Towards a Swift Multiagent SLAM System for Large-Scale Robotics Applications

by

Muhammad Usman Maqbool Bhutta

A Thesis Submitted to
The Hong Kong University of Science and Technology
in Partial Fulfillment of the Requirements for
the Degree of Doctor of Philosophy
in the Department of Electronic and Computer Engineering

February 2021, Hong Kong

Authorization

I hereby declare that I am the sole author of the thesis.

I authorize the Hong Kong University of Science and Technology to lend this thesis to other institutions or individuals for the purpose of scholarly research.

I further authorize the Hong Kong University of Science and Technology to reproduce the thesis by photocopying or by other means, in total or in part, at the request of other institutions or individuals for the purpose of scholarly research.

Muhammad Usman Maqbool Bhutta

February 2021

Towards a Swift Multiagent SLAM System for Large-Scale Robotics Applications

by

Muhammad Usman Maqbool Bhutta

This is to certify that I have examined the above PhD thesis and have found that it is complete and satisfactory in all respects, and that any and all revisions required by the thesis examination committee have been made.

Professor Ming Liu (ECE), Thesis Supervisor

Professor Bertram Emil Shi (ECE), Head of ECE Department

Thesis Examination Committee

1. Professor Ming Liu (Thesis Supervisor)	Department of Electronic and Computer Engineering
2. Professor Wei Zhang	Department of Electronic and Computer Engineering
3. Professor Yiwen Wang	Department of Electronic and Computer Engineering
4. Professor Qiong Luo	Department of Computer Science & Engineering
5. Professor Xiong Rong (External Examiner)	College of Control Science and Engineering, Zhejiang University, Hangzhou, Zhejiang, China.

Department of Electronic and Computer Engineering
The Hong Kong University of Science and Technology
February 2021

Acknowledgment

I would like to express my deep gratitude to my supervisor, Prof. Ming Liu who has given me a lot of advice and kindly support in my research during the years of my PhD study. I would like to thank The Punjab Educational Endowment Fund (PEEF) for providing me with the CMMS PhD scholarship award so I had the valuable opportunity to study here. This thesis and the work it embodies would not have been possible without the continual support of a number of individuals and organizations. I would like to take this opportunity to express my gratitude to everyone who has contributed to the fulfillment of this thesis. I want to gratefully acknowledge and express my appreciation to my supervisor, Prof. Ming Liu, for his invaluable guidance, full support and patience through my PhD studies. His outstanding attributes such as humility, diligence, enthusiasm and dedication, together with his amazing academic insights and on-going belief were constant sources of motivation for me. I am truly privileged to have had him as my supervisor. I would like to thank Prof. Roland Siegwart and Dr. Cesar Dario Cadena Lerma from the Autonomous Systems Lab, ETH Zürich, Switzerland for kindly welcoming me to visit their research group. The discussions with them stimulated many interesting ideas in my research area of place recognition. Thanks must go to The Hong Kong University of Science and Technology for providing the top up support to my PhD scholarship and overseas research scholarship along with a wealth of other supports during the most painful time of my life. It has been my great pleasure to carry out my PhD study in the RAM-LAB @ HKUST research group. I would like to thank everyone for making it friendly and providing desirable environment for research discussions. Special thanks should be given to the staff in the department administration for all their assistance and patience. I am deeply thankful to my amazing parents and for the way they raised me, without which it would be impossible to be who I am at the moment. Last, I would like to dedicate this thesis and everything I do to my love, who is always with me.

To my family...

Contents

List of Figures

List of Tables

Towards a Swift Multiagent SLAM System for Large-Scale Robotics Applications

Muhammad Usman Maqbool Bhutta

Department of Electronic and Computer Engineering

The Hong Kong University of Science and Technology

Abstract

Abstract goes here.

Chapter 1

Introduction

1.1 Introduction-1

1.2 Introduction-2

1.3 Introduction-3

1.4 Thesis Contributions

Some text...

1.4.1 Contribution 1: TITLE

Some text... Our contributions are as follows:

- Some text....

- We introduce Some text....
- Some text....
- Some text....

1.4.2 Contribution 2: TITLE

Some text... Our contributions include:

- Some text...
- Some text...
- Some text...
- Some text...

