

論文 / 著書情報
Article / Book Information

題目(和文)	人間 - ロボット相互干渉における深層環境情報へのマルチロボット行動適応
Title(English)	Multi-Robot Behavior Adaptation to Deep Level Environmental Information in Humans-Robots Interaction
著者(和文)	陳略峰
Author(English)	Chen Luefeng
出典(和文)	学位:博士(工学), 学位授与機関:東京工業大学, 報告番号:甲第9873号, 授与年月日:2015年3月26日, 学位の種別:課程博士, 審査員:廣田 薫,柴田 崇徳,室伏 俊明,長谷川 修,小野 功
Citation(English)	Degree:, Conferring organization: Tokyo Institute of Technology, Report number:甲第9873号, Conferred date:2015/3/26, Degree Type:Course doctor, Examiner:,,,,
学位種別(和文)	博士論文
Type(English)	Doctoral Thesis

Multi-Robot Behavior Adaptation to Deep Level Environmental Information in Humans-Robots Interaction

Chen Luefeng

Supervisor: Prof. Kaoru Hirota

Doctoral Thesis

Tokyo Institute of Technology

Interdisciplinary Graduate School of Science and Engineering

Department of Computational Intelligence and Systems Science

March 2015

Abstract

Multi-robot behavior adaptation attracts more and more attention due to the rapid development of large-scale human-robot interaction, which consists of numbers of humans and robots, and each robot owns limited capabilities in terms of sensing, computation, adaptation, and communication. Moreover, as cognitive science is widely used in the humans-robots interaction, understanding of deep level environmental information (i.e., emotion, communication atmosphere, and human intention) becomes more and more popular, especially representation of communication atmosphere and human intention understanding.

A multi-robot behavior adaptation mechanism in humans-robots interaction is proposed to adjust robots' behavior to communication atmosphere, where fuzzy production rule based friend-Q learning (FPRFQ) is introduced. It aims to shorten the response time of robots and decrease the social distance between humans and robots to realize the smooth communication of robots and humans. Experiments on robots/humans interaction are performed in a virtual communication atmosphere environment. Results show that robots adapt well by saving 44 and 482 learning steps compared to that by friend-Q learning (FQ) and independent learning (IL), respectively; additionally, the distance between human-generated atmosphere and robot-generated atmosphere is 3 times and 10 times shorter than the FQ and the IL, respectively.

For concerning a large scale humans-robots interaction, which consists of cooperation, neutral, and competition, not only global communication atmosphere, but also local communication atmosphere are existed. a multi-robot behavior adaptation mechanism

based on cooperative-neutral-competitive fuzzy-Q learning (CNCFQ) is proposed for coordinating local communication atmospheres in humans-robots interaction, in which the communication atmosphere is represented by a two-layer fuzzy fusion model and is visualized by shape-color-fill-wave graphics. It aims to realize smooth communication between humans and robots in the local and global communication atmosphere coexistent interaction by decreasing the response time of robots and social distance between humans and robots, as well as visualizing the communication atmosphere. Experiments on multi-robot behavior adaptation are performed in a virtual home party environment. Results show that the proposal saves 47 and 103 learning steps (i.e., the learning rate is increased by 72% and 85%) compared to fuzzy production rule based friend-Q learning (FPRFQ) and friend-Q learning (FQ), respectively; the distance between human-generated atmosphere and robot-generated atmosphere is 3 times and 7 times shorter than the FPRFQ and the FQ, respectively. Additionally, subjective estimation of graphic visualization of the atmosphere through questionnaire obtains 85.5% accuracy for shape, 77.2% for color, 65.3% for fill, and 91.7% for wave.

With the development of deep level information understanding, especially intention understanding becomes necessary and feasible, an intention understanding model based on two-layer fuzzy support vector regression (TLFSVR) is proposed in human-robot interaction, where fuzzy c-means clustering is used to classify the input data, and intention understanding is mainly obtained by emotion, with identification information such as age, gender, and nationality. It aims to realize the transparent communication by understanding customers' order intentions at a bar, in such a way that the social relationship between bar staffs and customers becomes smooth. To demonstrate the aptness of intention understanding model, experiments are designed in term of relationship between emotion-age-gender-nationality and order intention. Results show that the proposal obtains an intention understanding accuracy of 71%/72%/82% with clusters number $C=2/3/6$ (according to different genders/ages/nationalities), which is 25%/26%/36% and 45%/46%/56% higher than that of SVR and back propagation neural networks (BPN-N), respectively; the computational time of proposal is about 0.388s/0.699s/0.845s with

clusters number $C=2/3/6$, while 1.031s for SVR and 2.789s for BPNN. Additionally, the preliminary application experiment is performed in the developing human-robot interaction system, called mascot robot system (MRS), where the experiment is performed in a scenario of “drinking at a bar”, result shows that the bar lady robot obtains an accuracy of 77.7% for understanding customers’ order intentions and receives a satisfaction evaluation of “satisfied”.

After intention understanding, how robots adjust their behavior to human intentions is not only critical for successful communication, but also a new and interesting challenge. A multi-robot behavior adaptation mechanism to human intention is proposed in human-robot interaction, where the information-driven fuzzy friend-Q learning (IDFFQ) is used to generate an optimal behavior-selection policy, and intention understanding is done mainly based on human emotions. It aims to endow robots with the human-oriented interaction capabilities in understanding and adapting their behaviors to human intentions, which decreases the response time of robots by embedding the humans identification information (e.g., religion) for behavior selection, and increases the satisfaction of humans by considering humans deep-level information (e.g., intention and emotion), so as to make interactions run smoothly. Experiments are performed in a scenario of drinking at a bar. Results show that the learning steps of the proposal is 51 steps less than that of the fuzzy production rule based friend-Q learning (FPRFQ), and reports robot time (computational time) is about 1/4 of the time consumed in FPRFQ. Additionally, emotion recognition and intention understanding receive an accuracy of 80.4% and 85.71%, respectively. Moreover, subjective evaluation of customers through questionnaire obtains a satisfaction of “satisfied”.

The proposed behavior adaptation mechanisms are being applied to our developing humans-robots interaction system, called mascot robot system (MRS), and basic experimental results are shown in home party scenario with five eye robots and four humans. According to the preliminary application of intention understanding model, the proposals lay the basic foundations towards behavior adaptation of business communication atmosphere and customers’ intention understanding.

Acknowledgements

This thesis represents not only my work at the keyboard, it is a milestone in more than three years of work at TITECH and specifically within the Hirota laboratory. Since my first day on October 5th, 2011, I have felt at home at TITECH. Throughout these years, I kept challenging myself with unremitting efforts and persistence, and taking every difficulty as the chance to make me better. At the same time, I have learned that there are those who build eye robots and those who use them, and my passion is in establishing the mechanisms used in robot behavior adaptation research. These thesis is the result of work by dozens of people, who I wish to thank. But this thesis is also the result of many experiences I have encountered at TITECH from dozens of remarkable individuals who I also wish to acknowledge. I really cherish these years forever that extremely enriched my life, looking back, a thousand thanks to all those people who have accompanied me through the doctoral period.

First and foremost, I would like to express my sincere gratitude to my advisor and supervisor, Prof. Kaoru Hirota, for not only his continuous support of my Ph.D. study and research, but for also his patience, motivation, enthusiasm, and immense knowledge to help me in my daily life. His guidance helped me a lot in writing journal paper, and I remember he used to say something like "Chen-kun, maybe someday you could make your own theory, named Chen's theorem" to encourage me to create new idea of the research, I could not have imagined having a better advisor and mentor for my Ph.D. study.

My appreciation goes to Prof. Fangyan Dong, for her thoughtful kindness both in my research and daily life all the time, she is so powerful that make everything possible. I also

want to deeply thank Ms. Harumi Hoshino for assisting professors in everyday running of our laboratory.

I would like to appreciate my examiners, Profs. Takanori Shibata, Toshiaki Murofushi, Osamu Hasegawa, and Isao Ono for their precious time, deserve special commendation for their professional comments and criticisms, which helped to shape my research and improve the content of this thesis to a large extent.

I want to deeply thank the robotics group members at Hirota laboratory, Drs. Zhentao Liu, Yongyong Tang, Jiajun Lu. Ever since, Dr. Zhentao Liu has supported me academically and emotionally through the rough road to finish this thesis. Thanks to him I had the opportunity to know Fuzzy Atmosfield (FA). He helped me come up with the thesis topic, and during the most difficult times when writing this thesis, he gave me the moral support and the freedom I needed to move on.

I am also grateful for the past members of Hirota laboratory who deserve my special appreciation include: Drs. Fei Yan, Janet P. Betancourt, Abdullah M. Iliyasu, Bo Sun, Martin L. Tangel, Chastine Fatichah, and Mrs. Zhihong Min, Li Bai. Among the present members of the lab, I must acknowledge the friendship and comradeship in Japan, Drs. Tianyu Li, Kazuhiro Ohnishi, Jesus A. Garcia Sanchez, Toshihiro Akamatsu, Atsushi Shibata, Hidenori Sakaniwa, and I am thankful to Dr. Maslina B. ZOLKEPLI for always patiently helping me English proofreading. I would also like to thank other members at Hirota laboratory, both Japanese and international students, for their continuing support and never ending patience with my constant requests of assistance.

Many thanks to my friends, Dr. Tianwen Lv, and Mr. Chuangping Ye, for helping me stay positive through these years. Their support and consideration helped me overcome frustration and stay focused on my doctoral study.

Last but not the least, no amount of words can convey my deepest gratitude to my family and a special mention to Dr. Min Ding. Their trust, understanding, love, support, and encouragement have strengthened my determination in following my dream.

Contents

1	Introduction	1
2	Acquisition of Deep Level Environmental Information	7
2.1	Introduction	7
2.2	Emotion Recognition	8
2.2.1	Emotion Space	8
2.2.2	Emotion Recognition based on facial expression and speech	9
2.3	Representation of Communication Atmosphere	10
2.3.1	Fuzzy Atmosfield (FA)	10
2.3.2	Communication Atmosphere Representation in the FA by Using Emotion Information	12
2.4	Intention Understanding	14
2.4.1	Two-Layer Fuzzy Support Vector Regression (TLFSVR)	14
2.4.2	Intention Understanding based on TLFSVR by Using Emotion In- formation	18
2.5	Chapter Summary	19
3	Adapting Multi-Robot Behavior to Communication Atmosphere	21
3.1	Introduction	21
3.2	Fuzzy Production Rule based Friend-Q Learning (FPRFQ)	23
3.3	FPRFQ based Multi-Robot Behavior Adaptation Mechanism	25

3.3.1	Architecture of Behavior Adaptation Mechanism	25
3.3.2	Fuzzy Classification	27
3.3.3	FPRFQ based Behavior Adaptation	28
3.3.4	Memory Retrieval for Behavior Adaptation	30
3.4	Experiments on Multi-Robot Behavior Adaptation	31
3.4.1	Environment Setting	31
3.4.2	Experiment on Adaptation to Unfamiliar Atmosphere	33
3.4.3	Experiment on Adaptation to Unfamiliar-Familiar Mixed Atmosphere	35
3.5	Chapter Summary	40
4	Multi-Robot Behavior Adaptation to Local and Global Communication Atmosphere	43
4.1	Introduction	43
4.2	Cooperative-Neutral-Competitive Fuzzy-Q Learning (CNCFQ)	47
4.3	Representation of Local and Global Atmosphere	50
4.3.1	Two-Layer Fuzzy Fusion Model for Atmosphere Generation	50
4.3.2	Graphic Visualization for Atmosphere Representation	52
4.4	CNCFQ based Multi-Robot Behavior Adaptation Mechanism	53
4.4.1	Architecture of Behavior Adaptation	53
4.4.2	Questionnaire for Relationship Identification	55
4.4.3	Fuzzy Classification	56
4.4.4	CNCFQ based Behavior Adaptation	57
4.4.5	Memory Retrieval for Behavior Adaptation	59
4.5	Experiments on Multi-Robot Behavior Adaptation to Local and Global Coexistent Communication Atmosphere	60
4.5.1	Environment Setting	60
4.5.2	Experiment on Adaptation to Local and Global Coexistent Human-Generated Atmosphere	62
4.5.3	Evaluation of Graphic Visualization for Atmosphere Representation	66
4.6	Chapter Summary	70

5	Intention Understanding by Using Emotion-Age-Gender-Nationality Information	73
5.1	Introduction	73
5.2	Characteristics Analysis	76
5.3	Intention Understanding Model	82
5.3.1	Architecture of Intention Understanding Model	82
5.3.2	Information Acquisition	83
5.3.3	TLFSVR based Intention Understanding	85
5.3.4	Memory Retrieval for Intention Understanding	86
5.3.5	Intention generation by fuzzy inference	87
5.4	Experiments on Intention Understanding of Costumer's Order at a Bar . .	87
5.4.1	Environment Setting	87
5.4.2	Data perpetration	89
5.4.3	Experiment on Intention Understanding of Costumer's Order at a Bar	92
5.4.4	Experiment by Using Mascot Robot System	97
5.5	Chapter Summary	102
6	Multi-Robot Behavior Adaptation to Humans' Intention	105
6.1	Introduction	105
6.2	Information-Driven Fuzzy Friend-Q Learning (IDFFQ)	107
6.3	IDFFQ based Multi-Robot Behavior Adaptation Mechanism	111
6.3.1	Architecture of Behavior Adaptation Mechanism	111
6.3.2	Information Processing of Emotion and Intention	112
6.3.3	IDFFQ based Behavior Adaptation	113
6.4	Experiments on Behavior Adaptation to Costumers' Order Intention at a Bar	116
6.4.1	Experiment Setting	116
6.4.2	Data Preparation	118
6.4.3	Experiment on Emotion-Intention Understanding	119

6.4.4	Experiment on Behavior Adaptation to Costumers' Order Intention at a Bar	121
6.5	Chapter Summary	126
7	Conclusion	129
7.1	Conclusion of Multi-Robot Behavior Adaptation to Deep Level Environ- mental Information	129
7.2	Future Perspectives of Multi-Robot Behavior Adaptation Mechanisms and Potential Application	134
	Bibliography	136
	Related Publications	149

List of Figures

1.1	Outline of Thesis.	6
2.1	The APA emotion space.	9
2.2	Scheme of the emotion recognition.	10
2.3	FA represented in a 3D fuzzy cubic space.	11
2.4	Scheme of the FA.	12
2.5	Two-layer fuzzy support vector regression (TLFSVR).	17
2.6	Scheme of emotion based intention understanding.	19
3.1	Fuzzy production rule based friend-Q learning (FPRFQ).	24
3.2	FPRFQ based adaptation mechanism.	26
3.3	Membership functions of 3D FA axes.	27
3.4	Division of FA.	32
3.5	Average steps to complete once task.	34
3.6	Average reward to complete once task.	35
3.7	Social distance to complete once task.	35
3.8	Average steps to complete once task.	36
3.9	Average reward to complete once task.	37
3.10	Social distance to complete once task.	37
3.11	Candidate areas of emotional behaviors for atmosphere (VFR, VL, VCA).	38
3.12	Scenario of “greeting guests at the door”.	39
3.13	Scenario of “query of train timetable”.	40

3.14 Scenario of “watching TV”.	40
4.1 Fuzzy Atmosfield expressed in 3D space.	45
4.2 Cooperative-neutral-competitive fuzzy-Q learning (CNCFQ).	48
4.3 Two-layer fusion model of atmosphere generation.	51
4.4 Membership functions of intensity.	52
4.5 Visualization of atmosphere.	54
4.6 Architecture of CNCFQ based behavior adaptation mechanism.	55
4.7 Membership functions of 3D FA axes.	56
4.8 Virtual environment.	61
4.9 Division of FA.	61
4.10 Average steps to complete one task.	64
4.11 Average reward to complete one task.	64
4.12 Social distance to complete one task.	65
4.13 International communication.	69
4.14 Network system of MRS.	70
5.1 Identification information of volunteers.	76
5.2 Presentation of brainstorming.	77
5.3 Correlation between influence factors and intention.	79
5.4 Dendrogram of intention related factors relationship.	80
5.5 Boxplots of intention correlation by age/gender/nationality.	82
5.6 TLFSVR based intention understanding model.	83
5.7 Pleasure-Arousal emotion plan.	84
5.8 Network system of MRS.	89
5.9 Information acquisition.	90
5.10 Emotion recognition.	91
5.11 Training data for simulation.	92
5.12 Testing data for simulation.	93

5.13	Output curves and output error of proposed TLFSVR, SVR, and BPNN with $C=2$	94
5.14	Output curves and output error of proposed TLFSVR, SVR, and BPNN with $C=3$	95
5.15	Output curves and output error of proposed TLFSVR, SVR, and BPNN with $C=6$	96
5.16	Scenario of “drinking at a bar”.	97
5.17	Boxplots of responses by situations.	99
5.18	Comparative experiment between simulation and MRS.	102
6.1	Information-driven fuzzy friend-Q learning (IDFFQ).	109
6.2	IDFFQ based behavior adaptation mechanism.	111
6.3	Membership function of emotion plane.	115
6.4	Volunteers’ emotion.	118
6.5	Training data for intention understanding.	119
6.6	Testing data of 8 volunteers for intention understanding.	120
6.7	Searching map.	120
6.8	Result of emotion recognition for 8 volunteers.	121
6.9	Result of intention understanding for 8 volunteers.	121
6.10	Average steps to complete a task.	123
6.11	Average reward to complete a task.	123
6.12	Average computational time to complete a task.	124
6.13	MRS network system.	126

List of Tables

3.1	Result comparison of adaption to unfamiliar atmosphere	34
3.2	Result comparison of adaption to unfamiliar-familiar mixed atmosphere . .	36
4.1	Results comparison for adaptation to local and global atmosphere coexis- tent communication	65
4.2	Questionnaire for evaluation of graphic visualization of communication at- mosphere	66
4.3	Questionnaire for comparison between two kinds of graphic visualization .	68
5.1	Result analysis of brainstorming	78
5.2	Questionnaire for correlation between influences factors and intention . . .	78
5.3	Basic information of volunteers	79
5.4	Correlation coefficients among related factors	81
5.5	Data samples of ANOVA	81
5.6	ANOVA summary of (a)/(b)/(c)	81
5.7	Look-up table for output	87
5.8	Questionnaire for relationship between emotion and intention to order at a bar	90
5.9	Analysis of variance	92
5.10	Result comparison of simulation for intention understanding	95
5.11	Identification information and emotion-intention information of volunteers	97
5.12	Questionnaires for customer feedback	98

5.13	Response of questionnaires	99
5.14	ANOVA summary of (a)/(b)	100
5.15	Result of application for intention understanding	100
5.16	Result analysis of Comparative experiment	101
6.1	Results comparison for adaptation to 8 volunteers' intention	124
6.2	Questionnaire for customer satisfaction level	124

Chapter 1

Introduction

As a relatively young discipline, human-robot interaction (HRI) attracts a lot of attention due to the increasing availability of complicated intelligent robots. With the rapid development of robotics, especially the humanoid robots, people's exposure to these kinds of robots in daily lives is becoming more and more popular, e.g. robotic toys at home. Moreover, robots are increasingly being developed to the real world application areas, such as household appliances (robotic vacuum cleaners or lawn movers), robotic shopping guider in supermarket, elder care robots, disaster relief robots, and other assistive or educational applications. Under these circumstances, interaction between people and robots becomes increasingly socially situated and multi-faceted [1], which indicates not only the benefits of having robots interact with people, but also the need for the interactions to be smooth and natural.

Due to the requirement of more intelligent daily life, as well as the emergence of the humanoid robots, the scale of human-robot interaction becomes bigger and bigger. A large-scale human-robot interaction consist of numbers of humans and robots, where each robot owns limited capabilities in terms of sensing, computation, adaptation, and communication. In our daily communication, humans may subconsciously adapt their behaviors to communication partners, comparably, in human-robot interaction, not only humans but also robots are expected to adapt to their partners, thus behavior adaptation is being intensively studied [2][3]. In this case, robots need to recognize and respond appropriately to humans' social signals including speech, gesture, and facial expression

[4][5][6][7]. Reinforcement learning [8, 9] is widely used in robots behavior adaptation according to individual gesture [10], speech [11], and environment objects [12].

With the development of large-scale human-robot interaction, there are a lot of environmental information during the communication, including surface level environmental information and deep level environmental information. Some kinds of information can be audible and visible, even directly detectable by sensors, these kind of information are defined as surface level environmental information, such as location, speech, gesture, facial expression, and so on. However, Some kinds of information are inaudible and invisible, they are usually complex physiological/psychological phenomena of humans and need to comprehensive by using cognitive ability, these kinds of information are defined as deep level environmental information, such as emotion, communication atmosphere, intention, and so on. Most of the research focus on speech recognition, gesture recognition, facial expression recognition [13, 14, 15], and the deep level information need to pay more attention.

In the area of behavior adaptation to deep level environmental information, human emotional state based robot behavior adaptation attracts a lot of attention [16][17]. Emotion interaction module based HRI framework is used to robot behavior adaptation in [18], where cognitive interaction is used for cooperative sharing of tasks. In [19], a comprehensive, time-varying affective model is designed for adapting the robot behavior, according to generation and adaptation of emotion. In our previous researches, surface level environmental information based emotion recognition [20][21][22]and concept of atmosphere [23][24] are introduced to apply in the humans-robots interaction, but the communication between humans and robots are still not so smooth, some problems are given as follow,

- Most of the research focus on one robot to one human interaction, not so much about the multi-robot behavior adaptation in humans-robots interaction.
- Although the emotion recognition catches a lot of attention, other two kinds of deep level environmental information, i.e., communication atmosphere and human's intention, are seldom mentioned in acquisition, especially human intention understanding.

- Robot behavior adaptation is mainly concerned with surface level environmental information, i.e., speech, gesture, and facial expression, with the introduction of the deep level environmental information, i.e., emotion, atmosphere, and intention, how the robots do behaviors adapt to the humans based on the deep level environmental information is a new challenge and very interesting topic.

The thesis aims to adapt robots' behavior to humans by adequately understanding of humans' deep level environmental information, in such a way that the communication between humans and robots becomes more smoothly. It is organized as follows,

In Chapter 2, acquisition of deep level environmental information is introduced, including the emotion, communication atmosphere, and human intention. At first, emotion recognition is realized using weighted fusion fuzzy inference based on the bimodal information, i.e., facial expression and speech. Emotion states are represented in the Affinity-Pleasure-Arousal (APA) space. Then, the concept of the Fuzzy Atmosfield (FA) is introduced to represent the communication atmosphere based on the individual emotion. In addition, an intention understanding model based on the two-layer fuzzy support vector regression (TLFSVR) is introduced to deeply comprehend people's inner thoughts using individual emotion.

In Chapter 3, a behavior adaptation mechanism in humans-robots interaction is proposed to adjust robots' behavior to communication atmosphere, where fuzzy production rule based friend-Q learning (FPRFQ) is introduced. Firstly, A fuzzy production rule based friend-Q learning method (FPRFQ) is introduced to handle the uncertainties and continuous applications, i.e., human-robot interaction, in which the discrete lookup Q-table in friend-Q learning is replaced by fuzzy production rules with Q-values, and a behavior generating approach based on nearby exploration is used. Based on the FPRFQ, a behavior adaptation mechanism is proposed to solve the robots' behavior adaptation problem. The communication atmosphere information represented in the FA [25] is used as the input for the reward function, where it is defined based on the Euclidean distance between human-generated atmosphere and robot-generated atmosphere in such a way that the reward increases as the distance decreases. Then each robot dynamically adjusts

its behavior to the communication atmosphere by maximizing the reward.

In Chapter 4, multi-robot behavior adaptation mechanism based on cooperative-neutral-competitive fuzzy-Q learning(CNCFQ) is proposed for coordinating local communication atmospheres in humans-robots interaction, in which the communication atmosphere is represented by a two-layer fuzzy fusion model and is visualized by shape-color-fill-wave graphics. First, A cooperative-neutral-competitive fuzzy-Q learning (CNCFQ) method is introduced to handle the complex communication, i.e., cooperation, neutral, and competition mixed humans-robots interaction, in which the maximum based updating Q function (i.e., in the FPRFQ [26]) is replaced by the maximum-minimax based updating Q functions and the relationships among agents are identified as friends, neutral, or foes. Based on CNCFQ, a behavior adaptation mechanism is proposed to multiple robots for adapting behaviors to the local and global coexistent communication atmosphere. The communication atmosphere represented in the FA [23] is generated by using the two-layer fuzzy fusion model (i.e., local fusion layer and global fusion layer), and the global communication atmosphere is used as the input for the reward function, where it is defined based on the Euclidean distance between global human-generated atmosphere and global robot-generated atmosphere in such a way that the reward increases as the distance decreases. Using the proposed mechanism, robots firstly identify the opponents as friends, neutral, or foes, then robots dynamically adjust their behavior to the communication atmosphere by maximizing the reward based on the maximum algorithm (in the case of cooperation or neutral) or minimax algorithm (in the case of competition).

In Chapter 5, an intention understanding model based on two-layer fuzzy support vector regression (TLFSVR) is proposed in human-robot interaction, where fuzzy c-means clustering is used to classify the input data, and intention understanding is mainly obtained by emotion, with identification information such as age, gender, and nationality. Since unfamiliar issues becomes familiar through cognitive process, robot is firstly able to distinguish between familiar and unfamiliar people. In the case of familiar people, robot may understand his/her intention through memory, where the memory retrieval is based on the fuzzy production rules; in the case of unfamiliar people, an intention understanding

model based on the two-layer fuzzy support vector regression (TLFSVR) is introduced to deeply comprehend people's inner thoughts. The TLFSVR is an extension of support vector regression (SVR) [14], in which two-layer consists of local learning layer and global learning layer. Local learning layer includes fuzzy c-means based input data classification and SVR based learning, and global learning layer outputs the intention of TLFSVR based on fuzzy weighted average algorithm. In view of the local learning [27], the training data is split into multiple subsets, so that the traditional one SVR is extended to multiple SVRs, for example, considering various people with different ages, different genders, and different nationalities, which may lead to the problem of health, gender hobby, religion, and so on. Different regression models reflect different ages/genders/nationalities, which is better than only one regression model covers all situations. Moreover, age, gender, and nationality play important role in intention understanding, for example they affect people's intention to purchase products or services [28], accept new things [29], and handle risky situation [30]. Thus, three components of human property, i.e., age, gender, and nationality, are used for intention understanding beside emotion, then training data can be classified according to the identification informations. Finally, fuzzy inference based on identification informations is used to generate the intention, which is either output of memory retrieval or the TLFSVR.

In Chapter 6, a multi-robot behavior adaptation mechanism to human intention is proposed in human-robot interaction, where the information-driven fuzzy friend-Q learning (IDFFQ) is used to generate an optimal behavior-selection policy, and intention understanding is done mainly based on human emotions. First, an information-driven fuzzy friend-Q learning (IDFFQ) method is introduced, where Q-values based behavior selection is replaced by mixed information (including identification information such as religion, as well as Q-values) based behavior selection. Based on the IDFFQ, a behavior adaptation mechanism is proposed to multiple robots for adapting behaviors to peoples' intention. Intention is understood by using the two-layer fuzzy support vector regression model (TLFSVR) in [31] and reward function reflects the adaptation level of robots, it is defined not only according to the expected location transformed from deep-level information (i.e.,

intention and emotion), but also the human satisfaction level, in such a way that, reward raises as location approaches and satisfaction increases. Using the proposed mechanism, information robots firstly collect people identification information (i.e., age, gender, nationality, and religion) and understand deep-level information (i.e., emotion and intention), then task robots dynamically adjust their behaviors to intention by maximizing reward based on deep-level information and Q-values or only Q-values.

In Chapter 7, the proposals and related works in the thesis are summarized, in addition, new extensions and applications are discussed for the future work.

The road map visualizes the dependence among the individual chapters, and summarizes the thesis organization as shown in Figure 1.1.

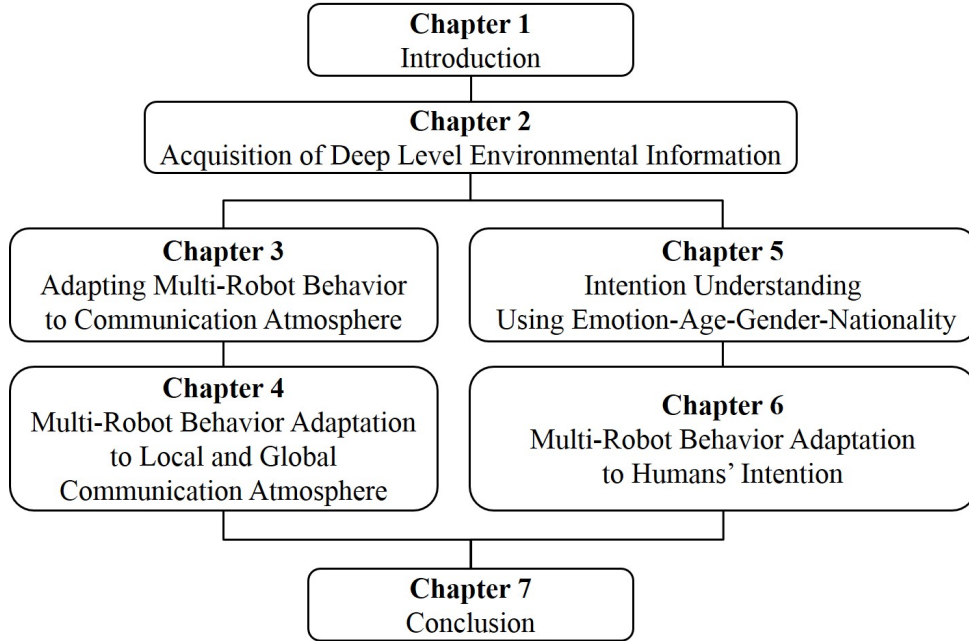


Figure 1.1: *Outline of Thesis.*

Chapter 2

Acquisition of Deep Level Environmental Information

2.1 Introduction

With the development of large scale human-robot interaction, i.e., one human to one robot communication is extending to multi-human to multi-robot interaction, there are a lot of environmental information during the communication, including surface level environmental information and deep level environmental information. Some kinds of information can be audible and visible, even directly detectable by sensors, these kind of information are defined as surface level environmental information, such as location, speech, gesture, facial expression, and so on. However, Some kinds of information are inaudible and invisible, they are usually complex physiological/psychological phenomena of humans and need to comprehensive by using cognitive ability, these kinds of information are defined as deep level environmental information, such as emotion, communication atmosphere, intention, and so on. Most of the research focus on speech recognition, gesture recognition, facial expression recognition [13, 32, 15], and the deep level information need to pay more attention.

To recognize human emotion, multiple channels of information are used, including facial expression, speech, gesture, and so on [33, 34, 35]. In consideration of small scale human-robot interaction (i.e., one to one or several to several), some attempts try

to define the communication atmosphere [36, 37], as well as recognition of communication atmosphere [38]. In the case of large scale communication (i.e., many to many), concept of Fuzzy Atmosfield (FA) is proposed to represent communication atmosphere and estimated by the flexible fusion of individual emotion in our previous works [23, 24]. Intention understanding is essential for many aspects of cognition, culture and social, including image intention understanding [39], speech intention understanding [40], and gesture intention understanding [32]. Although speech and gesture play important roles in intention understanding, they cannot intuitive reflect people's internal states. Emotion is an important missing link in the intention-behavior gap [41], and it is essential for social behavior [42]. How to acquire the communication atmosphere and intention based on emotion information are signification and useful for smooth human-robot interaction.

Section 2.2 shows the Affinity-Pleasure-Arousal (APA) emotion space, in addition, emotion recognition is given based on the facial expression and speech. Concept of Fuzzy Atmosfield (FA) is introduced to represent communication atmosphere in section 2.3, as well as the estimation of communication atmosphere by the flexible fusion of individual emotion. In section 2.4, An intention understanding model is proposed to deeply comprehend people's inner thoughts according to the emotion.

2.2 Emotion Recognition

2.2.1 Emotion Space

Emotion space is commonly used as a tool for quantitative analysis of emotion, and it can express emotional states in two dimensions, three dimensions, and so on. To express the emotional states of humans, a 3D emotion space, i.e., Affinity-Pleasure-Arousal (APA) emotion space [43] is employed, as shown in Fig. 2.1. Using APA emotion space, it is not only possible to express fixed emotional states but also takes in consideration rapid variations of the emotional state due to time involving communication [43].

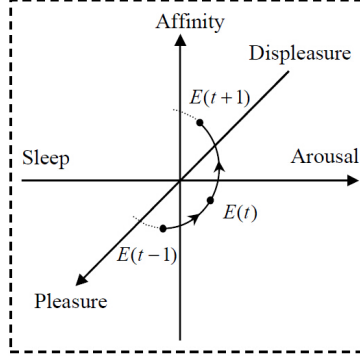


Figure 2.1: The APA emotion space.

The emotional state in APA emotion space is defined as

$$E = (e_{affinity}, e_{pleasure}, e_{arousal}), \forall e \in [-1, 1] \quad (2.1)$$

2.2.2 Emotion Recognition based on facial expression and speech

For understanding deep level environmental information, emotion recognition is a supported module based on our previous study [22, 24], which is realized using weighted fusion fuzzy inference based on the bimodal information, i.e., facial expression and speech, as shown in Fig. 2.2. Emotion is recognized by fusing semantic cues from speech and facial expression, and represented in the Affinity-Pleasure-Arousal (APA) space, mainly consisting of three steps:

Step 1: Initial state of emotion is recognized based on the facial expression and semantic cues from speech performed by communicators using weighted fusion.

Step 2: Starting from the initial/previous state of emotion, transition of emotional states in APA emotion space is calculated based on acoustic features of speech by using a fuzzy inference system.

Step 3: At the sampling time $t + 1$, if the scene changes, then jump to *Step 1*, or else jump to *Step 2* to calculate the transition of emotional state.

The number of sub-steps of *Step 3* is based on duration of a scene, for example, there are five sub-steps in a scene that lasts ten seconds, in which the sampling period is two seconds. Thus, scene change is important for determining when to recalculate initial state

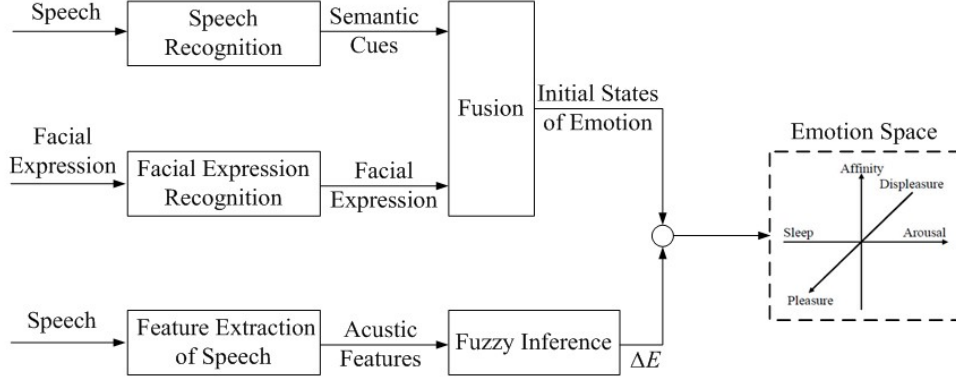


Figure 2.2: Scheme of the emotion recognition.

of an emotion, i.e., jump to *Step 1*. And in one scene, emotional state can be calculated as

$$E(t) = \begin{cases} E_{Initial} + \Delta E, & t = 1 \\ E(t-1) + \Delta E, & t > 1 \end{cases} \quad (2.2)$$

where $E_{Initial}$ is initial state of emotion, ΔE is the transition of emotional state. And if $e_{iInitial} + \Delta e_i > 1$ or $e_i(t-1) + \Delta e_i > 1$, then $e_i(t) = 1$; if $e_{iInitial} + \Delta e_i < -1$ or $e_i(t-1) + \Delta e_i < -1$, then $e_i(t) = -1$, i.e., “affinity”, “pleasure”, or “arousal”.

Six primary emotions [36], i.e., “happiness”, “surprise”, “fear”, “anger”, “disgust”, and “sadness”, together with “neutral” are employed as the initial state of emotion in *Step 1*.

2.3 Representation of Communication Atmosphere

2.3.1 Fuzzy Atmosfield (FA)

The term Atmosfield (a contraction from the words “atmosphere” and “field”) is proposed to represent the existence of the space, which although invisible, is supposed to exist by occupying space and containing energy. The Atmosfield reflects not only the state of the atmosphere surrounding us during interactive communication, but also affect the emotional states of humans to some extent. Therefore, it is defined as a kind of

“psychological field”, producing psychological feelings that can influence the process and results of human behavior.

Due to the psychological characteristics of the Atmosfield, it is difficult to establish an exact mathematical model of it as some classical fields such as electrical or magnetic field. According to [23], taking the uncertainty and vagueness of the atmosphere into account, which needs subjective comprehension to determine its attributes and fuzzy logic to determine the reasoning from atmosphere related factors to the FA. All the atmospheres that exist in real life could be represented by the three pairs of attributes of the FA, and also by a 3D fuzzy cubic space of the FA, as shown in Fig. 2.3.

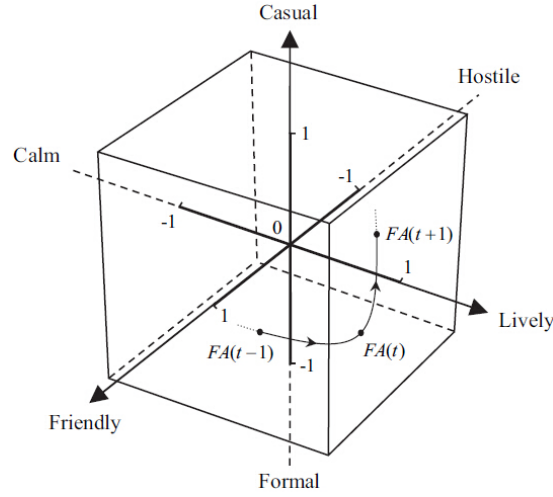


Figure 2.3: FA represented in a 3D fuzzy cubic space.

The state of the FA in this 3D fuzzy cubic space is represented as

$$FA = (a_{friendly}, a_{lively}, a_{casual}), \forall a \in [-1, 1] \quad (2.3)$$

where FA is the state of the FA, and $a_{friendly}$, a_{lively} , and a_{casual} are the values on “Friendly-Hostile”, “Lively-Calm”, and “Casual-Formal” axes, respectively. They represent extremely friendly, extremely lively, and extremely casual, respectively, when the value on each axis is equal to 1. When the value on each axis is equal to -1, they represent extremely hostile, extremely calm, and extremely formal, respectively. And the value of FA at the origin represents the neutral state.

2.3.2 Communication Atmosphere Representation in the FA by Using Emotion Information

The FA based on emotional states using fuzzy inference system and expert rules is introduced, where the scheme of the FA is illustrated in Fig. 2.4. There are three steps to calculating and displaying the state of the FA. First of all, the emotional states (represented in the APA emotion space) of all participants involved in the interactive communication are obtained by emotion recognition system. Secondly, the information fusion module of emotional states is proposed. Here, fuzzy inference system is employed to establish the fuzzy relationship between the FA and the emotion space, and with the utilization of speech volume, expert rules are arranged to realize the mapping from n individual emotion to the FA. Finally, the state of the FA is calculated and represented in the 3D fuzzy cubic space.

