in effectively training a convolutional neural network (CNN) for off-road driving. Furthermore, while technical details of the CNN's architecture and training procedure are briefly mentioned, the paper overlooks potential biases or limitations inherent in the dataset used for training and the evaluation metrics employed. Additionally, while experiments may have been conducted to evaluate the system's performance, the paper lacks detailed analysis of performance metrics and comparisons with alternative methods. Without addressing these concerns, the practical utility and reliability of the proposed vision-based system for autonomous off-road driving may be questionable, especially in real-world scenarios with diverse and unpredictable environments. Chen Chen et al. [33] addresses the issue of low-light picture augmentation, aiming to improve photos taken in dim lighting using a deep learning-based method. While the research proposes a promising approach, critical examination is necessary to assess its practical implications and limitations. Although the authors highlight the limitations of conventional techniques for low-light image enhancement and emphasize the importance of raising image quality in such conditions, the paper lacks in-depth discussion on potential challenges or limitations in effectively training a deep convolutional neural network (CNN) for this task. Furthermore, while technical details of the CNN's architecture and training process are briefly mentioned, the paper overlooks potential biases or limitations inherent in the benchmark datasets used for evaluation and comparison. Additionally, while experiments may have been conducted to evaluate the method's efficacy, the paper lacks detailed analysis of performance metrics and comparisons with state-of-the-art methods. Without addressing these concerns, the practical utility and reliability of the proposed deep learning-based approach for low-light image enhancement may be questionable, especially in real-world scenarios with diverse lighting conditions and image characteristics.

2.4 ROS2 and Gazebo based research papers

Steven Macenski et al.'s [26] examine the design philosophies and objectives that guided the creation of ROS 2, highlighting significant advancements and distinctions from ROS. They offer insights into ROS 2's fundamental design, describing its modular and distributed nature as well as its support for various message and communication paradigms. The ROS 2 support for multi-robot agents is covered in the research. The authors look at the practical uses of ROS 2 in real-world circumstances in addition to the technical aspects. They go over case studies and instances of ROS 2 deployments that have been successful, showing how it has been used in industries like manufacturing, healthcare, agriculture, and autonomous cars. The study most likely offers insights into the difficulties encountered in these deployments and how ROS 2 resolves them, taking into account issues with interoperability, security, and safety. The research is significant because it offers a thorough understanding of ROS 2, which has become a crucial framework for the creation of robotic systems. The paper is a significant resource for researchers, developers, and practitioners interested in leveraging ROS 2 for their robotic projects by outlining its design principles, architecture, and real-world applications. It clarifies the improvements achieved in ROS 2 and emphasises their significance in enabling sophisticated and varied robotic systems in practical settings. Robots may now perform autonomous navigation using deep reinforcement learning (RL), according to Mauro Martini et al.'s [27] introduction of a modular framework called PIC4rl-gym. This work will be made accessible as an arXiv preprint in 2022. Within the Robot Operating System 2 (ROS 2) ecosystem, the authors want to overcome the difficulties of creating and testing RL-based navigation algorithms on robotic platforms.

For autonomous Navigation RL algorithms were tested using modular framework. Modules for perception, state representation, action execution, and reward feedback are probably included. The design of the framework, interactions between the various modules, and procedures for training and assessing RL agents in virtual or actual contexts may all be covered in the article. The PIC4rl-gym framework's usefulness in enabling autonomous navigation with deep RL was probably tested by the authors. Training RL agents in virtual environments and applying the learned policies to actual robots may be part of the evaluation process. Success rates, efficiency of navigation, and generalisation potential may all be performance indicators. The integration of the framework with current ROS 2 tools and libraries may also be included in the paper. The research is significant because it offers a modular architecture that makes it easier to create and assess RL-based navigation algorithms within the ROS 2 environment. PIC4rl-gym enables researchers and developers to use deep RL for autonomous navigation challenges on robotic systems by providing standardised interfaces and components. Vittorio Mayellaro [28], submitted at Politecnico di Torino in 2022, focuses on the development of a robotic system capable of autonomously navigating indoors while being aware of the presence of people. The framework's flexibility and modularity also support customization and adaptation to different robot architectures and environments. The goal of the research is to tackle the unique difficulties associated with enabling a sanitising robot to function safely and productively in dynamic contexts where human presence is an important factor. A summary of the issue statement and the significance of person-aware navigation for preserving safety and enhancing the adoption of robotic systems in public settings are probably the first two sections of the dissertation. The field of robotics' literature and methods for detecting, following, and interacting with people may be discussed by the author. The Robot Operating System 2 (ROS 2) framework may be used by the author's study technique in order to develop the autonomous navigation system. The design and implementation of perception algorithms to find and follow individuals, as well as their integration with the robot's navigation system, may be covered in detail in the dissertation. The author probably ran experiments in true-to-life interior settings to gauge how well the developed technology performed. The experimental setup, including the hardware and sensors used, as well as the metrics used to gauge the system's performance in person-aware navigation, may be included in the dissertation. The author may also point out any difficulties that arose throughout the tests and suggest possible directions for future development. The research is significant because it addresses the requirement for person-aware autonomous navigation in the setting of an indoor sanitising robot. The research contributes to the larger objective of integrating robots into public settings for different uses by taking into account the presence of people and ensuring safe and socially acceptable interactions. The dissertation probably offers suggestions and insights for integrating person-aware navigation technologies into similar robotic platforms. Noé Pérez-Higueras et al.'s introduction of the HuNavSim simulation framework, set to be available as an arXiv publication in 2023, aims to tackle the challenges of assessing human-aware robot navigation algorithms within the Robot Operating System

2 (ROS 2) ecosystem. While the research proposes a valuable tool for evaluating navigation algorithms, critical scrutiny is essential to understand its practical implications and limitations. Although the authors outline the importance of addressing human presence in navigation scenarios and discuss the simulation environment's components, the paper lacks detailed discussion on potential limitations or biases in the simulation models used, particularly regarding human behavior simulations. Furthermore, while the paper briefly mentions the benchmarking features and potential applications of HuNavSim, it overlooks potential challenges in accurately assessing algorithm performance and generalizing results to real-world scenarios. Additionally, while the paper may highlight advantages of the framework, it lacks critical evaluation of its shortcomings or potential areas for improvement. Without addressing these concerns, the practical utility and reliability of HuNavSim for evaluating human-aware navigation algorithms may be questioned, particularly in scenarios with complex human-robot interactions and diverse environmental conditions. Shu-juan Fang's research [30] on the simulation study of an indoor orbital inspection robot for the International Conference on Intelligent Traffic Systems and Smart Cities (ITSSC 2021) utilizing the Gazebo simulation platform presents a valuable opportunity to assess the effectiveness of robotic systems for indoor inspections. While the research offers insights into the potential uses of such robots and the challenges associated with their deployment in enclosed spaces, critical examination is necessary to understand its practical implications and limitations fully. Despite the description of the robot's concept and its applications in various fields, the paper lacks detailed discussion on potential limitations or biases in the simulation models used, particularly regarding the representation of complex indoor environments. Furthermore, while the author may discuss the implementation of control algorithms and navigational plans within the simulation framework, there is a lack of critical evaluation of the simulation's ability to accurately represent real-world scenarios and conditions. Additionally, while simulation studies may have been conducted to evaluate the robot's performance, the paper lacks detailed analysis of performance metrics and comparisons with real-world deployment scenarios. Without addressing these concerns, the

practical utility and reliability of the simulation study for informing the design and deployment of indoor inspection robots may be questioned, particularly in scenarios with diverse indoor environments and inspection requirements.2.5 Research papers based on infrared data utilization

2.5 Research papers based on infrared data utilization

Matteo et al.'s [31] exploration of cross-spectral neural radiance fields presents an intriguing approach to scene representation and synthesis, yet critical examination reveals certain limitations and areas for improvement. While the researchers aim to enhance the processing capabilities of neural radiance fields for multi-modal input data, the paper lacks detailed discussion on potential challenges and drawbacks associated with this approach. Despite outlining the difficulties current techniques face in handling diverse spectrum inputs, there is limited exploration of potential biases or inaccuracies introduced by the cross-spectral representation. Furthermore, while the advantages of combining various spectral modalities are highlighted, there is a lack of critical analysis on the practical implications and limitations of the proposed method for real-world applications, such as scene interpretation in challenging environments or under varying lighting conditions. Additionally, the paper may benefit from more comprehensive evaluation and comparison with existing techniques, including qualitative and quantitative assessments of scene synthesis quality and performance across different spectral domains. Without addressing these concerns, the practical utility and reliability of cross-spectral neural radiance fields for diverse applications may be questioned, limiting their potential impact in fields such as computer vision, robotics, and remote sensing. The Conference on Control, Automation, and Systems paper by Sooyong Lee and Jae-Bok Song [34] focuses on the localization of mobile robots utilising infrared light reflecting landmarks. In order to improve the localization precision of mobile robots in indoor situations, the researchers suggest a technique that makes use of landmarks with infrared reflectors. The first section of the study will probably cover the significance of precise localization for mobile robots as well as the difficulties presented by interior environments, such as the limited availability of GPS and inconsistent sensor data. The authors might draw attention to the necessity for further localization techniques that can get around these difficulties and produce accurate position estimates. The suggested method makes use of landmarks that reflect infrared light and are positioned at recognised locations in the surrounding area. With the help of infrared sensors, the mobile robot can identify landmarks and calculate their distances from one another. The infrared sensors' architecture and implementation, as well as the technique employed to glean landmark data from sensor observations, may be discussed by the authors. The researchers probably ran tests with a mobile robot inside areas with various arrangements of landmarks to gauge the efficacy of the proposed strategy. However, the experimental setup, including the hardware and sensor combinations employed, may be presented in the publication in more detailed manner. The metrics and evaluation criteria used to determine the localization accuracy attained by the authors' method may also be discussed. Quantitative evaluations of the localization accuracy attained by the suggested strategy in comparison to other localization techniques will probably be included in the findings and analysis portion of the study. The authors may go over the benefits and drawbacks of their approach, giving information about how well-suited and reliable it is for use in practical situations. The finding is significant because it provides a cutting-edge method for localising mobile robots using landmarks that reflect infrared light. The authors offer a technique that can improve the precision and dependability of robot localization in indoor settings where conventional methods might not be enough by utilising the special characteristics of infrared light. The work makes a contribution to the field of robotics and automation by offering an important method for enhancing the localization and navigation of mobile robots. Matteo Poggi et al.'s [35] proposal for synthesising images across multiple spectral domains through neural radiance fields is undoubtedly innovative, but a critical examination reveals several areas that warrant further consideration. While the researchers acknowledge the challenge of scarce data availability in specific spectral domains and propose a solution leveraging data from other domains, the paper may lack sufficient exploration of potential biases or limitations introduced by this approach. Although the significance of cross-spectral image synthesis is highlighted, there appears to be a gap in discussing the potential drawbacks or inaccuracies inherent in extrapolating data across different spectral regions. Moreover, while the method's design and training process are outlined, there is limited discussion on the robustness and generalizability of the approach across diverse datasets and real-world scenarios. The paper could benefit from more rigorous evaluation and comparison with existing techniques, including comprehensive quantitative assessments of image synthesis quality and performance in various spectral domains. Without addressing these concerns, the practical utility and reliability of the proposed method for synthesising images across multiple spectral domains may be questioned, potentially limiting its impact and adoption in fields such as surveillance, medical imaging, and remote sensing.

2.6 Research Gaps

The literature study listed above includes details on the work that has already been done in the fields of navigation and map representation. Researchers have used NeRF-based and neural network-based strategies to enhance the navigational process. Major research in the NeRF field demonstrates how NeRF is superior to other scene representation techniques [14], which may eventually lead to better map creation for robot navigation. In this approach, NeRF has been trained most frequently using visual cameras and related data. Also, utilization of NeRF technology with Blender simulator for generating simulation environment for an aerial robot [15] was remarkable and could be regarded as one of the inspirational steps towards using NeRF for SLAM and Autonomous Navigation. Another important contribution in the domain of NeRF for robotics was by M. Tancik, et al [24], where only RGB camera is utilized and NeRF is integrated with ROS2 for aerial robot. This study fills the gap of integrating ROS2 and Gazebo ignition simulator with NeRF technology for better and more realistic environment representation particularly for differential drive robots. Furthermore, assessment of different configurations and relative comparisons of multiple sensors (LiDAR, depth Camera, Infrared) for different indoor scenarios is another unique aspect of this study.

2.7 Research Synthesis Table

Paper	Year	Authors	Main Contribution	Limitations
Sivakumar et al. [6]	2021	Sivakumar, Arun Narenthiran, et al.	Propose a learned visual navigation approach for under-canopy agricultural robots	Limited to under-canopy environments; may not generalize well to other scenarios
Gupta et al. [7]	2017	Gupta, Saurabh, et al.	Introduce a method that unifies map and landmark-based representations for visual navigation	Does not consider dynamic environments or real-time adaptation
Mo et al. [8]	2018	Mo, Kaichun, et al.	Present the AdobeIndoorNav dataset for deep reinforcement learning-based real-world indoor robot visual navigation	Dataset may have limited diversity or coverage of indoor environments
Qiu et al. [9]	2020	Qiu, Yiding, Anwesan Pal, and Henrik I. Christensen.	Explore target-driven visual navigation by leveraging object relationships	Limited to target-driven scenarios; may not handle complex navigation tasks
Nguyen et al. [10]	2019	Nguyen, Khanh, and Hal Daumé III.	Propose a multimodal assistance approach for visual navigation using imitation learning	Relies on retrospective curiosity-encouraging imitation learning, which may have limitations in capturing diverse human behaviors and providing real-time assistance
Zhu et al. [11]	2021	Zhu, Fengda, et al.	Present a survey on deep learning techniques for	Does not provide novel contributions; limited to

			embodied vision navigation	summarizing existing approaches
Wang et al. [12]	2021	Wang, Haiyang, et al.	Introduce collaborative visual navigation techniques	Limited evaluation of collaborative navigation methods in real-world scenarios
Chen et al. [13]	2019	Chen, Kevin, et al.	Propose a behavioral approach to visual navigation using graph localization networks	Limited evaluation of the approach's performance in complex and dynamic environments
Mildenhall et al. [14]	2021	Mildenhall, Ben, et al.	Introduce the NeRF model for representing scenes as neural radiance fields	Limited discussion on practical limitations and challenges of implementing NeRF in real-world scenarios
Adamkiewicz et[15]	2022	Adamkiewicz, Michal, et al.	Present a vision-only navigation approach in a neural radiance world	Limited evaluation in real-world environments; may not handle dynamic or unknown obstacles
Tang et al. [16]	2022	Tang, Yang, et al.	Provide a survey on perception and navigation in autonomous systems	Does not propose novel contributions; limited to summarizing existing research
Sun et al. [17]	2022	Sun, Jiankai, et al.	Propose a transformer-based object localization method within neural radiance fields	Limited evaluation on challenging scenarios or complex objects
Byravan et al. [18]	2022	Byravan, Arunkumar, et al.	Propose Nerf2Real, a method for transferring vision-guided bipedal motion skills from simulation to the real world using neural radiance fields	Limited to bipedal motion skills transfer and may not generalize well to other types of robot motion
Gao et al. [19]	2022	Gao, Kyle, et al.	Provide a comprehensive review	Does not propose novel contributions;

			of neural radiance fields in 3D vision	limited to summarizing existing research
Maggio et al. [20]	2022	Maggio, Dominic, et al.	Introduce Loc-NeRF, a Monte Carlo Localization method using neural radiance fields	Limited evaluation in complex and dynamic environments
Tong et al. [21]	2022	Tong, Mukun, Charles Dawson, and Chuchu Fan	Propose an approach to enforce safety for vision-based controllers using Control Barrier Functions and neural radiance fields	Limited evaluation in safety-critical scenarios or real-world environments
Klenk et al. [22]	2023	Klenk, Simon, et al.	Introduce E-Nerf, a method for neural radiance fields using a moving event camera	Limited evaluation in real-world scenarios; may not handle dynamic and fast-moving objects well
Chen et al. [23]	2023	Chen, Timothy, Preston Culbertson, and Mac Schwager	Propose CATNIPS, a collision avoidance approach using neural implicit probabilistic scenes	Limited evaluation in complex and cluttered environments; may not handle real-time collision avoidance
Tancik et al. [24]	2022	Tancik, Matthew, et al.	Present Block-NeRF, a scalable neural view synthesis method for large scenes	Limited discussion on practical limitations and scalability challenges of Block-NeRF
Lin and Yi [25]	2022	Lin, Kevin, and Brent Yi	Introduce active view planning for radiance fields	Limited evaluation in dynamic environments or scenarios with unknown obstacles
Macenski et al. [26]	2022	Macenski, Steven, et al.	Discuss the design, architecture, and uses of Robot Operating System 2 (ROS 2)	Does not propose novel contributions; limited to discussing ROS 2
Martini et al. [27]	2022	Martini, Mauro, et al.	Introduce PIC4rl-gym, a ROS 2 modular	Limited evaluation of the framework's

			framework for Robots Autonomous Navigation with Deep Reinforcement Learning	performance and scalability
Mayellaro [28]	2022	Mayellaro, Vittorio	Present a person-aware autonomous navigation approach for an indoor sanitizing robot in ROS 2	Limited evaluation in diverse environments or real-world scenarios
Pérez-Higueras et al. [29]	2023	Pérez-Higueras, Noé, et al.	Introduce HuNavSim, a ROS 2 Human Navigation Simulator for Benchmarking Human-Aware Robot Navigation	Limited evaluation of the simulator's realism and generalizability
Fang [30]	2022	Fang, Shu-juan	Conduct a simulation analysis of an indoor orbital inspection robot based on Gazebo	Limited to simulation analysis and does not provide empirical results
Poggi et al. [31]	2022	Poggi, Matteo, et al.	Present Cross-Spectral Neural Radiance Fields, a method for synthesizing radiance fields from cross-spectral images	Limited evaluation in challenging lighting and environmental conditions
Hadsell et al. [32]	2009	Hadsell, Raia, et al.	Discuss the learning of long-range vision for autonomous off-road driving	Limited to off-road driving scenarios and may not generalize well to other environments
Chen et al. [33]	2018	Chen, Chen, et al.	Introduce a method for learning to see in the	Limited to low-light imaging scenarios and may not generalize well to other vision

			dark	tasks
Lee and Song [34]	2007	Lee, Sooyong, and Jae-Bok Song	Propose mobile robot localization using infrared light reflecting landmarks	Limited to localization using specific landmarks and may not be applicable in all environments
Poggi et al. [35]	2022	Poggi, Matteo, et al.	Discuss Cross-Spectral Neural Radiance Fields and their applications in synthesizing radiance fields from cross-spectral images	Limited to summarizing existing research and does not propose novel contributions

TABLE-1: RESEARCH SYNTHESIS

2.8 Scope and Limitation of Research

The research focuses specifically on enhancing robot navigation in indoor regions using LiDAR, depth camera and infrared cameras. The creation and assessment of SLAM methods and navigation algorithms using said sensors' data via NeRF are included in the scope. The Robot Operating System 2 (ROS2) framework and the Gazebo Ignition simulator are used in the research, development, testing, and assessment. In order to evaluate the effectiveness of the navigation systems, the scope includes utilising the functionalities offered by ROS2 and the realistic simulation environment of Gazebo Ignition. The research compares pixels and voxel-based simulation of world with NeRF based simulated continuous representation of the real world. The scope includes assessing how well various strategies work in terms of avoiding obstacles, safety, and time lag. The comparison sheds light on the advantages and disadvantages. The focus is on building navigation system that can function well in situations where restricted visibility presents difficulties for robots in avoiding obstacles.

The research's conclusions and findings may be limited to the specific settings and circumstances considered in the study, potentially restricting their application to other contexts with differing features or challenges. To ensure broader applicability, real-world trials should be integrated into the research to validate performance across various settings. Additionally, the study's reliance on simulated sensors introduces constraints and potential limitations related to sensor accuracy and range, which can impact data quality and navigation system effectiveness. It's essential to acknowledge these sensor restrictions and account for potential noise or interference in the data. Furthermore, the computational intensity of NeRF training presents another limitation, necessitating consideration of computational resources and hardware limitations during real-time navigation. Optimal results may require high-power GPUs or TPUs due to the computational burden involved. Simulated settings may also fail to accurately represent the complexity and variability of real-world conditions, including differences in lighting, object appearances, and environmental elements. Such discrepancies can affect navigation system performance, highlighting the importance of considering real-world variability in the research. Additionally, while the study's findings provide valuable insights, their generalizability to other robotic platforms, sensor configurations, or specific application areas may vary. Further research is needed to determine the extent to which these findings apply to different contexts, emphasizing the importance of clearly defining the research scope. Understanding and addressing these limitations are essential when analysing the research findings and considering their potential implications for improving navigation systems in real-world scenarios.

2.9 Research Objectives

The scope of the research is very broad but considering the limited time constraint of MS thesis following Objectives are considered:

1. To analyse different navigation scenarios that will help a robot navigate across challenging terrain.

2. Examine NeRF generated scene while travelling through low-light conditions or indoor areas using different sensor combinations.

3. Choosing the best sensor configuration and software setup for navigation in various contexts by considering near real world scenario of Navigational Environment during R&D phase.

4. Filling the previous research gap related to undertaking NeRF implementation with ROS2, Gazebo Ignition for differential drive robots with multiple sensors configurations.

Chapter 3: Research Methodology

3.1 Research issues

The research endeavours to determine the optimal robotic setup, encompassing both software and hardware components, tailored specifically for navigating indoor terrain. To achieve this goal, a series of key issues will be addressed. Firstly, the study will focus on creating a realistic simulation environment using Gazebo Ignition software capable of accurately replicating varying levels of lighting intensity characteristic of dark areas. Secondly, methods for generating diverse terrains and obstacles within the simulation will be explored to comprehensively assess the navigation capabilities of the robotic system across different scenarios. Thirdly, the research will delve into the development of the robotic system's functionality using ROS2 and its synchronization with Gazebo Ignition for seamless simulation and testing. Additionally, efforts will be made to identify and integrate optimal plugins for infrared, depth camera, and LiDAR sensors with the robotic platform to enhance environmental perception and navigation precision. Furthermore, the study will address the methodology for manually navigating the robot within simulated environments to collect sensor data crucial for map generation. Moreover, the utilization of various NeRF modules to construct detailed 3D scenes from real-world data and their integration into the Gazebo Ignition simulator will be investigated. Finally, a comparative analysis will be conducted to determine the most suitable configuration model based on performance metrics and requirements, thus providing valuable insights into the ideal robotic setup for navigating low light terrains effectively. Through tackling these research questions, the study aims to advance robotic navigation technology, particularly in challenging environmental conditions, ultimately enhancing the performance and reliability of autonomous robotic systems.

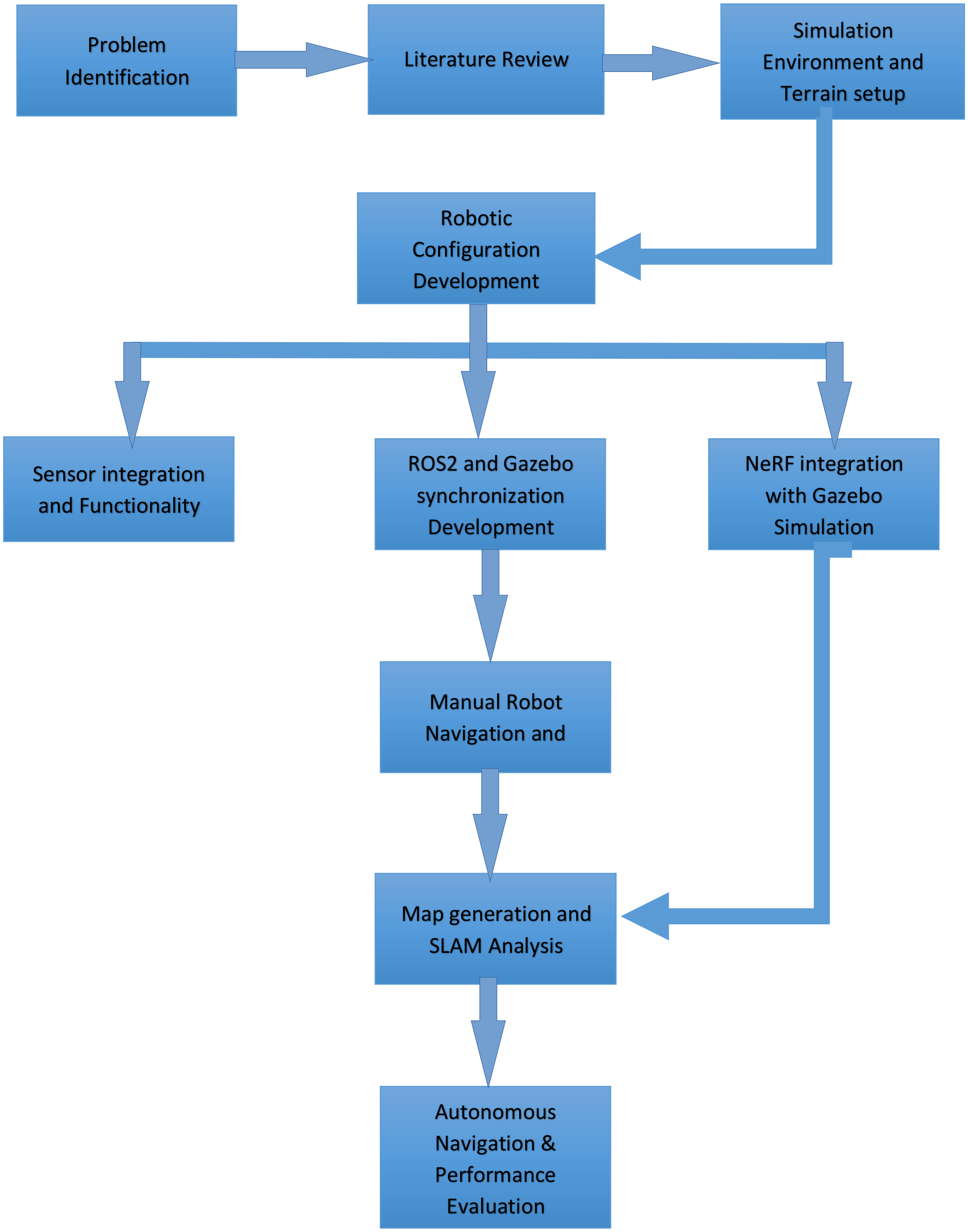
3.2 Research Methodology

The goal of the study is to improve a robot's visual navigation abilities in low-luminance or indoor environments. This is done by starting with a clear problem statement. The study's main goal is to fill a knowledge gap regarding the usage of neural radiation fields (NeRF) with diverse sensors for navigation. A simulation environment utilising the Gazebo Ignition simulator is built up to conduct the experiments. In this regard, both built-in gazebo world (ideal world) and NeRF based world integrated with gazebo have been considered given at Annex- “A”. This makes it possible to design two different terrains with obstacle. Using different robotic configurations discussed earlier, the robot platform is manually manoeuvred while environmental data and impediments are captured during SLAM operation for generating Maps to assist Navigation. Conclusively, six maps are generated based on different terrains and robotic configurations. The operation of SLAM is analysed on the basis of quality of occupancy grids as it’s a qualitative measure so marking is done on the basis of 1-5 scale, where 5 regarded as best and 1 regarded as poor.

Afterwards, Autonomous Navigation of the robot is performed and analysed based on provided Maps. In his operation capability of robot to avoid obstacle while reaching multiple destination is evaluated, it’s also an ability that can be measured qualitatively by watching recorded videos of Performed Experiment so it is also based on the same scheme adopted for map assessment. In this respect, linear & angular velocities are measured and assessed quantitively and configuration with highest average linear velocity is considered as most suitable and vice-versa. Finally approximate least time elapsed during Navigation by any configuration is considered as a favourable factor.

3.3 Research Process Flowchart

A more detailed flowchart illustrating the research process described in the above paragraph is mentioned on the following page:


Problem Identification
Literature Review
Simulation Environment and Terrain setup
Robotic Configuration Development
Sensor integration and Functionality
ROS2 and Gazebo synchronization Development
NeRF integration with Gazebo Simulation
Manual Robot Navigation and
Map generation and SLAM Analysis
Autonomous Navigation & Performance Evaluation

Chapter 4: Tools and Techniques

4.1 ROS2

For creating robot software Robot Operating System 2, or ROS2, is a standardized platform. A set of tools, libraries, and conventions are offered as part of this open-source effort to make it easier to design sophisticated robotic systems. Performance, scalability, and robustness issues that ROS previously encountered have been resolved by upgrading the ROS1 to ROS2.

