Creating Conceptual Maps by Natural Language Descriptions for Human-Friendly Robot Navigation

Hitihami Mudiyanselage Ravindu Tharaka Bandara

(178068R)

Thesis submitted in partial fulfillment of the requirements for the degree Master of Science in Electrical Engineering

Department of Electrical Engineering

University of Moratuwa

Sri Lanka

August 2019

DECLARATION

I declare that this is my own work and this dissertation does not incorporate without acknowledgment any material previously submitted for a Degree or Diploma in any other University or institute of higher learning and to the best of my knowledge and belief it does not contain any material previously published or written by another person except where the acknowledgment is made in the text.

Also, I hereby grant to University of Moratuwa the non-exclusive right to reproduce and distribute my dissertation, in whole or in part in print, electronic or other medium. I retain the right to use this content in whole or part in future works (such as articles or books).

Signature: Date:

The above candidate has carried out research for the MSc thesis under my supervision.

Signature of the Supervisor(s): Date:

Dr. A.G.B.P. Jayasekara

Dr.D.P. Chandima

Abstract

The living standards of elderly people are uplifted through developing assistive robots who are capable of supporting their daily activities by possessing the competency to provide companionship to human beings. Human activities are frequently related to navigational tasks and human tend to use descriptions which include natural language phrases and the terms which describe the distance related uncertainties such as little, large, near, far, small, close to describe about spatial information. Therefore assistive robots should be capable of analyzing and understanding descriptions which include natural language phrases. The best and foremost navigation skills of a robot are represented with the competency of virtual imagination related to an unknown environment. Subsequently the quality of virtual imagination of the robot should be improved with the experiences as human beings. The quality of understanding the user appropriately, efficiently and effectively plays a vital role in order to expand their knowledge as well as the experiences collected from day to day performances and conversations. Accordingly, the robot should contain a memory and an advance knowledge base including the information regarding objects that it experiences in day to day activities. The requirement of a user should be clearly configured with the capability of processing data in order to grasp the relationships among attributes of objects. Obviously, a robot should possess the competency to interpret spatial information in the mode of uncertain terms. Significantly the robot should be capable enough to enhance the knowledge through effective communication with the user.

This research proposes a procedure to understand spatial information in a description with the uncertain terms and creates a conceptual map in a robot memory which can be linked with spatial cognitive map for purposeful, effective and human friendly navigation task. The proposed method is consisted with creation of cognitive object maps and cognitive spatial maps. Further, both maps are created based on information conveyed through interactive conversations between the user, robot and the vocal descriptions. The conceptual maps are created by amalgamating the spatial and object cognitive maps. Moreover, the conceptual map also can be updated using the conversations occurred in the interactions among robot and the user. Furthermore, this research proposes a procedure to enhance the capability of virtual imagination of a service robot while understanding the information regarding the uncertainties of an object size with the aid of artificial neural network. In addition to that gesture command identification method and vocal navigation command identification method are also implemented. Moreover, navigation commands are categorized and studied for the unique attributes such as uncertainty, incompleteness, inconsistency and unpredictability to improve human robot interaction.

The proposed method is validated by the experiments using MI Rob platform (Version 2.0). Further, a software platform was introduced by integrating the implementation of proposed methods. Software platform is consisted of processing spatial maps for room boundary establishment to object identification considering spatial information acquired through descriptions and conversations. System generated data was compared with data derived from human studies and analyzed using statistical methods in order to validate system accuracy. The proposed system can be used to enhance the human-robot interactions and to perform navigational tasks more effectively and purposefully in a previously unknown environment.

***Keywords*-Virtual Maps, Human-Robot Interaction, Conceptual Maps**

DEDICATION

To my beloved parents

ACKNOWLEDGMENTS

I would like to express my deepest appreciation to my supervisors Dr. A. G. B. P. Jayasekara and Dr D. P. Chandima who convincingly conveyed the best attitude on behalf of my research and progress. Their guidance and persistence showed the path of success in regards to this dissertation. Moreover I would like to express my sincere gratitude to the panelists Prof. Nalin Wickckramarachchi and Prof. Ruwan Gopura for the valuable comments provided throughout the years.

