

RTAB-Map Onboard COBRA

Team 12

Henry Noyes, Nishad Rajhans, Mayank
Polagoni, Sai Kiran Athaluri



Northeastern
University



RTAB Mapping

Real-Time Appearance-Based Mapping, or RTAB, is an effective technique for simultaneous localization and mapping (SLAM) in computer vision and robotics. It allows robots to localize themselves inside an environment and create a map of it.

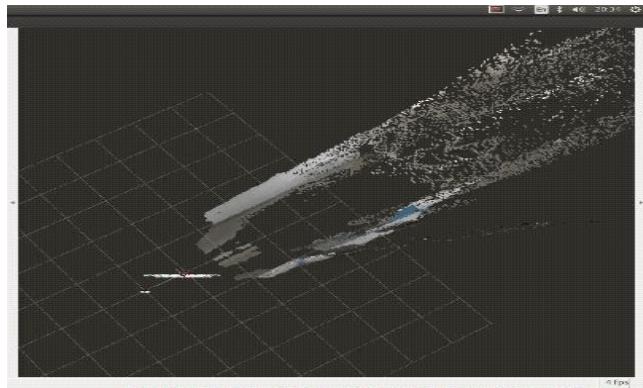
The input data for RTAB MAP:

Cameras:

- **RGB Cameras:** Standard color cameras for 2D feature extraction.
- **RGB-D Cameras:** Cameras that provide both color and depth information (e.g., Intel RealSense, Microsoft Kinect).
- **Stereo Cameras:** Two cameras providing stereo images for depth estimation (e.g., ZED cameras).

LiDAR: Point cloud data from LiDAR sensors can be used for precise 3D mapping and localization.

IMU: Inertial Measurement Unit (IMU) provides accelerometer and gyroscope data for motion estimation and odometry.





Default feature description:

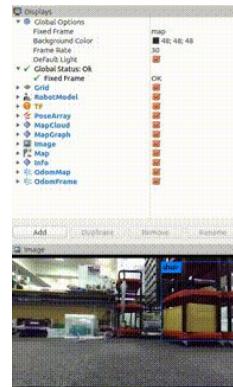
The Default feature description algorithm used for the RTAB slam algorithm is SIFT, (Scale-Invariant Feature Transform) as the default feature detector, which is generally a good balance between accuracy and performance.

Bag-of-Words (BoW) for Place Recognition

- Not a feature detector, but a mechanism for comparing visual features.
- Converts extracted features (e.g., ORB, SIFT, SURF) into visual words for efficient matching and loop closure detection.

Structure of Factor Graphs:

1. **Nodes:** Represent variables to be estimated (e.g., robot poses, landmark locations).
2. **Factors:** Represent constraints or relationships between variables.
 - Examples: Odometry measurements, loop closures, or sensor observations.
3. **Edges:** Connect nodes to factors, forming a bipartite graph.



Graph Optimization:

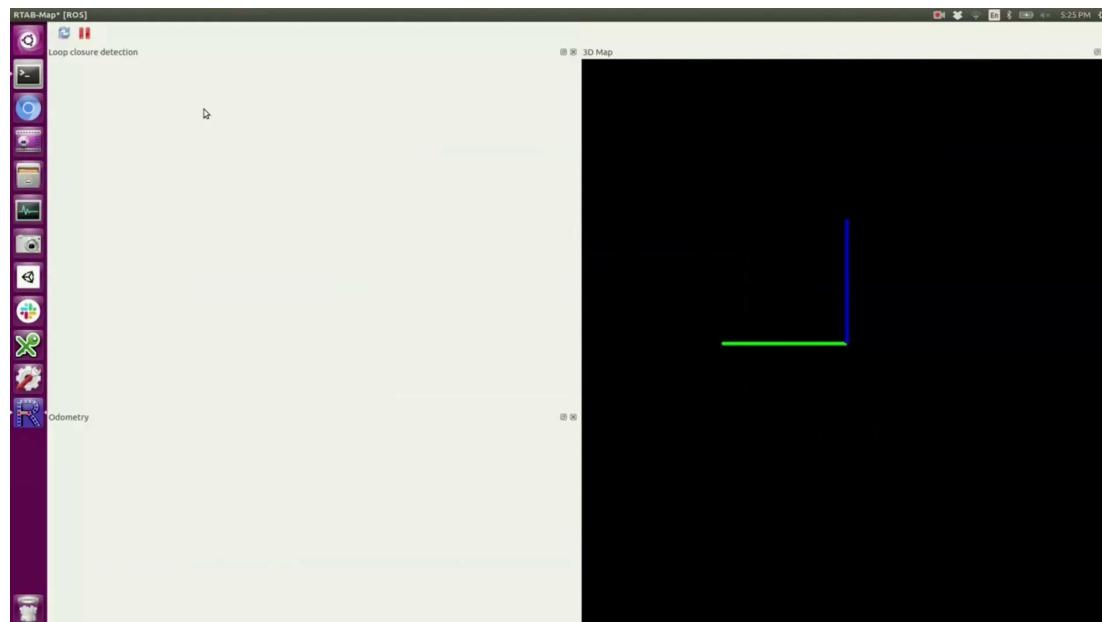
- RTAB-Map uses solvers like **g2o** or **Ceres Solver** to minimize the errors in the graph.
- This step aligns the map and corrects drift by adjusting poses based on odometry and loop closures.

