

3D non-rigid SLAM in minimally invasive surgery

by

Jingwei Song

A thesis submitted in partial fulfilment for the
degree of Doctor of Philosophy

at the
Centre for Autonomous Systems
Faculty of Engineering and Information Technology
University of Technology Sydney

March 2020

Declaration of Authorship

I certify that the work in this thesis has not previously been submitted for a degree nor has it been submitted as part of requirements for a degree except as fully acknowledged within the text.

I also certify that the thesis has been written by me. Any help that I have received in my research work and the preparation of the thesis itself has been acknowledged. In addition, I certify that all information sources and literature used are indicated in the thesis.

This research is supported by an Australian Government Research Training Program Scholarship.

Signed: _____
Production Note: Signature removed prior to publication.

Date: 22 March 2020

UNIVERSITY OF TECHNOLOGY SYDNEY

Abstract

Faculty of Engineering and Information Technology

Centre for Autonomous Systems

Doctor of Philosophy

by Jingwei Song

Aiming at reducing trauma and morbidity associated with large incisions in open surgery, minimally invasive surgery (MIS) has been widely acquired in clinical practice as a powerful tool enabling patients with less pain, shorter hospital stay, and fewer complications. However, MIS narrows the surgeon's field of view which confines visual information when implementing MIS. Therefore, a stereoscope or monocular scope is an essential tool for capturing and transmitting 2D images during the procedure.

Although numbers of special sensors including laser, structured light, time-of-flight cameras have been applied or investigated in MIS, RGB scope is still widely applied in the intro-operative system because it is non-invasive and cheap to be installed. Thus it is an important topic to rebuild and visualize the latest deformed shape of soft-tissue surfaces to mitigate tissue damages from stereo or monocular scopes. This research aims at proposing innovative robocentric simultaneous localization and mapping (SLAM) algorithm for deformable dense reconstruction of soft-tissue surfaces using a sequence of images obtained from a stereoscope or monocular camera. In this paper, we try to solve the problem by introducing a warping field based on the embedded deformation (ED) nodes which makes full use of the 3D shapes recovered from consecutive pairs of stereo images by deforming the last updated model to the current live model. Our robocentric SLAM system (off-line and tested on stereo videos) can: (1) Incrementally build a live model by progressively fusing new observations with vivid accurate texture. (2) Estimate the deformed shape of

the unobserved region with the principle As-Rigid-As-Possible. (3) Perform the dynamic model shape deformation. (4) Estimate the current relative pose between the soft-tissue and the scope.

We further improve and optimize the proposed robocentric deformable SLAM algorithm to MIS-SLAM: a complete real-time large scale robocentric dense deformable SLAM system with stereoscope in MIS based on heterogeneous computing by making full use of CPU and GPU. Idled CPU is used to perform ORB-SLAM for providing robust global pose. Strategies are taken to integrate modules from CPU and GPU. We solve the key problem raised in previous work, that is, fast movement of scope and blurry images make the scope tracking fail. Benefiting from improved localization, MIS-SLAM can achieve large scale scope localizing and dense mapping in real-time. It transforms and deforms the current model and incrementally fuses new observation while keeping the vivid texture. In-vivo experiments conducted on publicly available datasets presented in the form of videos demonstrate the feasibility and practicality of MIS-SLAM for potential clinical purpose.

In MIS-SLAM, however, it remains challenging to keep constant speed in deformation nodes parameter estimation when the model grows larger. In practice, the processing time grows rapidly in accordance with the expansion of the maps. Therefore, we propose an approach to decouple nodes of deformation graph in large scale robocentric dense deformable SLAM and keep the estimation time to be constant. We discover that only partial deformable nodes in the graph are connected to visible points. Based on this principle, the sparsity of the original Hessian matrix is utilized to split parameter estimation into two independent steps. With this new formulation, we achieve faster parameter estimation with amortized computation complexity reduced from $O(n^2)$ to closing $O(1)$. As a result, the computation cost barely increases as the map keeps growing. By our strategy, the bottleneck of limited computation in estimating deformation field in large scale environment has been overcome. The effectiveness is validated by experiments, featuring large scale deformation scenarios.

