Performance Enhancement of RTAB-map Utilizing ORB-SLAM2 Depth Information

Project Progress Overview

Islam Ali
iaali@ualberta.ca

Zifei Jiang
zifei.jiang@ualberta.ca

UNIVERSITY OF ALBERTA

Agenda

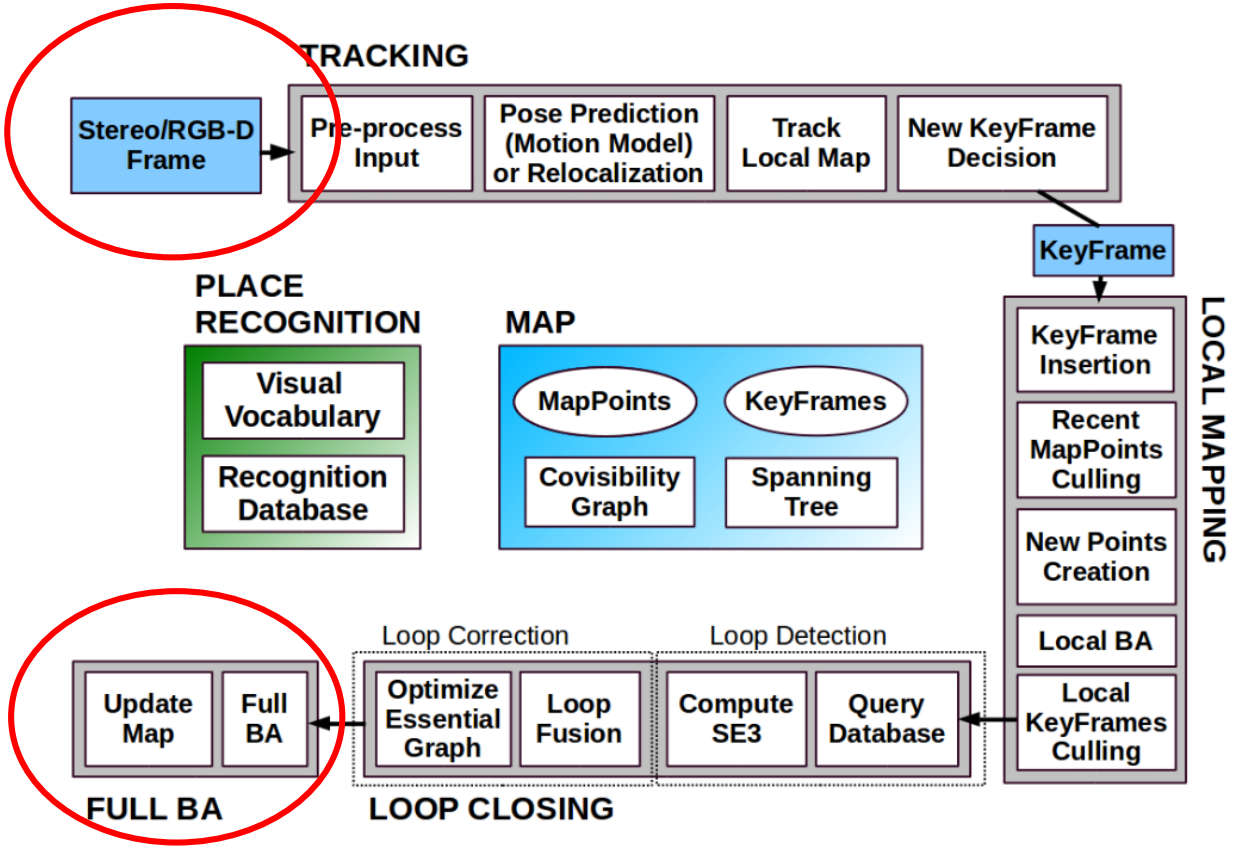
- **Introduction**
- **Background :** RTAB-Map
- **Background :** ORB-SLAM2
- **RTAB-Map Vs. ORB-SLAM2**
- **Problem Definition**
  - Revisiting
- **System Block Diagram**
  - Abstract Integration Structure
  - Detailed Block Diagram
- **Initial Baseline Results**
  - KITTI & TUM datasets Results
  - Discussion
- **Open Questions**

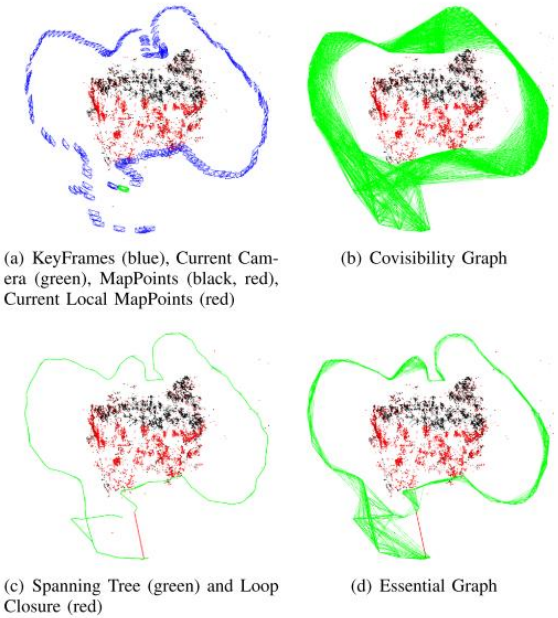
UNIVERSITY OF ALBERTA

Introduction

- In this presentation, we discuss the details of both ORB-SLAM2 and RTAB-Map as well as the integration details between the two.
- Additionally, we present some of the baseline results and discuss their indications.
- Finally, we conclude with some open questions that are yet to be answered in the course of the project.

