INVESTIGATIONS ON HUMAN PERCEPTUAL MAPS USING A STEREO-VISION MOBILE ROBOT

ENG SWEE KHENG

UNIVERSITI SAINS MALAYSIA

2018

INVESTIGATIONS ON HUMAN PERCEPTUAL MAPS USING A STEREO-VISION MOBILE ROBOT

by

ENG SWEE KHENG

Thesis submitted in fulfillment of the requirements for the degree of Doctor of Philosophy

June 2018

ACKNOWLEDGEMENTS

To write this thesis is a great endeavor, and I could not have reached the finish line without the encouragement, advice, and support, especially from my supervisors and family. Many people have contributed towards the completion of this thesis. It is an honour for me to place on record, my sense of gratitude to one and all, who directly or indirectly, have lent their hand in this venture.

I would like to express my very profound gratitude to my supervisors, Associate Professor Khoo Bee Ee and Professor Lim Chee Peng for their guidance and encouragement throughout the research. They consistently provide advice and help me to situate my research in right context, and ensure that I work towards the outcome. Their help are greatly appreciated.

I wish to extend my gratitude to Professor Yeap Wai Kiang for his support and insightful suggestions. Further thanks go to the Ministry of Higher Education (MOHE) and Universiti Malaysia Perlis (UniMAP) for the financial support of my study.

I would like to thank those people who spent their time sharing their knowledge with me to improve my research work, especially Wee Chuen, Tow Leong, Nick, Mr. Nor Azhar bin Zabidin, Mr. Amir bin Hamid, Madam.Mazlina bt Ghani, You Long, Jing Hui, Jing Rui, Belinda, Wei Zeng, Dr. Anwar, Sze Sing, and Wei Hong.

Applauds and appreciations are dedicated to my friends: Jason, Brother Lim, Lee Teng, Susan, Wei Sheik, Adrian, Earn Tzeh, Chong, Saad, and Hafiz for their support and friendship. I would like to thank also all others who have directly or indirectly helped me in this journey.

Lastly, I wish to express my very deep gratitude to my parent, wife, and other family members, for their patience in providing me unfailing support and continuous encouragement throughout my studies.

TABLE OF CONTENTS

	Page
ACKNOWLEDGEMENT	ii
TABLE OF CONTENTS	iv
LIST OF TABLES	viii
LIST OF FIGURES	ix
LIST OF ABBREVIATIONS	xv
LIST OF SYMBOLS	xvi
ABSTRAK	xvii
ABSTRACT	xix
CHAPTER ONE : INTRODUCTION	
1.1 Background	1
1.2 Problem Statement and Motivation	3
1.3 Research Aim and Objectives	6
1.4 Research Scope	8
1.5 Thesis Outlines	9
CHAPTER TWO : LITERATURE REVIEW	
2.1 Introduction	10
2.2 Human Perceptual Maps	11
2.2.1 Sholl (2001) model	13
2.2.2 Wang and Spelke (2000,2002) model	13
2.2.3 McNamara et al. (2003) model	14
2.2.4 Discussion	15

2.3 Robot Mapping - Simultaneous Localization and Mapping 19

2.3.1 Discussion 21

2.4 Implementation of Yeap's Theory of Human Perceptual Mapping 21

2.4.1 Discussion 23

2.5 Implementation of cognitive mapping models using robots 24

2.5.1 Bio-Inspired Cognitive Map - Spatial Semantic Hierarchy Model 24

2.5.2 Neural-Inspired Cognitive Maps 27

2.5.2(a) The RatSLAM model by Milford and Wyeth (2009) 28

2.5.2(b) The neural network model of Hafner (2001) 32

2.5.3 Discussion 34

2.6 Summary 36

CHAPTER THREE : METHODOLOGY

3.1 Research Methodology 38

3.2 Motivation 40

3.3 Building the HPM Model 43

3.4 Stereo vision images 46

3.4.1 Pre-processing Step 46

3.4.1(a) Determine each step of the robot movement from two appropriate frames and detect the locations of the corresponding edge points of two views 46