4.1.1 Relevance of ROS2

Modularity: ROS2 has a modular architecture that enables developers to create and integrate discrete parts known as nodes that interact with one another via a publish-subscribe messaging system. This modular design makes it possible to reuse, maintain, and scale code. ROS2 facilitates the creation of distributed systems, in which nodes can run on several machines and interact over a network. Building large-scale robotic systems that need distributed computing capabilities will benefit from this feature. **Interoperability:** ROS2 supports a variety of programming languages, enabling programmers to create nodes in their chosen language (such as C++, Python, and others) and yet have them communicate with each other without any issues. The collaboration and integration of various software libraries and components are encouraged by this interoperability. Real-time Capabilities: ROS2 adds real-time features that make it possible to build robots with precise timing and control requirements. This is especially crucial in fields or applications that require low-latency communication, such as safety-critical ones. Industrial Status of Robot Development: ROS2 has been widely used in both research and industrial contexts and has made great progress in the field of robotics. Advanced robotic systems are being developed utilising ROS2 by a significant number of robotics companies, including both established businesses and start-ups. ROS2's improved performance, scalability, and real-time capabilities, which make it suited for a variety of applications like industrial automation, autonomous vehicles, healthcare robotics, and more, are what are driving its industrial adoption. ROS2 offers a variety of navigation-related features and packages that are necessary for robot navigation tasks. ROS2's main navigational attributes are as follows:

Map-Making: ROS2 offers mapping packages [42] like g-mapping and mapper that let robots map their surroundings using sensor data (such as LiDAR, cameras, etc.). SLAM toolbox is an innovative breakthrough that allows mapping of highly vast and dynamic environments. The discussed tool is based on appreciable algorithms [43] based on “Efficient spares pose adjustment based on 2D Mapping”. With autonomous cars, SLAM (simultaneous localization and mapping) allows you to simultaneously create a map and locate your vehicle on it. Through the use of SLAM techniques, the vehicle can map out uncharted territory. Engineers utilize the map data to complete jobs like route bulding for operations. Preparing and avoiding roadblocks. For many years, SLAM has been the focus of technical study. However, thanks to significant advancements in computer processing speed and the accessibility of inexpensive sensors like cameras and laser range finders, SLAM is being applied practically in an increasing variety of sectors. Many other uses for SLAM exist, such guiding a group of mobile robots to arrange shelves in a warehouse, parking a self-driving car in a vacant space, or using a drone to deliver a gift in an unfamiliar area. In general, two categories of technological elements are employed to accomplish SLAM. The first kind, which mostly depends on the sensors being utilized, is sensor signal processing, which includes front-end processing. Pose-graph optimization, which includes sensor-agnostic back-end processing, is the second kind [44].

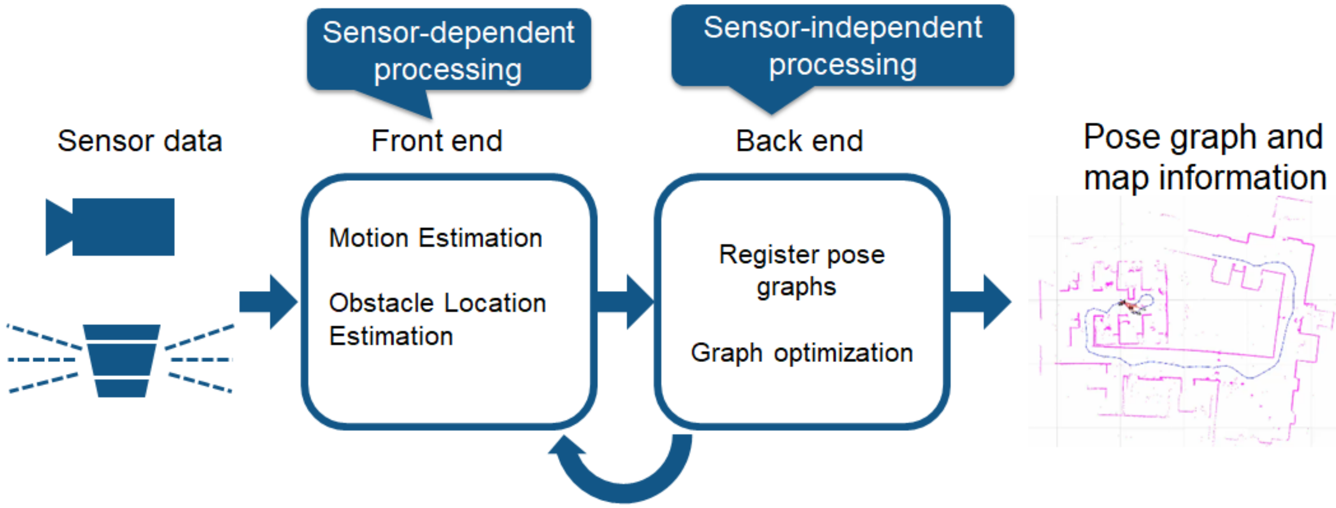


Figure-1: TYPICAL SLAM PROCESS FOR MAP BUILDING [44]

visual SLAM (or vSLAM) uses images acquired from cameras and other image sensors. Compound eye cameras (stereo and multiple cameras), RGB-D cameras (depth and ToF cameras), and basic cameras (wide angle, fish-eye, and spherical cameras) can all be used with visual SLAM. Cheap cameras allow for the low-cost implementation of visual SLAM. Additionally, cameras can be used to recognize landmarks (already measured positions) because to their vast volume of information. Graph-based optimization and landmark identification can be used to achieve flexible SLAM implementation. When vSLAM only employs one camera as a sensor, it's known as monocular SLAM, which makes depth definition difficult. This can be resolved by either merging the information from the camera with another sensor, like an inertial measurement unit (IMU), which can measure physical characteristics like velocity and orientation, or by identifying AR markers, checkerboards, or other known items in the image for localization. Structure from motion (SfM), visual odometry, and bundle adjustment are examples of vSLAM-related technology. Two general categories can be used to group visual SLAM systems. Sparse approaches use algorithms like PTAM and ORB-SLAM to match feature points in images. Dense techniques employ algorithms like DTAM, LSD-SLAM, DSO, and SVO and rely on the overall brightness of the images [44]. Lidar is a technique that predominantly relies on a laser sensor (or distance sensor). Lasers are considerably more accurate than cameras, ToF, and other sensors. They are employed in applications involving fast-moving vehicles like self-driving automobiles and drones. The laser sensors typically provide 2D (x, y) or 3D (x, y, z) point cloud data as output values. The laser sensor point cloud offers accurate distance measurements and is highly efficient for map creation using SLAM. In general, the process of estimating movement involves systematically comparing and aligning point clouds. The measured displacement (distance traveled) is utilized for the localization of the vehicle. Lidar point cloud matching involves the utilization of registration procedures, such as the iterative closest point (ICP) and normal distributions transform (NDT) algorithms. Point cloud maps, whether 2D or 3D, can be depicted as either a grid map or a voxel map. However, point clouds lack the same level of detail as images in terms of density and may not always offer enough distinctive characteristics for matching purposes. In locations with few obstructions, the process of aligning the point clouds becomes challenging, perhaps leading to the loss of the vehicle's position tracking. Furthermore, point cloud matching typically necessitates substantial computational resources, therefore necessitating the optimization of processes to enhance performance. In light of these difficulties, the process of localizing autonomous cars may include the integration of other measurement outcomes, such as wheel odometry, global navigation satellite system (GNSS), and IMU data. Warehouse robots often utilize 2D lidar SLAM, while UAVs and autonomous driving systems can employ SLAM with 3-D lidar point clouds[44]. A point cloud is a set of data points in three-dimensional space, where each point represents the X, Y, and Z coordinates of a spot on the surface of a real-world item. Together, these points create a complete map of the object's surface [45]. Lidar scanners, stereo cameras, and time-of-flight cameras are frequently used to generate point clouds.

Point clouds can be categorized into two groups based on the type of data they contain:

Structured point clouds are organized in a grid-like fashion, similar to how image data is stored. The format is M x N x C, where M represents the number of rows, N represents the number of columns, and C represents the number of channels. Typically, stereo cameras and time of flight cameras provide structured point clouds. It provides details regarding the connection between neighboring points.

Unorganized point clouds are not organized into rows and columns. The Format is MxC, where M is the number of points in the point cloud and C number of channels. Typical lidar sensors produce unorganized point clouds. An unorganized point cloud can be converted to an organized point cloud by projecting onto a sphere.

Localization: Robots can estimate their pose (position and orientation) inside a specified map using localization packages provided by ROS2, such as AMCL (Adaptive Monte Carlo Localization). For proper navigation and path tracking, localization is essential.

Autonomous Navigation/Route planning: Move base, one of the packages in ROS2 [46], uses global and local route planning algorithms to produce a path that will get the robot to its goal without colliding with other objects. The robot's present attitude, the environment's map, and any obstructions picked up by sensors are all taken into account while determining the robot's route. ROS2 offers tools and algorithms to facilitate the detection and avoidance of obstacles during robot navigation. It is based on well-known algorithms such A* or Djikstras expansion. In order to avoid collisions, the robot must be able to detect objects in its environment using sensor data from devices like cameras or LiDAR. Robot navigation tasks are easier to regulate and carry out thanks to ROS2. It enables the integration of controllers, such as model predictive controllers or PID controllers, to manage the robot's mobility and provide accurate and seamless navigation.

Overall, the navigation capabilities and tools provided by ROS2 allow programmers to create complex robot navigation systems. A robust platform for creating and implementing navigation algorithms in a variety of robotic applications, ROS2's modular and distributed architecture, interoperability, and real-time capabilities [36] all contribute to this framework's effectiveness.

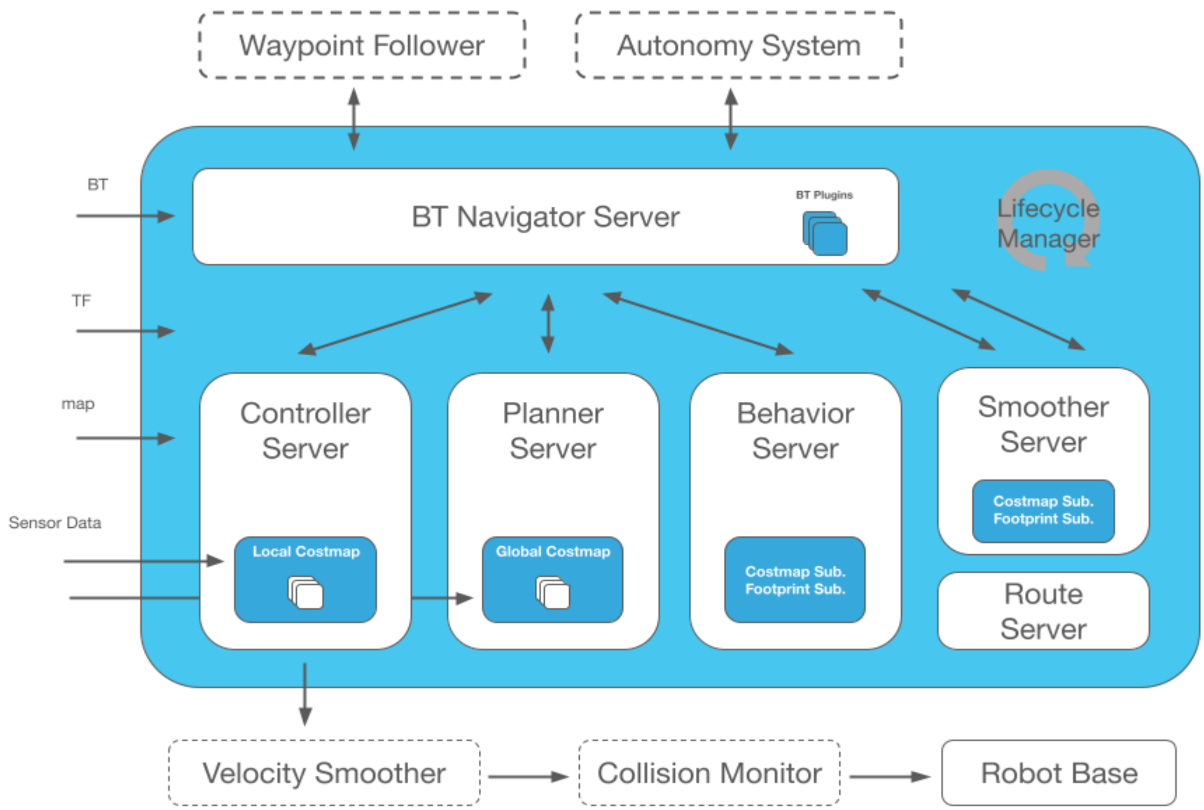


Figure-2: ARCHITECTURE OF NAVIGATION SYSTEM IN ROS2 [46]

Navigation2 [47] in ROS2 employs a modular node-based structure to handle various aspects of robot navigation. The core nodes include the planner server for high-level path planning, controller server for low-level control, and recoveries for managing unexpected situations. It relies on a 2D cost map generated by the costmap_2d node, incorporating data from sensors like laser scans. Localization is handled by the AMCL node, estimating the robot's pose based on sensor input and a known map. A behaviour tree manages decision-making processes, and there are distinct global and local planners for path determination and motion control. The map server provides the map to other nodes, and communication between nodes is facilitated by ROS2 interfaces. The architecture is designed to be extensible through a plugin system, allowing customization of different components, and dynamic reconfiguration enables parameter adjustments during runtime.

4.2 Gazebo Ignition

A robust and well-liked physics-based robot simulator is Gazebo Ignition. It is an improvement over the Gazebo simulator, offering more capabilities and better performance for simulations of robotics that are actually realistic. Gazebo ignition is a crucial tool for navigation simulation since it is made to faithfully replicate robot dynamics, sensors, and interaction with the environment.

Observable Effects of Gazebo Ignition are as under:

Actual Simulation: A physics engine from Gazebo Ignition correctly simulates robot behaviour and interactions with the environment. Developers can test and validate navigational algorithms using a variety of supported sensors, including cameras, LiDAR, and infrared sensors.

Environment Customization: For building unique and sophisticated simulated settings, Gazebo Ignition provides a wide range of customization possibilities. Users can import real-world maps and models as well as define their own terrains, structures, and other items. Due to their adaptability, navigation simulations can be performed in both indoor and outdoor settings with varying lighting.

Sensor Integration: The simulation environment may easily incorporate sensors thanks to Gazebo Ignition. To assess the effectiveness of navigation algorithms, developers might mimic sensor data such as camera images, LiDAR point clouds, or infrared readings. Prior to their use in the real world, robot navigation systems must be tested and optimised using this capacity.

ROS connection: Developers may take advantage of the extensive ecosystem of ROS packages for robot navigation thanks to Gazebo Ignition's great connection with ROS and ROS2. It enables smooth communication and control between the simulated robot and external ROS nodes by supporting data interchange between Gazebo and ROS.

Gazebo ignition is commonly used in the robotics industry for navigational simulation. Industrial Status in Robot Navigational Simulation It is used to test robot behaviours, validate navigation algorithms, and simulate sophisticated robotic systems by robotics researchers, developers, and businesses. In a variety of industries, including autonomous vehicles, industrial automation, and service robotics, Gazebo Ignition is the preferred option for simulating and evaluating robot navigation due to the availability of realistic physics-based simulations and the integration capabilities with ROS and ROS2. It is a significant asset in the field of robot navigational simulation due to its broad community support, ongoing development, and integration with other robotic technologies [37].

4.3 Neural Radiance Fields (NeRF)

NeRF (Neural Radiance Fields) is a technique used in computer vision and robotics for 3D scene representation and view synthesis [14]. It allows for the generation of highly detailed and realistic 3D scenes from 2D images or sparse 3D data. NeRF models the scene as a continuous volumetric function, capturing both the geometry and appearance of the objects in the scene.

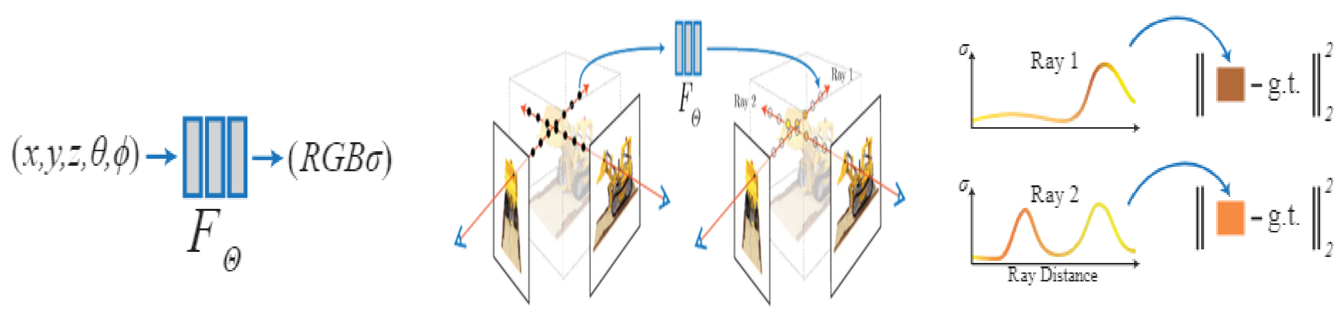


FIGURE-3: Neural Radiance Fields (NeRF) architecture [14]

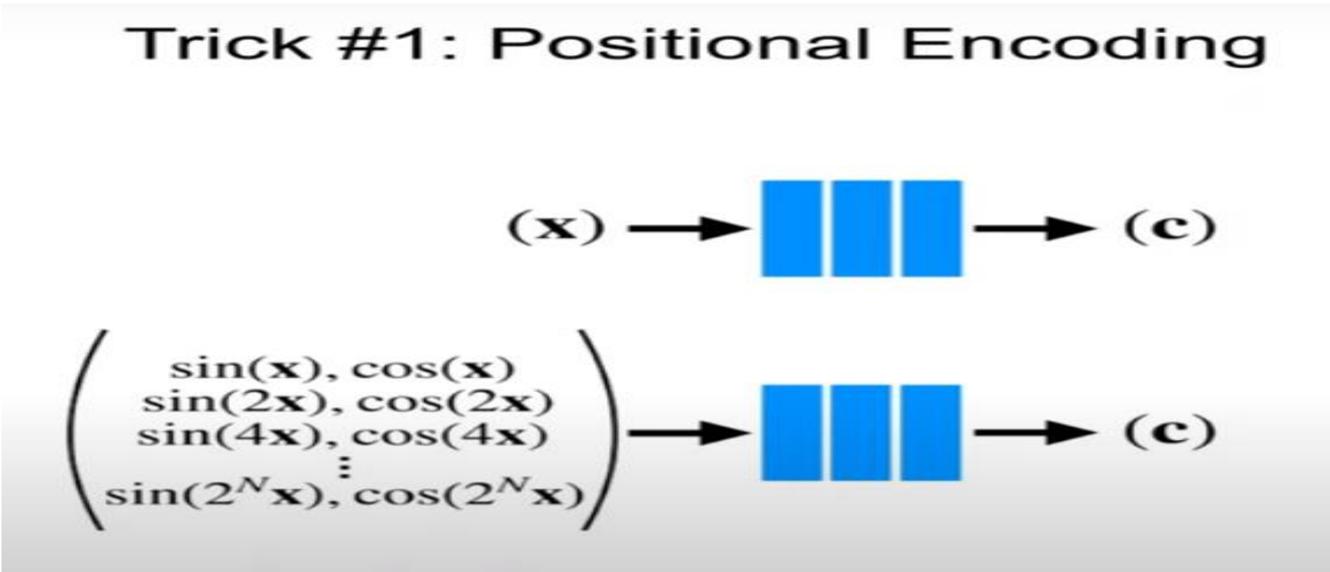


FIGURE 4: POSITIONAL ENCODING TECHNIQUE FOR NeRF [14]

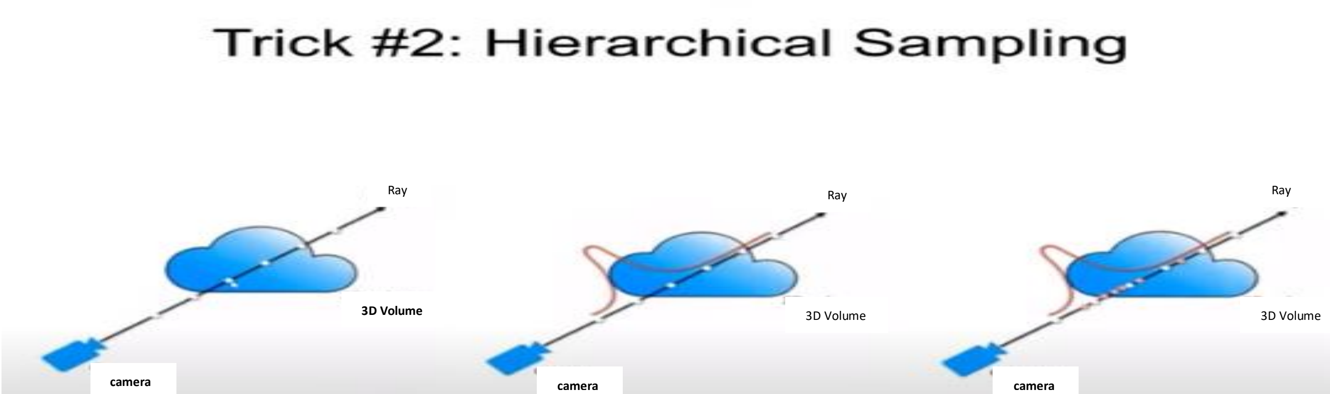


FIGURE 5: HIERARCHIAL TECHNIQUE FOR NeRF [14]

NeRF (Neural Radiance Fields) is a method for 3D scene modelling and view synthesis in computer vision and robotics [14]. It permits the creation of incredibly accurate and detailed 3D scenes from 2D photos or limited 3D data. NeRF captures the geometry and look of the scene's objects by modelling the scene as a continuous volumetric function. By increasing the dimensionality of the input through the Fourier transform, which improves the performance of the neural network [38], and by using hierarchical sampling of the input points to produce a smart selection of data rather than the entire scene [39], Figures 2 and 3 show the two crucial steps taken to improve the conventional volumetric rendering technique. NeRF's capacity to produce fresh views of the scene, allowing the robot to see its surroundings from various angles without physically moving, is one of its key advantages in the context of robot navigation [15]. This view synthesis capacity allows the robot to predict and reason about its environment, which is very helpful for tasks like path planning, obstacle avoidance, and scene understanding. Robots can benefit from more precise and thorough scene representations by adopting NeRF for navigation, which will enhance perception and decision-making abilities [15]. NeRF can help a robot move more safely and effectively by improving its capacity to localise itself in the environment, comprehend the 3D structure of the surroundings, and understand their layout. NeRF also has benefits in terms of computing efficiency [14]. NeRF operates on a continuous function as opposed to conventional grid- or voxel-based representations, which negates the requirement for grid-based computations and discretization. Because of this, processing is quicker and more effective, making it appropriate for real-time applications like robot navigation. NeRF's capacity to offer highly accurate scene representations, enable view synthesis for all-encompassing vision, and increase computational efficiency is what makes it significant for robot navigation overall [15]. Robots can travel in complex and dynamic situations with greater situational awareness and decision-making abilities by utilising NeRF.

4.4 Light intensity for different Regions[41]

Condition (Outdoor)	Illumination (lux)	Condition (Indoor)	Illumination (lux)
Full Daylight	10752	Public areas with dark surroundings	20-50
Sunlight	107527	Simple orientation for short visits	50-100
Overcast Day	1075	Areas with traffic and corridors - stairways, escalators and travellators - lifts - storage space	100
Very Dark Day	107	Working areas where visual tasks are only occasionally performed	100-150
Twilight	10.8	Warehouses, homes, theatres, archives, loading bays	150
Deep Twilight	1.08	Coffee break room, technical facilities, ball-mill areas, pulp plants, waiting rooms	200
Full Moon	0.108	Easy office work	250
Quarter Moon	0.0108	Class rooms	300
Starlight	0.0011		
Overcast Night	0.0001		

TABLE 2. ILLUMINATION VALUES

Chapter 5: COMPARATIVE ANALYSIS AND RESULTS

This Section comprises of details regarding the six different experiments performed for the subject research study. In every experiment SLAM and Autonomous Navigation operations of the differential drive robot were conducted for analysis and comparison among the experiments. First three experiments are performed in an ideal simulated house with ground, walls and obstacles built entirely in Gazebo simulator as shown at Appendix- ‘A’. The aforementioned house and its obstacles are combination of multiple grids and pixels.

Last three experiments were conducted in an environment that consists of floor and few obstacles taken from real world then trained by the NeRF (Neural Radiance Fields) system to include every single detail and making the experience as near as possible to the real world. For immediate comparison some area of ground and obstacle are directly taken from Gazebo software as well. The world for last three experiments is shown at Appendix- ‘B’. Initially, for every experiment, map is built during SLAM operation and then recorded map becomes the base for respective Autonomous Navigation. In this regard, comparison of linear and angular velocities given by command velocity node and followed by Autonomous Navigation node respectively are considered as suitable parameters for assessment. Secondly, ability of the robot to avoid obstacle is another important factor that is accounted, it can better be visualized by recorded videos of navigation against each Experiment. Third, quality of different maps is examined visually. Fourth, time elapsed in Autonomous Navigation by a robot for a particular sensor configuration and environmental conditions of simulated worlds are noted through videos. All the results are mentioned for each experiment in the form of graphs, charts, figures, tables and videos. All aforementioned parameters are examined for getting an insight and establishing a logical comparison among the six-experiment performed and ultimately to establish a conclusion at the end of this section.

5.1 Experiment-1

This experiment is based on the simulation of a wooden house as discussed earlier. In this experiment the configuration of the robot is mainly based on utilization of laser camera for scanning during SLAM for map building. Afterwards the Autonomous Navigation is performed on the basis of recorded map.

5.1.1 Experiment-1 SLAM

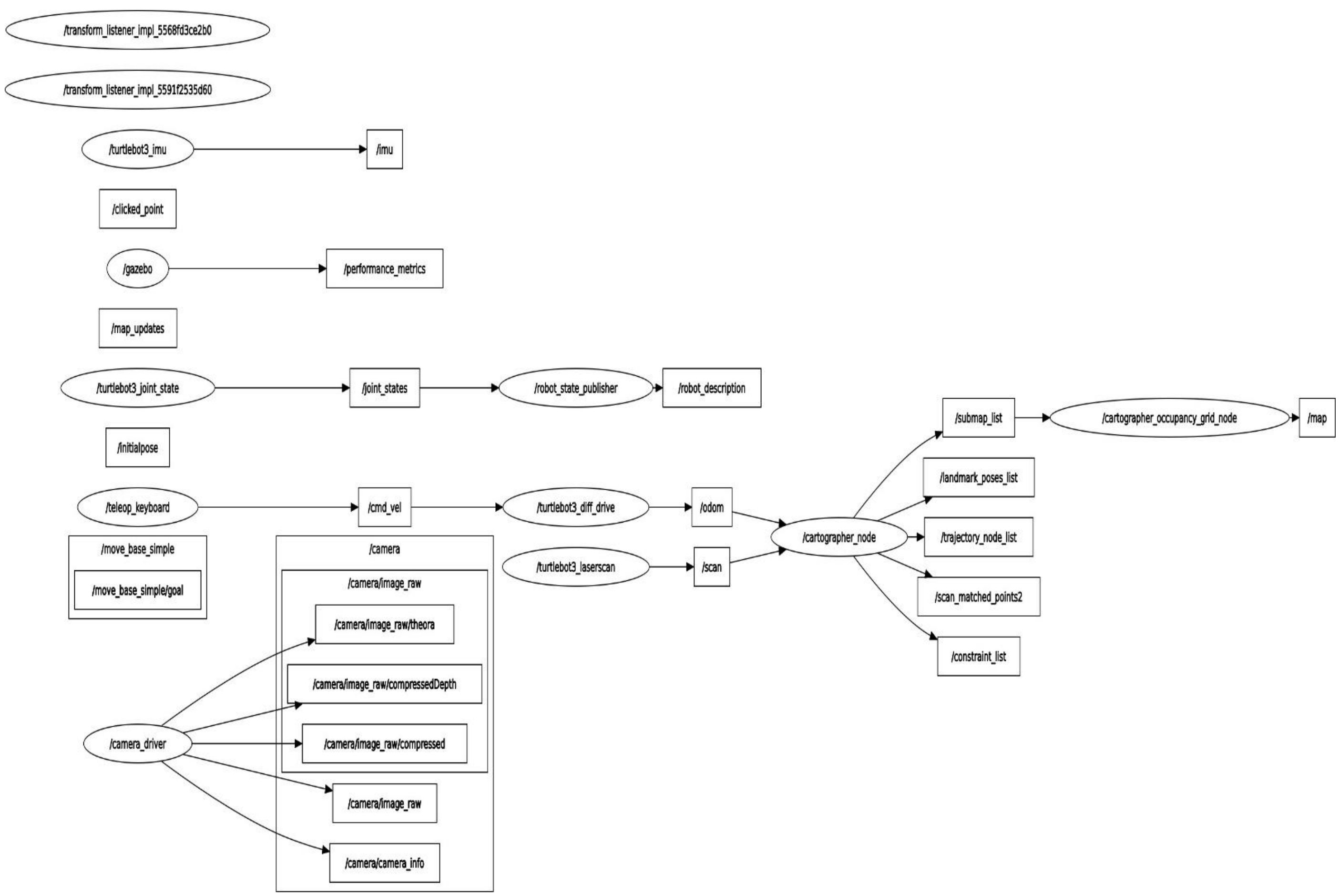


FIGURE-6: GRAPH OF SLAM OPERATION EXPERIMENT-1

The given graph reflects the flow of activities during the SLAM operation performed in Experiment-1. The rectangular blocks in the graph are topics from or to which the data is published or subscribed by the nodes. Mainly it consists of Gazebo package connecting with ROS through an inherent bridge. Also, it shows how the joint states of the robot is being published and subscribed for assessment and description of differential drive robot status at any particular instant. Also, teleop_keyboard node alongwith its topics is making the movement of the robot possible through velocity and direction commands given by the user for scanning the area. Odometry data and laserscan data is being utilized by the cartographer node for making maps with publishing recording its performance parameters for map building by cartographer occupancy_grid_node. Simple camera node is being installed for just checking the direction of robot movement during zoom out mode of SLAM activity without any image recording.