Description for the RTAB Mapping Dataset

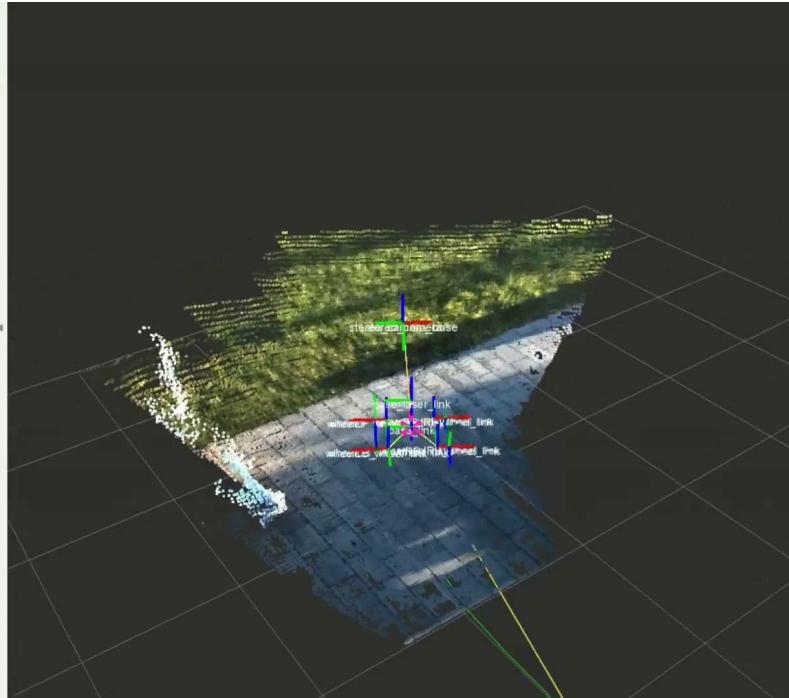
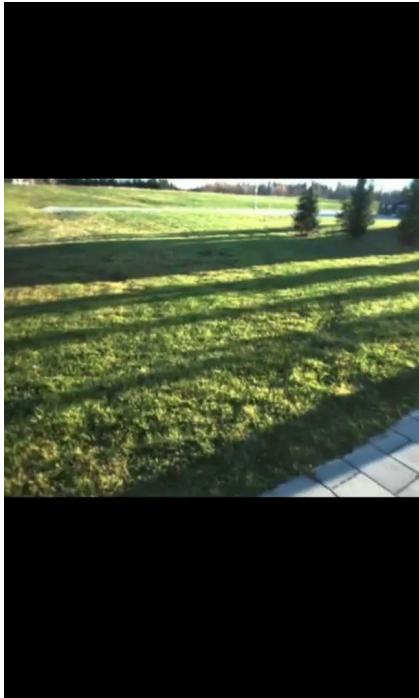
For this RTAB-Map sequence on freiburg1_xyz from the TUM dataset, the camera was pointed at a typical desk in an office environment. This sequence contains only translatory motions along the principal axes of the camera, while the orientation was kept (mostly) fixed.