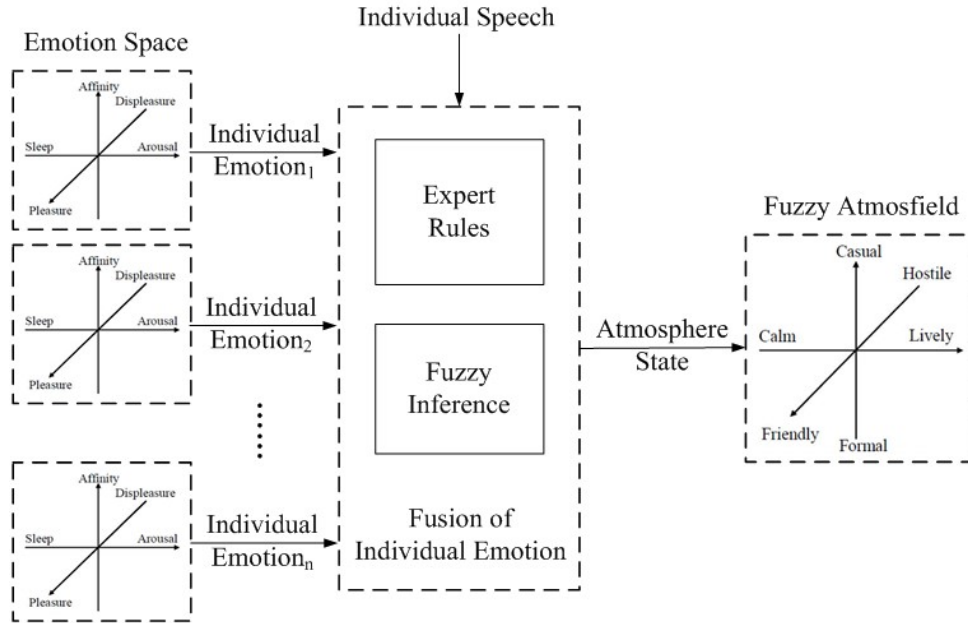


Figure 2.4: Scheme of the FA.

To construct the model of the FA in multi-participant interaction, three hypotheses are put forward as:

- The FA is supposed to be composed of individual emotion, environmental factors such as background music and other atmosphere-related factors involved during the

communication, all of which can be represented by emotional states in the APA emotion space. In addition, the emotional states are assumed to be independent, namely, the effects of mutual coupling among the emotional states are neglected for estimating the FA .

- The variation of the states of the FA is a continuous process, and the FA is time-dependent as well as the emotional state, which can be represented by $FA(t-1)$, $FA(t)$, and $FA(t+1)$ as shown in Fig. 2.3.
- When no new emotional element comes up, the intensity of the FA will decrease (i.e. it will not disappear instantaneously) until finally it reaches the neutral state as time goes by.

As mentioned above, the state of the FA is supposed to be a continuous and time-related variable, in other words, at time t , the transfer of the state of the FA is not only related to the state of the FA at time $t-1$, but also to the new emotional states of communicators stimulated by environmental factors. Therefore, the formula of the FA can be defined as

$$E(t) = \begin{cases} f(E_1(t), \dots, E_n(t)), & t = 1 \\ (1 - \lambda)FA(t-1)\gamma + \lambda f(E_1(t), \dots, E_n(t)), & t = 2, 3, \dots, m \end{cases} \quad (2.4)$$

where f is the function of emotional states E at time t ; n is the number of individuals involved in the communication scenario; γ is a monotonically decreasing function, $0 \leq \gamma \leq 1$, such as an exponential function $\exp(-kT)$, k is a positive parameter defining the decreasing speed of $FA(t-1)$, and T is the sampling period for calculating the FA ; λ is the correlation factor, $0 \leq \lambda \leq 1$; when $t = 0$, FA is in the origin of the FA space as the initial state.

2.4 Intention Understanding

2.4.1 Two-Layer Fuzzy Support Vector Regression (TLFSVR)

1. Support vector regression (SVR)

As a powerful and promising learning machine, support vector machines (SVMs) is famous for solving problems characterized by small samples, nonlinearity, high dimension, and local minima [44][45][46]. SVM employs structural risk minimization (SRM) principle to achieve better generalization ability, so that it provides higher performance than traditional empirical risk minimization (ERM) based learning machines, e.g., neural networks [47]. There are two main categories for SVM, which are support vector classification (SVC) and support vector regression (SVR). The goal of SVC is to construct an optimal separating hyperplane in a higher dimensional space by maximizing the margin between the separating hyperplane and classification data. Maximizing the margin is a quadratic programming (QP) problem and can be solved from its dual problem by introducing Lagrangian multipliers. Using the dot product functions called kernels, SVC finds the optimal hyperplane in feature space, which means kernel functions play a vital role in classification since they determine feature spaces where data sample are classified, and it may directly affect performances of the SVM classification. A combination of a few input points called support vectors can be used to generate the solution of the optimal hyperplane.

With the introduction of ε -insensitive loss function, the SVM is extended to solve regression problems, called support vector regression (SVR) [29]. While SVM is mostly used to perform classification by determining the maximum margin separation hyperplane between two classes, SVR tries the inverse, i.e., to find the optimal regression hyperplane so that most training samples lie within an ε -margin around this hyperplane. SVR is applied to various applications such as optimal control [48], time-series prediction [49], interval regression analysis [50]. Suppose training data set is a set of l pairs of vectors and the set is described as $D = \{(X_i, y_i) | X_i \in R^n, y_i \in R, i = 1, 2, \dots, l\}$, Considering

the simplest case of linear SVR, the basic function is given as

$$f(X) = \langle W, X \rangle + b \quad (2.5)$$

where the function $f(X)$ involves taking the dot product of the input vector X with a model parameter W and adding the scalar bias b .

The aim of ε -SVR is to find a function $f(X)$ that has at most ε deviation from the actually obtained targets y_i for all the training data, and at the same time as flat as possible by minimizing the norm of W . The problem can be described as convex optimization problem and written as

$$\begin{aligned} \min \quad & \frac{1}{2} \|W\|^2 \\ \text{subject to} \quad & y_i - \langle W, X_i \rangle - b \leq \varepsilon \\ & \langle W, X_i \rangle + b - y_i \leq \varepsilon. \end{aligned} \quad (2.6)$$

Sometimes constraints may not always be satisfied, that means not all pairs (X_i, y_i) fit with ε precision. Slack variables ξ_i and ξ_i^* are introduced to cope with the otherwise infeasible constraints, and Eq. (2.6) becomes

$$\begin{aligned} \min_{W, b, \xi_i, \xi_i^*} \quad & \frac{1}{2} \|W\|^2 + \gamma \sum_{i=1}^l (\xi_i + \xi_i^*) \\ \text{subject to} \quad & y_i - \langle W, X_i \rangle - b \leq \varepsilon + \xi_i \\ & \langle W, X_i \rangle + b - y_i \leq \varepsilon + \xi_i^* \\ & \xi_i, \xi_i^* \geq 0, i = 1, 2, \dots, l \end{aligned} \quad (2.7)$$

where γ is a scalar regularization parameter that penalizes the amount of slack used (smaller values of γ yield more outlier rejection).

Retrieving the optimal hyperplane in Eq. (2.7) is a QP problem, which can be solved by constructing a Lagrangian and transformed into the dual as Eq. (2.8), where α_i and α_i^* are the Lagrange multipliers. All the dot product in Eq. (2.8) can be replaced by kernel function and the regression function is reconstructed as

$$\begin{aligned}
 & \max_{\alpha, \alpha^*} \quad -\frac{1}{2} \sum_{i,j=1}^l (\alpha_i - \alpha_i^*) (\alpha_j - \alpha_j^*) \langle X_i, X_j \rangle - \varepsilon \sum_{i=1}^l (\alpha_i + \alpha_i^*) + \sum_{j=1}^l y_j (\alpha_j - \alpha_j^*) \\
 & \text{subject to} \quad \sum_{i=1}^l (\alpha_i - \alpha_i^*) = 0 \quad \text{and} \quad \alpha_i, \alpha_i^* \in [0, \gamma]
 \end{aligned} \tag{2.8}$$

$$f(X) = \sum_{i=1}^l (\alpha_i - \alpha_i^*) K(X_i, X) + b \tag{2.9}$$

where the kernel function is the Gaussian kernel,

$$K(x_i, x_j) = \exp\left(-\frac{(x_i - x_j)^2}{2\sigma^2}\right). \tag{2.10}$$

2. Two-layer fuzzy support vector regression (TLFSVR)

SVR is being an extensive application in the field of human-robot interaction [51], and it meets some challenges, one is the difficulty to guarantee the good local modeling performance of the obtained model, which means only one regression hyperplane cannot be guaranteed most training samples lie within it. For example, a training data sample is collected from different ages, considering that people in different ages may have their own hobbies, different regression models reflect different ages maybe better than only one regression model that involves all the ages. Local learning algorithm is attempt to locally adjust the capacity of the model to the properties of the training data subset [30], and based on the idea of local learning, two-layer fuzzy support vector regression (TLFSVR) is proposed to handle the evenly distributed problem that the training data sample cover the entire array. TLFSVR is an extension of the SVR, where local learning is fused with the SVR, as well as fuzzy logic, and two-layer includes local learning layer and global learning layer, as shown in Fig. 2.5.

First, the training data set are split into several subsets by using the fuzzy c-means (FCM) algorithm [49, 50]. The algorithm is based on the minimization of the following objective function,

$$f(U, V) = \sum_{i=1}^l \sum_{k=1}^C (\mu_{ik})^m \|X_i - V_k\|^2 \tag{2.11}$$

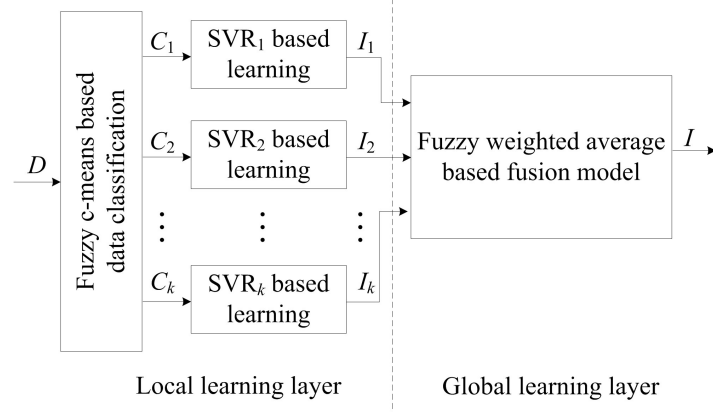


Figure 2.5: Two-layer fuzzy support vector regression (TLFSVR).

where $l \in N$ is the number of training data, $C \in [2, l)$ is the number of clusters, $m \in [1, +\infty)$ is a weighting exponent, $\mu_{ik} \in [0, 1]$ is the degree of membership of training data X_i in the cluster k , V_k is the center of cluster k . Approximate optimization of Eq. (2.11) by the FCM algorithm is based on iteration, with the update of membership and the cluster centers V_k by

$$\mu_{ik} = \left(\sum_{k=1}^C \left(\frac{d_{ij}}{d_{ik}} \right)^{\frac{2}{m-1}} \right)^{-1}, 1 \leq i \leq l, 1 \leq j \leq C \quad (2.12)$$

$$V_k = \frac{\sum_{i=1}^l (\mu_{ik})^m X_i}{\sum_{i=1}^l (\mu_{ik})^m}, 1 \leq j \leq C \quad (2.13)$$

where $d_{ik} = \|X_i - V_k\|$.

The iteration will stop when termination criterion is satisfied, i.e., $\|X_i - V_k\| \leq \epsilon$, where $V = (v_1, v_2, \dots, v_C)$, and ϵ is the given sensitivity threshold.

Then the spread width is calculated as

$$\delta_k^o = \sqrt{\frac{\sum_{i=1}^l (\mu_{ik})^m \|x_i^o - v_i^o\|^2}{\sum_{i=1}^l (\mu_{ik})^m}}, 1 \leq k \leq C, 1 \leq o \leq N. \quad (2.14)$$

According to the obtained centers and spread width, the training data sample is split

$$D_k = \{(X_i, y_i) \mid v_k^o - \eta\delta_k^o \leq x_i^o \leq v_k^o + \eta\delta_k^o\}, 1 \leq i \leq l, 1 \leq o \leq N, 1 \leq k \leq C \quad (2.15)$$

into training subsets as Eq. (2.15), where η is a constant for controlling the overlap region of the training subsets and the size of training subsets increases as η increases, as well as computational time.

Second, regression function of each cluster is constructed by Eq. (2.9) as follow

$$SVR_k = \sum_{i=1}^{l_k} (\alpha_{i,k}^* - \alpha_{i,k}) k(X_i, X) + b_k, X \in D_k, 1 \leq k \leq C \quad (2.16)$$

where l_k denotes the number of training data in the k th training subset, the parameters $\alpha_{i,k}^*$, $\alpha_{i,k}$ and b_k are obtained by the SVR for the k th training subset.

Finally, after the local learning, global learning for output is calculated based on the fuzzy weighted average algorithm [52] fusion model with triangle membership functions and SVRs. The membership function of the fuzzy weighted average algorithm is built as

$$\begin{aligned} A_k(X_i) &= a_k^1(x_i^1) a_k^2(x_i^2) \cdots a_k^o(x_i^o) a_k^o(x_i^o) \\ &= \max \left(\min \left(\frac{x_i^o(v_k^o - \eta\delta_k^o)}{v_k^o - (v_k^o - \eta\delta_k^o)}, \frac{(v_k^o + \eta\delta_k^o) - x_i^o}{(v_k^o + \eta\delta_k^o) - v_k^o} \right), 0 \right), 1 \leq k \leq C \end{aligned} \quad (2.17)$$

where $A_k(X_i)$ is the fuzzy weight of the k th SVR.

Then, the global output of the proposed TLFSVR is based on the fusion model which is calculated using fuzzy weighted average algorithm as

$$y_i(X_i) = \frac{\sum_{k=1}^C A_k(X_i) SVR_k(X_i)}{\sum_{k=1}^C A_k(X_i)}, 1 \leq i \leq l, 1 \leq k \leq C. \quad (2.18)$$

2.4.2 Intention Understanding based on TLFSVR by Using Emotion Information

Based on the proposed TLFSVR, scheme of intention understanding is shown in Fig. 2.6, mainly consisting of four steps:

Step 1: Individual emotion state E_i is recognized based on the facial expression and semantic cues from speech performed by communicators as introduced in 2.2.

Step 2: Clustering the similar emotion E_i by using FCM algorithm according to Eqs. (2.12,2.13,2.14), then n clusters of emotion C_{En} is obtained as shown in Eq. (2.15).

Step 3: Intention understand of each cluster is learned by using SVR according to Eq. (2.16).

Step 4: Intention understanding of the TLFSVR is outputted base on the fuzzy weighted average algorithm according to Eq. (2.17,2.18).

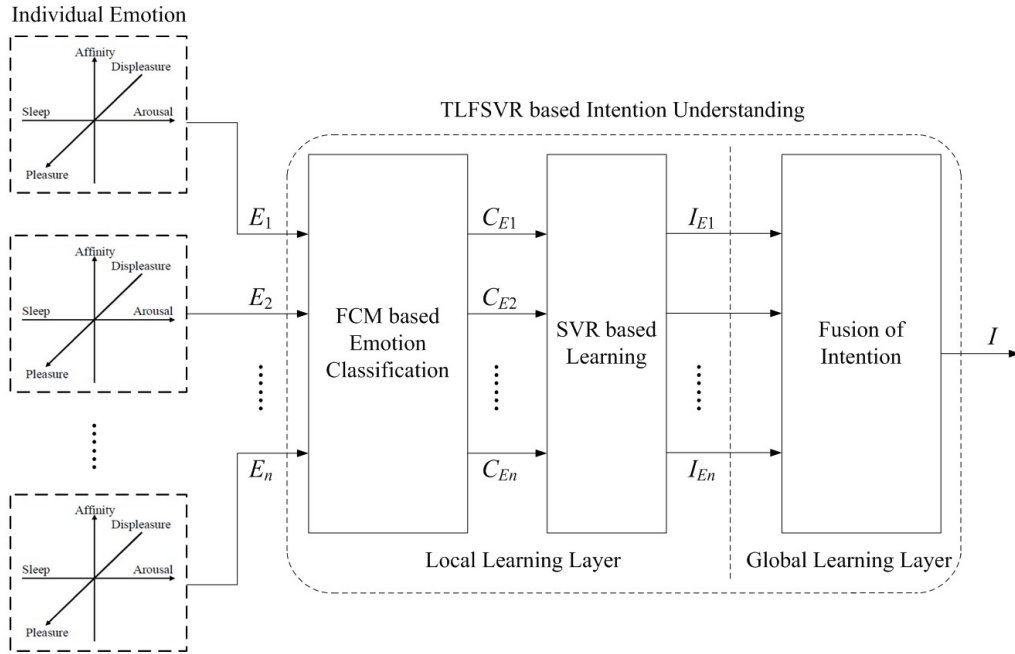


Figure 2.6: Scheme of emotion based intention understanding.

2.5 Chapter Summary

To understand deep level environmental information, especially communication atmosphere and intention, emotion recognition is realized using weighted fusion fuzzy inference based on the bimodal information, i.e., facial expression and speech. Emotion states are represented in the Affinity-Pleasure-Arousal (APA) space, six primary emotions, i.e., “happiness”, “surprise”, “fear”, “anger”, “disgust”, and “sadness”, together with “neutral” are employed as the initial state of emotion. Moreover, Emotion recognition is a

supported module for acquiring the communication atmosphere and intention.

To reflect the uncertainty and subjectivity of communication atmosphere, the concept of the Fuzzy Atmosfield (FA) is introduced to represent the communication atmosphere based on the individual emotion, where the term “Atmosfield” is a combination of “atmosphere” and “field”, reflecting not only the invisible existence of the atmosphere during communication, but also subjective feelings (e.g., effect degree of the atmosphere) of communicators. In addition, a 3D fuzzy cubic space is introduced to represent the states of the FA, where the three coordinates are defined as “Friendly-Hostile”, “Lively-Calm”, and “Casual-Formal” axes with the range from -1 to +1, respectively.

An intention understanding model based on the two-layer fuzzy support vector regression (TLFSVR) is introduced to deeply comprehend people’s inner thoughts using individual emotion. The TLFSVR consists of two layers, including local learning layer and global learning layer. Local learning layer includes fuzzy c-means based input data classification and SVR based learning, after classification, similar emotions are clustering together, which leads to reduce the dimension calculation. According to the identification information (e.g., age, gender, and nationality), global learning layer outputs the intention of human based on fuzzy weighted average algorithm.

In terms of future work, not only the global communication atmosphere, but also the local communication atmosphere could be represented, in addition, it is being applied to represent the communication atmosphere for behavior adaptation of robots in human-robot interaction. Intention understanding model is being extended to some application, such as understanding customers’ order intentions at a bar, at the same time, it is being designed as the supported module of multi-robot behavior adaptation to human intention.

Chapter 3

Adapting Multi-Robot Behavior to Communication Atmosphere

3.1 Introduction

In humans-robots interaction, humans subconsciously adapt their behaviors to communication partners to make interactions run smoothly, actually, not only humans but also robots are expected to adapt to their partners, thus behavior adaptation has been intensively studied [2][3]. Reinforcement learning [8, 9, 53, 54, 55] is widely used in robots behavior adaptation according to individual gesture [10], speech [11], and environment objects [12]. Moreover, emotion plays a key role in human-robot interaction [56]. Emotion based behavior adaptation for autonomous robot is discussed in human-robot interaction [57, 58]. Human speech based robot mentality expression is presented [43].

Since many humans to many robots interaction becomes more widespread, more attention is being paid to communication atmosphere consisted of individual emotion. Most researches, e.g., [8, 9, 53, 54, 55, 10, 11, 12, 56, 57, 58, 43], address mainly the gesture, speech, or emotion, by not paying so much consideration to atmosphere. The communication atmosphere is visually represented using fuzzy atmosfield (FA) in a 3D continuous space that is also useful for humans-robots interaction where uncertainties do exist [25], but not discussed much about the behavior adaptation. Since communication atmosphere is generated from gesture, speech, and emotion, how to use this integrated information

for robot behavior adaptation is important for smooth communication, and is also a new challenge in humans-robots interaction.

Friend-Q learning [54] meets great challenges in real applications, one is the difficulty in handling unexpected situations, and another is the problem of huge dimension of state space. Specifically, the state space of Friend-Q learning is discrete, which cannot represent the situation between two connected states. Fuzzy logic is well-known for successfully applying to uncertain robot control and a discrete algorithm can be generalized to continuous state space through fuzzy logic. Thus, the state space of Friend-Q learning using fuzzy logic can cover all the continuous states rather than limited discrete states, hence the continuous variation of the communication atmosphere can be observed and the small Euclidean distance between learning value and target value is possible to be recognized, which makes robot-generated atmosphere close to human-generated atmosphere. A fuzzy production rule based friend-Q learning method (FPRFQ) is introduced to handle the uncertainties and continuous applications, i.e., human-robot interaction, in which the discrete lookup Q-table in friend-Q learning is replaced by fuzzy production rules with Q-values, and a behavior generating approach based on nearby exploration is used. Based on the FPRFQ, a behavior adaptation mechanism is proposed to solve the robots' behavior adaptation problem. The communication atmosphere information represented in the FA [25] is used as the input for the reward function, where it is defined based on the Euclidean distance between human-generated atmosphere and robot-generated atmosphere in such a way that the reward increases as the distance decreases. Then each robot dynamically adjusts its behavior to the communication atmosphere by maximizing the reward.

Smooth communication of robots and humans means human to human like natural communication in such a way that there is no burden for human to pay attention to a machine or a robot. It is characterized by two indices, i.e., response time of robots and social distance between humans and robots, where the response time is determined by the learning step, and the social distance is defined by the Euclidean distance between human-generated atmosphere and robot-generated atmosphere. To reduce the response time of robots, prior knowledge and learning memory are embedded into the fuzzy production

rules to decrease the learning steps, so that the response time of robots is shortened. For decreasing the social distance between humans and robots, the behavior adaptation mechanism searches for behaviors of robots to maximize the reward.

To show the effectiveness of the proposed behavior adaptation mechanism, a virtual environment is created by MATLAB, where two virtual eye robots (eye movement is the only behavior but simple and typical) and human-generated atmosphere are included. Since unfamiliar issue usually becomes familiar in human cognition process, the communication atmosphere may be classified into two cases, i.e., one is unfamiliar communication atmosphere and the other is unfamiliar-familiar mixed communication atmosphere. Dynamic adjustment of eye movement (of each virtual eye robot) to the two typical atmospheres are confirmed in the simulation experiments, where human-generated atmosphere is introduced in home party scenario [20] and a PC is employed with a dual core processor (2.8 GHz), memory (2.99 GB), Windows 7 system, and MATLAB language.

3.2 Fuzzy Production Rule based Friend-Q Learning (FPRFQ)

Robots' reinforcement learning has been investigated for the interaction between robots and their environment, e.g., in humans-robots interaction, concerning robot states, behaviors, and rewards. The Q-learning algorithm [53] maintains an estimate $Q(s, a)$ value while taking behavior a in state s , and an optimal policy is found by maximizing rewards. Similarly, friend-Q learning [54] maintains an estimate $FQ(s, a_i)$ value while taking behavior a_i of robot i in state s and the $FQ(s, a_i)$ is defined as

$$FQ_{k+1}(s, a_i) = (1 - \alpha_k)FQ_k(s, a_i) + \alpha_k[r_{k+1} + \gamma V_k(s')] \quad (3.1)$$

where i is the number of robots, a_i is the selective behavior of robot i in state s , $FQ_k(s, a_i) \in \mathbb{R}$ is the k -th estimate value of the optimal Q-value, r_{k+1} is the reward, $\alpha_k \in [0,1]$ is a learning rate, $\gamma \in [0,1]$ is a discount factor, $V_k(s') = \max_{a'_i} FQ_k(s', a'_i)$, which

represents the maximum Q-value of all possible behaviors in the next state s' , and a'_i is the possible behavior of robot i in the next state s' . There are 125 kinds of atmosphere states and 25 kinds of emotional behaviors in the experiment.

For discrete applications, Q-values are listed in a Q-table [53], where each item represents a state-behavior pair. For continuous applications, the learning becomes more difficult, as it is hard to list all continuous values in a Q-table. Fuzzy production rule based friend-Q learning (FPRFQ) is an extension of the friend-Q learning method [54] for a continuous environment, where the Q-table is replaced by fuzzy production rule, as shown in Fig. 3.1.

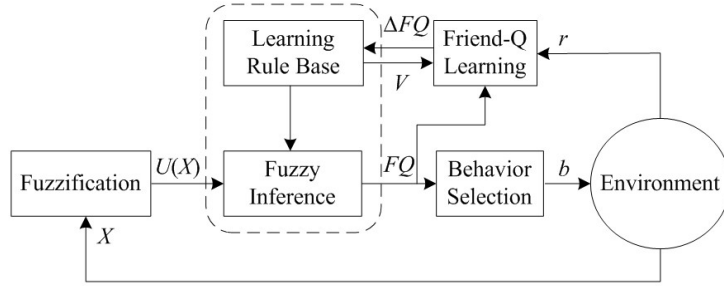


Figure 3.1: Fuzzy production rule based friend-Q learning (FPRFQ).

The FPRFQ in Fig. 3.1 mainly consists of four steps.

Firstly, fuzzification of each input X , e.g., communication atmosphere state, is done by trial and error using membership function in triangular and trapezoidal forms. Secondly, the “IF-THEN” fuzzy production rules of fuzzy inference are Takagi-Sugeno (TS) fuzzy rules designed by observing the relationship between the state and the behavior as

$$R_i : \text{If } X = S_i \text{ Then } RB_j = \text{behavior}_{kj} \text{ with } FQ_{kj} \quad (3.2)$$

where i is number of the learning rules, X is input variable, S_i is fuzzy state, RB_j is behavior set of j -th robot, behavior_{kj} is k -th possible behavior of j -th robot, and FQ_{kj} is a Q-value $\in \mathbb{R}$ fixed for each behavior_{ki} .

Thirdly, each robot selects the behavior b . Each robot has to find the best behavior in the selective learning rule. To export the set of possible behaviors, the behavior of each robot is usually generated by the Boltzmann function [59] which calculates the possibility

of each possible behavior. Since Boltzmann function is random exploration, behaviors will be discovered extremely slowly. Thus, Boltzmann function based nearby exploration is introduced for behavior selection, where the next possible behavior is related to the previous behavior, but not all the 25 kinds of emotional behaviors, e.g., if the previous behavior is anger behavior, then the next possible behavior may be disgust behavior, fear behavior, or surprise behavior. And the probability that a behavior is selected depends on how it is compared to the Q-values of other possible behaviors. If one of the Q-values is much higher, the corresponding behavior is most likely to be taken. The probability is defined as

$$p(s, a_i) = \frac{e^{\frac{\max_{a_i} FQ(s, a_i)}{T}}}{\sum_{i=1}^n e^{\frac{FQ(s, a_i)}{T}}} \quad (3.3)$$

where $p(s, a_i) \in [0,1]$ is the probability of selecting behavior a_i in state s , $T \in \mathbb{R}^+$ is the temperature parameter which can be decreased over time to decrease exploration, and n is the number of possible behaviors in state s . Finally, the Q-values are updated to the learning rule base by using friend-Q learning (FQ) according to the reward calculated by the adaptation level of robots, and the updating function of Q-values is defined as

$$\Delta FQ_{k+1}(s, a_i) = \alpha_k \left[r_{k+1} - FQ_k(s, a_i) + \gamma \max_{a'_i} FQ_k(s', a'_i) \right] \quad (3.4)$$

where $\Delta FQ_{k+1}(s, a_i) = FQ_{k+1}(s, a_i) - FQ_k(s, a_i)$.

3.3 FPRFQ based Multi-Robot Behavior Adaptation Mechanism

3.3.1 Architecture of Behavior Adaptation Mechanism

Base on the proposed FPRFQ and the 3D fuzzy atmosfield (FA) [25], a FPRFQ based behavior adaptation mechanism is proposed as shown in Fig. 3.2, and it mainly includes three modules, i.e., fuzzy classification, behavior adaptation learning, and behavior adaptation memory. In addition, the interaction environment mainly consists of commu-

nication atmosphere (human-generated atmosphere HA and robot-generated atmosphere RA) representation and robot behavior actuator and is supposed to be established.

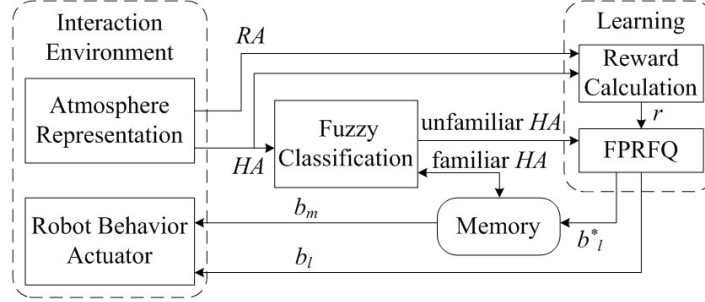


Figure 3.2: *FPRFQ based adaptation mechanism.*

Fuzzy classification module is used to visualize the human-generated atmosphere by using the fuzzy linguistic variables, e.g., the familiar or unfamiliar human-generated atmosphere HA is classified.

Learning module is activated for behavior adaptation to the unfamiliar human-generated atmosphere, which consists of reward calculation and FPRFQ. After the learning process, the most adaptive behavior b_l^* is recorded in the memory module. The human-generated atmosphere and robot-generated atmosphere from the interaction environment are used to calculate the reward. The output of the learning module is learning behavior b_l . In the case of familiar human-generated atmosphere, an adaptive behavior b_m of each robot is selected from the memory module.

Generally speaking, in human life, unfamiliar issue usually becomes familiar through cognition process, so in some degree, the human-generated atmosphere may be classified into two cases, i.e., one is unfamiliar human-generated atmosphere and the other is unfamiliar-familiar mixed human-generated atmosphere. The proposed behavior adaptation mechanism adjusts the policy according to different human-generated atmosphere, i.e., if the human-generated atmosphere is familiar, the memory module is used for adaptation, else the learning module is used.

3.3.2 Fuzzy Classification

The atmosphere state is given as a point in the 3D fuzzy atmosfield (FA) which is defined by “Friendly-Hostile,” “Lively-Calm,” and “Casual-Formal” axes [25]. Fuzzification is necessary to get the fuzzy state of the human-generated atmosphere as well as to determine the adaptation policy whether via the learning module or the memory module.

By using fuzzy linguistic variables, each axis value in the 3D FA is classified into five levels as shown in Fig. 3.3, where the trapezoidal membership functions are used to remove fuzziness from the extreme input. Five linguistic variables are Very Friendly (VFR), Friendly (FR), Neutral (NT), Hostile (H), and Very Hostile (VH) for “Friendly-Hostile”; Very Lively (VL), Lively (L), Neutral (NT), Calm (C), and Very Calm (VC) for “Lively-Calm”; Very Casual (VCA), Casual (CA), Neutral (NT), Formal (F), and Very Formal (VF) for “Casual-Formal.”

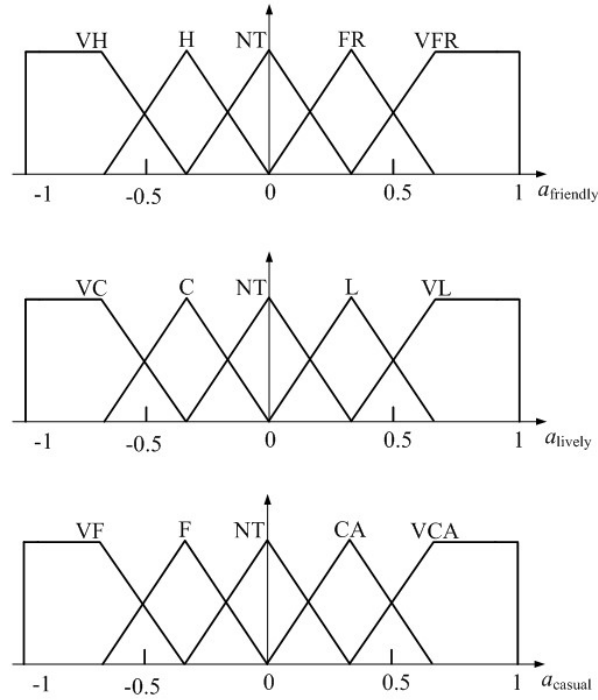


Figure 3.3: Membership functions of 3D FA axes.

After fuzzification of the human-generated atmosphere based on membership functions in Fig. 3.3, the fuzzy state of the human-generated atmosphere is classified into two cases, i.e., unfamiliar HA and familiar HA. If it is a familiar human-generated at-

mosphere, memory module is selected for behavior adaptation; else the FPRFQ based learning module is selected to learn where behavior of each robot is the best adaptive to the unfamiliar human-generated atmosphere.

3.3.3 FPRFQ based Behavior Adaptation

Social distance cannot be irreducible to spatial or biologic distance [60], at least four kinds of social distance, i.e., affective distance, normative distance, interactive distance, and culture distance. If the social distance focuses on affectivity, it is associated with affective distance, i.e., how much or little understanding of others. For example, in humans-robots interaction, if a robot develops specific emotional behaviors toward humans, it should already be able to identify the human-generated atmosphere consisted of individual emotion. The reward function is given for the social distance between humans and robots. The social distance is defined as the Euclidean distance between human-generated atmosphere and robot-generated atmosphere in 3D fuzzy atmosfield [25] as

$$distance(HA, RA) = \sqrt{\sum_{i=1}^3 (HA_i - RA_i)^2} \quad (3.5)$$

where $distance \in [0, 3.46]$, HA_i is i -th coordinate value of human-generated atmosphere state, and RA_i is i -th coordinate value of robot-generated atmosphere state.

The reward function r is defined as

$$r(distance) = \begin{cases} 20 & distance \leq a \\ 1 & a < distance \leq b \\ -20 & distance > b \end{cases} \quad (3.6)$$

where the threshold a and b are determined by the experiment, i.e., trial and error such as $a = 0.3$ and $b = 1$ in the experiment. The reward function is defined based on the Euclidean distance between human-generated atmosphere and robot-generated atmosphere in such a way that the reward increases when the distance becomes shorter.

In the learning process, each robot dynamically adjusts its behavior to the unfamil-

iar human-generated atmosphere by maximizing the reward. Using the FPRFQ, robots cooperate with each other to learn iteratively by trial and error until they receive the expected reward with the best adaptive behavior for the human-generated atmosphere.

The FPRFQ based learning adaptation mainly consists of 7 steps:

Step 1: Initialization

Parameters, i.e., learning rate α (initial value = 1), discount factor γ (= 0.99), temperature parameter T (= 1000), and the learning step n (= N) are initialized.

Step 2: Fuzzification

Fuzzification of the human-generated atmosphere is done by the fuzzy classification module.

Step 3: Fuzzy inference

The possible behaviors of each robot are generated by fuzzy inference according to the unfamiliar human-generated atmosphere and fuzzy production rules.

Since each axis of the 3D FA is classified into five levels by fuzzy linguistic variables, the human-generated atmosphere contains 125 ($= 5^3$) states, e.g., (VC, VH, VF), (NT, NT, NT), and (VL, VFR, VCA). Hence 125 fuzzy production rules in the learning rule base are obtained for all 125 states of the human-generated atmosphere such as

$$\text{If } X = (\text{VC}, \text{VH}, \text{VF}) \text{ Then } RB_j = \text{behavior}_{kj} \text{ with } FQ_{kj} \quad (3.7)$$

where i is the number of learning rules, X is input variable, RB_j is behavior set of j -th robot, behavior_{kj} is k -th possible behavior of j -th robot, and FQ_{kj} is a Q-value $\in (-\infty, +\infty)$ fixed for each behavior_{kj} .

Step 4: Behavior selection

The behavior with the highest Q-value is selected for each robot according to the Boltzmann function in (3.3).

Step 5: Q-values updating

The Q-value of the selective behavior in each robot is updated and saved to the learning rule base by using updating function (3.4) according to the obtained reward.

Step 6: If human-generated atmosphere is equal to robot-generated atmosphere, then

learning stops, and the best learning behavior b_l^* of each robot is outputted; else jump to *Step 7*.

Step 7: Let $n = n+1$. If learning step $n < N$, then jump to *Step 1*; else jump to *Step 6*.

3.3.4 Memory Retrieval for Behavior Adaptation

The memory module not only stores the successful experience, i.e., effective behaviors to adapt to the human-generated atmosphere of robots to the memory rule base, but also generates behavior of each robot in the case of familiar human-generated atmosphere. The fuzzy production rules of the memory rule base are Mamdani fuzzy rules designed as

$$R_i : \quad \text{If } X = S_i \quad \text{Then } behavior_j \quad \text{with } SR_i \quad (3.8)$$

where i is the number of rules, X is input, S_i is familiar human-generated atmosphere state, $behavior_j$ is the behavior of robot j , and SR_i means success rate of rule i . For example, in a two-robot-interaction, one of the memory rules is first selected, which is defined as “If $X = (VL, VFR, VCA)$ then $behavior_1$ is happy behavior and $behavior_2$ is happy behavior with $SR_i = 1$.”

Several behaviors can lead to the expected human-generated atmosphere. Taking into account this problem, the behavior of each robot is selected based on the success rate of the behavior. If robots' behaviors are found to be successful to adapt to a particular human-generated atmosphere according to the FPRFQ, then a rule is added to the memory rule base. Subsequently, the success rate ($SR \in [0,1]$) of the rule is updated when the same rule is used again and is calculated in terms of time-weighted statistics as

$$SR = \frac{ST}{ST + FT} \quad (3.9)$$

where $ST \in N$ is the number of successes and $FT \in N$ is the number of failures. Increase or decrease of the rules depends on two thresholds, i.e., the upper bound ($T_{upper} = 1$) and the lower bound ($T_{lower} = 0.6$). The success rate of effective rule must between these two

boundaries, and others should be deleted.

3.4 Experiments on Multi-Robot Behavior Adaptation

3.4.1 Environment Setting

Autonomous robots may be not only like humans, but also equal to humans, meaning that humans-robots smooth communication is great necessary. Since practical application of smooth communication between humans and robots is not so easy, and taking the complexity of mechanism of the eye robots movements into account, at first, simulation experiments are performed, which not only can run repeatedly to simplify debugging of the mechanism of eye robots, but also can simulate various atmospheres that seldom are experienced in real life.

To show the effectiveness of the proposed behavior adaptation mechanism, a virtual environment is created by MATLAB to simulate the many robots and many humans interaction, where two virtual eye robots and human-generated atmosphere are included. The experiment using two virtual robots is a representative case of multi-robot, and each robot can execute different behaviors at the same time. Virtual robots are eye robots because eye movement can reflect the internal emotion of individuals so that has a strong impact on communication atmosphere and it is the typical part of a humanoid robot. In the experiment, robots' behavior is eye movement only, and the human-generated atmosphere is represented by the 3D FA [25].