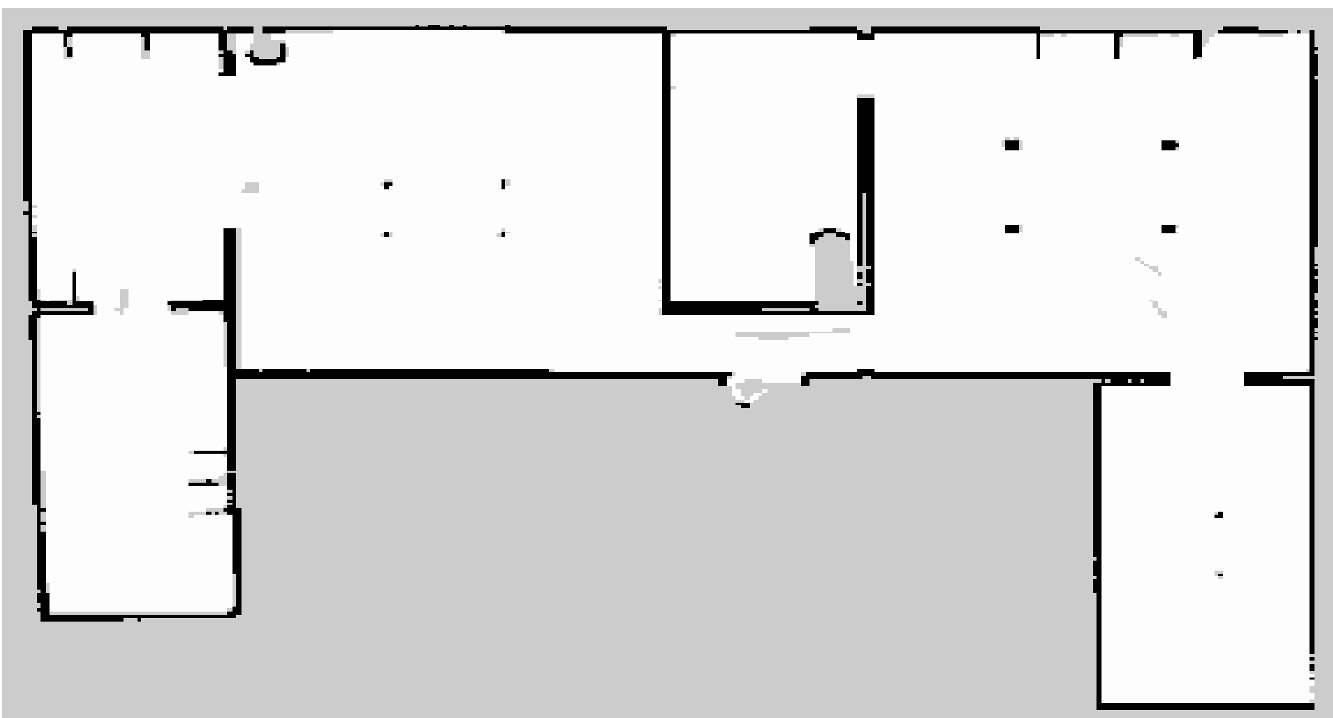


FIGURE-7: MAP GENERATED EXPERIMENT-1

Figure-7 shows the map of house laser scanned by the robotic configuration utilized in Experiment-1. White area shows the areas without obstacle that are safe for robot navigation. Whereas, dark areas represent walls, base of tables and obstacles as placed at various location of the houses. Overall map scanned by the laser is of high quality with minimum missed scanned areas.

5.1.2 SAMPLE SLAM DATA

```
image: laser_normal.pgm
mode: trinary
resolution: 0.05
origin: [-8.15, -6.1, 0]
negate: 0
occupied thresh: 0.65
free thresh: 0.25
```

5.1.3 EXPERIMENT-I AUTONOMOUS NAVIGATION

Autonomous Navigation related to experiment-1 is based on the map acquired during the SLAM phase of Experiment-1 as discussed earlier. The flow of operation is enclosed herewith as Appendix- 'E'.

Enclosed graph shows the too complex functionality of different nodes and topics for making the autonomous navigation successful. In this regard a simplified way of understanding the key concepts behind this operation is given as under:

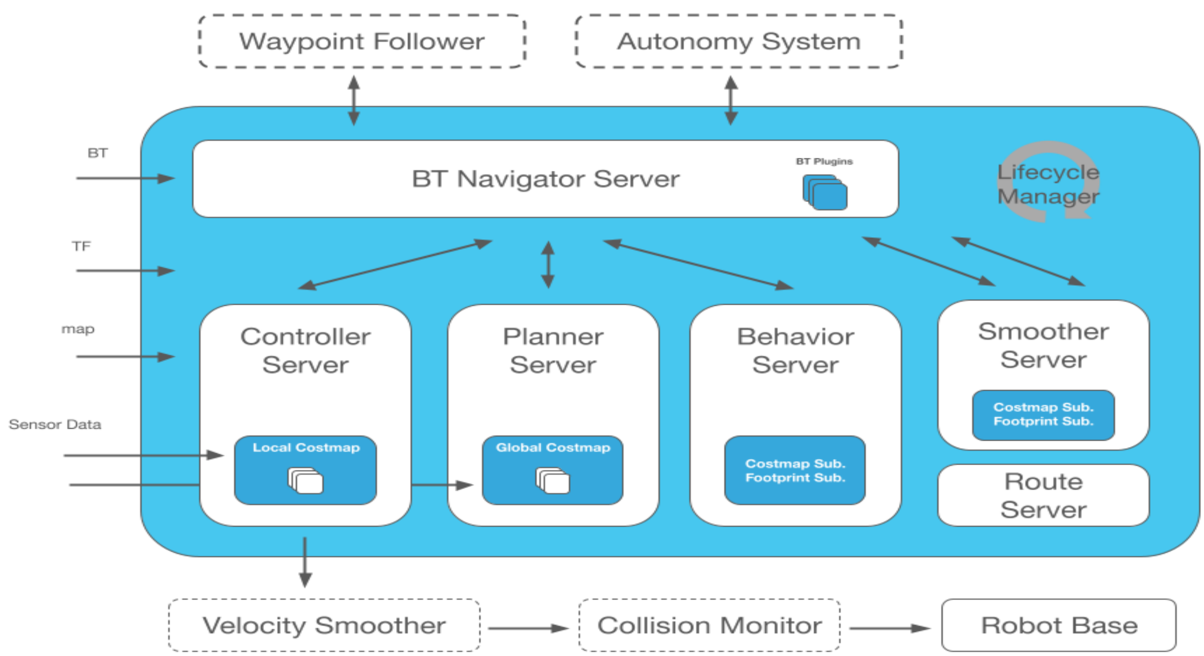


Figure-8: AUTONOMOUS NAVIGATION ARCHITECTURE [46]

- Manage the loading, serving, and storage of maps with a Map Server. Localize the robot on the map (AMCL)
- Perform robot localization on the map using the Adaptive Monte Carlo Localization (AMCL) algorithm. Control the robot as it follows the path (Nav2 Controller)
- Design a route from point A to point B while avoiding obstacles with the Nav2 Planner. Convert sensor data into a costmap representation of the world (Nav2 Costmap 2D)
- Build complicated robot behaviours using behaviour trees (Nav2 Behaviour Trees and BT Navigator)
- Convert the data collected by the sensor into a representation of the world known as a costmap using Nav2 Costmap 2D. Follow sequential waypoints (Nav2 Waypoint Follower)
- Construct intricate robot behaviors by utilizing behavior trees, such as Nav2 Behavior Trees and BT Navigator.
- Plugins are available to facilitate the integration of your own customized algorithms and behaviors into the Nav2 Core system.
- Monitor unprocessed sensor data to detect an impending accident or hazardous condition (accident Monitor).
- Python3 API for interacting with Nav2 in a pythonic manner, specifically through the use of Simple Commander.
- A smoother is applied to output velocities to ensure the dynamic feasibility of commands (velocity).
- Sensor data in this case is provided by laser sensor

5.1.4 COMPARISON OF LINEAR CMD-VELOCITY VS LINEAR NAVIGATIONAL VELOCITY WITH LASER SCAN

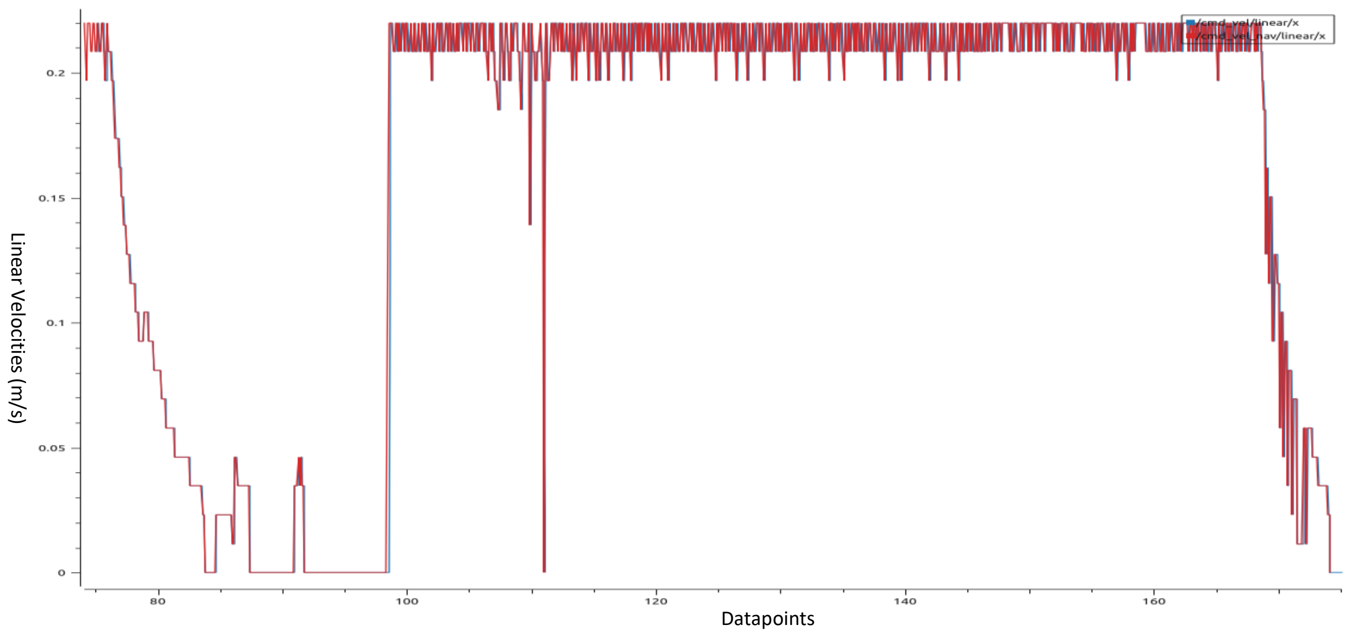
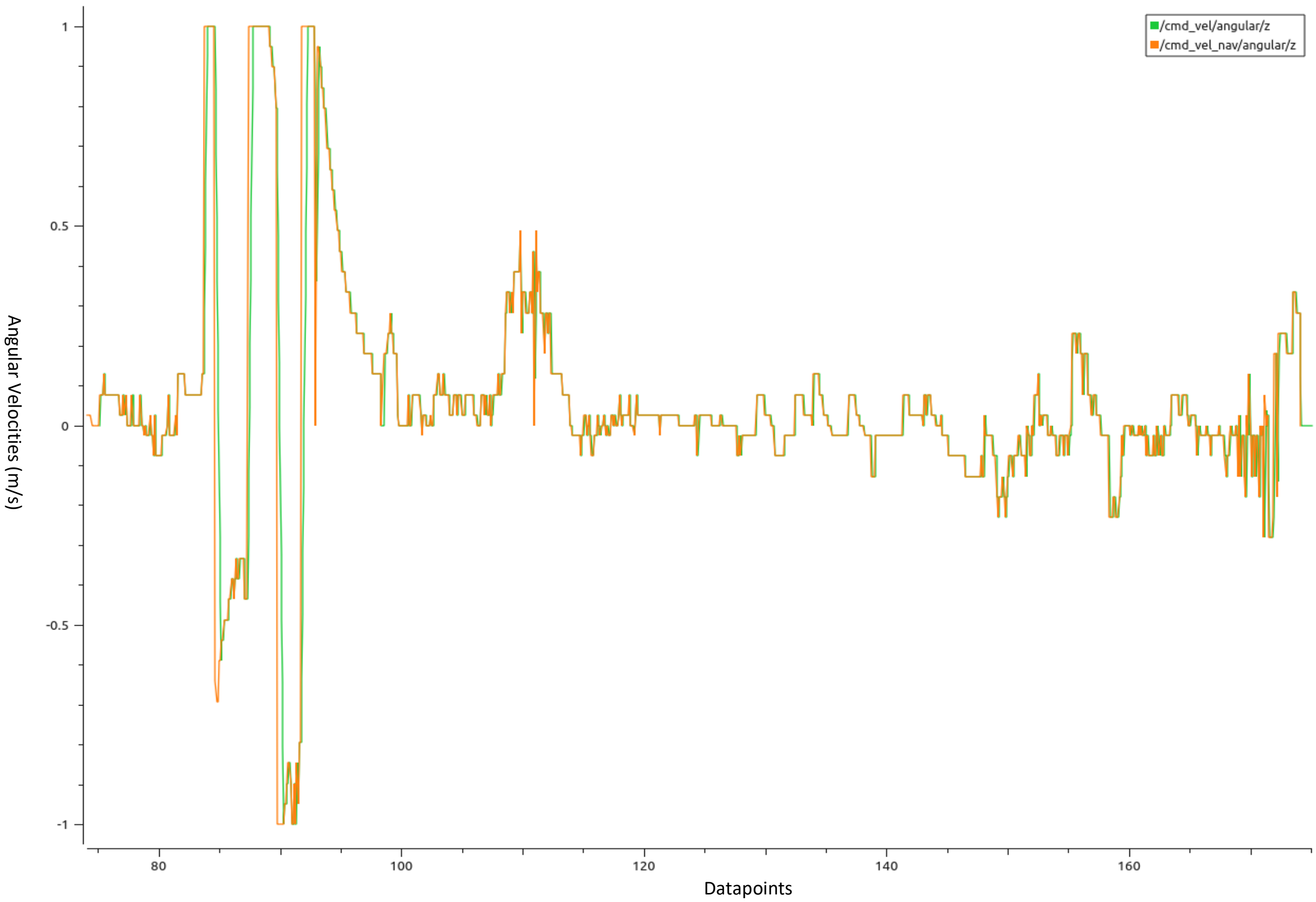


FIGURE-9: CMD VS NAVIGATTION LINEAR VELOCITIES IN EXPERIMENT-1

In the above graph vertical axis shows the velocities of robot in meter/sec with 0m/s lowest and 0.26m/s the highest. whereas horizontal axis shows the datapoints recorded during the autonomous navigation related to velocity. As the map acquired by the global planner is quite rich and also the surfaces of the gazebo world in this scenario are very ideal- (free of any irregularities and unexpected cracks and bumps on the floor. Therefore, both global planners driven by global costmap and local planner driven by laser made local costmap are performing optimally. On the basis of observed graph, it can be analysed that except few points of horizontal data points from 80 to 90, the linear velocity graph is approximately overlapping the linear velocity navigational graph. In addition, ups and down in the graphs indicates the changes in velocity while moving and approaching different locations or waypoints set by the user for autonomous navigation as could be seen in the related video of Experiment-1. Overall, the robot traverse and maintained on average at velocity of 0.16 m/s linearly and 0.05m/s in angular domain.

5.1.5 COMPARISON OF ANGULAR CMD-VELOCITIES VS ANGULAR NAVIGATIONAL VELOCITY WITH LASER SCAN



FIGUR-10: CMD VS NAVIGATTION ANGULAR VELOCITIES IN EXPERIMENT-1

This graph shows the angular velocities achieved during the navigation w.r.t z-axis pointing upward. The highest and lowest peaks between datapoint 90 and 95 are as a result of sharp turns robot made during the path following. Overall angular velocities in terms of command and navigation remains synchronize leading to smooth navigation of the robot.

5.1.6 OVERALL STATISTICS OF COMPARISON ANGULAR AND LINEAR CMD-VELOCITIES VS NAVIGATIONAL VELOCITIES WITH LASER SCANNER

S. No	Velocities (m/s)	Min	Max	Average
1	/cmd_vel/angular/z	-1	1	0.054616
2	/cmd_vel/linear/x	0	0.22	0.163274
3	/cmd_vel_nav/angular/z	-1	1	0.055811
4	/cmd_vel_nav/linear/x	0	0.22	0.165248

TABLE-3: EXPRIMENT 1 VELOCITIES STATISTICS

5.2 Experiment-2

In a similar fashion Experiment-2 is performed by following steps adapted earlier for Experiment-1. However, in this particular experiment, SLAM and Navigation are performed using depth-camera despite of laser scanner.

5.2.1 Experiment-2 SLAM results

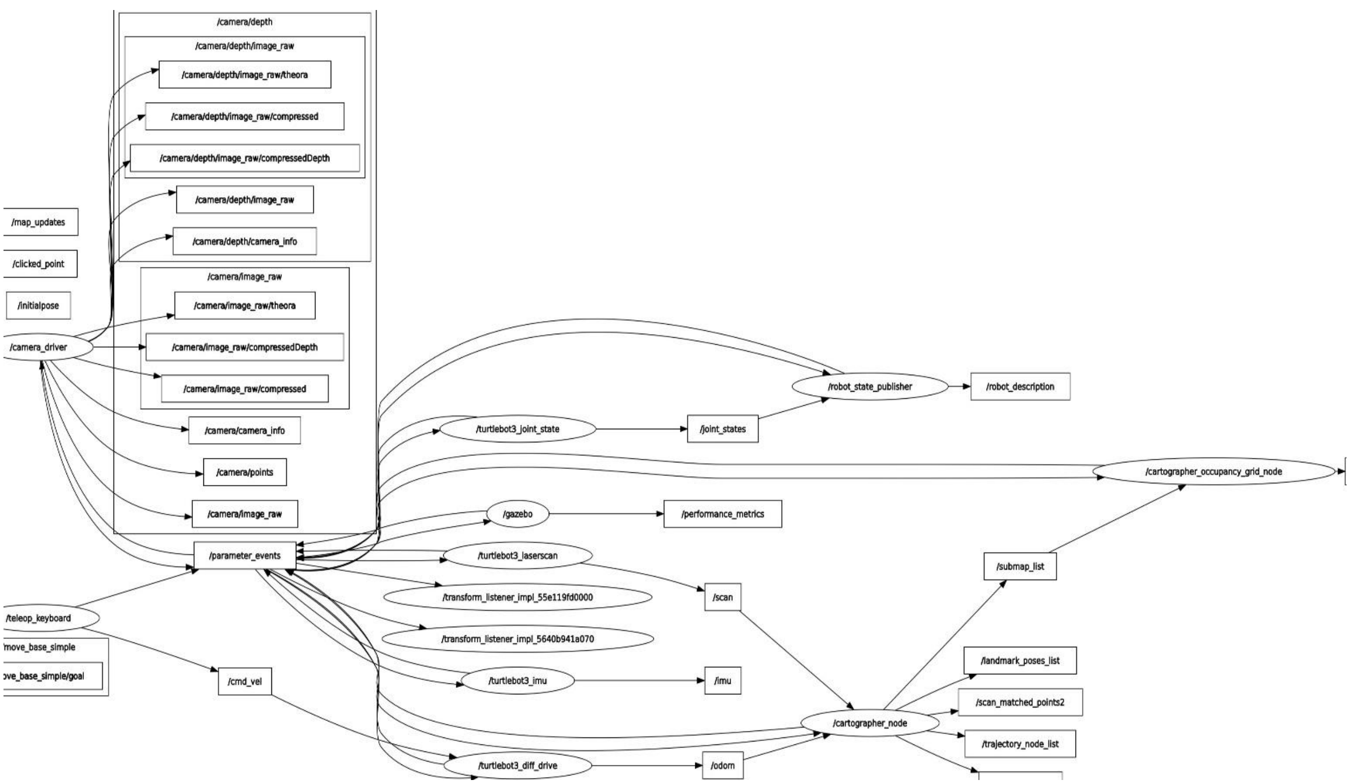


Figure-11: GRAPH OF SLAM OPERATION RELATED WITH EXPERIMENT-2

Above Graph shows the involvement of depth camera related nodes, drivers and topics in delivering the final product that is occupancy grid map.

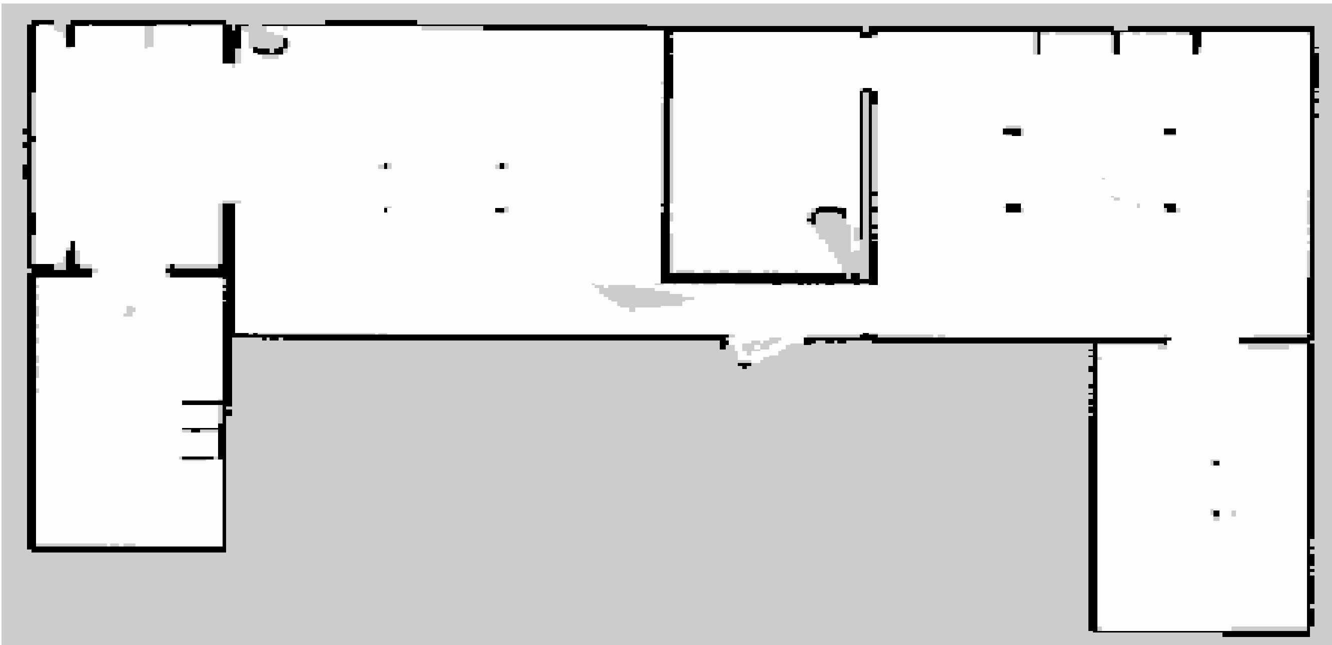


FIGURE-12: MAP GENERATED IN EXPERIMENT-2

Figure-12 shows the map of house scanned by depth camera based robotic configuration utilized in Experiment-2. Overall quality of map can be regarded as good but not superior than the laser-based scanning of Experiment-1 at some locations marked as grey rather than white.

5.2.2 EXPERIMENT-2 AUTONOMOUS NAVIGATION

Autonomous Navigation related to experiment-2 is based on the map acquired during the SLAM phase of Experiment-2 as discussed earlier. The flow of operation is enclosed herewith as Appendix-‘F’.

5.2.3 COMPARISON OF LINEAR CMD-VELOCITY VS LINEAR NAVIGATIONAL VELOCITY WITH DEPTH CAMERA

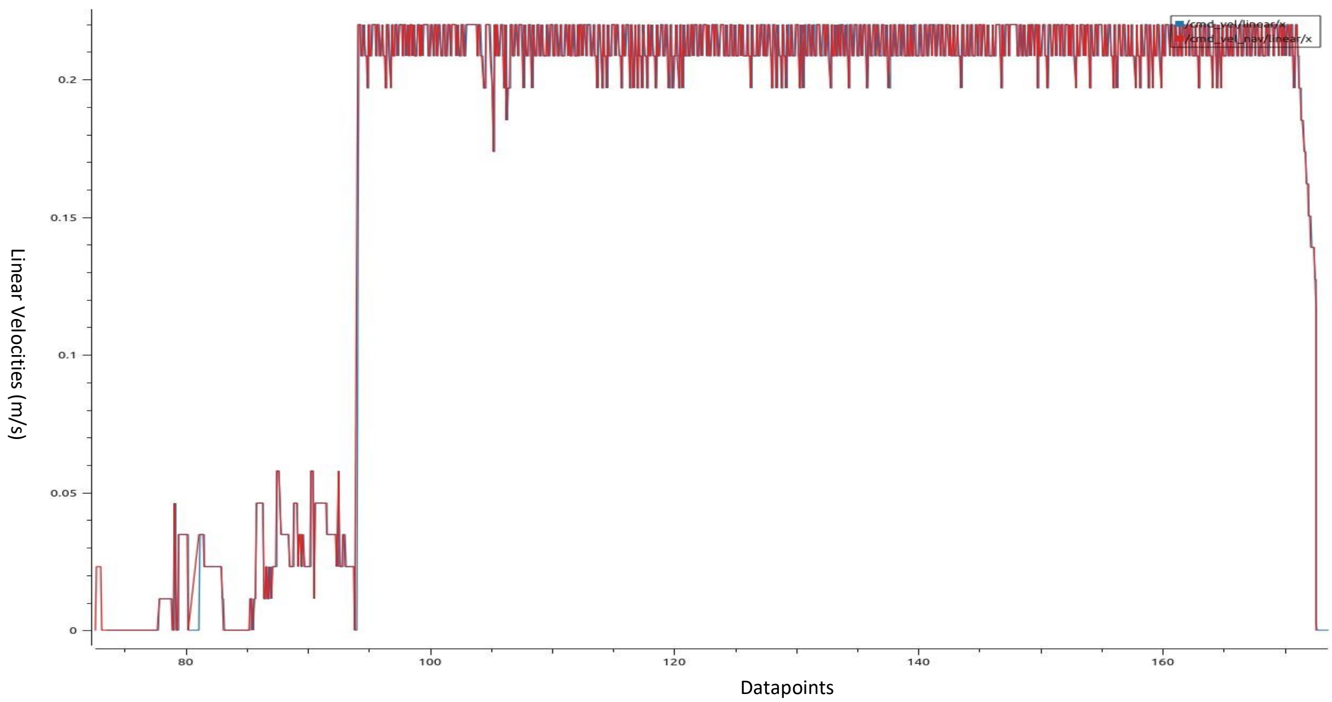
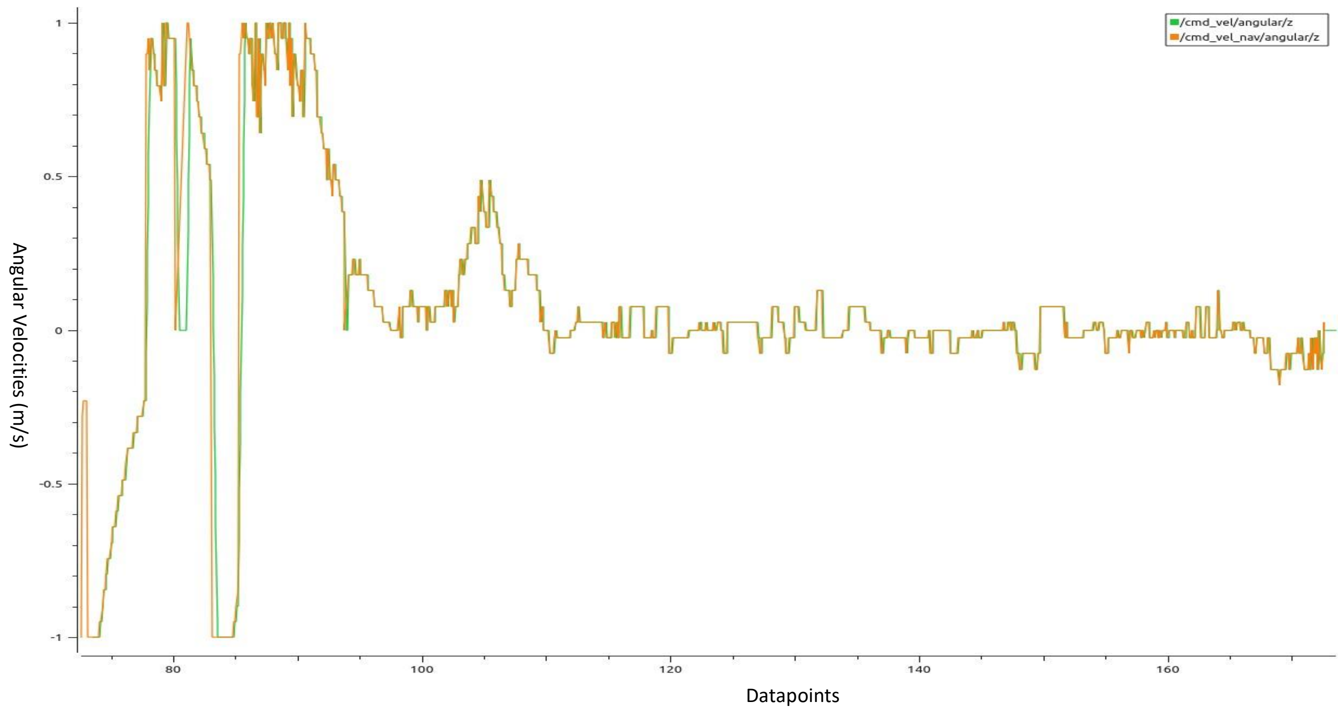


FIGURE-13: CMD VS NAVIGATTION LINEAR VELOCITIES IN EXPERIMENT-2

In case of Experiment-2, an anomaly is detected at datapoint 84 where navigational velocity couldn't synchronize with cmd velocity. overall, the trained of following cmd velocity by navigational linear velocity is synchronized. The average linear velocity in this case is 0.17m/s, however the variation of velocity is comparatively less leading to smoother Navigation as compared to laser-based Navigation. Furthermore, no intermediate falls and peaks are observed as well.