3.4.1(b) Retrieving of the odometry information 54

3.4.1(c) The procedure for yielding the equation (3.3) 58

3.4.1(d) Procedure for retrieving the locations (in cm) of the edge points 59

3.4.2 Tracked Reference Objects 53

3.4.2(a) Clustering 63

3.5 Tracking and Recognizing the Tracked Reference Objects 67

3.6 Expanding the Map 70

3.7 Remove redundant TROs for boundary computation 76

3.8 Summary 79

CHAPTER FOUR : RESULTS AND DISCUSSION

4.1 Overview 80

4.2 The results of the pre-possessing step 81

4.2.1 The results of each robot step 81

4.2.2 The results of odometry information between two views 88

4.2.3 Remarks on the pre-processing results 91

4.3 Results of tracked reference object (TRO) creation process 97

4.4 Results of the mapping process 99

4.5 Results of Indoor Scene 1– a half-route environment 103

4.6 Results of Indoor Scene 2– a full-route environment 109

4.6.1 Remarks of the indoor (half and full-route) environments 113

4.7 Results of outdoor scene 1 - water fountain 114

4.8 Results of outdoor scene 2 – a car park 120

4.8.1 Remarks of the outdoor (water fountain and car park) environments 124

4.9 Remarks of all tested indoor and two outdoor environments 125

4.10 Summary 128

CHAPTER FIVE : CONCLUSION AND FUTURE WORKS

5.1 Conclusions and Contributions of the Research 129

5.2 Suggestions for Future Work 132

REFERENCES 134

APPENDICES

Appendix A - Bumblebee 2 Stereo Vision Camera

Appendix B - Depth Information

B.1 Determine the Depth Information (Z Coordinate)

Appendix C - Specification of Mobile Robot

Appendix D - Angle of Rotation

Appendix E - The Information of Noise

LIST OF PUBLICATIONS

LIST OF TABLES

		Page
Table 2.1	Summary of three competing models	17
Table 2.2	Implementation of Yeap's Human Perceptual Mapping	23
Table 2.3	Implementation of cognitive models using robots	34
Table 3.1	The locations (pixel and cm) of some corresponding edge points in the previous and the current views.	54
Table 3.2	The result of Mean_Angel of rotation per pixel.	59
Table 3.3	Some of the inputs and outputs used for yielding the equation (3.3).	59
Table 3.4	Configuration of the Bumblebee 2 Stereo Vision Camera.	62
Table 4.1	Quantitative data for Figure 4.1(a to d).	84

LIST OF FIGURES

		Page
Figure 1.1	Sholl (2001) model: the dot with a cross indicates the position and orientation of the viewer in the map. The view and the human perceptual map are egocentric representations, while the cognitive map is an allocentric representation.	3
Figure 1.2	An indoor test environment and the robot's path.	6
Figure 2.1	Wang and Spelke (2002) model: No allocentric map is computed.	14
Figure 2.2	McNamara et al. (2003) model: Individual allocentric maps are remembered (indicate by circles inside the cognitive map), and the most recent one (indicate by dashed circle) includes the allocentric position of the self (denoted by a dot). The dashed circle superposed on the view indicates possible egocentric updating of a few objects that have move out of view.	15
Figure 2.3	A tentative process supporting spatial cognition. (Source: Yeap, 2014).	18
Figure 2.4	The framework of Spatial Semantic Hierarchy model. Dependencies is denoted using close-headed arrows and the open-headed arrows denotes the potential information flow without dependency. (Source: Kuipers, 2000).	25
Figure 2.5	The framework of HSSH model. (Source: Beeson, et al., 2010).	27
Figure 2.6	The RatSLAM model. Each local view cell is associated with a distinct visual scene in the environment and becomes active when that scene is perceived by the robot. A 3-D CAN forms the pose cells, where active pose cells encode the estimate of the pose of robot. Active local view and pose cells drive the creation of experience nodes in the experience map. (Source: Wyeth and Milford, 2009).	29
Figure 2.7	Visual interface to the local view cells. (Source: Wyeth and Milford, 2009).	29
Figure 2.8	The develop robot system in an indoor environment. (Source: Wyeth and Milford, 2009).	30

