THE UNIVERSITY of ADELAIDE

Robust Estimation in Computer Vision: Optimisation Methods and Applications

by

Trung Thanh Pham

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

in the

Faculty of Engineering, Computer and Mathematical Sciences
School of Computer Science

May 2014

This dissertation is dedicated to my parents,

my brothers and my sisters
for their unconditional love and endless support.

Contents

Contents i

List of Figures v

List of Tables vii

List of Algorithms viii

Abstract ix

Declaration x

Publications xi

Acknowledgements xii

1 Introduction 1
- 1.1 Robust Estimation in Vision & Challenges . . . 3
- 1.2 Research Contributions . . . 6
- 1.3 Thesis Outline . . . 7

2 Optimisation Algorithms for Robust Estimation 9
- 2.1 Robust Fitting Criterion . . . 10
  - 2.1.1 Goodness of Fit . . . 10
  - 2.1.2 Model Complexity . . . 10
  - 2.1.3 Other Requirements . . . 11
  - 2.1.4 Objective Functions . . . 11
- 2.2 Optimisation Algorithms for Model Selection . . . 12
  - 2.2.1 Tabu Search . . . 13
  - 2.2.2 Branch and Bound . . . 15
  - 2.2.3 Linear Programming Relaxation . . . 17
  - 2.2.4 Message Passing . . . 18
  - 2.2.5 Graph-cuts and Expansion Move Algorithm . . . 19
  - 2.2.6 Quadratic Programming . . . 21
  - 2.2.7 Discussion . . . 23

2.3 Random Hypothesis Generation . . . 24
2.3.1 Elemental Subsets . . . 24
2.3.2 Random Sampling . . . 25
2.3.3 Guided Sampling . . . 27
2.3.4 Discussion . . . 33
2.4 Summary . . . 34

3 Simultaneous Hypothesis Sampling and Model Selection 35
3.1 MCMC Methods . . . 36
3.2 Model Fitting with Reversible Jump MCMC . . . 38
3.2.1 Objective Function . . . 38
3.2.2 Simulated Annealing Optimisation . . . 38
3.2.3 Reversible Jump MCMC . . . 40
3.2.3.1 Birth and Death Moves . . . 41
3.2.3.2 Local Update . . . 41
3.3 Adaptive MCMC Proposal . . . 42
3.3.1 The Multi-GS Algorithm . . . 42
3.3.2 Adaptive Proposal using Multi-GS . . . 44
3.3.3 Convergence . . . 44
3.3.3.1 Invariance . . . 45
3.3.3.2 Diminishing Adaptation . . . 45
3.3.3.3 Uniform Ergodicity . . . 47
3.4 Experiments . . . 47
3.4.1 Synthetic data . . . 47
3.4.2 Real data . . . 50
3.5 Summary . . . 52

4 Rapid Hypothesis Sampling using Random Cluster Models 54
4.1 A Graph Cuts based Simulated Annealing Approach . . . 56
4.1.1 Energy Function . . . 57
4.1.2 Energy Optimisation . . . 58
4.1.2.1 Model Selection Optimisation . . . 58
4.1.2.2 Label Optimisation . . . 59
4.1.2.3 Local Refinement of Model Parameters . . . 60
4.2 The Random Cluster Model . . . 60
4.2.1 Sampling Larger-than-minimal Subsets . . . 61
4.2.2 Hypothesis Generation using the RCM . . . 62
4.2.3 Validity of Spatial Smoothness Assumption . . . 62
4.2.4 Online Learning of The RCM . . . 63
4.2.5 The Swendsen-Wang Method for Label Optimisation . . . 65
4.3 Experiments . . . 66
4.3.1 Larger-than-minimal Subsets vs. Minimal Subsets . . . 66
4.3.2 The RCM vs. Local Refinement and Global Exploration . . . 67
4.4 Summary . . . 69