ROS bag file (file size: approx. 0.42GB)

- Ground-truth trajectory length: 7.112m
- Avg. translational velocity: 0.244m/s
- Avg. angular velocity: 8.920deg/s



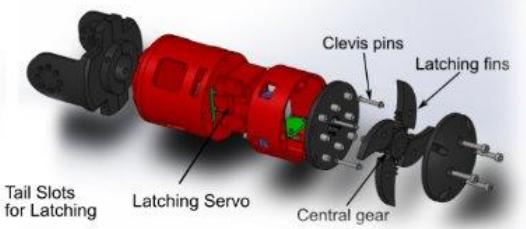
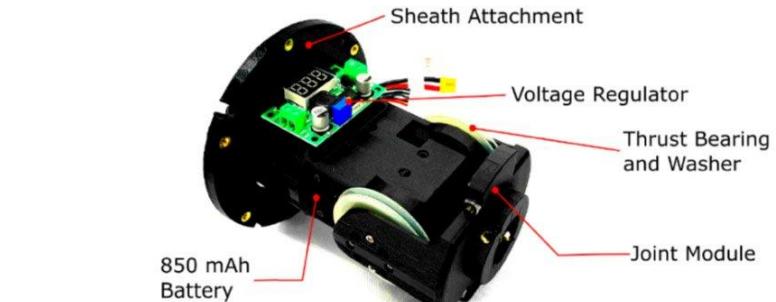
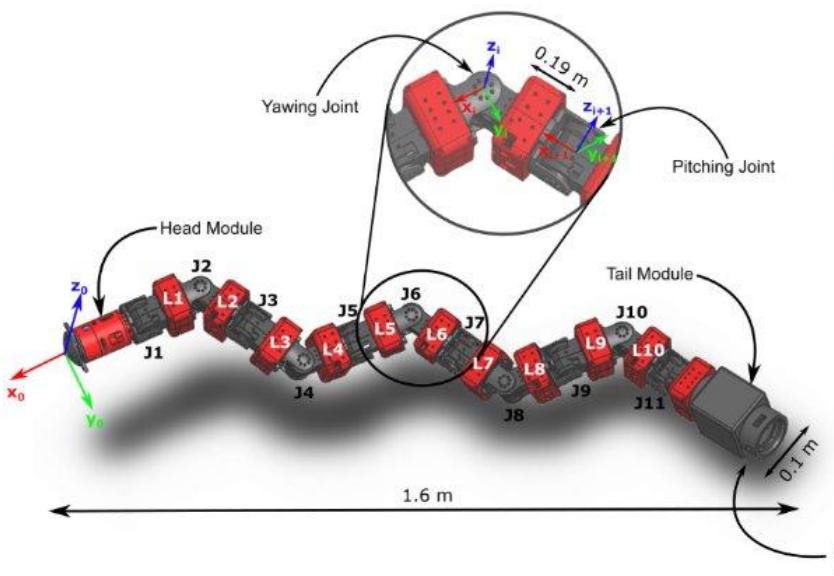
This video highlights RTAB-Map's ability to map outdoor areas using stereo camera data in a robust and efficient way.



COBRA Hardware



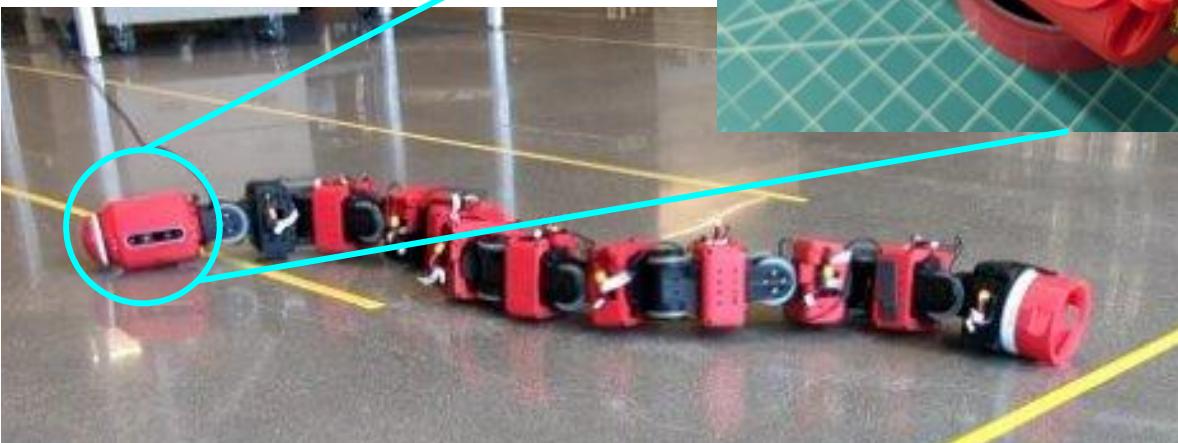
- 11 Joint Modules
 - 12V Dynamixel Servo
 - 850mAh LiPo
- Printed in PLA and Onyx (carbon fiber composite)
- 1.6 meters; 7kg