Background : ORB-SLAM2





(a) KeyFrames (blue), Current Camera (green), MapPoints (black, red), Current Local MapPoints (red)

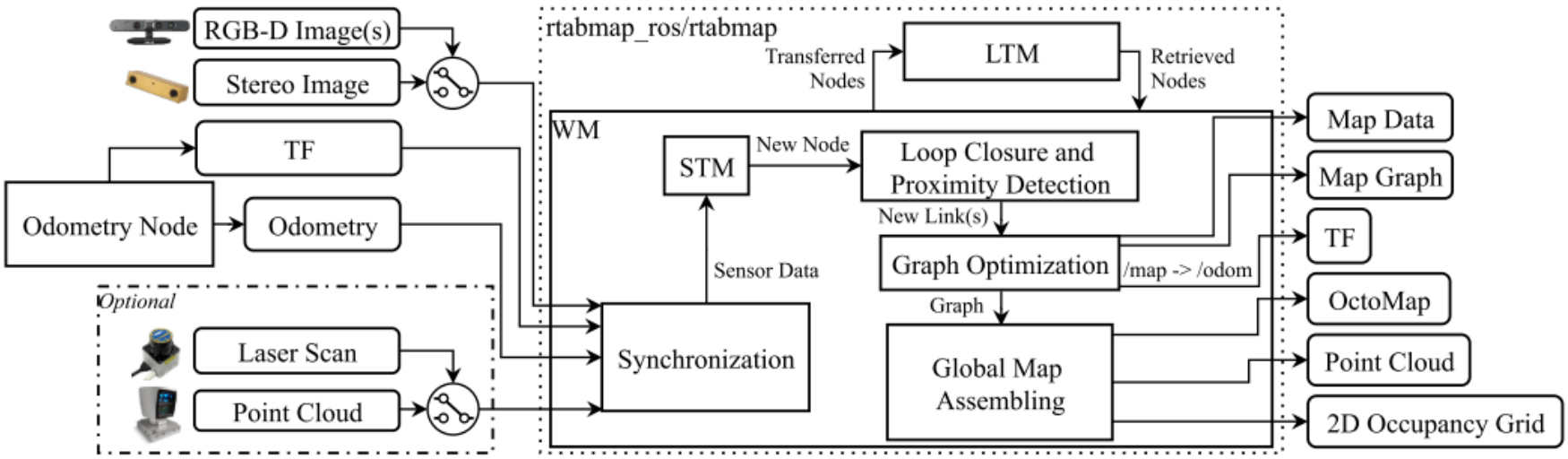
(b) Covisibility Graph

(c) Spanning Tree (green) and Loop Closure (red)

(d) Essential Graph

- Stereo Image enable easier map initialization
- ORB-SLAM using both Homography, Fundamental matrix in visual odometry.
- Local BA optimize a set of covisible frames and all map points seen in these frames
- Full BA can be really costly and may not influence the result a lot.
- When doing loop closing correction, orb-slam optimizes essential graph.

Background : RTAB-Map



- The main character of RTAB-Map is Pose Graph Slam system, which enable the separation of Odometry and Map
- In RTAB-Map paper, ORB-SLAM2 has been implemented as a Odometry part, and it out perform other method in the KITTI dataset.
- Using ICP to estimate motion is easily to realize. RGB-D Slam in 2014 used ICP[1].
- One big advantage of RTAB-Map system is the memory management. Many previous Slam system building dense map can only in small scale environments.[2]
- RTAB-Map uses ray tracing to build a 3D grid map

[1] "3D Mapping with an RGB-D Camera", *F. Endres, J. Hess, J. Sturm, D. Cremers, W. Burgard*, IEEE Transactions on Robotics, 2014.
[2] T. Whelan, R. F. Salas-Moreno, B. Glocker, A. J. Davison, and S. Leutenegger, “ElasticFusion: Real-time dense SLAM and light source estimation,” Int. J. Robot. Res., vol. 35, no. 14, pp. 1697–1716, 2016.
[3]

RTAB-Map Vs. ORB-SLAM2

- **A number of studies were conducted to compare SLAM systems including ORB-SLAM 1&2 and RTAB-Map:**

[1] Ragot, N., Khemmar, R., Pokala, A., Rossi, R., & Ertaud, J. Y. (2019, July). Benchmark of Visual SLAM Algorithms: ORB-SLAM2 vs RTAB-Map. In *2019 Eighth International Conference on Emerging Security Technologies (EST)* (pp. 1-6). IEEE.

[2] Giubilato, R., Chiodini, S., Pertile, M., & Debei, S. (2018, June). An experimental comparison of ros-compatible stereo visual slam methods for planetary rovers. In *2018 5th IEEE International Workshop on Metrology for AeroSpace (MetroAeroSpace)* (pp. 386-391). IEEE.

