Indoor Place Classification for Intelligent Mobile Systems

By

Lei Shi

Thesis submitted as a requirement for the degree of

Doctor of Philosophy

School of Electrical, Mechanical and Mechatronic Systems

Faculty of Engineering and Information Technology

University of Technology, Sydney (UTS)

June, 2013

Declaration of Authorship

I, Lei Shi, 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.

Signed:

Date:

Acknowledgements

I would like to take the opportunity to express my gratitude to all the people who have offered me help and support during the past three and a half years of my candidature. First and foremost, I would like to give my sincere thanks to my supervisor Dr. Sarath Kodagoda for his guidance, inspiration, numerous hours of discussion, and the opportunities he has presented me. From him I learned not only the knowledge but also the research methodology, which will benefit me throughout my career. I would also like to thank Prof. Gamini Dissanayake and Dr. Rami Khushaba for their inspiration, continuous support, and their direct contributions to my research.

I wish to acknowledge Dr. Oscar Martinez Mozos at the University of Lincoln, and Dr. Andrzej Pronobis at the Royal Institute of Technology who have been in the forefront of the research on spatial understanding for intelligent mobile systems. I benefit greatly from communicating with them, and I appreciate their contributions to the community including publications, codes and publicly accessible data sets.

I appreciate the strong support from Dr. Mark Schmidt at the École normale supérieure and Prof. Massimo Piccardi at UTS, for the discussions from the feasibility to the implementation details of the theoretical ideas I proposed. I also would like to convey special thanks to Prof. Dieter Fox at the University of Washington and Dr. Cyrill Stachniss at the University of Freiburg, who supportively confirmed the ground truth on the floor plans of their laboratories; to Dr. Dejan Pangercic at the Technische Universität München, who spent his precious time in solving my problems with deploying a software package; and to many other researchers, book authors and software owners who generously helped me with my various questions.

Many thanks go to Dr. Alen Alempijevic, Dr. Ravindra Ranasinghe, Dr. Stephan Sehestedt, Dr. Liang Zhao, Mr. Gibson Hu, Mr. Lakshitha Dantanarayana, and my many colleagues who gave me hands-on help on setting up the experimental environment to collect and process data; to Ms. Helen Chan at the UTS library who has been assisting me with utilising library services; and to the administrative staff at UTS for maintaining an organised, efficient and comfortable academic environment.

Additional thanks go to my parents and friends, who are always proud of me for my achievements, and supported and cared for me throughout these years.

I extend my gratitude to the Australian Commonwealth Government for the exemption of my tuition fee and for assisting with my general living costs through the APA scholarship; to the ARC Centre of Excellence for Autonomous Systems (CAS) for subsidising my living expenses; and to the CAS and the Faculty of Engineering and Information Technology (UTS FEIT) for sponsoring me in attending local and international conferences.

Finally, I would like to thank Ms. Monique Hohnberg for proofreading my thesis and some other publications.

Publications

Journal Papers:

1. **L. Shi**, S. Kodagoda, “Towards Generalization of Semi-supervised Place Classification over Generalised Voronoi Graph”, *Robotics and Autonomous Systems*, vol. 61, no. 8, pp. 785-796, Aug. 2013.

Conference Papers:

1. **L. Shi**, S. Kodagoda, and M. Piccardi, “Towards simultaneous place classification and object detection based on conditional random field with multiple cues”, in *Proc. IEEE/RSJ Int. Conf. Intell. Robots Syst. (IROS)*, Tokyo, 2013, pp. 2806 -2811.
2. **L. Shi**, S. Kodagoda, R. Khushaba and G. Dissanayake, “Application of CRF and SVM based Semi-supervised Learning for Semantic Labeling of Environments”, in *Proc. Int. Conf. Control Autom. Robot. Vis. (ICARCV)*, Guangzhou, 2012, pp. 835-840.
3. **L. Shi**, S. Kodagoda and G. Dissanayake, “Application of Semi-supervised Learning with Voronoi Graph for Place Classification”, in *Proc. IEEE/RSJ Int. Conf. Intell. Robots Syst. (IROS)*, Vilamoura, 2012, pp. 2991-2996.
4. R. N. Khushaba, **L. Shi**, and S. Kodagoda, "Time-Dependent Spectral Features for Limb Position Invariant Myoelectric Pattern Recognition", in *Proc. Int. Symp. Commun. Inf. Technol. (ISCIT)*, Gold Coast, 2012, pp. 1020-1025.
5. **L. Shi**, S. Kodagoda and R. Ranasinghe, “Fast Indoor Scene Classification Using 3D Point Clouds”, in *Proc. Australasian Conf. Robot. Autom. (ACRA)*, Melbourne, 2011.
6. **L. Shi**, S. Kodagoda and G. Dissanayake, “Laser Range Data Based Semantic Labeling of Places”, in *Proc. IEEE/RSJ Int. Conf. Intell. Robots Syst. (IROS)*, Taipei, 2010, pp. 5941-5946.