Each virtual eye robot is designed with a pair of eyeballs and eyelids, and eye movement is defined as the combination of eyelid motion and ocular motion. The eye movement of eye robot is determined according to the Affinity-Pleasure-Arousal 3D emotion space [43]. In the experiment, each virtual eye robot has 25 kinds of emotional behaviors, e.g., happy eye movement, surprise eye movement, and sad eye movement, for adaptation to the 125 kinds of human-generated atmosphere. The human-generated atmosphere is

given as input in the simulation experiment, e.g., from (very calm, very hostile, very formal) to (neutral, neutral, neutral). According to the real application in the developing mascot robot system (MAS), the atmosphere state is estimated by using the model of fuzzy atmosfield with the fusion of individual emotions using fuzzy logic [25][23], in the simulation experiment, the weight of i -th individual emotional state contributing to the human-generated atmosphere, i.e., w_i is set to $1/n$, where n is the number of individuals. Individual emotion is recognized based on emotional speech and gesture [23]. Rewards in the learning process are calculated according to the Euclidean distance between the human-generated atmosphere and the robot-generated atmosphere, as shown in (3.5) and (3.6).

To verify the effectiveness of the proposed mechanism over the whole area of the FA [25], the FA is divided into eight octants, i.e., from octant I to octant VIII, shown as Fig. 3.4. Each octant can be familiar or unfamiliar atmosphere area. In the experiments, dynamic adjustment of eye movement (of each virtual eye robot) to the unfamiliar human-generated atmosphere and the unfamiliar-familiar human-generated atmosphere are both confirmed in the simulation experiments, where a PC is employed with a dual core processor (2.8 GHz), a 2.99 GB memory, Windows 7 system, and MATLAB language.

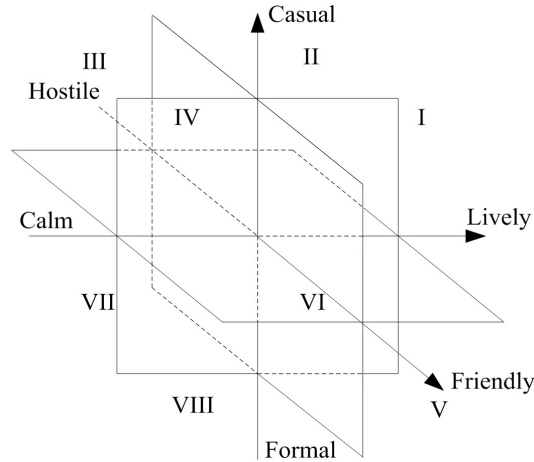


Figure 3.4: *Division of FA.*

3.4.2 Experiment on Adaptation to Unfamiliar Atmosphere

To validate the performance of the proposed behavior mechanism for adapting to unfamiliar human-generated atmosphere, a comparative experiment among the FPRFQ, the FQ [54], and the IL [55] is carried out in the virtual atmosphere environment. The two virtual robots dynamically adjust the eye movements to adapt to the unfamiliar human-generated atmosphere, i.e., from (very calm, very hostile, very formal) in octant VII of the FA to the neighborhood of the origin, i.e., (neutral, neutral, neutral). The discounted parameter γ in each algorithm is set to 0.99. In the FPRFQ, the FQ and the IL, the learning rate α (initial value = 1) is set to a decreasing function. To guarantee an optimal policy, the temperature of Boltzmann function T is set to 1000.

In the experiment, the three methods are repeated 5 times in total. The average learning steps, the average reward, and the social distance (sum of the Euclidean distance between human-generated atmosphere and robot-generated atmosphere in all steps) for completing once task (until robot-generated atmosphere is equal to human-generated atmosphere) are used as the indexes for evaluating the performance of the three methods. Smooth communication is not only for achieving behavioral tasks with a short response time, but also for building social relationships to increase familiarity between humans and robots [61]. Thus, smooth communication between humans and robots is characterized by two indices, i.e., response time of robots and social distance between humans and robots. Response time is determined by the learning step which is from 0 to $+\infty$, and social distance is defined by the Euclidean distance between human-generated atmosphere and robot-generated atmosphere which is from 0 to 3.46 (the maximum distance between two points in the 3D fuzzy atmosfield space). In the experiments, a communication is regarded as smooth for completing a task of robot adaptation (i.e., robot-generated atmosphere is equal to human-generated atmosphere) when it meets the following conditions as “if response time $< t_{sm}$ (average learning steps $< n_{sm}$) and social distance (sum in all steps) $< d_{sm}$ then smooth communication is achieved,” where t_{sm} , n_{sm} , and d_{sm} are the upper bound of response time, average learning steps, and social distance, respectively, in the unfamiliar atmosphere adaptation experiment, $t_{sm} = 0.3s$, $n_{sm} = 25$, and $d_{sm} = 35$; in

the unfamiliar-familiar mixed atmosphere adaptation experiment, $t_{sm} = 0.6s$, $n_{sm} = 50$, and $d_{sm} = 115$. All the upper bound values are obtained by trial and error, i.e., each 20 times experiments for the FPRPQ, the FQ, and the IL.

According to Figs. 3.5, 3.6, and 3.7, data analysis of results comparison among the FPRFQ, the FQ and the IL are showed in Table 4.1, the proposed mechanism saves 47 learning steps and gets 11 times the reward compared with the IL. Robots' response time is about 0.289 s by using the proposed mechanism, which saves 0.073s and 0.205s than the FQ and the IL, respectively. The adaptation of the proposed mechanism has better performance than the FQ and the IL in average steps and reward, which is due to embedding of prior knowledge and learning memory into the fuzzy production rules to decrease the number of learning steps in the proposed behavior mechanism, and that are not included in the other two methods. Moreover, the social distance between human-generated atmosphere and robot-generated atmosphere is 2 times and 4 times shorter than that of the FQ and the IL, which shows that the proposed behavior adaptation mechanism decreases the social distance between humans and robots.

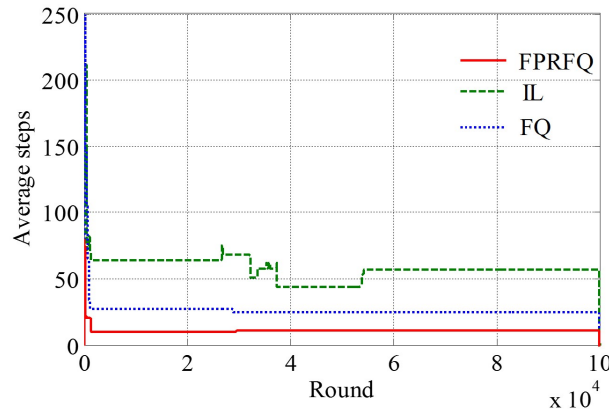


Figure 3.5: Average steps to complete once task.

Table 3.1: Result comparison of adaption to unfamiliar atmosphere

	FPRFQ/IL/FQ		
	Mean	Min	Max
Average steps	11/58/26	10/44/25	20/83/43
Average reward	9/-91/-57	-30/-130/-98	11/-61/-49
Social distance	33/112.3/59.9	31.5/98.9/47.9	44.1/2215/77.8

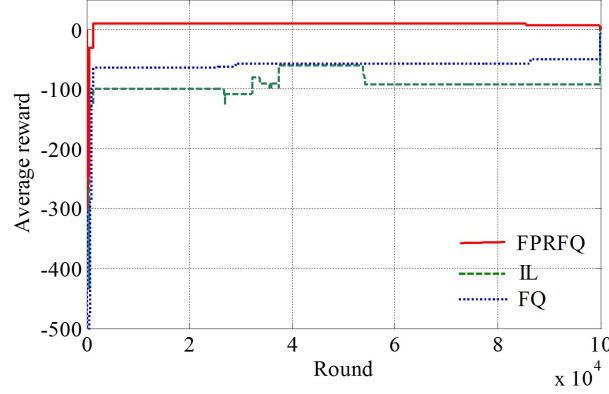


Figure 3.6: Average reward to complete once task.

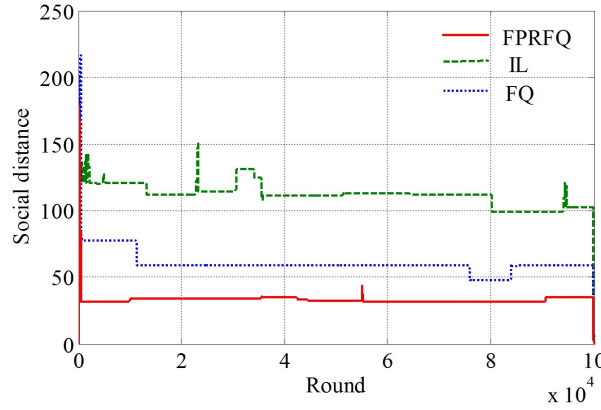


Figure 3.7: Social distance to complete once task.

3.4.3 Experiment on Adaptation to Unfamiliar-Familiar Mixed Atmosphere

To validate the performance of the proposed behavior adaptation mechanism for adapting to unfamiliar-familiar mixed atmosphere, a comparative experiment among the FPRFQ, the FQ, and the IL is carried out in virtual atmosphere environment.

The task is to adapt to the unfamiliar-familiar mixed atmosphere from (very calm, very hostile, very formal) in octant VII of the FA to (very lively, very friendly, very casual) in octant I of the FA via (neutral, neutral, neutral) i.e., the origin of the FA by dynamically adjusting the eye movement of each virtual robot. The discounted parameter γ in each algorithm is set to 0.99. In all the three methods, the learning rate α (initial value = 1) is set to a decreasing function. To guarantee an optimal policy, the temperature of

Boltzmann function T is set to 1000.

According to Figs. 3.8, 3.9, and 3.10, the data analysis of result comparison among the three methods is shown in Table 3.2, where the FPRFQ saves 44 learning steps and obtains 2 times reward compared with the FQ and it also saves 482 steps and gets 7 times the reward compared with the IL. Robots' response time is about 0.458s by using the proposed mechanism, which saves 0.156s and 0.435s than FQ and IL, respectively. The proposed behavior adaptation mechanism has a better performance than other two methods in average steps and average reward, which is due to the embedding of the prior knowledge and learning memory into the fuzzy production rules to decrease the number of learning steps. Moreover, the social distance between the human-generated atmosphere and the robot-generated atmosphere is 3 times and 10 times shorter than that of the FQ and the IL, which verifies that the adaptation mechanism decreases the social distance between humans and robots in the case of unfamiliar-familiar mixed communication atmosphere.

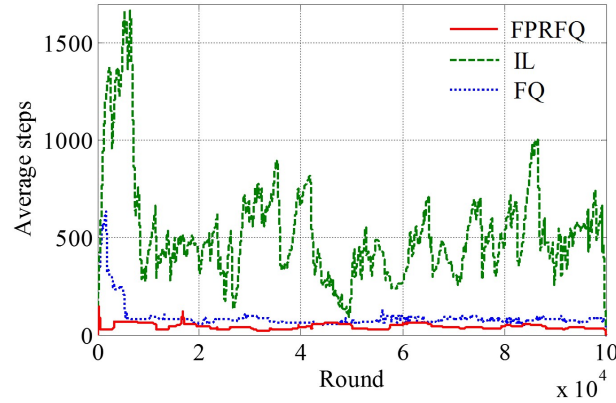


Figure 3.8: Average steps to complete once task.

Table 3.2: Result comparison of adaption to unfamiliar-familiar mixed atmosphere

	FPRFQ/IL/FQ		
	Mean	Min	Max
Average steps	43/525/87	23/93/48	120/1666/636
Average reward	-233/-1571/-57	-763/-5653/-4923	-101/-234/-229
Social distance	113/1199/321	31/166/92	382/3806/2215

According to the experimental results of adaptation to two typical communication atmospheres, the robots' response time is reduced in both experiments by using the pro-

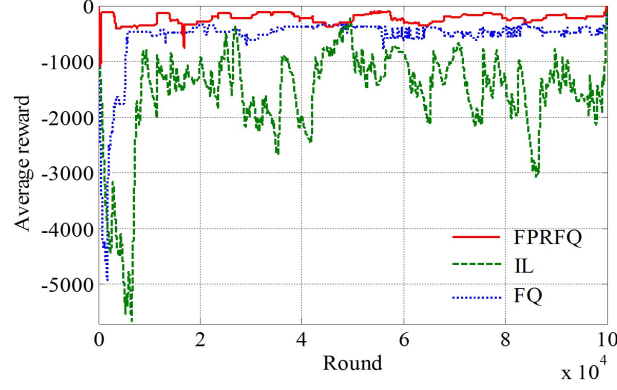


Figure 3.9: Average reward to complete once task.

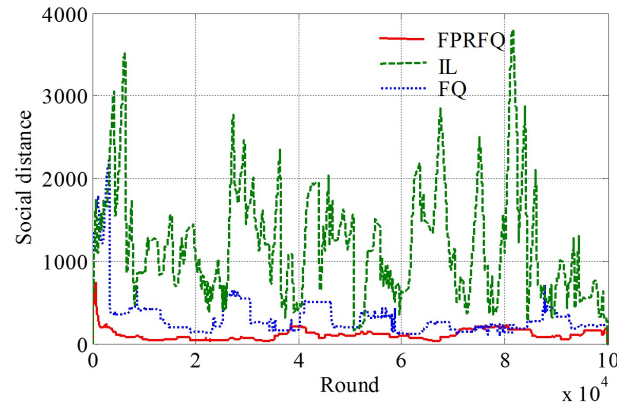


Figure 3.10: Social distance to complete once task.

posed behavior adaptation mechanism. This is because the proposed behavior adaptation mechanism shows better performance than other two methods (i.e., Friend Q-learning (FQ) and Independent Learning (IL)) in terms of average steps and average reward. Two main reasons for the superior performance compared to the FQ and the IL are given. The first reason is embedding of prior knowledge into the fuzzy production rules for the FQ. According to the human-generated atmosphere state, e.g., (VL, VFR, VCA), using the FQ or the IL, the virtual eye robot selects one appropriate behavior from all the 25 kinds of emotional behaviors. Using the proposed FPRFQ based behavior adaptation mechanism, the virtual eye robot just selects the behavior from the neighborhood of the happy behavior according the prior knowledge that the friendly and relaxed atmosphere is related to happy behavior, and in Affinity-Pleasure-Arousal 3D emotion space [43] the candidate area of emotional behaviors for the atmosphere (VL, VFR, VCA) is shown in

Fig. 3.11. Therefore, the virtual eye robots just select from several candidate behaviors not all the 25 kinds of emotional behaviors so that it saves the learning time.

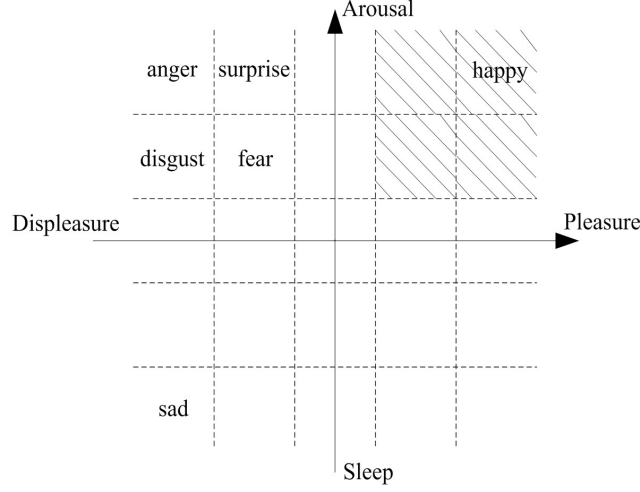


Figure 3.11: Candidate areas of emotional behaviors for atmosphere (VFR, VL, VCA).

The second reason is embedding of learning memory into the fuzzy production rules for the FQ. Using the FQ or the IL, the virtual eye robot always selects behavior from the 25 kinds of emotional behaviors adapting to the human-generated atmosphere, regardless of whether it is familiar or unfamiliar. A memory module is introduced in the proposed behavior adaptation mechanism in order to reduce learning steps in the case of familiar atmosphere. By using the memory module, the virtual eye robot can directly find the corresponding behavior from their memory, saving the learning steps as well.

Moreover, the virtual eye robots' response time is sub-second in the simulation, even in the real robots adaptation, e.g., in the MRS (five eye robots and four humans) [20], the CPU of robot is ARM STM32F103 (72 MHz) and about 30 times slower than that of the PC employed in the experiment. In this case, the response time is still in seconds, indicating that the proposed mechanism can reduce the response time of robots in real application of humans-robots communication.

In both experiments, the social distance between humans and robots becomes closer, which means the whole communication atmosphere becomes more harmonious and is for smooth communication, this is because both humans and robots generate similar atmospheres that can be combined together very well. In the experiment, the whole

space of the 3D FA (Fuzzy Atmosfield) [25] is divided into eight octants. The proposed mechanism adapts to the atmosphere in the seventh octant (calm, hostile, formal area) in unfamiliar case of the first experiment and in familiar case of the second experiment, and adapts to the first octant (lively, friendly, casual area) in unfamiliar case of the second experiment. It also confirms the adaptation to the atmosphere in the neighborhood of the origin. Since the adaptation of the proposed mechanism is confirmed in the first octant (all positive), the seventh octant (all negative), and the neighborhood of the origin (all directions), the effectiveness of the proposed adaptation mechanism may be confirmed in the general octant case.

Basic experiment on eye robots' eye movement (emotional movement [43]) adaptation is developed in MRS [20] with four humans and five eye robots by using the proposed behavior adaptation mechanism, where an idealized home party environment is present without background music or noises.

Figs 3.12, 3.13, and 3.14 present three scenarios. In Fig. 3.12, the human-generated atmosphere is $(0, 0, 0.3)$ and fuzzification is (neutral, neutral, casual), after adaptation, eye robot do some eye movement to reflect the emotion (approximately zero, approximately zero). In Fig. 3.13, the human-generated atmosphere is $(0.6, 0.6, 0.3)$ and fuzzification is (lively, hostile, casual) and the emotion is (negative small, positive small). In Fig. 3.14, the human-generated atmosphere is $(0.5, 0.8, 0)$ and fuzzification is (very lively, friendly, neutral) and the emotion is (positive small, positive large). Results show that eye robots well adapt to the atmosphere by executing appropriate behavior based on the proposed mechanism.

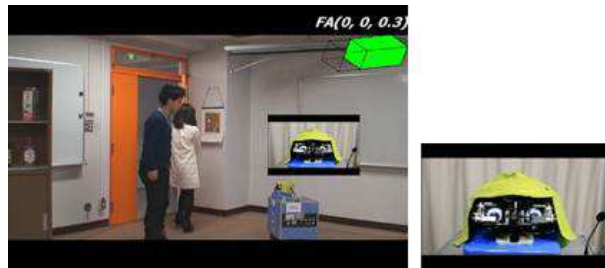


Figure 3.12: Scenario of “greeting guests at the door”.

Future research on the proposed mechanism would be interesting in the following



Figure 3.13: Scenario of “query of train timetable”.



Figure 3.14: Scenario of “watching TV”.

ideas, one is to enrich the adaptation mechanism for humans, besides robots’ behavior adaptation; the other is music adaptation to the atmosphere, where music such as that from the violin [22] is considered as an independent emotional object as well as the humans and robots. For actual application, the multi-agent Q-learning method has already been applied to the humanoid robot, e.g., Robovie II [10] and Nico [62], indicating that the proposed improved multi-agent Q-learning, i.e., FPRFQ based behavior adaptation mechanism may be welcomed to humanoid robots’ behavior adaptation in humans-robots interaction.

3.5 Chapter Summary

First experiment in unfamiliar communication atmosphere, the adaptation of the proposed mechanism reduces 15 learning steps and 47 learning steps than that of friend-Q learning (FQ) [54] and the independent learning (IL) [55], respectively, where robots’ response time is about 0.289s using the proposed mechanism, by saving 0.073s and 0.205s than that of the FQ and the IL, respectively. Meanwhile, the proposed mechanism obtains about 6 times reward and 11 times reward than that of the FQ and the IL, re-

spectively, where the distance between human-generated atmosphere and robot-generated atmosphere is 2 times and 4 times shorter than that of the FQ and the IL, respectively.

Second experiment in unfamiliar-familiar mixed communication atmosphere, the adaptation of the proposed mechanism reduces 44 learning steps than that of the FQ, and 482 learning steps than that of the IL, where robots' response time is about 0.458s using the proposed mechanism, by saving 0.156s and 0.435s than that of the FQ and the IL, respectively. Moreover, the proposed mechanism obtains about 2 times the reward than that of the FQ and 7 times the reward than that of the IL, respectively, where the distance between the human-generated atmosphere and the robot-generated atmosphere is 3 times and 10 times shorter than that of the FQ and the IL, respectively.

According to the results of adaptation to these two typical communication atmospheres, the robots' response time is sub-second in the simulation, even in real robots adaptation, e.g., in the MRS (five eye robots and four humans) [20], the CPU of robot is ARM STM32F103 (72 MHz) and about 30 times slower than that of the PC employed in the experiment, but the response time is still in seconds, indicating that the proposed mechanism can reduce the response time of robots in real application among at most 9 humans and robots. Moreover, with the development of the CPU computational speed, a humans-robots interaction environment can be divided into several areas using area distribution control. And in each area, there can be one robot as minimum and nine robots as maximum for humans-robots interaction. The proposed behavior adaptation mechanism of robot is used for each area separately, and in a case of 9 areas environment, the proposal can reduce the response time of robots among at most $81(9 \times 9)$ the number of humans and robots.

Additionally, the social distance between humans and robots becomes closer in the both experiments, which could contribute to realization of smooth communication to some extent. Generally speaking, if the 3D FA (Fuzzy Atmosfield) [25] is employed to express the communication atmosphere, the whole atmosphere space is divided into eight octants. The proposed mechanism adapts to the atmosphere in the seventh octant (hostile, calm, formal area) in unfamiliar case of the first experiment and in familiar case of the second

experiment, and adapts to the first octant (friendly, lively, casual area) in unfamiliar case of the second experiment. It also confirms the adaptation to the atmosphere in the neighborhood of the origin. Since the adaptation of the proposed mechanism is confirmed in the first octant (all positive), the seventh octant (all negative), and the neighborhood of origin (all direction), the effectiveness of the proposed adaptation mechanism may be confirmed in the general octant case, and the application experiments are ongoing in general humans-robots interaction scheme.

In addition, basic experiment is applied to the MRS [20], where a home party environment is created with five eye robots and four humans. Results show that eye robots adapt well to the communication atmosphere in several scenarios, e.g., “greeting guests at the door,” “query of train timetable,” and “watching TV,” and the eye robots execute different eye movements to represent different emotion in each scenario, which demonstrates the basic feasibility of the proposal in actual humans-robots interaction. Moreover, future study on the proposed mechanism would be interesting to enrich the adaptation mechanism for humans (not only robots), and consider the fact that the influence factors of adaptation to atmosphere is not only humans and robots, but also musical instruments, e.g., the violin [22]. For the actual application, the multi-agent Q-learning method has already been applied in the humanoid robot, e.g., Robovie II [10] and Nico [62], indicating that the proposed improved multi-agent Q-learning (FPRFQ) based behavior adaptation mechanism to be welcomed to apply to humanoid robots’ behavior adaptation in humans-robots interaction.

Chapter 4

Multi-Robot Behavior Adaptation to Local and Global Communication Atmosphere

4.1 Introduction

With the development of robotic technology, especially the emergence of humanoid robot, increasing attention is paid to domestic environments and assistive tasks in humans-robots interaction (HRI). HRI involves not only communication between humans and robots, but also a dynamic process that contains an observation of others' behaviors and an estimation of others intention from observation [63][64]. To realize an anticipated interaction or cooperation in HRI, learning and adaptation are indeed necessary in the robotics community [65]. As sociable partners, it is not enough to achieve task-based goals only, and robots need to perform naturally to obtain an appropriate social behavior and response. In this case, robots should recognize, understand, and respond appropriately to humans' social signals including speech, gesture, and gaze[4][5][6][7]. Reinforcement learning algorithm is suitable for robot behavior adaptation, for example, a policy gradient reinforcement learning based behavior adaptation mechanism is introduced to make interactions run smoothly [10]. A virtual agent learns how to behave according to the conversations of other participants based on reinforcement learning in a computer network [66].

Cognitive science as one of the main concerns in affective computing [67], is an im-

portant topic for the interactive communication such as harmonious humans-robots interaction, where human emotional state based robot behavior adaptation attracts a lot of attention [16][17]. Emotion interaction module based HRI framework is used to robot behavior adaptation in [18], where cognitive interaction is used for cooperative sharing of tasks. In [19], a comprehensive, time-varying affective model is designed for adapting the robot behavior, according to generation and adaptation of emotion.

For concerning a large scale humans-robots interaction (i.e., multiple humans to multiple robots interaction) [68], besides commonly used information such as speech, gesture, and emotion, another very significant information is introduced for humans-robots interaction, i.e., communication atmosphere. Atmosfield [23] – a contraction of Atmosphere and Field – is a concept that expresses the existence of communication atmosphere in the space surrounding us. Due to the ambiguity and uncertainty of the Atmosfield, it is further defined as Fuzzy Atmosfield (FA). The FA is described by 3D space with “Friendly-Hostile”, “Lively-Calm”, and “Casual-Formal” axes, as shown in Fig. 4.1, which is generated by emotional states of humans and improves the smooth communication between humans and robots [24]. The state of the FA in the 3D space is described as $FA = (afriendly, alively, acasual)$, where FA is the state of the FA, and *afriendly*, *alively*, and *acasual* are the values on “Friendly-Hostile,” “Lively-Calm,” and “Casual-Formal” axes respectively. [23][24] mainly focus on the generation of the communication atmosphere, without so much about behavior adaptation. In the previous research, multiple robots behaviors adaptation to communication atmosphere in HRI uses fuzzy production rules based friend-Q learning (FPRFQ), where the discrete lookup Q-table in friend-Q learning [54] is replaced by fuzzy production rules with Q-values [26]. In [26], the FPRFQ is used for cooperative situation of humans-robots interaction, which means that all the humans and robots are considered as a whole gathering or a group, and the robots are designed to adapt to the whole communication atmosphere, i.e., global communication atmosphere. In real life, however, more situations do exist, e.g., in a football match or a discussion meeting, individuals may support different teams or have different opinions. In these complex situations where cooperation, neutral, and competition coexist, only using global communication

atmosphere for robots behaviors adaptation is not enough. Thus, local communication atmosphere i.e., group-based atmosphere is introduced to distinguish it from the global communication atmosphere and enrich the atmosphere information. Then, how the robots adapt to the local and global coexistent communication atmosphere becomes an interesting problem.

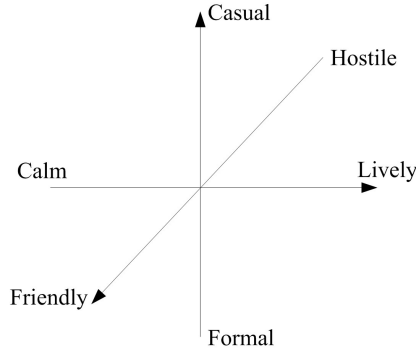


Figure 4.1: *Fuzzy Atmosfield expressed in 3D space.*

A cooperative-neutral-competitive fuzzy-Q learning (CNCFQ) method is introduced to handle the complex communication, i.e., cooperation, neutral, and competition mixed humans-robots interaction, in which the maximum based updating Q function (i.e., in the FPRFQ [54]) is replaced by the maximum-minimax based updating Q functions and the relationships among agents are identified as friends, neutral, or foes. Based on CNCFQ, a behavior adaptation mechanism is proposed to multiple robots for adapting behaviors to the local and global coexistent communication atmosphere. The communication atmosphere represented in the FA [23] is generated by using the two-layer fuzzy fusion model (i.e., local fusion layer and global fusion layer), and the global communication atmosphere is used as the input for the reward function, where it is defined based on the Euclidean distance between global human-generated atmosphere and global robot-generated atmosphere in such a way that the reward increases as the distance decreases. Using the proposed mechanism, robots firstly identify the opponents as friends, neutral, or foes, then robots dynamically adjust their behavior to the communication atmosphere by maximizing the reward based on the maximum algorithm (in the case of cooperation or neutral) or minimax algorithm (in the case of competition).

The proposal enables robots to adapt their behaviors to the communication atmosphere so that a smooth humans-robots interaction could be realized. Smooth communication between robots and humans is interpreted as that there is no burden for human to pay attention to a machine or a robot [26]. To realize smooth communication, response time of robots and social distance between humans and robots are two important indexes, where the response time is determined by the learning steps, and the social distance is defined by the Euclidean distance between global human-generated atmosphere and global robot-generated atmosphere. Learning step is one of the key factors influencing the learning rate [69], which means how far the robots responses to the humans. Smooth communicate makes agents know each other more easily since the response time is reduced, also the communication atmosphere is easy to adapt because of the short social distance.

The local and global communication atmosphere is visualized through the easy-to-understand computer graphics, i.e., shape-color-fill-wave graphics. The visualization offers a more visual and expressive way for users to communicate both abstract and concrete ideas, moreover, it not only gives a better sense of how friendly/lively/casual the atmosphere is according to the shape-color-fill graphics, but also can directly feel the intensity of the atmosphere from the wave graphic.

To show the effectiveness of the proposed behavior adaptation mechanism, a virtual home party environment is created by MATLAB, which is a scenario of watching football match (watching football match is popular and typical for friends, neutral, and foes coexistent situation). There are 6 virtual eye robots (eye movement is the only behavior but simple and typical) and 6 virtual human agents, and they are divided into different groups. Global communication atmosphere consists of several local communication atmospheres generated by each group. Dynamic adjustment of eye movement (of each virtual eye robot) to the local and global coexistent communication atmospheres is confirmed in the simulation experiments, also the questionnaires are done for evaluating the graphic visualization of the communication atmosphere, where a PC is employed with a dual core processor (2.8GHz), memory (2.99GB), Windows 7 system, and MATLAB language.

4.2 Cooperative-Neutral-Competitive Fuzzy-Q Learning (CNCFQ)

Multi-agent reinforcement learning is widely investigated in multi-player games. Stochastic games (SG) or Markov games extends one-state Matrix Game (MG) to multi-state MGs by modelling state transitions with a Markov Decision Process (MDP) [70]. Under the SG framework, Hu and William extended the minimax-Q learning algorithm for zero-sum games [71] to general-sum games, and a Nash Q-learning algorithm [72] is proposed to some constrained games. Due to the multiple equilibriums, a friend-Q learning (FQ) is used [54]. For handling the environment uncertainty and large state space in robotic application, fuzzy logic is fused with the discrete Q-learning [53] to generalize the continuous state space.

The previous proposed fuzzy production rule based friend-Q learning (FPRFQ) [26] is the fuzzy extension of the FQ, where the discrete lookup Q-table is replaced by “IF-THEN” fuzzy production rules. It mainly consists of four steps, i.e., fuzzification, fuzzy production rule selection, robots behavior selection, and Q-values updating. The updating function is defined as

$$\Delta FQ_{k+1}(s, a_i) = \alpha_k \left[r_{k+1} - FQ_k(s, a_i) + \gamma \max_{a'_i} FQ_k(s', a'_i) \right] \quad (4.1)$$

where i is the robot number, a_i is the selective behavior of i th robot in state s , $FQ_k(s, a_i) \in \mathbb{R}$ is the k th estimate value of the optimal Q-value, r_{k+1} is the reward, $\alpha_k \in [0, 1]$ is learning rate, $\gamma \in [0, 1]$ is discount factor, $\max_{a'_i} FQ_k(s', a'_i)$ represents the maximum Q-value of all possible behaviors in the next state s' , a'_i is the possible behavior of i th robot in the next state s' , and $\Delta FQ_{k+1}(s, a_i) = FQ_k(s, a_i) - FQ_{k+1}(s, a_i)$.

The FPRFQ focuses on solving the cooperative problem in humans-robots interaction, where all the agents are friends. In real life, there are more relationships other than friends, e.g., foes and neutral. Cooperative-neutral-competitive fuzzy-Q learning (CNCFQ) is proposed for complicated interaction involving friends, foes, and neutrals. It is an extension of the FPRFQ, where maximum based updating Q function is replaced

by maximum-minimax based updating Q functions, as shown in Fig. 4.2, and mainly consists of five steps.

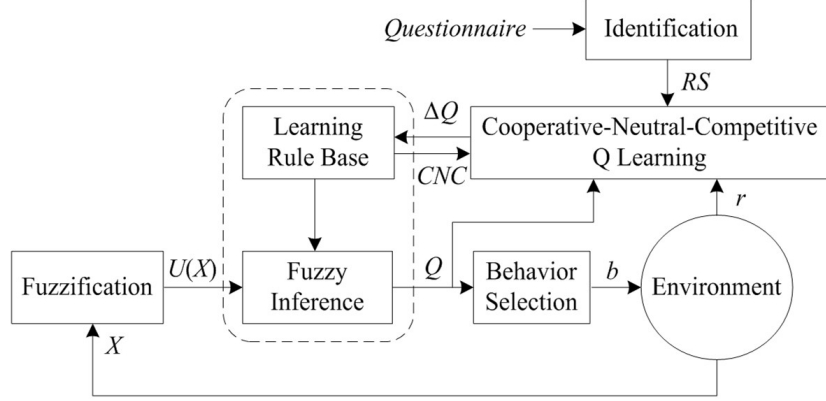


Figure 4.2: Cooperative-neutral-competitive fuzzy-Q learning (CNCFQ).

Firstly, identification of the relationships RS among agents, i.e., friend, neutral, or foe, which is obtained by questionnaires (e.g., when watching a football match, the question is “which team do you support? A or B.”)

Secondly, rule selection consists of fuzzification and fuzzy inference. Fuzzification is the process of changing a real scalar of each input X (e.g., communication atmosphere state) into a fuzzy value $U(X)$. Membership functions are triangular forms, and the selection process is done by trial and error. Then, the “IF-THEN” fuzzy production rules are based on Takagi- Sugeno (TS) fuzzy rules designed by trial and error, which are represented in the form of relationship between the state and the behavior as

$$R_{ij} : \text{If } U(X) = S_i \quad \text{Then} \quad RB_{kj} = \text{behavior}_{lk} \quad \text{with} \quad Q_{lk} \quad (4.2)$$

where i is the learning rule number, j is the group number, $U(X)$ is fuzzification of the input variable, S_i is fuzzy state, RB_{kj} is behavior set of k th robot in j th group, behavior_{lk} is l th possible behavior of k th robot, and Q_{lk} is a Q-value $\in \mathbb{R}$ fixed for each behavior_{lk} .

Thirdly, each robot has to pick up the best behavior b from the selective learning rule. In general, the behavior of each robot is generated by the explore/exploit policies, e.g., ϵ -greedy [73] and Boltzmann function [59]. Boltzmann exploration is widely used, and it can cover greedy case by adjusting the parameter (i.e., temperature), thus, the

Boltzmann function is used for behavior generation. Boltzmann exploration is based on the principle that the frequency of robot behavior should be proportional to its average reward and the key idea of pursuit is to maintain an explicit probability distribution over the behaviors, and to search directly in the probability distributions space of Q-values. If one of the Q-values is much higher, the corresponding behavior is most likely to be taken. The probability is defined as

$$p(s, a_i) = \frac{e^{\frac{\max_{a_i} Q(s, a_i)}{T}}}{\sum_{i=1}^n e^{\frac{Q(s, a_i)}{T}}} \quad (4.3)$$

where $p(s, a_i) \in [0, 1]$ is the probability of selecting behavior a_i in state s , n is the number of possible behaviors in state s , and $T \in \mathbb{R}^+$ is the temperature parameter which can be decreased over time to decrease exploration, when $T = 0$, Boltzmann exploration acts like pure greedy, as T tends to infinity, robots select behaviors uniformly at random.

Finally, according to the relationships among agents and the reward obtained by the adaptation level of robots, Q-values are calculated and updated to the learning rule base. The maximum-minimax based updating function of Q-values is defined as

$$\Delta Q_{k+1}^j(s, a_{ij}) = \alpha_k \left[r_{k+1}^k - Q_k^j(s, a_{ij}) + \gamma \text{CNC}_{a'_{ij}}^j(s', a'_{ij}) \right] \quad (4.4)$$

where j is the group number, e.g., A vs B in a football match, A's and B's fans form 1st and 2nd group, respectively, and $\Delta Q_{k+1}^j(s, a_{ij}) = Q_k^j(s, a_{ij}) - Q_{k+1}^j(s, a_{ij})$.

More specifically, for different relationships RS (i.e., friend, neutral, or foes), the CNC function is different and if the RS is friend or neutral, they are all in the same group, $j=1$, and the Q-values is defined as

$$\Delta Q_{k+1}(s, a_i) = \alpha_k \left[r_{k+1} - Q_k(s, a_i) + \gamma \max_{a'_i} Q_k(s', a'_i) \right] \quad (4.5)$$

and if the RS is foe, there are two groups in the game, $j=1, 2$, then

$$\begin{aligned}\Delta Q_{k+1}^1(s, a_{i1}, a_{i2}) &= \alpha_k \left[r_{k+1}^1 - Q_k^1(s, a_{i1}, a_{i2}) + \gamma \max_{a'_{i1} \in RB_{i1}} \min_{a'_{i2} \in RB_{i2}} Q(s', a'_{i1}, a'_{i2}) \right] \\ \Delta Q_{k+1}^2(s, a_{i1}, a_{i2}) &= \alpha_k \left[r_{k+1}^2 - Q_k^2(s, a_{i1}, a_{i2}) + \gamma \max_{a'_{i2} \in RB_{i2}} \min_{a'_{i1} \in RB_{i1}} Q(s', a'_{i1}, a'_{i2}) \right]\end{aligned}\quad (4.6)$$

where RB_{i1} and RB_{i2} are behavior sets of i th robot in the 1st group and the 2nd group, respectively.

4.3 Representation of Local and Global Atmosphere

4.3.1 Two-Layer Fuzzy Fusion Model for Atmosphere Generation

For small scale humans-robots interaction, e.g., 5 agents including 3 humans and 2 robots, the global communication atmosphere is generated in the 3D fuzzy atmosfield (FA) based on the individual emotion [23]. As the scale becomes larger, e.g., more than 10 agents, the global communication atmosphere alone is not enough to cover all the situations in the whole environment, especially in the case of friends-neutral- foes coexistent communication. It may be not suitable for generating only one atmosphere, thus, local atmosphere is a good supplement to enrich the communication atmosphere for behavior adaptation. A two-layer fuzzy fusion model is introduced to generate the communication atmosphere in FA, as shown in Fig. 4.3.