I am extremely grateful to my colleagues Viraj Muthugala, chapa sirithunge ,Sahan priyanayana, Baghya samarakoon, Sachi Natasha for the indispensable suggestions provided to complete the thesis.

My heartfelt thanks to parents for their love and continuous assistance given to make this dissertation a success.

Finally, special thanks goes to my dearest friends Sandaruwan Bassnayaka, S.T. Pahattuge, K.D.M. Jayawardhana, T.M. Senarath, B.L.P.M. Balasooriya and Aruna Buddhika for being supportive to complete my thesis successfully.

This work was supported by the University of Moratuwa Senate Research Grant SRC/CAP/17/03 and SRC/LT/2018/20. Furthermore, some of the conference publications originated from this thesis were partially supported by travel grants awarded by IEEE Systems, Man, and Cybernetics Society.

TABLE OF CONTENTS

Declaration i

Abstract ii

Dedication iv

Acknowledgments v

Table of Contents xi

List of Figures xvii

List of Tables xix

1 Introduction 1

1.1 Introduction 1

1.2 Problem statement 2

1.3 Thesis Overview 3

2 Literature Review 6

2.1 Navigation 9

2.2 Conversation Models 11

2.3 Navigation Commands 14

2.4 Cognitive Mapping 16

3 System Design **21**

3.1 Hardware Platform 22

3.2 System overview 23

3.3 Software Platform 26

4 Evaluation of language description **27**

4.0.1 Human study I - Identification of Sentence Patterns 27

4.0.2 Description 27

4.0.3 Sentence patterns 29

4.0.4 Uncertain terms 30

4.0.5 Discussion - sentence patterns 31

4.0.6 New object identification 32

4.0.7 Position identification 33

4.0.8 Human study II - Virtual placement of objects 35

4.0.9 Evaluation of uncertainty of position 36

4.1 Results 37

4.1.1 Experimental setup 37

4.1.2 Experimental methodology 37

4.1.3 Experimental results 40

5 Conceptual Space **49**

5.1 Introduction 49

5.2 Identification of factors 50

5.2.1 Certainty Level - C 50

5.2.2 Experiment I-Identification of the impact of object size when interpreting uncertain term “near” 50

5.2.3 Experiment II- Identification of certainty levels related with uncertain term “near” 50

5.3 Conceptual Dimensions 52

5.3.1 Conceptual factors in distance 54

5.4 Impact of object Size in Certainty level- C 54

5.5 Position identification 56

5.6 Results and Discussion 56

5.6.1 Experimental setup 56

6 Cognitive Maps **63**

6.1 Introduction 63

6.2 Background 63

6.3 System overview for the cognitive map creation module 64

6.4 Conversation Management Module - (CMM) 65

6.5 Creation of Cognitive Map 70

6.6 Uncertain Term Interpretation 71

6.7 Results and Discussion 72

7 Voice command categorization **76**

7.1 Attributes in Conversations 76

7.1.1 Unpredictability 76

7.1.2 Incompleteness 76

7.1.3 Inconsistency 77

7.1.4 Uncertainty 77

7.2 Human study - Understanding the attributes of human-human vocal communication 79

7.3 Human study II- Communication preference understanding 82

8 Identification of Navigation Command **86**

8.1 Introduction 86

8.2 System Overview for navigation command identification 86

8.3 Navigation command interpretation 88

8.3.1 Direct navigational Command Classification 88

8.3.2 Creation of cognitive map 88

8.3.3 Robot Eye - Base Coordination 89
8.3.4 Hand Gesture Identification 91
8.4 Conversation management module - CMM 94
8.5 Result and discussion 94

9 Cognitive Object Maps **99**

9.1 System Overview 99
9.1.1 Object Library 100
9.1.2 Object Memory 100
9.2 Attributes 101
9.3 Object Categorisation 102
9.4 User impact on object 103
9.5 Conversation Management Module - (CMM) 105
9.6 Result and Discussion 109
9.7 Conclusion 112

