EECE 5554:

ROBOTIC SENSING AND NAVIGATION

FINAL PROJECT

RTAB-MAP SLAM

Presented by:

Atharva Jamsandekar

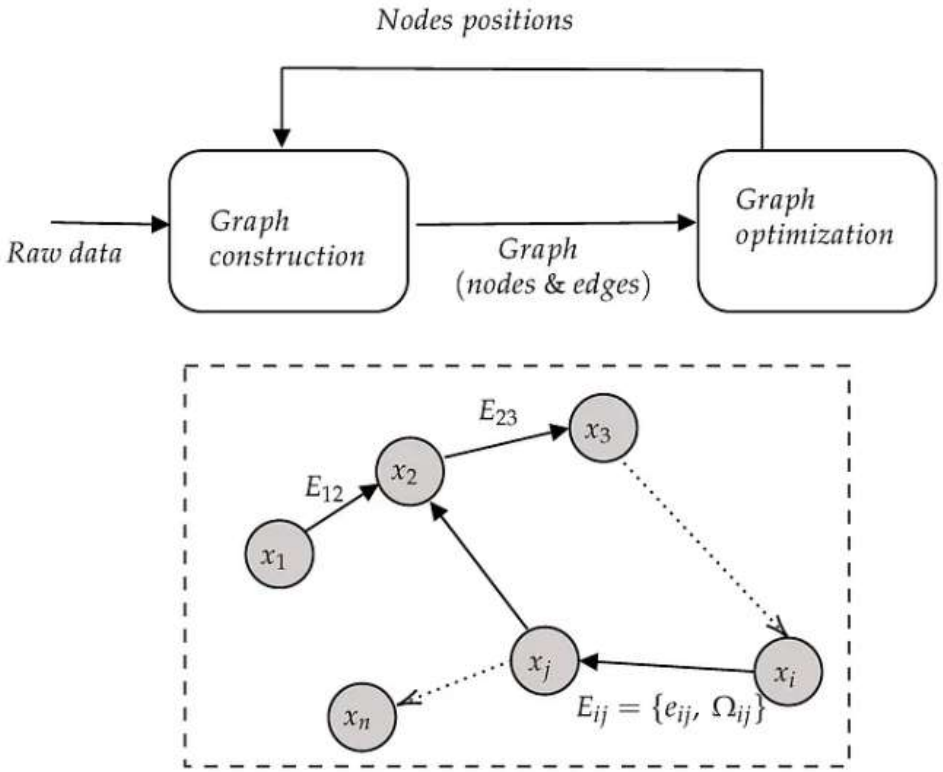
Anusha Manohar

Febin Wilson

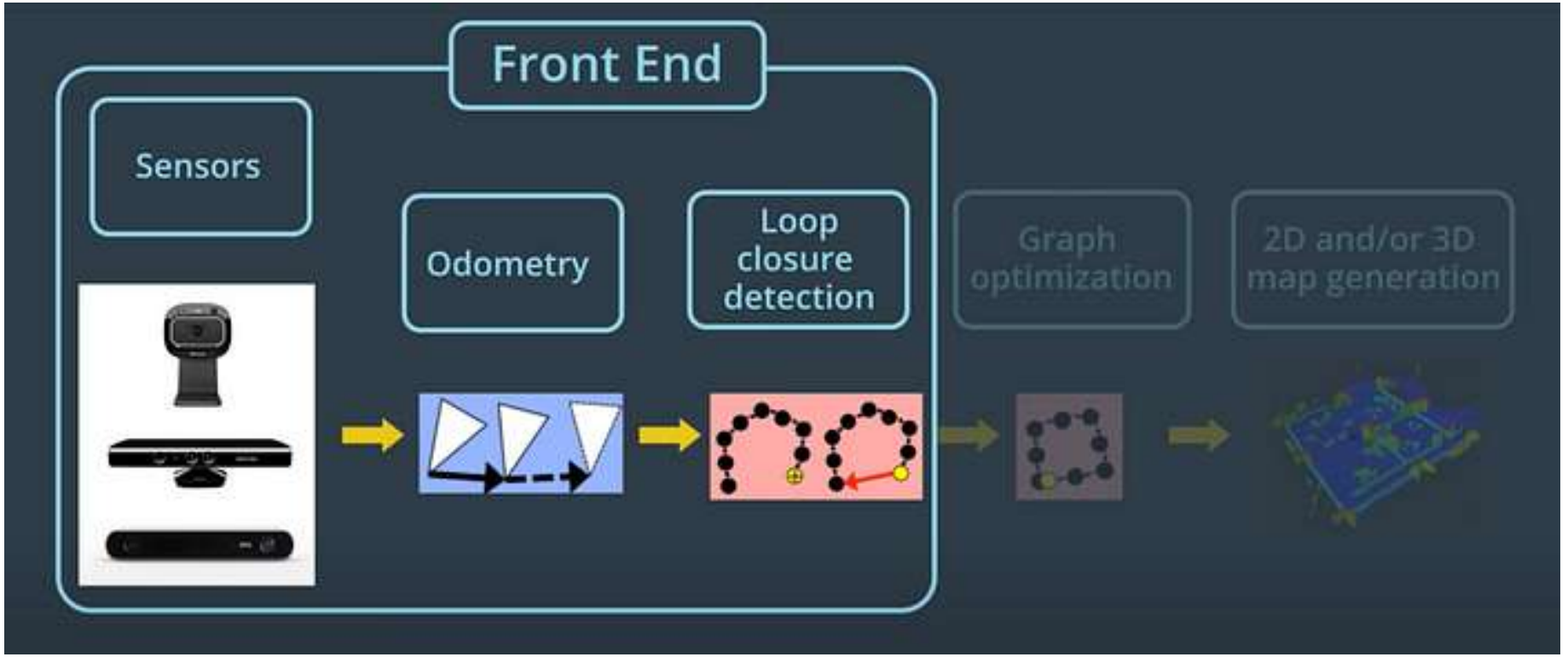
Ruchik Jani

SLAM BASICS

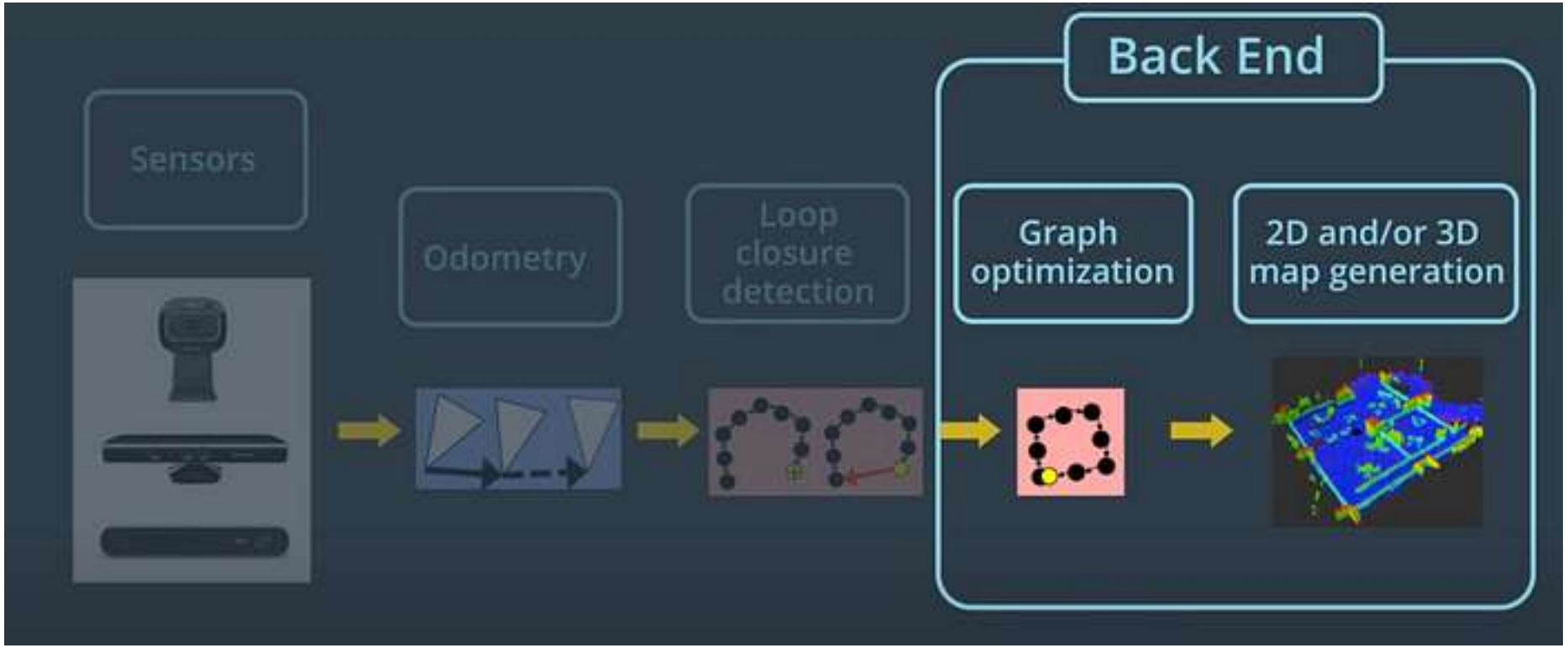
- SLAM (Simultaneous Localization and Mapping) is an algorithm that enables a robot or autonomous vehicle to simultaneously create a map of its environment and determine its own location within that environment in real time.
- Localization pinpoints a robot's pose in a known map using sensors like cameras or LiDAR.
- Mapping involves creating or updating a 2D/3D model of the environment as the robot moves, using different sensors.
- RTAB-Mapping, short for Real-Time Appearance-Based Mapping, is a graph-based SLAM approach.