7. **L. Shi**, S. Kodagoda and G. Dissanayake, “Environment Classification and Semantic Grid Map Building Based on Laser Range Finder Data”, in *workshop IEEE/RSJ Int. Conf. Intell. Robots Syst. (IROS)*, Taipei, 2010.
8. **L. Shi**, S. Kodagoda and G. Dissanayake, “Multi-class Classification for Semantic Labeling of Places”, in *Proc. Int. Conf. Control Autom. Robot. Vis. (ICARCV)*, Singapore, 2010, pp. 2307-2312.
9. **L. Shi**, S. Kodagoda and G. Dissanayake, “Semantic Grid Map Building”, in *Proc. Australasian Conf. Robot. Autom. (ACRA)*, Brisbane, 2010.

Contents

1 Introduction 1

1.1 Background 1

1.2 Prospects and Problems 2

1.3 Motivations and Objectives 3

1.4 Principal Contributions 4

1.5 Organisation of this Thesis 5

2 Background and Literature Review 7

2.1 Place Classification and Related Research 7

2.1.1 Introduction 7

2.1.2 Applications, Requirements, and Benchmark 8

2.1.3 Sensory Modalities 11

2.1.4 Machine Learning in Place Classification 15

2.2 Supervised Learning 18

2.2.1 Introduction 18

2.2.2 Classifiers for Independent Prediction 21

2.2.2.1 Logistic Regression 21

2.2.2.2 Support Vector Machine 24

2.2.3 Graphical Models for Structured Prediction 29

2.2.3.1 Pairwise Conditional Random Field 29

2.2.3.2 Parameter Estimation and Inference 33

2.3 Semi-supervised Learning 34

2.4 Labels and Feature Manipulation 37

2.4.1 Labels: the Target Concepts 37

2.4.2 Feature Construction and Selection 37

2.5 Summary 41

3 Classification Scheme for Independent Prediction 42
3.1 Introduction 42
3.2 Feature Construction 44
3.2.1 Features from 2D Laser Range Data 44
3.2.2 Features from Kinect™ Data 47
3.3 Feature Selection 49
3.4 Experiments, Results and Discussion 51
3.4.1 2D Laser Range Data based Binary Classification 51
3.4.1.1 Experimental Setup 51
3.4.1.2 Manual Feature Selection 52
3.4.1.3 Place Classification in a Binary Scenario 54
3.4.1.4 Conclusion 55
3.4.2 2D and 3D Range Data based Multi-class Classification 56
3.4.2.1 Experimental Setup 57
3.4.2.2 DEFS Feature Selection 60
3.4.2.3 Place Classification in a Multi-class Scenario 63
3.4.2.4 Effect of Coverage 65
3.4.2.5 Conclusion 66
3.5 Summary 67
4 Classification Scheme for Structured Prediction 69
4.1 Introduction 69
4.2 Generalised Voronoi Graph 70
4.3 Conditional Random Field for Place Classification 71
4.4 Semi-supervised Learning 73
4.4.1 Co-training-like Semi-supervised Learning 74
4.4.2 CRF with Partial Labelled Data 76
4.5 Node and Edge Features 78
4.6 Experimental Setup 80