According to Fig. 4.3, there are two layers in the model, one is the local fusion layer and the other is the global fusion layer. Suppose that the communication environment is divided into m groups based on the previous identification process, and the local atmosphere LA is generated by fuzzy rules based weight fusion function in the local fusion layer according to the individual emotion and defined as

$$LA_i = \sum_{j=1}^{n_1} w_{Hj} \cdot \underset{\sim}{defuzzzy}(E_{Hj} \circ R) + \sum_{k=1}^{n_2} w_{Rk} \cdot \underset{\sim}{defuzzzy}(E_{Rk} \circ R) \quad (4.7)$$

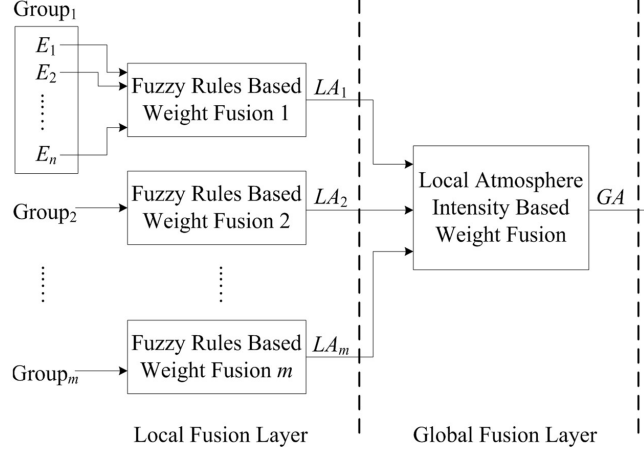


Figure 4.3: Two-layer fusion model of atmosphere generation.

where i is the group number, j is the human agent number in i th group, k is the robot agent number in i th group, n_1 is the total number of robot agent in i th group, n_2 is the total number of human agent in i th group, w is the weight of each emotional state contributing to the local atmosphere LA , $\forall w \in [0, 1]$, $\sum_{j=1} n_1 w_{Hj} + \sum_{k=1} n_2 w_{Rk} = 1$, R is the relationship between individual emotion and the LA , which is obtained by using a fuzzy inference system with 75 fuzzy rules [23], $\tilde{E}$ is a fuzzy set of the individual emotion, the center of gravity method is used as a defuzzification method in *defuzzy*. Emotion recognition using multiple channels of information fusion based on bimodal information, i.e., speech and gesture [24].

In humans-robots interaction, the communication ways for emotional expression mainly include speech, facial expression, and body language, especially, speech is the highly effective method for communicating emotion [2] and the sound of speech influences the communication atmosphere significantly [23], in other words, the sound determines the existence and state of an atmosphere to a great extent. Therefore, speech volume is employed to compute the weight of each emotional state as

$$w_i = \frac{v_i}{\sum_{j=1}^n v_j}, i, j = 1, 2, \dots, n \quad (4.8)$$

where $v \in [0 \text{ dB}, 90 \text{ dB}]$ is the speech volume.

In the global fusion layer, the global atmosphere GA is generated based on the intensity of the local atmosphere and defined as

$$GA_i = \sum_{j=1}^m \lambda_i LA_j \quad (4.9)$$

where λ is the intensity of the local atmosphere, $\forall \lambda \in [0, 1]$, $\sum_{j=1}^m \lambda_j = 1$.

Considering the factors affect the GA , besides the speech volume, the number of the agents in a group is also related, since it seems easier to influence other group if the size of the group is bigger. Thus, the intensity of the local atmosphere is calculated according to the speech volume and the number of the agents in a group and defined as

$$\lambda_i = \frac{\hat{v}_i + \hat{s}_i}{\sum_{j=1}^m (\hat{v}_i + \hat{s}_i)}, i, j = 1, 2, \dots, m \quad (4.10)$$

where $\hat{v}$ is the normalization of v , $\hat{s}$ is the normalization of s , and s is the number of the agents in a group.

To more vividly express the intensity of local atmosphere, “Weak-Strong” axis is used to fuzzy description, as shown in Fig. 4.4, where λ is divided into seven fuzzy level based on the triangular membership functions, i.e., Extremely Strong (ES), Very Strong (VS), Strong (S), Neutral (NT), Weak (W), Very Weak (VW), and Extremely Weak (EW).

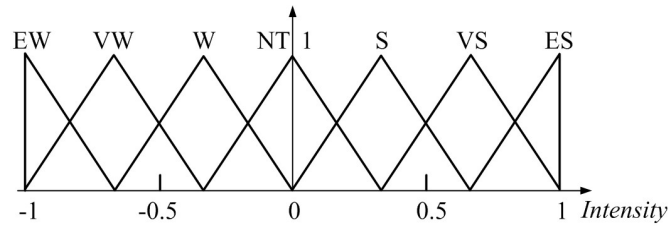


Figure 4.4: Membership functions of intensity.

4.3.2 Graphic Visualization for Atmosphere Representation

The local and global atmosphere represented in the 3D FA space are recorded as either data (i.e., coordinate value) or texture (i.e., friendly/lively/casual), which cannot

give a direct feeling of how the atmosphere is. Graphic is more appealing, also easier to comprehend and remember rather than the data or pure texture. For easy observation and understanding by human, a shape-color-length graphical representation for the FA is introduced in [23], where shape, color, and length are employed to represent the states of FA on “Friendly-Hostile”, “Lively-Calm”, and “Casual-Formal” axes, respectively.

As the interaction scale becomes larger, all the participants may be classified into several groups according to their interest, thus, not only the global atmosphere (generated by all groups) exists, but also the local atmosphere (generated by each group). Graphic visualization in [23] gives the information of how friendly/lively/casual the atmosphere is, for the local and global coexistent atmosphere. However, more information is necessary, e.g., which is the dominant group in the communication. To represent how strong the local atmosphere is, intensity of the local atmosphere mentioned in 4.3.1 is added. Shape-color-fill-wave graphic visualization is introduced, where shape, color, and fill are employed to represent the states of the FA on “Friendly-Hostile”, “Lively-Calm”, and “Casual-Formal” axes, respectively, wave is used to represent the intensity of the local atmosphere since intensity is always related to the wave which amplitude expresses how strong the atmosphere is, as shown in Fig. 4.5.

4.4 CNCFQ based Multi-Robot Behavior Adaptation Mechanism

4.4.1 Architecture of Behavior Adaptation

Based on the proposed CNCFQ and local-global communication atmosphere represented in fuzzy atmosfield (FA) [23], a multi-robot behavior adaptation mechanism is proposed in humans-robots interaction as shown in Fig. 4.6. The adaptation architecture reflects empirical findings accumulated in cognitive (memory) and reinforcement learning suggesting the relationship of action execution and environment. The basic idea is that the behavior adaptation mechanism may support the complex interaction (e.g., co-

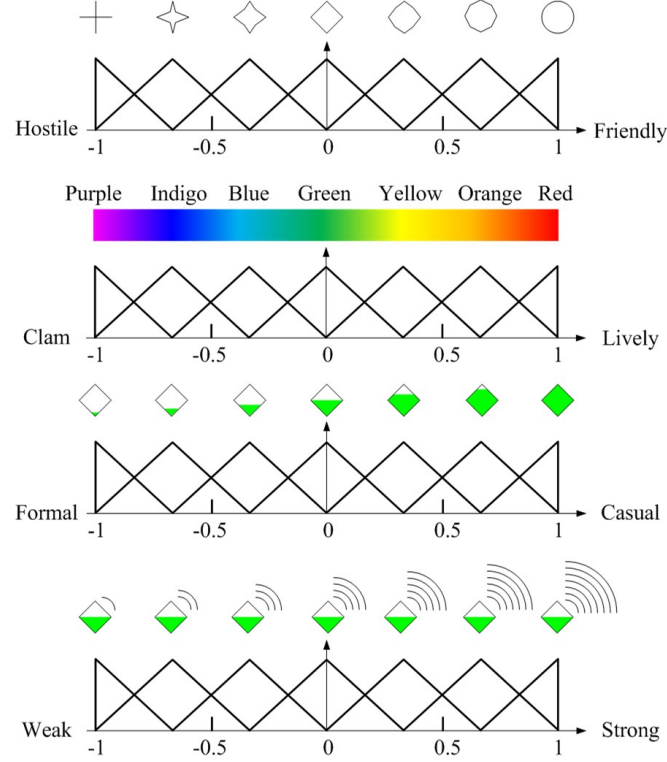


Figure 4.5: Visualization of atmosphere.

operation, neutral, and competition mixed humans-robots interaction) where not only the global communication atmosphere exists but also the local communication atmosphere. It mainly includes three modules, i.e., fuzzy classification, behavior learning, and memory. In addition, the interaction environment consists of questionnaire, local-global atmosphere (i.e., local/global human-generated atmosphere LHA/GHA and local/global robot-generated atmosphere LRA/GRA) representation, and robot behavior actuator.

Questionnaire is a survey including the question Q and the answers A . It is used to identify the relationships among agents.

Fuzzy classification module is used to distinguish the global human-generated atmosphere into two types by using the fuzzy linguistic variables, i.e., familiar or unfamiliar global human-generated atmosphere GHA . The human-generated atmosphere can be classified into these two types to some extent, because unfamiliar issue usually becomes familiar through cognition process in human life.

Local-global atmosphere representation module focuses on generating and visualization of atmosphere. Atmosphere is generated by using two-layer fuzzy fusion model while

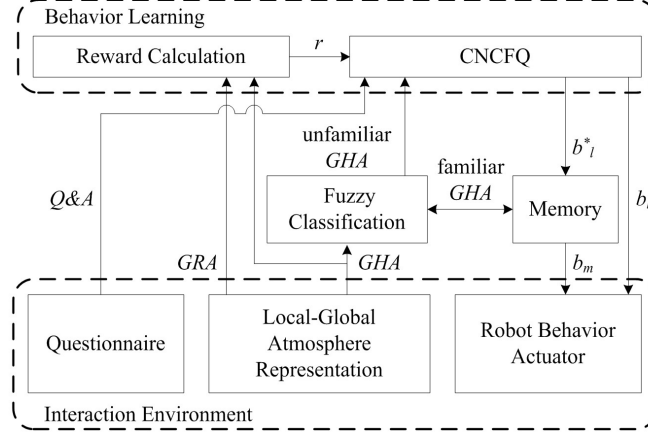


Figure 4.6: Architecture of CNCFQ based behavior adaptation mechanism.

visualized by graphical methods.

Behavior learning module consists of reward calculation and the cooperative-neutral-competitive fuzzy-Q learning (CNCFQ). The global human-generated atmosphere and global robot-generated atmosphere are used as the input. After the learning process, learning behavior b_l is transferred to the robot behavior actuator, and the most adaptive behavior b_l^* is recorded in the memory module. In the case of familiar global human-generated atmosphere, an adaptive behavior b_m of each robot is selected from the memory module. The proposed behavior adaptation mechanism adjusts the policy according to different global human-generated atmosphere, i.e., if GHA is familiar, the memory module is used for adaptation, or else through the learning module.

4.4.2 Questionnaire for Relationship Identification

For identifying the relationships among the participants in the interaction, questionnaire is developed. It is designed by opened-ended (responding by a specific piece of information) and closed-ended (responding by picking an answer from a given number of options) mixed type, which consists of a number of questions Q and their corresponding answers A .

4.4.3 Fuzzy Classification

Fuzzy classification mainly includes fuzzification and classification which is used to determine the adaptation policy via the learning module or the memory module. The atmosphere state is a point in the 3D fuzzy atmosfield (FA) which is characterized with “Friendly-Hostile”, “Lively-Calm”, and “Casual- Formal” axes [23].

By using fuzzy linguistic variables, each axis value in the 3D FA is classified into seven levels as shown in Fig. 4.7, where the triangular membership functions are used to remove fuzziness from the extreme input. Seven linguistic variables are Extremely Friendly (EFR), Very Friendly (VFR), Friendly (FR), Neutral (NT), Hostile (H), Very Hostile (VH), and Extremely Hostile (EH) for “Friendly-Hostile”; Extremely Lively (EL), Very Lively (VL), Lively (L), Neutral (NT), Calm (C), Very Calm (VC), and Extremely Calm (EC) for “Lively-Calm”; Extremely Casual (ECA), Very Casual (VCA), Casual (CA), Neutral (NT), Formal (F), Very Formal (VF), and Extremely Formal (EF) for “Casual-Formal”.

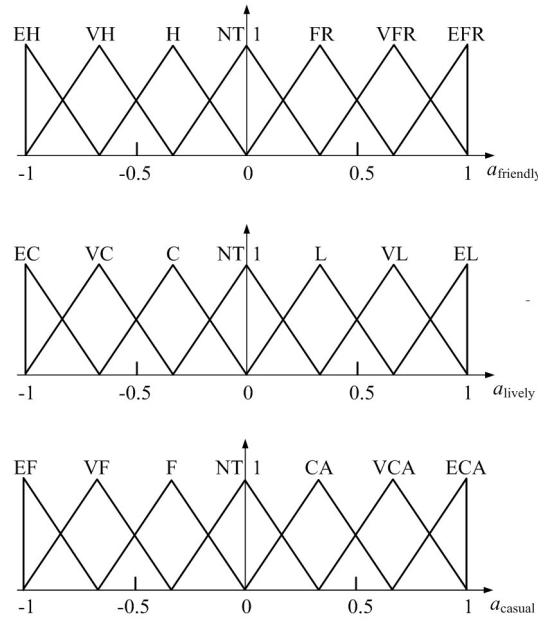


Figure 4.7: Membership functions of 3D FA axes.

After fuzzification, the fuzzy state of the global human-generated atmosphere (GHA) is obtained. Then the fuzzy state is classified into two cases according to the memory, i.e., unfamiliar GHA and familiar GHA . If it is a familiar GHA , memory module is used

for behavior adaptation or else the CNCFQ based learning module is used where the best adaptive behavior of each robot is selected to the unfamiliar *GHA*.

4.4.4 CNCFQ based Behavior Adaptation

Generally speaking, social distance describes the distance between different groups of society and is opposed to locational distance. In sociological literature, social distance can be conceptualized in affective distance, how much or little the members of a group feel for another group [60]. For example, in humans-robots interaction, robots are designed to estimate and understand the humans' atmosphere, and even adapt to it, namely the social distance is getting closer. The social distance is the Euclidean distance between global human-generated atmosphere and global robot-generated atmosphere in 3D fuzzy atmosfield [23] and defined as

$$distance(GHA, GRA) = \sqrt{\sum_{i=1}^3 (GHA_i - GRA_i)^2} \quad (4.11)$$

where distance $(GHA, GRA) \in [0, 3.46]$, GHA_i is i th coordinate value of global human-generated atmosphere state, and GRA_i is i th coordinate value of robot-generated atmosphere state.

The reward function r is defined based on the social distance between humans and robots as following

$$r(distance) = \begin{cases} 20 & distance \leq a \\ 1 & a < distance \leq b \\ -20 & distance > b \end{cases} \quad (4.12)$$

where the threshold a and b are determined by the experiment, i.e., trial and error. $a = 0.3$ and $b = 1$ in the experiment. The reward increases as the distance becomes shorter.

In the learning process, each robot dynamically adjusts its behavior to the human-generated atmosphere by maximizing the reward. Using the CNCFQ, robots cooperate with each other to learn iteratively by trial and error until they receive the expected reward with the best adaptive behavior for the global human-generated atmosphere. The

CNCFQ based learning adaptation mainly consists of 8 steps:

Step 1: Initialization.

Parameters, i.e., learning rate α (initial value = 1), discount factor γ (= 0.99), temperature parameter T (= 1000), and the learning step n (= N) are initialized.

Step 2: Identification.

To identify the relationships RS among the agents as friend, neutral, or foe by using the questionnaire.

Step 3: Fuzzification.

Fuzzification of the human-generated atmosphere is done by the fuzzy classification module.

Step 4: Fuzzy inference.

The possible behavior of each robot is generated by fuzzy inference according to unfamiliar global human-generated atmosphere and fuzzy production rules.

Since each axis of the 3D FA is classified into seven levels by fuzzy linguistic variables, the global human-generated atmosphere contains 343 ($= 7^3$) states, e.g., (EH, EC, EF), (NT, NT, NT), and (VFR, VL, VCA). Hence 343 fuzzy production rules in the learning rule base are obtained such as

$$\text{If } X = (\text{EH}, \text{EC}, \text{EF}) \text{ Then } RB_{kj} = \text{behavior}_{lk} \text{ with } FQ_{lk} \quad (4.13)$$

where j is the group number, X is fuzzification of the input variable, RB_{kj} is behavior set of k th robot in j th group, behavior_{lk} is l th possible behavior of k th robot, and Q_{lk} is a Q -value $\in \mathbb{R}$ fixed for each behavior_{lk} .

Step 5: Behavior selection.

The behavior with the highest Q -value is selected for each robot according to the Boltzmann function in (4.3).

Step 6: Q -values updating.

The Q -value of the selective behavior in each robot is updated and saved to the learning rule base by using updating function (4.5) (RS is friend or neutral) or (4.6) (RS is foe) according to the obtained reward.

Step 7: If global human-generated atmosphere is equal to global robot-generated atmosphere, then learning stops, and the best learning behavior b_l^* of each robot is outputted; else jump to Step 8.

Step 8: Let $n = n + 1$. If learning step $n < N$, then jump to Step 2; else jump to Step 1.

4.4.5 Memory Retrieval for Behavior Adaptation

The memory module stores successful experiences, i.e., effective behaviors to adapt to the human-generated atmosphere of robots to the memory rule base. Since unfamiliar issue usually becomes familiar in human cognition process, the communication atmosphere may be classified into two cases, i.e., one is unfamiliar communication atmosphere and the other is familiar communication atmosphere. Memory module is used to generate behavior of each robot in the case of familiar global human-generated atmosphere. The fuzzy production rules of the memory rule base are Mamdani fuzzy rules designed as

$$R_i : \quad \text{If } X = S_i \quad \text{Then } behavior_{ki} \quad \text{with } SR_i \quad (4.14)$$

where i is the number of rules, X is input, S_i is familiar global human-generated atmosphere state, $behavior_{kj}$ is the behavior of k th robot in j th group, and SR_i means success rate of rule i . For example, in a three-group-interaction, each group has two robots, one of the memory rules is firstly selected, which is defined as “If $X = (EH, EL, EF)$ Then $behavior_{11}$ is happy behavior, $behavior_{21}$ is anger behavior, $behavior_{12}$ is surprise behavior and $behavior_{22}$ is surprise behavior, $behavior_{13}$ is fear behavior and $behavior_{31}$ is fear behavior with $SR_i = 1$.

Several behaviors may be suitable for the expected global human-generated atmosphere. Taking this issue into account, the behavior of each robot is selected based on the success rate of the behavior. If robots' behaviors are found to be successful to adapt to a particular human-generated atmosphere according to the CNCFQ, then a rule is added to the memory rule base. Subsequently, the success rate ($SR \in [0, 1]$) of the rule is updated

when the same rule is used again and is calculated in terms of time-weighted statistics as

$$SR = \frac{ST}{ST + FT} \quad (4.15)$$

where $ST \in N$ is the number of successes and $FT \in N$ is the number of failures. Increase or decrease of the rules depends on two thresholds, i.e., the upper bound ($T_{upper} = 1$) and the lower bound ($T_{lower} = 0.6$). The success rate of effective rule must be between these two boundaries, otherwise, it should be deleted.

4.5 Experiments on Multi-Robot Behavior Adaptation to Local and Global Coexistent Communication Atmosphere

4.5.1 Environment Setting

To verify the proposed behavior adaptation mechanism, a virtual home party environment is created by MATLAB to simulate many robots to many humans interaction, which is a scenario of watching football match, six virtual eye robots and six virtual human agents are included, as shown in Fig. 4.8. Taking the complexity of mechanism of the eye robots movements into account, simulation experiments are performed. It not only can run repeatedly to simplify debugging of the mechanism of eye robots, but also can simulate various atmospheres that are seldom experienced in real life.

Scenario of watching football match is chosen since football match is popular all around the world, and the three situations of friend, neutral, and foe commonly coexist in a football game. The virtual robots are eye robots because eye movement is a typical part of a humanoid robot that can reflect the internal emotion of individuals, and it has a strong impact on communication atmosphere. Robots' behavior is eye movement which is determined according to the Affinity-Pleasure-Arousal 3D emotion space [43]. The human-generated atmosphere is represented in the 3D FA [23] divided into eight

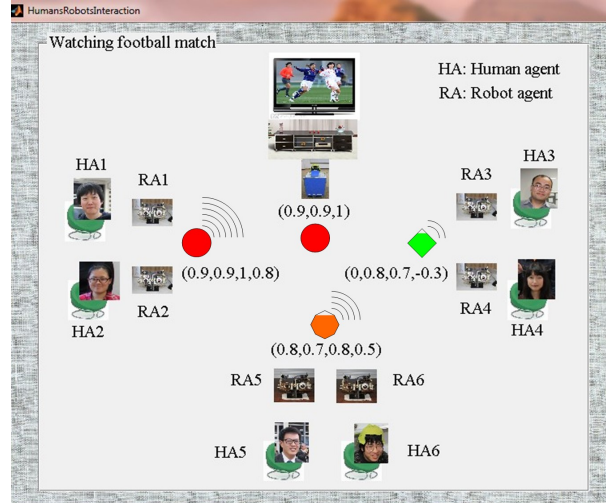


Figure 4.8: *Virtual environment.*

quadrants, i.e., from quadrant I to quadrant VIII, as shown in Fig. 4.9.

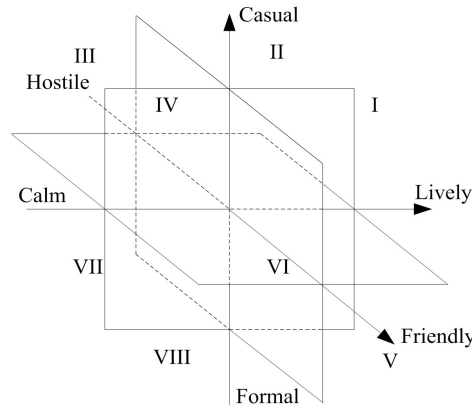


Figure 4.9: *Division of FA.*

In the experiment, each virtual eye robot uses 25 kinds of emotional behaviors [43], e.g., happy, surprise, and sad eyemovement, adapt to the 343 kinds of global human-generated to atmosphere. The global human-generated atmosphere is given as input according to the atmosphere changed when watching a real football match at home. The people generate a friendly, lively, and casual atmosphere at the beginning, and then it turns stiff because they support different teams. Accordingly, communication atmosphere changes from (extremely friendly, extremely lively, extremely casual) to (neutral, neutral, neutral), and finally to (extremely hostile, extremely calm, extremely formal). Individual emotion is chosen directly from the 3D emotion space [43]. Learning Rewards are calculated

according to the Euclidean distance between the global human-generated atmosphere and the global robot-generated atmosphere, as shown in 4.11 and 4.12.

Dynamic adjustment of eye movement (of each virtual eye robot) to the local and global human-generated atmosphere is confirmed in the simulation experiments, as well as the subjective estimation for graphic visualization of atmosphere. The experiments are done in a PC which contains a dual core processor (2.8GHz), a 2.99GB memory, Windows 7 system, and MATLAB language.

4.5.2 Experiment on Adaptation to Local and Global Coexistent Human-Generated Atmosphere

To validate the performance of the proposed behavior mechanism for adapting to local and global coexistent human-generated atmosphere, a comparative experiment among the CNCFQ, the FPRFQ [26], and the FQ [54] is carried out in a virtual watching football match environment.

Suppose the opponents in the football match are team A and team B, six virtual robots are divided into three groups (i.e., group supports team A, group supports team B, and group remain neutral) according to the questionnaire (Q_1 is “which team do you support?” A_1 is “A or B”; Q_2 is “what is your agent number” A_2 is the answer by agent), where each group has two virtual robots. They dynamically adjust the eye movements to adapt to the global human-generated atmosphere in the virtual watching football match environment, i.e., from (extremely friendly, extremely lively, extremely casual) in quadrant I of the FA to (neutral, neutral, neutral) in the neighborhood of the origin, finally to (extremely hostile, extremely calm, extremely formal) in quadrant VII of the FA, (neutral, neutral, neutral).

The goal of the experiment is to make the interaction between humans and robots run smoothly. Smooth communication is not only to achieve behavioral tasks with a short response time, but also for building social relationships between humans and robots [61]. It is characterized by two indices, i.e., response time of robots (determined by the learning step which is form 0 to $+\infty$) and social distance between humans and robots (defined

by the Euclidean distance between global human-generated atmosphere and global robot-generated atmosphere which is from 0 to 3.46, i.e., the maximum distance between two points in the 3D Fuzzy Atmosfield space and details in 4.3.1). For completing a task of robot adaptation (i.e., robots keep adapting until the global robot-generated atmosphere is equal to the global human-generated atmosphere), a communication is regarded as smooth when it meets the following conditions as if response time $< t_{sm}$ (average learning steps $< n_{sm}$) and social distance (sum in all steps) $< d_{sm}$ then smooth communication? where t_{sm} , n_{sm} , and d_{sm} are the upper bound of response time, average learning steps, and social distance, respectively.

Three methods are repeated 10 times and 10000 rounds in one time in the simulation experiment. The average learning steps, the average reward, and the social distance (sum of the Euclidean distance between human-generated atmosphere and robot-generated atmosphere in all steps) for completing once task (until global robot-generated atmosphere is equal to global human-generated atmosphere) are used as the indexes for evaluating the performance of the three methods. In the experiment, the discounted parameter γ in each algorithm is set to 0.99, the learning rate α (initial value = 1) is set to a decreasing function, i.e., $\frac{1}{(1+\alpha)^4}$. The temperature of Boltzmann function T is set to 1000 for guaranteeing an optimal policy. All the upper bound values for regarding a smooth communication are obtained by trial and error, i.e., each 30 times experiments for the three methods, that are $t_{sm} = 0.015s$, $n_{sm} = 25$, and $d_{sm} = 50$.

The data analysis of results comparison among the three methods is shown in Table 4.1 according to Fig. 4.10, Fig. 4.11 and Fig. 4.12, where the CNCFQ saves 47 learning steps (i.e., leaning rate is increased by 72%) and obtains 7 times reward compared with the FPRFQ and it also saves 103 steps (i.e., leaning rate is increased by 85%) and gets 15 times the reward compared with the FQ. Robots' response time is about 0.01s by using the proposed mechanism, which saves 0.025s and 0.056s compared to FPRFQ and FQ, respectively. Moreover, the social distance between the human-generated atmosphere and the robot-generated atmosphere is 3 times and 7 times shorter than that of the FPRFQ and the FQ, respectively.

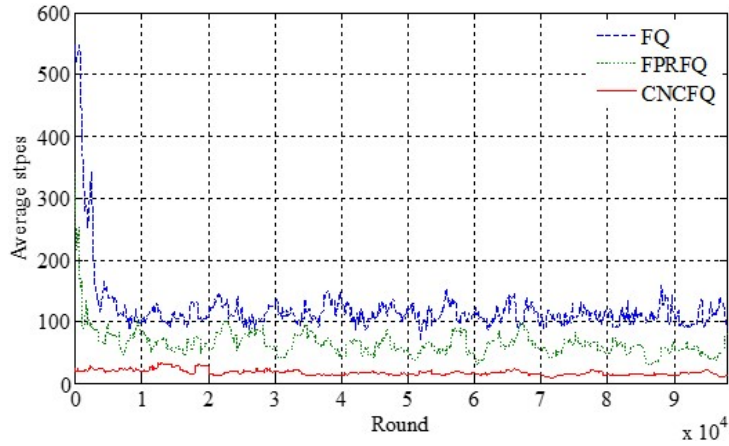


Figure 4.10: Average steps to complete one task.

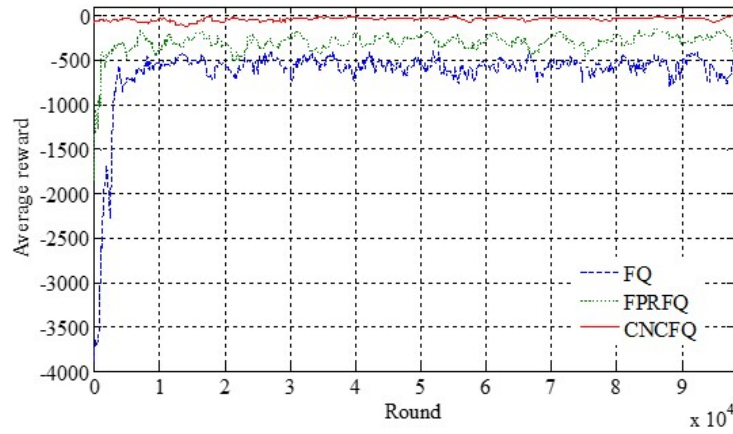


Figure 4.11: Average reward to complete one task.

The experimental results show that by using the proposed behavior adaptation mechanism, the robots' response time is reduced. There are mainly two reasons, one is adding the information of intensity of the local human-generated atmosphere for adaptation, not only global human-generated atmosphere used in FPRFQ and FQ, that is to say, the leading group is known so that other groups may keep up with him soon, which reduces the response time; the other is embedding of learning memory into the fuzzy production rules. Using the FQ, the virtual eye robot always selects behavior from all 25 kinds of emotional behaviors, regardless of familiar or unfamiliar global human-generated atmosphere. A memory module is introduced in the proposal to reduce learning steps in the case of familiar atmosphere, in other words, eye robot can directly find the corresponding behavior from their memory.

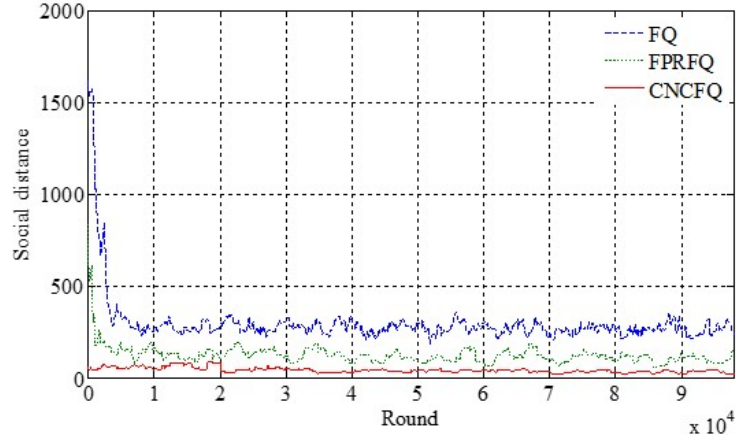


Figure 4.12: Social distance to complete one task.

Table 4.1: Results comparison for adaptation to local and global atmosphere coexistent communication

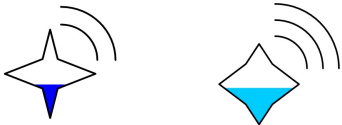
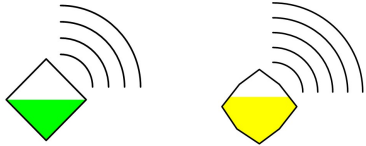
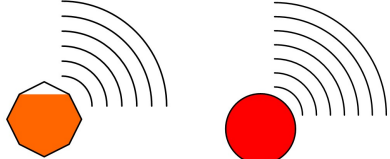
	CNCFQ/ FPRFQ / FQ		
	Mean	Min	Max
Average steps	18/65/121	10/30/69	33/340/548
Average reward	-43/-303/-630	-117/-1849/-3884	-4/-136/-392
Social distance	42/125/296	19/52/180	92/834/1617

The social distance between humans and robots becomes closer that means the robots generate similar atmospheres to humans', which makes the humans-robots interaction smoothly. In the experiment, the whole space of the 3D Fuzzy Atmosfield (FA) [68] is divided into eight quadrants. The proposed mechanism is confirmed in the first quadrant (friendly, lively, casual area), the eighth quadrant (hostile, calm, formal area), and the neighborhood of the origin. Since the adaptation of the proposed mechanism is confirmed in the first quadrant (all positive), the seventh quadrant (all negative), and the neighborhood of the origin (all directions), the effectiveness of the proposed adaptation mechanism may be confirmed in the general quadrant case.

4.5.3 Evaluation of Graphic Visualization for Atmosphere Representation

To evaluate the proposed graphic visualization of atmosphere, thirty-five volunteers (i.e., postgraduate students aged 20-35 years, from eight countries at different universities in Tokyo) completed a questionnaire about the relationship between three elements (i.e. shape, color, fill, and wave) and three attributes of the FA, as well as the relationship between wave and the intensity of the local atmosphere, as shown in Table 4.2. Twenty-one random graphics are provided for evaluation, in which cover all the seven fuzzy level of the FA and intensity of local atmosphere. For each question, there are seven options from which an answer is chosen.

Table 4.2: Questionnaire for evaluation of graphic visualization of communication atmosphere

Graphic visualization	Questions	Answers
	1. Does the shape represent hostile-friendly?	1-extremely hostile, 2-very hostile, 3-hostile, 4-neutral, 5-friendly, 6-very Friendly, 7-extremely friendly
	2. Does the color represent calm-lively?	1-extremely calm, 2-very calm, 3-calm, 4-neutral, 5-lively, 6-very lively, 7-extremely lively
	3. Does the fill represent formal-casual?	1-extremely formal, 2-very formal, 3-formal, 4-neutral, 5-casual, 6-very casual, 7-extremely casual
21 graphics in total	4. Does the wave represent weak-strong?	1-extremely weak, 2-very weak, 3-weak, 4-neutral, 5-strong, 6-very strong, 7-extremely strong

By comparing the answers from questionnaires with the meanings that are assigned to the shape, color, fill, and wave as mentioned in 4.3.2, the accurate rates, i.e., 85.5% for shape, 77.2% for color, and 65.3% for fill, and 91.7% for wave are obtained, which indicates that the wave is the easiest to be cognitive with the feelings of “weak” and

“strong”, followed by the shape. The reason why the evaluation result of the color is less than that of the shape and wave is due to the confusion between purple and red by some volunteers, they chose “extremely calm” for the feeling of red. The fill obtains the lowest accuracy since the filling rate is not obvious for the seven fuzzy level from “formal” to “casual”.

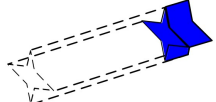
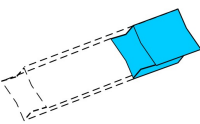
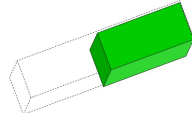
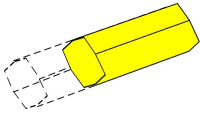
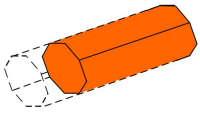
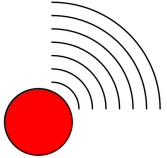
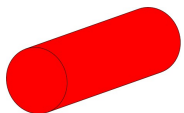
The comparison between the proposal and that in [23] is carried out to verify the effectiveness of the proposed shape-color-fill-wave graphic visualization of the atmosphere, where thirty-five volunteers (i.e., postgraduate students aged 20-35 years, from eight countries at different university in Tokyo) completed a questionnaire as shown in Table 4.3. Seven graphics are used to evaluation, including all seven fuzzy level in the FA space and the intensity of local atmosphere, compared with corresponding seven fuzzy level in the graphic visualization of [23]. For each question, there are two options from which an answer is chosen.

According to the result, more than 72.5% volunteers support the proposed shape-color-fill-wave graphic visualization of communication atmosphere. Some volunteers giving the reasons why they choose the proposal because of the wave for the intensity of local atmosphere is really impressive and more information is obtained.

People in the different countries may have their own culture and preferences, these resulted in different atmospheres in the communication, the proposed mechanism is being extended to develop for eye robots (in the developing humans-robots interaction system, called mascot robot system, MRS [24]) behavior adaptation to different country atmospheres in the international communication, where participants are from multiple countries, as shown in Fig. 4.13.

There are four fixed eye robots place on the furniture or appliances such as the TV, dart game, information monitor display, drink mini bar, and an autonomous mobile eye robot in the MRS. Robot Technology Middleware (RTM) is used to construct the network system of MRS called RTM-Network [21], as shown in Fig. 4.14. It is a two-layer network system consisted of information sensing and transfer layer as well as data processing layer. Each robot can be regarded as a network component and each function unit of

Table 4.3: Questionnaire for comparison between two kinds of graphic visualization

Graphic visualization		Questions	Answers
Proposal	Liu's [23]		
		1. Which is more impressive?	1-Proposal, 2- Liu's
			
			
		2. Which gives more information?	1-Proposal, 2- Liu's
			
			
			

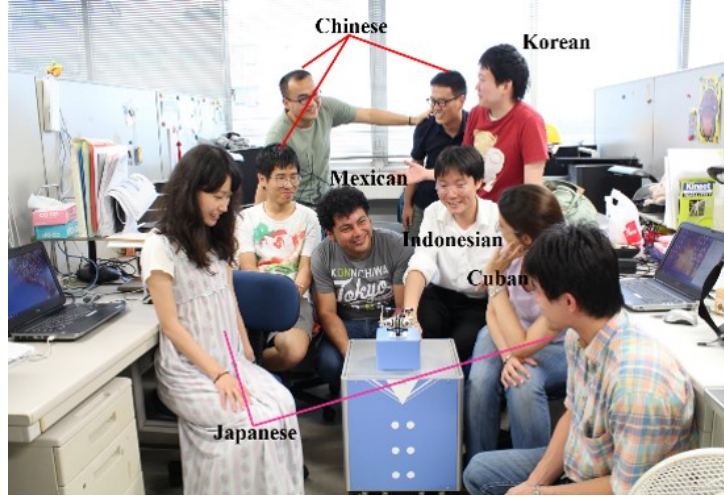


Figure 4.13: *International communication.*

the robot is called Robot Technology Component (RTC). There are eight modules in the RTM-Network, including the speech recognition module (SRM), gesture recognition module (GRM), eye-robot control module (ERCM), emotion processing module (EPM), display module, mobile robot control module, scenario server module, and information retrieval module [74]. The SRM sensing the speeches information through microphone and the GRM sensing the gestures information using Kinect. Moreover, the EPM obtains the recognized results from the SRM and the GRM, and then transmits them to the server module for the further processing, such as the calculation of individual emotions and communication atmosphere. The robots' behaviors are adapted according to the communication atmosphere, and the server module transmits eye movement instructions to the ERCM for expressing corresponding emotional behaviors of eye robots.

Moreover, as cognitive science is widely used in the humans-robots interaction, deep level understanding is getting more and more popular, interesting research topics may extend the adaptation mechanism not only for emotion or atmosphere, but also other hidden deep level information, e.g., intention [61]. In the real application of humans-robots interaction, the multi-agent learning algorithms (e.g., Minimax-Q Learning Algorithm and Nash-Q Learning Algorithm [75][76]) has been successfully applied in some humanoid robot, e.g., COMAN [77] and Nao [78][79], showing the evidence that the proposed C-NCFQ based behavior adaptation mechanism may be welcomed to employ to humanoid

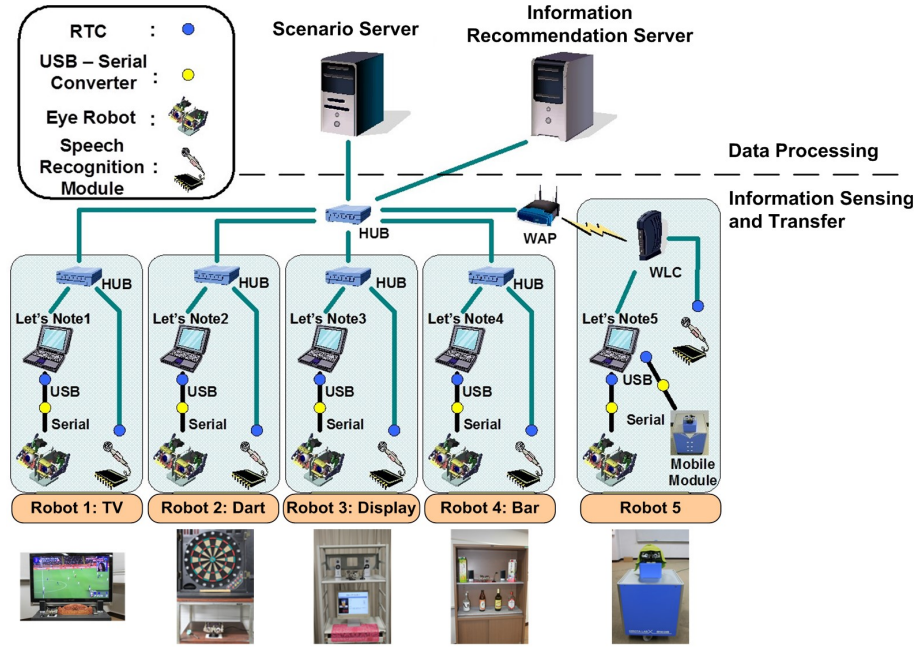


Figure 4.14: Network system of MRS.

robots' behavior adaptation in humans-robots interaction.