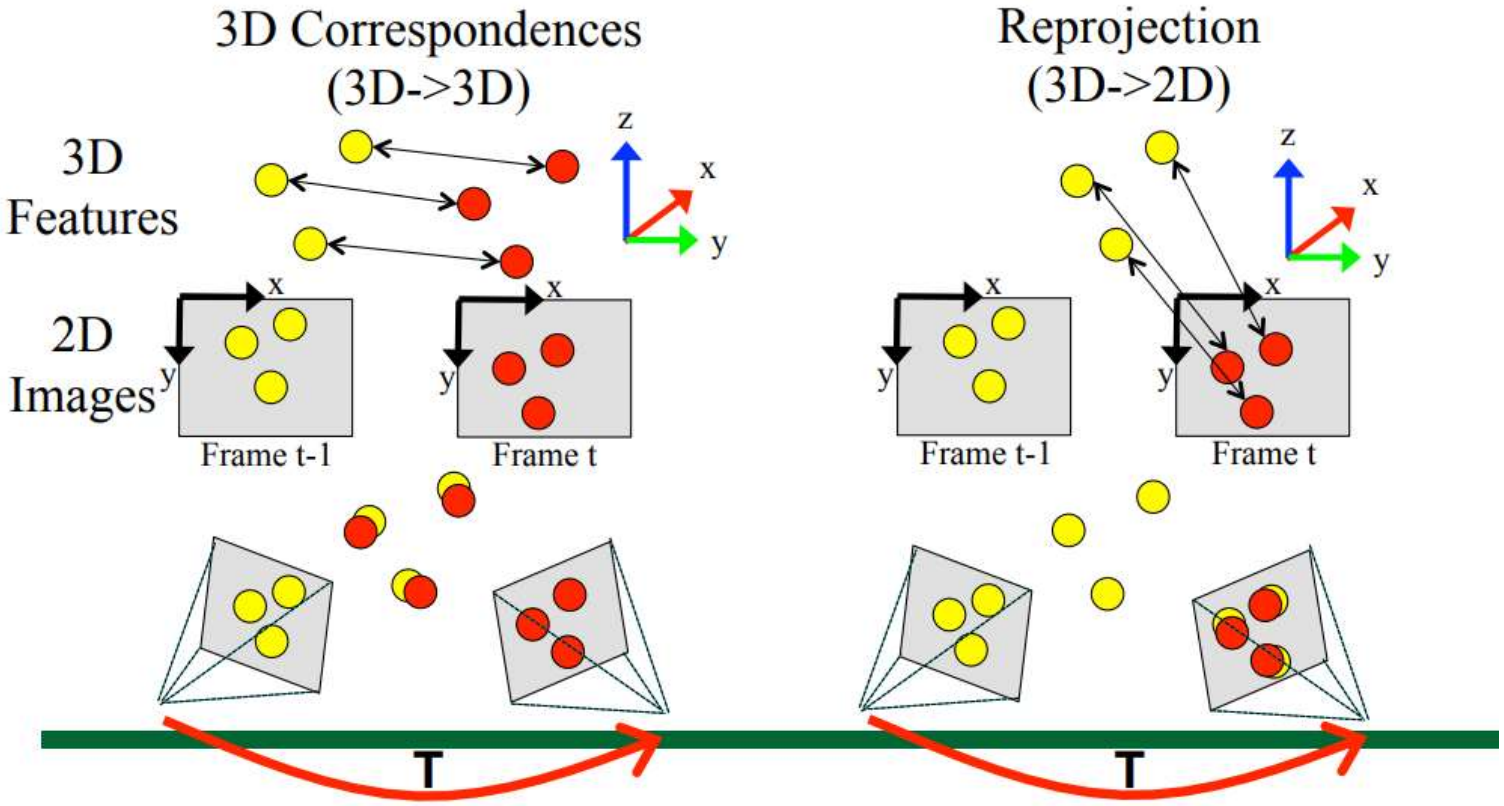
RTAB-MAP WORKING: FRONT END



RTAB-MAP WORKING: BACK END



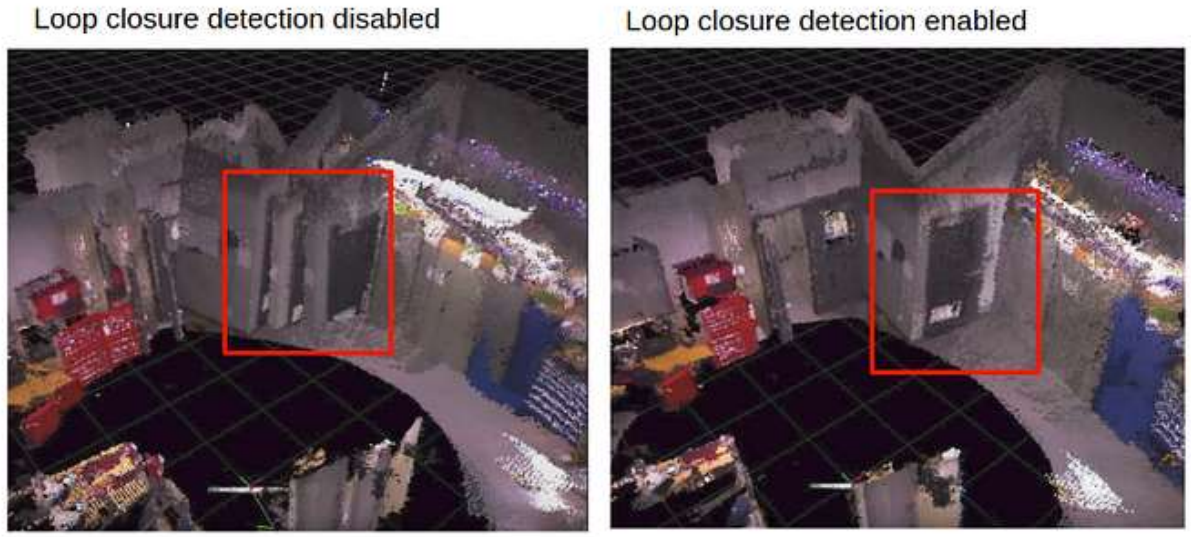
VISUAL ODOMETRY



Transformation Estimation T (x, y, z, roll, pitch, yaw)

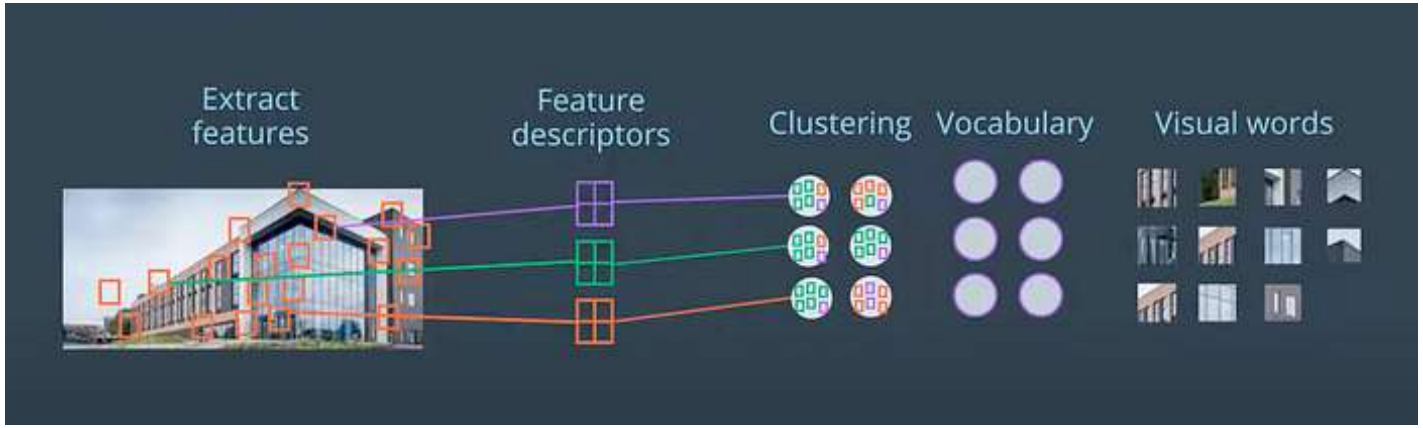
LOOP CLOSURE

- Loop closure is finding a match between the current and previously visited locations in SLAM.
- There are two types of loop closure detections: local and global.
- RTAB-Map uses global loop closures along with other techniques to ensure that the loop closure process happens in real-time.
- The loop closure is happening fast enough that the result can be obtained before the next camera images are acquired.