Figure 2.9 Top: Plan view of the indoor test environments. Bottom: The experience map. (Source: Wyeth and Milford, 2009). 31

Figure 2.10 Top: Aerial photo of St.Lucia. Bottom: The experience map. (Source: Wyeth and Milford, 2009). 32

Figure 2.11 A neural network structure for deciding a homing direction from two images. (Source: Hafner, 2001). 33

Figure 3.1 A flow chart for building the human perceptual map (HPM) model. 39

Figure 3.2 The platform of vision-based mobile robot. 42

Figure 3.3 The overview of the proposed perceptual map building procedure. 43

Figure 3.4 Creating a human perceptual map (HPM): V0…V5 are six robot viewing positions and the circles indicate the view boundary for V0, V2 and V4. The perceptual map consists of V0, V2 and V4. By triangulating the viewing position of V4 in V2 and then V2 in V0 using the common surfaces (solid line) respectively, one could create a global map of the three views. 44

Figure 3.5 The corresponding edges points between framen-1 and framen+2 using indirect and direct methods, respectively. 50

Figure 3.6 The results when the number of pre-defined values are set to 120 and 130, respectively. 51

Figure 3.7 Flow chart of determine each step of robot and the locations of the edges points. 52

Figure 3.8 An iterative method for finding the final edges: a) original image (F0); b) smoothed image (G0); c) edge pixels detected; d) a synthetic 3×9 subimage is generated from estimated features for every edge pixel; and e) sub-images are combined to generate a complete restored image (F1). (Source: Trujillo-Pino, 2013). 53

Figure 3.9 The remaining edge points with the corresponding numbers in the right image of the previous view (left) and the right image of current view (right). 53

Figure 3.10 The difference in the y pixel location between the corresponding edge points in the previous and current views. The horizontal and vertical axes represent number of matched pairs and the difference in the y pixel location between the corresponding matched pairs, respectively. 57

Figure 3.11	Matched pairs with a difference absolute value between the z coordinates of the edge points in the previous view and those in the current view.	57
Figure 3.12	Matched pairs with a difference absolute value between the x coordinates of the edge points in the previous view and those in the current view.	58
Figure 3.13	Camera geometry.	60
Figure 3.14	The edge points in an image. Green box denotes the robot position and the green arrow denotes the robot orientation.	63
Figure 3.15	The example of using 2-point line (red) for representing a group of reference points(blue) in a grid area.	65
Figure 3.16	Clustering the reference points using a grid-based method. Green box denotes the robot position and the green arrow denotes the robot orientation.	66
Figure 3.17	Reference objects with their respective ID (identification) numbers. Green box denotes the robot position and the green arrow denotes the robot orientation.	66
Figure 3.18	The matching between all *TRO'j* and *TROj* (pertaining to the current view coordinate system). The red and blue end points represent all *TRO'j* and *TROj*, respectively.	68
Figure 3.19	The results of remaining TROs (after matching) for *Rv* =0.5, 0.7, and 1.4, respectively.	69
Figure 3.20	Computing the spatial locations of reference objects (i.e., S2 and S3) close to a common tracked reference object. S1 is recognized as a common tracked reference object and S2 and S3 are coded using two pairs of vectors centred on the right end-point of S1. (Source: Azizul, 2014).	70
Figure 3.21	Pseudo-code for computing the position of each new TRO in the global map.	72
Figure 3.22	Computing the spatial locations of new tracked reference objects in the global map.	73
Figure 3.23	A floor plan in a clock-wise direction.	74
Figure 3.24	Changes in the perceptual map with time. The red lines indicates all the tracked reference objects. The blue lines indicate the path lines and the blue circles indicate the path points.	75