COBRA Hardware



- Main board: Jetson Orin NX 8GB
- Camera: Intel RealSense D435i
 - RGB 640x480 30Hz
 - Depth 640x480 15Hz
 - IMU 200Hz (fused)





- Collected data inside EXP Drone Cage
- Tested with varying environments and gaits





Fast sidewinding gait: difficult for SLAM algorithms due to choppy movement and motion blur



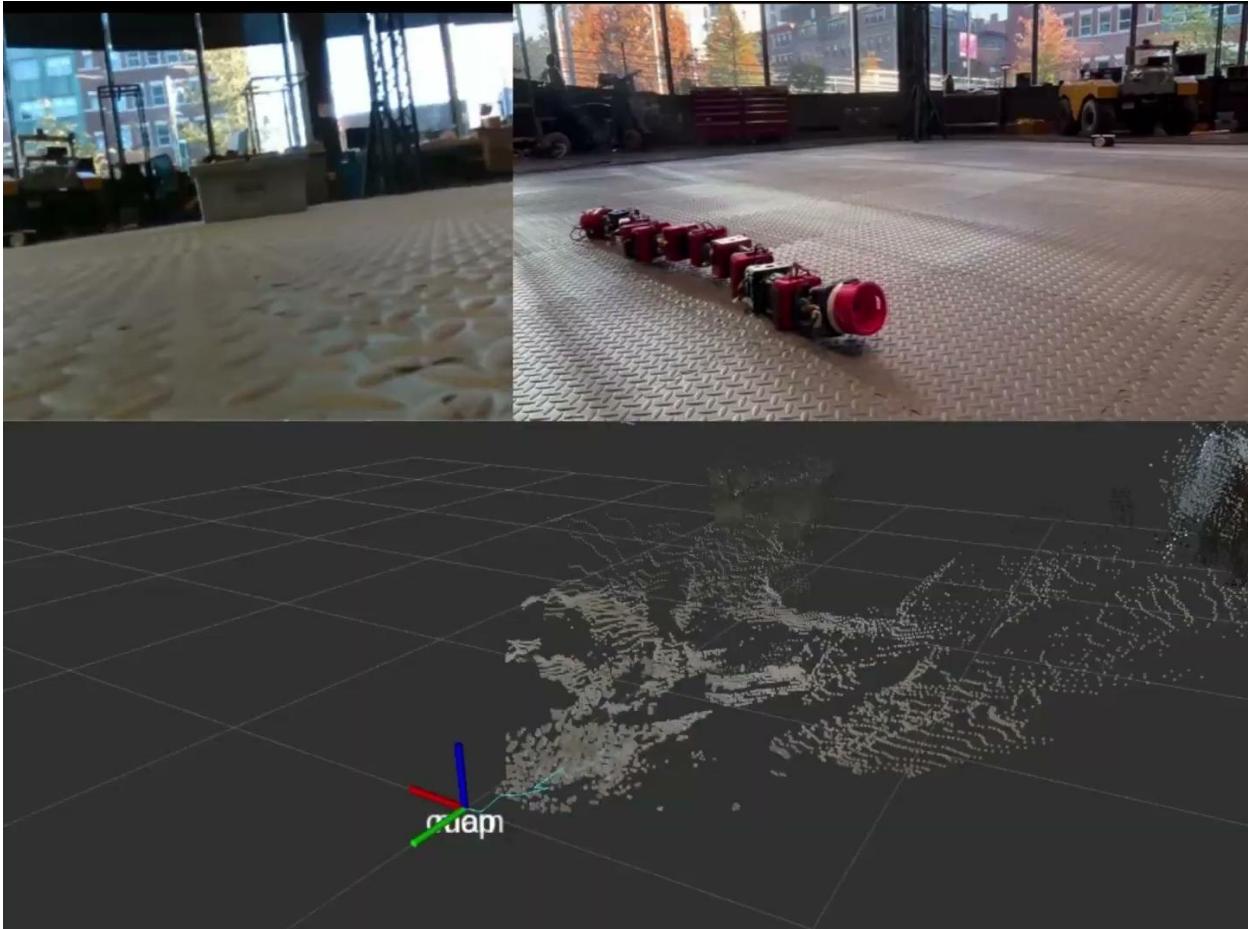
RTAB-Map Results



RTAB-Map output during slow sidewinding

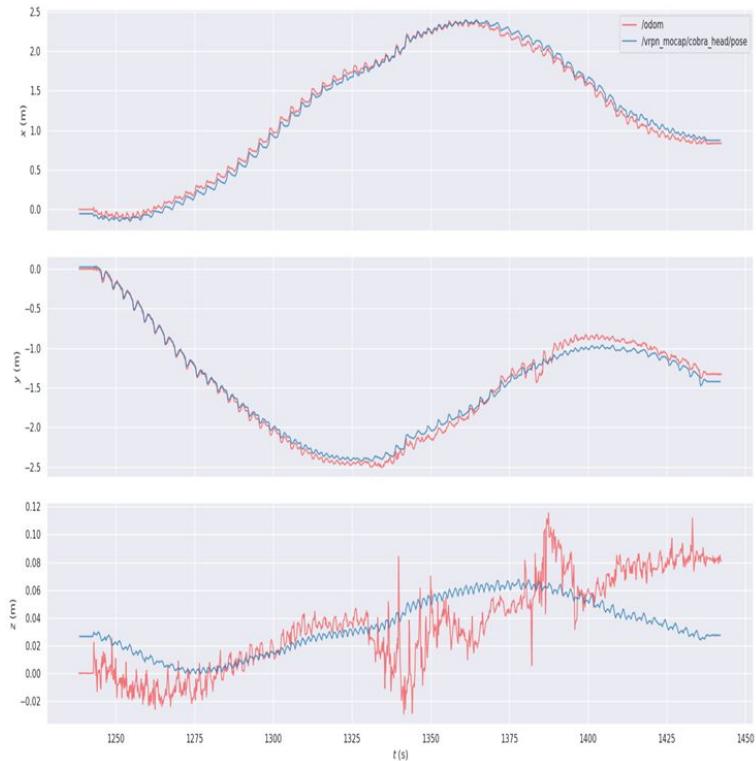
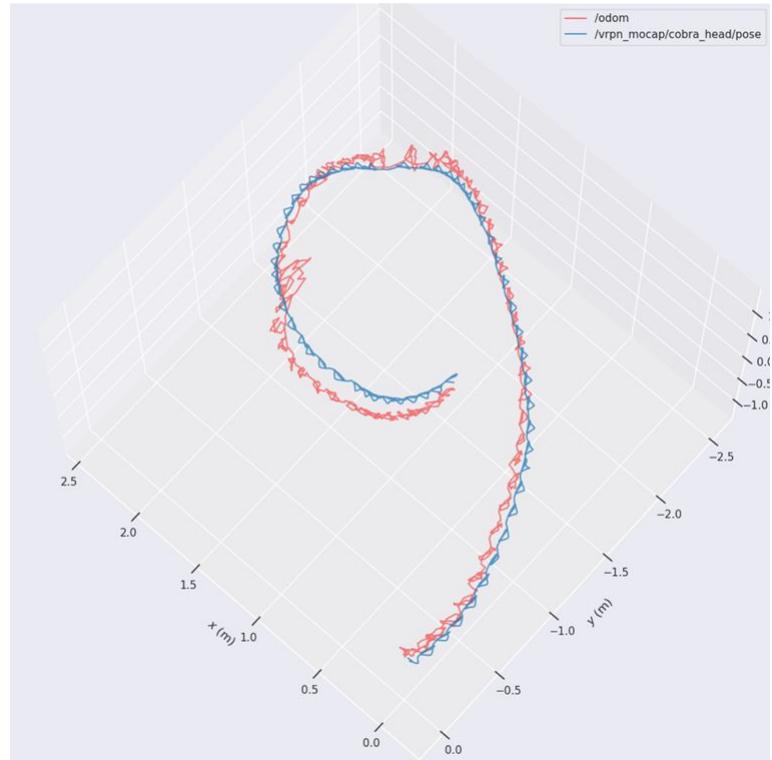
RTAB able to maintain tracking through entire 2 minute sequence