4.6 Chapter Summary

Experiments are performed in the virtual scenario of watching football match, where there are 6 virtual eye robots (eye movement is the only behavior but simple and typical) and 6 virtual human agents.

To confirm the validity of the proposed cooperative-neutral-competitive fuzzy-Q learning (CNCFQ) based behavior adaptation mechanism, a comparative experiment among the CNCFQ, the fuzzy production rule based friend-Q learning (FPRFQ) [26], and the friend-Q learning (FQ) [54] is carried out, where robots adapt to the local and global co-existent human-generated atmosphere. The adaptation of the proposed mechanism saves 47 learning steps (i.e., leaning rate is increased by 72%) compared to that of the FPRFQ, and 103 learning steps (i.e., leaning rate is increased by 85%) compared to that of the FQ, and the robots' response time is about 3 times and 6 times shorter than that of the FPRFQ and the FQ, respectively. Moreover, the proposed mechanism obtains about 7 times the reward compared to that of the FPRFQ and 15 times the reward compared

to that of the FQ, respectively, where the distance between the global human-generated atmosphere and the global robot-generated atmosphere is 3 times and 7 times shorter than that of the FPRFQ and the FQ, respectively.

According to the experiment results, a smooth humans-robots interaction is obtained since the response time is reduced and the social distance between humans and robots is closer. By knowing the intensity of each local human-generated atmosphere, in other words, the dominant group is determined, then other groups will quickly follow it soon so that the response time is reduced to make the interaction more smoothly. The proposal is confirmed by experiments in the case of cooperation, neutral, and competition, which may be extended to other interactive relationships in the communication. In the simulation, the response time of robots is millisecond (0.01s by using the proposal), even if in practical robots application, e.g., in the MRS (five eye robots and four humans) [24], the robots CPU is ARM STM32F103 (72MHz) and about 40 times slower than that of the PC employed in the experiment, but the response time is still in sub-second, indicating that the proposed mechanism can reduce the response time of robots in real application. In addition, the 3D FA (Fuzzy Atmosfield) [23] space is divided into eight quadrants. In the experiment, the global atmosphere changes from the seventh quadrant (i.e., hostile, calm, formal area) to the neighborhood of the origin (i.e., neutral, neutral, neutral area), and finally to the first quadrant (i.e., friendly, lively, casual area). Since the adaptation of the proposed mechanism is confirmed in the first quadrant (all positive), the seventh quadrant (all negative), and the neighborhood of origin (all direction), the effectiveness of the proposed adaptation mechanism may be confirmed in the general quadrant case.

Graphic visualization of the local and global atmosphere is evaluated by subjective estimation, where thirty-five volunteers from eight countries at different university in Japan who participate in the questionnaire. Results show that the wave obtains the highest accuracy of 91.7% since it is easy to understand the relationship between wave and the feeling of “weak” and “strong”, shape obtains the accuracy of 85.5%, the color obtains 77.2% accuracy, and the fill obtains the lowest accuracy of 65.3% caused by misunderstanding of the relationship between the length and the feelings of “casual” and

“formal”. Moreover, comparison between the proposed graphic visualization and that in [23] is carried out, 72.5% volunteers support the proposal since it is more impressive and gives more information (e.g., intensity of the local atmosphere).

In term of the future study, the proposed mechanism can be developed for robots behavior adaptation to international communication atmosphere generated by people from different countries. Besides emotion and atmosphere, it would be interesting to extent the adaptation mechanism to other hidden deep level information, e.g., intention [61]. In some real application of humans-robots interaction, multi-agent learning algorithms has been successfully applied in humanoid robots, e.g., COMAN [77] and Nao [78][79], from which we infer that the proposed CNCFQ based behavior adaptation mechanism would be welcomed to humanoid robots’ behavior adaptation in humans-robots interaction.

Chapter 5

Intention Understanding by Using Emotion-Age-Gender-Nationality Information

5.1 Introduction

As the gradual increase of the robotics community in social communication, robots that exist primarily to interact with people [80]. Humanoid robots, both cognitive and physical, become companions of humans, which open up an important research domain to develop human-robot interaction (HRI) [63]. Rapid advancement of robotic technology is bringing robots closer to everyday environments of people, such as home, school, and supermarket. Consequently, interaction between people and robots has become increasingly socially situated and multi-faceted [1]. Office roam-robot is designed to socially interact with people in daily lives [61] and socially assistive robot system is developed for elderly people in physical exercise [81], which indicate not only the benefits of having robots interact with people, but also the need for the interactions to be smooth and natural.

Robotic systems is developed to reflect the progress in different fields of robotics with regard to adaptation [10], learning [82], and autonomous behaviors [83]. Many researches conduct unveiling the importance of properly designed adaptive human-robot interaction strategies to adjust the robots' behaviors for smooth communication [26][84][85]. Still,

HRI is often not so easy because robots lack understanding of the people's internal states, e.g., intentions, rather than just social behaviors, and vice versa.

Intention understanding is essential for many aspects of cognition, culture and social, including image intention understanding [39], speech intention understanding [40], and gesture intention understanding [32]. Although speech and gesture play important roles in intention understanding, they cannot intuitively reflect people's internal states. Emotion is an important missing link in the intention-behavior gap [41], and it is essential for social behavior [42]. In addition, emotion is suggested as the driving force behind motivation on behavior. From a psychological perspective, emotion is more associated with people's internal state, i.e., intention. Therefore, emotion is supposed to be a key component to deeply reflect people's internal states, and how to use individual emotion for intention understanding becomes necessary and feasible, which is also a new challenge.

An intention understanding model is proposed to deeply comprehend people's inner thoughts according to the emotion, where age, gender, and nationality are additional informations for dealing with the problem of various people coexistence. Since unfamiliar issues become familiar through cognitive process, robot is firstly able to distinguish between familiar and unfamiliar people. In the case of familiar people, robot may understand his/her intention through memory, where the memory retrieval is based on the fuzzy production rules; in the case of unfamiliar people, an intention understanding model based on the two-layer fuzzy support vector regression (TLFSVR) is introduced to deeply comprehend people's inner thoughts. The TLFSVR is an extension of support vector regression (SVR) [14], in which two-layer consists of local learning layer and global learning layer. Local learning layer includes fuzzy c-means based input data classification and SVR based learning, and global learning layer outputs the intention of TLFSVR based on fuzzy weighted average algorithm. In view of the local learning [27], the training data is split into multiple subsets, so that the traditional one SVR is extended to multiple SVRs, for example, considering various people with different ages, different genders, and different nationalities, which may lead to the problem of health, gender hobby, religion, and so on. Different regression models reflect different ages/genders/nationalities, which

is better than only one regression model covers all situations. Moreover, age, gender, and nationality play important role in intention understanding, for example they affect people's intention to purchase products or services [28], accept new things [29], and handle risky situation [30]. Thus, three components of human property, i.e., age, gender, and nationality, are used for intention understanding beside emotion, then training data can be classified according to the identification informations. Finally, fuzzy inference based on identification informations is used to generate the intention, which is either output of memory retrieval or the TLFSVR.

The proposal enables robot to deeply understand people's intention, and robot can transparently communicates to the user and guess what he/she expects. Transparent communication means one accesses the comprehension and memory for understanding the internal states of the other, e.g., intention [86][87]. Transparently communicating to each other provides a new communication way which is from heart to heart, that is, our inner expectative intention is possible to carry out by emotion and some negative impact could be avoided based on different ages/genders/nationalities, which is a boost for smooth communication.

To show the effectiveness of the proposed intention understanding model, simulation and application experiments are done. Simulation experiment is designed to verify the accuracy of the proposal, where the simulation environment is built by Matlab in a PC employed with a dual core processor (2.8GHz), memory (2.99GB), and Windows 7 system. The training and testing data is collected from the information acquisition system, including the emotion recognition module and questionnaires module, which is created by C# in Visual Studio 2013. In addition, to confirm the validity and practicability of the proposed intention understanding model, application experiment of humans-robots interaction is conducted in the developing mascot robot system (MRS) [88], where a scenario of "drinking at bar" is performed by six salary men and two eye robots.

5.2 Characteristics Analysis

Using brainstorming, twenty-eight volunteers (i.e., different genders people aged 20-65 years, from 6 countries in Tokyo, as shown in Fig. 5.1) discuss topic like “what are the influence factors of intention understanding?”, and the presentation of the brainstorming is shown in Fig. 5.2, where 9 questions are asked by volunteers during the brainstorming, questions/responses and the information of questioners/responders including age, gender, nationality, and religion are given.

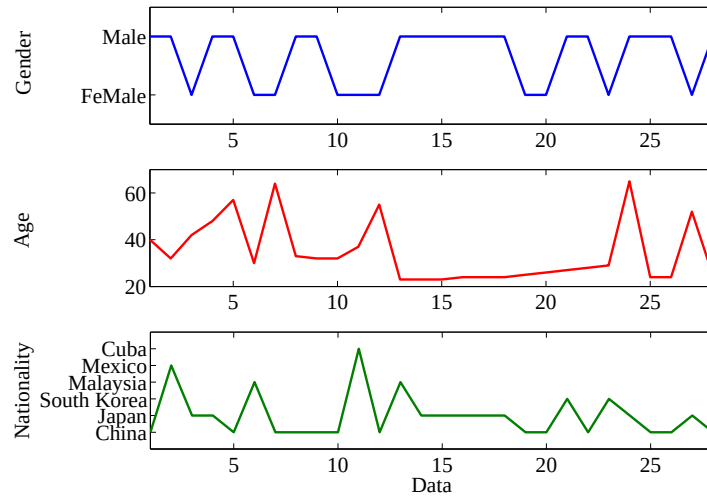
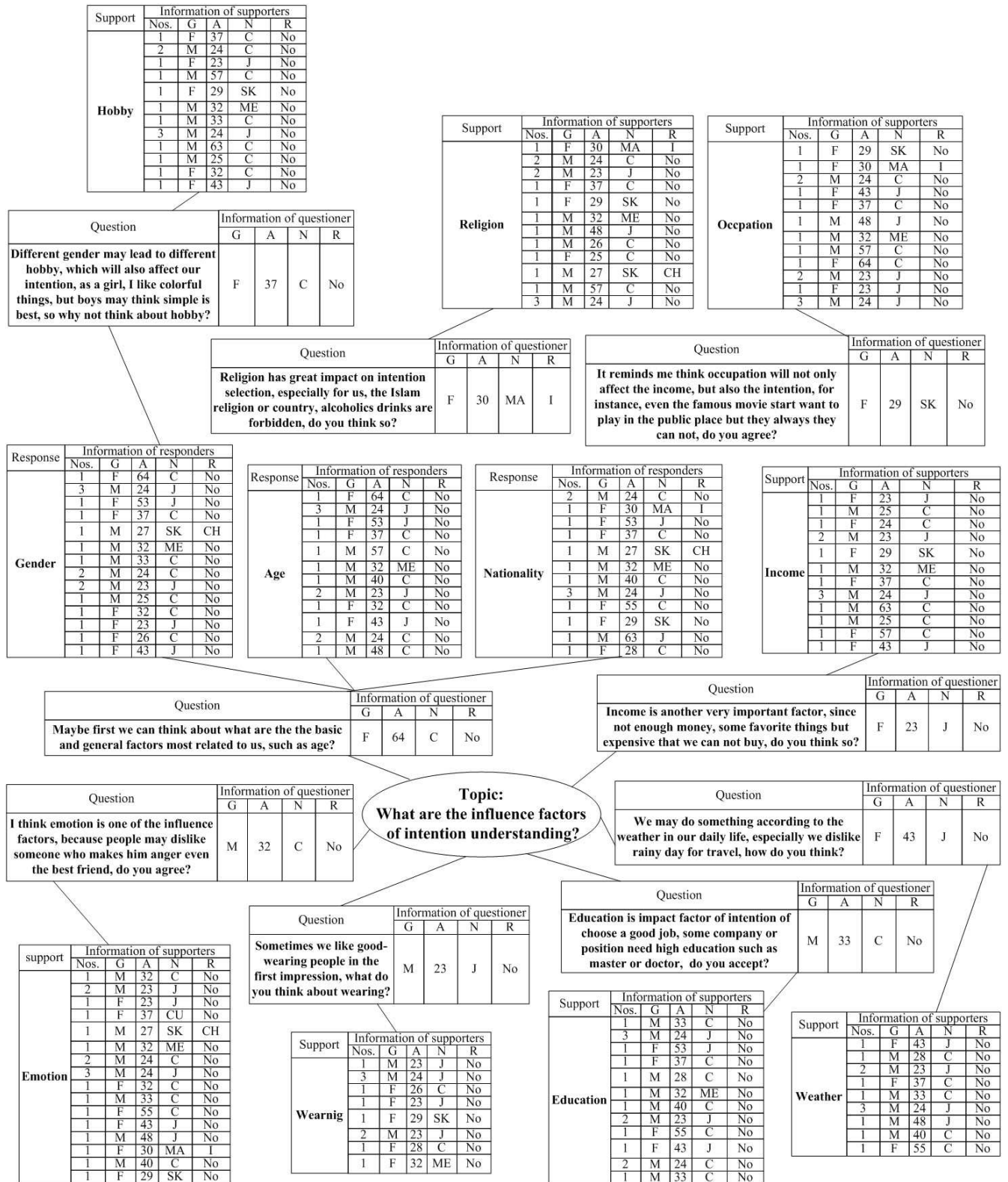


Figure 5.1: Identification information of volunteers.

According to the data analysis of brainstorming in Table 5.1, for influence factor “age”, supporters are from different ages in 20-65; for influence factor “gender”, supporters include both female and male; for influence factor “nationality”, supporters are from 6 different countries; for influence factor “religion”, supporters have different religions including Islam and Christian. Similarly to other factors, all the supporters are in different ages, genders, and nationalities, which indicates that the response of supporters are reasonable and significant, and 11 influence factors are collected from the brainstorming according to Table 5.1.

As the discussion continues, the participants’ passion slowly weaken, as well as the numbers. Moreover, the brainstorming is going to no-good situation, which the numbers of responders is less than half while discussing about the “weather”, even under 40% in



* A-Age, G-Gender, N-Nationality, R-Religion, F-Female, M-Male, C-China, J-Japan, SK-South Korea, MA-Malaysia, ME-Mexico, CU-Cuba, I-Islam, CH-Christian, No-No Religion

Figure 5.2: Presentation of brainstorming.

the discussion of “wearing”. In this case, the brainstorming need to stop, and 9 factors which the support rate higher than 50% will be chose as initial influence factors of human intention understanding, including age, gender, nationality, education, emotion, religion, occupation, income, and hobby.

Table 5.1: Result analysis of brainstorming

Influence factor	Support rate	Information of supporters								
		Age				Gender		Nationality	Religion	
		< 30	30 ~ 40	40 ~ 50	> 50	Female	Male		Islam	Christian
Age	57.1%	7	3	3	3	5	11	4	0	0
Gender	64.3%	11	4	1	2	7	11	5	0	1
Nationality	53.6%	8	3	1	3	6	9	6	1	1
Education	57.1%	8	4	2	2	4	12	4	0	0
Emotion	71.4%	10	6	3	1	7	13	6	1	1
Religion	57.1%	11	3	1	1	4	12	6	1	1
Occupation	57.1%	9	3	2	2	6	10	6	1	0
Income	53.6%	10	2	1	2	5	11	5	0	0
Hobby	53.6%	8	4	1	2	5	10	5	0	0
Weather	42.9%	6	2	3	1	3	8	3	0	0
Wearing	39.3%	10	1	0	0	5	6	4	0	0

To select the independent factors for intention understanding model, first, twenty-eight volunteers (identification information are shown in Fig. 5.1) complete questionnaires about the correlation between collected relevant factors and the intention, and the correlation coefficient is from 0 to 1, e.g., correlation coefficient of age and order intention is 0.9, as shown in Table 5.2. The answer of each volunteer about the correlation between influence factors and intention are shown in Fig. 5.3. According to the basic information of volunteers in Table 5.3, it shows that they are different ages, different genders, different nationalities, different educations, different emotion (based on Pleasure-Arousal plane as shown in Fig. 5.7), different religions, different occupations, and different income, which indicate the responses are signification in reflecting the correlation between influences factors and intention.

Table 5.2: Questionnaire for correlation between influences factors and intention

Questions	Answers
How big is the correlation between influences factors IF and your intention to do something? ($IF \in \{\text{Age, Gender, Nationality, Education, Emotion, Reliction, Occupation, Income, Hobby}\}$)	Any number from 0 to 1

Based on the correlation values in Fig. 5.3, nine intention related factors are classified into four clusters using the hierarchical clustering algorithm [89], as shown in Fig.5.4.

Fig. 5.4 shows the dendrogram of factor classification, y axis is cluster distance from

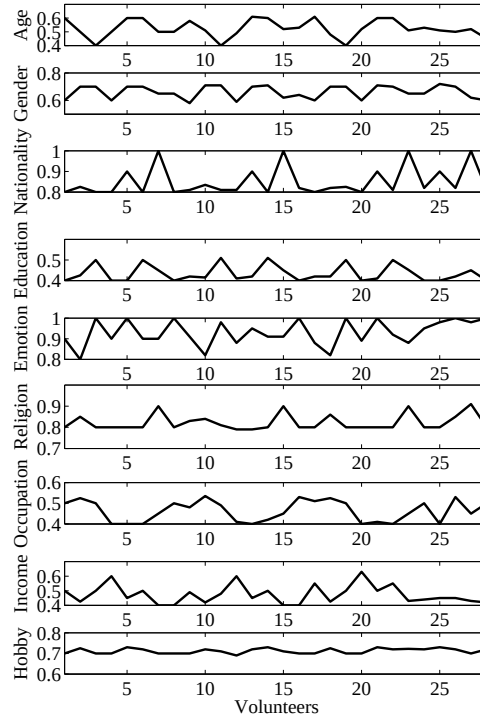


Figure 5.3: Correlation between influence factors and intention.

0 to 1, relationship become weak as the distance increase, with regard to this, 4 clusters are selected, where nationality and religion are classified in cluster 1; gender and hobby are classified in cluster 2; emotion is classified in cluster 3; age, education, occupation, and income are classified in cluster 4, and each cluster is relatively independent.

Table 5.4 shows the correlation coefficient between each factors. Inputs of intention understanding model is selected based on the principle of minimum correlation and different clusters. According to Fig.5.4 and Table 5.4, age, gender, nationality, and emotion

Table 5.3: Basic information of volunteers

Influence factor	Cluster	Numbers	Influence factor	Cluster	Numbers
Age	< 30	14	Education	Master	13
	30 ~ 40	6		Doctor	15
	40 ~ 50	3		Islam	1
	> 50	5	Religion	Christian	1
Gender	Female	11	Occupation	No	26
	Male	17		Student	22
	China	13		Teacher	3
Nationality	Japan	10		Salary men	3
	South Korea	2	Income	Scholarship	12
	Malaysia	1		salary	6
	Mexico	1		No	10
	Cube	1			

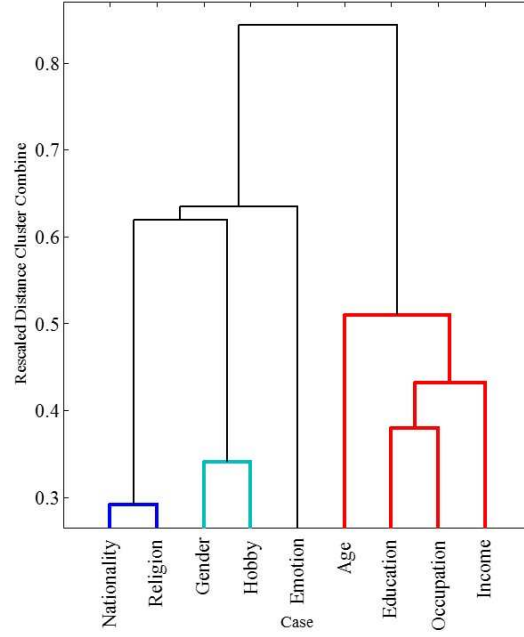


Figure 5.4: Dendrogram of intention related factors relationship.

are in different clusters, and the correlation coefficient between each other are -0.0742 (age and gender), 0.0576 (age and nationality), -0.1761 (age and emotion), 0.0115 (gender and nationality), 0.1322 (gender and emotion), and 0.0517 (nationality and emotion), respectively, which indicates that age, gender, nationality, and emotion have little or no relationship to each other.

Moreover, a one-way ANOVA is used to compare the means of different age/gender/nationality groups to determine whether they are significantly different from each other, and the observation parameter are correlation values between age/gender/nationality and intention in Fig. 5.4. According to the data sample in Table 5.5, 3 groups are classified based on different ages, 2 groups are classified based on different genders, 2 group are classified based on the nationality of China or Japan, and each group has 6 data. The null hypothesis is all the age/gender/nationality groups have equal influence on the intention. By using one-way ANOVA, boxplots of intention correlation by age/gender/nationality are shown in Fig. 5.5, respectively. According to ANOVA summary in Table 5.6, all the P-values in 3 samples are lower than 0.05, thus the null hypothesis is rejected, that means age, gender, and nationality have signification influence on the intention. Since the independence and signification of age, gender, nationality, and emotion, they can be

chose as the inputs of the intention understanding model.

In addition, age is selected since not only it is the accumulation of changes in a person over time, where education, occupation, and income change with the age, but also some legal age are existed, such as drinking age and driving age, voting age and so on; gender refers to female and male, it is selected because a broader set of hobby can be obtained from the gender point of view; nationality is selected rather than religion, because it indicates the legal relationship between a person and a state, and the person should obey the state law, such as people are illegal to drink alcohol in some Muslim country, in such a way that, nationality includes religion. With these in mind, age, gender, nationality and emotion are suitable inputs for the intention understanding model.

Table 5.4: *Correlation coefficients among related factors*

	age	gender	nationality	education	emotion	religion	occupation	income	hobby
age	1.0000	-0.0742	0.0576	-0.2075	-0.1761	-0.1556	-0.4176	0.1355	0.2825
gender	-0.0742	1.0000	0.0115	0.4301	0.1322	-0.0871	-0.0191	-0.2819	0.6820
nationality	0.0576	0.0115	1.0000	0.0082	0.0517	0.7508	-0.2641	-0.4680	0.1320
education	-0.2075	0.4301	0.0082	1.0000	0.0527	0.0986	-0.0999	0.0834	0.0351
emotion	-0.1761	0.1322	0.0517	0.0527	1.0000	-0.2952	-0.0737	-0.1493	-0.0404
religion	-0.1556	-0.0871	0.7508	0.0986	-0.2952	1.0000	0.1847	-0.5078	0.0043
occupation	-0.4176	-0.0191	-0.2641	-0.0999	-0.0737	0.1847	1.0000	-0.4705	-0.1578
income	0.1355	-0.2819	-0.4680	0.0834	-0.1493	-0.5078	-0.4705	1.0000	-0.3308
hobby	0.2825	0.6820	0.1320	0.0351	-0.0404	0.0043	-0.1578	-0.3308	1.0000

Table 5.5: *Data samples of ANOVA*

Independent factor	Correlation between factor and intention						Information of responders					
Age	0.45	0.4	0.48	0.4	0.5	0.4	24	25	23	26	24	29
	0.5	0.51	0.49	0.52	0.53	0.51	32	33	32	37	32	30
	0.6	0.61	0.6	0.58	0.6	0.61	63	57	48	40	55	64
Gender	0.58	0.6	0.62	0.59	0.64	0.62	Male					
	0.71	0.7	0.72	0.65	0.71	0.65	Female					
Nationality	0.81	0.82	0.8	0.82	0.825	0.8	China					
	0.9	1	0.9	0.82	0.9	0.835	Japan					

Table 5.6: *ANOVA summary of (a)/(b)/(c)*

Source	SS	df	MS	F	P _{rob} > F
Columns	0.07874/0.02001/0.0192	2/1/1	0.03937/0.02001/0.0192	50.55/26.74/9.2	2.16016 × 10 ⁻⁷ /0.0004/0.0126
Error	0.01168/0.00748/0.02088	15/10/10	0.00078/0.00075/0.00209		
Total	0.09043/0.02749/0.04008	17/11/11			

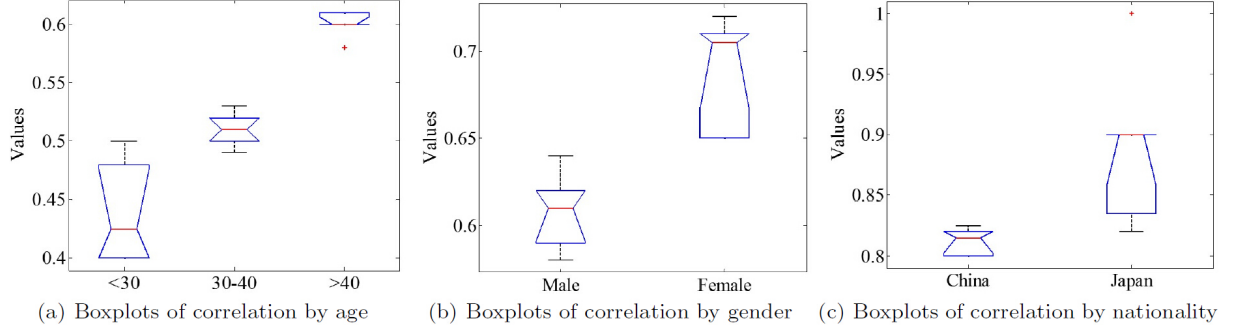


Figure 5.5: Boxplots of intention correlation by age/gender/nationality.

5.3 Intention Understanding Model

5.3.1 Architecture of Intention Understanding Model

Based on the TLFSVR introduced in Section 2.4.1 and characteristics analysis, an intention understanding model is proposed to comprehending peoples' inner thoughts in human-robot interaction as shown in Fig. 5.6. The understanding model reflects empirical findings accumulated in cognitive (memory), at the same time, machine learning and fuzzy logic indicate the relationship of human intention and emotion. The basic idea is that the intention understanding model may support the deeply level understanding based on the emotion, where age, gender, and nationality are the additional informations. It mainly includes four modules, i.e., ID mapping, fuzzy rules based memory retrieval, two-layer fuzzy support vector regression (TLFSVR) based intention understanding, and fuzzy inference based intention generation. In addition, the information acquisition consists of emotion recognition and questionnaire. Emotion recognition is a supported module for intention understanding, which can read people's emotion from facial expression and speech.

Questionnaire is a survey that contains the question Q and the answers A , where the identification information (i.e., ID, age, gender, and nationality) is collected as identification informations for intention understanding.

ID mapping module is used to distinguish between familiar and unfamiliar people based on the identity card (ID). Since unfamiliar issue usually becomes familiar through

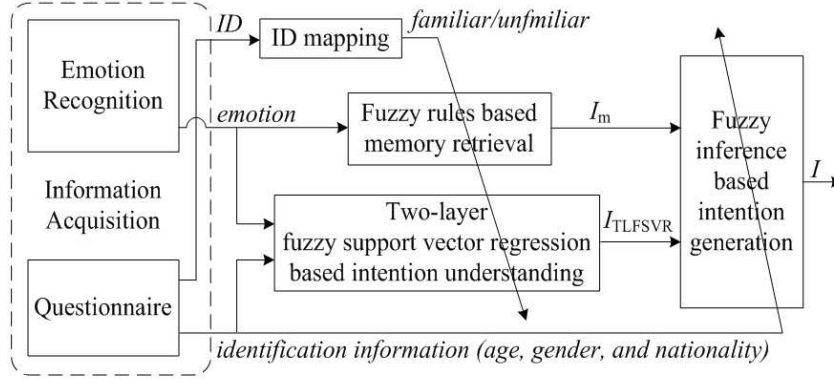


Figure 5.6: TLFSVR based intention understanding model.

cognition process in human life, to some extent, people can be classified into familiar and unfamiliar.

Fuzzy rules based memory retrieval module aims to understand familiar people's intention directly from memory.

TLFSVR based intention understanding module is a SVR learning machine fusion with fuzzy logic, which calculates the relationship from emotion to intention with the identification informations, i.e., age, gender, and nationality.

Fuzzy inference based intention generation module is based on the identification informations (i.e., age, gender, and nationality), which is fuzzy selection operation from the output of memory retrieval module and TLFSVR based intention understanding module by using "IF-THEN" fuzzy rules.

5.3.2 Information Acquisition

1. Emotion recognition

For understanding people's intention, emotion recognition is a supported module based on the previous study [23], which is realized using weighted fusion fuzzy inference based on the bimodal information, i.e., facial expression and speech. Emotion is recognized by fusing semantic cues from speech and facial expression, and represented in the Pleasure-Arousal plan, where happiness, surprise, fear, anger, disgust, sadness, and neutral are employed as the basic human emotion, totally 25 kinds of emotions, as shown in Fig. 5.7.

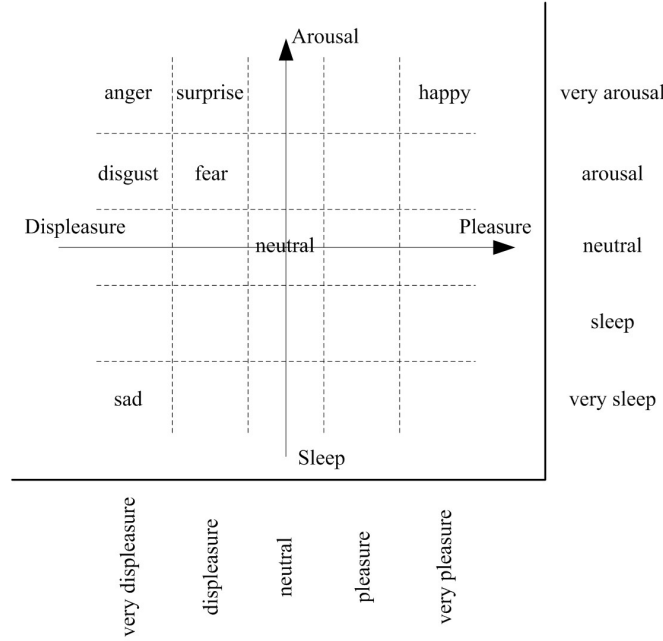


Figure 5.7: *Pleasure-Arousal emotion plan.*

2. Questionnaire

To collect the identification information (i.e., ID, age, gender, and nationality) for intention understanding, questionnaire is designed, which consists of a number of questions Q and their corresponding answers A .

In the questionnaire, questions Q contains “what is your name? / how old are you? / what is your gender? / what is your nationality?”, and answers A are opened-ended (responding by a specific piece of information). More specifically, ID is used to distinguish the people between familiar and unfamiliar in the human-robot interaction, while age, gender, and nationality are the identification informations for intention understanding.

3. ID mapping

ID mapping is designed to transform the people name into different types of people, and realized by fuzzy rules such as

$$\begin{aligned} \text{If } NA == \text{NULL} \quad \text{Then } T &= 0 \\ \text{If } NA \neq \text{NULL} \quad \text{Then } T &= 1 \text{ and } NA \end{aligned} \tag{5.1}$$

where NA is the people name, which is saved as string, and NULL means NA is not involved in the data base; T indicates the type of people, 0 means unfamiliar, and 1

means familiar. After ID mapping, people are classified into two situations according to the memory, i.e., unfamiliar people and familiar people. If he/she is a familiar people, fuzzy rules based memory retrieval module is used for intention understanding or else the TLFSVR based intention understanding module is used to generate human intention for the unfamiliar people.

5.3.3 TLFSVR based Intention Understanding

In the case of unfamiliar people, robot understands people's intention base on the introduced TLFSVR, which is the local learning algorithm with multiple regression functions. Using the TLFSVR, robot can handle the multiple classes of input, where different regression functions corresponding to different classes of input. The TLFSVR based learning adaptation mainly consists of 7 steps:

Step 1: Initialization of FCM clustering

Parameters, i.e., the number of clusters C , the sensitivity threshold ϵ , and the overlap constant η are set.

Step 2: FCM calculation

The membership u_{ik} is calculated by (2.12), the cluster centers V_k is calculated by (2.13), and the spread width δ_k^o is calculated by (2.14).

Step 3: If the stop criterion is not satisfied, then go to Step 2; otherwise, go to Step 4.

Step 4: FCM clustering

The training subsets D_k are obtained by (2.15) based on the cluster centers and spread width.

Step 5: Initialization of SVR

Parameters, i.e., the Gaussian kernel width σ and the scalar regularization parameter γ are set.

Step 6: SVR learning

Regression function of each cluster SVR_k is constructed by SVR learning approach.

Step 7: Intention output

$$R_{i \times j \times k} : \quad \text{If } x_1 == NA_i \text{ and } x_2 == E_{ij} \quad \text{Then } I_{jk} \text{ with } GR_{jk} \quad (5.2)$$

Intention understanding of the TLFSVR is outputted base on the fuzzy weighted average algorithm by using (2.18).

5.3.4 Memory Retrieval for Intention Understanding

Since unfamiliar issue usually becomes familiar in human cognition process, people may be classified into two cases, i.e., one is unfamiliar people and the other is familiar people. Memory retrieval module is used to generate intention in the case of familiar people. The memory module stores the personally information of people, e.g., name, emotion, and intention. The fuzzy production rules of the memory rule base are Mamdani fuzzy rules designed as where $i \times j \times k$ is the number of rules, x_1 and x_2 are inputs, NA_i is the i th people name, E_{ij} is j th emotion of i th people, I_{jk} is k th intention in j th emotion of i th people, and GR_{jk} is the generation rate of the corresponding intention. For example, a scenario of “drinking at a bar” in human-robot interaction, suppose robot is a bartender and a familiar customer named FC with happy emotion, who drinks the beer when felt happy last time, before ordering, since memory retrieval function of the robot, one of the fuzzy rules is selected, which is defined as “If $x_1 = \text{FC}$ and $x_2 = \text{happy}$ Then $I = \text{beer}$ with $GR=1$ ”.

Several intentions may be suitable for the expected emotion, taking into account this problem, fuzzy rule is selected based on the generation rate of the intention. If people do a new intention to the expected emotion, then a rule is added to the memory rule base. Subsequently, the generation rate ($GR \in [0, 1]$) of the rule is updated when the same rule is used again and is calculated in terms of time-weighted statistics as

$$GR = \frac{UT}{UT + OT} \quad (5.3)$$

where $UT \in N$ is the number of use times of k th intention in j th emotion, and $OT \in N$ is the number of use times of other intentions in j th emotion. Add or delete fuzzy rules depend on two thresholds, i.e., the upper bound ($B_{upper} = 1$) and the lower bound (B_{lower}

= 0.8). The generation rate of effective rule must be between these two boundaries, otherwise, it will be deleted.

5.3.5 Intention generation by fuzzy inference

Base on the identification informations, i.e., age, gender, and nationality, output of intention understanding model is designed by the look-up table, as shown in Table 5.7.

Table 5.7: Look-up table for output

Age	Nationality	
	Islam country or religion	Others
$< A_{lower}$	non-alcoholic drinks	non-alcoholic drinks
$A_{lower} \sim A_{upper}$	non-alcoholic drinks	I_m/I_{SVR}
$> A_{upper}$	non-alcoholic drinks	non-alcoholic drinks

where A_{lower} is the legal drinking age according to different countries, e.g., 18 in China, 20 in Japan, 19 in South Korea, and so on. A_{upper} is the healthy drinking age, e.g. low-risk drinking style for healthy life is begin from age 65 [52].

5.4 Experiments on Intention Understanding of Customer's Order at a Bar

5.4.1 Environment Setting

To validate the proposed intention understanding model, simulation experiment and application experiment are both done. Both simulation and application choose the scenario of “drinking at a bar” as the background, this is due to the prevalence of the bar culture in the word, especially for youth and salary man. In Japan, bar is endowed with a new form that is a type of Japanese drinking establishment, called Izakaya. Izakaya is a casual places of after-work drinking and very popular for the salary man, even becomes one part of their life. If the owner of an Izakaya understands each customer preferences, e.g., a right drink for specific mood, it would be very nice and deeply touched the tired salary man, by which it will also increase the turnover. The customer preferences prob-

lem seems to be a very interesting research, so the goal of the simulation and application experiments is understanding of the customer's intention to order at a bar based on the customer's emotion.

Taking the complexity of intention understanding process into account, simulation experiments are first performed, which not only can run repeatedly to simplify debugging of the intention understanding process, but also can simulate various emotions that are seldom experienced in real life. Simulation experiment is designed to verify the accuracy of the proposal, and is conducted in the Matlab environment, which is ran in a PC with dual core processor (2.8GHz), memory (2.99GB), and Windows 7 system. The training and testing data is collected from the information acquisition system, including the emotion recognition module and questionnaires module, which is created by C# in Visual Studio 2013.

The application experiment is developed to confirm the validity and practicability of the proposed intention understanding model in the developing Mascot Robot System (MRS) [90]. MRS is a typical application of humans-robots interaction, which is an information presentation system destined to be used mainly as an information terminal for communication with humans. MRS is now developing to the salary mans life for communication between robot and salary man [91].

In the MRS, five eye robots play different roles in the life of salary man, i.e., four fixed eye robots act as secretary, colleague, bar lady, and customer, respectively; an autonomous mobile eye robot acts as son of the salary man. Robot technology middleware (RTM) is used to construct the network system of MRS called RTM-Network, where each robot can be viewed as a network component and each function unit of every robot is called robot technology component (RTC), as shown in Fig. 5.8.

RTM-Network is a two-layer network system, including the information sensing and transfer layer as well as data processing layer. There are eight modules in the RTM-Network, including the speech recognition module (SRM), facial expression recognition module (FERM), gesture recognition module (GRM), eye-robot control module (ERCM), emotion processing module (EPM), display module, mobile robot control module, scenario

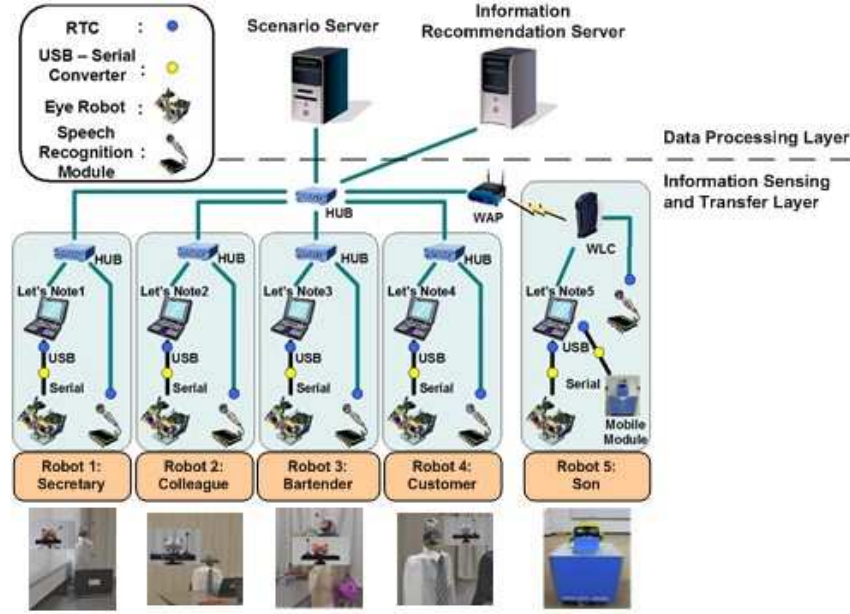


Figure 5.8: Network system of MRS.

server module, and information retrieval module [92]. The SRM sensing the speeches information through microphone, the FERM and GRM sensing the facial expressions and gestures information using Kinect. Moreover, the EPM obtains the recognized results from the SRM, the FERM, and the GRM, and then transmits them to the server module for the further processing, such as the calculation of individual emotions. Finally, each robot understands people's intention according to the obtained emotion. In the preliminary application, the FERM is mainly used for emotion recognition, and only six basic emotion including happiness, surprise, fear, anger, disgust, sadness are used to represent human emotions, since MRS has rich experience on recognition of these six basic emotion [22, 93, 23], the FERM module is supposed to recognize the six basic emotion exactly.