In addition to robocentric SLAM, this thesis also aims at developing a general SLAM which estimates the scope poses correctly. An elaborate observability analysis is conducted on

the ED graph. We demonstrate and prove that the ED graph widely used in such scenarios is unobservable and leads to multiple solutions unless suitable priors are provided. Example, as well as theoretical prove, are provided to show the ambiguity of ED graph and scope pose. Different from robocentric SLAM, in modeling non-rigid scenario with ED graph, motion priors of the deforming environment is essential to separate robot pose and deforming environment. The conclusion can be extrapolated to any free form deformation formulation. In guaranteeing the observability, this research proposes a preliminary deformable SLAM approach to estimate robot pose in complex environments that exhibits regular motion. A strategy that approximates deformed shape using a linear combination of several previous shapes is proposed to avoid the ambiguity in robot movement and rigid and non-rigid motions of the environment. Fisher information matrix rank analysis is performed to prove the effectiveness. Moreover, the proposed algorithm is validated using Monte Carlo simulations and real experiments. It is demonstrated that the new algorithm significantly outperforms conventional SLAM and ED based SLAM especially in scenarios where there is large deformation.

Acknowledgements

First and foremost I would like to express my sincere gratitude to my supervisor A/Prof. Shoudong Huang and Dr. Liang Zhao for the continuous support of my Ph.D study and related research, for their patience, motivation, and immense knowledge. Their guidance helped me in all the time of research and writing of this thesis. I hope I can continue to collaborate with you in the future.

Besides my supervisor, I would like to thank Dr. Teresa Vidal Calleja, A/Prof. Jaime Valls Miro, Prof. Gamini Dissanayake, and A/Prof. Sarath Kodagoda for their guidance and friendship, but also for the hard question which incited me to widen my research from various perspectives.

My sincere thanks also go to Dr. Antonio Agudo and Prof. Francesc Moreno-Noguer, who provided me an opportunity to join their team as an exchange student at Institut de Robòtica i Informàtica Industrial (IRI) of the Spanish Council for Scientific Research and the Technical University of Catalonia. And also for supporting me to win IRI-Maria de Maeztu intern scholarship. During my 4 months visit, I have terrific experience and memory in IRI as well as Barcelona. I am also grateful to Dr. Mitesh Patel for hosting me at FXPAL as an intern; the friendly environment and living in California touches me very much.

I acknowledge the scholarships International Research Training Program Scholarship (IRTP) funded by the Australian Government, Department of Education and Training to support 3.5 years for my research and make my study possible. I would also like to acknowledge Centre for Autonomous Systems and School of Mechanical and Mechatronic Engineering of the Faculty of Engineering and Information Technology at University of Technology Sydney for providing me funding to attend conferences held in London, Barcelona, Madrid, Brisbane, Vancouver and being an exchange student in Barcelona.

Many thanks go to my fellow friends and colleagues at the Centre for Autonomous Systems for the numerous stimulating and helpful discussions and their friendship. My special thanks go to Dr. Jun Wang for inspiring discussions on solving technical issues as well as reporting the state-to-art technologies. We work closely in achieving good publications. I also appreciate Dr. Raphael Falque for useful discussions on theories of deformation technologies. Raphael is senior to me and I benefit a lot from him. I would also like to thank my colleague Kristian Villavicencio for helping with the ElasticFusion experiments.

I thank my friends and labmates Karthick Thiyagarajan, Mahdi Hassan, Lakshitha Dantanarayana, Phillip Quin, Leo Shi, Nalika Ulapane, Buddhi Wijerathna, Asok Aravinda, Julien Collart, Alexander Virgona, Maani Ghaffari Jadidi, Kasra Khosoussi, Daobilige Su, Brendan Emery, Katherine Waldron, Cédric Le Gentil, Kanzhi Wu, Teng Zhang and many other colleagues. I thank my fellow labmates for the stimulating discussions and for all the fun we have had in the last four years, as well as their hands-on help on setting up the experimental environment to collect and process data; Many thanks to Herni Winarta and Katherine Waldron for their help on administrative work.

Additional thanks go to my other friends in Sydney, I enjoy the 3 years in this sunny city. The last but most important thanks give to my parents, who are always proud of me for my achievements, and supported and cared for me throughout these years.

