

UNIVERSITY OF TECHNOLOGY SYDNEY  
Faculty of Engineering and Information Technology

# **Active Perception for Deformable 3D Pointclouds**

by

**Behnam Maleki**

A THESIS SUBMITTED  
IN PARTIAL FULFILLMENT OF THE  
REQUIREMENTS FOR THE DEGREE

**Doctor of Philosophy**

Centre for Autonomous Systems

December 2019

## Certificate of Original Authorship

I, Behnam Maleki declare that this thesis, is submitted in fulfilment of the requirements for the award of Doctor of Philosophy, in the School of Mechanical and Mechatronic Engineering/Faculty of Engineering and IT at the University of Technology Sydney.

This thesis is wholly my own work unless otherwise referenced or acknowledged. In addition, I certify that all information sources and literature used are indicated in the thesis.

This document has not been submitted for qualifications at any other academic institution.

This research is supported by the Australian Government Research Training Program.

Signature:

Production Note:  
Signature removed prior to publication.

Date: 17/12/2019

## ABSTRACT

Supplying precise and comprehensive representation of an object by assembly of pointclouds ultimately assists a robot in enhancing the reliability of its perception. An efficient data acquisition approach is to steer a depth sensor in 3D space actively to be positioned in the best (*optimal*) viewpoints to scan the desirable parts of the object and then align (*register*) and integrate the captured scans effectively and seamlessly to reconstruct a 3D model with high fidelity.

As the first contribution, we propose an optimization on a manifold approach to find the optimal position and orientation (*pose*) of a depth sensor in continuous 3D space. It has been demonstrated previously that precise measurement by a depth sensor is achieved when it is gazing at the object perpendicularly. Accordingly, the proposed terms of the objective function are to align the main axis of the depth sensor towards parts of interest while also prioritising areas with higher task-relevant information, such as curvature. The resulting poses achieved by this method conform to numerical and visual evaluations on several objects with a significantly less computation compared to the-state-of-the-art .

Reconstructing objects with high fidelity necessitates dealing with a variety of scenarios which can be differentiated in terms of temporal configurations and articulation of objects in the scene, namely rigidity or non-rigidity. Arguably the most challenging scenario is where a single depth sensor is scanning a texture-less object that is deforming non-rigidly. Under these conditions, apart from the computational overhead, most of the mesh reconstruction methods fail to yield satisfactory results. Moreover, there is not sufficient visual features on the surface to be extracted for correspondence.

Given these limitations, this thesis, as the second contribution, proposes a non-rigid registration for mesh-free and color-free pointclouds based on the *soft partitioning* concept. The soft patches (partitions) as the features are, then, equipped with local descriptors to provide a metric for association. Assuming that the global defor-

mation of the object is the aggregation of local rigid transformations, this association is refined by measuring the deviation of each potentially corresponding soft-patch and its neighborhood from a rigidity metric defined by the As-Rigid-As-Possible algorithm. The established local correspondences are assigned with transformations that are subsequently propagated to the nearby points. Experimental results demonstrate the capabilities of this framework in handling large deformations and highly articulated objects.

Fusing the aligned pointclouds, a 3D model of the targeted object is incrementally developed, and this model, coupled with the current scan, contributes to a formulation for selecting the next region of interest leading to the next optimal viewpoint. Unlike the conventional approaches regarding deformable objects (which take a great model where the extent of the object is seen and then it deforms), our proposed pipeline explores beyond the bounds of the current acquired frame and reconstructed model and continuously evolves it leveraging an exploration and exploitation strategy. The application of the devised framework for reconstruction is demonstrated on rigid and non-rigid objects demonstrating high fidelity to the original shape.

Dedication

To my family.