5.4.2 Data perpetration

Data preparation is designed based on the scenario of "drinking at a bar". Data sample for simulation experiment is collected from thirty-two volunteers by using the information acquisition system, as shown in Fig. 5.9. Information acquisition system is created by C# in Visual Studio 2013, which includes the emotion recognition module and questionnaires module.

Questionnaires are about the relationship between emotion and intention to order at a bar, at the same time, write down the identification information, such as age, gender, nationality, as shown in Table 5.8. For easy understanding and selecting, questionnaires is transformed into questionnaires module in Fig. 5.9, including the identification information and order intention selection. The order intentions mainly includes the top six popular drink in the Izakaya.

Emotion recognition is realized by facial expression, where twenty-five kinds of emotions in total according to the Pleasure-Arousal emotion space, as shown in Fig.5.10. By using the emotion recognition module, it can not only directly give the recognition result, but also can add any one of the 25 kinds of emotion and the corresponding facial expression states to the Access 2013 database. Since the emotion recognition module has the function of emotion recording, in the case of having the facial expression data records of volunteers, emotion recognition accuracy is guaranteed and will not affect the intention understanding accuracy.

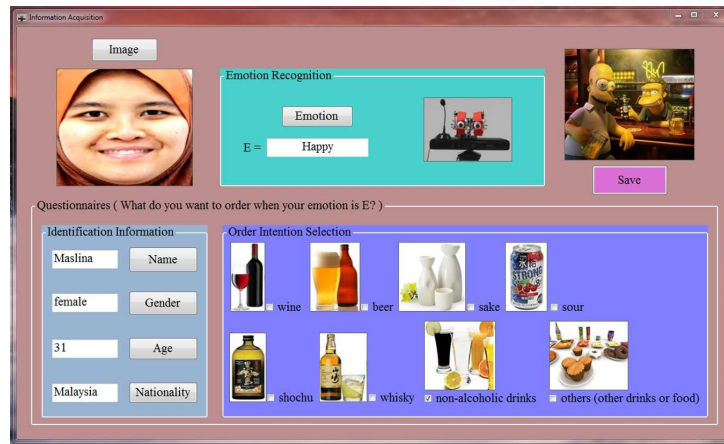


Figure 5.9: Information acquisition.

Table 5.8: Questionnaire for relationship between emotion and intention to order at a bar

Questions	Answers
What do you want to order when your emotion is E ? ($E \in \{\text{anger, disgust, sadness, surprise, fear, happiness, etc.}\}$, 25 questions corresponding to 25 kinds of emotions in total, as shown in Fig. 5.7)	1-Wine, 2-beer, 3-sake, 4-sour, 5-shochu, 6-whisky, 7-non-alcoholic drink, 8-others (other drinks or food)

Volunteers contains five salary men aged 30 to 65 years old and seventeen postgraduate



Figure 5.10: *Emotion recognition.*

students aged 20 to 35 years. They are in different genders, and from six countries (i.e., China, Japan, South Korea, Malaysia, Cuba, and Mexico) in Tokyo. Moreover, to get more valid results, more data are collected by asking ten actual customers at the Izakaya named Watami in Tokyo, and they are aged 20-60 years old, including 6 male and 4 female, mainly from China and Japan. They ordered randomly, and if it is beyond the popular six drinks as shown in Table 5.8, the intention will be collected as 7-non-alcoholic drink or 8-others (other drinks or food). A total of 800 groups of data are collected from thirty-two volunteers, where 600 groups for training and 200 groups for testing, as shown in Fig. 5.11 and Fig. 5.12.

N-way ANOVA is used to examine the influence of emotion/age/gender/nationality on intention by using the collected data, where the independent variables are emotion, age, gender, and nationality, and the dependent variable is intention. The null hypothesis is all the emotion/age/gender/nationality groups have equal influence on the intention. According to analysis of variance in Table 5.9, all the P-values in 4 independent variables

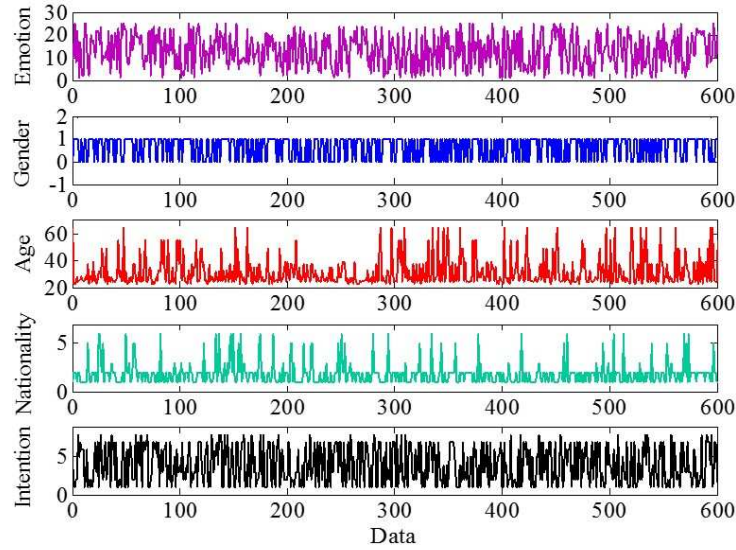


Figure 5.11: Training data for simulation.

are lower than 0.05, thus the null hypothesis can be rejected, that means emotion, age, gender, and nationality have signification influence on the intention. Since the collected data can proof the statistical significance of emotion, age, gender, nationality, they can be applied to the simulation.

Table 5.9: Analysis of variance

Source	Sum Sq.	d.f.	Mean Sq.	F	$P_{\text{rob}} > F$
Emotion	492.62	24	20.256	5.92	1.1102×10^{-16}
Age	763.79	13	58.753	16.93	0
Gender	111.44	1	111.435	32.12	2.0632×10^{-8}
Nationality	484.08	3	161.359	46.51	0
Error	2626.5	757	3.47		
Total	4186.87	799			

5.4.3 Experiment on Intention Understanding of Costumer's Order at a Bar

To validate the performance of the proposed intention understanding model, a comparative experiment among the TLFSVR, the SVR, and the BPNN is carried out in simulation, aiming to generate the intention by emotion and discuss the influence of age/gender/nationality on the intention understanding.

Simulation experiment is simulating the situation of “drinking at the bar” based on the

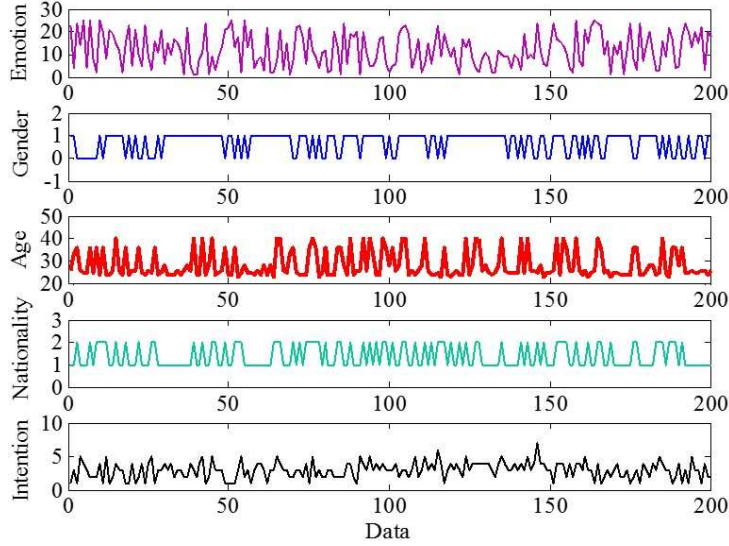


Figure 5.12: Testing data for simulation.

collected data of emotion and order (intention), as well as the identification information (i.e., age, gender, and nationality). The goal is to realize the transparent communication (communication from heart to heart), that is, our inner expectative intention is possible to carry out, and is a boost for smooth communication. In other words, correctly obtaining the customers order by using the proposed intention understanding model base on the customer's emotion and identification information. According to the goal, four indexes are used to evaluate the performance, including accuracy, computational time, average absolute relative deviation (AARD), correlation coefficient, where AARD is defined as

$$\%AARD = \frac{100}{N} \sum_{i=1}^N \left| \frac{y_{cal} - y_{exp}}{y_{exp}} \right| \quad (5.4)$$

where N denotes the number of the testing data, y_{cal} represents the output of the proposal, y_{exp} is the actual value from information acquisition system.

In the first experiment, the number of clusters C for the FCM is chosen as 2 according to the gender (i.e., female and male). In the second experiment, the number of clusters C for the FCM is chosen as 3 according to the age distribution (in general, people aged <30 belong to student, aged 30-55 belong to worker, and age >55 belong to retire). In the third experiment, the number of clusters C for the FCM is chosen as 6 according to the nationality (volunteers for data collection are from 6 different countries). Generally

speaking, the overlap parameter η can be selected as 2.5, 3.5, 4, and 5 corresponding to the dimension of input variables are 1, 2, 4, and 8, respectively. According to the input variables are 4 in the experiment, the overlap constant η and the sensitivity threshold ϵ for the FCM are set to 3 and 1, respectively. For each SVR, the Gaussian kernel width σ is suggested as $(0.1 \sim 0.5) \times (\text{rang of the inputs})$, since the inputs are all normalized, σ is chosen as 0.5, and the scalar regularization parameter γ are chosen as 200. BPNN has three layers (input layer, hidden layer, and output layer units are 4, 12, and 1, respectively), where learning rate is 0.25, and inertia coefficient is 0.05.

Fig. 5.13, Fig. 5.14, and Fig. 5.15 show the output curves and output error of the proposed TLFSVR, SVR, and BPNN with $C=2, 3, 6$, respectively, and the data analysis of result comparison among the three methods is showed in Table 5.10. According to simulation results, while $C=2/3/6$, the proposed TLFSVR obtains an accuracy of 70%/72%/80%, which is 23%/26%/33% and 35.5%/37.5%/45.5% higher than SVR and back propagation neural networks (BPNN), respectively. Computational time is about 0.976s/0.935s/0.67s by using the proposed intention understanding model, indicating that the computing speed is about 1.9/2/2.8 and 3.6/3.7/5.2 times faster than the SVR and BPNN, respectively.

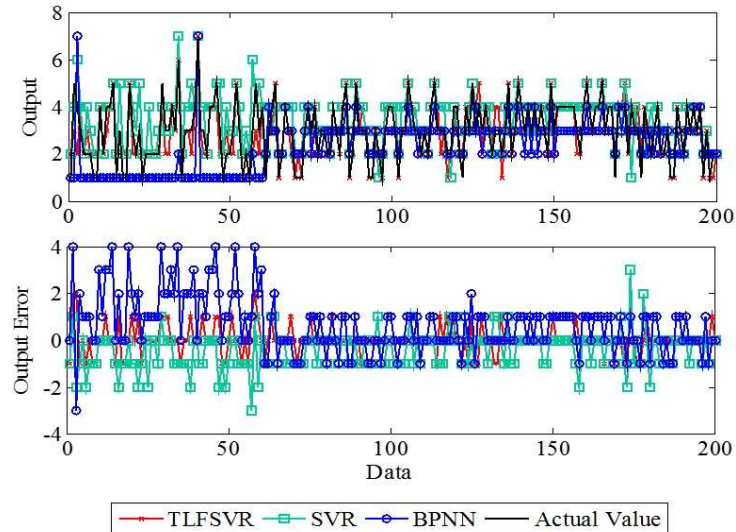


Figure 5.13: Output curves and output error of proposed TLFSVR, SVR, and BPNN with $C=2$.

In the simulation examples, it can be found that the accuracy of the proposed under-

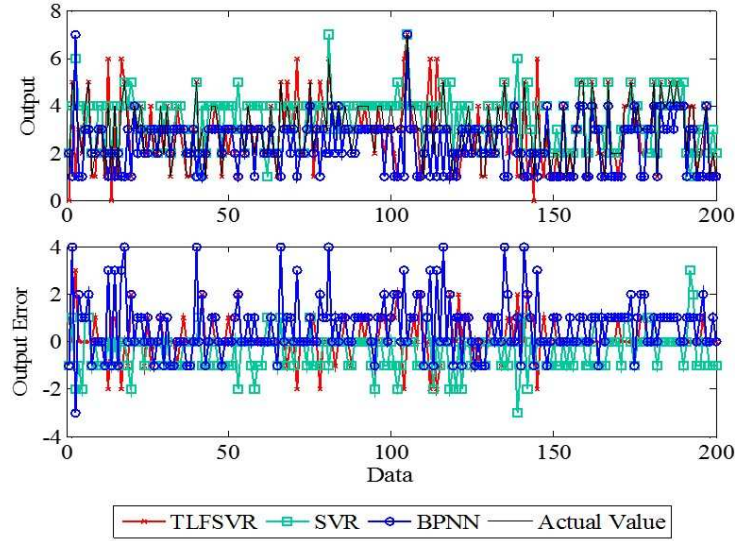


Figure 5.14: Output curves and output error of proposed TLFSVR, SVR, and BPNN with $C=3$.

Table 5.10: Result comparison of simulation for intention understanding

Method		Index			
		Accuracy (%)	Computational time (s)	AARD (%)	Correlation
TLFSVR	C=2	70	0.976	15.450	0.869
	C=3	72	0.935	10.417	0.915
	C=6	80	0.67	10.405	0.928
SVR		47	1.889	31.208	0.771
BPNN		34.5	3.505	33.767	0.456

standing model is higher than that of the SVR and the BPNN, leading to the smaller AARD and higher correlation coefficient between the output and the actual value. Smaller AARD and higher correlation indicate that the TLFSVR based intention understanding model better matches the actual situation. The increment of clusters number C improves the understanding accuracy to some extent. The overlap parameter ϵ is an important parameter in the proposed understanding model, which becomes larger as the dimension of the training data becomes higher. Since the dimension of the input value is four (i.e., emotion, age, gender, and nationality) in the experiments, the overlap parameter ϵ can be set large to cover all the training data, so that the accuracy can be guarantee. Moreover, besides emotion, identification information including age, gender, and nationality are added as supplement for understanding humans' intention according to Table 5.7, especially to deal with miss or false recognition on the hidden emotion such as "curious",

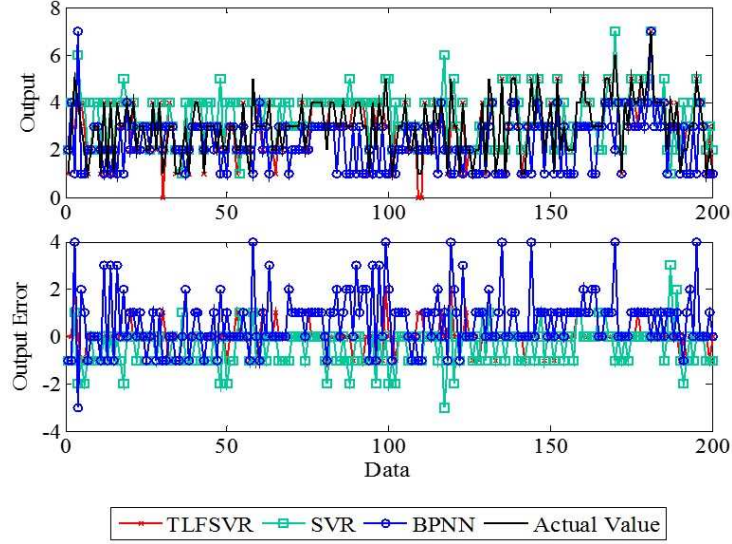


Figure 5.15: Output curves and output error of proposed TLFSVR, SVR, and BPNN with $C=6$.

“forced smile”, “cute”, and so on. For example, in the case of people from Islam country, whatever the emotion, miss or false, they will not choose alcoholic drinks, and by using the proposed intention understanding model, a non-alcoholic drinks is selected according to the nationality information. In addition, the ID information (people’s name) is recording by using the proposal, in the case of familiar ID, the favorite drink (frequently order) is selected. All of these supplementary information are used to reduce the impact of miss or false emotion recognition and guarantee the intention understanding accuracy to some extent.

The computational time mainly includes some processes such as the split of the training data into C training subsets using the FCM clustering algorithm, constructing of all SVRS, and the calculation of the output with a fuzzy weighted average algorithm. The computational time is affected by the computational complexity of the proposal. The computational complexity is explicitly expressed in terms of the number of data and computation scale, for example, n is the number of the training data, n_i is the number of the training data in the i th training subset. C is the number of clusters, and m is the numbers of layers in BPNN, then computational complexity of the SVR is set as $O(n^2)$; for the proposal, there are C times of the SVR conducted with d_i training data, that is $O(\sum_{i=1}^C n_i^2)$; the computational complexity is $O(m \times n^2)$ for BPNN. Since computational

complexity of the proposal is simplest among these three methods, the computational time of it is less than that of the SVR and BPNN.

Table 5.11: Identification information and emotion-intention information of volunteers

Group	ID	Identification information			Emotion	Intention of drink
		Age	Gender	Nationality		
Age-different	salary man A	24	male	China	anger	beer
	salary man B	33	male	China	sad	sake
Gender-different	salary man A	25	female	Japan	happy	whisky
	salary man B	25	male	Japan	disgust	sour
Nationality-different	salary man A	28	male	China	surprise	sour
	salary man B	28	male	Japan	fear	shochu

5.4.4 Experiment by Using Mascot Robot System

Using the developing mascot robot system (MRS) in our lab [88], the proposed intention understanding model based on the emotion is evaluated by performing a scenario of “drinking at a bar” with salary men (played by postgraduates) and two eye robots (i.e., bar lady robot and colleague robot), as shown in Fig. 5.16. In the scenario of “drinking at a bar”, bar lady robot needs to understanding the customers’ intentions to order based on the emotions. There groups of volunteers are invited to experience the scenario, which are age-different group (i.e., two male aged 24 and 33 both from China), gender-different group (i.e., one male and one female both aged 25 from Japan), and nationality-different group (i.e., one male from Japan and the other from China both age 28).



Figure 5.16: Scenario of “drinking at a bar”.

In the experiment, six primary emotions mentioned, i.e., happiness, surprise, fear, anger, disgust, sadness are used to represent human emotions. In the dialogue between

Table 5.12: Questionnaires for customer feedback

Questions	Answers
1. Is this (the drink given by bar lady robot) you want in your mind?	1-Yes, 2-no
2. How much is your satisfaction?	1-Very dissatisfied, 2-dissatisfied, 3-normal, 4-satisfied, 5-very satisfied

humans and robots, emotional keywords and phrases are used for expressing human emotions, such as “good job” and “well done” for happiness; “really” and “oh my god” for surprise; “worry about” and “work hard” for fear; “shut up” and “keep away from me” for anger; “bad” and “hate” for disgust; “mistake” and “stress” for sadness. In addition, emotional facial expressions by using facial movement are necessary. Moreover, several gestures are performed, such as “toast”, “nod”, “victory pose”, “farewell”, and “hand-clap” for happiness; “squatting with hands over head” for sadness; “face covering” for fear.

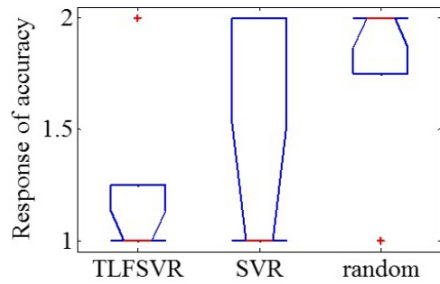
Instead of saying what would like to drink, each customer expresses the emotional behavior to prompt the bar lady robot. The emotional behavior includes the emotional speech facial expression, and gestures, where the interaction between customers and robots requires clear speech and standard facial expressions as well as gestures, or else it will be miss recognized.

To verify the proposal, three situations are applied by using MRS, including bar lady robot served drinks by the proposed TLFSVR based intention understanding model, SVR based intention understanding model, and randomly. After the experience the scenario of “drinking at a bar”, the identification information and emotion-intention information are obtain as shown in Table 5.11. At the same time, the volunteer of each groups needs to answer the questionnaire to feedback, as shown in Table 5.12. Table 5.13 shows the response of volunteers after experience the three situations.

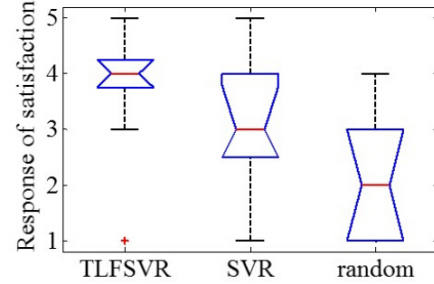
To show the statistical significations of the response, a one-way ANOVA is used to compare the means of different accuracy (intention understanding)/satisfaction in the three different situations to determine whether they are significantly different from each other. The null hypothesis is all the three situations have equal influence on intention

Table 5.13: Response of questionnaires

Group	ID	Bar lady served drinks by TLFSVR/SVR/random in MRS	
		Responses for accuracy	Responses for satisfaction
Age-different	salary man A	1/1/1	4/4/4
	salary man B	2/1/2	3/3/2
	colleague robot	1/2/2	5/1/1
Gender-different	salary man A	1/2/2	4/3/2
	salary man B	1/1/2	4/4/3
	colleague robot	2/2/2	1/1/1
Nationality-different	salary man A	1/1/2	4/3/2
	salary man B	1/1/1	4/3/3
	colleague robot	1/1/2	5/5/1



(a) Boxplots of accuracy by situations



(b) Boxplots of satisfaction by situations

Figure 5.17: Boxplots of responses by situations.

understanding accuracy and satisfaction. By using one-way ANOVA, boxplots of response by situations are shown in Fig. 5.17, respectively. According to ANOVA summary in Table 5.14, all the P-values are lower than 0.05, thus the null hypothesis is rejected, that means different situations have signification influence on intention understanding accuracy and satisfaction.

According to the result of questionnaires in Table 5.13, Table 5.15 shows the understanding accuracy and satisfaction of each group, where response of question 1 is used to calculate the understanding accuracy of the customers' intention to order, answer "Yes" for 100% and "No" for 0%. For a robot, if the accuracy is satisfaction of the bar lady is 5-very satisfied, otherwise 1-very dissatisfied. For human, the satisfaction is based on the understanding accuracy and the service attitude (subjective evaluation by human). According to Table 5.15, using the proposed TLFSVR based intention understanding model, the bar lady robot obtains an average accuracy of 77.8% which is 11% and 50% higher than that of SVR model and served randomly, respectively; the average satisfaction level

is about 4-satisfied which is one level and two levels higher than that of SVR model and served randomly, respectively.

Table 5.14: ANOVA summary of (a)/(b)

Source	SS	df	MS	F	P _{rob} > F
Columns	1.55556/12.5185	2/2	0.77778/6.25926	3.65/4.36	0.0412/0.0242
Error	5.11111/34.4444	24/24	0.21296/1.43519		
Total	6.66667/46.963	26/26			

Table 5.15: Result of application for intention understanding

Group	ID	Bar lady served drinks by TLFSVR/SVR/random in MRS	
		Accuracy (%)	Satisfaction
Age-different	salary man A	100/100/100	4/4/4
	salary man B	0/100/0	3/3/2
	colleague robot	100/0/0	5/1/1
Gender-different	salary man A	100/0/0	4/3/2
	salary man B	100/100/0	4/4/3
	colleague robot	0/0/0	1/1/1
Nationality-different	salary man A	100/100/0	4/3/2
	salary man B	100/100/100	4/3/3
	colleague robot	100/100/0	5/5/1
		77.8/66.7/22.2 (Average)	4/3/2 (Average)

To discuss whether the results of preliminary application by using MRS were consistent with the results of simulation, a comparative experiment are done. In the simulation, training data is the same as before as shown in Fig. 5.11, and testing data is obtain from MRS as shown in Table 5.11. According to the comparative result in Fig. 5.18 and data analysis in Table 5.16, the prediction accuracy is 83.3%, that means the simulation agrees well with the preliminary application, and the correlation between them are 0.9608, indicating that the simulation can well reflect the change trends of the real situation in preliminary application.

The results of the preliminary application indicate that humans can transparently communicate with robots in some degree, this is because the bar lady robot can probably understand the customer requirements, and understanding is foundation of the communication, if get to know each other deeper, communication can start from heart to heart, then a transparent communication is naturally formed. Transparent communication becomes more and more important in our daily life, especially for building trust from each other, for example, the application at a bar using the proposal, the bar lady robot try

to understand the customer, actually they are building trust mechanism during the understanding process, and the customers feedback high satisfaction is improvement for the communication. If the communication is transparent, that means communication without mental disorder in such a way that the distance between people is shorten, this is also a boost for communicating smoothly.

In the application experiment, bar lady robot gets a proper accuracy which satisfies the customer, and this benefits from the proposed intention understanding model, which takes into account people in the different ages, different genders, and different countries may has their own culture, resulted in different requirements. Reference from the preliminary application, the proposed intention understanding model can be extended to an ordering system in the bar for the business communication, where the business men are from different countries with different ages, different genders, and different culture backgrounds. Restaurant is similar to the bar, both have order system, based on the proposed intention understanding model, besides age, gender, nationality, more identification information are needed to add for the restaurant's order system, such as dietary opinions and food taboos. In addition, since people may has different intentions in different scenarios, a general intention understanding model is not easy to build. But in the proposed model, age, gender, and nationality are used as the general information, even other information can be added, which indicates the proposal could be a general model for the intention understanding, and one of the possible extension is group discussion, another important environmental information - communication atmosphere should be considerate in humans-robots interaction according to our previous research [22] in such a way that it would an interesting topic to extent the application of the proposal to more general scenario by fusing other identification or environmental information.

Table 5.16: *Result analysis of Comparative experiment*

Prediction accuracy (%)	Correlation
83.3	0.9608

Moreover, deep level understanding is being popular in human-robot interaction, and research topics would be interesting to extent the previous multi-robot behavior adapta-

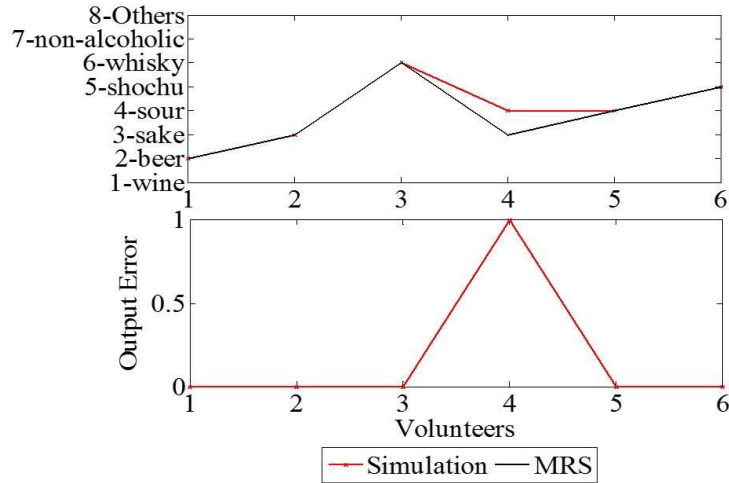


Figure 5.18: Comparative experiment between simulation and MRS.

tion mechanisms [26, 84] from emotion and atmosphere to intention. As the development of the cognitive science, in the real application of human-robot interaction, intention understanding has been successfully applied in some autonomous robots, e.g., Pioneer 2DX robot understands people’s walking direction from experience using the Hidden Markov Models [91] and the ARoS (ActivMedia Robotics Operating System) robot as sociable partners in collaborative joint activity based on cognitively understanding of people’s actions [92], showing the evidence that the proposed TLFSVR based intention understanding model may be an effective way to autonomous robots for deeply understanding peoples’ inner thoughts in human-robot interaction.

5.5 Chapter Summary

A two-layer fuzzy SVR based intention understanding model is proposed to deeply comprehending people’s inner thoughts. To validate the proposed intention understanding model, simulation experiment and application are both designed based on the background of “drinking at a bar”.

To confirm the validity of the proposed two-layer fuzzy support vector regression (TLFSVR) based intention understanding model, a comparative simulation experiment among the TLFSVR, SVR, and back propagation neural networks (BPNN) is carried out, where training/testing data are collected from questionnaire about relationship between

emotion and intention to order at a bar, as well as the identification informations (i.e., age, gender, and nationality). According to simulation results, while clusters number $C=2/3/6$, the TLFSVR obtains an accuracy of 70%/72%/80%, which is 23%/26%/33% and 35.5%/37.5%/45.5% higher than SVR and back propagation neural networks (BPNN), respectively. Computational time is about 0.976s/0.935s/0.67s by using the proposed intention understanding model, indicating that computing speed is about 1.9/2/2.8 and 3.6/3.7/5.2 times faster than the SVR and BPNN, respectively.

According to the simulation experiment results, it can be found that the accuracy of the proposed understanding model is higher than that of the SVR and the BPNN, leading to the smaller AARD and higher correlation coefficient between the output and the actual value. Smaller AARD and higher correlation indicate that the TLFSVR based intention understanding model better matches the actual situation. The clusters number C and the overlap parameter η is two important parameter in the proposed understanding model, which guarantee the accuracy. The computational time is affected by the computational complexity of the proposal, which is explicitly expressed according to the number of data and computation scale, for example, n is the number of the training data, n_i is the number of the training data in the i th training subset. C is the number of clusters, and m is the numbers of layers in BPNN, then computational complexity of the SVR, the proposed TLFSVR, and the BPNN is $O(n^2)$, $O(\sum_{i=1}^C n_i^2)$, and $O(m \times n^2)$, respectively. Since computational complexity of the proposal is simplest among these three methods, the computational time is less than that of the SVR and BPNN.

The application experiment is developed to evaluate the validity and practicability of the proposed intention understanding model in the developing humans-robots interaction system, called mascot robot system (MRS) in our lab [88], where a scenario of “drinking at a bar” is performed with six salary men (played by postgraduates) and two eye robots (the bar lady robot and colleague robot). Six salary men are separated into three groups (i.e., age-different, gender-different, and nationality-different) to experience the scenario of “drinking at a bar”, after that, they are requested to answer the questionnaire to evaluate the satisfaction of the bar lady’s service. According to the application experiment result,

the bar lady robot obtains an average accuracy of 77.8% and the average satisfaction level is about 4-satisfied using the proposed TLFSVR based intention understanding model. Bar lady robot gets an proper accuracy which satisfies the customer, and this benefits from the proposed understanding model, which takes into account people in the different ages, different genders, and different countries may lead to different requirements. The results of the preliminary application also indicate that humans can transparently communicate with robots in some degree, this is because the bar lady robot can probably understand the customer requirements, and understanding is foundation of the communication, if get to know each other deeper, communication can start from heart to heart, then a transparent communication is naturally formed. If the communication is transparent, that means communication without mental disorder so that the distance between people is shorten, this is also good for communicating smoothly. Moreover, a comparative experiment between simulation and MRS are developed to discuss whether the results of preliminary application by using MRS were consistent with the results of simulation, According to the comparative result, the prediction accuracy by simulation and the correlation between them are 83.3% and 0.9608, respectively, indicating that the simulation can well reflect the change trends of the real situation in preliminary application.

Inspired by the preliminary application, for future research, the proposed intention understanding model can be developed as an ordering system in the bar for the business communication, where business could be international. Besides emotion and atmosphere, multi-robot behavior adaptation to intention is also a potential study base on the proposed intention understanding model. In some real application of human-robot interaction, intention understanding has been successfully applied in some autonomous robots, e.g., Pioneer 2DX robot [91] and ARoS [92], from which we infer that the proposed TLFSVR based intention understanding model would be an effective way to autonomous robots for deeply understanding peoples' inner thoughts in human-robot interaction.

Chapter 6

Multi-Robot Behavior Adaptation to Humans' Intention

6.1 Introduction

Robotics technology is slowly expanded from industrial applications and academic research into highly sophisticated android robots with humanoid appearance [65]. The inevitable tendency of humanoid robots in human-robot interaction (HRI) [94] brings robots into home [95], office [61], and other areas associated with our daily life [10]. To facilitate the interaction, robots need to be able to adapt their behaviors to human partners. In the case of learning and adaptation on robots' side, information-driven theory is applied to robotics technology in the self-organization of robots' behavior [96], where there are three visible and effective surface-level information in the communication, i.e., speech, gesture, and facial expression. From the view of natural communication between humans and robots, emotion [97], atmosphere [23], and intention [93] play signification roles in HRI. These kinds of deep-level information are complex physiological phenomena and often tightly coupled to surface-level information.

Most researches focus on behavior adaptation to surface-level information, for example, autonomous robot adapts motor skills to its partner's hitting motion in playing table tennis [98]; an integrated motion synthesis framework is designed for safely adapting robot behavior to joint human's action [99]; a cognitive system is introduced to enable the iCub

robot with speech understanding capability [100]. For deep-level information, as widely used of affective computing, emotion is increasingly being applied to determine control policy of behavior adaptation [101]. Taking large scale HRI into consideration, fuzzy production rule based friend-Q learning (FPRFQ) and cooperative-neutral-competitive fuzzy-Q learning (CNCFQ) for behavior adaptation mechanism are proposed in our previous researches [26][84], and they are used to adapt multi-robot behaviors to communication atmosphere in HRI. Intention understanding is essential for communication and directly connects to human inner thoughts, and it is realized based on emotion in our previous research [31]. After intention understanding, how robots adjust their behavior to human intentions is not only critical for successful communication, but also a new and interesting challenge.

Reinforcement Learning (RL) becomes popular for solving behavior adaptation problems in HRI. A popular Q-learning algorithm uses reward and/or punishment from environment information (e.g., expected location, goods, etc.) as input of the reward function [102]. Most of these works implicitly assume that robots will finish their goal regardless of humans' feelings. We argue that RL can be reformulated to more effectively incorporate people identification information and deep-level information in the HRI. To address this, an information-driven fuzzy friend-Q learning (IDFFQ) method is introduced, where Q-values based behavior selection is replaced by mixed information (including identification information such as religion, as well as Q-values) based behavior selection. Based on the IDFFQ, a behavior adaptation mechanism is proposed to multiple robots for adapting behaviors to peoples' intention. Intention is understood by using the two-layer fuzzy support vector regression model (TLFSVR) in [31] and reward function reflects the adaptation level of robots, it is defined not only according to the expected location transformed from deep-level information (i.e., intention and emotion), but also the human satisfaction level, in such a way that, reward raises as location approaches and satisfaction increases. Using the proposed mechanism, information robots firstly collect people identification information (i.e., age, gender, nationality, and religion) and understand deep-level information (i.e., emotion and intention), then task robots dynamically adjust their behaviors

to intention by maximizing reward based on deep-level information and Q-values or only Q-values.

The proposed behavior adaptation mechanism makes benefits on both sides (i.e., robots and humans) by addressing the robots' response time and humans' satisfaction. Response time (RT) is usually time sensitive in collaborate robotic tasks, and also determines the potentials to extend the proposal to more large scale interaction. For decreasing RT in our proposal, with the religion information, robots may soon pick up non-alcoholic drink for the Muslim regardless of the intention. Moreover, human satisfaction is the key criterion of services, especially for the customers in the business communication, and also indicates the practicality of the proposal from the human side. Our proposal emphasizes on human-oriented interaction, that is to say, focusing on human side, including identification information (e.g., religion) and deep-level information (e.g., intention and emotion). Making use of these two characteristics gives a big boost to the relationship between humans and robots, so that the communication may run smoothly.

To validate the proposal, simulation experiments on multi-robot behavior adaptation to intention are performed based on a scenario of drinking at bar, where eight volunteers are invited as customers and PC is used as the robots' processor. Experiments consist of emotion-intention understanding and behavior adaptation. The training and testing data is collected from eight volunteers, including photos and questionnaires. Moreover, Emotion-intention understanding experiment is developed by C# in Visual Studio 2013 and MATLAB, and behavior adaptation experiment is built by Matlab, where the environment is a PC with a dual core processor (2.8GHz), memory (2.99GB), and Windows 7 system.

6.2 Information-Driven Fuzzy Friend-Q Learning (IDF-FQ)

Multi-agent reinforcement learning (MARL) [103] enables autonomous entities (e.g., robots) to discover optimal behaviors using trial-and-error interactions with their sharing

environment. Among MARL algorithms, Q-learning attracts a great deal of attention [104], however, the interaction environment is no longer stationary in HRI. For dynamic environment, a minimax-Q Learning algorithm is proposed to zero-sum games [72], and Nash-Q learning is the extension for general-sum games [54], by comparison to Nash-Q learning, the friend-or-foe Q-learning does not require learning estimates to the Q-functions of opponents [53]. In considering of the environment uncertainty and large state space in HRI, our previous proposed fuzzy production rule based friend-Q learning (FPRFQ) [26] generalizes the discrete friend Q-learning to the continuous state space, where the discrete lookup Q-table is replaced by “IF-THEN” fuzzy production rules. The FPRFQ mainly consists of four steps, i.e., state fuzzification, fuzzy production rule generation, behaviors selection, and Q-values updating. The updating function is defined as

$$\Delta FQ_{k+1}(s, b_i) = \alpha_k \left[r_{k+1} - FQ_k(s, b_i) + \gamma \max_{b'_i} FQ_k(s', b'_i) \right] \quad (6.1)$$

where i is the number of robot, b_i is the selective behavior of i th robot in state s , $\alpha \in [0, 1]$ is learning rate, r_{k+1} is the reward, $FQ_k(s, b_i) \in \mathbb{R}$ is the k th estimate value of the optimal Q-value, $\gamma \in [0, 1]$ is discount factor, $\max_{b'_i} FQ_k(s', b'_i)$ represents the maximum Q-value of all possible behaviors in the next state s' , b'_i is the possible behavior of i th robot in the next states', and $\Delta FQ_{k+1}(s, b_i) = FQ_k(s, b_i) - FQ_{k+1}(s, b_i)$.

The FPRFQ pays close attention to the target state, that is achieving the goal ignores some useful process states, but in fact, some process states greatly promote the communication, for instance, identification information such as age and religion have a great impact on eating habits, are nice addition to robots' knowledge. To make full use of the effective information in HRI, information-driven fuzzy friend-Q learning (IDFFQ) algorithm is introduced, where Q-values based behavior selection is replaced by mixed information (including identification information, i.e., religion, and Q-values) based behavior selection, as shown in Fig. 6.1, mainly consists of four steps.