5.2.4 COMPARISON OF ANGULAR CMD-VELOCITY VS ANGULAR NAVIGATIONAL VELOCITY WITH DEPTH CAMERA



FIGUR-14: CMD VS NAVIGATTION ANGULAR VELOCITIES IN EXPERIMENT-2

In Experiment-2 the average angular velocities remain mostly around 0.07m/s. It means more targeted movement was achieved as compared to Experiment-1 as evident in the video of Experiment-2. Despite there are sudden variation and large peaks that could be observed between 0 and 85 datapoints with same anomaly as in case of linear velocity at point 84.

5.2.5 OVERALL STATISTICS OF COMPARISON ANGULAR AND LINEAR CMD-VELOCITIES VS NAVIGATIONAL VELOCITIES WITH DEPTH CAMERA

S. No	Velocities (m/s)	Min	Max	Average
1	/cmd_vel/angular/z	-1	1	0.074799
2	/cmd_vel/linear/x	0	0.22	0.170072
3	/cmd_vel_nav/angular/z	-1	1	0.072069
4	/cmd_vel_nav/linear/x	0	0.22	0.171644

TABLE-4: EXPERIMENT-2 VELOCITIES STATISTICS

5.3 Experiment-3

Experiment-3 is performed by following steps adapted last two Experiments. However, in this particular SLAM and Autonomous Navigation are performed using Infrared-camera.

5.3.1 Experiment-3 SLAM results

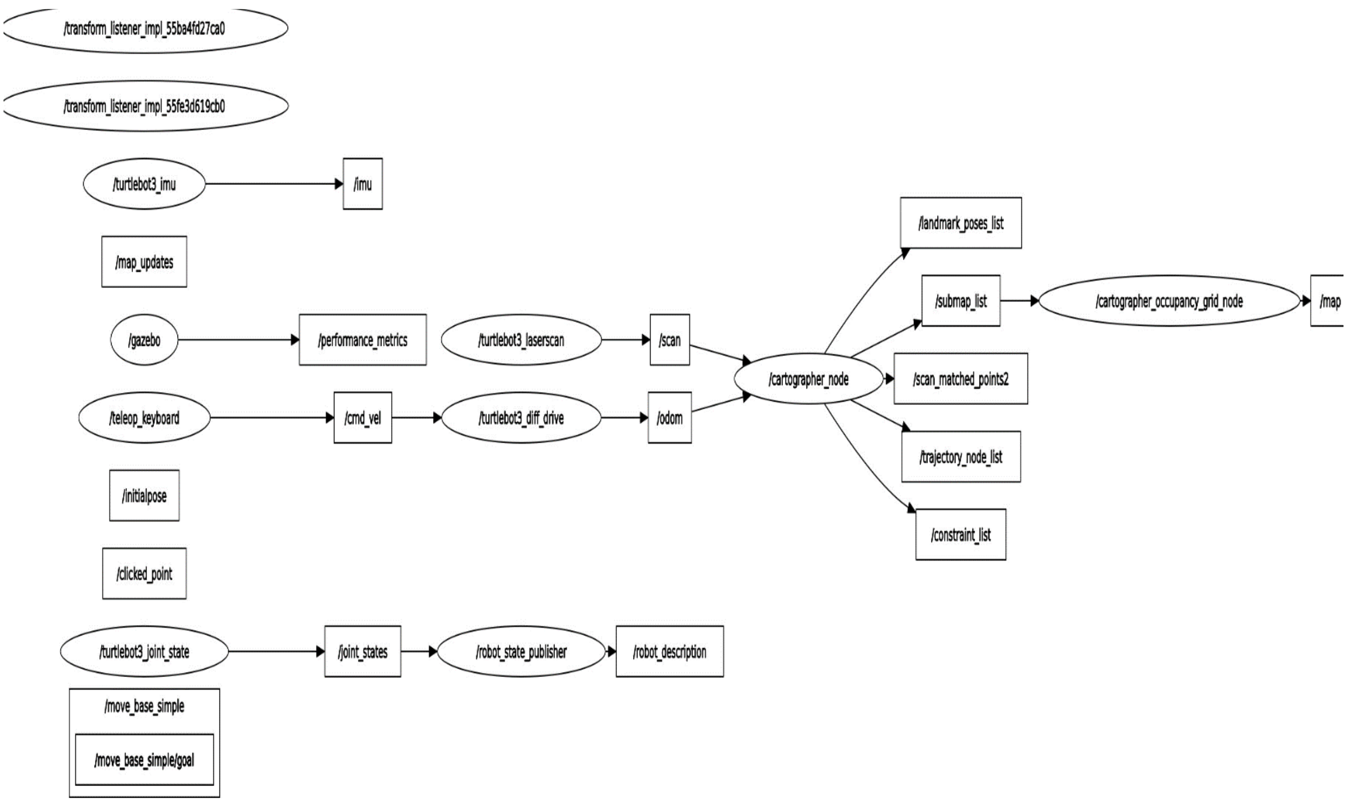


Figure-15: GRAPH OF SLAM OPERATION EXPERIMENT-3

Above Graph shows the involvement of infrared-camera related nodes, drivers and topics in delivering the final product that is occupancy grid map.

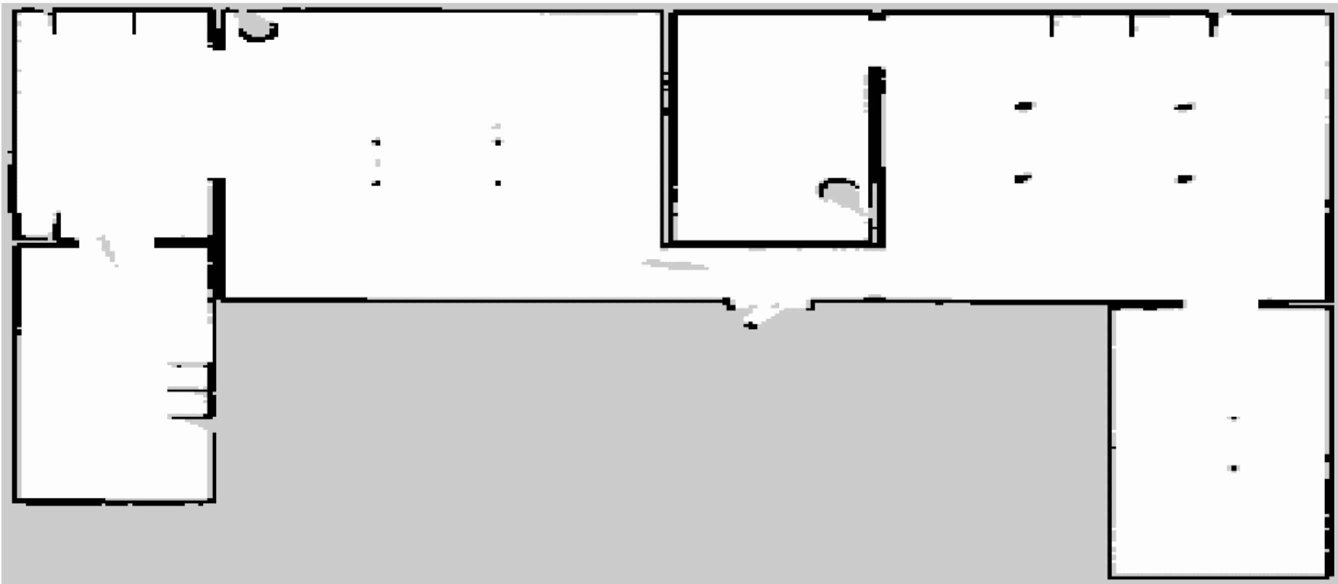


FIGURE-16: MAP GENERATED IN EXPERIMENT-3

Above Figure shows the map of house scanned by infrared camera based robotic configuration utilized in Experiment-3. Overall quality of map can be regarded as same as of that laser-camera.

5.3.2 EXPERIMENT-3 AUTONOMOUS NAVIGATION

Autonomous Navigation related to experiment-3 is based on the map acquired during the SLAM phase of Experiment-3 as discussed earlier. The flow of operation enclosed herewith as Appendix-'G'.

5.3.3 COMPARISON OF LINEAR CMD-VELOCITY VS LINEAR NAVIGATIONAL VELOCITY WITH INFRARED CAMERA

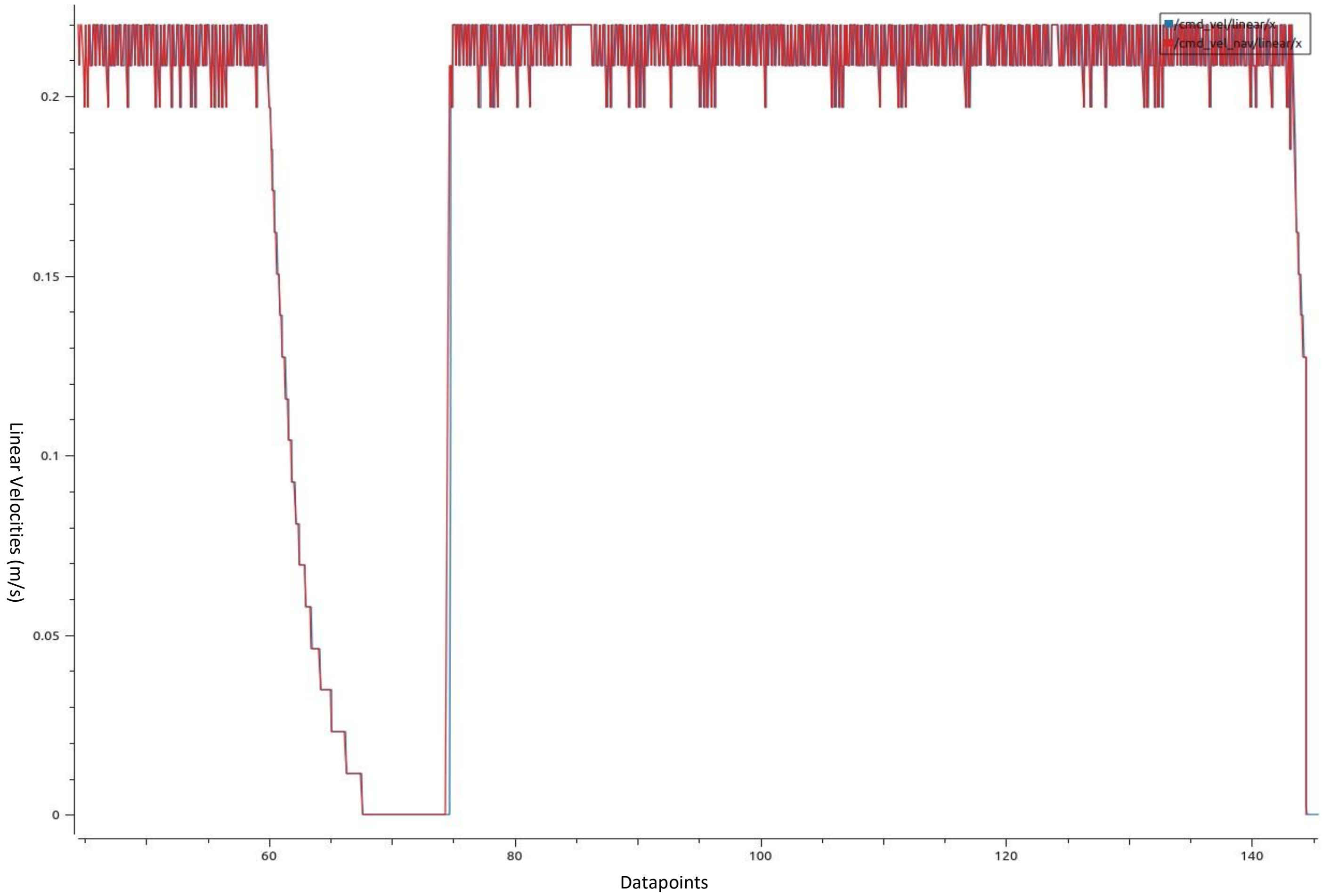


FIGURE-17: CMD VS NAVIGATTION LINEAR VELOCITIES IN EXPERIMENT-3

In case of Experiment-3, a sudden decline in velocity is detected from point 62 to 70 datapoint. whereas navigational velocity couldn’t be synchronized by a negligible margin with cmd velocity near 76 of the datapoint. overall, the trained of following cmd velocity by navigational linear velocity is synchronized. The average velocity in this case is 0.186 m/s. Furthermore, no intermediate falls and peaks are observed as well except for aforementioned points.

5.3.4 COMPARISON OF ANGULAR CMD-VELOCITY VS ANGULAR NAVIGATIONAL VELOCITY WITH DEPTH CAMERA

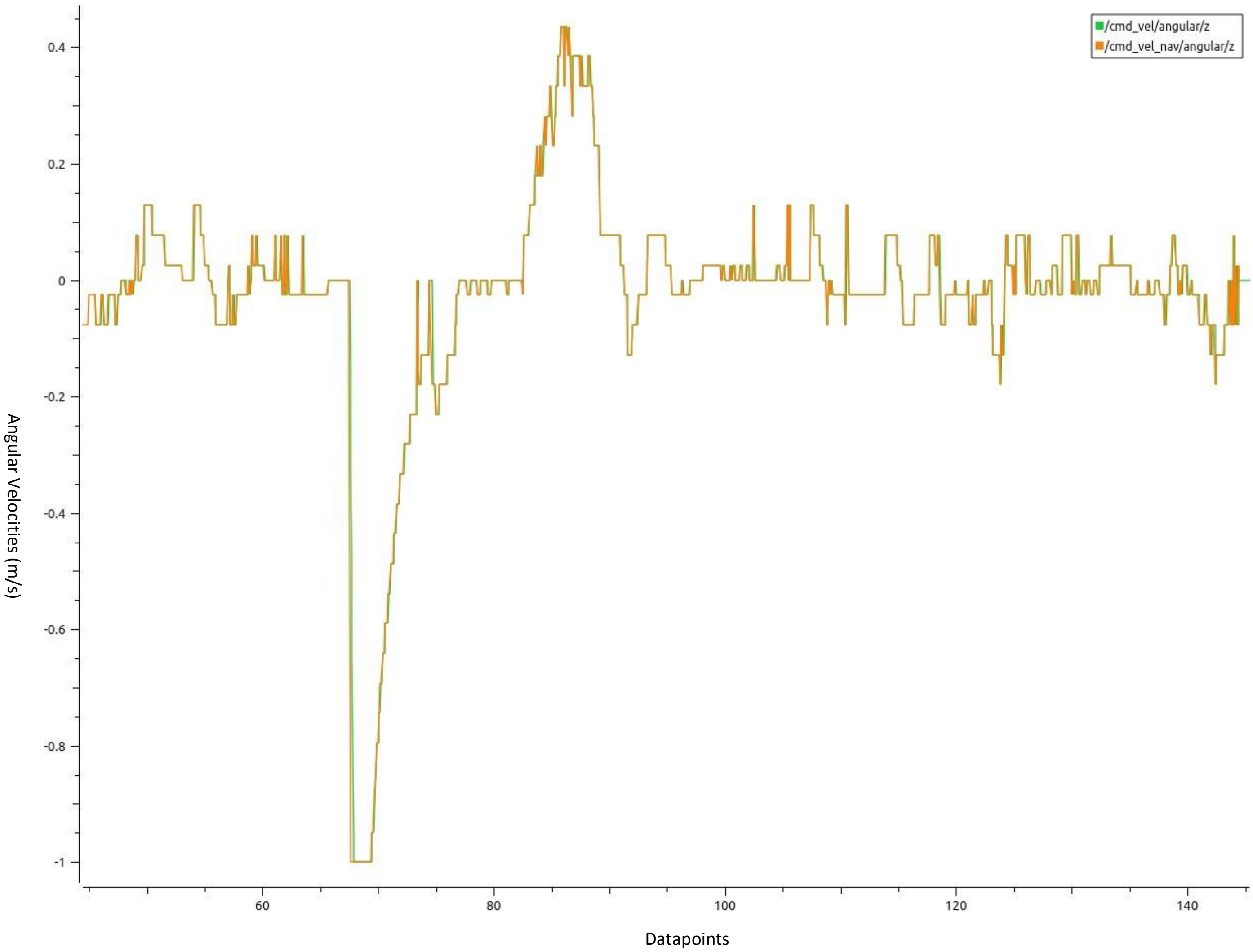


FIGURE-18:CMD VS NAVIGATTION LINEAR VELOCITIES IN EXPERIMENT-3

In Experiment-3 the average angular velocities remain mostly around -0.02 m/s. It means more targeted movement was achieved as compared to Experiment-1. Despite there are sudden variation and large peaks that could be observed between 68 to 70 datapoints.

5.3.5 OVERALL STATISTICS OF COMPARISON ANGULAR AND LINEAR CMD-VELOCITIES VS NAVIGATIONAL VELOCITIES WITH DEPTH CAMERA

S. No	Velocities (m/s)	Min	Max	Average
1	/cmd_vel/angular/z	-1	0.435897	-0.024372
2	/cmd_vel/linear/x	0	0.22	0.184123
3	/cmd_vel_nav/angular/z	-1	0.435897	-0.026435
4	/cmd_vel_nav/linear/x	0	0.22	0.186096

TABLE-5: EXPERIMENT-3 VELOCITIES STATISTICS

Following last three experiments (4,5 and 6) are performed in NeRF generated world one by one by using Laser camera then Depth camera and at the end with Infrared camera.

5.4 Experiment-4

Experiment-4 is performed in NeRF generated world. Furthermore, in this particular SLAM and Autonomous Navigation are performed using laser-camera.

5.4.1 Experiment-4 SLAM results

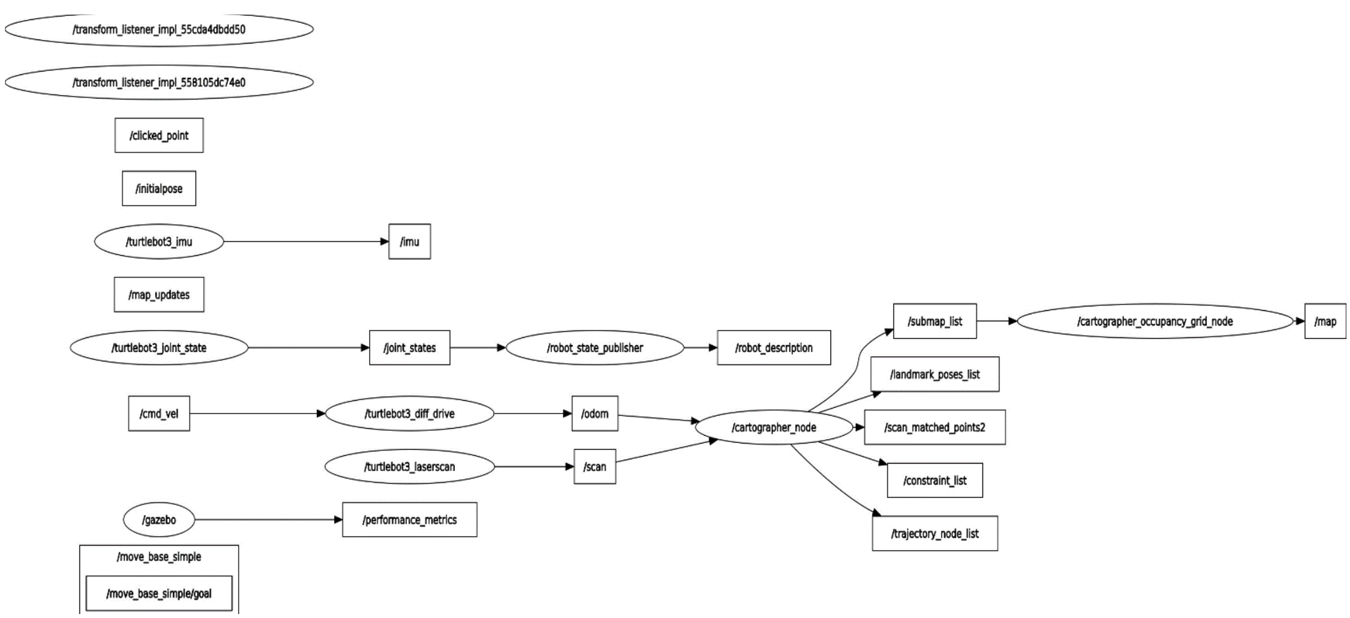


Figure-19: GRAPH OF SLAM OPERATION EXPERIMENT-4

Above Graph shows the involvement of laser camera related nodes, drivers and topics in delivering the final product that is occupancy grid map.

FIGURE-20: MAP GENERATED IN EXPERIMENT-4

Figure-20 shows the map of NeRF world scanned by laser camera based robotic configuration utilized in Experiment-4. Overall quality of map can be regarded as same as of that depth-camera but not superior than the laser-based scanning of Experiment-1 because of imperfect land having bumps and cracks. A clear differentiation in map quality can be seen between left (NeRF land) and right (ideal land).

5.4.2 EXPERIMENT-4 AUTONOMOUS NAVIGATION

Autonomous Navigation related to experiment-4 is based on the map acquired during the SLAM phase of Experiment-4 as discussed earlier. The flow of operation is enclosed herewith as Appendix-‘E’.

5.4.3 COMPARISON OF LINEAR CMD-VELOCITY VS LINEAR NAVIGATIONAL VELOCITY WITH LASER CAMERA IN NeRF WORLD

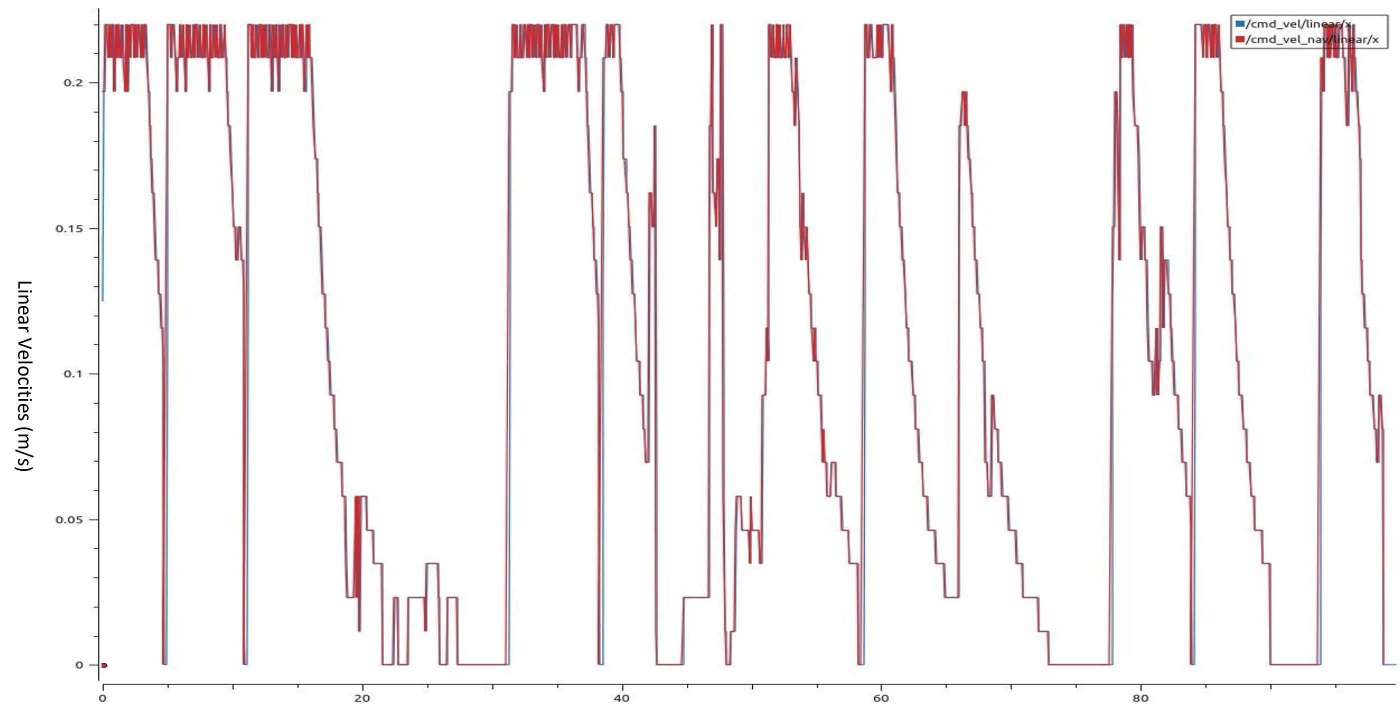


FIGURE-21: CMD VS NAVIGATTION LINEAR VELOCITIES IN EXPERIMENT-4

In case of Experiment-4, many sudden decline and falls could be blatantly seen. It is because of more time taken by the robot in taking decision for executing command velocities received. However, one positive aspect is the synchronization of cmd and navigational velocities throughout the experiment. This time lag for taking decision is due to rough NeRF generated land covering minute details of surface cracks with NeRF generated blue cylinder as shown in Map-2. The average linear velocity remains between 0.10 m/s to 0.11 m/s.

5.4.4 COMPARISON OF ANGULAR CMD-VELOCITY VS ANGULAR NAVIGATIONAL VELOCITY WITH DEPTH CAMERA

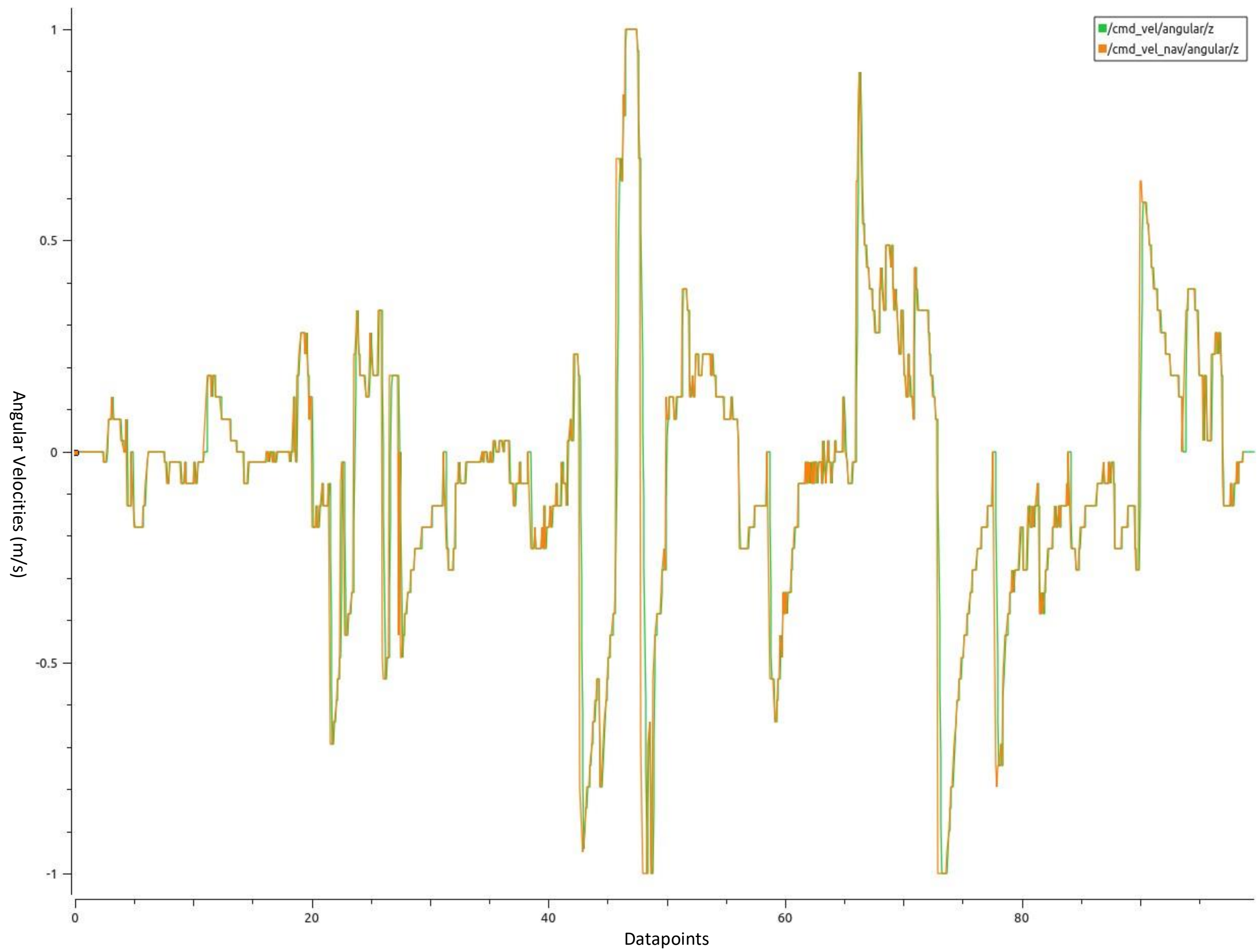


FIGURE-22: CMD VS NAVIGATTION ANGULAR VELOCITIES IN EXPERIMENT-4

In Experiment-4 the average angular velocities remain mostly around -0.04 m/s to -0.05 m/s. However, synchronization of cmd velocities and navigational velocities can be observed as a positive aspect.

5.4.5 OVERALL STATISTICS OF COMPARISON ANGULAR AND LINEAR CMD-VELOCITIES VS NAVIGATIONAL VELOCITIES WITH DEPTH CAMERA

S. No	Velocities (m/s)	Min	Max	Average
1	/cmd_vel/angular/z	-1	1	-0.04602
2	/cmd_vel/linear/x	0	0.22	0.107851
3	/cmd_vel_nav/angular/z	-1	1	-0.05298
4	/cmd_vel_nav/linear/x	0	0.22	0.110371

TABLE-6: EXPERIMENT-4 VELOCITIES STATISTICS

5.5 Experiment-5

Experiment-5 is performed in NeRF generated world. Furthermore, in this particular SLAM and Autonomous Navigation are performed using depth-camera.