[3] Filipenko, M., & Afanasyev, I. (2018, September). Comparison of various slam systems for mobile robot in an indoor environment. In *2018 International Conference on Intelligent Systems (IS)* (pp. 400-407). IEEE.

[4] Ibragimov, I. Z., & Afanasyev, I. M. (2017, October). Comparison of ros-based visual slam methods in homogeneous indoor environment. In *2017 14th Workshop on Positioning, Navigation and Communications (WPNC)* (pp. 1-6). IEEE.

[5] da Silva, B. M., Xavier, R. S., do Nascimento, T. P., & Gonsalves, L. M. (2017, November). Experimental evaluation of ROS compatible SLAM algorithms for RGB-D sensors. In *2017 Latin American Robotics Symposium (LARS) and 2017 Brazilian Symposium on Robotics (SBR)* (pp. 1-6). IEEE.

[6] Gaspar, A. R., Nunes, A., Pinto, A., & Matos, A. (2017, November). Comparative Study of Visual Odometry and SLAM Techniques. In *Iberian Robotics conference* (pp. 463-474). Springer, Cham.

RTAB-Map Vs. ORB-SLAM2

- **A number of studies were conducted to compare SLAM systems including ORB-SLAM 1&2 and RTAB-Map:**

[1] Ragot, N., Khemmar, R., Pokala, A., Rossi, R., & Ertaud, J. Y. (2019, July). Benchmark of Visual SLAM Algorithms: ORB-SLAM2 vs RTAB-Map. In *2019 Eighth International Conference on Emerging Security Technologies (EST)* (pp. 1-6). IEEE.

[2] Giubilato, R., Chiodini, S., Pertile, M., & Debei, S. (2018, June). An experimental comparison of ros-compatible stereo visual slam methods for planetary rovers. In *2018 5th IEEE International Workshop on Metrology for AeroSpace (MetroAeroSpace)* (pp. 386-391). IEEE.

[3] Filipenko, M., & Afanasyev, I. (2018, September). Comparison of various slam systems for mobile robot in an indoor environment. In *2018 International Conference on Intelligent Systems (IS)* (pp. 400-407). IEEE.

[4] Ibragimov, I. Z., & Afanasyev, I. M. (2017, October). Comparison of ros-based visual slam methods in homogeneous indoor environment. In *2017 14th Workshop on Positioning, Navigation and Communications (WPNC)* (pp. 1-6). IEEE.

[5] da Silva, B. M., Xavier, R. S., do Nascimento, T. P., & Gonsalves, L. M. (2017, November). Experimental evaluation of ROS compatible SLAM algorithms for RGB-D sensors. In *2017 Latin American Robotics Symposium (LARS) and 2017 Brazilian Symposium on Robotics (SBR)* (pp. 1-6). IEEE.

[6] Gaspar, A. R., Nunes, A., Pinto, A., & Matos, A. (2017, November). Comparative Study of Visual Odometry and SLAM Techniques. In *Iberian Robotics conference* (pp. 463-474). Springer, Cham.

RTAB-Map Vs. ORB-SLAM2

Comparison of ROS-based Visual SLAM methods in homogeneous indoor environment

- **Setup:**
  - LRF, monocular camera, stereo camera, Kinect 2.0 depth sensor.
  - Processing was done on mounted laptop device.
- **Testing data:**
  - Self gathered dataset included navigating the robot along segmented straight lines.
- **Evaluation:**
  - **HectorSLAM** Lidar data are treated as ground truth.
  - **Max. and Avg. deviation** of trajectory was used as performance metric.

RTAB-Map Vs. ORB-SLAM2

Comparison of ROS-based Visual SLAM methods in homogeneous indoor environment

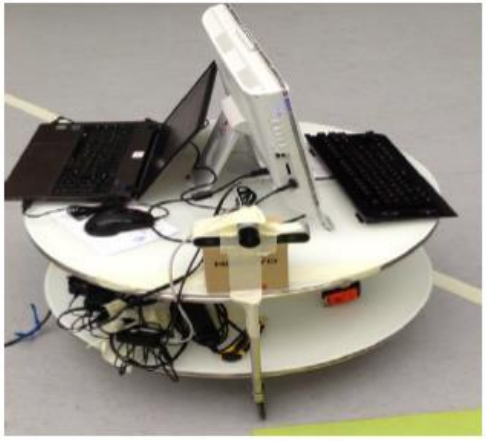


Figure 1. Human-operated prototype of Unmanned Ground Vehicle (UGV).

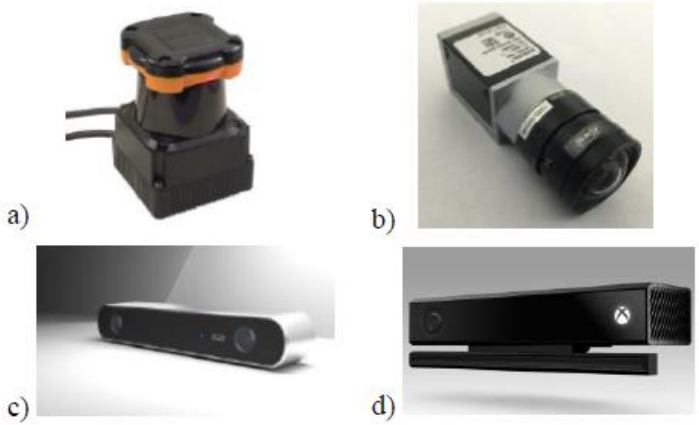


Figure 2. Sensors of the UGV prototype: a) Hokuyo UTM-30LX laser rangefinder. b) Basler acA2000-50gc camera. c) Stereolabs ZED camera. d) Microsoft Kinect 2.0 depth sensor. Courtesy of sensors manufacturers.