10 Creation of Conceptual Maps **113**

10.1 Introduction 113
10.2 System Overview for conceptual map creation module 115
10.3 Spatial and virtual map creating process 116

10.3.1 RGB image-map to Binary image-map conversion algorithm 116

10.3.2 Continuity enhancing to minimize edge discontinuities in the map 117

10.3.3 Space filling algorithm for spaces consist of complete boundaries. 119

10.3.4 Rectangular room boundary establishing algorithm 120

10.3.5 VISMALC (Visual and Spatial Map Link Connector) algorithm 123

10.4 Result and discussion 125

10.4.1 Experiment setup 125

11 Virtual Map Improvement **127**

11.1 Introduction 127

11.2 System Overview 129

11.3 Virtual Object Modeler-VOM 131

11.4 Result and Discussion 138

12 Conclusions **148**

Bibliography **164**

LIST OF FIGURES

1.1 Steps for the proposed navigation system implementation 5

2.1 What is navigation ? 10

2.2 Research Focus 20

3.1 System Overview 21

3.2 MIRob Platform developed by Intelligent Service Robotics-Group University of Moratuwa 22

3.3 Overall Functionality of proposed system 23

3.4 Interface of the monitoring software- sentence pattern identification step 24

3.5 System Overview for Description evaluation 25

4.1 The map of the environment designed for human study explained in section 3.10. 35

4.2 Results of the human study explained in section 3.10. used to find values for α and β 36

4.3 Explaining descriptions to the robot 39

4.4 Boxplots for Description 1 object positioning 43

4.5 Locations of the objects from description 1 44

4.6 Boxplots for description 2 data 47

4.7 Locations of the objects from description 2 48

5.1 Experiment I results. (a) participants movements relative to table and (b) participants movements with pen and book 51

5.2 Certainty levels for Experiment II. (a) certainty levels calculated for chair and table movements. (b) certainty levels calculated for book and pen movements 53

5.3 (a) and (b) certainty level calculated for Experiment I based on proposed method. (c) middle range identification based on certainty level 57

5.4 (a) and (b) shows results for Experiment III 58

5.5 Experiment IV setup 60

6.1 System Overview for the Cognitive map creation module 64

6.2 positions defined for square shaped room 66

6.3 Conversation flow of the question category. Conversation starts by the robot by giving R6 response to the user. Based on that user's response robot will select the suitable category as shown in the diagram. Each and every output is given as a code shown in TABLE 6.1,6.2 and 6.3. Circle shapes represent the responses of the robot while rectangular shapes represent the users response. Diamond shapes are for decision making situations handles by SIP 68

6.4 Conversation flow for Descriptions 69

6.5 Conversation flow for undefined responses of user 69

6.6 Spatial Map of experiment setup 70

6.7 Cognitive Map for fig. 6.6. 71

6.8 Spatial Map of experiment setup 73

6.9 Cognitive Map for fig. 6.8. 73

6.10 User having conversation with the robot 75

7.1 Uncertainty associated with navigation commands in human-human and robot-human interaction 80

7.2 Attribute usage in human-human and robot-human interaction . 80

7.3 User rating Vs navigation command calculated based on results of human study II 83

7.4 Navigation command Vs number of participants identified the command as a Human response 84

7.5 Navigation command with irregularities and without irregularities distribution over 30 commands used in Human study II 85

7.6 Venn diagram for user preferences over voice commands 85

8.1 System Overview for navigation command identification 87

8.2 Cognitive map to represent direct navigational commands 89

8.3 The frame transformation between kinect frame and the neck base frame 90

8.4 11.1, 10.2(b) has been captured while giving command “go there” and “come here”. 8.4(c) and 8.4(d) shows the gestures recognise by Kinect using skeleton tracking feature during capture of 11.1 and 10.2(b) 91