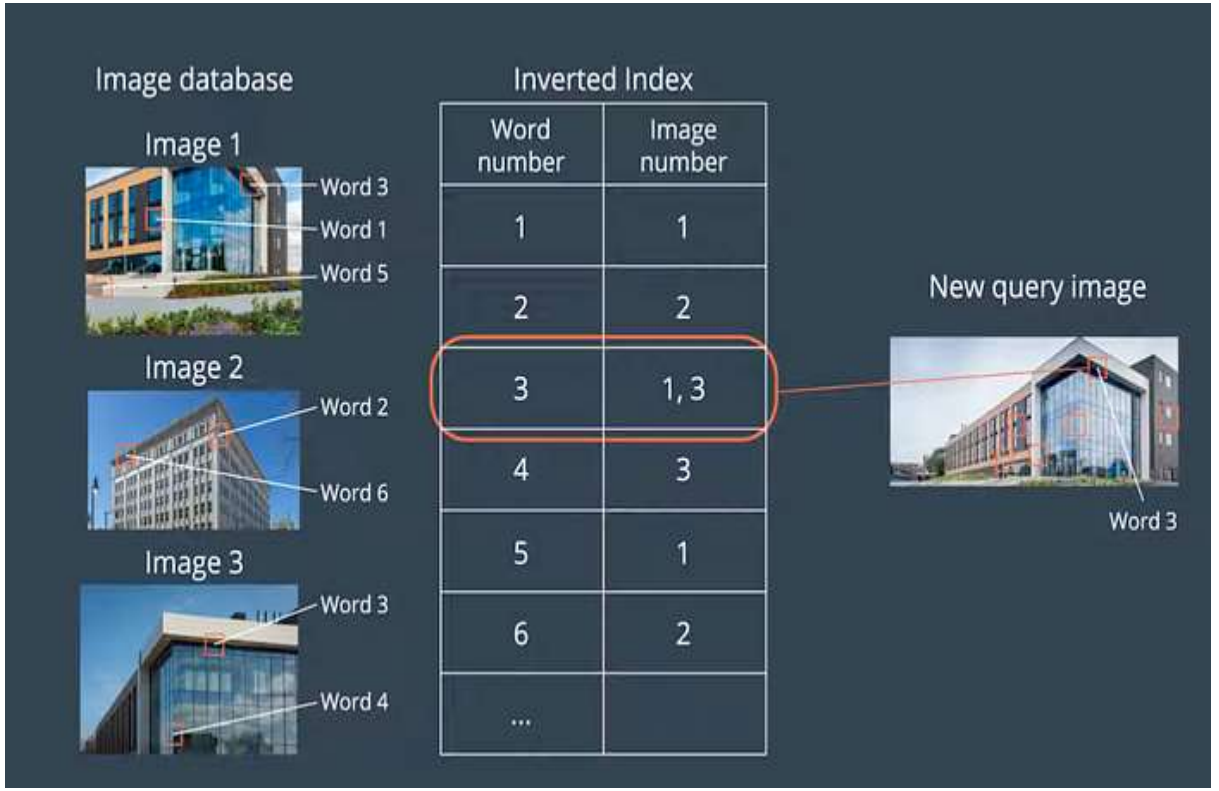
BAG OF WORDS

- In RTAB-Mapping, the default method used to extract features from an image is called Speeded Up Robust Features or SURF.
- A feature descriptor is a unique and robust representation of the pixels that make up a feature.
- Vocabulary is used for faster comparison of feature descriptors.
- When a feature descriptor is mapped to one in the vocabulary, it is called quantization.
- When all features in an image are quantized, the image is now a bag of words.



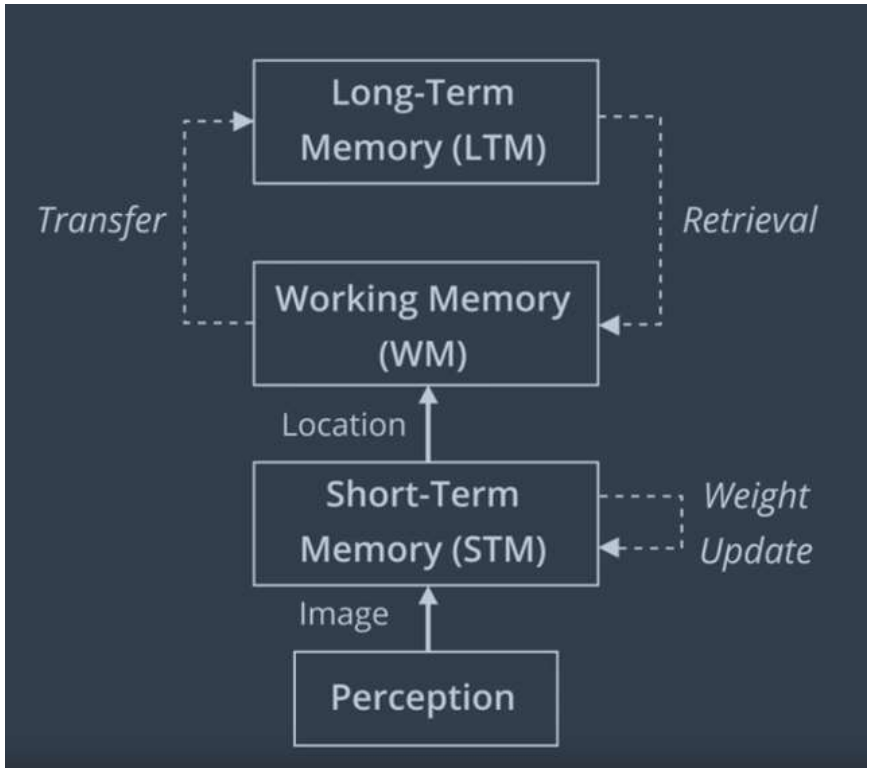
BAG OF WORDS

- Images are compared by assigning matching scores based on shared visual words.
- An inverted index keeps track of words and associated images.
- Increased scores occur when words are seen in images, and a Bayesian filter assesses scores.
- A loop closure is detected when the hypothesis surpasses a predefined threshold.



MEMORY MANAGEMENT

- RTAB-Map employs a memory management approach for efficient loop closure detection in real time. It can be used for long-term and large-scale environment mapping.
- The strategy involves retaining recent and frequently visited locations in the robot's Working Memory while transferring others to Long-Term Memory.
- When a new image is acquired, a new node is created in the **Short-Term Memory (STM)**.

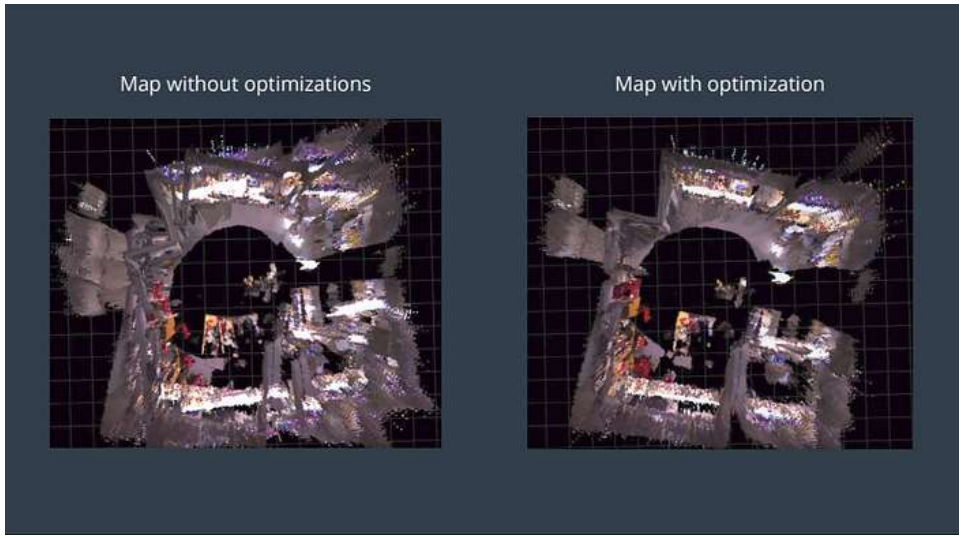


MEMORY MANAGEMENT

- When a node is formed, features are extracted and matched to the vocabulary, forming a bag of words for that node. Node weights in the STM are determined by the time spent in a location, with longer durations receiving higher weights. If consecutive images are similar, the first node's weight increases by one, and no new node is generated for the second image.
- The STM has a fixed size of S. When the STM reaches S nodes, the oldest node is moved to WM to be considered for loop closure detection.
- Loop closure happens in the WM. The Working Memory (WM) in RTAB-Map has a fixed size determined by a time limit, T. When processing time reaches T, older and less-weighted nodes are moved from WM to Long-Term Memory (LTM), maintaining a nearly constant WM size.
- LTM, distinct from WM, is not involved in loop closure detection or graph optimization.
- In the event of loop closure, neighbours from the LTM of an old node can be retrieved to WM, a process known as retrieval.