5.5.1 Experiment-5 SLAM results

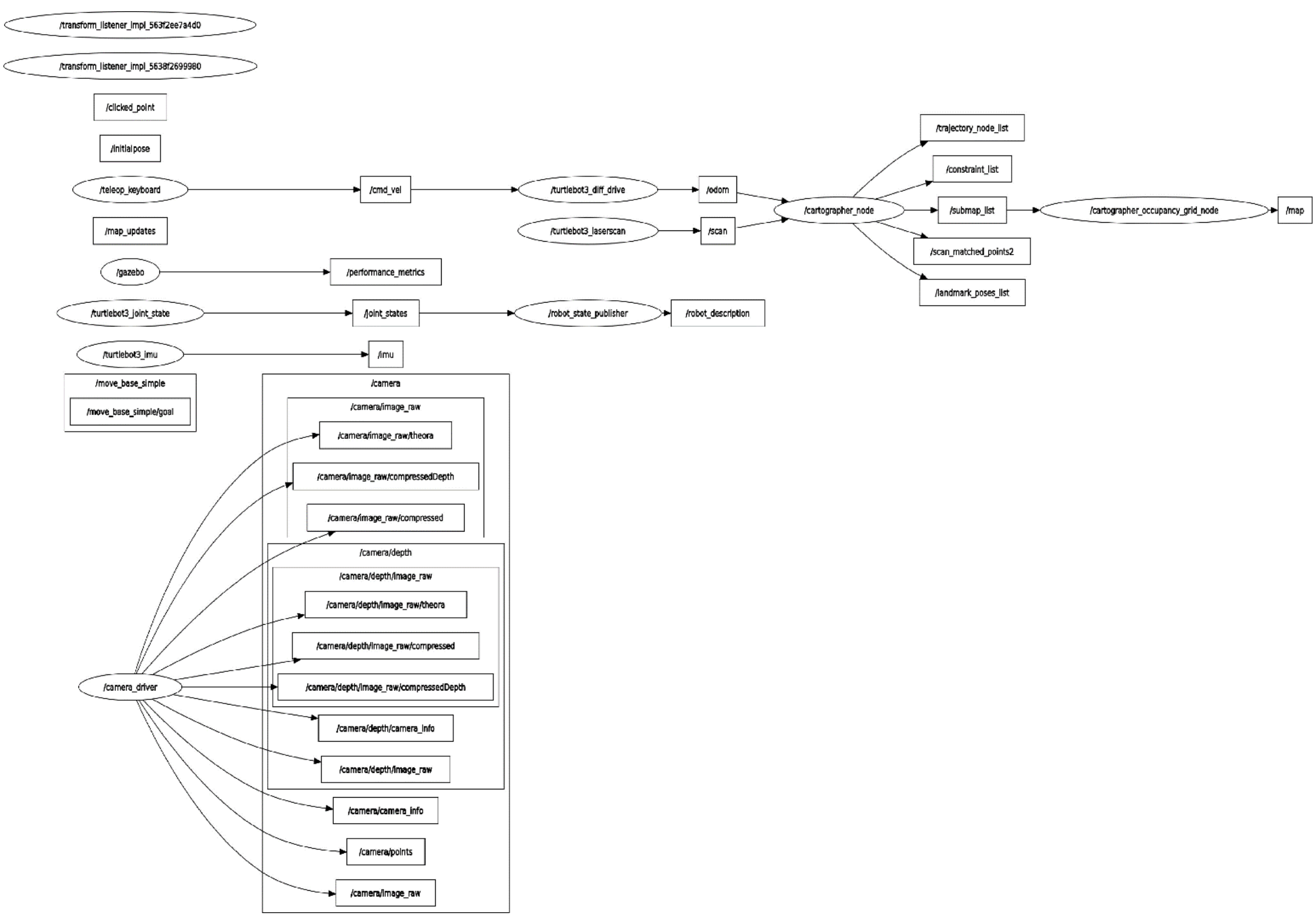


FIGURE-23: GRAPH OF SLAM OPERATION EXPERIMENT-5

Above Graph shows the involvement of depth camera related nodes, drivers and topics in delivering the final product that is occupancy grid map.

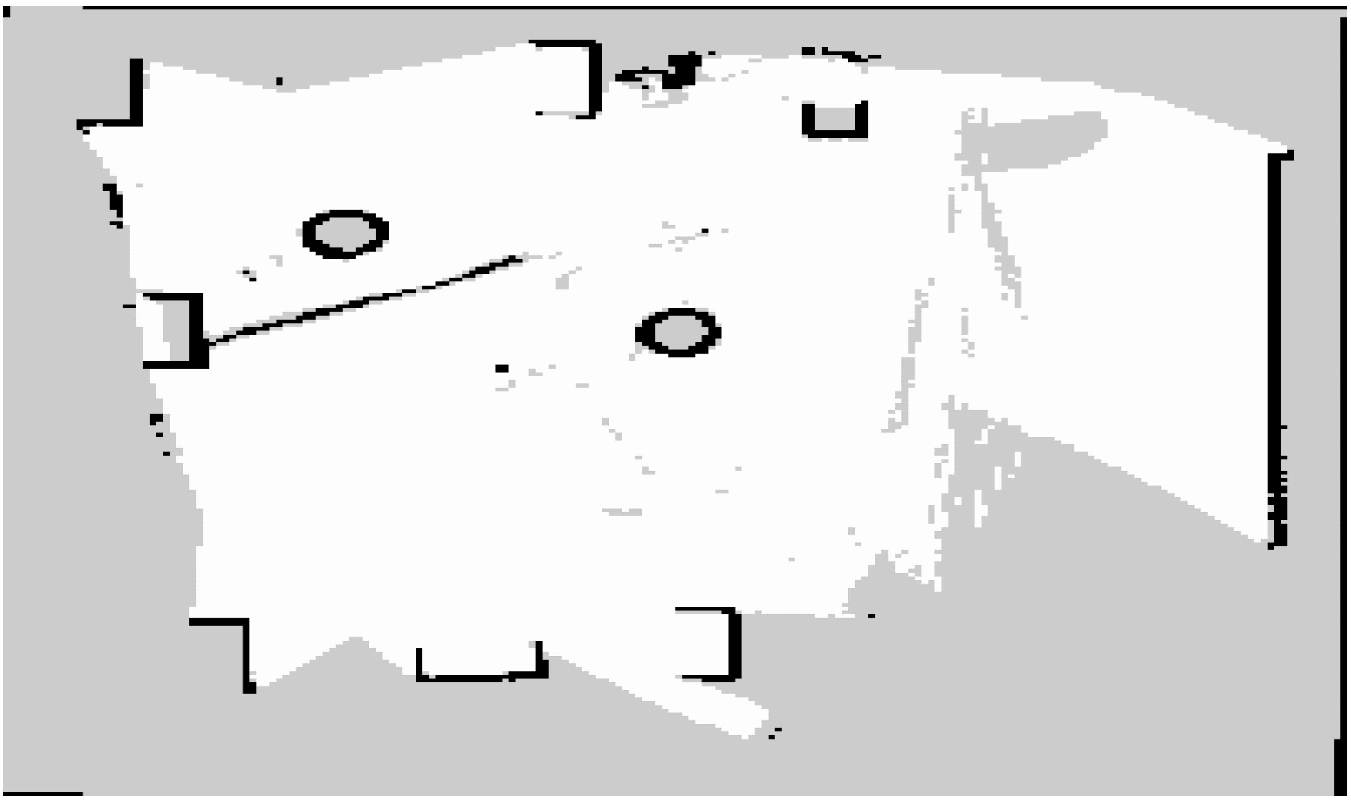


FIGURE-24: MAP GENERATED IN EXPERIMENT-5

Figure-24 shows the map of NeRF world scanned by depth camera based robotic configuration utilized in Experiment-5. Overall quality of map can be regarded as worse as compared to laser scanning of NeRF world.

5.5.2 EXPERIMENT-5 AUTONOMOUS NAVIGATION

Autonomous Navigation related to experiment-5 is based on the map acquired during the SLAM phase of Experiment-5 as discussed earlier. The flow of operation is enclosed herewith as Appendix- ‘F’.

5.5.3 COMPARISON OF LINEAR CMD-VELOCITY VS LINEAR NAVIGATIONAL VELOCITY WITH INFRARED CAMERA

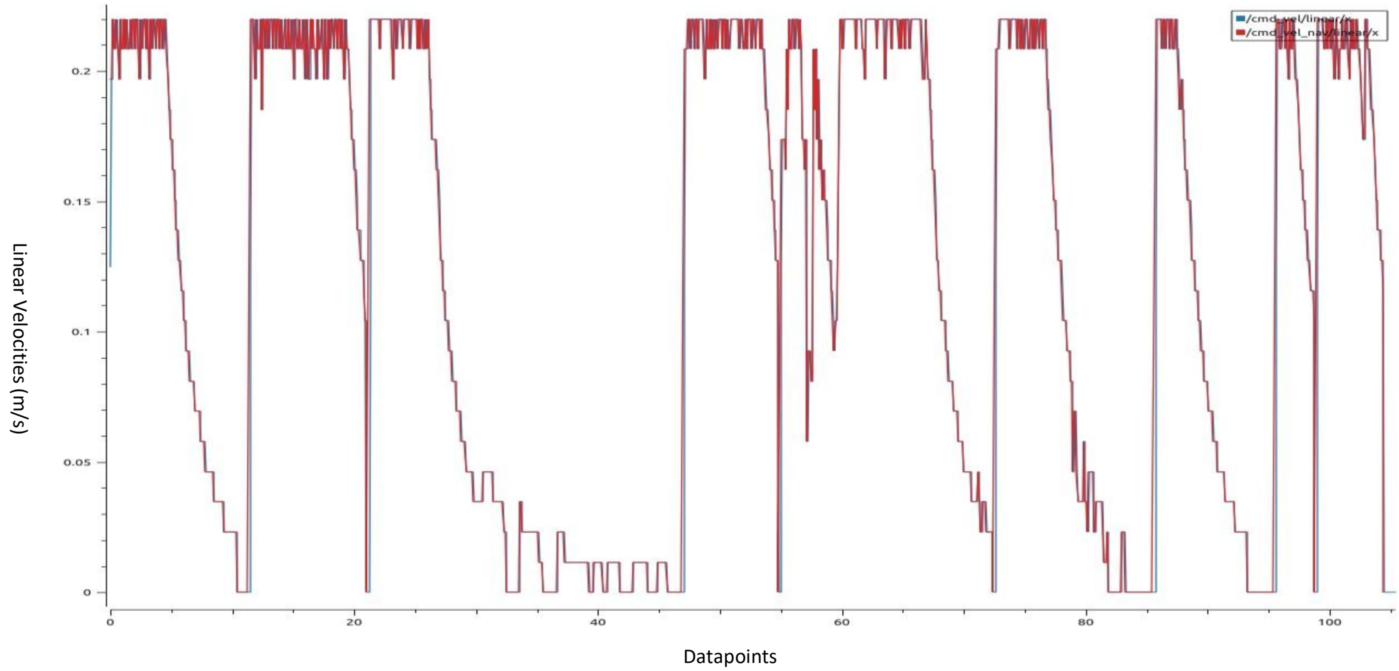


FIGURE-25: CMD VS NAVIGATTION LINEAR VELOCITIES IN EXPERIMENT-5

In case of Experiment-5, many sudden decline and falls could be blatantly seen. It is because of more time taken by the robot in taking decision for executing command velocities received. However, span of variations are less as compared to laser-based scanning of NeRF world. Another positive aspect is the synchronization of cmd and navigational velocities throughout the experiment. This time lag for taking decision is due to rough NeRF generated land covering minute details of surface cracks with NeRF generated blue cylinder as shown in Map-2. The average linear velocity remains around 0.12 m/s.

5.5.4 COMPARISON OF ANGULAR CMD-VELOCITY VS ANGULAR NAVIGATIONAL VELOCITY WITH DEPTH CAMERA

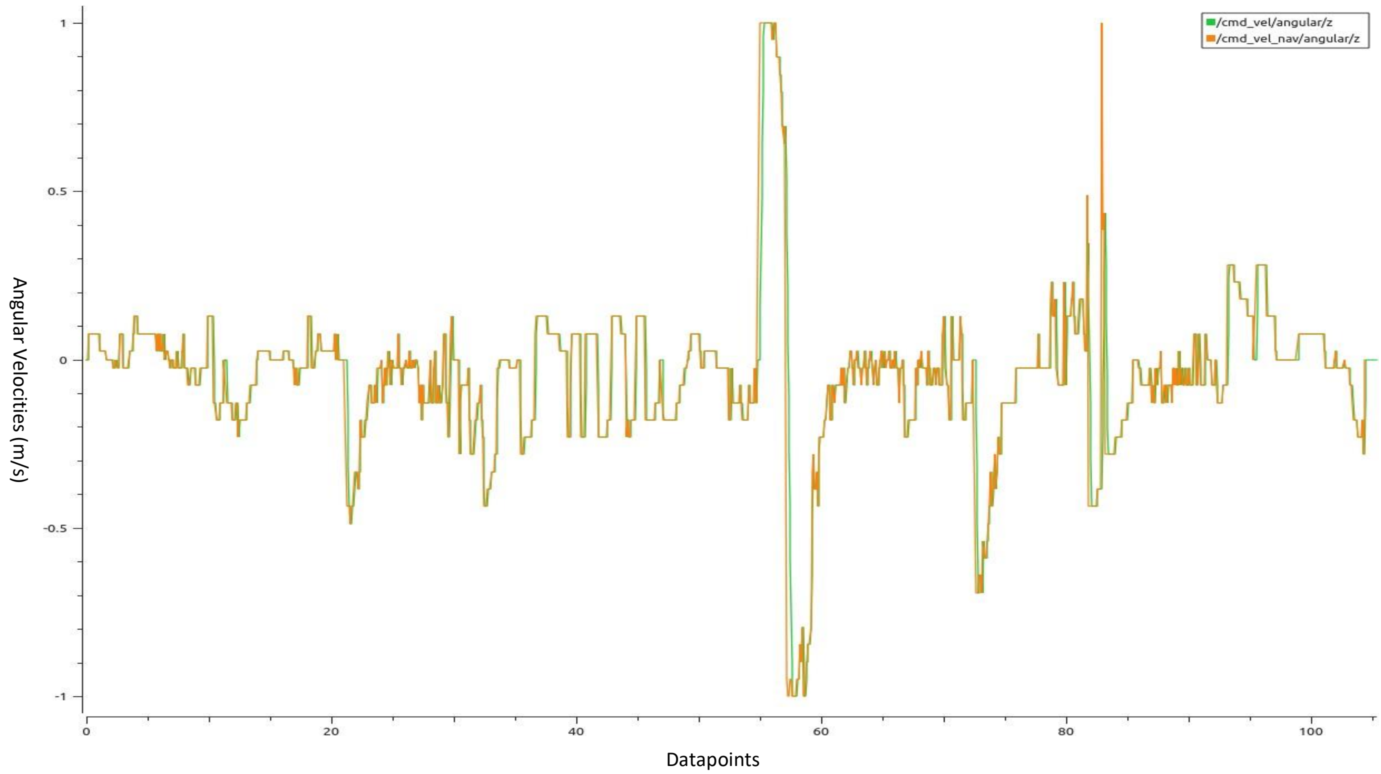


FIGURE-26: CMD VS NAVIGATTION ANGULAR VELOCITIES IN EXPERIMENT-5

In Experiment-5 the average angular velocities remain mostly around -0.04m/s to -0.05 m/s.. However, overall synchronization of cmd velocities and navigational velocities can be observed as a positive aspect.

5.5.5 OVERALL STATISTICS OF COMPARISON ANGULAR AND LINEAR CMD-VELOCITIES VS NAVIGATIONAL VELOCITIES WITH DEPTH CAMERA

S. No	Velocities (m/s)	Min	Max	Average
1	/cmd_vel/angular/z	-1	1	-0.046255
2	/cmd_vel/linear/x	0	0.22	0.124128
3	/cmd_vel_nav/angular/z	-1	1	-0.050666
4	/cmd_vel_nav/linear/x	0	0.22	0.126767

TABLE-7: EXPERIMENT-5 VELOCITIES STATISTICS

5.6 Experiment-6

Experiment-6 is performed in NeRF generated world. Furthermore, in this particular SLAM and Autonomous Navigation are performed using infrared-camera.

5.6.1 Experiment-6 SLAM results

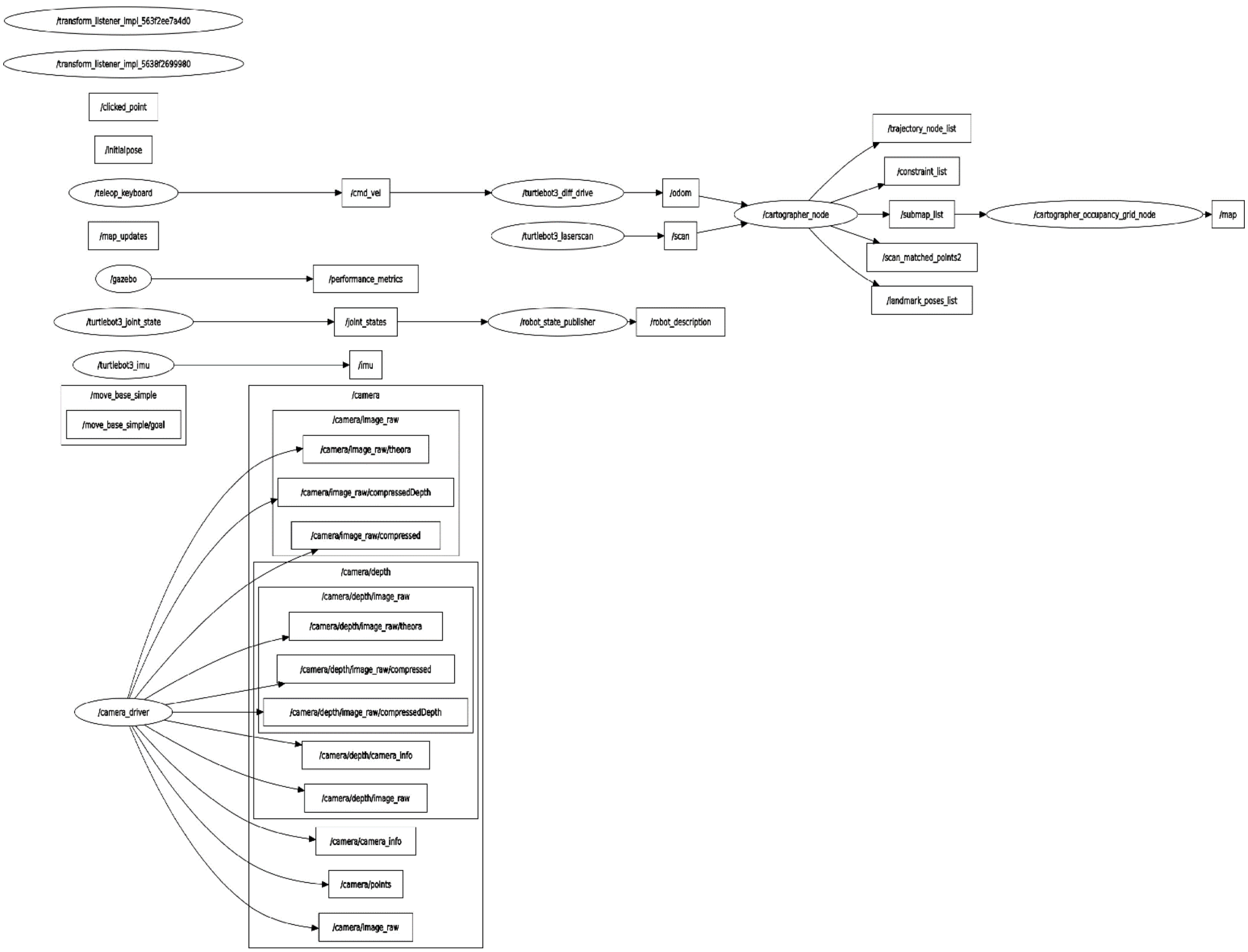


FIGURE-27: GRAPH OF SLAM OPERATION EXPERIMENT-6

Above Graph shows the involvement of infrared camera related nodes, drivers and topics in delivering the final product that is occupancy grid map.

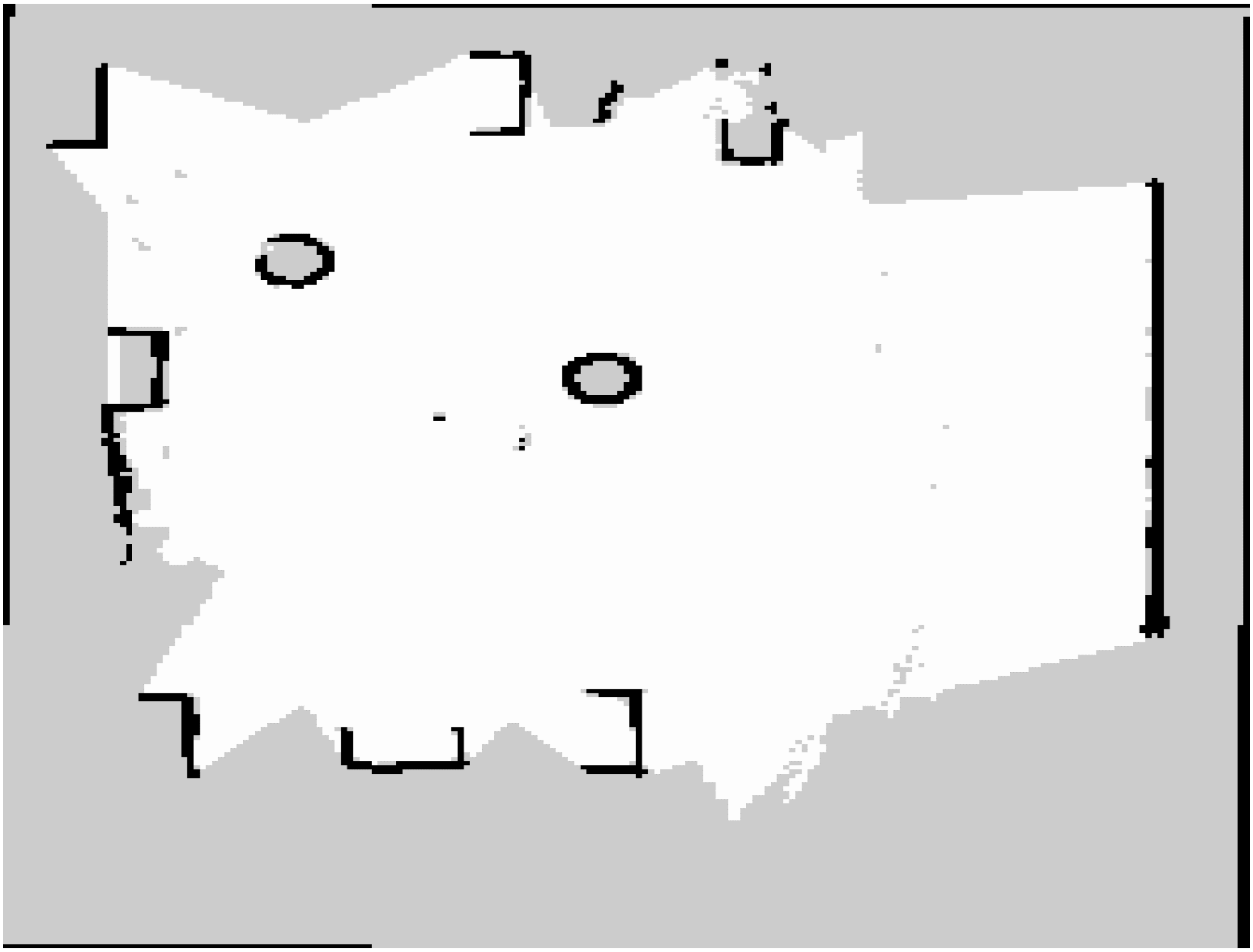


FIGURE-28: MAP GENERATED IN EXPERIMENT-6

Figure-28 shows the map of NeRF world scanned by infrared camera based robotic configuration utilized in Experiment-6. Overall quality of map can be regarded as best as compared to laser and depth camera-based maps defined in last two sections.

5.6.2 EXPERIMENT-6 AUTONOMOUS NAVIGATION

Autonomous Navigation related to experiment-6 is based on the map acquired during the SLAM phase of Experiment-6 as discussed earlier. The flow of operation is enclosed herewith as Appendix-'G'.

5.6.3 COMPARISON OF LINEAR CMD-VELOCITY VS LINEAR NAVIGATIONAL VELOCITY WITH INFRARED CAMERA

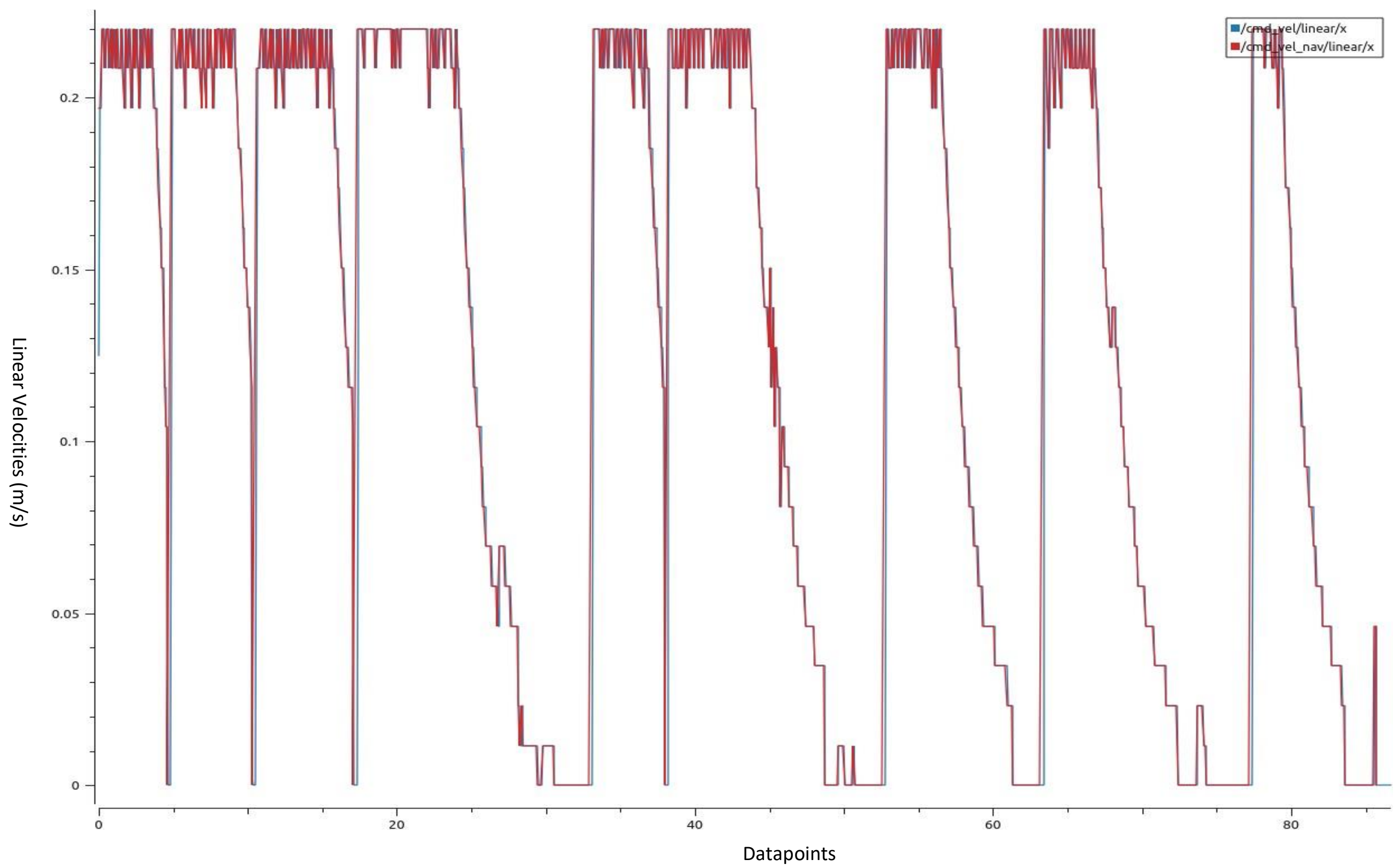


FIGURE-29: CMD VS NAVIGATTION LINEAR VELOCITIES IN EXPERIMENT-6

In case of Experiment-5, many sudden decline and falls could be clearly seen. It is because of more time taken by the robot in taking decision for executing command velocities received. However, span of variations are less as compared to laser-based scanning of NeRF world. Another positive aspect is the synchronization of cmd and navigational velocities throughout the experiment. This time lag for taking decision is due to rough NeRF generated land covering minute details of surface cracks with NeRF generated blue cylinder as shown in Map-2. The average linear velocity remains around 0.12 m/s to 0.13 m/s.