Figure 3. Workspace for UGV prototype motion along white line.

RTAB-Map Vs. ORB-SLAM2

Comparison of ROS-based Visual SLAM methods in homogeneous indoor environment

- **Key Findings:**
  - ORB-SLAM point cloud is very sparse and has very little details.
  - ORB-SLAM created less amount of outliers.
  - RTAB-Map created a dense point cloud with a small number of outliers but greater than ORB-SLAM
  - RTAB-Map showed repeated surfaces in the generated map.
  - RTAB-Map suffered odometry noise.
  - **ORB-SLAM both average and maximum trajectory deviation was less than that of RTAB-Map while RTAB-Map map quality was better due to its density.**

UNIVERSITY OF ALBERTA

RTAB-Map Vs. ORB-SLAM2

Comparative Study of Visual Odometry and SLAM Techniques

- **Setup:**
  - ORB-SLAM2: 16G RAM, 8 CPUs, 240G SSD, 512G HDD
  - RTAB-Map: Virtual machine, 7G RAM, 4 CPUs, 512G SSD
- **Testing data:**
  - KITTI and MIT State Center datasets were used for evaluation.
- **Evaluation:**
  - Ground truth data provided with each dataset was used.
  - **Max. and Avg. errors** of trajectory was used as performance metric.
  - **CPU utilization and execution time** were reported as indicators for efficiency.

RTAB-Map Vs. ORB-SLAM2

Comparative Study of Visual Odometry and SLAM Techniques

- **Key Findings:**
  - ORB-SLAM2 generated better trajectories in most of the cases.
  - ORB-SLAM2 can lose a lot of loop closure opportunities when compared to other methods.
  - RTAB-Map showed better efficiency and smaller execution time when depending on internal odometry options, however, higher CPU utilization.
  - RTAB-Map is dependent on the environment and can suffer empty space anomalies in indoor environments.

UNIVERSITY OF ALBERTA

Problem Definition: Revisiting

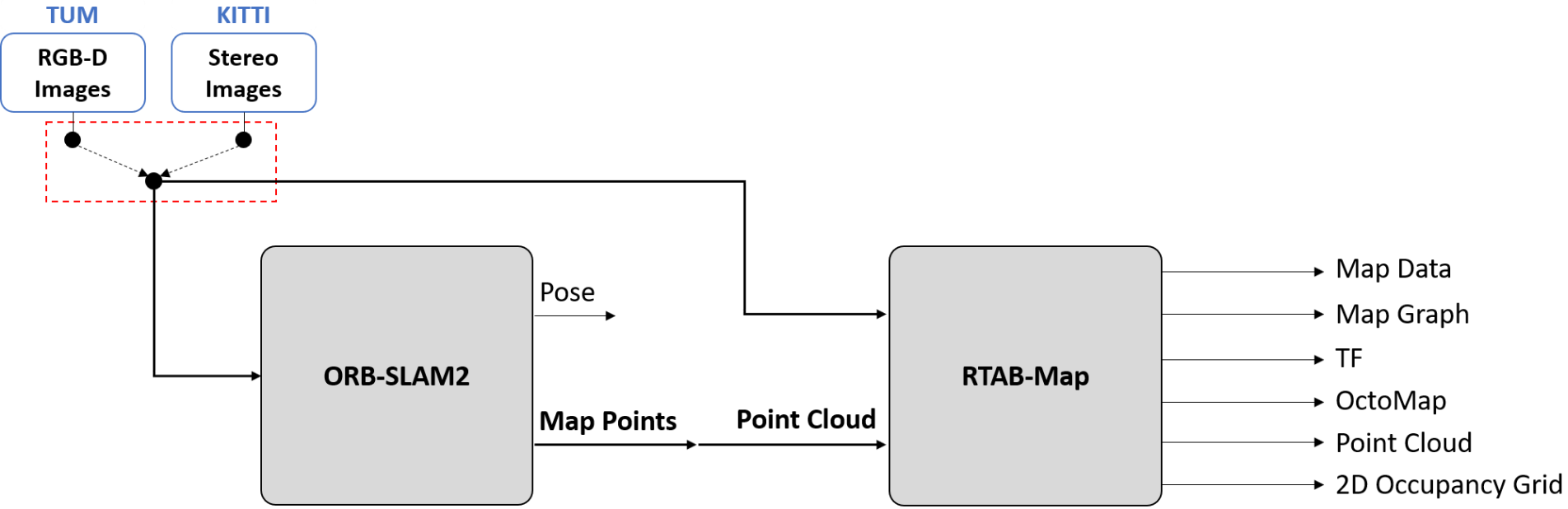
- **Both ORB-SLAM2 and RTAB-Map are state-of-the-art SLAM algorithms but each has performance issues under certain conditions due to its architecture.**

- **ORB-SLAM2 optimizes both camera poses and map points (2 levels of optimization) while RTAB-Map only optimize camera poses.**