GRAPH OPTIMIZATION

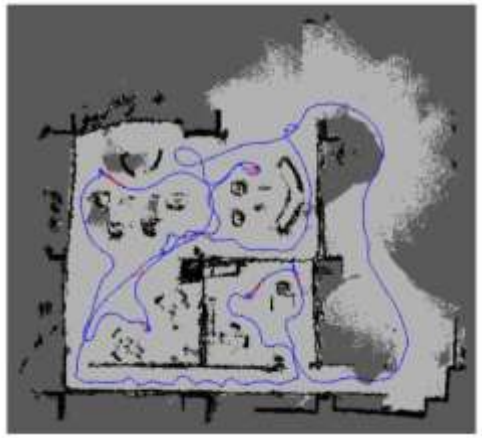
- When a loop closure hypothesis is accepted, a new constraint is added to the map's graph, then a graph optimizer minimizes the errors in the map.
- RTAB-Map supports 3 different graph optimizations: Tree-based network optimizer or TORO, General Graph Optimization or G2O and GTSAM (Smoothing and Mapping).
- All these optimizations use node positions and link transformations as constraints. When a loop closure is detected, errors introduced by the odometry can be transferred to all links, correcting the map.



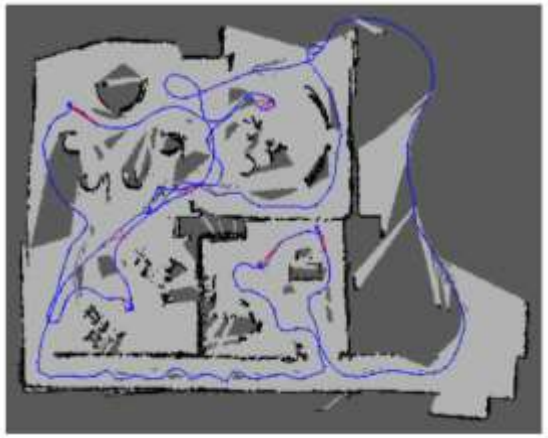
GTSAM

- GTSAM (Georgia Tech Smoothing and Mapping):
- C++ library for robotics and computer vision.
- Solves SLAM and sensor fusion problems.
- Uses factor graphs for flexible representation.
- Supports Bayesian filtering (Extended Kalman Filter - EKF, incremental SAM).
- Efficient sparse linear algebra for large-scale problems.
- Open-source for academic and industrial use.

MAP GENERATION

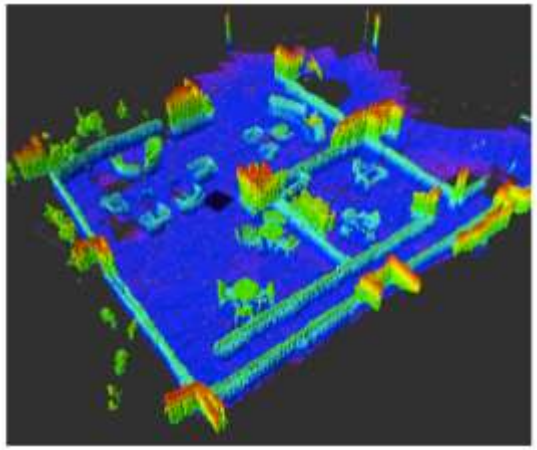


2D (PROJECTION)



2D LASER SCANS

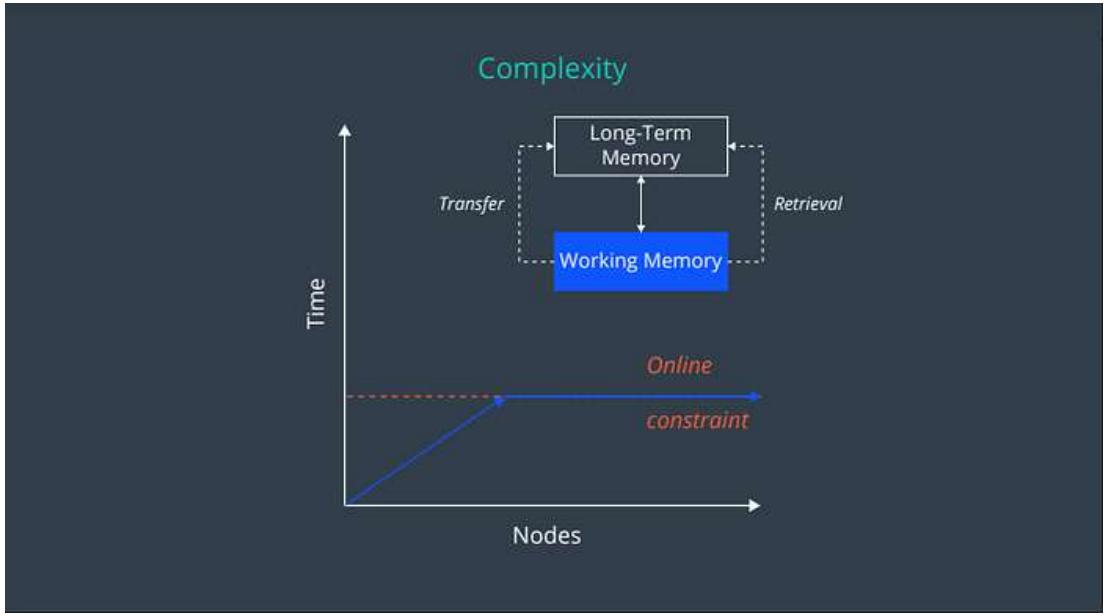
3D (POINT CLOUD)



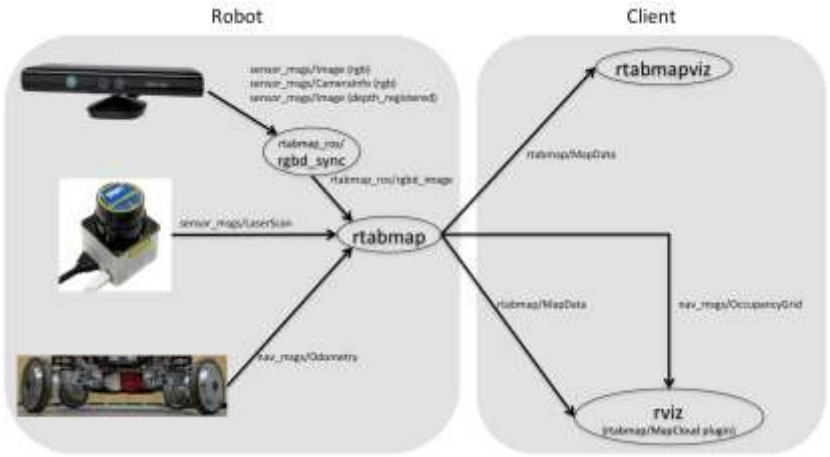
3D (OCTOMAP)

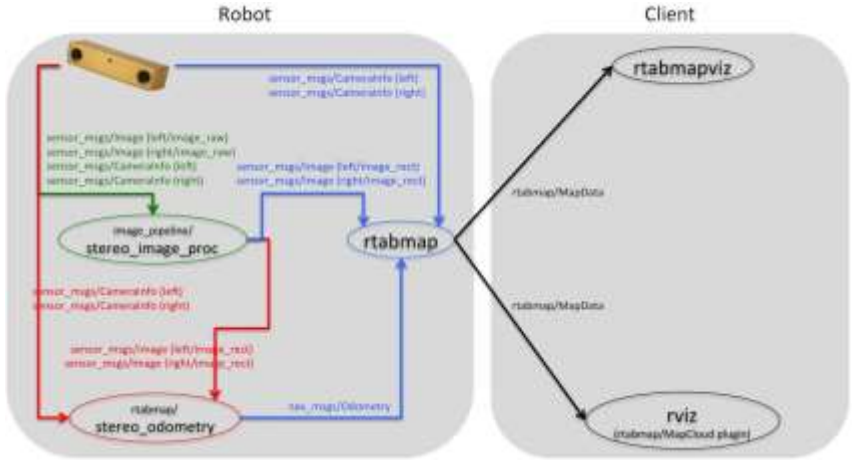
COMPLEXITY

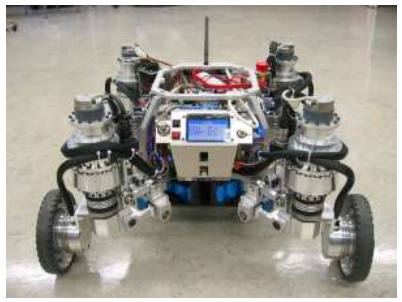
- Graph-SLAM complexity is linear, according to the number of nodes, which increases according to the size of the map.
- By providing constraints associated with how many nodes are processed for loop closure by memory management, the time complexity becomes constant in RTAB-Map.