4.7 Results and Discussion 83
4.7.1 Fully Supervised Comprehensive Training 83
4.7.2 Fully Supervised Leave-one-out Training 83
4.7.3 Fully Supervised Leave-many-out Training 85
4.7.4 Semi-supervised Leave-many-out Training 87
4.7.5 Conclusion 91
4.8 Summary 92
5 Simultaneous Place Classification and Object Detection 93
5.1 Introduction 93
5.2 Candidate Models for SPCOD 94
5.3 High-level Features of an Image 98
5.4 Experimental Setup 100
5.5 Results and Discussion 103
5.5.1 Performance of Place Classification 103
5.5.2 Performance of Object Detection 104
5.5.3 Overall Performance of SPCOD 106
5.5.4 Object Localisation 107
5.5.5 Conclusion 109
5.6 Summary 110
6 Conclusions and Future Research 112
6.1 Thesis Summary 112
6.2 Thesis Conclusion 113
6.3 Recommendations for Future Research 115
Bibliography 117

List of Tables

Table 3.1: Performance of Different Feature Combinations 53
Table 3.2: Performance of the Logistic Regression Classifier 54
Table 3.3: Performance of the Logistic Regression Classifier on Noisy Raw Data 55
Table 4.1: FSCT results from SVM and CRFoGVG models 84
Table 4.2: Performance of LOO training using SVM and CRFoGVG 84
Table 4.3: Performance of SVM trained on *FR79* with LMO training 85
Table 4.4: Performance of SVM trained on *Intellab* with LMO training 86
Table 4.5: Performance of CRFoGVG trained on *FR79* with LMO training 86
Table 4.6: Performance of CRFoGVG trained on *Intellab* with LMO training 87
Table 4.7: Performance of SPCoGVG trained on *FR79* 88
Table 4.8: Performance of SPCoGVG trained on *Intellab* 88
Table 4.9: Performance of Different Classifiers in LMO scenario 91
Table 5.1: Overall Place Classification Accuracies 103
Table 5.2: Object Classification Accuracies Expressed in Precision 105
Table 5.3: Object Classification Accuracies Expressed in Recall 105
Table 5.4: Object-specific Classification Accuracies Expressed in Precision 105
Table 5.5: Object-specific Classification Accuracies Expressed in Recall 106
Table 5.6: System Performance Expressed in Overall Accuracies 107

List of Figures

Figure 2.1: A conceptual hierarchical mapping scheme 9
Figure 2.2: Two common sensors and a comparison of their horizontal field of views . 12
Figure 2.3: The Kinect™ and other functionally identical sensors 12
Figure 2.4: I.I.D. data for supervised learning 18
Figure 2.5: Data structure of a graph 19
Figure 2.6: Samples for undirected graphical models 19
Figure 2.7: A simplified flow chart of the supervised learning 20
Figure 2.8: An illustration of the hard-margin SVM 25
Figure 2.9: An illustration of the soft-margin SVM 28
Figure 2.10: The relationship between some typical graphical models 30
Figure 2.11: An illustration of the feature manipulation techniques 38
Figure 3.1: Typical scanning laser range finder data visualised in different forms 44
Figure 3.2: The indoor environment for 2D laser range data collection 52
Figure 3.3: Test results of the binary classification 54
Figure 3.4: Real OG maps of two indoor environments 56
Figure 3.5: LISA robot (Lightweight Integrated Social Autobot) 57
Figure 3.6: Blueprint of the indoor environment at level 6, building 2, UTS 58
Figure 3.7: Examples of 360° 3D / 2D instance pairs from each class 59
Figure 3.8: The results of feature selection on features from 3D range data 60
Figure 3.9: The results of feature selection on features from 2D range data 61
Figure 3.10: Test accuracies on 360° 3D data set 64
Figure 3.11: Test accuracies on 360° 2D data set 64
Figure 3.12: Test accuracies of different coverage patterns 65
Figure 4.1: An example of the Voronoi diagram 71
Figure 4.2: An example of GVG on the grid map 71

Figure 4.3: An example of CRF over GVG 72
Figure 4.4: The learning framework of fully supervised CRFoGVG 72
Figure 4.5: Schematic flow chart of the proposed framework SPCoGVG 74
Figure 4.6: An example of CRF_{PL} 76
Figure 4.7: Real-world grid maps from different international universities. 81
Figure 4.8: Test results of the of the SPCoGVG framework 89
Figure 4.9: Comparisons of experimental results shown in box-and-whisker diagrams 90
Figure 5.1: Single sensory modality place classification model 95
Figure 5.2: Multiple sensory modalities SPCOD with individual object feature 96
Figure 5.3: Multiple sensory modalities SPCOD with object feature pool 98
Figure 5.4: The process of constructing the OFP for the M-SPCOD-OFP 100
Figure 5.5: The blueprint of the data collection environment 102
Figure 5.6: Examples of object localisation in an office environment 108
Figure 5.7: Examples of object localisation in bathroom and stairwell 108
Figure 5.8: Examples of object localisation in kitchen 109
Figure 5.9: Examples of object localisation in printer room 109
Figure 5.10: A generalised M-SPCOD-OFP model 110