8.5 Direction Identification using hand gesture 93

8.6 conversation flow- Each and every output is given as a code shown in TABLE 8.1, 8.3 and 8.2. Circle shapes represent the responses of the user while rectangular shapes represent the robot’s responses. Diamond shapes are for decision making situations. 95

8.7 Robot positions during the experiment. Red circles show the initial position of the robot before receiving a command and blue circles show the position of the robot after executing the command . . . 98

8.8 User satisfaction levels for navigational commands used in experiment. Navigationa commands are given in Table. 8.4 98

9.1 System Overview 100

9.2 Example Cognitive Map 104

9.3 Conversation flow. Each and every output is given as a code shown in TABLE 8.1 and 8.2. Circle shapes represent the responses of the user while rectangular shapes represent the robot’s responses. Diamond shapes are for decision making situations. 106

9.4 Cognitive Map 110

10.1 (a)- System Overview, (b)-Process flow of laser maps 118

10.2 (a) Map produced by the laser scanner mapping system, (b) Binary map produced by the Color to Binary Image Convertor ,(c) Map produced by the Edge Continuity Corrector. , (d) Room boundaries defined by the Room Boundary Establisher, (e)Free space inside the room filled by Free Space Filler, (f) Free space outside the room is filled by Outside Filler, (g) Object map after inversing the space filled map, (h) Calculated centroid of the objects and approximations to the standard shapes. 124

10.3 Virtual spatial map 125

10.4 Cognitive map 126

10.5 Robot positions during experiment session. N denotes the command number shown in TABLE 10.1 126

11.1 System Overview 130

11.2 Process flow of map creation 131

11.3 Artificial Neural Network For object dimension identification . . . 132

11.4 Shapes used for virtual map improvements 132

11.5 (a)-Fuzzy sets for first Dimensional Parameter : Generated based on the values K_1^S , K_1^M , K_1^L which are generated by ANN as described,(b)-Fuzzy sets for second Dimensional Parameter : Generated based on the values K_2^S , K_2^M , K_2^L which are generated by ANN as described, (c)-Fuzzy output to give a combine result for an object size based on two object dimensional parameters. . . . 134

11.6 (a) Actual map of environment, (b) Actual map after overlaped with virtual map, (c) Processed map with identified objects . . . 139

11.7 Enhanced virtual map 140

11.8 Interface of the monitoring software - edge improving process . . . 141

11.9 Interface of the monitoring software - edge smoothing process . . 142

11.10Interface of the monitoring software - Room boundary establishment process 143

11.11Interface of the monitoring software - Room space identification process 144

11.12Interface of the monitoring software object verifying process . . . 145

11.13Interface of the monitoring software - object detecting process . . 146

11.14Interface of the monitoring software - object mapping process . . 147

LIST OF TABLES

4.1 Example Sentences And outputs 33

4.2 α and β values 36

4.3 Output of Instruction Identifier 40

4.4 Input of VSDI 40

4.5 Output of System and Human Study description 1 42

4.6 Output of Instruction Identifier - description 2 43

4.7 Input of VSDI - description 2 45

4.8 Output of System and Human Study description 2 46

5.1 Experiment III result 59

5.2 Experiment IV result 61

5.3 Comparison between Experiment IV results 62

6.1 Defined Questions of the CMM 67

6.2 Defined Responses of the CMM 67

6.3 Defined Descriptions of the CMM 67

7.1 Composition of identified attributes in vocal navigation commands 80

7.2 Results obtained from Human Study II 82

8.1 Speech outputs for conversation - Questions 94

8.2 Speech outputs for conversation - Responses 96

8.3 Speech outputs for conversation - Navigational command 96

8.4 user satisfaction level for Navigational commands 97

9.1 Speech outputs for conversation - Questions 108

9.2 Speech outputs for conversation - Responses 108

10.1 Navigation commands 126

11.1 ANN input table 132

11.2 ANN input table 134

11.3 ANN input table 134

11.4 Confusion Matrix for map given in Fig 11.6(b) 138

11.5 Kappa analysis for confusion matrix given in Table 11.4 138