Lack of objects in environment -> sparse point cloud





Benchmarking RTAB-Map odometry vs OptiTrack ground truth (Evo)





Benchmarking RTAB-Map odometry vs OptiTrack ground truth (Evo)

```
Aligning using Umeyama's method...
Rotation of alignment:
[[ 1.0000000e+00  3.20056481e-16 -1.04083409e-17]
 [-1.80194401e-16  1.00000000e+00  0.00000000e+00]
 [ 1.04083409e-17 -3.46944695e-18  1.00000000e+00]]
Translation of alignment:
[ 1.77635684e-15  0.00000000e+00 -5.55111512e-17]
Scale correction: 1.0
-----
Compared 1093 absolute pose pairs.
Calculating APE for translation part pose relation...
-----
APE w.r.t. translation part (m)
(with SE(3) Umeyama alignment)

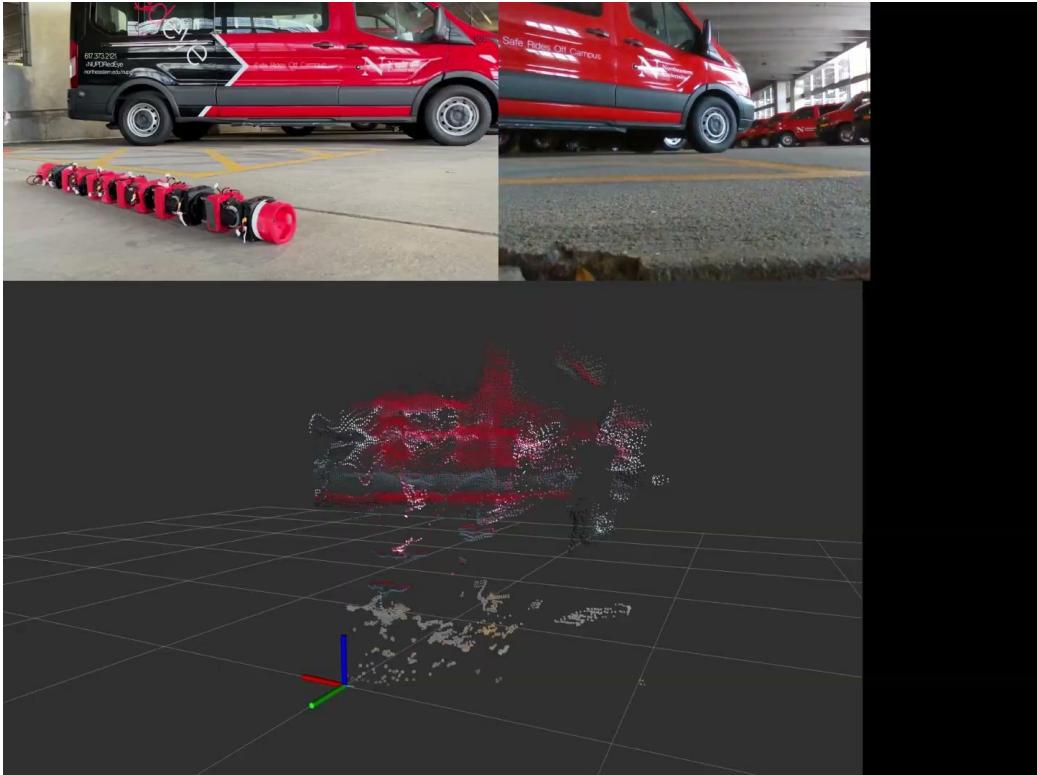
      max      0.204878
      mean     0.093487
      median   0.087104
      min      0.022170
      rmse     0.099023
      sse       10.717460
      std      0.032646
```

```
RPE w.r.t. translation part (m)
for delta = 1 (frames) using consecutive pairs
(with SE(3) Umeyama alignment)
```

max	0.219738
mean	0.027118
median	0.016765
min	0.000043
rmse	0.040701
sse	1.808947
std	0.030350



- Benchmark RTAB-Map mapping using external LiDAR
- Further evaluate strengths/limitations of RTAB with different gaits, environments, etc.
- Implement techniques to improve performance during faster sidewinding





Thank You! Questions?