Acronyms and Abbreviations

2D	Two-dimensional
3D	Three-dimensional
AMN	Associative Markov Network
CENTRIST	Census Transform Histogram
COLD	COsy Localisation Database
CRF	Condition Random Field
CRFoGVG	Condition Random Field over the Generalised Voronoi Graph
CRF_{PL}	CRF having Partially Labelled Data
CRFH	Compose Receptive Field Histograms
DAGSVM	Directed Acyclic Graph SVM
DEFS	Differential Evolution based Feature Subset Selection
DNF	Desired Number of Features
EEG	Electroencephalography
FSCT	Fully Supervised Comprehensive Training
GVD	Generalised Voronoi Diagram
GVG	Generalised Voronoi Graph
HMM	Hidden Markov Model
HOG	Histogram of Oriented Gradients
HOUP	Histogram of Oriented Uniform Patterns
HRI	Human-robot Interaction
I.I.D.	Independent and Identically Distributed
IR	Infrared
ISSA	Indoor-space Scene Analysis
LBP	Loopy Belief Propagation
LDA	Linear Discriminant Analysis

LISA	Lightweight Integrated Social Autobot
LMO	Leave-many-out
LOO	Leave-one-out
MRF	Markov Random Field
M-SPCOD-IOF	Multiple Sensory Modalities SPCOD with Individual Object Feature
M-SPCOD-OFP	Multiple Sensory Modalities SPCOD with Object Feature Pool
NF	Number of Features
NLP	Natural Language Processing
Non-I.I.D.	Non Independent and Identically Distributed
NP	Number of Population
OB	Object Bank
OFP	Object Feature Pool
OG Map	Occupancy Grid Map
OISVM	Online Independent-SVM
PCA	Principal Component Analysis
SIFT	Scale-Invariant Feature Transform
SLAM	Simultaneous Localisation and Mapping
S-PC	Single Sensor Modality Place Classification
SPCOD	Simultaneous Place Classification and Object Detection
SPCoGVG	Semi-Supervised Place Classification over a Generalised Voronoi Graph
SPM	Spatial Pyramid Matching
SVM	Support Vector Machine
SVM-DAS	SVM-based Discriminative Accumulation Scheme
VD	Voronoi Diagram
VRF	Voronoi Random Field

Abstract

Place classification is an emerging theme in the study of human-robot interaction which requires common understanding of human-defined concepts between the humans and machines. The requirement posts a significant challenge to the current intelligent mobile systems which are more likely to be operating in absolute coordinate systems, and hence unaware of the semantic labels. Aimed at filling this gap, the objective of the research is to develop an approach for intelligent mobile systems to understand and label the indoor environments in a holistic way based on the sensory observations.

Focusing on commonly available sensors and machine learning based solutions which play a significant role in the research of place classification, solutions to train a machine to assign unknown instances with concepts understandable to human beings, like room, office and corridor, in both independent and structured prediction ways, have been proposed in this research. The solution modelling dependencies between random variables, which takes the spatial relationship between observations into consideration, is further extended by integrating the logical coexistence of the objects and the places to provide the machine with the additional object detection ability. The main techniques involve logistic regression, support vector machine, and conditional random field, in both supervised and semi-supervised learning frameworks.

Experiments in a variety of environments show convincing place classification results through machine learning based approaches on data collected with either single or multiple sensory modalities; modelling spatial dependencies and introducing semi-supervised learning paradigm further improve the accuracy of the prediction and the generalisation ability of the system; and vision-based object detection can be seamlessly integrated into the learning framework to enhance the discrimination ability and the flexibility of the system.

The contributions of this research lie in the in-depth studies on the place classification solutions with independent predictions, the improvements on the generalisation ability of the system through semi-supervised learning paradigm, the formulation of training a conditional random field with partially labelled data, and the integration of multiple cues in two sensory modalities to improve the system's functionality. It is anticipated that the

findings of this research will significantly enhance the current capabilities of the human robot interaction and robot-environment interaction.