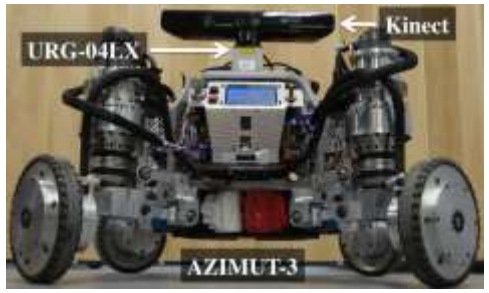


VERSATILITY



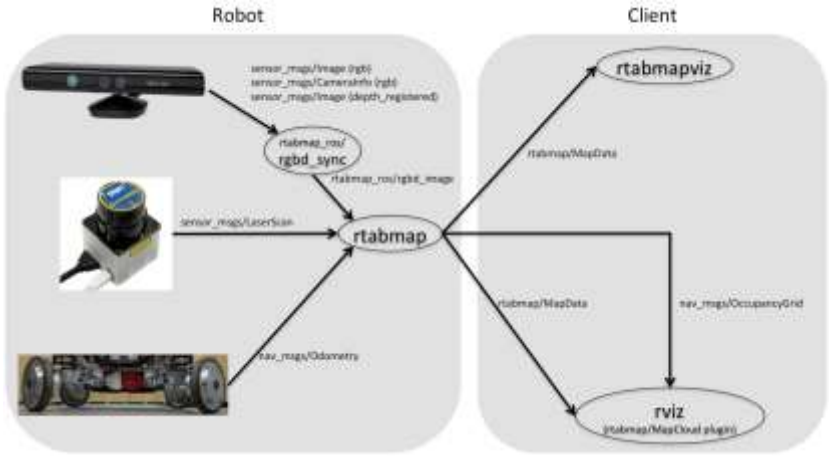




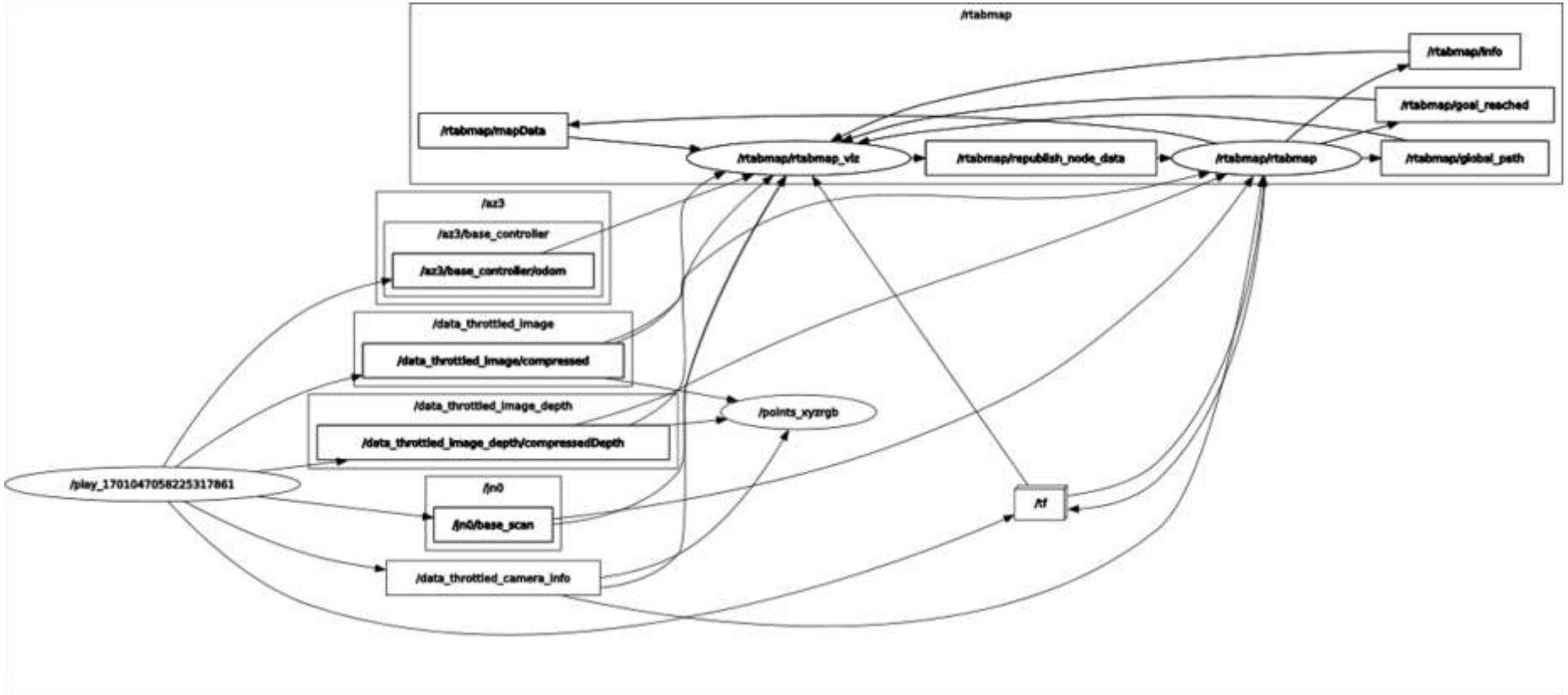


DEMO: DEVELOPER DATASET

- Location: Corridors within a building with rooms at the start and end of the dataset video.
- Sensor Input Configuration: RGB-D Camera + LiDAR + IMU + Wheel Odometry
- Lighting Conditions: Brightly-lit corridors and rooms.
- Human Crowd Conditions: Only 1 person encountered in the middle of the video who was standing.