5.6.4 COMPARISON OF ANGULAR CMD-VELOCITY VS ANGULAR NAVIGATIONAL VELOCITY WITH DEPTH CAMERA

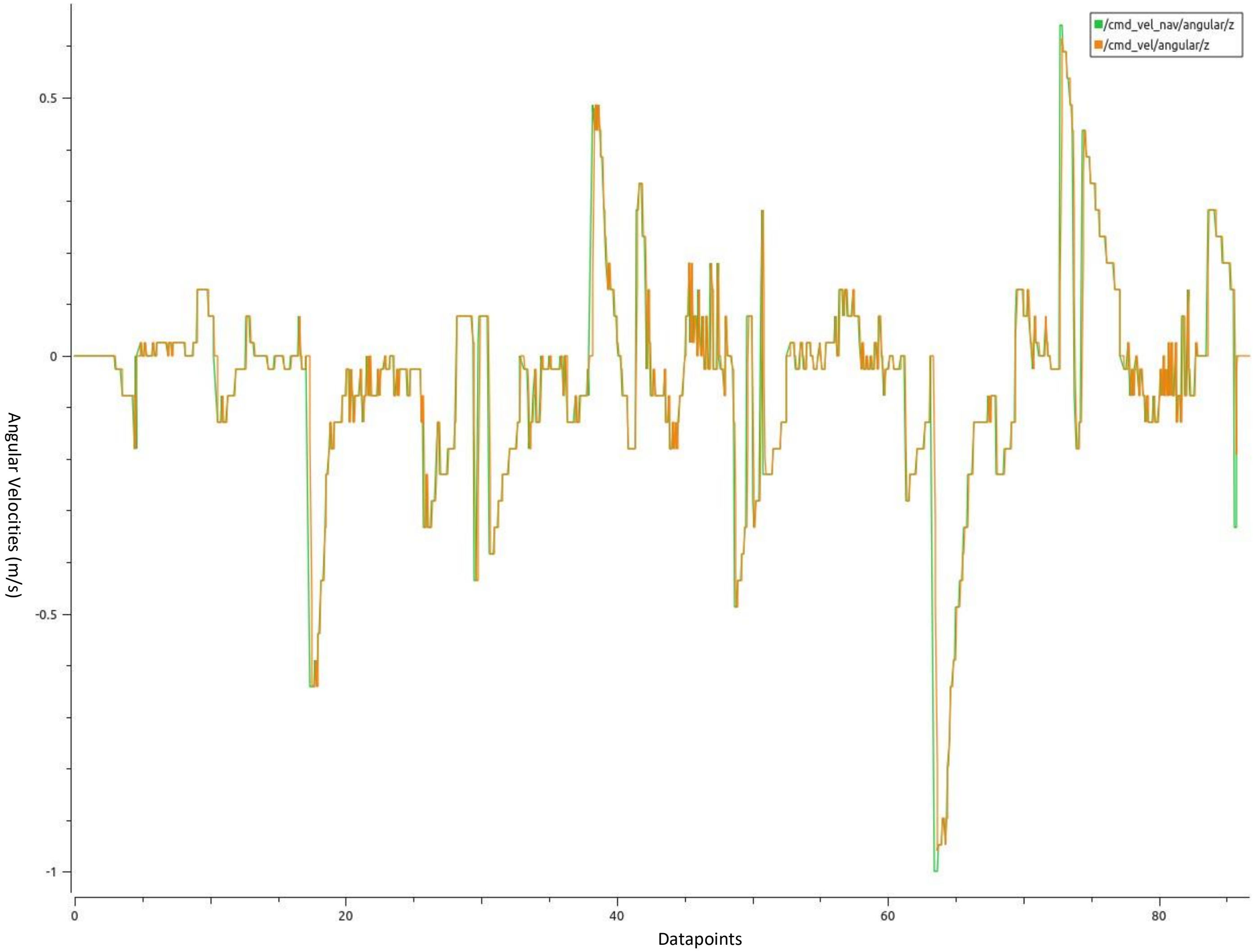


FIGURE-30: CMD VS NAVIGATTION ANGULAR VELOCITIES IN EXPERIMENT-6

In Experiment-6 the average angular velocities remain mostly around -0.04m/s to -0.05 m/s. However, overall synchronization of cmd velocities and navigational velocities can be observed as a positive aspect.

5.6.5 OVERALL STATISTICS OF COMPARISON ANGULAR AND LINEAR CMD-VELOCITIES VS NAVIGATIONAL VELOCITIES WITH DEPTH CAMERA

S. No	Velocities (m/s)	Min	Max	Average
1	/cmd_vel/angular/z	-0.96	0.614359	-0.04483
2	/cmd_vel/linear/x	0	0.22	0.126997
3	/cmd_vel_nav/angular/z	-1	0.641026	-0.04809
4	/cmd_vel_nav/linear/x	0	0.22	0.130277

TABLE-8: EXPERIMENT-6 VELOCITIES STATISTICS

5.7 RESULTS AND DISCUSSION

As discussed earlier in research methodology how the aforementioned six experiments are evaluated. Therefore, a table below shows the comparison of all parameters for subject experiments:

	Map Quality*	Object Avoidance*	Time elapsed during Navigation (sec)	Average Linear Navigational Velocity (m/s)
Experiment-1	5	5	180	0.16
Experiment-2	4	5	193	0.17
Experiment-3	5	5	128	0.18
Experiment-4	4	3	120	0.11
Experiment-5	3	4	95	0.12
Experiment-6	5	4	87	0.13

TABLE-9: SUMMARY OF ALL EXPERIMENTAL RESULTS

* Map Quality and Object Avoidance are qualitative measures and subjective in nature. The system is based on scale from 1 to 5. Where 5 indicates Excellent and 1 indicates poor and scores of 4,3,2 are Good, Average and below Average respectively.

The table-9 shows some interesting results as it can be noticed that in terms of linear velocities where infrared sensor configurations are used in Experiment-3 and Experiment-6 respectively, the highest values are obtained. Subsequently least time is taken to achieve the given Navigational goals. In addition, object avoidance and Map building qualities of said sensor configurations are also highest. Following the infrared configuration, the results of depth camera configuration are better than that of laser configuration except in Map building domain.

On the other hand, comparing built-in vs NeRF assisted simulation scenarios, in Experiment-3, the simulation scenario uses pixels to represent the environment. This means that the simulated surfaces are relatively smooth and lack fine details. However, in Experiment-6, the simulation scenario employs NeRF, which captures very detailed information about the surfaces, including tiny cracks and imperfections. As a result, when running SLAM and navigation operations in these different scenarios, there are noticeable differences. Such as in Experiment-3, where the environment is represented with pixels, SLAM operations are quicker because the environment is simpler and smoother. Though, in Experiment-6, where NeRF is used to capture intricate surface details, SLAM operations take longer because the system has to process all those fine details. However, despite the longer processing time, the navigation operations and path planning in Experiment-6 benefit from this detailed information. Because NeRF captures minute details of the environment, the navigation system can create more accurate representations and make better decisions about how to navigate through it. Also, it is observed during Navigation in NeRF world scenarios that intermittent Navigational targets should be given in shorter distances to gain optimum Navigational performance with less computational burden. Thus, aforementioned experiments support the hypothesis that incorporating the NeRF technology could better help differential drive robots in understanding the real-world scenarios and making response accordingly as compared to built-in pixel or voxel-based representation of the simulation software, while undertaking SLAM and Autonomous Navigation in ROS2. Furthermore, it is also evident from comparing different sensor configurations that infrared sensors could potentially be utilized as a better alternative to Lidar and depth camera in indoor regions with low light intensity.

Chapter 6: Conclusion and Recommendation

6.1 CONCLUSION

In conclusion, the use of NeRF (Neural Radiance Fields) in robotic navigation holds great potential for improving the capabilities of autonomous robots. By leveraging NeRF, robots can benefit from advanced 3D scene reconstruction and representation, allowing them to navigate and interact with their environment more effectively. Additionally, the utilization of IR cameras could provide a comparatively economical solution for navigating in darker regions with low luminous intensity levels. These sensors, either used independently or in combination with other sensors, can enhance the robot's perception and enable it to navigate safely and efficiently in challenging lighting conditions. The combination of ROS2 (Robot Operating System 2) and Gazebo, along with NeRF modules may offer an efficient and cost-effective platform for conducting research and development in robotics. ROS2 provides a robust framework for integrating various components and modules, facilitating the development and deployment of complex robotic systems. Gazebo, as a powerful simulation environment, allows researchers to simulate and test their navigation algorithms and strategies in a realistic and controlled manner. The research's conclusions and findings might be unique to the settings and circumstances taken into account in the study. Furthermore, computational power limitation is a notable hurdle for getting optimum results. Simulated settings can sometime be unable to accurately represent the complexity and variety of real-world circumstances. Potential differences in lighting, object appearances, and environmental elements that may affect the operation of the navigation systems in use should be taken into consideration in the research. Overall, the use of NeRF, along various sensors and the combination of ROS2 and Gazebo simulator presents a promising approach for enhancing robotic navigation capabilities. These technologies and methodologies can contribute to improved perception, mapping, obstacle avoidance, and overall navigation performance of autonomous robots, enabling them to operate more effectively and efficiently in various environments, including low light indoor regions. Continued research and development in this area can lead to significant advancements in robotic navigation, bringing us closer to the realization of fully autonomous and adaptive robotic systems.

6.2 Recommendations and Future Directions

Based on the findings and insights discussed, there are several key recommendations and future directions for enhancing robotic navigation using NeRF with different sensors, and the combination of ROS2, and Gazebo Simulator. Firstly, further research and development should be conducted to explore the full potential of NeRF in robotic navigation. This includes investigating ways to improve the scalability, efficiency, and robustness of NeRF algorithms. Additionally, exploring the integration of NeRF with other sensor modalities and navigation techniques can lead to more comprehensive and accurate scene understanding for robotic navigation. Secondly, the utilization of IR sensors trained on NeRF shows promise for navigation in darker regions with low luminous intensity levels. To maximize their performance and capabilities, it is recommended to optimize the training process on NeRF and explore sensor fusion techniques with other sensors such as LiDAR and visual cameras. Developing advanced algorithms for obstacle detection and localization based on the combined sensor data can further enhance navigation performance. Thirdly, the combination of ROS2, Gazebo, and NeRF provides an efficient and cost-effective platform for robotic research and development. To leverage its potential, it is recommended to continue improving the integration and compatibility between these tools. This involves enhancing the ROS2 framework for seamless communication and data exchange with Gazebo simulations, optimizing the simulation capabilities of Gazebo for realistic navigation scenarios, and expanding the functionalities and support for Pytorch-based models and algorithms within the framework. Lastly, collaborative research efforts and standardization play a crucial role in advancing robotic navigation. Encouraging collaboration and knowledge sharing among researchers, developers, and industry experts can lead to accelerated progress in the field. Establishing standardized datasets, benchmarks, and evaluation metrics can facilitate comparative studies and enable the development of robust navigation algorithms. By working together, the field can identify common challenges and formulate best practices in utilizing NeRF, IR+PIR sensors, and the ROS2+Gazebo+Pytorch framework for robotic navigation. In conclusion, by focusing on these recommendations and future directions, the field of robotic navigation can continue to evolve and make significant advancements in improving the navigation capabilities of autonomous robots. The integration of advanced scene representation techniques, sensor fusion, and efficient development frameworks holds great promise for enabling robots to navigate more effectively and autonomously in a wide range of environments.

References

[1] J. Zhang, F. Keramat, X. Yu, D. M. Hernández, J. P. Queralta, T. Westerlund, “Distributed robotic systems in the edge-cloud continuum with ros 2: A review on novel architectures and technology readiness.” Seventh International Conference on Fog and Mobile Edge Computing (FMEC). IEEE, 2022. doi: 10.1109/FMEC57183.2022.10062523.

[2] T. Blass, A. Hamann, R. Lange, D. Ziegenbein, B. B. Brandenburg, “Automatic latency management for ros 2: Benefits, challenges, and open problems.” 27th Real-Time and Embedded Technology and Applications Symposium (RTAS). IEEE. IEEE, 2021. doi: 10.1109/RTAS52030.2021.00029.

[3] Z. Wang, et al., *‘NeRF--: Neural radiance fields without known camera parameters.’ arXiv Preprint ArXiv:2102.07064*, 2021.

[4] T. Kattenborn, J. Leitloff, F. Schiefer, S. Hinz, “Review on Convolutional Neural Networks (CNN) in vegetation remote sensing,” *ISPRS J. Photogramm.*, vol. 173, pp. 24–49, 2021. doi: 10.1016/j.isprsjprs.2020.12.010.

[5] F. Gul, W. Rahiman, and S. S. N. Nazli Alhady, “A comprehensive study for robot navigation techniques,” *Cogent Eng.*, vol. 6, no. 1, p. 1632046, 2019. doi: 10.1080/23311916.2019.1632046.

[6] A. N. Sivakumar, et al., “Learned visual navigation for under-canopy agricultural robots,” arXiv Preprint ArXiv:2107.02792, 2021. doi: 10.15607/RSS.2021.XVII.019.

[7] S. Gupta, et al., “Unifying map and landmark based representations for visual navigation,” arXiv Preprint ArXiv:1712.08125, 2017.

[8] K. Mo, et al., “The adobeindoornav dataset: Towards deep reinforcement learning based real-world indoor robot visual navigation,” arXiv Preprint ArXiv:1802.08824, 2018.

[9] Qiu, Yiding, A. Pal, and H. I. Christensen, “Target driven visual navigation exploiting object relationships arXiv Preprint ArXiv:2003.067492,” vol. 7, 2020.

[10] K. Nguyen and H. Daumé III, “Help, anna! visual navigation with natural multimodal assistance via retrospective curiosity-encouraging imitation learning,” arXiv Preprint ArXiv:1909.01871, 2019. doi: 10.18653/v1/D19-1063.

[11] F. Zhu, et al., “Deep learning for embodied vision navigation: A survey,” arXiv Preprint ArXiv:2108.04097, 2021.

[12] H. Wang, et al., “Collaborative visual navigation,” arXiv Preprint ArXiv:2107.01151, 2021.

[13] K. Chen, et al., “A behavioural approach to visual navigation with graph localization networks,” arXiv Preprint ArXiv:1903.00445, 2019.

[14] Ben Mildenhall, et al., “NeRF: Representing scenes as neural radiance fields for view synthesis,” *Commun. ACM*, vol. 65, no. 1, pp. 99–106, 2021.

[15] M. Adamkiewicz, et al., “Vision-only robot navigation in a neural radiance world,” *IEEE Robot. Autom. Lett.*, vol. 7, no. 2, pp. 4606–4613, 2022. doi: 10.1109/LRA.2022.3150497.

[16] Y. Tang, et al., “Perception and navigation in autonomous systems in the era of learning: A survey,” *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 34, no. 12, 9604–9624, 2023. doi: 10.1109/TNNLS.2022.3167688.

[17] J. Sun, et al., “NeRF-loc: Transformer-based object localization within neural radiance fields,”, *IEEE Robot. Autom. Lett.*, vol. 8, no. 8, 5244–5250, 2022. doi: 10.1109/LRA.2023.3293308.

[18] A. Byravan, et al., “Nerf2real: Sim2real transfer of vision-guided bipedal motion skills using neural radiance fields,” arXiv Preprint ArXiv:2210.04932, 2022. doi: 10.1109/ICRA48891.2023.10161544.

[19] K. Gao, et al., “Nerf: Neural radiance field in 3d vision, a comprehensive review,” arXiv Preprint ArXiv:2210.00379, 2022.

[20] D. Maggio, M. Abate, J. Shi, C. Mario, L. Carlone, “Loc-NeRF: Monte Carlo localization using neural radiance fields,” arXiv Preprint ArXiv:2209.09050, 2022. doi: 10.1109/ICRA48891.2023.10160782.

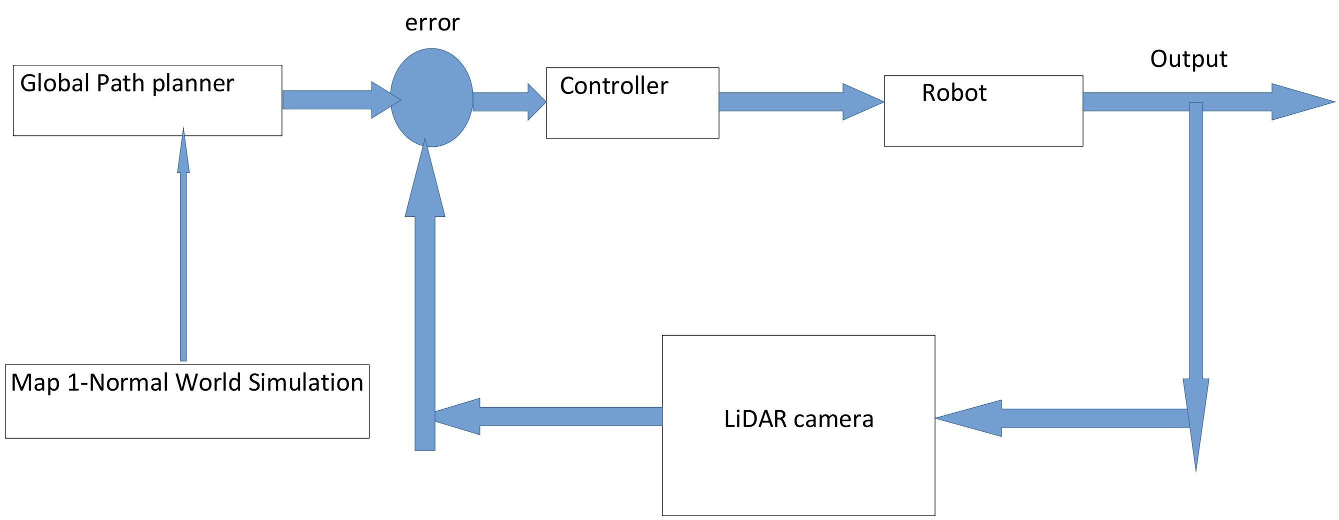
[21] M. Tong, C. Dawson, C. Fan, “Enforcing safety for vision-based controllers via Control Barrier Functions and Neural Radiance Fields,” arXiv Preprint ArXiv:2209.12266, 2022. doi: 10.1109/ICRA48891.2023.10161482.
[22] S. Klenk, L. Koestler, D. Scaramuzza, D. Cremers, “E-nerf: Neural radiance fields from a moving event camera,” *IEEE Robot. Autom. Lett.*, vol. 8, no. 3, pp. 1587–1594, 2023. doi: 10.1109/LRA.2023.3240646.
[23] T. Chen, P. Culbertson, and M. Schwager, “CATNIPS: collision avoidance through neural implicit probabilistic scenes arXiv Preprint ArXiv:2302.12931,”, *IEEE Trans. Robot.*, 1–18, 2023. doi: 10.1109/TRO.2024.3386394.
[24] M. Tancik, et al., “Block-nerf: Scalable large scene neural view synthesis,”, in *Proc. IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2022, 8238–8248. doi: 10.1109/CVPR52688.2022.00807.
[25] K. Lin and B. Yi, “Active view planning for radiance fields,” *Robot. Sci. Syst.*, 2022.
[26] S. Macenski, T. Foote, B. Gerkey, C. Lalancette, W. Woodall, “Robot Operating System 2: Design, architecture, and uses in the wild,” *Sci. Robot.*, vol. 7, no. 66, p. eabm6074, 2022. doi: 10.1126/scirobotics.abm6074.
[27] M. Martini, A. Eirale, S. Cerrato, M. Chiaberge, “PIC4rl-gym: A ROS2 modular framework for Robots Autonomous Navigation with Deep Reinforcement Learning,” arXiv Preprint ArXiv:2211.10714, 2022. doi: 10.1109/ICCCR56747.2023.10193996.
[28] V. Mayellaro, Person-aware autonomous navigation for an indoor sanitizing robot in ROS2, [Diss]. Politecnico di Torino, 2022.
[29] N. Pérez-Higueras, R. Otero, F. Caballero, L. Merino, “HuNavSim: A ROS 2 human navigation simulator for benchmarking human-aware robot navigation,”, *IEEE Robot. Autom. Lett.*, vol. 8, no. 11, 7130–7137, 2023. doi: 10.1109/LRA.2023.3316072.
[30] S.-J. Fang, “Simulation analysis of indoor orbital inspection robot based on Gazebo.” International Conference on Intelligent Traffic Systems and Smart City (ITSSC 2021), vol. 12165. SPIE, 2022. doi: 10.1117/12.2628608.
[31] M. Poggi, P. Z. Ramirez, F. Tosi, S. Salti, S. Mattoccia, L. D. Stefano, “Cross-spectral neural radiance fields,” arXiv Preprint ArXiv:2209.00648, 2022. doi: 10.1109/3DV57658.2022.00071.
[32] R. Hadsell, et al., “Learning long-range vision for autonomous off-road driving,” *J. Field Robot.*, vol. 26, no. 2, pp. 120–144, 2009. doi: 10.1002/rob.20276.
[33] C. Chen, Q. Chen, J. Xu, V. Koltun, “Learning to see in the dark,”, in *Proc. IEEE Conference on Computer Vision and Pattern Recognition*, 2018, 3291–3300. doi: 10.1109/CVPR.2018.00347.
[34] S. Lee and J. B. Song, “Mobile robot localization using infrared light reflecting landmarks.” International Conference on Control., Automation and Systems. IEEE, 2007.
[35] M. Poggi, P. Z. Ramirez, F. Tosi, S. Salti, S. Mattoccia, L. D. Stefano, “Cross-spectral neural radiance fields,” arXiv Preprint ArXiv:2209.00648, 2022. doi: 10.1109/3DV57658.2022.00071.
[36] “ROS2 documentation”. Available at: https://docs.ros.org/en/foxy/index.html.
[37] “Gazebo Ignition website”. Available at: https://gazebosim.org/api/gazebo/4.0/tutorials.html.
[38] N. Rahaman, et al., *‘On the Spectral Bias of Neural Networks. arXiv e-prints, page.’ arXiv Preprint ArXiv:1806.08734*, 2018.
[39] M. Levoy, “Efficient ray tracing of volume data,” *ACM Trans. Graph.*, vol. 9, no. 3, pp. 245–261, 1990. doi: 10.1145/78964.78965.
[40] A. Krizhevsky, et al., “Imagenet classification with deep convolutional neural networks,” *Adv. Neural Inf. Process. Syst.*, vol. 25, pp. 1097–1105, 2012.
[41] *ROS2 Documentation*. Available at: https://www.engineeringtoolbox.com.
[42] S. Macenski and I. Jambrecic, “SLAM Toolbox: SLAM for the dynamic world,” *J. Open Source Softw.*, vol. 6, no. 61, p. 2783, 2021. doi: 10.21105/joss.02783.

[43] K. Konolige, G. Grisetti, R. Kümmerle, W. Burgard, B. Limketkai, and R. Vincent, "Efficient sparse pose adjustment for 2D mapping,", in *Proc. IEEE/RSJ International Conference on Intelligent Robots and Systems*. IEEE, 2010, pp. 22–29. doi: 10.1109/IROS.2010.5649043.
[44] Mathsworks. Available at: https://www.mathworks.com/discovery/slam.html.
[45] MathWorks. Available at: https://www.mathworks.com/discovery/point-coud.html.
[46] S. Macenski, F. Martín, R. White, and J. Clavero, *The Marathon 2: A Navigation System*. IEEE, 2020. doi: 10.1109/IROS45743.2020.9341207.
[47] "Navigation2 stack documentation". Available at: https://navigation.ros.org/plugins/index.html.