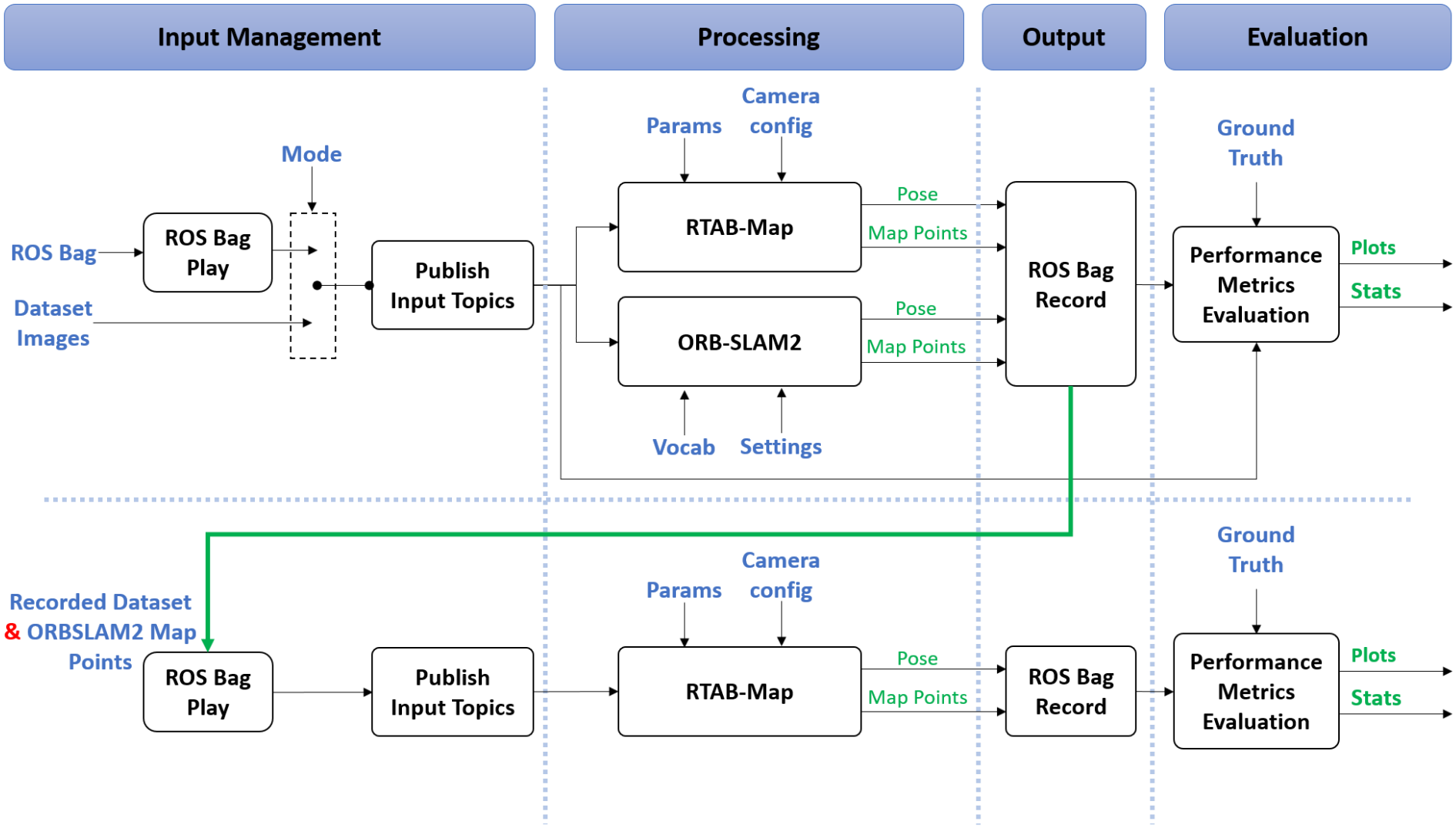
- **Integration of both algorithms can yield more reliable solution and can achieve better performance under wider range of scenarios.**

System Block Diagram: Abstract

- **In a nutshell, we are trying to utilize the optimized map points from ORB-SLAM2 in RTAB-Map in order to enhance its accuracy.**
- **Input controls are dependent on the dataset under test (DUT) and is sent to both ORB-SLAM2 and RTAB-Map.**