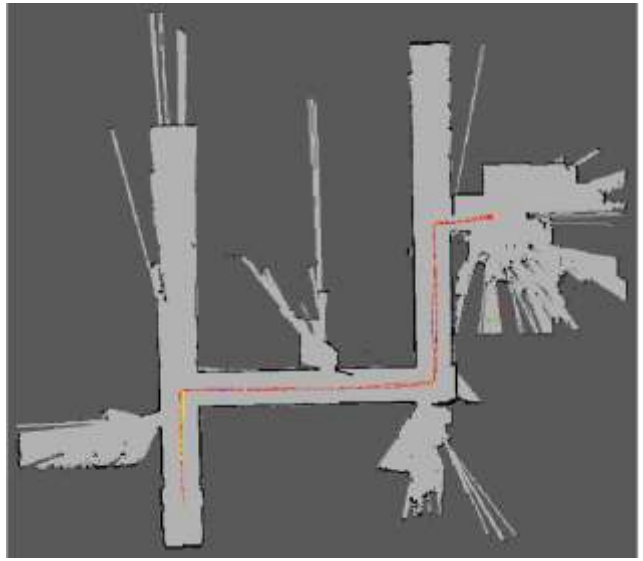


DEMO: DEVELOPER DATASET

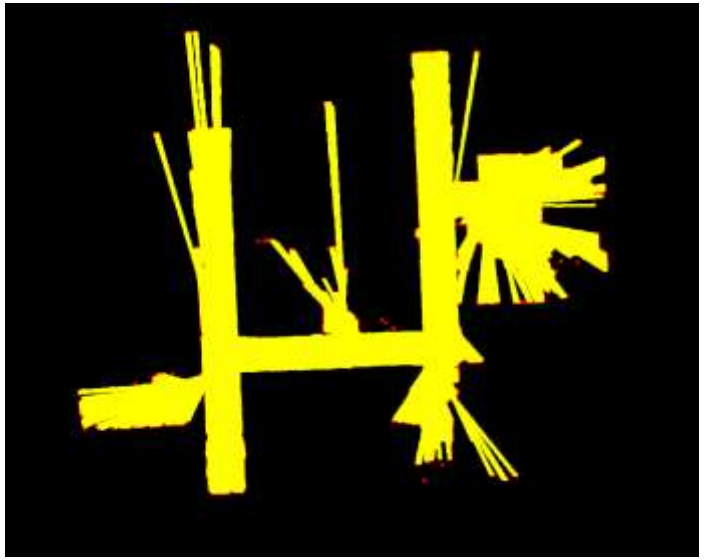


DEMO: DEVELOPER DATASET

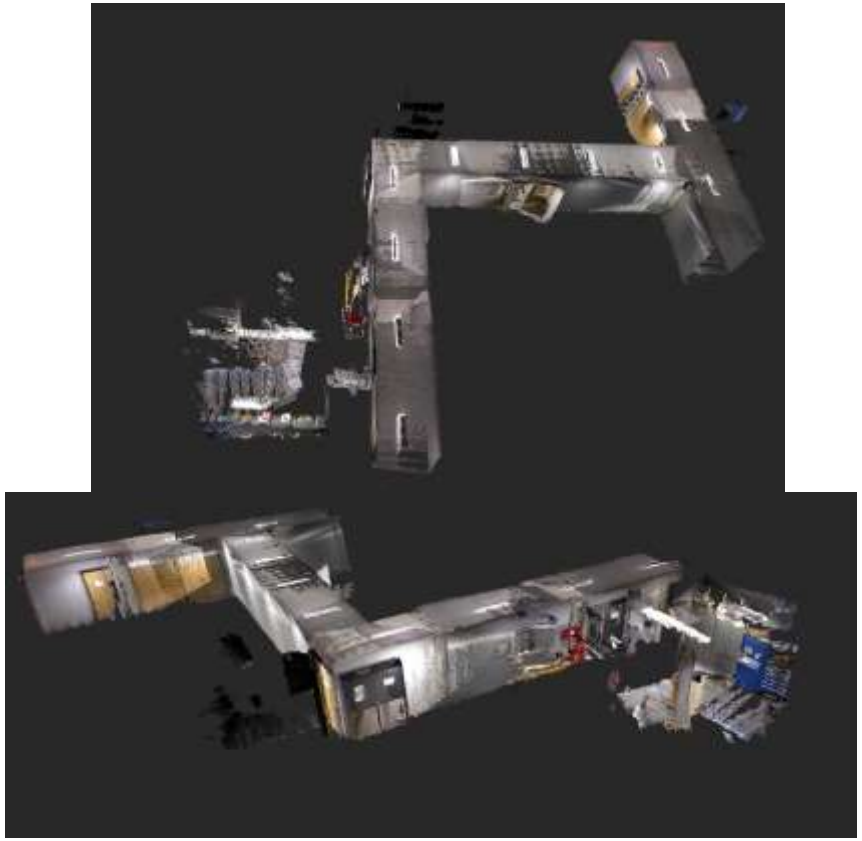
RESULTS: DEVELOPER DATASET



Trajectory in 2D Projection Map



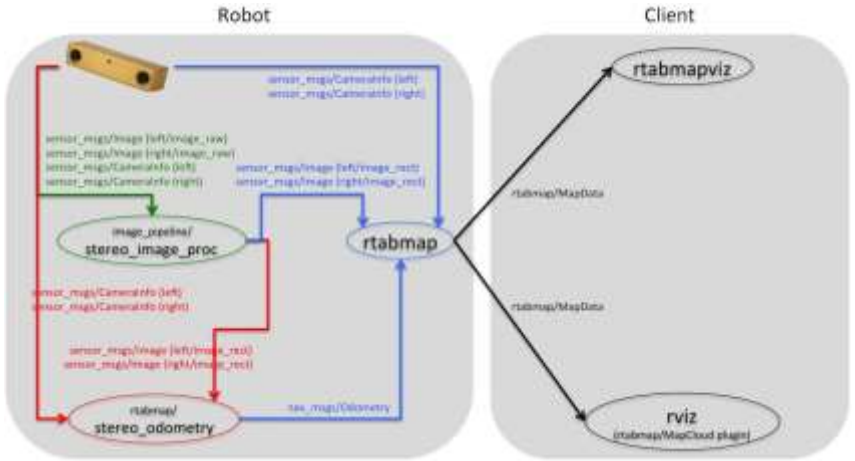
Occupancy Grid Map



3D Map in two views

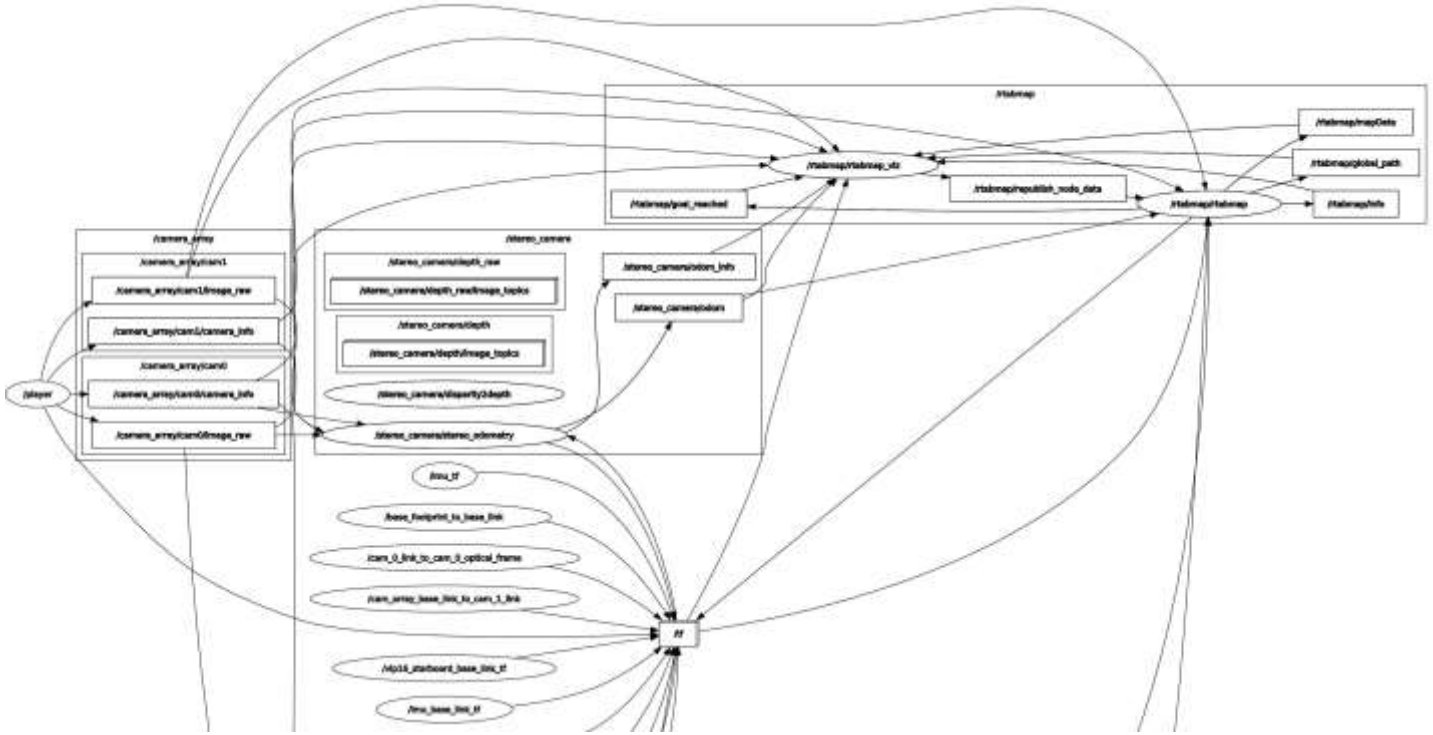
DEMO: NEU OUTDOOR DATASET

- Location: Data collected using NUANCE Autonomous Car while driving on Newbury Street, Boston.
- Sensor Input Configuration: Stereo Camera
- Lighting Conditions: Cloudy weather conditions with insufficient lighting.
- Human Crowd Conditions: 2 people and a cyclist crossing the street.