## Acknowledgements

During my Ph.D. candidature, I experienced lots of ups and downs; it kicked off a bit tough until I landed in the safe hands of a superb team. I believe that after working three years intimately with your PhD supervisor, you start gradually resemble them as your role model consciously or subconsciously; And I was lucky that I encountered the one who I ever liked to work with, Alen. I was lucky that I met two game-changers during this time, first my amazing friend, Kasra, who directed me to CAS, and then my main supervisor, Alen who not only changed my attitude towards research but taught me how to think better. During my candidature, I enjoyed learning from an exceptional person who boosted my standards by her exclusive constructive strategy. During this time, I was lucky to work with a perfect team of the good cop (Alen), the bad cop (Teresa), the prison guard (my honest and caring buddy Raphael)! Also, I am appreciative of my CAS colleagues who helped me a lot (to-be-dad-soon) Alex, Laki, Phil, Fred, Cedric, Mahdi, Julien, and excellent academics, Shoudong, Liang, Jaime, and David Eager. I also wish to thank Stuart Perry, who facilitated my transition to CAS.

I want to express my deepest gratitude to my beloved family, my kind-hearted mom, Mahnaz, and my fantastic dad, Saeid, who supported me unconditionally and always believed in me, my amazing brothers Behkam and Behdad. I really appreciate the supports of my in-laws, Mina, Heshmatollah, Pedram, Kimia, Tjasa, and little Kaia!

Last but not least, to one who the words are not enough to express my feelings to her. All this journey was so favorable and enjoyable thanks to my beloved spiritual wife, Hasti, who not only supported me endlessly but also inspired me by her solid character and unconditional love.

Behnam Maleki, Sydney, Australia, 2019.

## List of Publications

- I. Maleki, B., Ebrahimnezhad, H., Xu, M. and He, X., 2015, August. Hand gesture recognition for a virtual mouse application using geometric feature of finger's trajectories. In Proceedings of the 7th International Conference on Internet Multimedia Computing and Service (p. 73). ACM. (Selected as the Best Student Paper)
- II. Maleki, B., Alempijevic, A. and Vidal-Calleja, T., 2019, January. Continuous Optimization Framework for Depth Sensor Viewpoint Selection. In Workshop on the Algorithmic Foundations of Robotics.
- III. Maleki, B., Falque R., Vidal-Calleja, T. and Alempijevic, A., 2020, SPaM: Soft Patch Matching for Non-rigid Point Cloud Registration, Submitted to IEEE International Conference on Robotics and Automation 2020.

# Contents

Certificate	ii
Abstract	iii
Dedication	v
Acknowledgments	vi
List of Publications	vii
List of Figures	xiii
List of Tables	xvi
<b>1 Introduction</b>	<b>1</b>
1.1 Motivation . . . . .	4
1.2 Objectives . . . . .	6
1.3 Contributions . . . . .	6
1.4 Structure of the Thesis . . . . .	7
1.5 General Notation . . . . .	8
<b>2 Literature Review</b>	<b>10</b>
2.1 Background and Foundations . . . . .	10
2.1.1 Three-Dimensional Geometry . . . . .	10
2.1.1.1 Pose: Position and Orientation . . . . .	11
2.1.1.2 Lie Transformation Groups . . . . .	12
2.1.2 Surface Discretization by Point Cloud . . . . .	13



2.1.3	From Points to Clouds . . . . .	13
2.1.3.1	Depth Sensors . . . . .	14
2.1.3.2	Intrinsic and Extrinsic Parameters . . . . .	15
2.1.4	Geometrical Properties of Point Cloud . . . . .	17
2.1.4.1	Vertex Normals . . . . .	17
2.1.4.2	Vertex Curvature . . . . .	18
2.1.4.3	3D Shape Descriptor . . . . .	20
2.1.5	Optimization on Manifold . . . . .	21
2.2	Related Works . . . . .	24
2.2.1	Optimal Viewpoint of Sensor . . . . .	24
2.2.2	Non-rigid Point Cloud Registration . . . . .	27
2.2.3	3D Shape Reconstruction and Active Vision . . . . .	34
2.2.4	Summary . . . . .	44
<b>3</b>	<b>Continuous Viewpoint Optimization</b>	<b>46</b>
3.1	Overview . . . . .	46
3.2	Optimal Viewpoint Formulation . . . . .	48
3.2.1	Angular Constraints . . . . .	50
3.2.2	Distance Constraint . . . . .	52
3.2.3	Objective Function . . . . .	52
3.2.4	Camera Field of View Constraint . . . . .	53
3.3	Continuous Optimization on Manifold . . . . .	54
3.4	Experimental Results and Evaluation . . . . .	56
3.5	Summary . . . . .	64
<b>4</b>	<b>Non-rigid pointcloud Registration</b>	<b>66</b>