1.4.3 Contribution 3: TITLE

Some text... The contribution of this research is four-fold, as follows:

- Some text...
- Some text...
- Some text...
- Some text...

1.5 Thesis Overview

The remainder of this thesis is organized as follows. Chapter ...

1.6 Publications

1.6.1 Journal Papers

- M. U. M. Bhutta, M. Kuse, R. Fan, Y. Liu, and M. Liu, "Loop-Box: Multiagent direct SLAM triggered by single loop closure for large-scale mapping," *IEEE Transactions on Cybernetics*, 2020.
- M. U. M. Bhutta, Y. Sun, and M. Liu, "Why-So-Deep: Image correspondence verification by probabilistic spatial landmarks elevation for visual place recognition." Manuscript submitted for publication.

1.6.2 Conference Papers

- M. U. M. Bhutta and M. Liu, "PCR-Pro: 3d sparse and different scale point clouds registration and robust estimation of information matrix for pose graph SLAM," in *2018 IEEE 8th Annual International Conference on CYBER Technology in Automation, Control, and Intelligent Systems (CYBER)*, pp. 354–359, 2018.

1.6.3 Other Publications during Study

- M. U. M. Bhutta, S. Aslam, P. Yun, J. Jiao, and M. Liu, "Smart-Inspect: Micro scale localization and classification of smartphone glass defects for industrial automation," in *2020 IEEE/RSJ International Conference on*

Intelligent Robots and Systems, IROS 2020, Institute of Electrical and Electronics Engineers Inc., 2020.

- H. Ma, Y. Ma, J. Jiao, M. U. M. Bhutta, M. J. Bocus, L. Wang, M. Liu, and R. Fan, "Multiple lane detection algorithm based on optimised dense disparity map estimation," in *2018 IEEE International Conference on Imaging Systems and Techniques (IST)*, pp. 1–5, 2018.

1.7 Related Material and Demo Videos

- "PCR-Pro": https://sites.google.com/view/pcr-pro.
- "Loop-Box": https://usmanmaqbool.github.io/loop-box.
- "MAQBOOL": https://usmanmaqbool.github.io/why-so-deep.
- "Smart-Inspect": https://usmanmaqbool.github.io/smart-inspect.

1.8 References

[1] R. Arandjelovic, P. Gronat, A. Torii, T. Pajdla, and J. Sivic, "NetVLAD: CNN architecture for weakly supervised place recognition," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 5297–5307, 2016.

[2] M. U. M. Bhutta, M. Kuse, R. Fan, Y. Liu, and M. Liu, "Loop-Box: Multi-agent direct SLAM triggered by single loop closure for large-scale mapping," *IEEE Transactions on Cybernetics*, 2020.

[3] M. U. M. Bhutta, Y. Sun, and M. Liu, "Why-So-Deep: Image correspondence verification by probabilistic spatial landmarks elevation for visual place recognition." Manuscript submitted for publication.

[4] M. U. M. Bhutta and M. Liu, "PCR-Pro: 3d sparse and different scale point clouds registration and robust estimation of information matrix for pose graph SLAM," in *2018 IEEE 8th Annual International Conference on CYBER Technology in Automation, Control, and Intelligent Systems (CYBER)*, pp. 354–359, 2018.

[5] M. U. M. Bhutta, S. Aslam, P. Yun, J. Jiao, and M. Liu, "Smart-Inspect: Micro scale localization and classification of smartphone glass defects for industrial automation," in *2020 IEEE/RSJ International Conference on Intelligent Robots and Systems, IROS 2020*, Institute of Electrical and Electronics Engineers Inc., 2020.

[6] H. Ma, Y. Ma, J. Jiao, M. U. M. Bhutta, M. J. Bocus, L. Wang, M. Liu, and R. Fan, "Multiple lane detection algorithm based on optimised dense dis-

parity map estimation," in *2018 IEEE International Conference on Imaging Systems and Techniques (IST)*, pp. 1–5, 2018.

Chapter 2

Scientific Background and Literature Review

This chapter includes all your thesis literature review.

2.1 SLAM

2.1.1 Graph SLAM

.