Contents

Declaration of Authorship	i
Abstract	ii
Acknowledgements	v
List of Figures	x
List of Tables	xiv
Acronyms & Abbreviations	xv
Nomenclature	xvii
1 Introduction	1
1.1 Motivation	4
1.2 Research aims	5
1.3 The structure of the thesis	7
1.4 Publications	8
2 Related works	10
2.1 Sensors and single frame shape recovery approaches	10
2.1.1 Stereo shape recovery	11
2.1.2 Shape from shading	12
2.1.3 Electromagnetic tracking device for navigation	13
2.2 Rigid SLAM in non-rigid environment.	15
2.2.1 Sparse rigid SLAM in MIS	15
2.2.2 Dense rigid SLAM in MIS	16
2.3 3D non-rigid SLAM	17
2.3.1 Non-rigid RGB-D SLAM	17
2.3.2 Implementing ED graph SLAM in large scale SLAM	19
2.3.3 Template based non-rigid SLAM	20
2.4 Ambiguity in deformable surface 3D motions	21

3	Modeling soft-tissue deformation with ED graph	26
3.1	Revisit ED deformation graph	27
3.2	Template based SLAM with ED deformation graph	30
3.2.1	The framework of template based structure	31
3.2.2	Technical details	32
3.3	Robocentric template free SLAM with ED deformation graph	33
3.3.1	The robocentric template free SLAM framework	33
3.3.2	Technical details	35
3.4	Results and discussion	38
3.4.1	Template based approach	38
3.4.2	Template free approach	43
3.5	Chapter summary	52
4	MIS-SLAM: A complete robocentric SLAM system for MIS scenario	53
4.1	Overview of MIS-SLAM	54
4.2	Depth estimation, sparse key correspondences and global rigid transformation	55
4.3	Deformation field estimation	56
4.4	Model update with new observation	58
4.5	Results and discussion	62
4.5.1	Robustness enhancement	63
4.5.2	Deforming the model and fusing new depth	65
4.5.3	GPU implementation and computational cost	65
4.5.4	Validation using simulation and ex-vivo experiments	65
4.5.5	Limitations and discussions	66
4.6	Chapter summary	67
5	Efficient two step optimization in ED based SLAM	68
5.1	Efficient two step optimization	69
5.1.1	Matrix form of ED graph deformation	69
5.1.2	Sparsity in ED graph formulation	71
5.1.3	Lossy two-level optimization	72
5.1.4	Connection with marginalization and information loss	74
5.2	Results and discussion	78
5.2.1	Qualitative ED deformation comparisons	78
5.2.2	Time complexity comparisons	79
5.2.3	Accuracy comparisons	81
5.3	Chapter summary	85
6	A time series SLAM algorithm for deformable environment	86
6.1	Observability analysis of ED based SLAM	87
6.1.1	Qualitative analysis of ED based SLAM formulation	88
6.1.2	Prove of unobservability in ED based SLAM formulation	90
6.2	Priori based SLAM formulation	94
6.2.1	Prediction modelling	97
6.2.2	Observability analysis	98

6.3	Results and discussion	102
6.3.1	Monte Carlo simulations	103
6.3.2	SLAM in deformable soft-tissues	105
6.3.3	Observability test	105
6.4	Chapter summary	106
7	Conclusions and future work	108
	Appendices	110
	Bibliography	110

List of Figures

1.1	The difference between open surgery and MIS.	3
2.1	Depth estimation from stereo images.	12
2.2	Illustration of Lambertian reflectance model.	13
2.3	NDI Aurora $\text{\textcircled{R}}$ EM tracking system [1]	14
2.4	The framework of Handeye Calibration.	15
2.5	A demonstration of 3D reconstruction of ElasticFusion.	24
2.6	An example of FEM based curve surface representation.	25
2.7	A toy model demonstrating ambiguity between camera and soft-tissue. Red dots constitute the heart and blue dots are the observation from camera. . .	25
3.1	A toy example of an ED graph. The red circles are ED nodes, say node j , encoding a geometric position \mathbf{g}_j , and an affine transformation given by \mathbf{A}_j and \mathbf{t}_j . The blue triangle is a vertex, that can be deformed from \mathbf{v}_i to $\tilde{\mathbf{v}}_i$, through the impact of its neighboring ED nodes.	28
3.2	The framework of the proposed template based deformable soft-tissue reconstruction based on DFF and pre-operative CT model.	30
3.3	(A) is the DFF volume recording distance field values. (B) is a section of the volume. The black line is an example planar and each voxel records its distance to the planar.	32
3.4	The framework of the robocentric template free SLAM with ED deformation graph.	34
3.5	(a) to (e) are the simulations of generating the depth scan observation from the deformed liver model. The blue points are the simulated depth observations. The points in red is the deformed model.	40
3.6	The results of model-to-scan registration colored by the matching error (mm) which is directly obtained from the DFF. (a)-(d) are selected error map from the heart model; (e)-(h) are selected error map from the right kidney model; (i)-(l) are selected error map from the liver model.	41
3.7	The comparison between the deformed models recovered from the robocentric template free SLAM and the ground truth used for generating the depth observations, by using the heart model (a) - (c), the kidney model (d) - (f) and the liver model (g) - (i) respectively. The models in green are the ground truth, while the models in white are the recovered soft-tissues. .	42