5 Robust Two-view Geometry Estimation and Applications 71
5.1 Two-view Geometry . . . 71

5.1.1 Fundamental Matrix 72
5.1.2 Multiple Fundamental Matrices for Dynamic Scenes 73
5.1.3 Homography Matrix 73
5.1.4 Multiple Homography Matrices 74
5.1.5 Estimating Image-to-image Transformations 74
5.2 Two-view Motion Segmentation 75
5.2.1 Problem Statement 75
5.2.2 Estimation Methods 76
5.2.3 Experimental Design 76
5.2.3.1 Datasets 76
5.2.3.2 Benchmarking 77
5.2.3.3 Error Metric 77
5.2.4 Results 78
5.3 Two-view Multi-homography Detection 79
5.3.1 Problem Statement 79
5.3.2 Datasets 80
5.3.3 Results 80
5.4 Augmented Reality using Homography Mappings 82
5.4.1 Offline Preprocessing 84
5.4.2 Online Multiple Homography Detection 85
5.4.3 Plane Matching and Image Augmentation 85
5.5 Summary 86

6 Fitting with Geometric Priors 91
6.1 High-Level Priors in Vision 92
6.2 Robust Fitting with Geometric Priors 94
6.2.1 Energy Function 94
6.2.2 Energy Normalisation 96
6.2.3 Energy Minimisation 96
6.2.4 Interacting Geometric Costs vs. Category Costs 100
6.3 Experiments and Applications 101
6.3.1 Two-view Multi-homography Estimation 101
6.3.2 Vanishing Point Detection in a Manhattan World. 104
6.3.3 Circular Object Detection in Images 110
6.3.3.1 Fitting Circles 111
6.3.3.2 Fitting Disks 112
6.3.4 Planar Surface Reconstruction from 3D Point Clouds 114
6.3.5 Tuning Parameter Sensitivity 115
6.4 Summary 116

7 Conclusions and Future Work 119
7.1 Summary of Contributions 119
7.2 Future Research Directions 121

Bibliography **123**

List of Figures

1.1 Typical computer vision applications which rely on robust estimation. . . 2
1.2 An example of multiple circle fitting data. 4

2.1 An illustration of the over-fitting issue. 11
2.2 Robust parameter estimation using a global optimisation approach. 13
2.3 A binomial solution tree for Branch-and-Bound. 16
2.4 Factor graph construction for multi-model fitting 19
2.5 The exponential increase of the number of required hypotheses against the outlier ratio. 26
2.6 The exponential increase of the number of required hypotheses against the size of the minimal subsets. 26
2.7 Different types of "all-inlier" minimal subsets for the line and circle fitting. 28

3.1 Global optimisation using a simulated annealing approach. 39
3.2 2D point cloud generated from 5 lines. 48
3.3 The evolution of data similarities. 49
3.4 The evolution of ARJMC as the number of iterations increases. 49
3.5 The evolution of the objective function value as ARJMC progresses. . . . 49
3.6 The evolution of the estimated number of structures as ARJMC progresses. 50
3.7 Performance comparison between ARJMC and FLOSS on a motion segmentation problem. 51

4.1 Line hypotheses fitted on a minimal and larger-than-minimal subset. . . . 55
4.2 A conceptual illustration of the RCM method on a multiple line fitting problem. 56
4.3 Demonstrating the RCM hypothesis generator on a multi-homography fitting problem. 56
4.4 Qualitative comparison of hypotheses progressively generated using RANSAC, Multi-GS and RCM on a multiple line fitting problem. 63
4.5 Matrix of edge probabilities for the multiple line fitting task. 65
4.6 Multi-homography detection on Raglan Castle. 67
4.7 Evolution of objective value by Algorithm 4.1 using different sampling schemes. 68
4.8 The probability of hitting all-inlier subsets for different sampling methods as Algorithm 4.1 progresses. 68
4.9 Evolution of objective value by Algorithm 4.1 using different sampling schemes with local refinement and exploration. 69

5.1 Two-view geometry and fundamental matrix. 72

5.2 Two-view geometry and homography matrix. . . . 73
5.3 Correspondence extraction between two views. . . . 75
5.4 Comparing efficiency on motion segmentation on a subset of the data. . . 80
5.5 Comparing efficiency on homography detection on a subset of the data. . 82
5.6 Using our RCMSA fitting method in an augmented reality application. . . 83
5.7 Preprocessing for an automatic augmented reality application. . . . 84
5.8 Associating planes and estimated homographies. . . . 85
5.9 Qualitative motion segmentation on *breadcartoychips* dataset. . . . 87
5.10 Qualitative motion segmentation on *cubebreadtoychips* dataset. . . . 87
5.11 Qualitative motion segmentation on *breadtoycar* dataset. . . . 88
5.12 Qualitative motion segmentation using *biscuitbookbox* dataset. . . . 88
5.13 Qualitative multi-homography detection using *Johnsona* dataset. . . . 89
5.14 Qualitative multi-homography detection using *Merton College 1* dataset. 89
5.15 Qualitative multi-homography detection using *Unionhouse* dataset. . . . 90
5.16 Two examples where RCMSA fails to recover all the structures on multi-homography detection. . . . 90