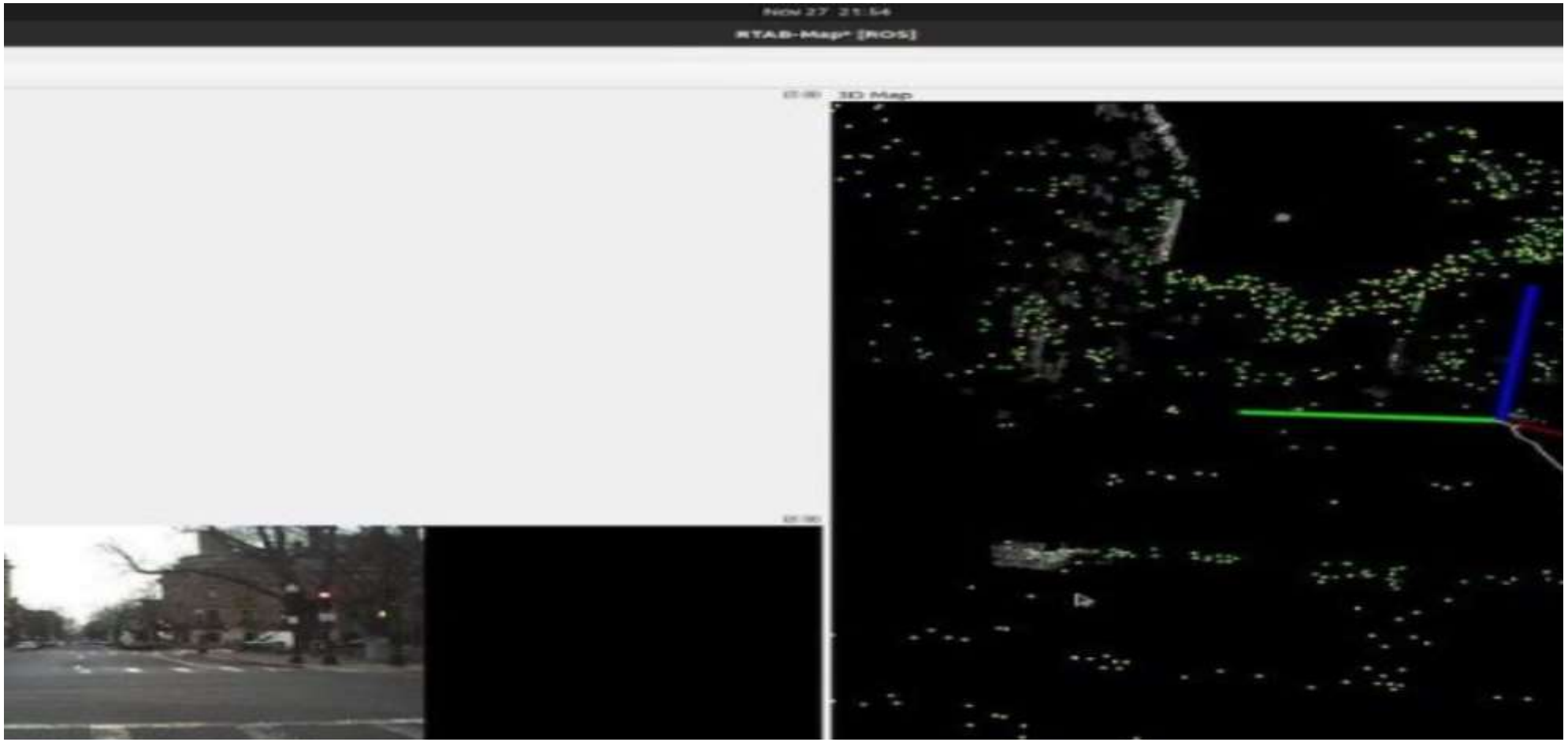


DEMO: NEU OUTDOOR DATASET

CAR_IR_RGB_Lidar

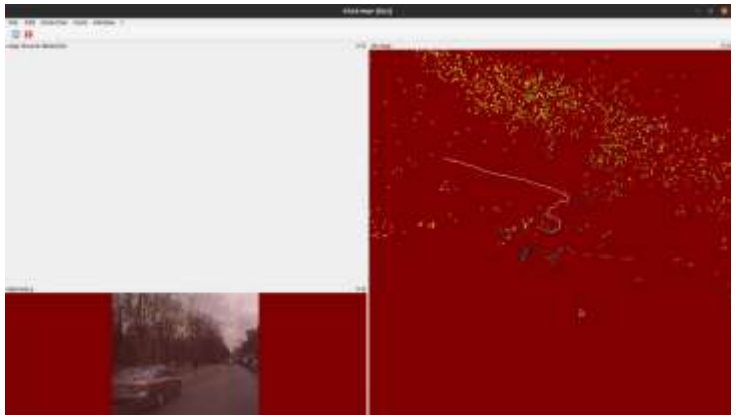


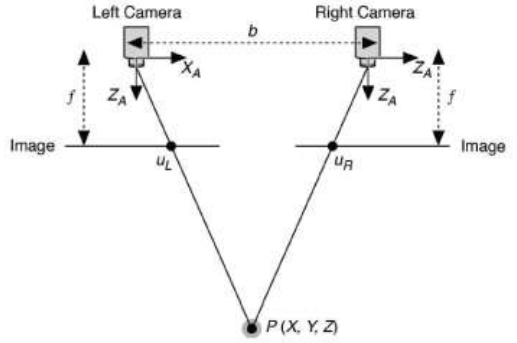
NEU OUTDOOR DATASET



ANALYSIS: NEU OUTDOOR DATASET

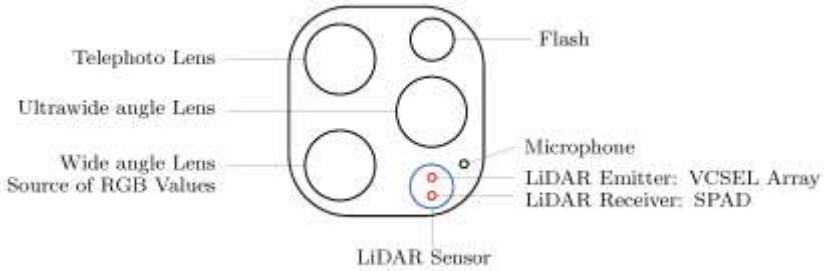
1. When the background turns RED, it means that odometry cannot be computed:
   - **Featureless environments:** White walls, uniform texture, dark areas
   - **"Empty space" environments:** Pointing the sensor over a clear area where there are no objects within sensor depth range.
   - **Fast camera motion for the framerate**: If the camera is moved too fast, there will be less features to match between successive frames.
2. Inability to develop a 3D Map
3. Lack of Loop Closure detection
4. Stereo Camera baseline constraint on the Depth Range