3.8	The comparison between the dense SURF and SIFT using stereo videos of abdomen wall. Results imply that dense SURF can generate more key points which are critical in soft-tissue matching while SIFT produce less or even no correspondences.	46
3.9	The comparison between pipeline with and without dense SURF constraint (Left is with constraint while right is without). Significant errors happen either in texture or in topologies without SURF constraint.	47
3.10	Non-rigid reconstruction of different soft tissues using in-vivo datasets. Illustrated are the sequences of 3D reconstructions. The five videos are (from top to bottom): abdomen (1), abdomen wall, liver, abdomen (2) and abdomen (3).	48
3.11	The Axial, Coronal, Sagittal and 3D views of the deformed model and ground truth at the last frame (heart). The red points denote the scan of the last frame.	50
3.12	Ex-vivo validation with the two Hamlyn validation datasets: Silicon heart phantoms deforming with cardiac motion and associated CT scans. The upper figure is the time series of average error. The lower figures are the reconstructed geometry and corresponding error maps measured by distance to ground truth.	51
3.13	Comparison with VolumeDeform approach. Left is the result of the robot-centric template free SLAM. The right is VolumeDeform's result. Note the difference in texture (letters on the T-shirt).	51
4.1	The framework of MIS-SLAM. CPU is responsible for ORB-SLAM, uploading features, rigid and start a visualization module. GPU processes depth estimation, registration, fusion and visualization.	54
4.2	Examples of depth and smoothed depth.	56
4.3	MIS-SLAM processes 3 in-vivo datasets. Figures present the whole constructed model at different frames. The three videos are (from top to bottom): Abdomen wall (1), abdomen (2) and abdomen example (3).	61
4.4	Comparisons between Section 3.3.2 (First row) and the proposed MIS-SLAM (Second row).	63
4.5	The Axial, Coronal, Sagittal and 3D views of the deformed model and ground truth at the last frame (liver). The red points denote the scan of the last frame.	64
5.1	Illustrated are the spatial relations of the visible points and the node graph. (a) is the latest reconstruction. (b) shows both the model (red) and the target depth (blue). (c) is the ED nodes and their edges. (d) presents the ED nodes and the target depth.	72
5.2	(a) is an example of Jacobian. Empty block means the element in this block are zero. (b) is re-ordered Jacobian.	73
5.3	(a) is Jacobian of E_{rot} with regard to all nodes. (b) is Jacobian of E_{reg} with regard to all nodes.	76
5.4	Two types of nodes and edges. Green nodes are the PR nodes and purple nodes are PI nodes.	77