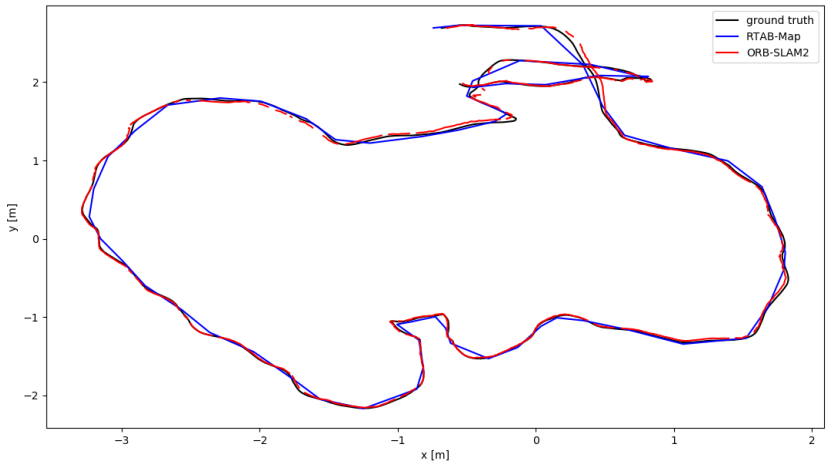


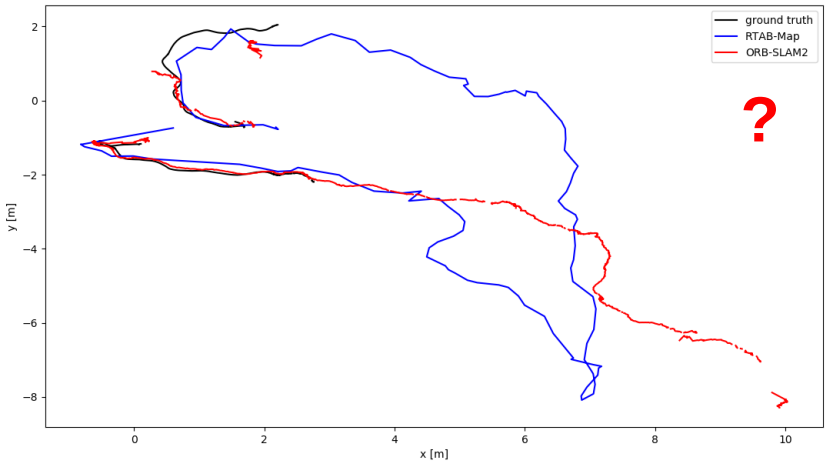
System Block Diagram: Detailed

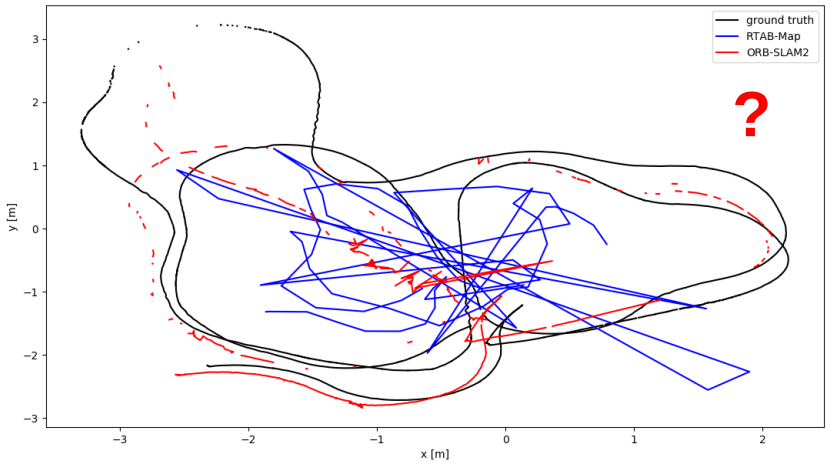


Initial Baseline Results

- **TUM RGB-D Plots:**
  - Freiburg2 trajectories have large errors when compared to freiburg3.
  - Different maybe in the un-distortion applied only to freiburg3.