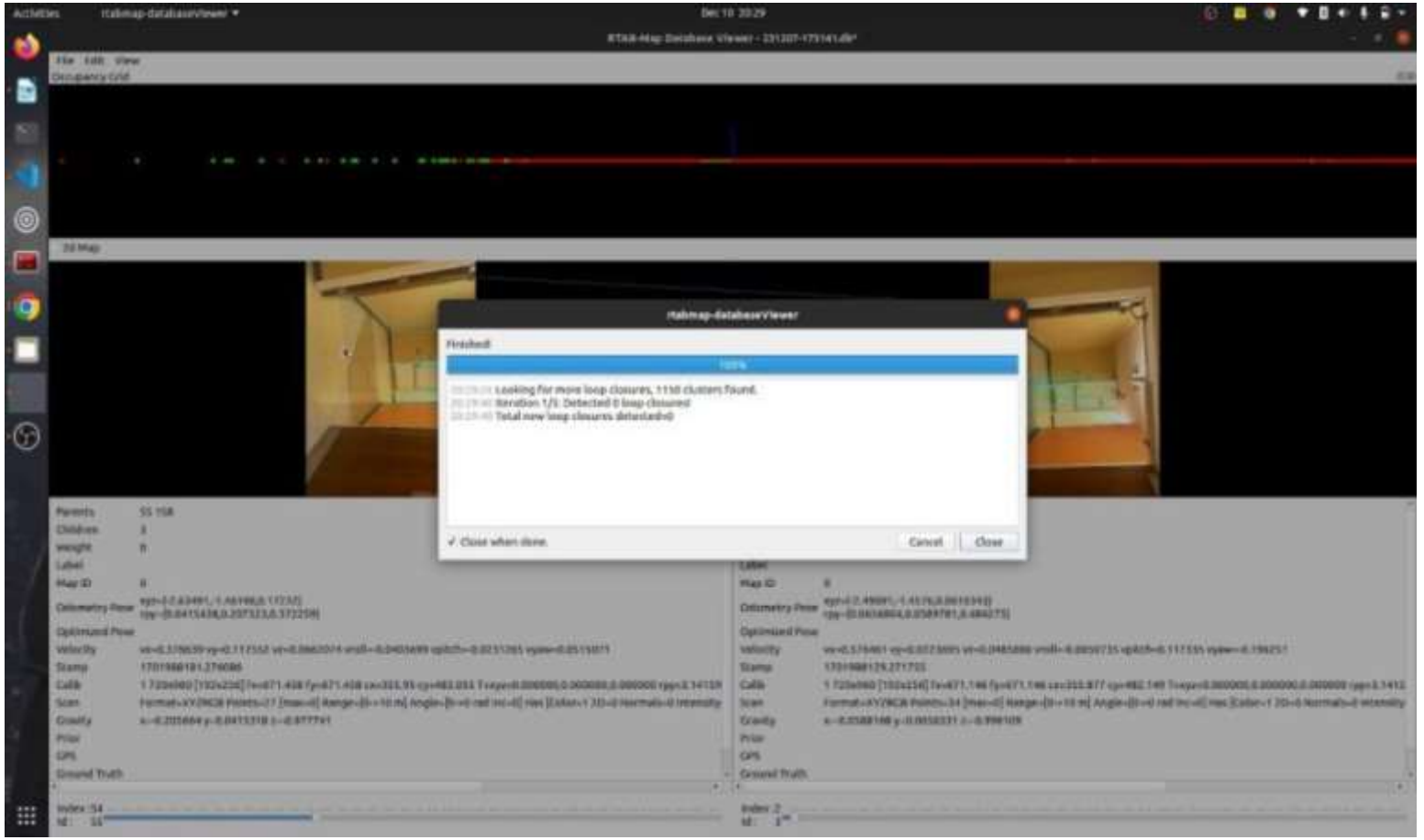


NEU INDOOR DATASET

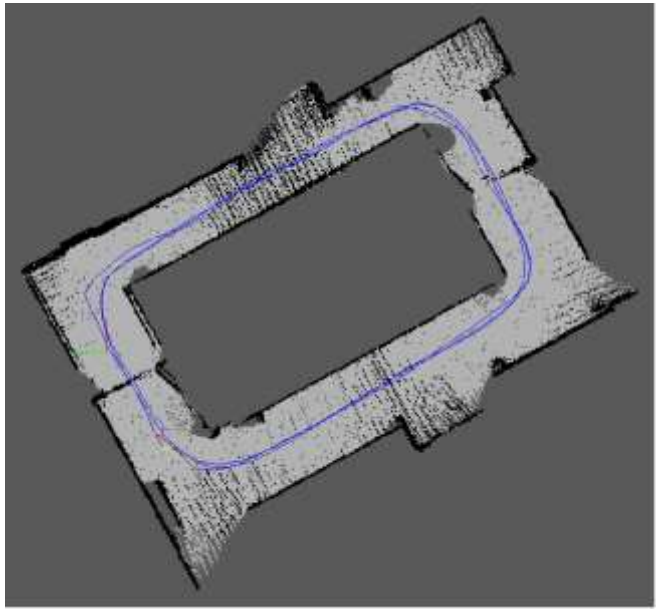
- Location: Snell Library Basement Corridor at NEU.
- Sensor Input Configuration: iPhone in-built camera + LiDAR
- Lighting Conditions: Well-lit corridor with bright overhead lights.
- Human Crowd Conditions: Only one person was sitting down near a classroom.



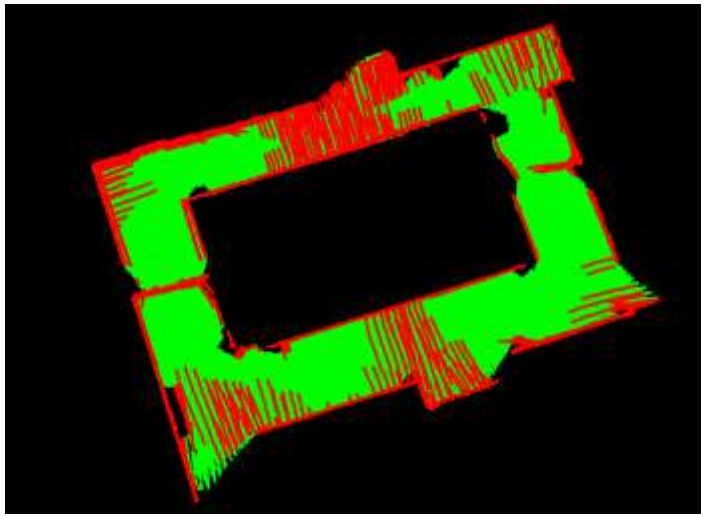
RESULTS: NEU INDOOR DATASET



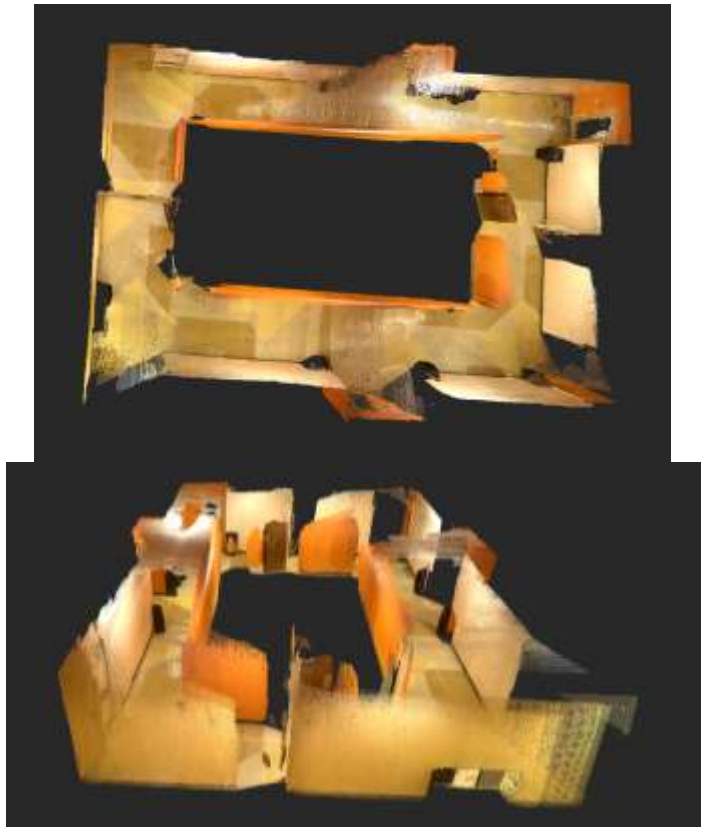
RESULTS: NEU INDOOR DATASET



Trajectory in 2D Projection Map



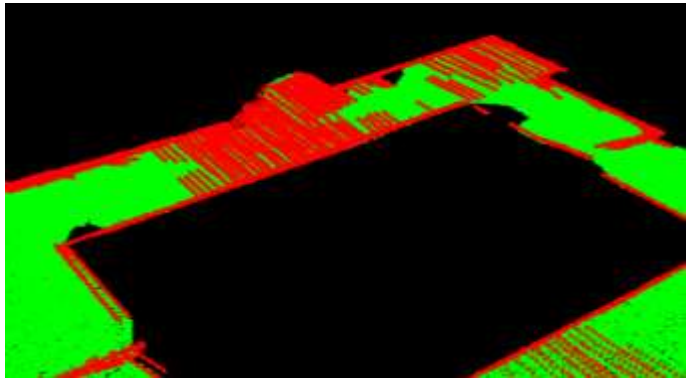
Occupancy Grid Map

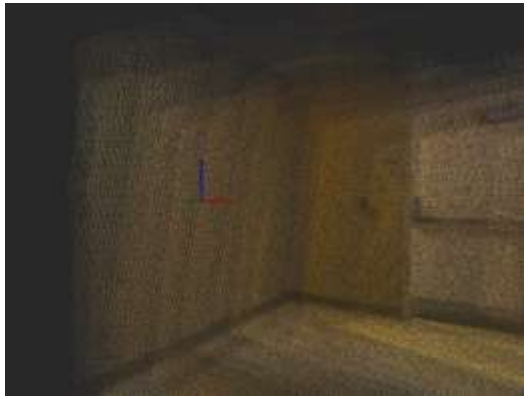


3D Map in two views

ANALYSIS: NEU INDOOR DATASET

1. A 3D Map was developed successfully for all elements within the camera's field of view.
2. Sufficient Loop Closures are detected for 3D Map formation
3. Occupancy Grid Map Performance:
    - The red grids in the image show occupied cells in the pathway despite being empty.
4. Local Loop Closures on Texture-less Elements:
    - The local loop closures are completely dependent on the non-visual odometry as there are insufficient features for Visual Odometry calculations between subsequent frames.





CONCLUSION

- The actions taken in the developer dataset are ideal for creating a perfect map for RTAB-Map SLAM. The developer makes effort to pan the camera at multiple locations to extract more features for additional loop closures.
- The RTAB SLAM works far better on NEU Indoor Dataset as compared to NEU Outdoor Dataset. This happens because of many different factors:

  i) Distinct and discernible features are spread across the entire indoor map like doors, walls, corners, dustbins, etc. which are used in feature matching and tracking.

  ii) Lighting conditions are much better and can be controlled in the indoor location which aids in feature detection.

  iii) There are no moving people or objects in the indoor dataset. RTAB Map works best in a static environment.

REFERENCES

- RTAB-Map as an open-source lidar and visual simultaneous localization and mapping library for large-scale and long-term online operation – Mathieu Labbe & François Michaud
- WikiROS – rtabmap_ros
- Github – introlab – rtabmap_ros
- Simultaneous Localization and Mapping (SLAM) with RTAB-Map - Mathieu Labbé - Université de Sherbrooke
- Introduction to 3D SLAM with RTAB-Map - Medium - Shiva Chandrachary
- Github – introlab – rtabmap_ros

 Repository  Report