Figure 3.25	Flow chart of how to remove the redundant TROs.	77
Figure 3.26	The initial global map with the identified inner (green line) and outer (red line) tracked reference objects. The blue lines indicate the path lines and the blue circles indicate the path points.	77
Figure 3.27	The global map after removing the redundancy inner (green line) and outer (red line) tracked reference objects. The blue line indicates the path line and the blue circles indicate all path points.	78
Figure 3.28	The computed boundary of the global map. The blue, green and red lines indicate the path line, the inner tracked reference objects, and the outer tracked reference objects, respectively.	78
Figure 4.1	The results of matched edge points for case (a).	85
Figure 4.2	The results of matched edge points for case (b).	86
Figure 4.3	The results of matched edge points for case (c).	87
Figure 4.4	The result of curve fitting tool of Equation (3.3).	88
Figure 4.5	Results of odometry information between two views in half route environment.	93
Figure 4.6	Results of odometry information between two views in full route environment.	94
Figure 4.7	Results of odometry information between two views in water fountain.	95
Figure 4.8	Results of odometry information between two views in car park.	96
Figure 4.9	The results of the TRO creation process. The green box denotes the robot position and the green arrow denotes the robot orientation.	98
Figure 4.10	The results of mapping process.	101
Figure 4.11	The floor plan of the half-route environment. (a) clockwise direction; (b) anti-clockwise direction.	105
Figure 4.12	The initial global map of the half-route environment. The blue and red lines indicate the path lines and the tracked reference objects, respectively. (a) clockwise direction; (b) anti-clockwise direction.	106

Figure 4.13	The global map of the half-route environment with redundant objects removed. The blue, green, and red lines indicate the path lines, the inner tracked reference objects and the outer tracked reference objects, respectively. (a) clockwise direction; (b) anti-clockwise direction.	107
Figure 4.14	The computed boundary of the global map for the half-route environment. The blue, green, and red lines indicate the path lines, the inner tracked reference objects, and the outer tracked reference objects, respectively. (a) clockwise direction; (b) anti-clockwise direction.	108
Figure 4.15	The floor plan of the full-route environment. (a) clockwise direction; (b) anti-clockwise direction.	110
Figure 4.16	The initial global map of the full-route environment. The blue and red lines indicate the path lines and the tracked reference objects. (a) clockwise direction; (b) anti-clockwise direction.	111
Figure 4.17	The global map of the full-route environment with redundant objects removed. The blue, green, and red lines indicate the path lines, inner tracked reference objects and outer tracked reference objects, respectively. (a) clockwise direction; (b) anti-clockwise direction.	112
Figure 4.18	The computed boundary of the global map for the full-route environment. The blue, green, and red lines indicate the path lines, inner tracked reference objects, and outer tracked reference objects, respectively. (a) clockwise direction; (b) anti-clockwise direction.	113
Figure 4.19	The view of the water fountain. (a) clockwise direction; (b) anti-clockwise direction.	116
Figure 4.20	The initial global map of the water fountain. The blue and red lines indicate the path lines and the tracked reference objects, respectively. (a) clockwise direction; (b) anti-clockwise direction.	117
Figure 4.21	The global map of the water fountain with redundant objects removed. The blue, green, and red lines indicate the path lines, the inner tracked reference objects, and the outer tracked reference objects, respectively. (a) clockwise direction; (b) anti-clockwise direction.	118

Figure 4.22 The computed boundary of the global map for the water fountain. The blue, green, and red lines indicate the path lines, the inner tracked reference objects, and the outer tracked reference objects, respectively. (a) clockwise direction; (b) anti-clockwise direction. 119

Figure 4.23 The view of the car park. a) with clockwise direction b) with anti-clockwise direction. 121

Figure 4.24 The initial global map of the car park. The blue line and red line indicate the path line and the tracked reference objects, respectively. a) with clockwise direction b) with anti-clockwise direction. 122

Figure 4.25 The global map of the water fountain with redundancy objects removed. The blue, green, and red lines indicate the path line, inner tracked reference objects and outer tracked reference objects, respectively. a) with clockwise direction; b) with anti clockwise direction. 123

Figure 4.26 The computed boundary of the global map for the car park. The blue, green, and red lines indicate the path line, the inner tracked reference objects and outer tracked reference objects, respectively. a) with clockwise direction; b) with anti-clockwise direction. 124

Figure 4.27 Map generated by Albot1 for three tests. The route is indicated by arrows (left). The floor plans and the corresponding maps are shown at left and right, respectively. (Source: Hossain, 2014). 127