Firstly, information transfer of the environment information I . Using this module, input values are divided into two kinds, the state information I_S and the addition information I_A , where I_S is input set of IDFFQ and I_A contains the rest of information for

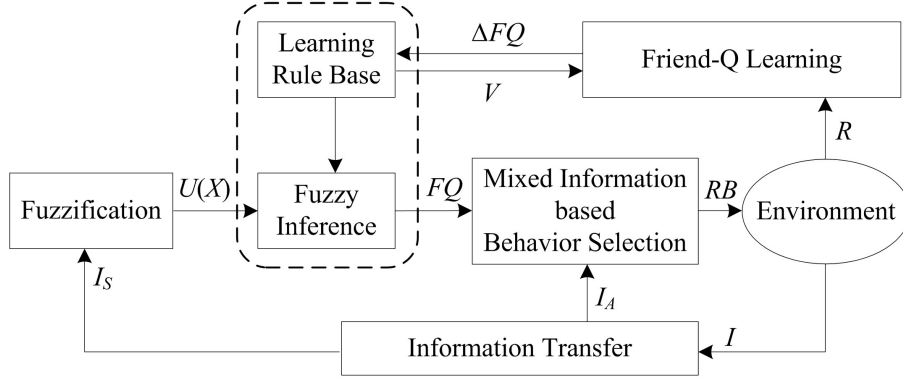


Figure 6.1: Information-driven fuzzy friend-Q learning (IDFFQ).

behavior selection.

Secondly, fuzzy rules generation consist of fuzzification and fuzzy inference. Fuzzification is the process of transforming crisp input set I_S (e.g., human intention and emotion) into linguistic terms of fuzzy sets (e.g., wine in intention, happy in emotion) by membership function. Then, the “IF-THEN” fuzzy production rules of fuzzy inference are designed by observing the relationship between the state and the behavior as

$$R_i : \text{If } I_S = S_i \text{ Then } RB_j = \{b_{kj} \text{ with } FQ_{kj}\} \quad (6.2)$$

where i is the number of learning rule, I_S is input set, S_i is fuzzy state set, RB_j is behavior set of j th robot, b_{kj} is k th possible behavior of j th robot, and FQ_{kj} is a Q-value $\in \mathbb{R}$ fixed for each b_{kj} .

Thirdly, each robot selects behavior b based on the mixed information FQ and IA . Generally speaking, the explore/exploit policies such as ϵ -greedy [73] and Boltzmann function [59] are used to selected behavior. Based on the principle that the frequency of robot behavior should be proportional to its reward, Boltzmann function maintains an explicit probability distribution over the behaviors, and searches behaviors directly in the probability distributions space of Q-values. If one of the Q-values is much higher, the

corresponding behavior is most likely to be taken. The probability is defined as

$$p(s, b_i) = \frac{e^{\frac{\max_{b_i} FQ(s, b_i)}{T}}}{\sum_{i=1}^n e^{\frac{FQ(s, b_i)}{T}}} \quad (6.3)$$

where $p(s, b_i) \in [0, 1]$ is the probability of selecting behavior b_i in state s , n is the number of possible behaviors in state s , $T \in \mathbb{R}^+$ is the temperature parameter which is decreased over time to decrease exploration.

Boltzmann exploration is widely used, and it can cover greedy case by adjusting the parameter (i.e., temperature T), when $T = 0$, Boltzmann exploration acts like pure greedy, as T tends to infinity, robots select behaviors uniformly at random. Using the Boltzmann exploration, robots select behaviors just according to possibility of the Q-values which directly relates to the states of robots. In real HRI application, for example, in HRI, not only robot states, but also human states maybe immediately affect the robot behavior, for example, a robot selects appropriate drink according to customer's religion, hence, function of Boltzmann exploration combined with human addition information IA is used to behavior selection, and (6.3) is revised as

$$p(s, b_i) = \begin{cases} \frac{e^{\frac{\max_{b_i} FQ(s, b_i)}{T}}}{\sum_{i=1}^n e^{\frac{FQ(s, b_i)}{T}}}, & I_A = \emptyset \\ f(I_A), & I_A \neq \emptyset \end{cases} \quad (6.4)$$

where $f(I_A)$ is a mapping from possibility $p(s, b_i)$ to addition human information I_A , it can be a function or some rules. Finally, based on the reward calculated by the adaptation level of robots, the Q-values are updated to the learning rule base by using updating function in (6.1).

6.3 IDFFQ based Multi-Robot Behavior Adaptation Mechanism

6.3.1 Architecture of Behavior Adaptation Mechanism

Based on the proposed IDFFQ and emotion-age-gender-nationality information based intention understanding model in [31], a behavior adaptation mechanism is proposed to multi-robot in human-robot interaction. The mechanism reflects the influence of human identification information (i.e., age, gender, nationality, and religion) and deep-level information (i.e., intention and emotion) in robots behavior adaptation. Reinforcement learning indicates the relationship of behavior execution and the HRI environment. The basic idea is multiple robots cooperate with each other to achieve humans' expective goal, and all the robots mainly separate into two kinds, including information robots and task robots. The architecture mainly contains two parts, i.e., information processing and behavior adaptation, as shown in Fig. 6.2.

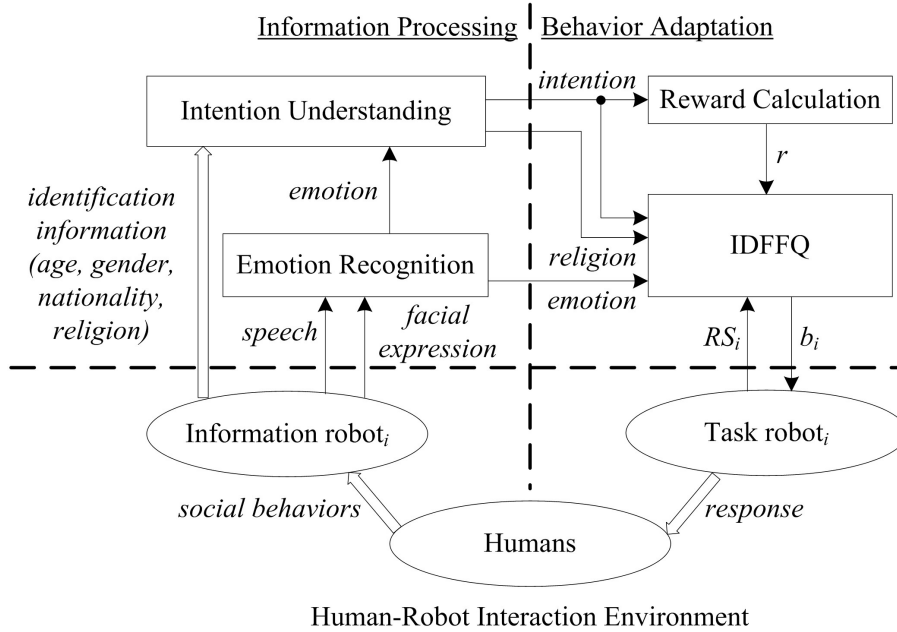


Figure 6.2: IDFFQ based behavior adaptation mechanism.

Information processing carried out by information robot is used to understand humans' social behaviors. Emotion recognition module recognizes human emotion by facial

expression and speech, then intention understanding is realized based on emotion and identification information (i.e., age, gender, and nationality).

Behavior adaptation is realized by task robots, which consists of reward calculation and information-driven fuzzy friend-Q learning (IDFFQ) based behavior adaptation. Humans' states HS and robots' state RS are used as the inputs, where HS contains the identification information such as religion, as well as deep-level information such as intention and emotion; RS describes robots' own information such as location. Reward r_i is calculated according to the adaption level of the robots. After learning process, learning behavior b_i is transferred to the behavior actuator of the task robots, then they do the corresponding behavior to response humans.

6.3.2 Information Processing of Emotion and Intention

1. Emotion Recognition

Emotion recognition is a supported module based on the previous study [24], which is realized using weighted fusion fuzzy inference based on the bimodal information, i.e., facial expressions and speech, and is specified in the Pleasure-Arousal plane, where happiness, surprise, fear, neutral, disgust, sadness, and anger are employed as the basic human emotion.

Face recognition measures the overall facial structure and focuses on capturing, extracting, and analyzing the size and shape of eyes and mouth. For better recognition, skin texture analysis is added to turns patterns, unique lines, and spots appearance in a person's skin. Since different people from different country have their own language and accent, emotion recognition is mainly realized by facial expressions (more general), speech recognition is a complement for emotion.

2. Intention Understanding

Intention understanding is an important aspect for the successful communication among people, in our daily life, we rely heavily on this skill to "read" others' minds. In HRI, communication is often not so easy because robots lack understanding of people's inner thoughts. If robots want to effectively cooperate with humans, cognitive skills must

include mechanism for inferring intention.

Intention understanding is a supported module based on previous research [31] where a two-layer fuzzy support vector regression (TLFSVR) model is proposed to comprehend people's inner thoughts in HRI. It is realized based on the emotion with additional information such as age, gender, and nationality.

6.3.3 IDFFQ based Behavior Adaptation

Reward is used to evaluate the adaptation level of the task robots, and reward function is based on the tasks completion situation of robots and human satisfaction level in one interaction.

Tasks completion situation includes three states such as success, mistake, and failure. For instance, in a scenario of drinking at a bar, a robot is supposed to be a waiter, the task is to help customer order drink, a successful completion of task means robots select the right dink, mistake means picking up other drink but not the customer needs, and failure means picking up nothing. Highest reward will give to successful completion, encourage reward gives to mistake situation, and punishment goes to failure situation, details as

$$r_t = \begin{cases} 40, & \text{success} \\ -1, & \text{mistake} \\ -40, & \text{failure} \end{cases} \quad (6.5)$$

Human satisfaction level is obtain from human side using questionnaire, it is divided into 5 levels, including "very dissatisfied", "dissatisfied", "normal", "satisfied", and "very satisfied". Different reward is given to different satisfaction level as follow

$$r_h = \begin{cases} 40, & \text{very satisfied} \\ 20, & \text{satisfied} \\ -1, & \text{normal} \\ -20, & \text{dissatisfied} \\ -40, & \text{very dissatisfied} \end{cases} \quad (6.6)$$

Based on the reward obtained from tasks completion situation and human satisfaction level, the total reward r for the task robot is defined as

$$r = \begin{cases} r_t, & \text{simulation} \\ (r_t + r_h)/2, & \text{application} \end{cases} \quad (6.7)$$

In behavior adaptation process, using the IDFFQ, task robots cooperate with each other to learn iteratively by trial and error until they receive the expected reward with the best adaptive behavior to the human intention. The IDFFQ based behavior adaptation consists of 8 steps:

Step 1: Parameters initialization.

Parameters, i.e., the learning step n ($= N$), discount factor γ ($= 0.99$), learning rate α (initial value $= 1$), and temperature parameter T ($= 1000$) are initialized.

Step 2: Information transfer.

Environment information I are separated into two kinds (i.e., input states I_S and addition information I_A for behavior selection). According to Fig. 6.2, environment information make up of two parts, one is from information processing created by information robot, including human intention HI , emotion HE , and religion HR ; the other is robot's self-state.

For example, based on the scenario of drinking at the bar, one information robot and two task robots (i.e., waiter robot and music robot) are included, the tasks of waiter robot and music robot are to offer drink service based on the intention and music service based on the emotion, respectively. For waiter robot, its self-state is location RL , then $I_S = \{HI, RL\}$ and $I_A = \{HR\}$; for music robot, its self-state is number of playing music RP , then $I_S = \{HE, RP\}$ and $I_A = \{\emptyset\}$.

Step 3: Fuzzification.

Fuzzification are transforming crisp input set I_S into linguistic terms of fuzzy sets. For example, there are eight intentions of customer at a bar, crisp values (from 1 to 8) and the corresponding fuzzy linguistic variables are “1-wine”, “2-beer”, “3-sake”, “4-sour”, “5-shochu”, “6-whisky”, “7-non-alcoholic drink”, “8-others (other drinks or food)”.

Emotion is specified in the Pleasure-Arousal plane [23], and emotion fuzzy state is represented using the triangular and trapezoidal membership function which is obtained by trial and error, as shown in Fig. 6.3, where five linguistic variables for Pleasure-Displeasure axis, i.e., Very Pleasure (VP), Pleasure (P), Neutral (NT), Displeasure (D), and Very Displeasure (VD); for “Arousal-Sleep” axis, the linguistic variables are Very Arousal (VA), Arousal (A), Neutral (NT), Sleep (S), and Very Sleep (VS).

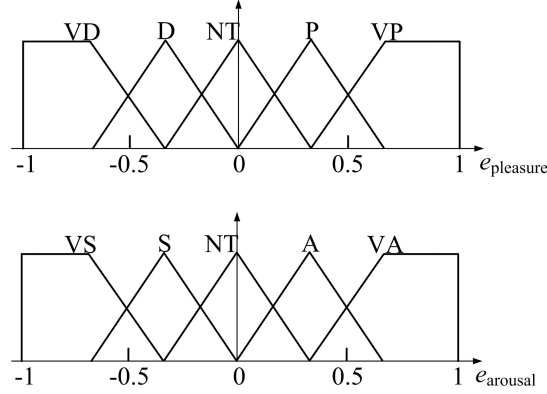


Figure 6.3: Membership function of emotion plane.

The customer has eight intentions at a bar, and the waiter robot locates according to the searching map, for instance, at a bar, a cabinet consist of several small grids and each grid with one of the customer's intention (e.g., wine), then the location can be related to the customer's intention. Hence, 64 (8×8) fuzzy rules are obtained for waiter robot and (6.2) is detail designed as

$$\text{If } I_S = \{HI \text{ and } RL\} \text{ Then } RB_j = \{b_{kj} \text{ with } FQ_{kj}\} \quad (6.8)$$

For music robot, emotion fuzzy state totally has 25 (55) kinds, but in our daily life, the most commonly used emotion are 6 basic emotion including happiness, surprise, fear, disgust, sadness, and anger, plus neutral. According to the previous study on emotion and music [22], violin music is used to convey 4 kinds of basic emotion. Hence, 28 (74) fuzzy rules are designed for music robot as

$$\text{If } I_S = \{HE \text{ and } RP\} \text{ Then } RB_j = \{b_{kj} \text{ with } FQ_{kj}\} \quad (6.9)$$

Step 5: Behavior selection.

The behavior for waiter robot is selected according to mixed information including human religion HR and Q-values. Based on the idea that Islam religion considers alcoholic consumption sinful, (6.4) is detail designed as

$$p(s, b_i) = \begin{cases} \frac{e^{\frac{\max_{b_i} FQ(s, b_i)}{T}}}{\sum_{i=1}^n e^{\frac{FQ(s, b_i)}{T}}}, & HR \neq \text{Muslim} \\ 0, & HR = \text{Muslim} \\ 1, & HR = \text{Muslim} \quad \& \quad \min_{b_i}(RL'_i - \overline{HR}) \end{cases} \quad (6.10)$$

where $\overline{HR}$ is the Muslim intention, always set as “7-non-alcoholic drink”, and $RL'_i \in 1, 2, 3, 4, 5, 6, 7, 8$ is robot's location after the behavior b_i .

For music robot, behavior is generated according to (6.4), where the behavior with highest Q-value has more possibility to be selected.

Step 6: Q-values updating.

The Q-value of the selective behavior in each robot is updated and saved to the learning rule base by using updating function (6.1) according to the obtained reward.

Step 7: If $RL = HI / \overline{HR}$ (in the case of waiter robot) or $RP = HE'$ (in the case of music robot, HE' is the selecting music according to emotion HE), then learning stops, and the best learning behavior b_i^* of each robot is outputted; else jump to *Step 8*.

Step 8: Let $n = n + 1$. If learning step $n < N$, then jump to *Step 2*; else jump to *Step 1*.

6.4 Experiments on Behavior Adaptation to Costumer's Order Intention at a Bar

6.4.1 Experiment Setting

Simulation experiments are carried out to validate the proposed behavior adaptation mechanism, including the emotion-intention understanding and behavior adaptation ex-

periment. Experiments use the scenario of “drinking at a bar” as the background, this is because the bar culture is very popular in the world. Bar is endowed with a new form called Izakaya in Japan, which is a casual place of after-work drinking and very popular for the salary man, even becomes one part of their life. If the owner of an Izakaya understands each customer preferences, it would be very nice for the tired salary man, also a good way to increase revenues. The problem of customer preferences seems to be a very interesting research, to solve this problem, emotion based intention understanding model is proposed in human-robot interaction [31], after understanding the intention, how service robots adapt to the customers' intentions at a bar is another attractive problem, so the goal of experiments is adapting service robots' behaviors to customers' order intentions at a bar based on identification information (i.e., religion) and deep-level information (i.e., intention and emotion).

Considering the complexity of behavior adaptation process to human intention, simulation experiments are first performed, which not only can run repeatedly to simplify debugging of the adaptation mechanism, but also can simulate various behaviors that are seldom experienced in real life. Moreover, in general, due to the large computing workload in emotion recognition, intention understanding, and behavior adaptation mechanism, ordinary robot cannot afford them, in this case, PC could be extended to an upper computer to complete complex intelligent algorithms, sequentially reducing robot's burden, thus, simulation on PC can be performed as similar as the application situation, which is designed to verify the accuracy of the proposal. Emotion-intention understanding experiments simulates the jobs of the information robot, includes emotion recognition developed by C# in Visual Studio 2013 and intention understanding created by MATLAB; behavior adaptation experiment simulates the jobs of the task robots, including the waiter robot for selecting drinks to human intention and music robot for playing songs to human emotion, both of them are built by MATLAB, where the environment is a PC with a dual core processor (2.8GHz), memory (2.99GB), and Windows 7 system.

6.4.2 Data Preparation

Data sample for emotion recognition are collected from photos, where eight volunteers are invited to express six basic emotions (i.e., happiness, surprise, fear, neutral, disgust, sadness, and anger) as well as neutral by facial expressions, 56 in total, as shown in Fig. 6.4.

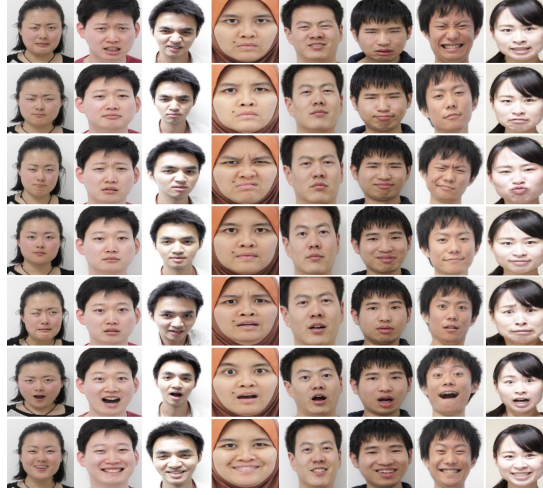


Figure 6.4: *Volunteers' emotion.*

They are from four countries (i.e., China, Japan, South Korea, and Malaysia) with different genders in Tokyo, and Malaysian is Islam religion. This data sample cover cases of different religions among different countries. Moreover, experiment uses seven basic emotion for recognition because they are dispersedly distributed in four quadrants of Pleasure-Arousal plane [23].

Data sample for intention understanding is collected from questionnaire, which is designed based on the scenario of drinking at a bar. Twenty-two volunteers are required to complete questionnaires about the relationship between 25 kinds of emotion (e.g., happiness, surprise, fear, etc.) and 8 intentions (i.e., wine, beer, sake, sour, shochu, whisky, non-alcoholic drink, and other drinks or food) to order at a bar, at the same time, write down the identification information, such as age, gender, nationality, and religion. A total of 550 groups of data are collected, where 350 groups for training and 56 groups (data corresponding to 7 basic emotions in eight volunteers) for testing, as shown in Fig. 6.5 and Fig. 6.6.

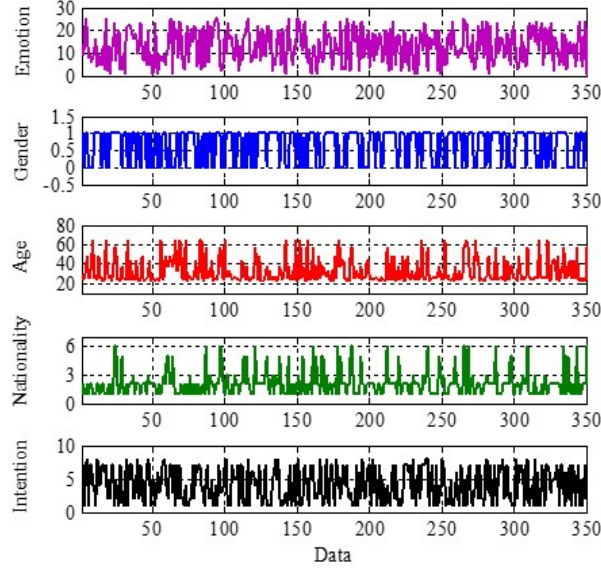


Figure 6.5: Training data for intention understanding.

Searching map for behavior adaptation is designed according to the cabinet in the scenario of drinking at a bar, as shown in Fig. 6.7.

In the map, human intentions are marked as number from 1 to 8, the corresponding meanings are “1-wine”, “2-beer”, “3-sake”, “4-sour”, “5-shochu”, “6-whisky”, “7-non-alcoholic drink”, and “8-others (other drinks or food) ”. The dashed box is the start point and the black boxes are forbidden areas.

6.4.3 Experiment on Emotion-Intention Understanding

Emotion-intention understanding experiment consist of emotion recognition and intention understanding, which are developed to support behavior adaptation experiment. Emotion recognition system mainly based on the facial expression of humans, including capturing, extracting, and analyzing the size and shape of eyes and mouth. 56 emotional photos from 8 volunteers are recognized by the system, experiment result is shown in Fig. 6.8, where the real values are compared to the recognition values.

According to data analyses, the recognition rate is 80.36%, where 45 photos are correctly recognized, 8 are wrong, and 3 are failed to recognize. Skin texture analysis is added to improve recognition by turning patterns, unique lines, and spots appearance in skin.

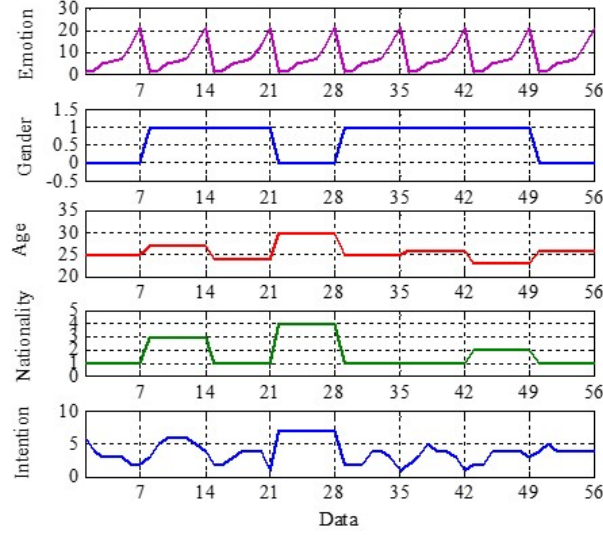


Figure 6.6: Testing data of 8 volunteers for intention understanding.



Figure 6.7: Searching map.

Intention understanding is based on the two-layer fuzzy support vector regression (TLFSVR) [31], mainly including fuzzy c-means (FCM) based data clustering and multi-SVR learning. In the experiment, inputs are emotion and identification information, i.e., age, gender, and nationality. The number of clusters for the FCM is chosen as 2 based on the gender (i.e., female and male), the overlap constant and the sensitivity threshold for the FCM are set to 3 and 1, respectively. For each SVR, the Gaussian kernel width is chosen as 0.5, and scalar regularization parameter are chosen as 200.

Fig. 6.9 show the output curves and errors of intention standing using TLFSVR. According to the data analysis of result, the TLFSVR obtains an accuracy of 85.71%, computational time is about 0.964s by using the proposed intention understanding model, correlation coefficient between the output and the real value is 0.9362, high correlation indicates that intention understanding model can well-match the actual situation.

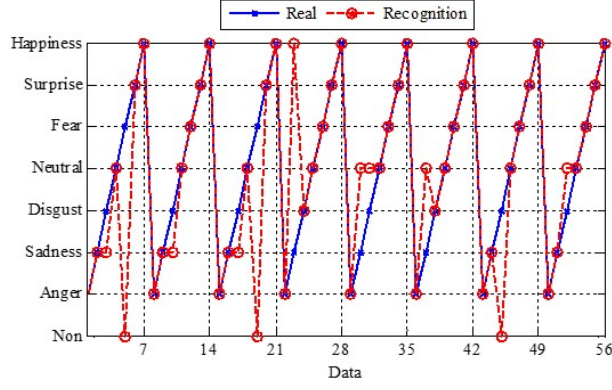


Figure 6.8: Result of emotion recognition for 8 volunteers.

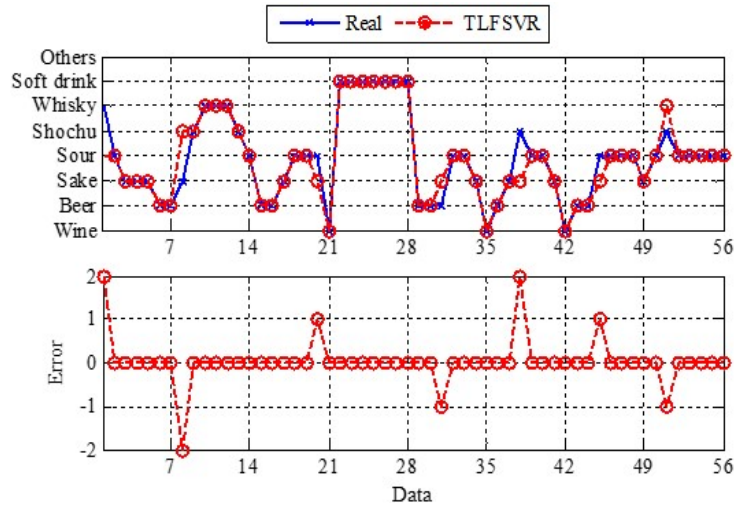


Figure 6.9: Result of intention understanding for 8 volunteers.

6.4.4 Experiment on Behavior Adaptation to Costumers' Order Intention at a Bar

To validate the performance of the proposed behavior mechanism for adapting to human intention, comparative experiments between the proposed IDFFQ with information and the FPRFQ [26] without information are carried out based on the senior of drinking at the bar.

Generally speaking, customers always drink alcohol listening to music at a bar, indicating that drink and music are two necessities of a bar. Supposed two service robots, one is waiter robot for selecting drinks to customers' intention and the other is music robot for playing music to adapt to customers' emotion. To simulate the task of waiter robot,

the customers' intentions contain 8 kinds such as "1-wine", "2-beer", "3-sake", "4-sour", "5-shochu", "6-whisky", "7-non-alcoholic drink", and "8-others (other drinks or food)", according to the searching map shown in Fig. 6.8, there are 5 kinds of behaviors in waiter robot, including "up", "down", "left", "right", and "pick up". To simulate the task of music robot, the customers' emotions consist of 6 basic emotion and neutral, according to [22], violin music are used to convey basic emotion, i.e., 1-"4th movement (Ode to Joy)" for happiness; 2-"Castle in the Sky" for sadness, 3-"Caprice No. 2" for surprise, fear, disgust, and anger ; 4-"Canon in D" for neutral. Thus, just one behavior in music robot, i.e., "playing next", and follows the sequence of "1-2-3-4-1".

The target is to improve the mechanism by simulation in the PC which can be extended as the upper computer of real robot, making human-robot interaction run smoothly. For the service robot at a bar, smooth communication is not only to achieve ordering tasks in a short response time, but also for building social relationships between humans and robots [95]. It is characterized by two indices, i.e., response time of robots (determined by computational time) and satisfaction level of customers (collected by questionnaire). Completing a task of waiter robot means robot keeps adapting until ordering is same as customer's intention, and for music robot means the playing music is the one reflects customer's emotion.

Three methods are repeated 2 times and 5000 rounds in one simulation. In the experiment, the discounted parameter γ in each algorithm is set to 0.99, the learning rate α (initial value = 1) is set to a decreasing function, i.e., $\frac{1}{(1+\alpha)^4}$. The temperature of Boltzmann function T is set to 1000 for guaranteeing an optimal policy.

According to the results of waiter robot behavior adaptation to the 8 volunteers' intention in Fig. 6.10, Fig. 6.11, and Fig. 6.12, the data analysis of results comparison between the two algorithms is shown in Table 6.1, where the average learning steps of the IDFFQ is 51 steps less than and the average reward is 3 times as big as that of FPRFQ. Robot's response time (defined by computation time) is about 1.12s using the proposed mechanism, which saves 3.09s and is 1/4 of the time consumed in FPRFQ.

The experimental results show that using the proposed behavior adaptation mecha-

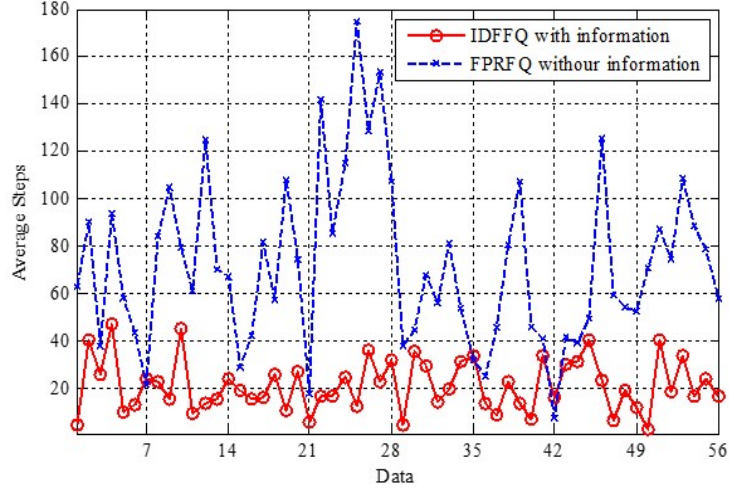


Figure 6.10: Average steps to complete a task.

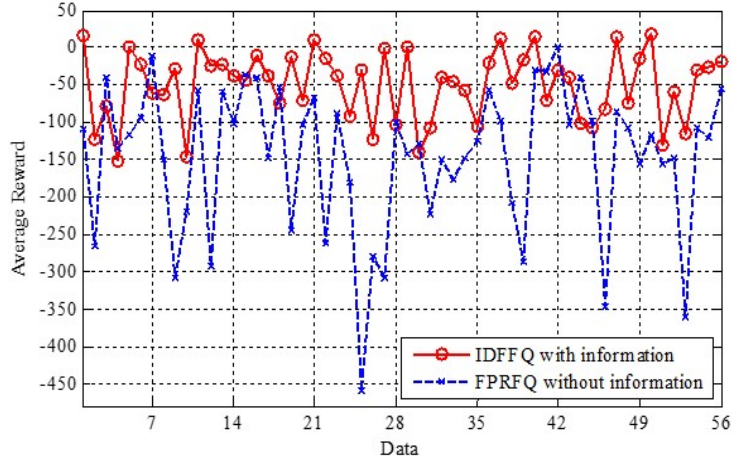


Figure 6.11: Average reward to complete a task.

nism, the robots' response time is reduced. There are mainly two reasons, one is adding the information robot to understand the deep-level information (i.e., human intention and emotion) for setting the behavior targets, that is to say, task robots get the goal directly from information robot, and no need to guest the customer's intention one by one; the other is embedding of identification information (i.e., religion) into behavior selection for the task robot, especially for Muslims, they just choose the non-alcoholic drink at a bar in any case, this useful information can quickly adjust the task goal to non-alcoholic drink in the case of Muslims.

According to the waiter robot's performance, as well as enjoying the violin music played by music robot, 8 volunteers give the satisfaction level S in five level, i.e., "1-very

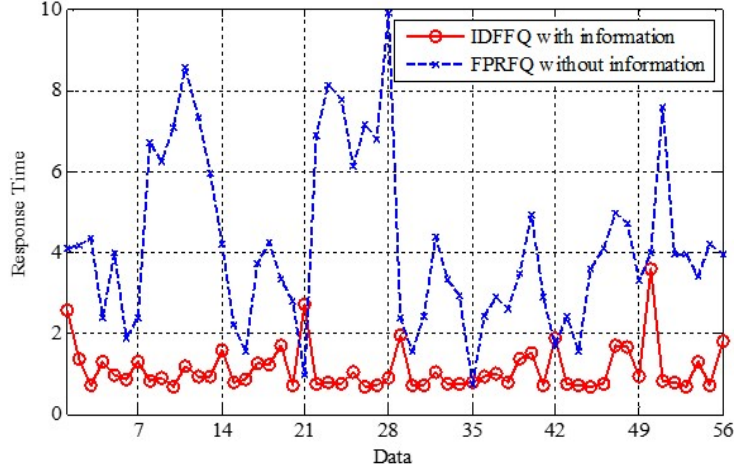


Figure 6.12: Average computational time to complete a task.

Table 6.1: Results comparison for adaptation to 8 volunteers' intention

	IDFFQ/FPRFQ		
	Mean	Min	Max
Average steps	22/73	3/8	43/175
Average reward	-49/-145	-152/-457	18/1.8
Average computational time (s)	1.12/4.21	0.67/0.76	3.60/9.91

dissatisfied”, “2-dissatisfied”, “3-normal”, “4-satisfied”, “5-very satisfied”, questionnaire is designed as shown in Table 6.2. where satisfaction level $S1$ in question 1 is according to

Table 6.2: Questionnaire for customer satisfaction level

Questions	Answers
1. Is the order (service by waiter robot) you want in your mind?	1-Yes, 2-no
2. What do you think about music	1-Very dissatisfied, 2-dissatisfied, 3-normal, 4-satisfied, 5-very satisfied

customer's answer, such as “yes” for “5-very satisfied” and “no” for “1-very dissatisfied”. For question 2, satisfaction level $S2$ is described as five answers corresponding to five satisfaction level, respectively. The total satisfaction level is calculated as $S = [(S1 + S2)/2]$.

According to the result, average satisfaction level of eight volunteers is “4-satisfied”, where six of them are “4-satisfied”, and the other two are “3-normal” and “5-very satisfied”, respectively. Some volunteers give the reason why they choose 5-very good for the music because of violin music make them feel more relaxed to enjoy the drink, and some

choose “2-bad” since the tone of violin is not their favorite musical instrument.

The preliminary application experiment is developing to investigate the practicability of the proposed intention understanding model in the developing Mascot Robot System (MRS) [21]. MRS is a typical application of humans-robots interaction, which is now developing to one day of salary man’s life with robot [93].

In MRS, one of the scenarios is drinking at a bar, three eye service robots play different roles, including information robot, waiter robot, and music robot. Robot Technology Middleware (RTM) is used to construct the network system called RTM-Network [21], as shown in Fig. 6.13. It is a two-layer network system consisted of information sensing and transfer layer as well as intelligent computation layer. RTM-network contains speech recognition module (SRM), gesture recognition module (GRM), eye-robot control module (ERCM), and intention computation module (IC). The SRM sensing the speeches information through microphone and the GRM sensing the gestures information using Kinect. Moreover, the recognition results from SRM and GRM will transfer to the upper system for the further processing. Based on the RTM-network, the proposed behavior adaptation mechanism can be comforted, for example, information robot recognizes the speech and facial expression by SRM and GRM module, at the same time, collects identification information (i.e., age, gender, nationality, and religion); then, transfers them to the IC module in the upper system (PC) for calculating the deep-level information (i.e., intention and emotion); finally, waiter robot and music robot adapt behaviors to the human intention and emotion, respectively.

In addition, as the huge information processing capability of the cloud computing, combination of the robotic technology and cloud computing may be very interesting, based on this, it is a good idea to extend the proposed information-driven behavior adaptation mechanism with cloud computing for cloud robotic [105]. For real application, some information based behavior reinforcement learning algorithms are successfully applied in some humanoid robots, e.g., Barrett WAM [106] learns to flip pancake under human guidance information and Nao [107] grasps an object with left hand inspired by the information of human cognition, indicating that the proposed IDFFQ based behavior

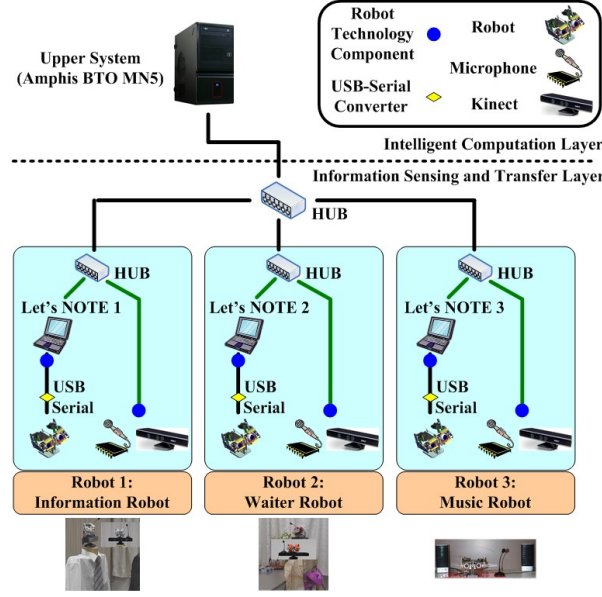


Figure 6.13: MRS network system.

adaptation mechanism may be available to humanoid robots for behavior adaptation in humans-robots interaction.

6.5 Chapter Summary

Experiments are performed base on a scenario of drinking at a bar, where eight volunteers are invited as customers and PC is used as the processor of the three robots, i.e., information robot, waiter robot, and music robot.

At first, emotion-intention understanding experiments are done by information robot, where emotion recognition by facial expressions (each volunteer expresses six basic emotion and neutral), and intention understanding (eight intentions includes wine, beer, sake, sour, shochu, whisky, non-alcoholic drink, and other drinks or food) based on emotion-age-gender-nationality are developed to support the behavior adaptation. According to experiment results, emotion recognition rate is 80.36%, where 45 photos are correctly recognized, 8 are wrong, and 3 are failed to recognize. To improve the recognition rate, skin texture analysis is added by turning patterns, unique lines, and spots appearance in skin. Moreover, Intention understanding obtains an accuracy of 85.71%, and correlation coefficient between output and real value is 0.9362, high correlation indicates that

intention understanding model can well-match actual situation.

To confirm the validity of the proposed information-driven fuzzy friend-Q learning (IDFFQ) based behavior adaptation mechanism, comparative experiments between the IDFFQ and fuzzy production rule based friend-Q learning (FPRFQ) [26] are carried out based on the emotion recognition and intention understanding, where waiter robot and music robot adapt to customers' intention (to order) and emotion (to music), respectively. The adaptation of the proposed mechanism saves 51 learning steps compared to that of the FPRFQ, and the robots' response time (defined by computation time) is about 1.12s by using the proposed mechanism, which saves 3.09s and is 1/4 of the time consumed in FPRFQ. Moreover, the reward of the proposal is about 3 times as big as that of FPRFQ. According to the waiter robot's performance, as well as music robot plays enjoyable violin music, eight volunteers give the average satisfaction level of "satisfied" to the service at a bar.