6.1 Two examples of computer vision applications where geometric priors should be effectively exploited to produce plausible results. . . . 92
6.2 Graph construction for the expansion move energy. . . . 100
6.3 Multiple homography estimation with geometric constraints. . . . 104
6.4 Approximated residual from an edge to a vanishing point. . . . 106
6.5 Quantitative vanishing point detection results using manually extracted lines. . . . 107
6.6 Quantitative vanishing point detection results using automatically extracted lines. . . . 108
6.7 Qualitative vanishing point detection results using automatically extracted lines. . . . 109
6.8 Circular object detection using an image of diatom. . . . 111
6.9 Circular object detection using an image of cells. . . . 112
6.10 Counting flamingoes via disk model fitting. . . . 113
6.11 Counting cells via disk model fitting. . . . 114
6.12 Fitting planes to *Pozzoveggiani* 3D point cloud. . . . 116
6.13 Fitting planes to *Piazza Dante* point cloud. . . . 117
6.14 Fitting planes to *Merton College 3* 3D point cloud. . . . 117

List of Tables

2.1 Different local-and-global search strategies for hypothesis sampling. 29

5.1 Summary statistics (median) over 50 repetitions on fundamental matrices detection. 79

5.2 Summary statistics (median) over 50 repetitions on multi-homography detection. 81

6.1 Quantitative results on two-view multi-homography estimation with geometric constraints. 105

List of Algorithms

2.1 Global optimisation based multi-structure robust estimation 12

2.2 Expansion move algorithm 21

2.3 RANSAC 25

2.4 The guided-MLESAC sampling algorithm 28

2.5 The LO-RANSAC sampling algorithm 30

2.6 The NAPSAC sampling algorithm 31

2.7 The Multi-GS guided sampling algorithm 32

3.1 Metropolis-Hastings algorithm. 37

3.2 Monte Carlo Markov Chain simulated annealing for geometric fitting. . . . 40

3.3 Reversible jump mixture of kernels MCMC 40

4.1 Graph cuts based annealing optimisation for geometric fitting. 59

4.2 Updating Edge Probabilities in RCM 64

6.1 Iterative hypothesis sampling and fitting for homography estimation. . . . 103

Abstract

Robust parameter estimation is an important area in computer vision that underpins many practical applications. Typically, the task is to estimate a generic model from unstructured observations, where the model and observed data may vary depending on the specific applications. In most cases, computer vision data inherently contains noisy measurements, multiple instances (structures) of a model, and outliers (i.e., points that do not belong to any structures). Unfortunately, standard techniques such as Least Squares (LS), Least Median Squares (LMS) are not robust to such kind of data.

Over the past decades, much research effort in computer vision has been devoted to proposing more robust and efficient estimators. Among those, the estimators based on global optimisation have attracted considerable attention and increasingly showed promising results. However these optimisation based methods still are faced with a number of issues. First, for tractability these robust techniques optimise robust objective functions over a collection of randomly sampled hypotheses using combinatorial methods. The trouble is that the adequacy of the hypothesis set could not be asserted prior to the optimisation, so the overall estimation could be misleading. In addition, the process of randomly sampling the hypothesis set is very time-consuming, especially for high-order models and complex data, thus generally decreasing the fitting efficiency. Moreover, to ease the optimisation, outliers are often assumed to distribute uniformly in the data space, and measurement noise is assumed to approximately follow a Gaussian distribution. Unfortunately, such assumptions are not always valid in practice.

The research conducted in this thesis follows the global optimisation approach, and makes three distinct contributions to the robust estimation field. First, we propose a novel fitting approach that simultaneously samples hypotheses and optimises the robust objective functions, such that the under- or over- hypothesis sampling issue can be avoided. In effect, our fitting approach can effectively minimise the wastage of the hypothesis sampling and objective optimisation. The second contribution is an unconventional sampling method based on Random Cluster Model (RCM) for rapidly generating accurate hypotheses. The RCM sampling method is effectively integrated into a continuous sampling-and-fitting framework to provide the superior fitting efficiency. Finally, the thesis offers a new robust estimation framework which seamlessly considers high-level geometric priors during the parameter estimation to enhance the robustness against non-uniform outliers and non-Gaussian noise. We validate the performance of the robust methods presented in this thesis on various computer vision applications ranging from estimating motions, planar homographies in image sequences to detecting geometric objects in images and 3D point clouds.