Bundle Adjustment

2.2 References

Chapter 3

Figures

3.1 Single Image Configuration

Use can use pdf image as well.

Figure 3.1: Single Image Configuration.

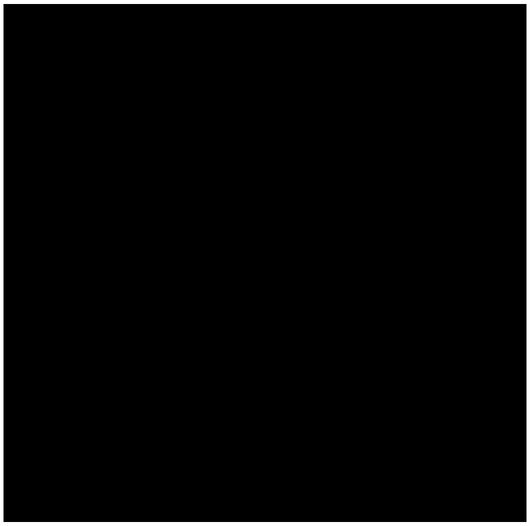
3.2 1-1 Configuration

(a) First Image

(b) Second Image.

Figure 3.2: 1-1 Configuration.

3.3 2-1 Configuration



(a) Image A.

(b) Image B.

(c) Image C.

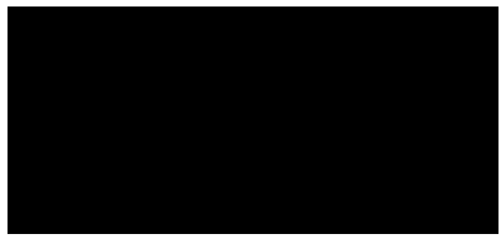
Figure 3.3: 2-1 Configuration.

3.4 4 x 4 Configuration

(a) Image 21.

(b) Image 23.

(c) Image 22.



(d) Image 24.

Figure 3.4: 4 x 4 Configuration.

3.5 3 x 2 Configuration

(a) Image 1.

(b) Image 2.

(c) Image 3.

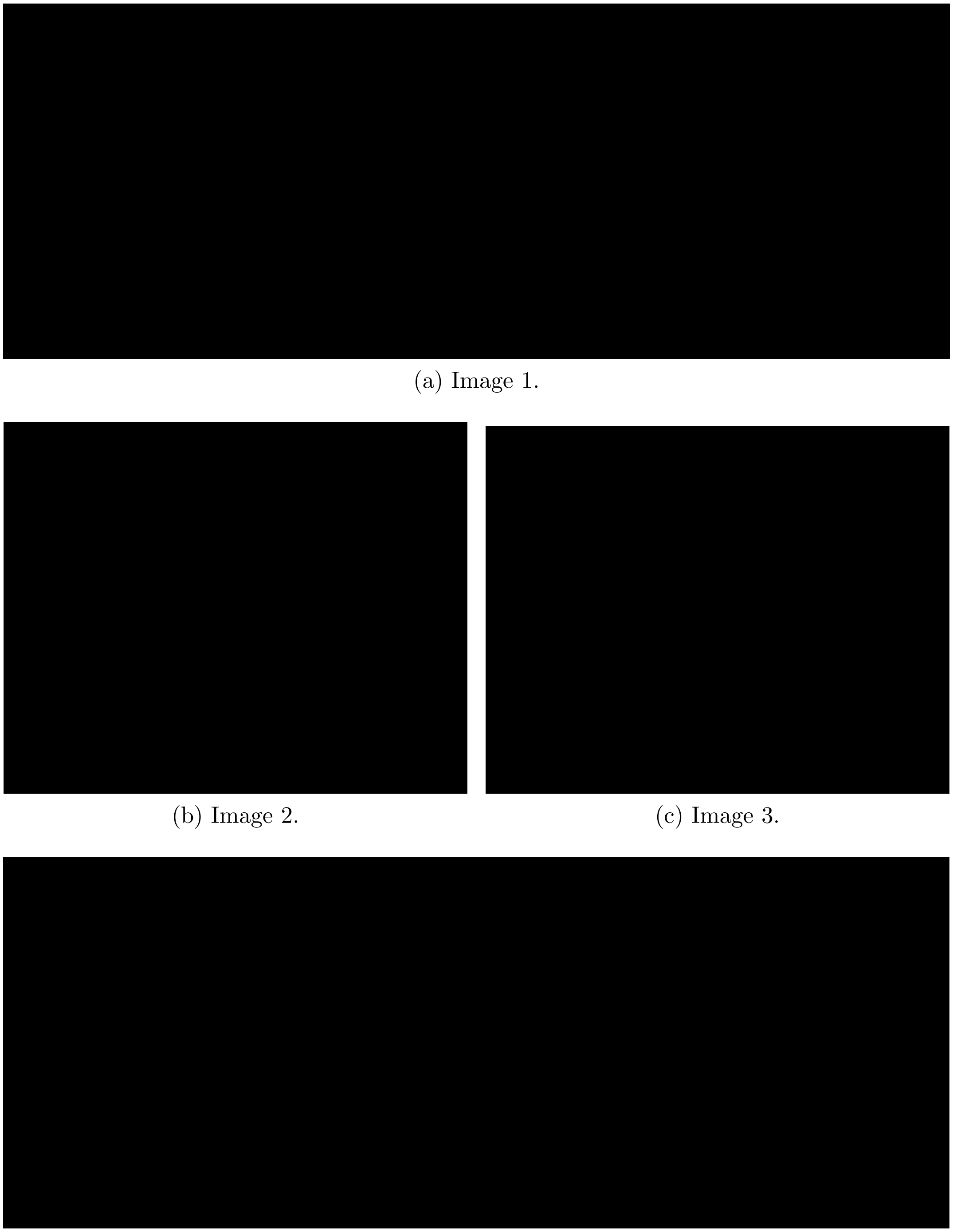
(d) Image 4.