4.1	Overview . . . . .	66
4.2	Soft Patch Description . . . . .	69
4.2.1	Soft Partitioning . . . . .	69
4.2.2	Property Assignment . . . . .	70
4.2.2.1	SHOT Vector . . . . .	70
4.2.2.2	Directional-Preserving Neighbourhood . . . . .	72
4.3	Soft Patch Matching . . . . .	73
4.3.1	Patch Similarity . . . . .	73
4.3.1.1	SHOT Vector distance . . . . .	73
4.3.1.2	Directional-preserving neighborhood SHOT distance . . . . .	74
4.3.2	Similarity Matrix . . . . .	75
4.3.3	Robust Correspondence Selection . . . . .	76
4.3.3.1	Non-rigidity error based on ARAP . . . . .	76
4.3.3.2	Corresponding Patch Solver . . . . .	77
4.3.3.3	ARAP Optimization Initializer . . . . .	80
4.3.3.4	ARAP Correspondence Filter . . . . .	81
4.3.3.5	Similarity Matrix Updater . . . . .	82
4.3.3.6	Stopping Criteria . . . . .	83
4.3.4	Target Deformation . . . . .	83
4.3.4.1	Correspondence-Wise Transformations . . . . .	83
4.3.4.2	Transformations Propagation . . . . .	84
4.3.4.3	Weighting Scheme . . . . .	85
4.4	Results . . . . .	90
4.4.1	Quantitative Evaluation of Registration . . . . .	90

4.4.2	Real-Data Experiments . . . . .	94
4.5	List of Parameters . . . . .	97
4.6	Summary . . . . .	100
<b>5</b>	<b>Active Deformable Object Reconstruction</b>	<b>101</b>
5.1	Overview . . . . .	101
5.2	Pre-processing of Pointclouds . . . . .	103
5.2.1	Vertex Normal Correction . . . . .	104
5.2.1.1	Intersection of Main-axis Ray and Pointcloud . . . . .	106
5.2.1.2	Algorithm for Vertex Normal Correction . . . . .	108
5.3	Next Best Viewpoint . . . . .	108
5.3.1	Boundary Band Estimation . . . . .	111
5.3.2	Next Region of Interest Selection . . . . .	112
5.3.2.1	Overlapping Term . . . . .	112
5.3.2.2	Curvature Term . . . . .	113
5.3.2.3	Model Closing Term . . . . .	113
5.3.3	NRoI Estimation . . . . .	116
5.3.4	Sensor Pose Estimation . . . . .	117
5.4	3D Object Reconstruction . . . . .	117
5.4.1	Registering Scans of object . . . . .	118
5.4.1.1	Reconstructed model in FoV of current sensor pose . . . . .	119
5.4.2	Fusion . . . . .	123
5.4.2.1	Determining Vertex Status: Fusible or Newly-emerged	126
5.4.2.2	Displacement of Newly-Emerged Point . . . . .	131
5.4.2.3	Normal Estimation of Fused Model . . . . .	139

5.5	Scanning Termination Criterion . . . . .	142
5.6	Results . . . . .	142
5.6.1	Rigid Experiment . . . . .	144
5.6.2	Non-rigid Experiment . . . . .	149
5.7	Summary . . . . .	156
<b>6</b>	<b>Conclusions</b>	<b>158</b>
6.1	Summary . . . . .	158
6.2	Future Work . . . . .	161
6.2.1	Including Uncertainty for Optimal Viewpoint . . . . .	161
6.2.2	Considering the sensor Restrictions in Viewpoint Optimization	161
6.2.3	Using Geodesic Distance . . . . .	161
6.2.4	A Metric for Local Density and Information of Point Cloud . .	161
6.2.5	Optimal Search Radius for the Proposed Pipelines . . . . .	162
6.2.6	Object Segmentation by Displacement Trajectory . . . . .	162
6.2.7	Reconstructing From Bigger and Faster Deformations . . . . .	163
	<b>Appendices</b>	<b>164</b>
A.1	Lemma . . . . .	164
	<b>Bibliography</b>	<b>166</b>