According to the experiment results, by considering humans' information, robots' response time is reduced and customers' satisfaction level is "satisfied", in such a way that, a human-oriented smooth human-robot interaction is obtained. There are two points for reducing time, one is adding information robot to understand the deep-level information (i.e., human intention and emotion) for setting the behavior targets, that is, task robots get the goal directly from information robot, and there is no need to guess the customer's intention one by one; the other is embedding of identification information (i.e., religion) into behavior selection for the task robot, especially for Muslims, they just choose the non-alcoholic drink at a bar in any case, this useful information can quickly adjust the task goal to non-alcoholic drink in the case of Islam religion. In addition, high satisfaction not only because of the waiter robot offer the correct order (at a bar) to customers in a short time, but also the enjoyment of the violin music played by the music robot.

The proposal is being extended to service robots for behavior adaptation to customers' intention to drink at a bar, where intelligent computation is completed in the upper computer system of MRS [93]. Based on the idea of cloud computing [105], the proposed mechanism is possible to be developed for cloud robotic behavior adaptation to human

intention. In real application, based on the information, some behavior reinforcement learning algorithms are successfully applied in some humanoid robots, e.g., Barrett WAM [106] and Nao [107], indicating that the proposed IDFFQ based behavior adaptation mechanism may be available to humanoid robots for behavior adaptation in human-robot interaction.

Chapter 7

Conclusion

7.1 Conclusion of Multi-Robot Behavior Adaptation to Deep Level Environmental Information

Multi-robot behavior adaptation mechanisms, i.e., fuzzy production rule based friend-Q learning, cooperative-neutral-competitive fuzzy-Q learning, and information-driven fuzzy friend-Q learning, are proposed in humans-robots interaction based on the deep level environmental information (i.e., emotion, atmosphere, and intention). In the thesis, main contributions and results are pointed out as follows:

- Acquisition of the emotion, communication atmosphere, and human intention are introduced in Chapter 2. Emotion recognition is realized using weighted fusion fuzzy inference based on the bimodal information, i.e., facial expression and speech. Emotion states are represented in the Affinity-Pleasure-Arousal (APA) space, six primary emotions, i.e., “happiness”, “surprise”, “fear”, “anger”, “disgust”, and “sadness”, together with “neutral” are employed as the initial state of emotion. Concept of the Fuzzy Atmosfield (FA) is introduced to represent the communication atmosphere based on the individual emotion, where a 3D fuzzy cubic space is introduced to represent the states of the FA, where the three coordinates are defined as “Friendly-Hostile”, “Lively-Calm”, and “Casual-Formal” axes with the range from -1 to +1, respectively. Based on the two-layer fuzzy support vector regression (TLFSVR), an

intention understanding model is introduced to deeply comprehend people's inner thoughts using individual emotion. The TLFSVR consists of two layers, including local learning layer and global learning layer. Local learning layer includes fuzzy c-means based input data classification and SVR based learning, after classification, similar emotions are clustering together, which leads to reduce the dimension calculation. According to the identification information (e.g., age, gender, and nationality), global learning layer outputs the intention of human based on fuzzy weighted average algorithm.

- Based on the fuzzy production rule based friend-Q learning (FPRFQ), a multi-robot behavior adaptation mechanism is proposed for adjusting robots' behavior to communication atmosphere in Chapter 3. To verify the proposed behavior adaptation mechanism, two experiments are developed. First experiment is robots behavior adaptation to unfamiliar communication atmosphere, robots' response time is about 0.289s using the proposed mechanism, by saving 0.073s and 0.205s than that of the FQ and the IL, respectively. Meanwhile, the distance between human-generated atmosphere and robot-generated atmosphere is 2 times and 4 times shorter than that of the FQ and the IL, respectively. Second experiment is robots behavior adaptation to unfamiliar-familiar mixed communication atmosphere, robots' response time is about 0.458s using the proposed mechanism, by saving 0.156s and 0.435s than that of the FQ and the IL, respectively. Moreover, the distance between the human-generated atmosphere and the robot-generated atmosphere is 3 times and 10 times shorter than that of the FQ and the IL, respectively. According to the results of adaptation to these two typical communication atmospheres, the robots' response time is sub-second in the simulation, even in real robots adaptation, e.g., in the MRS (five eye robots and four humans) [20], the CPU of robot is ARM STM32F103 (72 MHz) and about 30 times slower than that of the PC employed in the experiment, but the response time is still in seconds, indicating that the proposed mechanism can reduce the response time of robots in real application among at most 9 humans and robots.

- Based on cooperative-neutral-competitive fuzzy-Q learning (CNCFQ), a multi-robot behavior adaptation mechanism is proposed to coordinate local communication atmospheres in Chapter 4, where communication atmosphere is represented by a two-layer fuzzy fusion model and is visualized by shape-color-fill-wave graphics. To confirm the validity of the proposed CNCFQ based behavior adaptation mechanism, a comparative experiment among the CNCFQ, FPRFQ [26], and FQ [54] is carried out, where robots adapt to the local and global coexistent human-generated atmosphere. By using the proposed mechanism, robots' response time is about 3 times and 6 times shorter than that of the FPRFQ and the FQ, respectively. Moreover, the distance between the global human-generated atmosphere and the global robot-generated atmosphere is 3 times and 7 times shorter than that of the FPRFQ and the FQ, respectively. According to the experiment results, a smooth humans-robots interaction is obtained since the response time is reduced and the social distance between humans and robots is closer. The proposal is confirmed by experiments in the case of cooperation, neutral, and competition, which may be extended it to other interactive relationships in the communication. In addition, the 3D FA (Fuzzy Atmosfield) [23] space is divided into eight octants. In the experiment, since the adaptation of the proposal is confirmed in the first octant (all positive), the seventh octant (all negative), and the neighborhood of origin (all direction), the effectiveness of the proposed adaptation mechanism may be confirmed in the general quadrant case. Graphic visualization of the local and global atmosphere is evaluated by subjective estimation, results show that the wave obtains the highest accuracy of 91.7% since it is easy to understand the relationship between wave and the feeling of "weak" and "strong", shape obtains the accuracy of 85.5%, the color obtains 77.2% accuracy, and the fill obtains the lowest accuracy of 65.3% caused by misunderstanding of the relationship between the length and the feelings of "casual" and "formal". Moreover, comparison between the proposed graphic visualization and that in [23] is carried out, 72.5% volunteers support the proposal since it is more impressive and gives more information (e.g., intensity of the local atmosphere).

- Based on two-layer fuzzy support vector regression (TLFSVR), an intention understanding model is proposed in Chapter 5, where fuzzy c-means clustering is used to classify the input data, and intention understanding is mainly obtained by emotion, with identification information such as age, gender, and nationality. To confirm the validity of the proposed TLFSVR based intention understanding model, a comparative simulation experiment among the TLFSVR, SVR, and back propagation neural networks (BPNN) is carried out. According to simulation results, while clusters number $C=2/3/6$, the TLFSVR obtains an accuracy of 70%/72%/80%, which is 23%/26%/33% and 35.5%/37.5%/45.5% higher than SVR and back propagation neural networks (BPNN), respectively. Computational time is about 0.976s/0.935s/0.67s by using the proposed intention understanding model, indicating that computing speed is about 1.9/2/2.8 and 3.6/3.7/5.2 times faster than the SVR and BPNN, respectively. According to the results, it can be found that the accuracy of the proposed understanding model is higher than that of the SVR and the BPNN, leading to the smaller AARD and higher correlation coefficient between the output and the actual value. Smaller AARD and higher correlation indicate that the TLFSVR based intention understanding model better matches the actual situation. The application experiment is developed to evaluate the validity and practicability of the proposed intention understanding model in the developing humans-robots interaction system, called mascot robot system (MRS) in our lab [90], where a scenario of “drinking at a bar” is performed with six salary men (played by postgraduates) and two eye robots (the bar lady robot and colleague robot). According to the application experiment result, the bar lady robot obtains an average accuracy of 77.8% and the average satisfaction level is about 4-satisfied using the proposed TLFSVR based intention understanding model. Bar lady robot gets an proper accuracy which satisfies the customer, and this benefits from taking into account people in the different ages, different genders, and different countries may lead to different requirements. The results of the preliminary application also indicate that humans can transparently communicate with robots in some

degree, this is because the bar lady robot can probably understand the customer requirements, and understanding is foundation of the communication, if get to know each other deeper, communication can start from heart to heart, then a transparent communication is naturally formed. If the communication is transparent, that means communication without mental disorder so that the distance between people is shorten, this is also good for communicating smoothly.

- Based on information-driven fuzzy friend-Q learning (IDFFQ), a multi-robot behavior adaptation mechanism to human intention is proposed in Chapter 6, where intention understanding is done mainly based on human emotion. Experiments are performed base on a scenario of drinking at a bar, where eight volunteers are invited as customers and PC is used as the processor of the three robots, i.e., information robot, waiter robot, and music robot. At first, emotion-intention understanding experiments are done by information robot, where emotion recognition by facial expressions, and intention understanding based on emotion-age-gender-nationality are developed to support the behavior adaptation. According to experiment results, emotion recognition rate is 80.36%, where 45 photos are correctly recognized, 8 are wrong, and 3 are failed to recognize. To improve the recognition rate, skin texture analysis is added by turning patterns, unique lines, and spots appearance in skin. Moreover, Intention understanding obtains an accuracy of 85.71%, and correlation coefficient between output and real value is 0.9362, high correlation indicates that intention understanding model can well-match actual situation. To confirm the validity of the proposed IDFFQ based behavior adaptation mechanism, comparative experiments between the IDFFQ and FPRFQ [26] are carried out, where waiter robot and music robot adapt to customers' intention (to order) and emotion (to music), respectively. By using the proposed mechanism, robots' response time (defined by computation time) is about 1.12s, which saves 3.09s and is 1/4 of the time consumed in FPRFQ. Based on the waiter robot's performance, as well as music robot plays enjoyable violin music, eight volunteers give the average satisfaction level of "satisfied" to the service at a bar. According to the experiment results, by

considering humans' information, robots' response time is reduced and customers' satisfaction level is "satisfied", in such a way that, a human-oriented smooth human-robot interaction is obtained. There are two points for reducing time, one is adding information robot to understand the deep-level information (i.e., human intention and emotion) for setting the behavior targets, that is, task robots get the goal directly from information robot, and there is no need to guess the customer's intention one by one; the other is embedding of identification information (i.e., religion) into behavior selection for the task robot, especially for Muslims, they just choose the non-alcoholic drink at a bar in any case, this useful information can quickly adjust the task goal to non-alcoholic drink in the case of Islam religion. In addition, high satisfaction not only because of the waiter robot offer the correct order (at a bar) to customers in a short time, but also the enjoyment of the violin music played by the music robot.

7.2 Future Perspectives of Multi-Robot Behavior Adaptation Mechanisms and Potential Application

In terms of future work, the proposed multi-robot behavior adaptation mechanisms would be interesting in the following ideas, one is to enrich the adaptation mechanism for humans, besides robots' behavior adaptation; the other is music adaptation to the atmosphere, where music such as that from the violin [22] is considered as an independent emotional object as well as the humans and robots. Moreover, as cognitive science is widely used in the humans-robots interaction, deep level understanding is getting more and more popular, interesting research topics may extend the adaptation mechanism not only for emotion or atmosphere, but also other hidden deep level information, e.g., intention [61]. In addition, as the huge information processing capability of the cloud computing, combination of the robotic technology and cloud computing may be very interesting, based on this, it is a good idea to extend the proposed information-driven

behavior adaptation mechanism with cloud computing for cloud robotic [105].

In the real application of humans-robots interaction, the multi-agent learning algorithms (e.g., Minimax-Q Learning Algorithm and Nash-Q Learning Algorithm [75][76]) has been successfully applied in some humanoid robot, e.g., Robovie II [10], COMAN [77], Nico [62], and Nao [78][79], showing the evidence that the proposed behavior adaptation mechanisms may be welcomed to employ to humanoid robots' behavior adaptation in humans-robots interaction.

Moreover, deep level understanding is being popular in human-robot interaction, intention understanding has been successfully applied in some autonomous robots, e.g., Pioneer 2DX robot understands people's walking direction from experience using the Hidden Markov Models [91] and the ARoS (ActivMedia Robotics Operating System) robot as sociable partners in collaborative joint activity based on cognitively understanding of people's actions [92], showing the evidence that the proposed TLFSVR based intention understanding model may be an effective way to autonomous robots for deeply understanding peoples' inner thoughts in human-robot interaction. Reference from the preliminary application, the proposed intention understanding model can be extended to an ordering system in the bar for the business communication, where the business men are from different countries with different ages, different genders, and different culture backgrounds.

Bibliography

- [1] J. E. Young, J. Sung, A. Voids, E. Sharlin, T. Igarashi, H. I. Christensen, and R. E. Grinter. Evaluating human-robot interaction. *International Journal of Social Robotics*, 3(1):53–67, 2011.
- [2] T. Fong, I. Nourbakhsh, and K. Dautenhahn. A survey of socially interactive robots. *Robotics and Autonomous System*, 42(3):143–166, 2003.
- [3] S. Ikemoto, H. B. Amor, T. Minato, H. Ishiguro, and B. Jung. Physical interaction learning: Behavior adaptation in cooperative human-robot tasks involving physical contact. In *IEEE International Symposium on Robot and Human Interactive Communication*, pages 504–509, 2009.
- [4] M. E. Foster, S. Keizer, Z. Wang, and O. Lemon. Multi-agent reinforcement learning algorithm to handle beliefs of other agents’ policies and embedded beliefs. In *Proceedings of the fifth international joint conference on autonomous agents and multiagent systems*, pages 9–11, 2012.
- [5] C. Brezeal. Social interactions in HRI: The robot view. *IEEE Transactions on Systems, Man, and Cybernetics*, 34(2):181–186, 2004.
- [6] M. E. Sargn, Y. Yemez, E. Erzin, and A. M. Tekalp. Analysis of headgesture and prosody patterns for prosody-driven head-gesture animation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 30(8):1330–1345, 2008.

- [7] E. M. Chutorian and M. M. Trivedi. Head pose estimation in computer vision: A survey. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 31(4):607–626, 2009.
- [8] J. Chan and G. Nejat. A learning-based control architecture for an assistive robot providing social engagement during cognitively stimulating activities. In *IEEE International Conference on Robotics and Automation*, pages 3928–3933, 2011.
- [9] M. Malfaz and M. A. Salichs. Learning to deal with objects. In *IEEE International Symposium on Robot and Human Interactive Communication*, pages 504–509, 2001.
- [10] N. Mitsunaga, C. Smith, T. Kanda, H. Ishiguro, and N. Hagita. Adapting robot behavior for human-robot interaction. *IEEE Transactions on Robotics*, 24(4):911–916, 2008.
- [11] R. Cantrell, P. Schermerhorn, and M. Scheutz. Learning actions from human-robot dialogues. In *IEEE International Symposium on Robot and Human Interactive Communication*, pages 125–130, 2011.
- [12] M. A. Salichs and M. Malfaz. A new approach to modeling emotions and their use on a decision-making system for artificial agents. *IEEE Transactions on Affective Computing*, 3(1):56–68, 2012.
- [13] E. Martinson and W. Lawson. Learning speaker recognition models through human-robot interaction. In *IEEE International Conference on Robotics and Automation*, pages 3915–3920, Shanghai, China, 2011.
- [14] V. N. Vapnik. *The Nature of Statistical Learning Theory*. Springer, New York, USA, 2000.
- [15] F. Cid, J. A. Prado, P. Bustos, and P. Nunez. Learning speaker recognition models through human-robot interaction. In *IEEE/RSJ International Conference on Intelligent Robots and Systems*, pages 2188–2193, Tokyo, Japan, 2013.

-
- [16] R. W. Picard. Affective computing: Challenges. *International Journal of Human-Computer Studies*, 59(1-2):55–64, 2003.
- [17] Z. Shi, J. Wei, Z. Wang, and J. Tu. Affective transfer computing model based on attenuation emotion mechanism. *Journal on Multimodal User Interfaces*, 5(1-2):3–18, 2012.
- [18] K. W. Lee, H. R. Kim, W. C. Yoon, Y. S. Yoon, and D. S. Kwon. Designing a human-robot interaction framework for home service robot. In *IEEE international workshop on robots and human interactive communication*, pages 286–293, 2005.
- [19] S. Park, L. Moshkina, and R. C. Arkin. Mood as an affective component for robotic behavior with continuous adaptation via learning momentum. In *IEEE-RAS international conference on humanoid robots*, pages 340–345, 2010.
- [20] Y. Yamazaki, F. Y. Dong, Y. Masuda, Y. Uehara, P. Kormushev, H. A. Vu, P. Q. Le, and K. Hirota. Intent expression using eye robot formascot robot system. In *8th International Symposium on Advanced Intelligent Systems*, pages 576–580, 2007.
- [21] H. A. Vu, Y. Yamazaki, F. Y. Dong, and K. Hirota. Emotion recognition based on human gesture and speech information using RT middleware. In *IEEE international conference on fuzzy systems*, pages 787–791, Taipei, Taiwan, 2011.
- [22] Z. T. Liu, Z. Mu, L. F. Chen, P. Q. Le, C. Fatichah, Y. K. Tang, M. L. Tangel, F. Yan, K. Ohnishi, M. Yamaguchi, Y. Adachi, J. J. Lu, T. Y. Li, Y. Yamazaki, F. Y. Dong, and H. Kaoru. Emotion recognition of violin music based on strings music theory for mascot robot system. In *9th International Conference on Informatics in Control, Automation and Robotics*, pages 5–14, Rome, Italy, 2012.
- [23] Z. T. Liu, M. Wu, D. Y. Li, L. F. Chen, F. Y. Dong, Y. Yamazaki, and K. Hirota. Concept of fuzzy atmosfield for representing communication atmosphere and its application to humans-robots interaction. *Journal of Advanced Computational Intelligence and Intelligent Informatics*, 17(1):3–17, 2013.

-
- [24] Z. T. Liu, M. Wu, D. Y. Li, L. F. Chen, F. Y. Dong, Y. Yamazaki, and K. Hirota. Communication atmosphere in humans and robots interaction based on concept of fuzzy atmosfield generated by emotional states of humans and robots. *Journal of Automation, Mobile Robotics and Interlligent Systems*, 7(2):52–63, 2013.
- [25] Z. T. Liu, M. Wu, D. Y. Li, Y. Yamazaki, F. Y. Dong, and K. Hirota. Emotional states based 3-d fuzzy atmosfield for casual communicatio between humans and robots. In *IEEE International Conference on Fuzzy Systems*, pages 777–782, 2011.
- [26] L. F. Chen, Z. T. Liu, F. Y. Dong, Y. Yamazaki, M. Wu, and K. Hirota. Adapting multi-robot behavior to communication atmosphere in humans-robots interaction using fuzzy production rule based friend-q learning. *Journal of Adavanced Computational Intelligence and Intelligent Informatics*, 17(2):291–301, 2013.
- [27] L. Bottou and V. N. Vapnik. Local learning algorithms. *Neutal Computation*, 4(6):888–900, 1992.
- [28] A. Shafi and W. Vishanth. Understanding citizens’ behavioural intention in the adoption of e-government services in the state of qatar. In *European Conference on Information Systems*, pages 430–1952, Verona, Italy, 2009.
- [29] R. O. Orji. Impact of gender and nationality on acceptance of a digital library: An empirical validation of nationality based utaut using sem. *Journal of Emerging Trends in Computing and Information Sciences*, 1(2):68–79, 2010.
- [30] C. Holland and R. Hill. The effect of age, gender and driver status on pedestrians’ intentions to cross the road in risky situations intentions to cross the road in risky situations. *Accident Analysis and Prevention*, 39(2):224–237, 2007.
- [31] L. F. Chen, Z. T. Liu, M. Wu, F. Y. Dong, Y. Yamazaki, and K. Hirota. Emotion-age-gender-nationality based intention understanding in human-robot interaction using two-layer fuzzy support vector regression. *International Journal of Social Robotics*, 2015.

-
- [32] M. V. Bergh, D. Carton, R. D. Nijs, N. Mitsou, C. Landsiedel, K. Kuehnlenz, D. Wollherr, L. V. Gool, and M. Buss. Real-time 3d hand gesture interaction with a robot for understanding directions from humans. In *IEEE International Symposium on Robot and Human Interactive Communication*, pages 1945–1952, Atlanta, USA, 2011.
- [33] Y. Wang and L. Guan. Recognizing human emotional state from audiovisual signals. *IEEE Transactions on Multimedia*, 10(5):936–946, 2008.
- [34] M. L. Song, M. Y. You, N. Li, and C. Chen. A robust multimodal approach for emotion recognition. *Neurocomputing*, 71(10):1913–1920, 2008.
- [35] M. J. Han, J. H. Hsu, K. T. Song, and F. Y. Chang. A new information fusion method for bimodal robotic emotion recognition. *Journal of Computers*, 3(7):39–47, 2008.
- [36] T. M. Rutkowski and D. P. Mandic. *Artificial Intelligence for Human Computing*. Springer, Berlin, 2007.
- [37] T. M. Rutkowski, K. Kakusho, V. Kryssanov, and M. Minoh. *Knowledge-Based Intelligent Information and Engineering Systems*. Springer, Berlin, 2004.
- [38] Z. Zeng, M. Pantic, G. I. Roisman, and T. S. Huang. A survey of affect recognition methods - audio, visual, and spontaneous expressions. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 31(1):39–58, 2009.
- [39] A. Geiger, M. Lauer, and R. Urtasun. A generative model for 3d urban scene understanding from movable platforms. In *IEEE Conference on Computer Vision and Pattern Recognition*, pages 1945–1952, Colorado, USA, 2011.
- [40] M. J. Roberson, J. Bohg, G. L. Skantze, J. Gustafson, R. Carlson, B. Rasolzadeh, and D. Kragic. Enhanced visual scene understanding through human-robot dialog. In *IEEE/RSJ International Conference on Intelligent Robots and Systems*, pages 1945–1952, San Francisco, USA, 2011.

-
- [41] C. Mohiyeddini, R. Pauli, and S. Bauer. The role of emotion in bridging the intention behaviour gap: The case of sports participation. *Psychology of Sport and Exercise*, 10(2):226–234, 2009.
- [42] A. Olsson and K. N. Ochsner. The role of social cognition in emotion cognitive-emotional interactions. *Psychology of Sport and Exercise*, 12(2):65–71, 2007.
- [43] Y. Yamazaki, Y. Hatakeyama, F. Y. Dong, K. Nomoto, and K. Hirota. Fuzzy inference based mentality expression for eye robot in affinity pleasure-arousal space. *Journal of Advanced Computational Intelligence and Intelligent Informatics*, 12(3):304–313, 2008.
- [44] M. Moavenian and H. Khorrami. A qualitative comparison of artificial neural networks and support vector machines in ECG arrhythmias classification. *Expert Systems with Applications*, 37(4):3088–3093, 2010.
- [45] J. A. K. Suykens, J. Vandewalle, and B. D. Moor. Optimal control by least squares support vector machines. *Neural Networks*, 14(1):23–25, 2001.
- [46] S. Ismail, A. Shabri, and R. Samsudin. A hybrid model of self-organizing maps (SOM) and least square support vector machine (LSSVM) for time-series forecasting. *Expert Systems with Applications*, 38(8):10574–10578, 2011.
- [47] C. Huang. A reduced support vector machine approach for interval regression analysis. *Information Sciences*, 217:56–64, 2012.
- [48] J. P. Ferreira, M. M. Crisotomo, and A. P. Coimbra. Adaptive pd controller modeled via support vector regression for a biped robot. *IEEE Transactions on 941-949 Control Systems Technology*, 21(3):941–949, 2013.
- [49] J. C. Dunn. A fuzzy relative of the isodata process and its use in detecting compact well-separated clusters. *Journal of Cybernetics*, 3(3):32–57, 1973.
- [50] J. C. Bezdek. A convergence theorem for the fuzzy isodata clustering algorithms. *IEEE Transactions on Pattern Analysis Machine Intelligence*, PAMI-2(1):1–8, 1980.

-
- [51] K. C. Klauer and J. Musch. *Psychology of Evaluation: Affective Processes in Cognition and Emotion*. Lawrence Erlbaum Associates, Mahwah, USA, 2003.
- [52] G. Farchi, F. Fidanza, S. Giampaoli, S. Mariotti, and A. Menotti. Alcohol and survival in the italian rural cohorts of the seven countries study. *International Journal of Epidemiology*, 29:183–199, 2000.
- [53] C. J. H. Watkins and P. Dayan. Q-learning. *Machine Learning*, 8(3-4):279–292, 1992.
- [54] M. L. Littman. Friend-or-foe Q-learning in general-sum games. In *Proceeding of 18th Interational Conference on Machine Learning*, pages 322–328, Massachusetts, USA, 2001.
- [55] M. Tan. Muti-agent reinforcement learning: Independent vs. cooperative agents. In *Proceeding of 10th Interational Conference on Machine Learning*, pages 330–337, 1993.
- [56] M. Haring, N. Bee, and E. Andre. Creation and evaluation of emotion expression with body movement, sound and eye color for humanoid robots. In *IEEE International Symposium on Robot and Human Interactive Communication*, pages 204–209, 2011.
- [57] M. A. Salichs, R. Barber, A. M. Khamis, M. Malfaz, J. F. Gorostiza, R. Pacheco, R. Rivas, A. Corrales, E. Delgado, and D. Garca. Maggie: A robotic platform for human-robot social interaction. In *IEEE Conference on Robotics, Autonomous and Mechatronics*, pages 1–7, 2006.
- [58] M. Hashimoto, M. Yamano, and T. Usui. Effects of emotional synchronization in human-robot kansei communications. In *IEEE International Symposium on Robot and Human Interactive Communication*, pages 52–57, 2009.
- [59] S. Singh, T. Jaakkola, M. L. Littman, and C. Szepesvari. Convergence results for single-step on-policy reinforcement-learning algorithms. *Machine Learning*, 38(3):287–308, 2000.

-
- [60] N. Karakayali. Social distance and affective orientations. *Sociological Forum*, 24(3):538–562, 2009.
- [61] N. Mitssunaga, T. Miyashita, H. Ishiguro, K. Kogure, and N. Hagita. Robovie-IV: A communication robot interacting with people daily in an office. In *Proceedings of international conference on intelligent robots and system*, pages 5066–5072, Beijing, China, 2006.
- [62] E. S. Kim and B. Scassellati. Learning to refine behavior using prosodic feedback. In *Proceeding of the International Conference on Development and Learning*, pages 205–210, 2007.
- [63] R. R. Murphy, T. Nomura, A. Billard, and J. L. Burke. Human-robot interaction. *IEEE Robotics & Automation Magazine*, 17(2):85–89, 2010.
- [64] T. Makino and K. Aihara. Multi-agent reinforcement learning algorithm to handle beliefs of other agents’ policies and embedded beliefs. In *Proceedings of the fifth international joint conference on autonomous agents and multiagent systems*, pages 789–791, 2006.
- [65] S. Ikemoto, H. B. Amor, T. Minato, B. Jung, and H. Ishiguro. Physical human-robot interaction: Mutual learning and adaptation. *IEEE Robotics & Automation Magazine*, 19(4):24–35, 2012.
- [66] C. Isbell, C. R. Shelton, M. Kearns, S. Singhand, and P. Stone. A social reinforcement learning agent. In *Proceedings of the fifth international conference on autonomous agents*, pages 377–384, 2001.
- [67] J. Tao and T. Tan. *Affective computing: A review*. Spriger, Berlin, 2005.
- [68] A. D. Tews, M. J. Mataric, and G. S. Sukhatme. Scaling high level interactions between humans and robots. In *Proceedings of AAAI spring symposium on human interaction with autonomous systems in complex environments*, pages 196–202, 2003.

-
- [69] E. D. Eyal and M. Yishay. Learning rates for Q-learning. *Journal of Machine Learning Research*, 5:1–25, 2003.
- [70] J. Filar and K. Vrieze. *Competitive Markov Decision Processes*. Springer, Berlin, 1997.
- [71] M. L. Littman. Markov games as a framework for multiagent reinforcement learning. In *Proceedings of the 11th international conference on machine learning*, pages 157–163, 1994.
- [72] J. Hu and M. P. Wellman. Nash Q-learning for general-sum stochastic games. *Journal of Machine Learning Research*, 4:1039–1069, 2003.
- [73] L. P. Kaelbling, M. L. Littman, and A. W. Moore. Reinforcement learning: A survey. *Journal of Artificial Intelligence Research*, 4:237–285, 1996.
- [74] Y. K. Tang, H. A. Vu, P. Q. Le, D. Masano, O. H. Thet, C. Fatichah, Z. T. Liu, M. Yamaguchi, M. L. Tangel, F. Y. Dong, Y. Yamazaki, and K. Hirota. Multimodal gesture recognition for mascot robot system based on choquet integral using camera and 3d accelerometers fusion. *Journal of Advanced Computational Intelligence and Intelligent Informatics*, 5(5):563–572, 2011.
- [75] P. F. Dimoney and F. Warneken. The basis of shared intentions in human and robot cognition. *New Ideas in Psychology*, 29(3):260–274, 2011.
- [76] E. Yang and D. Gu. Multiagent reinforcement learning for multi-robot systems: A survey. Technical report, University of Essex, 2004.
- [77] P. Kormushev, S. Kormushev, and D. G. Caldwell. Reinforcement learning in robotics: Applications and real-world challenges. *Robotics*, 2(3):122–148, 2013.
- [78] N. N. Guerrero, C. Weber, P. Schroeter, and S. Wermter. Real-world reinforcement learning for autonomous humanoid robot docking. *Robotics and Autonomous Systems*, 60(11):1400–1407, 2012.

-
- [79] T. Hester, M. Quinlan, and P. Stone. Generalized model learning for reinforcement learning on a humanoid robot. In *IEEE international conference on robotics and automation anchorage*, pages 2369–2374, 2010.
- [80] R. Kirby, J. Forlizzi, and R. Simmons. Affective social robots. *Robotics and Autonomous Systems*, 58:322–332, 2010.
- [81] J. Fasola and M. Mataric. Using socially assistive humanrobot interaction to motivate physical exercise for older adults. *Proceedings of the IEEE*, 100(8):2512–2526, 2012.
- [82] S. Calinon, I. Sardellitti, and D. G. Caldwell. Learning-based control strategy for safe human-robot interaction exploiting task and robot redundancies. In *IEEE/RSJ International Conference on Intelligent Robots and Systems*, pages 18–22, Taipei, Taiwan, 2010.
- [83] F. Broz, I. Nourbakhsh, and R. Simmons. Planning for human-robot interaction in socially situated tasks. *International Journal of Social Robotics*, 5(2):193–241, 2013.
- [84] L. F. Chen, Z. T. Liu, F. Y. Dong, Y. Yamazaki, M. Wu, and K. Hirota. Multi-robot behavior adaptation to local and global communication atmosphere in humans-robots interaction. *Journal on Multimodal User Interfaces*, 8(3):289–303, 2014.
- [85] A. Drenner, M. Janssen, A. Kottas, A. Kossett, C. Carlson, R. Lloyd, and N. Panikolopoulos. Coordination and longevity in multi-robot teams involving miniature robots. *Journal of Intelligent Robotic Systems*, 72(2):263–284, 2013.
- [86] Keysar B. The illusory transparency of intent in: Linguistic perspective taking in text. *Cognitive Psychology*, 26(2):165–208, 1994.
- [87] M. Hanheide, M. Lohse, and H. Zender. Expectations, intentions, and actions in human-robot interaction. *International Journal of Social Robotics*, 4(2):107–108, 2012.

-
- [88] K. Hirota and F. Dong. Development of mascot robot system in nedo project. In *International IEEE Conference on Intelligent Systems*, pages 430–492, Varna, Bulgaria, 2008.
- [89] A. K. Jain and R. C. Dubes. *Algorithms for Clustering Data*. Prentice Hall, Englewood Cliffs, USA, 1988.
- [90] Y. Tarabalka, M. Fauvel, J. Chanussot, and J. A. Benediktsson. SVM- and MRF-based method for accurate classification of hyperspectral images. *IEEE Geoscience and Remote Sensing Letters*, 7(4):736–740, 2010.
- [91] R. Kelley, A. Tavakkoli, C. King, M. Nicolescu, M. Nicolescu, and G. Bebis. Understanding human intentions via hidden markov models in autonomous mobile robots. In *Proceedings of the 3rd ACM/IEEE International Conference on Human Robot Interaction*, pages 5–14, Amsterdam, Netherlands, 2008.
- [92] E. Bicho, W. Erhlagen, and E. Sousa. The power of prediction: Robots that read intentions. In *IEEE/RSJ International Conference on Intelligent Robots and Systems*, pages 5–14, Amsterdam, Netherlands, 2012.
- [93] Y. K. Tang, F. Y. Dong, Yuhki M., Y. Yamazaki, T. Shibata, and K. Hirota. Deep level situation understanding and its application to casual communication between robots and humans. In *International Conference on Informatics in Control, Automation and Robotics*, pages 5–14, Reykjavik, Iceland, 2013.
- [94] M. A. Goodrich and A. C. Schultz. Human-robot interaction: A survey. *Foundations and Trends in Human-Computer Interaction*, 1(3):203–275, 2007.
- [95] K. G. Kim, D. Choi, J. Y. Lee, J. M. Park, and B. J. You. Controlling a humanoid robot in home environment with a cognitive architectures. In *IEEE International Conference on Robotics and Biomimetics*, pages 1754–1759, Phuket, Thailand, 2011.
- [96] N. Ay, H. Bernigau, R. Der, and M. Prokopenko. Information-driven self-organization: The dynamical system approach to autonomous robot behavior. *IEEE Transactions on Robotics*, 131(3):161–179, 2012.

-
- [97] Y. Tian, T. Kanade, and J. F. Cohn. *Handbook of Face Recognition*. Springer, London, England, 2011.
- [98] K. Muelling, J. Kober, O. Kroemer, and J. Peters. Learning to select and generalize striking movements in robot table tennis. *The International Journal of Robotics Research*, 32(3):263–279, 2013.
- [99] E. A. Sisbot, L. F. M. Urias, X. Broqure, D. Sidobre, and R. Alami. Synthesizing robot motions adapted to human presence. *International Journal of Social Robotics*, 2(3):329–343, 2010.
- [100] V. Tikhonoff, A. Cangelosi, and G. Metta. Integration of speech and action in humanoid robots: icub simulation experiments. *IEEE Transactions on Autonomous Mental Development*, 3(1):17–29, 2011.
- [101] M. S. Erden. Emotional postures for the humanoid-robot nao. *International Journal of Social Robotics*, 5(4):441–456, 2013.
- [102] J. Kober and J. Peters. *Reinforcement Learning*. Springer, Berlin, Gemmer, 2012.
- [103] L. Busoniu, R. Babuska, and B. D. Schutter. *Innovations in Multi-Agent Systems and Applications-1*. Springer, Berlin, Gemmer, 2010.
- [104] M. L. Littman. Markov games as a framework for multi-agent reinforcement learning. In *Proceedings of the 11th international conference on machine learning*, pages 157–163, New Brunswick, USA, 1994.
- [105] G. Hu, W. P. Tay, and Y. Wen. Cloud robotics: Architecture, challenges and applications. *IEEE Network*, 26(3):21–28, 2012.
- [106] S. Calinon, P. Kormushev, and D. G. Caldwell. Compliant skills acquisition and multi-optima policy search with em-based reinforcement learning. *Robotics and Autonomous Systems*, 61(4):369–379, 2013.
- [107] M. Richter, Y. Sandamirskaya, and G. Schoner. A robotic architecture for action selection and behavioral organization inspired by human cognition. In *IEEE/RSJ*

International Conference on Intelligent Robots and Systems, pages 2457–2464, Algarve, Portugal, 2012.

Related Publications

• Journal Papers

- J1 Lue-Feng Chen**, Zhen-Tao Liu, Fang-Yan Dong, Yoichi Yamazaki, Min Wu, and Kaoru Hirota. Adapting Multi-Robot Behavior to Communication Atmosphere in Humans-Robots Interaction Using Fuzzy Production Rule based Friend-Q Learning. *Journal of Advanced Computational Intelligence and Intelligent Informatics*, 17(2):291-301, 2013.
- J2 Lue-Feng Chen**, Zhen-Tao Liu, Min Wu, Fang-Yan Dong, Yoichi Yamazaki, and Kaoru Hirota. Multi-robot Behavior Adaptation to Local and Global Communication Atmosphere in Humans-Robots Interaction. *Journal on Multimodal User Interfaces*, 8(3):289-303, 2014.
- J3 Lue-Feng Chen**, Zhen-Tao Liu, Min Wu, Min Ding, Fang-Yan Dong, and Kaoru Hirota. Emotion-Age-Gender-Nationality based Intention Understanding in Human-Robot Interaction Using Two-Layer Fuzzy Support Vector Regression. *International Journal of Social Robotics*, 2015.
- J4 Lue-Feng Chen**, Zhen-Tao Liu, Min Wu, Fang-Yan Dong, and Kaoru Hirota. Multi-Robot Behavior Adaptation to Humans' Intention in Human-Robot Interaction Using Information-Driven Fuzzy Friend-Q Learning. *Journal of Advanced Computational Intelligence and Intelligent Informatics*, 2015.
- J5** Zhen-Tao Liu, Min Wu, Dan-Yun Li, **Lue-Feng Chen**, Fang-Yan Dong, and Kaoru Hirota. Concept of Fuzzy Atmosfield for Representing Communication

Atmosphere and Its Application to Humans-Robots Interaction. *Journal of Advanced Computational Intelligence and Intelligent Informatics*, 17(1):3-17, 2013. **Young Researcher Award**

J6 Zhen-Tao Liu, Min Wu, Dan-Yun Li, **Lue-Feng Chen**, Fang-Yan Dong, and Kaoru Hirota. Communication Atmosphere in Humans and Robots Interaction based on the Concept of Fuzzy Atmosfield Generated by Emotional States of Humans and Robots. *Journal of Automation, Mobile Robotics and Intelligent Systems*, 7(2):52-63, 2013.

• International Conference Papers

C1 **Lue-Feng Chen**, Zhen-Tao Liu, Fang-Yan Dong, Yoichi Yamazaki, Min Wu, and Kaoru Hirota. Multi-Robot Behavior Adaptation to Communication Atmosphere in Humans-Robots Interaction Using Fuzzy Production Rule based Friend-Q Learning. In *International Symposium on Soft Computing Sponsored by ASPIRE League*, Tokyo, Japan, 2012. **Best Paper Award**

C2 **Lue-Feng Chen**, Zhen-Tao Liu, Min Wu, Fang-Yan Dong, and Kaoru Hirota. Information-Driven Multi-Robot Behavior Adaptation to Intention in Human-Robot Interaction. In *Proceedings of the Joint International Conference of ITCA2014&ISCIIA2014*, Changsha, China, 2014.

C3 Zhen-Tao Liu, Zhen Mu, **Lue-Feng Chen**, Quang Phuc Le, Fatichah Chastine, Yong-Kang Tang, Martin Leonard Tangel, Fei Yan, Kazuhiro Ohnishi, Masashi Yamaguchi, Yojiro Adachi, Jia-Jun Lu, Tian-Yu Li, Yoichi Yamazaki, Fang-Yan Dong, and Kaoru Hirota. Emotion Recognition of Violin Music based on Strings Music Theory for Mascot Robot System. In *9th International Conference on Informatics in Control, Automation and Robotics*, Rome, Italy, 2012.