(e) Image 5.

(f) Image 6.

Figure 3.5: (a), (b), and (c) and (d), (e), and (f) correspond to the 3 x 2 Configuration.

3.6 1-2-1 Configuration



(a) Image 1.

(b) Image 2.

(c) Image 3.

(d) Image 4.

Figure 3.6: 1-2-1 Configuration.

3.7 Appendix

3.7.1 Appendix subsection

Problem definition

3.8 References

Chapter 4

Equations and Algorithm

4.1 Equations and Algorithm

Let us assume γ_{ij} is the number of matched keypoints among two keyframes, i and j. These matches yield a distinct scale difference σ_{ij} depending on the number of matched keypoints γ_{ij}. The optimal scale difference σ^* will be

$$\sigma^* = \underset{\gamma}{\operatorname{argmax}} \frac{1}{2} |\gamma(\sigma_{ij}), \gamma(\sigma_{i'j'})|, \tag{4.1}$$

where σ_{ij} and $\sigma_{i'j'}$ are the two nearest points such that

$$|\sigma_{ij} - \sigma_{i'j'}| \leq \Delta^* \quad \forall \quad i, j, i' \text{ and } j' \in \mathbb{Z}^+, \quad \Delta^* \in \mathbb{R}. \tag{4.2}$$

The estimation is further explained in Algorithm 1.

Algorithm 1: Finest Tuning for Optimal Scale Estimation

Input: Matched keyframes $\mathbf{K}_{r_{ID}} = \{\mathbf{K}_s^i, \mathbf{K}_t^j\}$, poses ${}^w\mathbf{T}_{r_{ID}} = \{{}^w\mathbf{T}_s, {}^w\mathbf{T}_t\}$, point clouds $P_{r_{ID}}^{\mathcal{F}} = \{P_{s(i)}^{\mathcal{F}_s^i}, P_{t(j)}^{\mathcal{F}_t^j}\}$ with $i, j \in \mathbb{Z}^+$

Output: Optimal scale difference σ^*, initial guess relative transformation ${}^{si}\mathbf{T}_{ti}^{IG}$

initialization;

for $z = -1 : 1$ **do**

  $P_{s(i+z)}^{\mathcal{F}_w} = {}^w\mathbf{T}_{s(i+z)}(P_{s(i+z)}^{\mathcal{F}_{s(i+z)}})$;