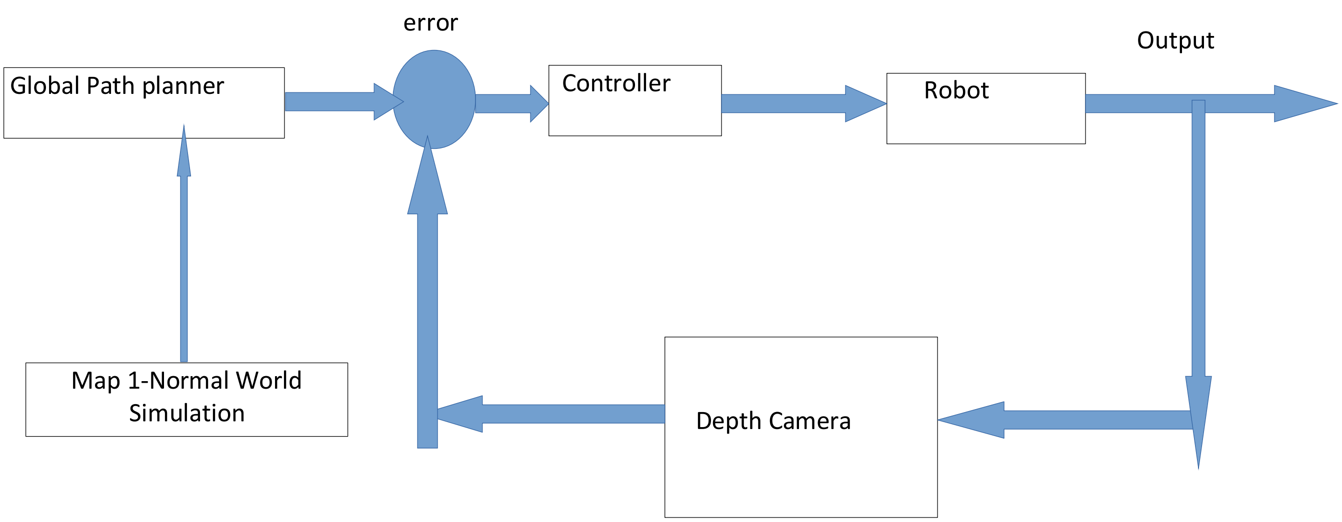
Appendix-A

EXPERIMENTAL SETUP FOR SIMULATION AND ANALYSIS

Experiment-1



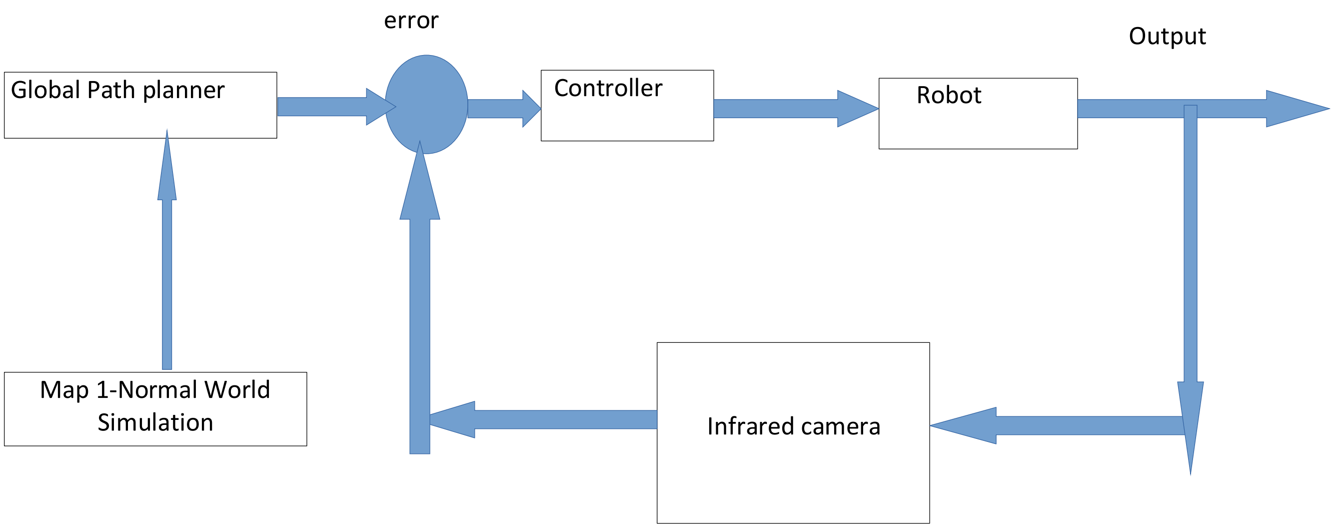
Experiment-2



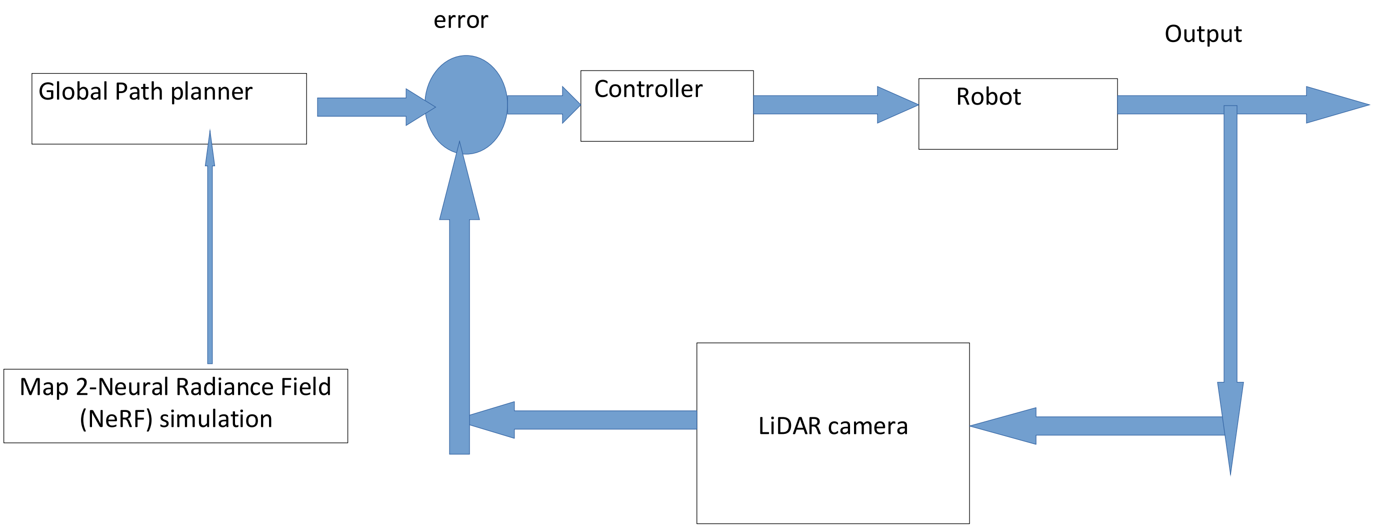
Appendix-A

EXPERIMENTAL SETUP FOR SIMULATION AND ANALYSIS

Experiment-3



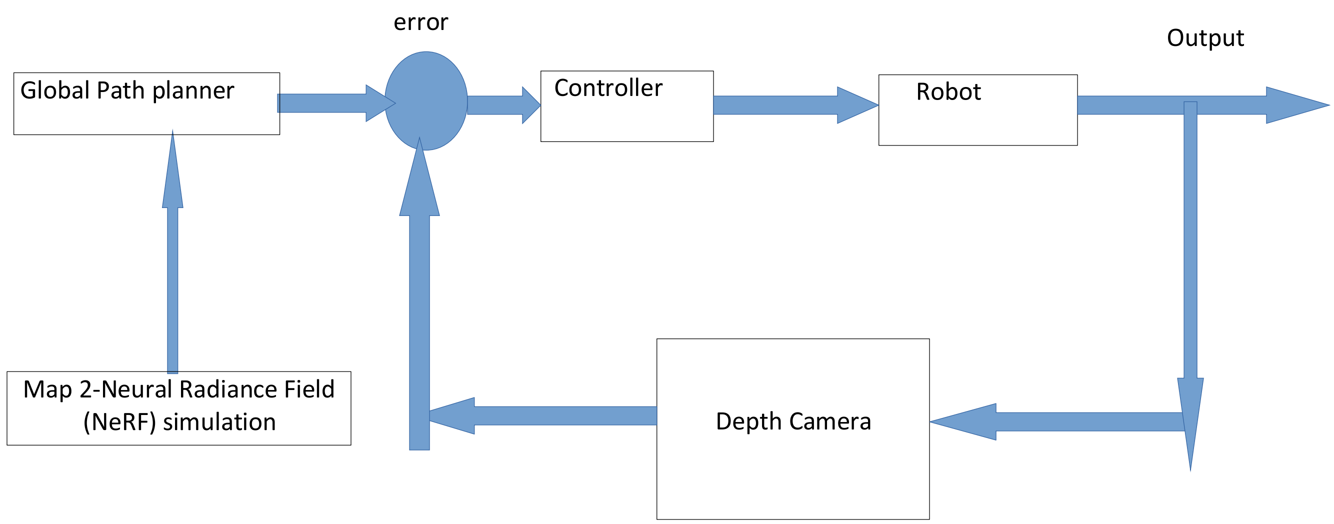
Experiment-4



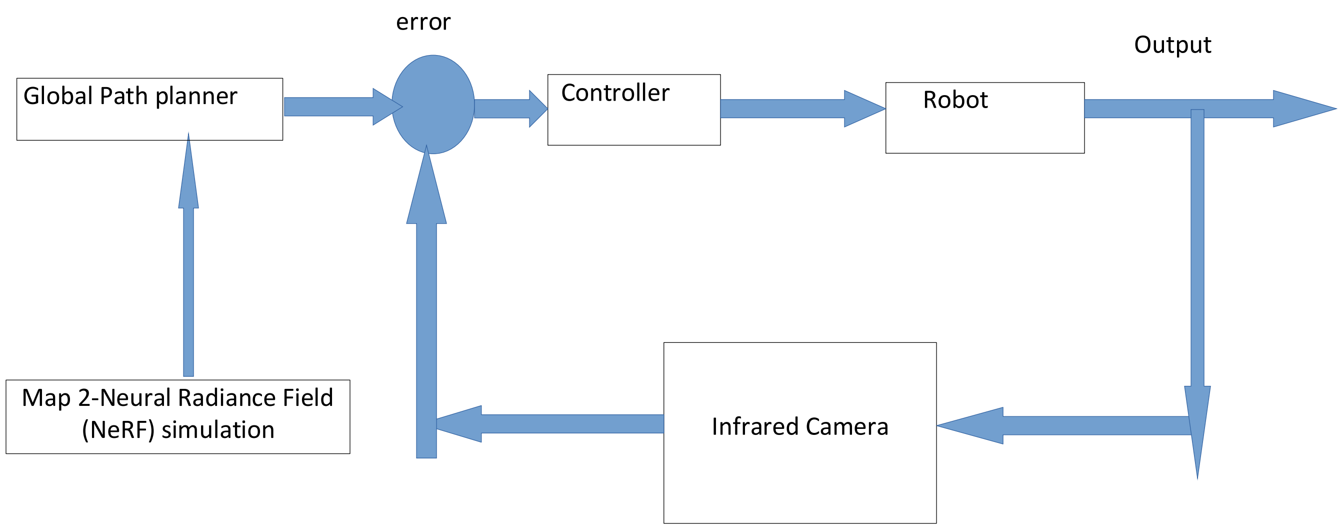
Appendix-A

EXPERIMENTAL SETUP FOR SIMULATION AND ANALYSIS

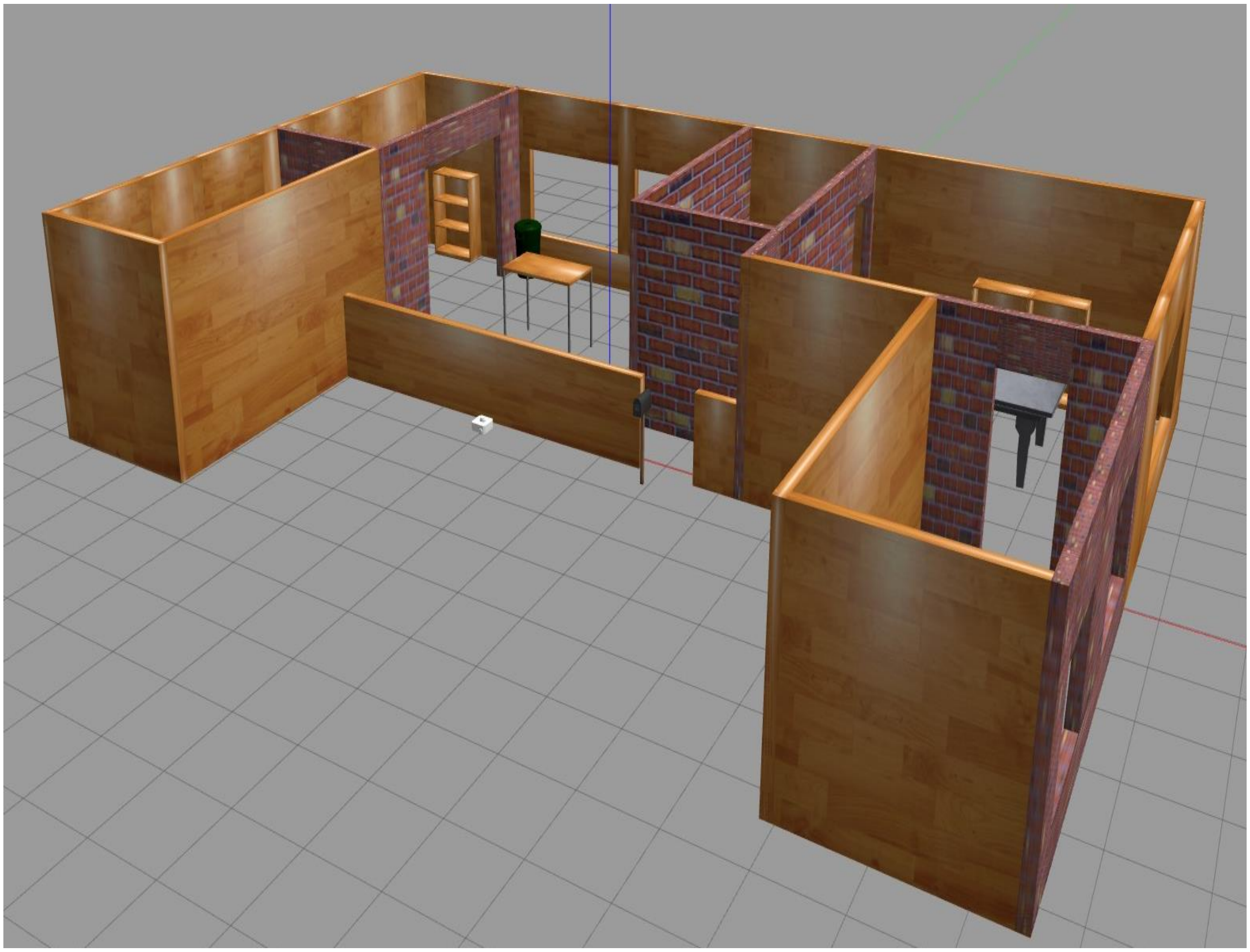
Experiment-5



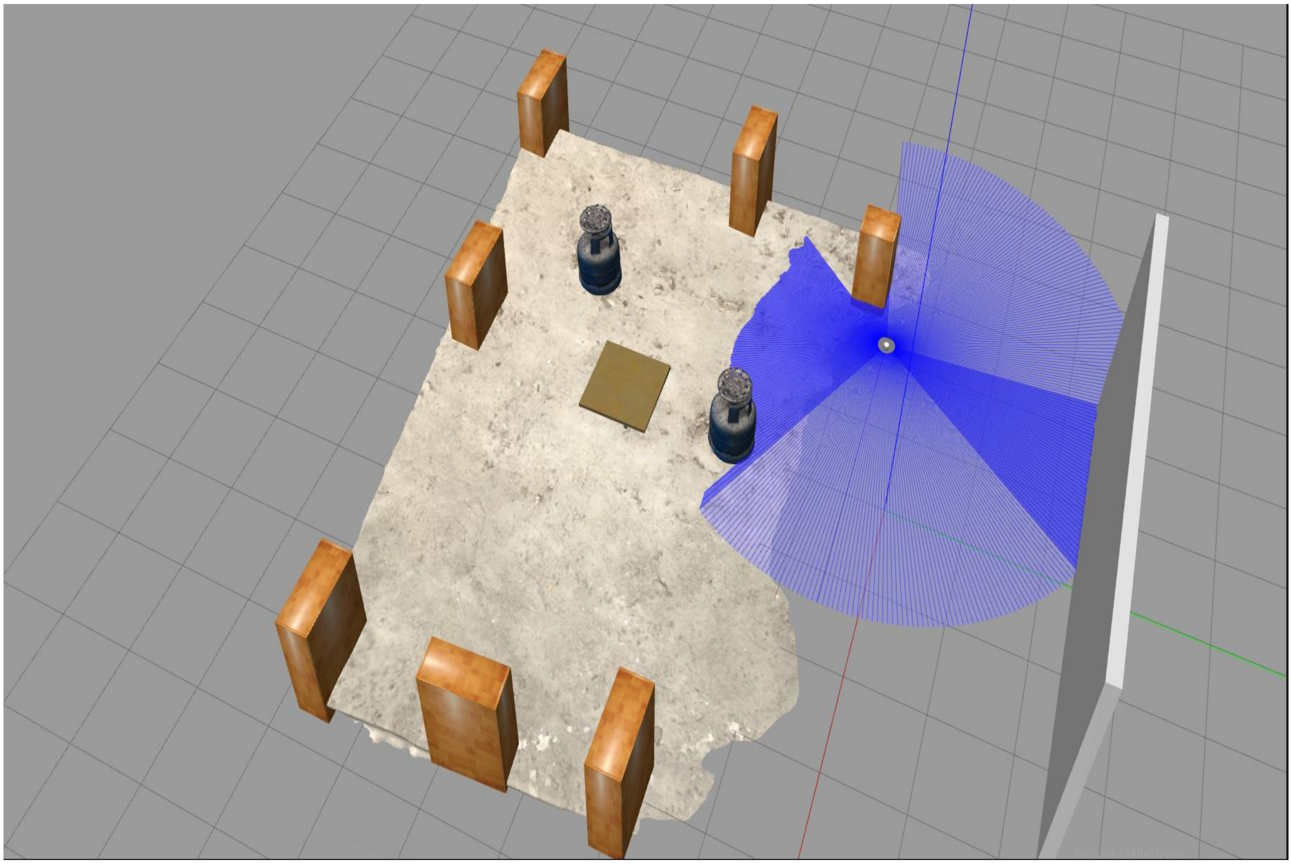
Experiment-6



APPENDIX-B
MAP-1 Entirely Generated in Gazebo-Ignition



APPENDIX-B
MAP-2 Generated via NeRF and attached within gazebo



APPENDIX-C
SENSORS SPECIFICAITONS USED IN EXPERIMENTS

Specification of laser Sensor used in Experiments


```
<sensor name="Laser" type="ray">
        <always_on>true</always_on>
        <visualize>true</visualize>
        <pose>-0.064 0 0.121 0 0 0</pose>
        <update_rate>5</update_rate>
        <ray>
          <scan>
            <horizontal>
              <samples>360</samples>
              <resolution>1.000000</resolution>
              <min_angle>0.000000</min_angle>
              <max_angle>6.280000</max_angle>
            </horizontal>
          </scan>
          <range>
            <min>0.120000</min>
            <max>3.5</max>
            <resolution>0.015000</resolution>
          </range>
          <noise>
            <type>gaussian</type>
            <mean>0.0</mean>
            <stddev>0.01</stddev>
          </noise>
        </ray>
        <plugin name="turtlebot3_laserscan" filename="libgazebo_ros_ray_sensor.so">
          <ros>
            <!-- <namespace>/tb3</namespace> -->
            <remapping>~/out:=scan</remapping>
          </ros>
          <output_type>sensor_msgs/LaserScan</output_type>
          <frame_name>base_scan</frame_name>
        </plugin>
      </sensor>
```


Specifications for both Depth and Infrared Cameras


```
<sensor name="camera" type="depth">
      <!--<sensor name="Infrared" type="ray">-->
        <always_on>true</always_on>
```

APPENDIX-C

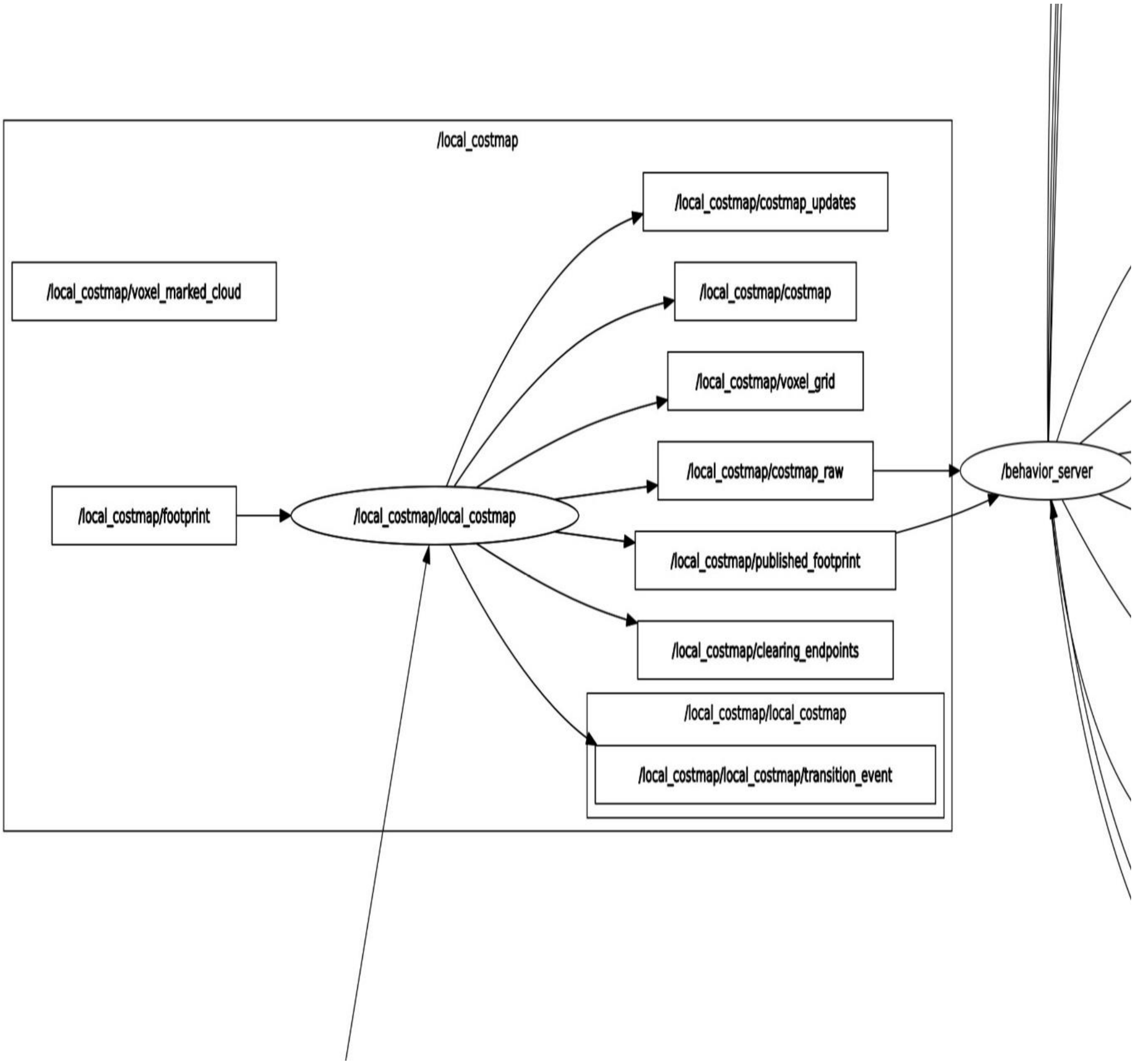
SENSORS SPECIFICAITONS USED IN EXPERIMENTS

```
<visualize>true</visualize>
    <update_rate>30</update_rate>
    <camera name="intel_realsense_r200">
      <horizontal_fov>1.02974</horizontal_fov>
      <image>
        <width>1920</width>
        <height>1080</height>
        <format>B8R8G8</format>
      </image>
      <!--<radiation>INFRARED</radiation>-->
      <topicName>proximity/front_left</topicName>
      <clip>
        <near>0.02</near>
        <far>300</far>
      </clip>
      <noise>
        <type>gaussian</type>
        <!-- Noise is sampled independently per pixel on each frame.
             That pixel's noise value is added to each of its color
             channels, which at that point lie in the range [0,1]. -->
        <mean>0.0</mean>
        <stddev>0.007</stddev>
      </noise>
    </camera>
      <plugin name="camera_driver" filename="libgazebo_ros_camera.so">
      <!--<plugin name="gazebo_ros_range" filename="libgazebo_ros_range.so">-->
        <ros>
          <!-- <namespace>test_cam</namespace> -->
          <!-- <remapping>image_raw:=image_demo</remapping> -->
          <!-- <remapping>camera_info:=camera_info_demo</remapping> -->
        </ros>
        <!-- camera_name>omit so it defaults to sensor name</camera_name-->
        <!-- frame_name>omit so it defaults to link name</frameName-->
        <!-- <hack_baseline>0.07</hack_baseline> -->
      </plugin>
   </sensor>
```

APPENDIX-D
WORK BREAKDOWN STRUCTURE (WBS)

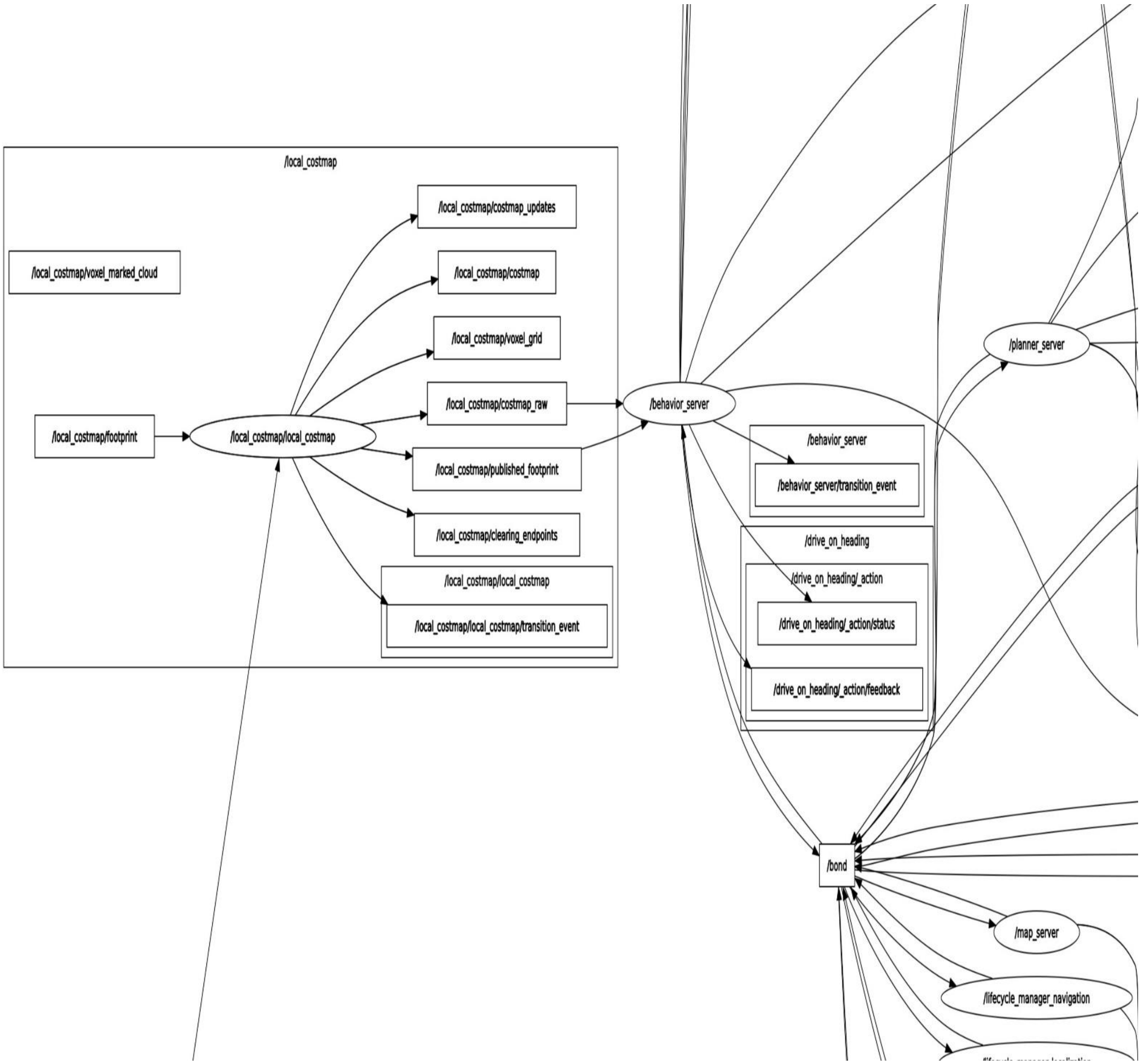
S. No	Main Topics	Sub Topics	Remarks
1a.	Literature Review and Theoretical aspects	Theoretical aspects of Neural Radiance fields (NeRF)and its potential implementation in the field of robotics for improving mapping and Autonomous Navigation	Completed
1b.		Simultaneous localization and Mapping and associated methods/algorithms (SLAM)	Completed
1c.		Navigation and associated methods/Algorithms	Completed
1d.		Sensors (LiDAR, rgb camera, rgb depth-camera and infrared camera)	Completed
2a.	Thesis Report Generation	Abstract	Completed
2b.		Introduction, problem statements, solution statements	Completed
2c.		Literature Review and comparison table on aforementioned topics	Completed
2d.		Conception of experimental setups for comparison and analysis of mapping techniques. It comprises classical vs NeRF based SLAM techniques for optimum Navigation.	Completed
2e		Result and Conclusion	Completed
2f.		Way Forward/future Direction	Completed
3a.	ROS2, Gazebo simulator and research implementation	Understanding the Linux operating system, shell and command line tool	Completed
3b.		Learning the advanced concept of python3 and C++ for ROS2 programming	Completed
3c.		Fundamentals of ROS2 like node, package, custom node, interaction with client library, message, services, topic, publisher, subscriber, rviz, gazebo-ros2 bridging etc.	Completed
3d.		Advanced model integration for mapping like SLAM toolbox	Completed
3e		Development of robot structure (links, joints and associated tfs) with integration of Lidar, camera, depth camera and infrared sensor	Completed
3f.		Mapping of gazebo world with obstacle with all three aforementioned sensors for further SLAM and autonomous navigation	Completed
3g.		Integration of Navigation library and map saver function for smooth navigation of robot in gazebo world	Completed
3i.		Development of NeRF maps and bridging of NeRF libraries with ros2 for further navigation and object detection/recognition	Completed
3j.		Results and comparative analysis of pixel or voxel-based Gazebo maps and NeRF integrated Gazebo maps in terms of SLAM and Autonomous Navigation for different robot's configuration.	Completed