Declaration

I certify that this work contains no material which has been accepted for the award of any other degree or diploma in any university or other tertiary institution and, to the best of my knowledge and belief, contains no material previously published or written by another person, except where due reference has been made in the text. In addition, I certify that no part of this work will, in the future, be used in a submission for any other degree or diploma in any university or other tertiary institution without the prior approval of the University of Adelaide and where applicable, any partner institution responsible for the joint-award of this degree.

I give consent to this copy of my thesis when deposited in the University Library, being made available for loan and photocopying, subject to the provisions of the Copyright Act 1968.

I also give permission for the digital version of my thesis to be made available on the web, via the University's digital research repository, the Library catalogue and also through web search engines, unless permission has been granted by the University to restrict access for a period of time.

Signed: ______________________________

Date: ______________________________

Publications

In carrying out my postgraduate research, a number of papers were published, accepted or are currently under revision. The content of this thesis is partially based on those papers, as listed below:

1. **T. T. Pham**, T.-J. Chin, J. Yu, and D. Suter, Simultaneous Sampling and Multi-Structure Fitting with Adaptive Reversible Jump MCMC, Advances of Neural Information Processing Systems (NIPS), pages 540–548, 2011.

2. **T. T. Pham**, T.-J. Chin, J. Yu, and D. Suter, The Random Cluster Model for Robust Geometric Fitting, In Proceedings of Computer Vision and Pattern Recognition (CVPR), pages 710-717, 2012.

3. **T. T. Pham**, T.-J. Chin, J. Yu, and D. Suter, The Random Cluster Model for Robust Geometric Fitting, IEEE Transactions on Pattern Analysis and Machine Intelligence (TPAMI) (Accepted Dec 2013).

4. **T. T. Pham**, T.-J. Chin, K. Schindler and D. Suter, Interacting Geometric Priors for Robust Multi-Model Fitting, IEEE Transactions on Image Processing (TIP) (Under revision).

Acknowledgements

I would like to gratefully acknowledge all of the people whose support and encouragement have helped me finish my PhD journey.

First and foremost, I would like to express the deepest appreciation to my supervisor, Prof. David Suter, who has brought me a passion for research, taught me the critical thinking and scientific writing skills. He has been very supportive since the days I contacted him for a postgraduate study. I am especially grateful to him for his incredible patience in discussing my research questions, reviewing my technical papers and particularly this thesis. I believe that without his supervision and constant help this dissertation would not have been possible.

I owe sincere thankfulness to my advisor, Dr. Tat-Jun Chin, for his invaluable help and guidance through the rough road to finish this thesis. He has given me constructive discussions about my research problems, taught me to question and debate the research techniques and results. He has enlightened me the fundamentals of writing and presentation. Without him, this thesis would never have been carried out.

I also very much appreciate Prof. Wojciech Chojnacki, A. Prof. Frank Neumann, Prof. Konrad Schindler and Dr. Jin Yu for the fruitful discussions and collaborative work. My gratitude is extended to Prof. Ba-Ngu Vo for sharing his priceless research and life experiences during my stay at the University of Western Australia. Many thanks also go to my mentor, Dr. Ben Appleton, who has shown me the fundamental tradeoff between industry and academia, and also kindly supervised me during my internship at Google.

I am grateful to the University of Adelaide for providing me the scholarship to pursue my PhD, to the school of computer science and the Australia Centre of Visual Technologies (ACVT) for a great research environment. In addition, I would like to express my sincere gratitude to all members in the school of computer science, especially my colleagues Dr. Hoi Sim Wong, Dr. Sakrapee Paisitkriangkrai, Guosheng Lin, Quoc-Huy Tran, Alvaro Parra Bustos, Julio Hernandez Zaragoza, Alex Chicowski, and Chao Zhang, whom I have had chances to talk and discuss all matters of research and life.

Last but not at least, this thesis would not have been existed without the unconditional love and endless support of my family. I dedicate this dissertation to my parents, whose love and hard-working have grown me up; to my brothers and sisters who have been always by my side.

Trung T. Pham
May, 2014