  $P_{t(j+z)}^{\mathcal{F}_w} = {}^w\mathbf{T}_{t(j+z)}(P_{t(j+z)}^{\mathcal{F}_{t(j+z)}})$;

  Function `PCR-Pro [1]`$(\mathbf{K}_{r_{ID}}, P_{r_{ID}}^{\mathcal{F}_w})$**:**

    Estimate volume ratio r_{vol} of $P_{s(i+z)}^{\mathcal{F}_w}, P_{t(j+z)}^{\mathcal{F}_w}$;

    ${}^{s(i+z)}\mathbf{T}_{t(j+z)}^{RC} \longleftarrow \gamma_z \longleftarrow \mathbf{K}_s^{i+z}, \mathbf{K}_t^{j+z}$;

    $\sigma_z \longleftarrow {}^{s(i+z)}\mathbf{T}_{t(j+z)}^{RC}, \gamma_z, {}^w\mathbf{T}_{s(i+z)}, {}^w\mathbf{T}_{t(j+z)}$;

    ${}^{s(i+z)}\mathbf{T}_{t(j+z)}^{IG} \longleftarrow \sigma_z, P_{s(i+z)}^{\mathcal{F}_w}, P_{t(j+z)}^{\mathcal{F}_w}$;

    return $\sigma_z, {}^{s(i+z)}\mathbf{T}_{t(j+z)}^{IG}$;

if $r_{vol} > 0.5$ **then**

  $\Delta^* = 5$;

  for $x = -1 : 1$ **do**

    for $y = -1 : 1$ **do**

      if $x \neq y$ **then**

        $\Delta = |\sigma_x - \sigma_y|$;

        if $\gamma^* < \gamma_{xy}$ *&&* $\Delta^* > \Delta$ *&&* $\Delta^* \neq 0$ **then**

          $\sigma^* = avg(\sigma_x, \sigma_y)$;

          $\Delta^* = \Delta$;

          $\gamma^* = \gamma_{xy}$;

else

  $\sigma^* = \sigma_{xy=00}$;

4.2 Appendix

4.3 References

[1] M. U. M. Bhutta and M. Liu, "PCR-Pro: 3d sparse and different scale point clouds registration and robust estimation of information matrix for pose graph SLAM," in *2018 IEEE 8th Annual International Conference on CYBER Technology in Automation, Control, and Intelligent Systems (CYBER)*, pp. 354–359, 2018.

Chapter 5

Tabls and Graphs

5.1 Table

Table 5.1 is shown below.

You can also create different style of latex table using link https://www.tablesgenerator.com/.

Table 5.1: Table I

Method	Type	Tested On		
		Recall@1	Recall@5	Recall@10
Method 1	Type 1	44.76	60.95	70.16
	Type 2	52.70	67.30	73.02
Method 2 [1]	Type 1	60.00	73.65	79.05
	Type 2	58.73	74.6	80.32
Method 3 [2]	Type 1	61.90	77.78	80.95
	Type 2	66.98	80.95	83.81
Method 4 (Ours)	Type 1	66.98	80.95	**85.71**
	Type 2	**67.30**	**81.27**	**85.71**