# List of Figures

1.1	The structure of proposed pipeline . . . . .	4
2.1	Two examples of pointcloud . . . . .	13
2.2	A taxonomy of 3D shape acquisition techniques . . . . .	14
2.3	Pipeline of a forward projection . . . . .	16
2.4	A planar curve embedded in the surface . . . . .	19
2.5	Signatures and histograms . . . . .	21
2.6	Signature structure for SHOT . . . . .	22
2.7	2-Sphere and an element on it . . . . .	23
2.8	Tangent space of a manifold . . . . .	23
2.9	Registration results for the CPD, RPM and ICP algorithms . . . . .	30
2.10	The autonomous six-DOF eye-in-hand system . . . . .	40
3.1	Constraints of objective function . . . . .	50
3.2	System Flowchart . . . . .	51
3.3	Interim poses of sensor during optimisation . . . . .	59
3.4	Optimal poses based on different information of interest . . . . .	60
3.5	Color map of cost function in different poses . . . . .	61
3.6	Optimal pose for Igea dataset . . . . .	62

4.1	The flowchart of devised framework . . . . .	68
4.2	Example of neighborhood map of patches . . . . .	75
4.3	Source and target pointcloud of Igea . . . . .	87
4.4	The established correspondences for Igea . . . . .	88
4.5	Comparing application of ED and AI weighting schemes . . . . .	89
4.6	A quantitative evaluation . . . . .	91
4.7	Six color-coded articulations of Human dataset captured by raytracing the associated original pointclouds. . . . .	92
4.8	Deviations from ground-truth in 5 experiments . . . . .	93
4.9	RMSE of Deviation of corresponding points . . . . .	93
4.10	Registration of two frames of the Umbrella dataset . . . . .	95
4.11	Color coded correspondences of the Sunflower dataset . . . . .	95
4.12	Registering two pointclouds of the Boxing dataset . . . . .	96
4.13	Qualitative comparison of our method (SPaM) with Surfel Warp . . . . .	98
4.14	An experiment on the Minion dataset . . . . .	99
5.1	The forward reconstruction pipeline . . . . .	102
5.2	The backward reconstruction pipeline . . . . .	103
5.3	The flipped normal problem . . . . .	105
5.4	The boundary band on Igea dataset . . . . .	111
5.5	Illustration of Model Closing term . . . . .	115
5.6	The reconstructed leg and the new scan . . . . .	118
5.7	The sensor pose and the reconstructed model . . . . .	119
5.8	The reconstructed model in the FoV of current sensor pose . . . . .	122
5.9	The correspondence map between two pointclouds of leg dataset . . . . .	123

5.10	The double surface problem on deforming leg . . . . .	125
5.11	A cylinder fit on the aggregated pointcloud . . . . .	126
5.12	The tranformed source point and its neighbours to origin . . . . .	129
5.13	The transformed source point and its fusible neighbors . . . . .	130
5.14	The fusible points and non-fusible points . . . . .	132
5.15	This displacement of selected points . . . . .	133
5.16	Different point groups in fusion process . . . . .	137
5.17	Three different groups of fused pointcloud . . . . .	138
5.18	The pointcloud of fused leg dataset . . . . .	138
5.19	Fusion of two plane pointclouds . . . . .	140
5.20	Fusion of two cap pointclouds . . . . .	141
5.21	The Igea model used in rigid experiments . . . . .	143
5.22	The optimal poses for Igea dataset . . . . .	145
5.23	The reconstructed model of Igea via the proposed pipeline . . . . .	146
5.24	The process of NRoI selection in iteration 4 . . . . .	148
5.25	The process of NRoI selection in iteration 10 . . . . .	148
5.26	The ground truth of Igea supserimposed onto the reconstrcuted model	149
5.27	The deviation error of the reconstructed model and the ground truth	150
5.28	The articulation of Leg dataset . . . . .	150
5.29	The boundary band for pose 5 . . . . .	151
5.30	The recosntrcuted model from two views . . . . .	152
5.31	The optimal poses of sensor during reconstrcution . . . . .	153
5.32	Superimposing the articulations into the model . . . . .	155
5.33	The difference between geodesic and euclidean distance . . . . .	156

## List of Tables

3.1	The details of the experiments on the cow back . . . . .	57
-----	--	----