LIST OF ABBREVIATIONS

CAN	Continuous Attractor Network
EM	Expectation Maximization
HPM	Human Perceptual Map
HSSH	Hybrid Spatial Semantic Hierarchy
LPM	Local Topological Map
NCC	Normalized Cross-Correlation
PF	Particle Filtering
RGB	Red, Green, Blue
SLAM	Simultaneous Localization and Mapping
SSH	Spatial Semantic Hierarchy
TRO	Tracked Reference Object
UAVs	Unmanned Aerial Vehicles

LIST OF SYMBOLS

Θ_c	Computed Angle of Rotation
Θ_e	Estimated Angle of Rotation
Θ_{turn}	Accumulated Turn-angle
F_0	Original Image
G_0	Smoothed Image
F_1	Synthetic Image
V_0	Starting View
V_n	New View
m	Slope
b	Offset
Tx	x Translations
Tz	z Translations
T	Transformation Matrix
$\bar{t}$	Mean of the descriptors
f	Image

PENYIASATAN TERHADAP PEMETAAN PERSEPSI MANUSIA DENGAN MENGGUNAKAN ROBOT BERGERAK PENGLIHATAN STEREO

ABSTRAK

Kognitif ruang adalah cabang psikologi kognitif mengenai pemerolehan, penyusunan, penggunaan, dan semakan pengetahuan tentang persekitaran ruang. Teori pengiraan baru untuk pemetaan kognitif ruang manusia telah dicadangkan dalam kesusasteraan dan dianalisis menggunakan robot mudah alih berasaskan laser. Berbeza dengan pendekatan SLAM (Penyetempatan dan Pemetaan Secara Serentak) yang membina peta persekitaran yang tepat dan sempurna, prosedur pembinaan peta persepsi manusia yang dicadangkan lebih mewakili pemetaan kognitif ruang dalam otak manusia, di mana peta persepsi persekitaran yang tidak tepat dan tidak lengkap boleh dibina dengan mudah. Langkah-langkah utama dalam metodologi adalah memperolehi imej-imej stereo penglihatan persekitaran, mewujudkan objek rujukan, menjejaki jumlah baki objek rujukan, dan mengembangkan peta apabila titik-titik had persekitaran dicapai. Sumbangan utama penyelidikan ini adalah penggunaan teknik penglihatan komputer dan algoritma pengiraan pemetaan pada robot mudah alih berasaskan stereo penglihatan untuk merumuskan peta persepsi manusia secara sistematik dan menilai peta persepsi manusia yang berkaitan dengan persekitaran dalaman dan persekitaran luaran secara komprehensif. Pengesahan peta persepsi manusia dengan menggunakan teknik berasaskan penglihatan adalah penting kerana dua sebab. Pertama, penglihatan memainkan peranan penting dalam pembangunan kognitif ruang manusia; Kedua, sistem penglihatan komputer kurang mahal dan kaya dengan maklumat dalam mewakili persekitaran. Secara khusus, teknik penglihatan komputer dibangunkan terlebih dahulu untuk menganalisis imej stereo yang berkaitan dan memperolehi maklumat anjakan robot bergerak, serta mewujudkan objek rujukan. Beberapa algoritma pengiraan pemetaan digunakan kemudian untuk membina persepsi manusia terhadap persekitaran dalam penyelidikan ini. Empat persekitaran dunia nyata iaitu dua persekitaran dalaman dan dua persekitaran luaran yang besar, dinilai secara empirik. Geometri ruang dari persekitaran pemeriksaan adalah berbeza-beza, dan persekitaran tertakluk kepada pelbagai kesan semula jadi termasuk pantulan dan hingar. Pantulan dan hingar terjadi di banyak bahagian imej. Oleh itu, algoritma tambahan dibangunkan untuk menyingkirkan pantulan dan hingar. Penyingkiran pantulan dan hingar ketara mengurangkan objek-objek rujukan (TROs) yang dibuat, untuk setiap pandangan semasa. Hasilnya menunjukkan bahawa teknik penglihatan komputer dan algoritma pengiraan pemetaan yang dicadangkan untuk pembinaan peta persepsi manusia adalah mantap dan berguna. Teknik penglihatan komputer yang dicadangkan dapat membina peta persepsi manusia yang tidak tepat dan tidak lengkap dengan perwakilan ruang yang baik untuk seluruh persekitaran. Peta yang tidak tepat dan tidak lengkap merujuk kepada peta yang dihasilkan tidak tepat dalam istilah metrik dan mempunyai permukaan yang hilang. Hasil kajian menunjukkan bahawa kedua-dua sistem berasaskan penglihatan dan laser dapat menghasilkan geometri ruang yang agak tepat bagi persekitaran yang diuji.