5.5	Connections of PI (purple) and PR (green) nodes. Left is the full connection while the PI and PR connections are cut in Level I optimization	77
5.6	Qualitative comparisons of the proposed strategy and original ED based deformation. The first shape is the original dolphin mesh. It shows the result of deformed shape (the last) along with the result of classical ED deformation (middle).	79
5.7	A toy model. One step optimization step from plain visible surface model (tiny dots) to warped surface (grid). PI nodes are in blue and PR nodes are in green. Camera remains static.	80
5.8	Test results relating to Fig. 5.7.	80
5.9	Processing time comparisons of model 6 (a), 20 (b) and 21 (c) in Hamlyn dataset. Blue lines are the batch optimization and red lines are the nodes decoupled optimization. Level I and Level II cannot be shown separately due to time consumption of Level II is extremely low.	82
5.10	Optimizing nodes comparisons in first level computation of model 6 (a), 20 (b) and 21 (c) in Hamlyn dataset. The red lines are the result of the decoupled optimization strategy while the blue lines are the original batch strategy.	83
5.11	Mean average error of 3D reconstruction (mm). From the first to last are Hamlyn dataset with ground truth, the synthetic heart, synthetic left kidney and synthetic stomach. Please refer to video for the synthetic data.	84
6.1	One step camera movement. Camera moves from \mathbf{p} to $\hat{\mathbf{p}}$. The movement is a mixture of camera transformation and deformation by ED node \mathbf{g} . The red line are the connecting edge from node \mathbf{g} to other nodes.	91
6.2	A typical feature deforming example. The ellipse deforms periodically depicted in ‘I, II and III’. The region within arrows are the visible region. The leftmost feature is not observed in phase ‘II’ and ‘III’. The rightmost feature is not observed in phase ‘I’.	99
6.3	A simple example of 2 steps camera movement. Different from SLAM in rigid scenario, the feature \mathbf{f} deforms in the space in position \mathbf{f}^1 , \mathbf{f}^2 and \mathbf{f}^3	100
6.4	The two figures is an example of Monte Carlo simulation. Display area is illustrated from different directions for visualization.	102
6.5	(a), (b) and (c) shows the camera moves randomly inside a deformable organ (Heart, stomach and liver). Red curves are the trajectories. Blue dots are the positions of the features and the attached quiver is the corresponding moving direction of each feature. Quiver only shows one step. Please refer to our video for the whole process.	103
6.6	Estimation errors of static SLAM, FEM, ED graph and the proposed time-series SLAM. Row 1 to 3 are the tests on scenarios of heart, liver and left lung. Column 1 to 3 are the RMSE of camera position X, camera position Y and camera heading.	103
6.7	Estimation errors of static SLAM, FEM, ED graph and the proposed time-series SLAM. Row 1 to 3 are the tests on scenarios of heart, liver and left lung. Column 1 to 3 are the RMSE of camera position X, camera position Y and camera heading.	104

6.8 Ground truth dataset from Hamlyn center. 106

List of Tables

3.1	Accuracy comparison between the proposed DFF approach and the back-projection approach (mm). Each value is calculated by averaging all the points of all the frames.	43
6.1	Pose and feature errors in Monte Carlo simulations.	96
6.2	Pose and feature errors of heart, stomach and lung.	97
6.3	Feature estimation accuracies (m) in three models. All the simulation noises (invariances) are set to be 0.1 m.	106

Acronyms & Abbreviations

CPU	Central Processing Unit
CT	Computed Tomography
CUDA	Compute Unified Device Architecture
DFD	Distance Field Function
ED	Embedded Deformation
EKF	Extended Kalman Filter
ELAS	Efficient Large-scale Stereo
EM sensor	Electromagnetic sensor
FEM	Finite Element Method
FIM	Fisher Information Matrix
GPGPU	General-Purpose computing on Graphics Processing Units
GPU	Graphical Processing Unit
ICP	Iterative Closest Point
MIS	Minimally Invasive Surgery
NRSfM	Non Rigid Structure from Motion
PI	Points Irrelevant

PR	Points Relevant
RANSAC	RANdom SAmples consensus
RMSE	Root-Mean-Square Error
SfM	Structure from Motion
SfS	Shape from Shading
SfT	Structure from Template
SIFT	Scale-invariant feature transform
SLAM	Simultaneous Localization And Mapping
SURF	Speeded Up Robust Feature
TSDF	Truncated Signed Distance Function
TSDW	Truncated Signed Distance Weight

Nomenclature

General Notations

O	Centers of the left image.
O'	Centers of the right image.
p	Pixel of the projected point on left image.
p'	Pixel of the projected point on right image.
u	First coordinate of 2D pixel.
v	Second coordinate of 2D pixel.
\mathbf{K}	Camera intrinsic matrix.
$\alpha(u, v)$	Surface albedo.
$\hat{I}(u, v)$	The measured pixel intensity at pixel (u, v) .
$O(\cdot)$	Computational complexity.
Ω	Linear elastic solid parameter.
λ and \mathbf{G}	Lam'e parameters that define the material elastic properties.
\mathcal{U}	Poisson's ratio.
\mathcal{E}	Young's modulus.
$a_{i,jj}$ and $a_{j,ij}$	Displacement vectors share the same edge.
\mathbf{g}_j	Position of node j .
\mathbf{A}_j	Affine matrix of node j .
\mathbf{v}	Position of a point.
$\tilde{\mathbf{v}}$	Target deformed vertex of \mathbf{v} .
\mathbf{R}_c	Global rotation of the scope.
\mathbf{T}_c	Global translation of the scope.