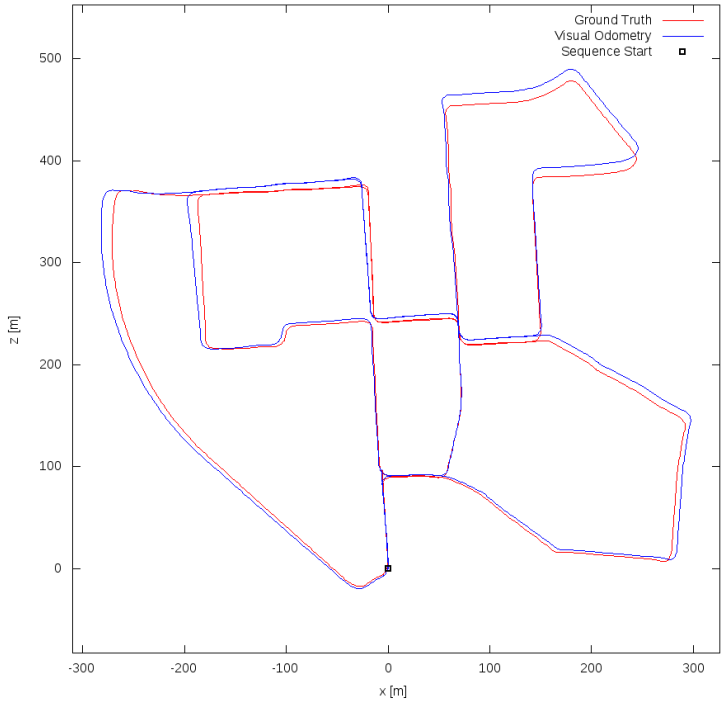
Initial Baseline Results

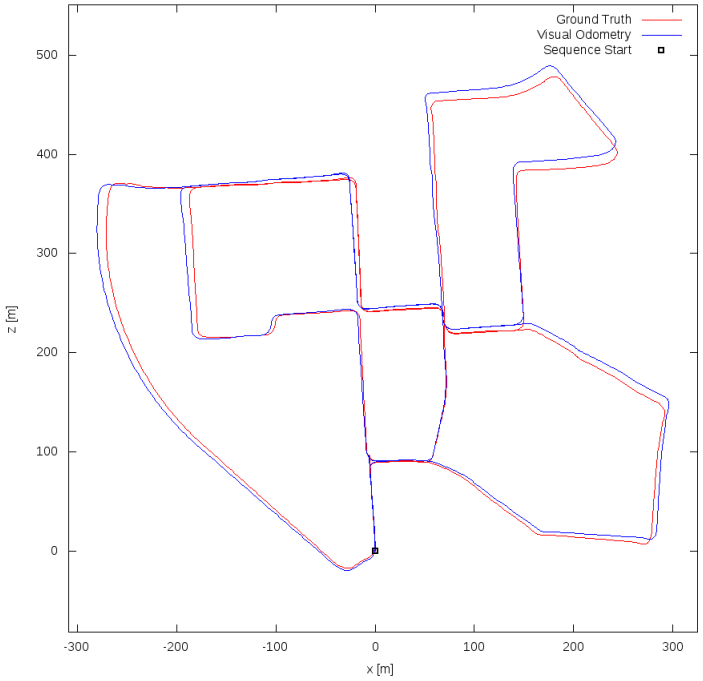
- **TUM RGB-D Dataset Quantitive Results**

Sequence	SLAM system	Absolute Trajectory Error (m)					
		RMSE	mean	median	std	min	max
Freiburg2 ? Large with loop	RTAB-Map	0.396	0.363	0.360	0.159	0.129	0.691
	ORB-SLAM2	0.155	0.124	0.102	0.093	0.015	0.830
Freiburg2 ? Pioneer slam	RTAB-Map	1.594	1.220	0.967	1.028	0.314	7.112
	ORB-SLAM2	0.526	0.350	0.230	0.385	0.033	3.065
Freiburg3 Long office household	RTAB-Map	0.043	0.040	0.039	0.013	0.004	0.074
	ORB-SLAM2	0.030	0.025	0.019	0.016	0.001	0.095

Initial Baseline Results

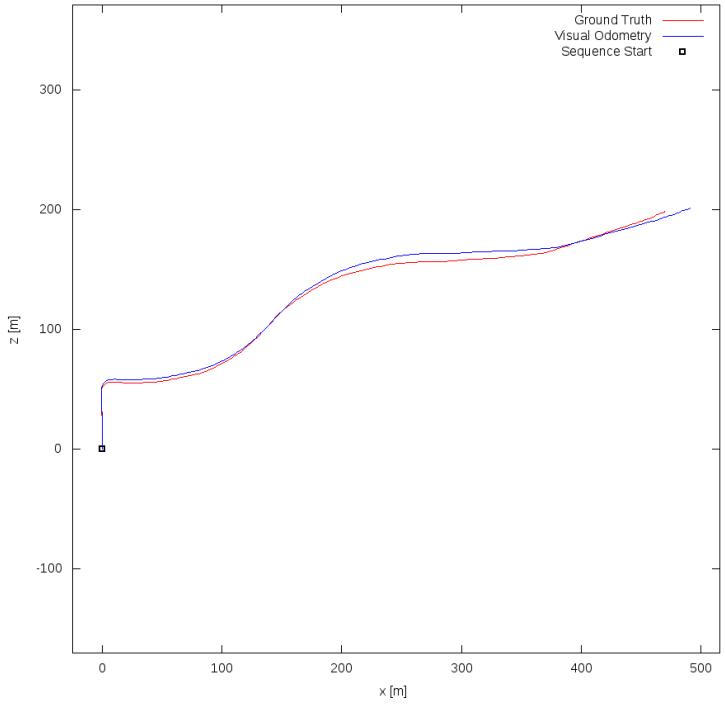
- **KITTI Stereo Plots:**
  - Sequence 00
  - RTAB-Map (left) & ORB-SLAM2 (right)