INVESTIGATIONS ON HUMAN PERCEPTUAL MAPS USING A STEREO-VISION MOBILE ROBOT

ABSTRACT

Spatial cognition is a branch of cognitive psychology concerning the acquisition, organization, utilization, and revision of knowledge about spatial environments. A new computational theory of human spatial cognitive mapping has been proposed in the literature, and analyzed using a laser-based mobile robot. In contrast with the well-established SLAM (Simultaneous Localization and Mapping) approach that creates a precise and complete map of the environment, the proposed human perceptual map building procedure is more representative of spatial cognitive mapping in the human brain, whereby an imprecise and incomplete perceptual map of an environment can be created easily. The key steps in the methodology are capturing stereo-vision images of the environment, creating the tracked reference objects (TROs), tracking the number of remaining TROs, and expanding the map when the limiting points of the environment are reached. The main contribution of this research is on the use of computer vision techniques and computational mapping algorithms on a stereo-vision mobile robot for formulating the human perceptual map systematically, and evaluating the resulting human perceptual maps pertaining to both indoor and outdoor environments comprehensively. Validating the human perceptual maps using vision-based techniques is important for two reasons. Firstly, vision plays an important role in the development of human spatial cognition; secondly, computer vision systems are less expensive and information-rich in representing an environment. Specifically, computer vision techniques are first developed for analyzing the associated stereo images and retrieving the displacement information of a mobile robot, as well as

creating the necessary tracked reference objects. A number of computational mapping algorithms are then employed to build a human perceptual map of the environment in this research. Four real-world environments, namely two large indoor and two large outdoor environments, are empirically evaluated. The spatial geometry of the test environments vary, and the environments are subject to various natural effects including reflection and noise. The reflection and noise occurrin many parts of the images. Therefore, additional algorithms are developed in order to remove the reflection and noise. The removal of reflection and noise significantly reduces the number of TROs createdfor every immediate view. The outcomes indicate that the proposed computer vision techniques and computational mapping algorithms for human perceptual map building are robust and useful. They are able to create imprecise and incomplete human perceptual maps with good spatial representation of the overall environments. The map is imprecise and incomplete in the sense that it is not accurate in metric terms and has perceived surfaces missing. It is shown that both vision-based and the laser-based systems are able to compute a reasonably accurate spatial geometry of the tested environment.

CHAPTER ONE

INTRODUCTION

1.1Background

Spatial cognition is a branch of cognitive psychology which is concerned with the acquisition, organization, utilization, and revision of knowledge about spatial environments (Freksa, 2004). It allows cognitive agents, e.g. humans, animals, or robots, to act and interact in space effectively and to communicate about spatial environments efficiently. The spatial and temporal cognitive capabilities allow humans to efficiently manage cognitive tasks, e.g. going to workplace or/and returning home, in everyday life (Nebel and Freksa, 2011, Freksa, 2004).

Researchers in the spatial cognition community infer one's internal representation of spatial knowledge pertaining to an explored environment as a 'cognitive map', a term first coined by Tolman (1948). The term was created by recording the behavior of a maze-running rat that was able to take short-cuts to a desired destination. In principle, cognitive mapping is a mental structuring mechanism involving the process of sensing, encoding, storing, and decoding knowledge that describes the relative locations and attributes of phenomena in one's spatial environment (Downs and Stea, 1973, Arthur and Passini, 1992).