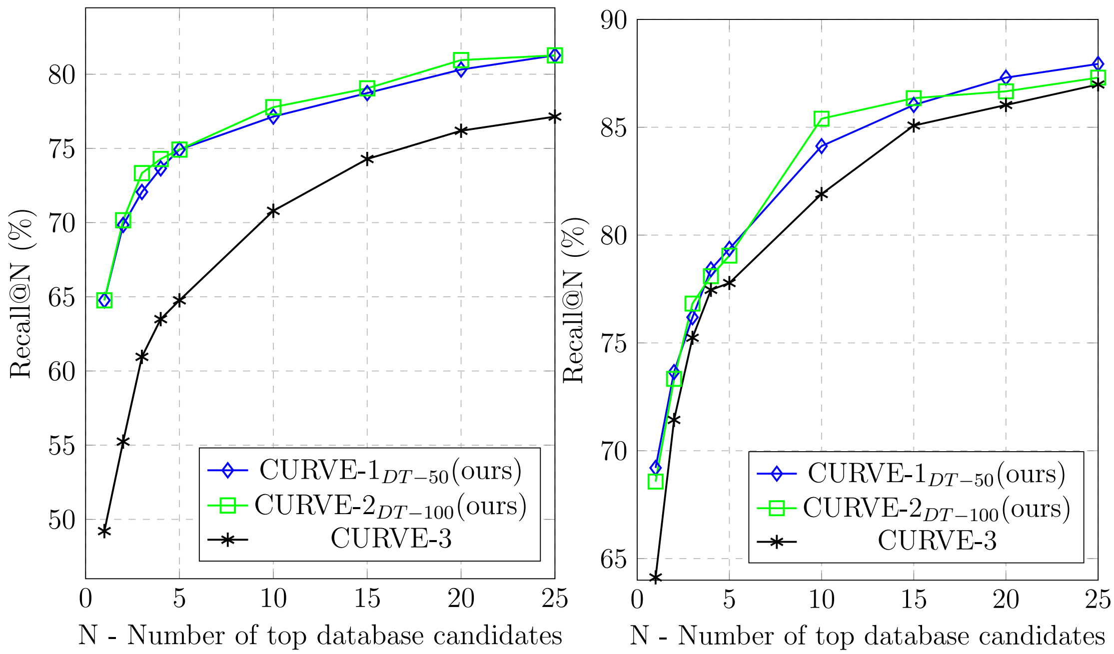
Table 5.2: Table II

Method	**Type**	**Number**	**Total Time Duration (sec)**		**Field 3**	
			Field 1	Field 2	Field 1 Time (sec)	Field 2
Method 1	Indoor	2	53.7	69	0.98	2.4922
Method 2	Outdoor	2	73	74	0.89	2.83519
Method 3	Outdoor	2	104	91	0.94	1.7857
Method 4	Outdoor	3	122	117	0.92	1.2786
Method 5			122	112	0.867	1.15

Table 5.3: Table III

Benchmarks	**Direction**	**Method 1** [3]		**Method 2** [4]		**Method 3 (Proposed)**	
		Time (sec)	RMSE (m)	Time (sec)	RMSE (m)	Time (sec)	RMSE (m)
Benchmark 1	Same	7.6	0.1503	5.1	0.0903	**2.53**	**0.0290**
Benchmark 2	Same	8.8	1.2453	5.5	0.9832	**2.93**	**0.0323**
Benchmark 3	Same	7.4	1.5594	5.04	0.0883	**2.47**	**0.0486**
Benchmark 4	Opposite	9	1.6	5.57	0.207	**3**	**0.175**
Benchmark 5	Same	11	0.3613	6.24	0.3073	**3.67**	**0.0712**

5.2 Tikz styles Graphs



(a) Example 1

(b) Example 1

Figure 5.1: **Note:** All the curves are in data folders and in 'dat' file. Using '**pgfplotsset**', you can also control the y-axis. '**tikz_styles.tex**' has all the further configuration.

5.3 References

[1] R. Arandjelovic, P. Gronat, A. Torii, T. Pajdla, and J. Sivic, "NetVLAD: CNN architecture for weakly supervised place recognition," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 5297–5307, 2016.

[2] Y. Zhu, J. Wang, L. Xie, and L. Zheng, "Attention-based pyramid aggregation network for visual place recognition," in *Proceedings of the 26th ACM international conference on Multimedia*, pp. 99–107, 2018.

[3] M. U. M. Bhutta and M. Liu, "PCR-Pro: 3d sparse and different scale point clouds registration and robust estimation of information matrix for pose graph SLAM," in *2018 IEEE 8th Annual International Conference on CYBER Technology in Automation, Control, and Intelligent Systems (CYBER)*, pp. 354–359, 2018.

[4] R. Kummerle, G. Grisetti, H. Strasdat, K. Konolige, and W. Burgard, "g^{2}o: A general framework for graph optimization," in *2011 IEEE International Conference on Robotics and Automation*, pp. 3607–3613, IEEE, May 2011.

Chapter 6

Conclusion and Future Work

Here is conclusion...

6.1 Summary of Contributions

6.1.1 Contribution 1

6.1.2 Contribution 1

6.2 Future Work and Challenges

6.2.1 Future Work 1

Part 1

Part 2