Appendix-'E'
Navigation operation graph

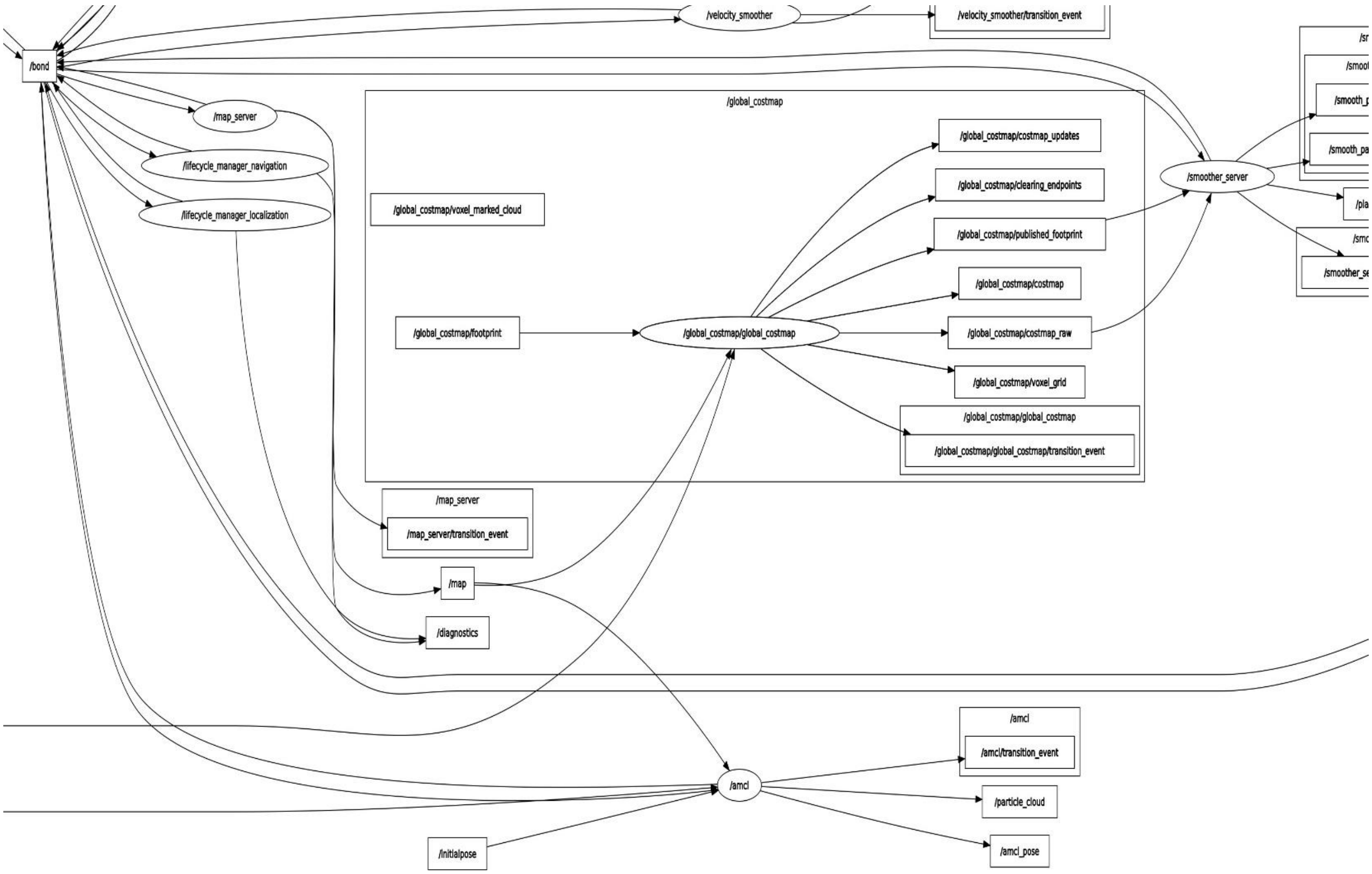


Appendix-'E'
Navigation operation graph

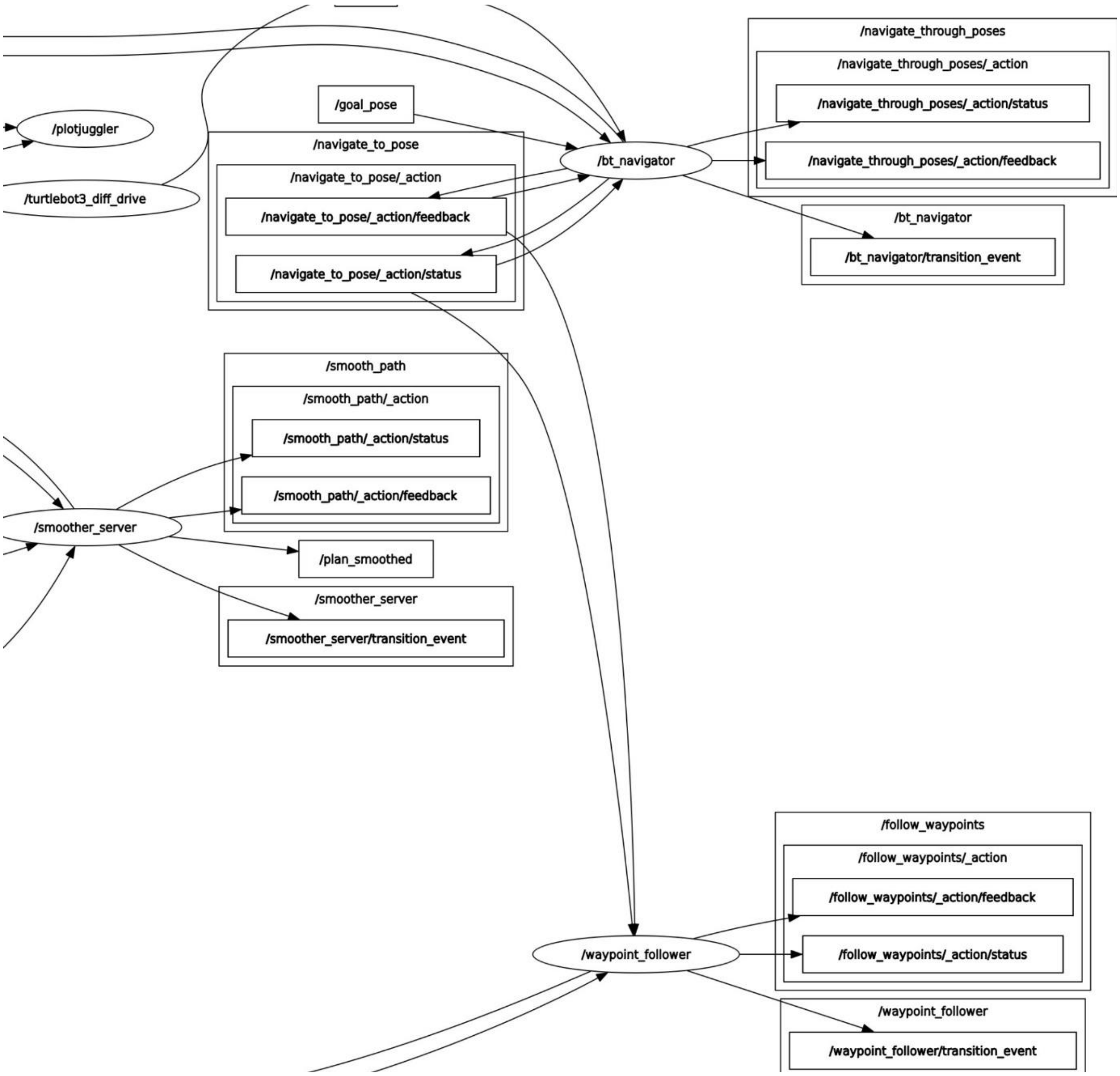
Appendix-'E'
Navigation operation graph



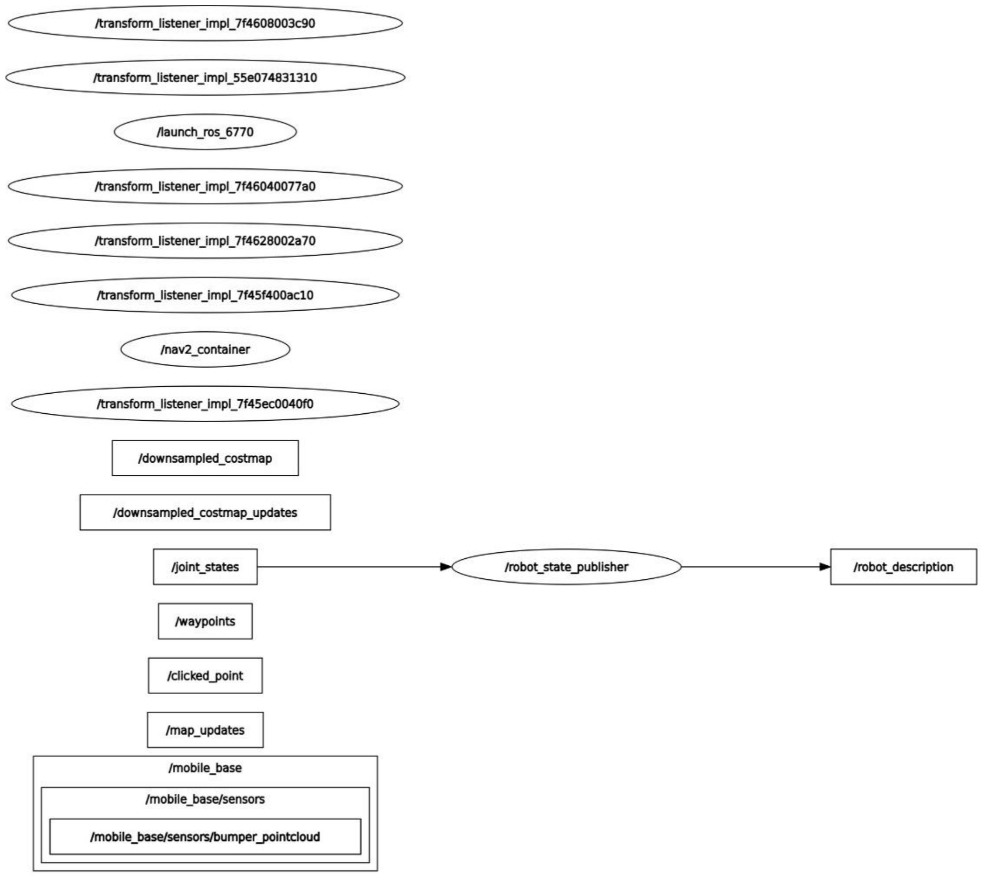
Appendix-'E'
Navigation operation graph



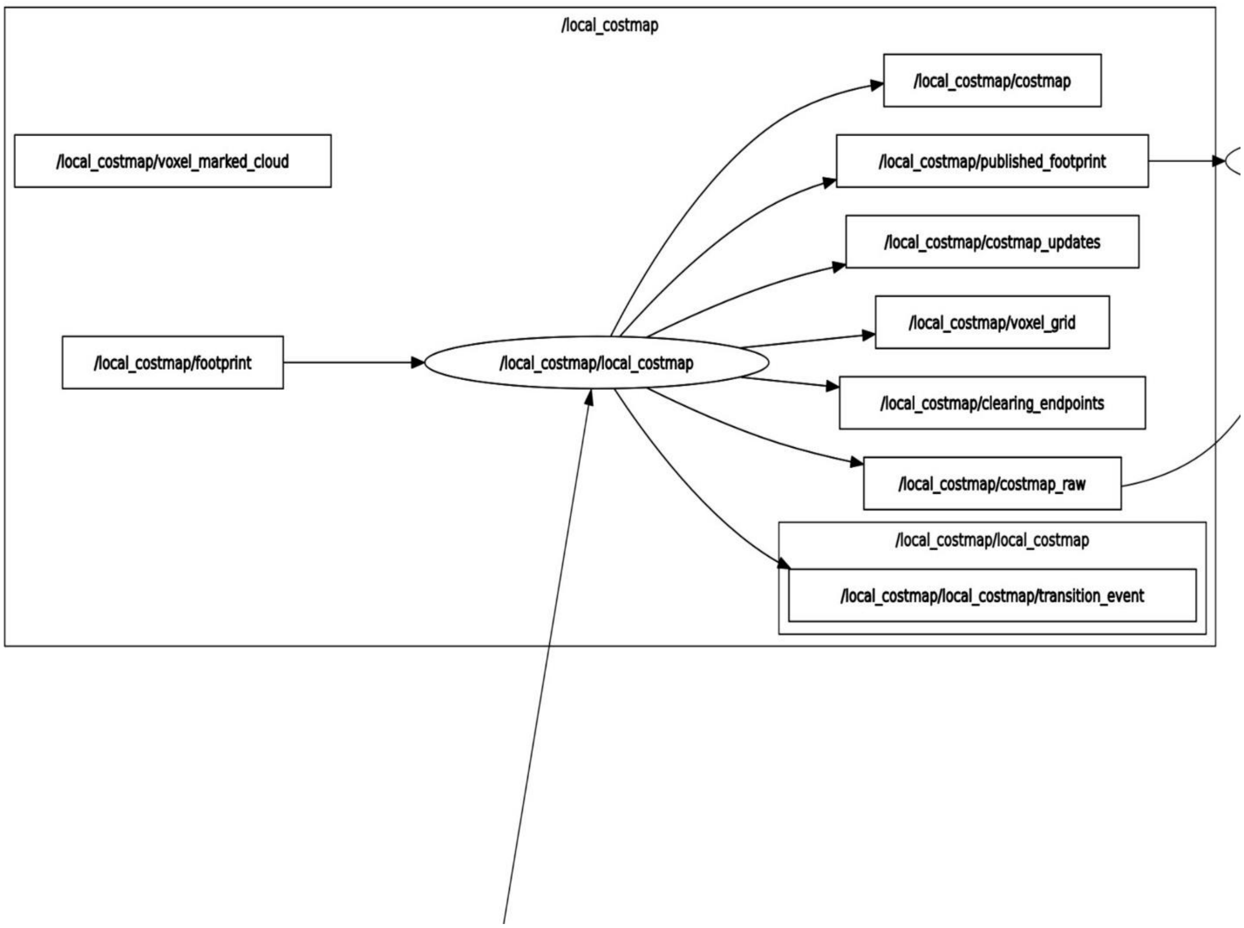
Appendix-'E'
Navigation operation graph



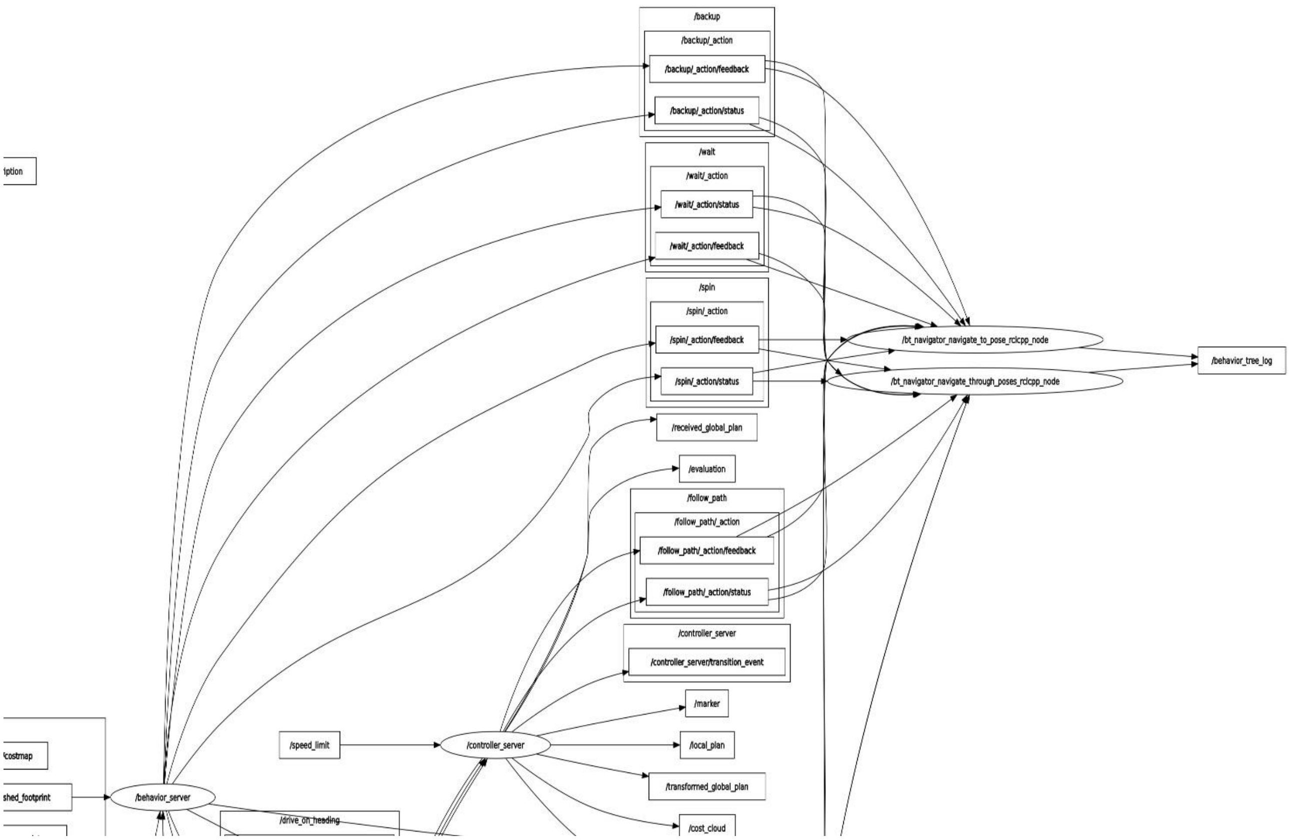
Appendix-'F'
Navigation operation graph



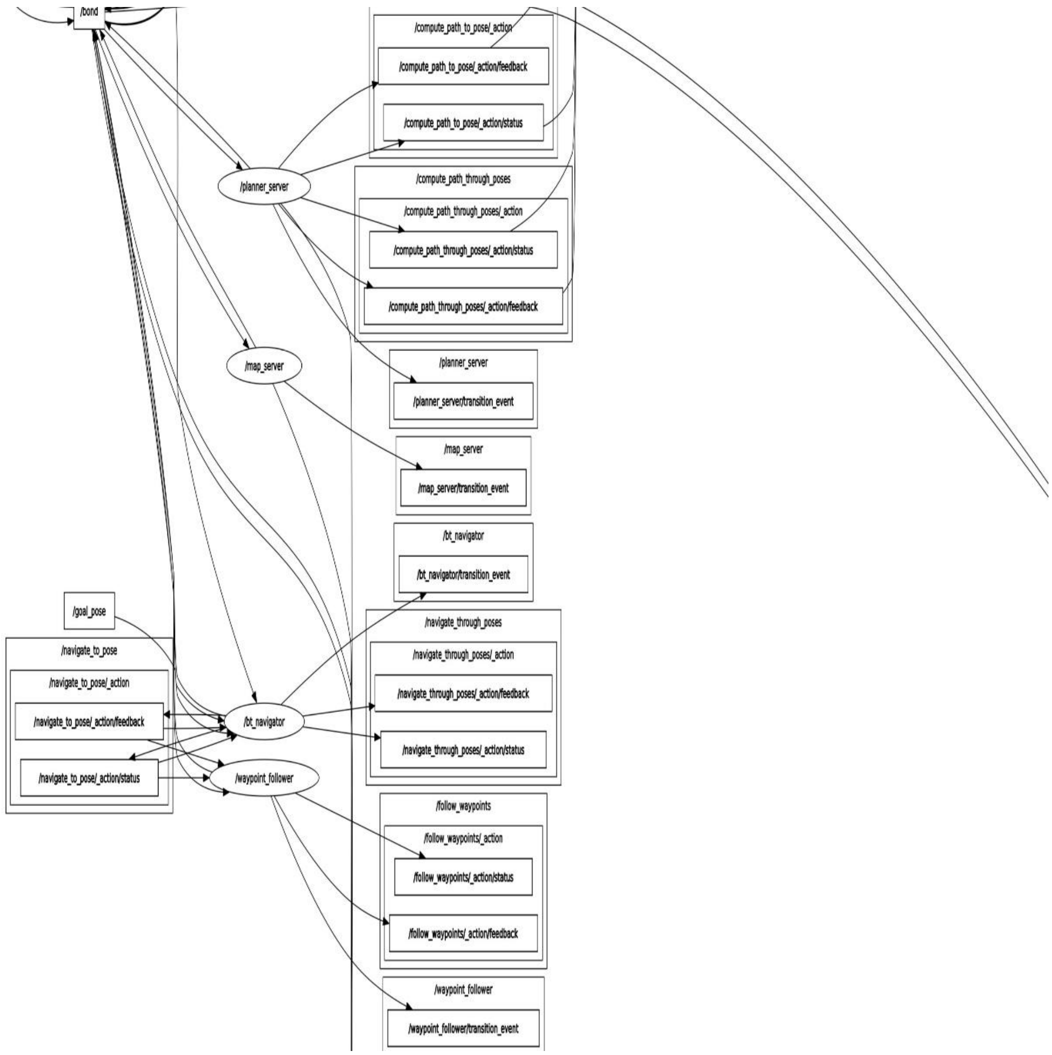
Appendix-'F'
Navigation operation graph



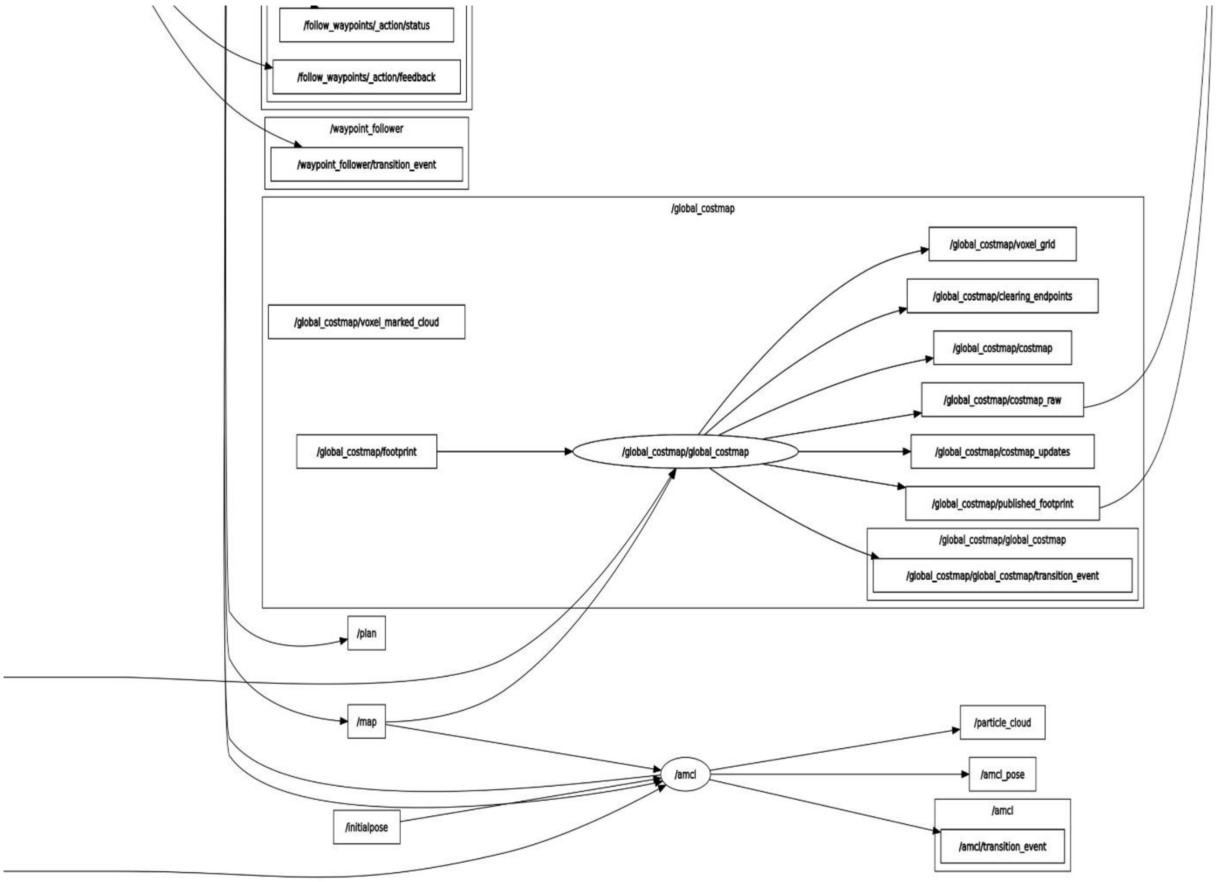
Appendix-'F'
Navigation operation graph



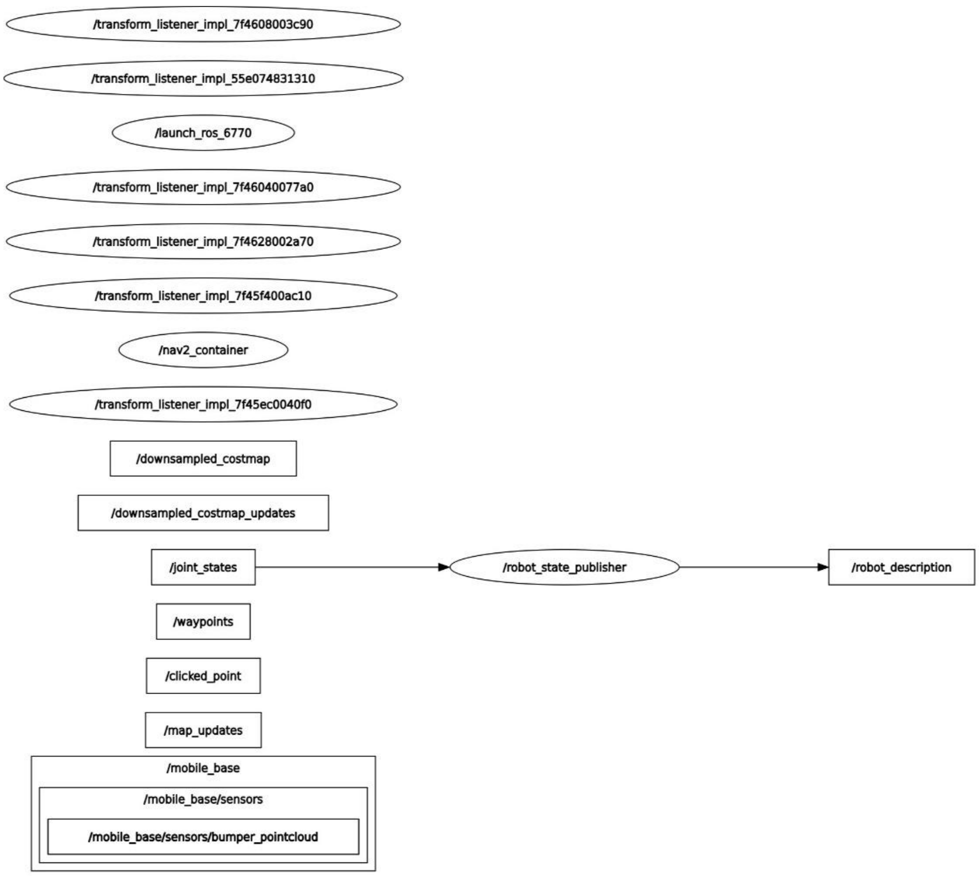
Appendix-'F'
Navigation operation graph



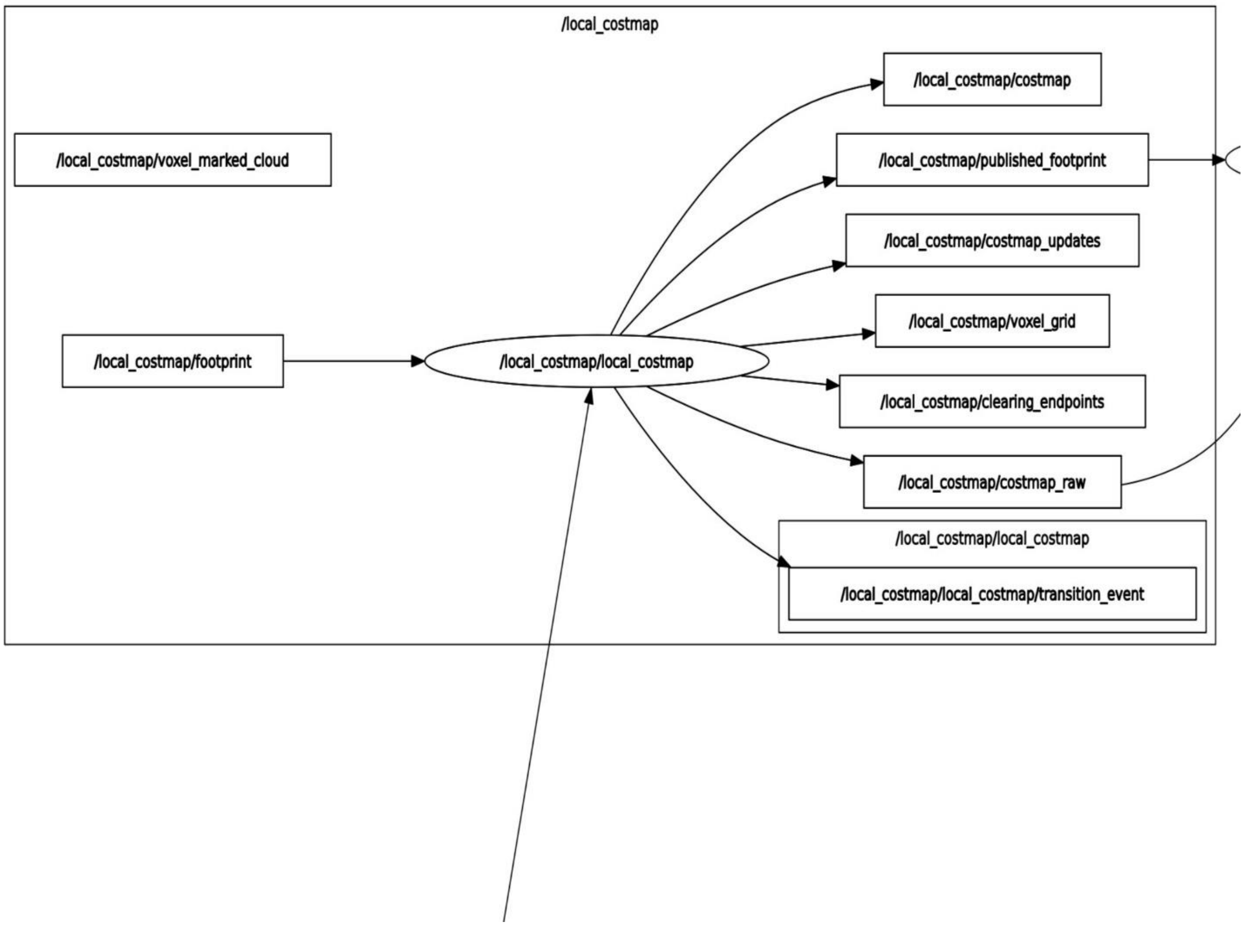
Appendix-'F'
Navigation operation graph



Appendix-‘G’
Navigation operation graph

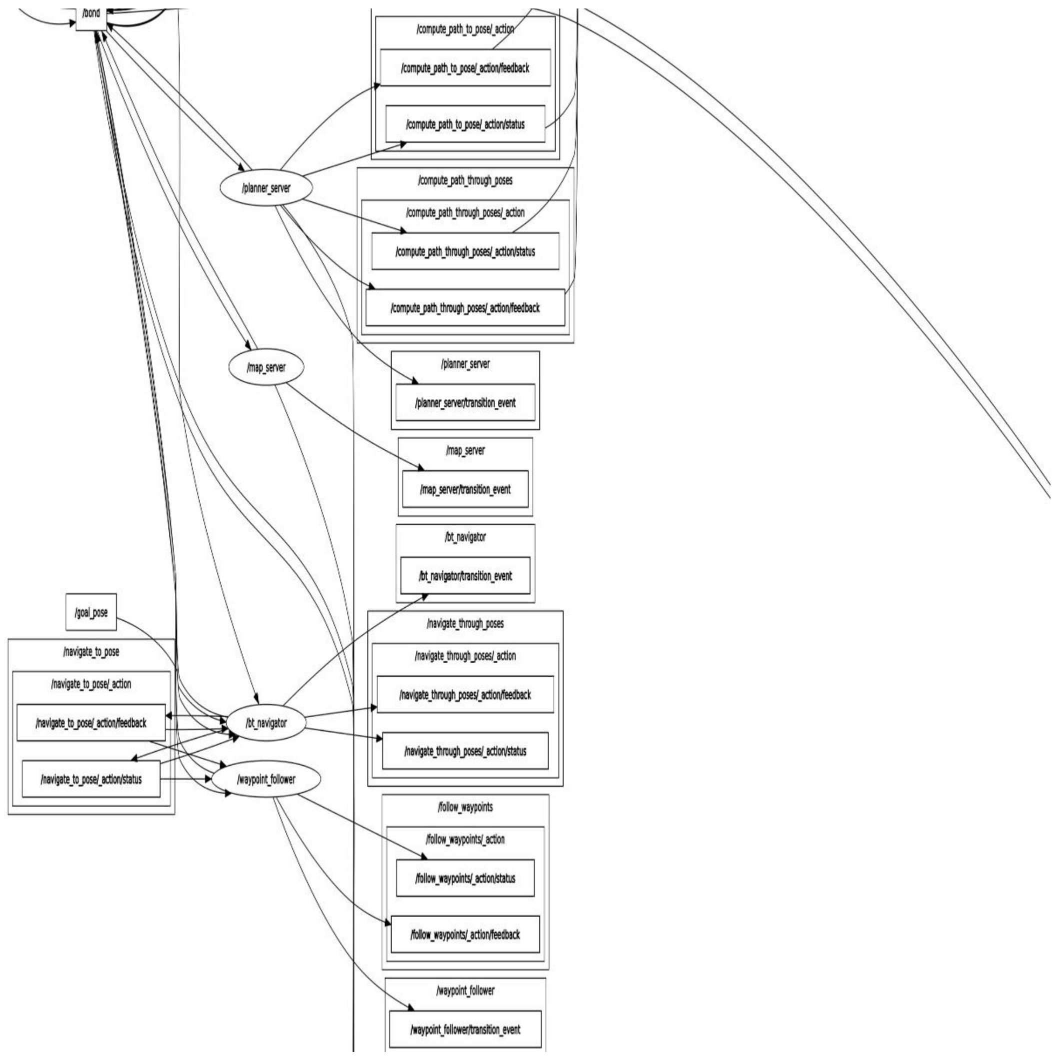


Appendix-'G'
Navigation operation graph



Appendix-'G'
Navigation operation graph

Appendix-'G'
Navigation operation graph



Appendix-'G'
Navigation operation graph