Since Tolman's (1948) work, researchers in cognitive psychology have carried out numerous experiments to investigate the nature of cognitive maps, e.g. Olton (1977); Siegel and White (1975); Presotto and Izar (2010); and Rosati and Hare (2013). Lynch (1960) carried out an empirical research on city planning and studied how urban residents orient themselves by means of mental maps. The mental maps consist of five inter-related components: paths, landmarks, nodes, edges, and districts. Their cognitive maps are the "images" of their city. In neurological studies, O'keefe and Nadel (1978) first outlined a spatial function of place-coded neurons in hippocampus to compute a cognitive map. The hippocampus of the human brain is regarded as the neural substrate of a cognitive map.

Despite attracting much interest, the notion of a cognitive map is still controversial (Bennett, 1996). Many studies, e.g. Tolman (1948); O'keefe and nadel (1978); and Gallistel (1990), have tried to define what it is. While it is widely accepted that the term "cognitive map" refers to the representation of one's environment, what is controversial is its map-like property that supposedly differentiates it from other known knowledge of one's environment (Yeap and Jefferies, 2000, Mackintosh, 2002, Yeap, 2014, Andrews and Beck, 2017). In conjunction with the notion of a cognitive map, a perceptual map is defined as a representation of the spatial layout of surfaces/objects perceived in one's immediate surroundings (Hossain et al., 2011, Yeap, 2011a). Therefore, much research focuses on integrating successive views and remembering the position of objects viewed, either relative to the self or within a fixed reference frame.

The perceptual map is used to maintain a perspective view of objects in one's immediate surroundings, while the cognitive map is used to create different perspectives on the remembered spatial arrangement of objects. A perceptual map acts as an interface between what is one's view and one's cognitive map. Figure 1.1 shows the Sholl (2001) model that depicts the relationship of a viewer, a perceptual map, and a cognitive map. On the other hands, One key aspect of cognitive mapping, as opposed to perceptual mapping, is the ability to do abstraction and use the knowledge abstracted to help solve spatial tasks (Hossain, 2014).

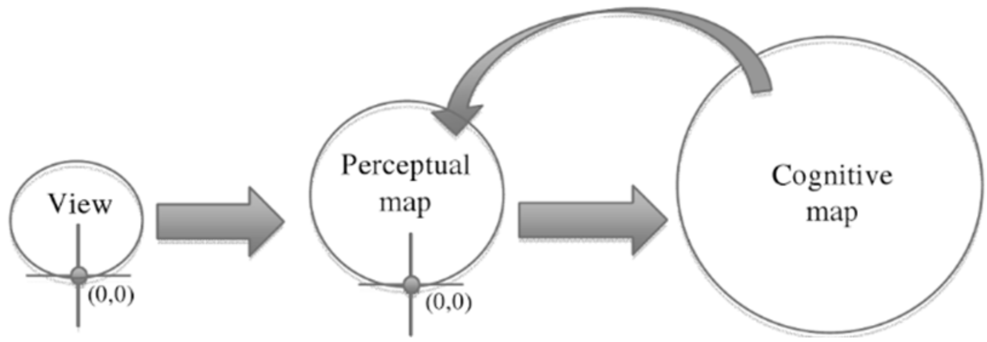

Figure 1.1 Sholl (2001) model: the dot with a cross indicates the position and orientation of the viewer in the map. The view and the human perceptual map are egocentric representations, while the cognitive map is an allocentric representation.

It is evident that humans have the capability of building a perceptual map of the environment. We are able to remember what is out of sight when we move forward or turn (Glennerster et al., 2009). Some investigations, e.g., Allen and Haun (2004) and Farrell and Robertson (1998), provide evidence to show that humans compute the perceptual map seamlessly and almost effortlessly. The computed map is accurate enough for humans to orient themselves in the environment. Many studies in spatial cognition, e.g. Burgess (2006); Wang and Spelke (2000); Zhang, Mou, and McNamara (2011); and Tatler and Land (2011) often assume that a perceptual map is computed by integrating successive views using a co-ordinate transformation method. As such, current research studies are focused on how to use the frame of references (egocentric and allocentric), and what representation can be computed from such a spatial cognition in general, and the perceptual map in particular.

1.2 Problem Statement and Motivation

A well-known problem of the co-ordinate transformation method is the computed perceptual map is easily distorted owing to errors in computing the turn and distance