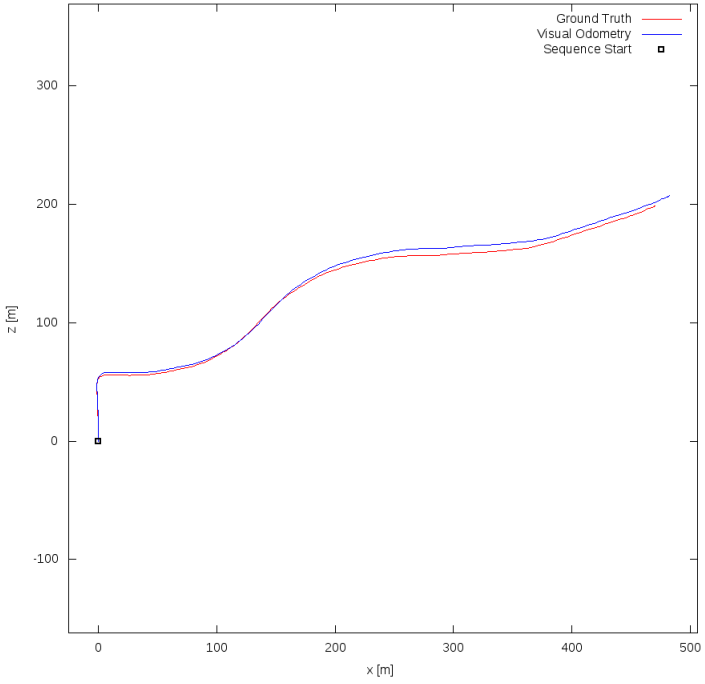




Initial Baseline Results

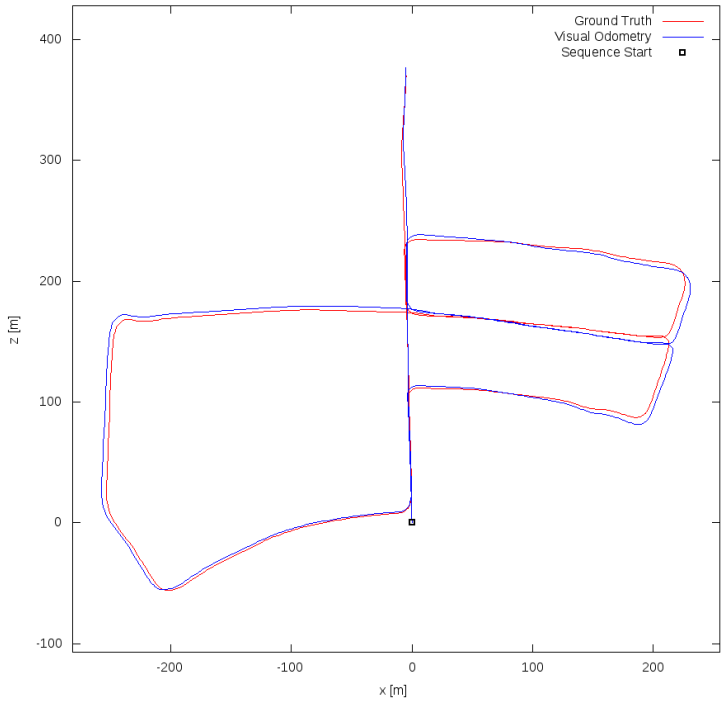
- **KITTI Stereo Plots:**
  - Sequence 03
  - RTAB-Map (left) & ORB-SLAM2 (right)

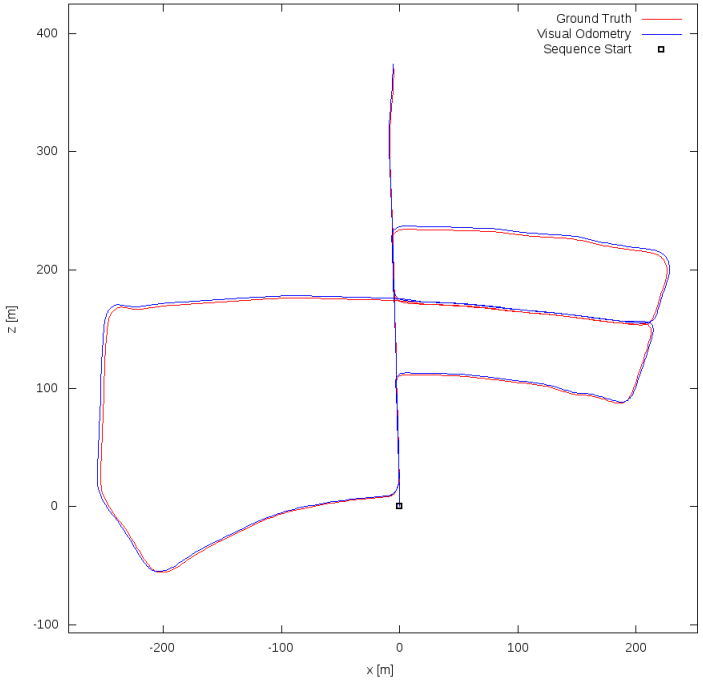




Initial Baseline Results

- **KITTI Stereo Plots:**
  - Sequence 05
  - RTAB-Map (left) & ORB-SLAM2 (right)





UNIVERSITY OF ALBERTA

Initial Baseline Results

- **KITTI Stereo Dataset Quantitive Results**

		Absolute Trajectory Error (m)					
Sequence	**SLAM system**	**RMSE**	**mean**	**median**	**std**	**min**	**max**
Sequence 00	**RTAB-Map**	3.210	6.189	5.845	3.211	0.092	12.426
	ORB-SLAM2	2.964	5.850	5.865	2.965	0.131	11.486
Sequence 03	**RTAB-Map**	6.476	10.329	10.190	6.488	0.096	21.573
	ORB-SLAM2	4.623	7.709	7.690	4.631	0.049	14.979
Sequence 05	**RTAB-Map**	1.868	3.705	3.600	1.869	0.024	7.171
	ORB-SLAM2	0.807	2.162	2.170	0.807	0.012	3.784

UNIVERSITY OF ALBERTA

Open Questions

1. ORB-SLAM2 generates sparse maps (point clouds), would this be sufficient for RTAB-Map to utilize in its operation?

2. Is the synchronization block in RTAB-Map enough to handle the synchronization between input images and point cloud?

3. Are KITTI (Stereo) and TUM (RGB-D) enough to evaluate the system and to explore its corner cases?

4. TUM Dataset, is the lack of distortion is the reason behind the degraded performance in freiburg2 sequences?

Questions

Thank You !