$w(\mathbf{v})$	The quantified weight for transforming \mathbf{v} exerted by each related ED node.
d_{max}	The maximum distance of the vertex to $k + 1$ nearest ED node.
m	The number of ED nodes.
$\mathbf{c}_1, \mathbf{c}_2$ and \mathbf{c}_3	The column vectors of \mathbf{A} .
E_{rot}	The affine matrix close to $SO(3)$.
E_{reg}	The sum the transformation errors from each ED node.
α_{jk}	The overlap influence of the two ED nodes.
$\mathbb{N}(j)$	The set of neighboring node to node \mathbf{g}_j .
$\mathcal{F}(\cdot)$	A general function defining a point to target distance.
$\mathcal{D}(\cdot)$	The corresponding voxel value recorded in DFF.
E_{data}	The sum distance error.
\mathbb{L}	The set for all the visible points.
\mathbb{D}	Depth scan.
$\Gamma(\cdot)$	Lift 2D depth pixel up to 3D point.
$\mathbb{H}(\cdot)$	Lift 2D normal pixel up to 3D normal.
$P(\mathbf{v})$	The projective function projecting 3D vertex to 2D pixel.
ϵ_d	Threshold for the distance.
ϵ_n	Threshold for the angle.
$\tilde{\mathbf{V}}_i$	The 3D points of current frame from last frame of ORB features.
\mathbf{V}_i	The 3D points of the deformed points from last frame of ORB features.
$\omega(\mathbf{v}_i)$	Weight of model point \mathbf{v}_i .
\mathbf{C}_i	Color of model point.
$d_{min}(\tilde{\mathbf{v}}_i)$	The minimum distance of model point.
$\tilde{\mathbf{v}}_i$	to its corresponding nodes.
ϵ	The average grid size of nodes.
$\tilde{\mathbf{v}}_i _z$	The value of point $\tilde{\mathbf{v}}_i$ on the z direction.
t_i	The time stamp of vertex \mathbf{v}_i .
\mathbf{n}_i	The normal of vertex \mathbf{v}_i .
\mathbf{s}_i	The boolean variable stability of vertex \mathbf{v}_i .

E_{corr}	Sum errors of distances between deformed key points and target key points.
E_r	The $SO(3)$ distance between estimated orientation and intilial orientation.
E_p	The Euclidean distance between estimated position and intilial position.
ϵ_{dv}	The threshold for extracting visible points based on distance.
ϵ_{nv}	The threshold for extracting visible points based on angle.
\mathbf{R}_c^n	Estimated rotation in step n .
\mathbf{T}_c^n	Estimated translation in step n .
$\tilde{\mathbf{R}}_c^n$	Rotation in step n from ORB-SLAM.
$\tilde{\mathbf{T}}_c^n$	Translation in step n from ORB-SLAM.
\mathbb{V}^n	Set of visible points in step n .
\mathbb{D}^n	Observed depth in step n .
\mathbb{P}^n	Fused point set.
ϵ_{df}	The threshold for fusing points based on distance.
ϵ_{nf}	The threshold for fusing visible points based on angle.
$\tilde{\mathbf{v}}_i^n _z$	The value of deformed point $\tilde{\mathbf{v}}_i^n$ in the z direction.
$\tilde{\mathbf{n}}_i^n$	The deformed normal of \mathbf{n}_i^n .
ω_{max}	The maximum weight for each model point.
\mathbf{P}	A group of predefined key source points.
$\tilde{\mathbf{P}}$	Deformed key points set of \mathbf{P} .
\otimes	The Kronecker product.
$\ \cdot\ _F^2$	The Frobenius norm.
X_i	The state vector is denoted as .
$skew(\cdot)$	The skew symmetric operator.
\mathbf{I}	A 3 by 3 identity matrix.
\odot	The Hadamard product.
\mathcal{H}_{ed}	Hessian matrix of ED formulation.
\mathbf{H}_1	One submatrix of \mathcal{H}_{ed} .
\mathbf{H}_2	One submatrix of \mathcal{H}_{ed} .

\mathbf{f}^n	A feature position in step n .
N	The number of features.
F	The number of steps.
\mathbf{B}	The combination of all valid features.
E_{obs}	The sum error of robot to feature observations.
\mathbf{m}_i^j	The observation from robot to location of feature i in step j .
$\mathcal{F}(\cdot)$	The estimated observation from robot pose to feature position.
E_f	The error between current feature and its estimation from historical locations.
E_{ini}	The initial robot pose keeps static in the period size t .
\ominus	Inverse retraction defining $SO(3)$ distance.
\mathbf{c}	The coefficient matrix .