博士論文

Research on
Affordance-Focused Learning and Generalization
through
Observation of Proper Handovers and Object Usages
in
Robot-Human Interactions

（ロボットと人のインタラクションにおける
適切な手渡し動作と物体操作の観察に基づく
アフォーダンスに着目した
学習と一般化に関する研究）

Wesley Patrick Chan
チャン　ウェスリー　パトリック

Abstract

Object handover is a common task arising frequently in many cooperative scenarios. Therefore, it is crucial that robots perform handovers well when working with people. However, determining the proper handover method for an object is a difficult problem since it varies depending on each object's affordances. Towards enabling effective human-robot cooperation, this thesis contributes a framework that enables robots to automatically determine handover methods for various objects by observing human handovers and object usages.

This thesis first documents a user study conducted to characterize and compare the handover orientations used by humans in different conditions. It puts forth the novel idea of object *affordance axes* for identifying patterns in handover orientations, and a distance minimizing method for computing mean handover orientation from a set of observations.

Next, this thesis presents an object grouping and classification method based on observed object usage for generalizing learned handover methods to new objects. Until now, a demonstrated method for generalizing handover methods to new object has been lacking. The presented method focuses on a set of action features extracted from the movement patterns and inter-object interactions observed during usage. An experiment demonstrates the effectiveness of the method on grouping objects and then classifying new objects and computing proper handover methods for them.

The described framework for learning and generalizing handover methods is implemented onto a Kawada Industries HRP2V robot, and this thesis also documents the verification experiments. The implementation in this thesis overcomes the robot perception challenge of identifying a held object's pose at handover by detecting the object at the pre-occluded state and tracking its pose using a sequential Monte Carlo method. Results show that the framework allows robots to learn handover methods from demonstrations and compute proper handover methods for new objects. This is the first demonstrated system capable of automatically learning and generalizing handover methods from observations. Finally, integration into a household service robot application shows how this work this can enhance the capabilities of robots working in the real world by enabling them to work effectively with humans.

Through enabling better human-robot object handovers, this thesis contributes towards improving the interaction between humans and robots, thus, allowing safer, more natural, and more efficient human-robot cooperation.

Preface

The research documented in this thesis was conducted under the supervision of Professor Masayuki Inaba. The user study presented in Chapter 4 was carried out in collaboration with Professor Elizabeth Croft and Mr. Mathew Pan, and the study has been approved by the University of British Columbia Behavioural Research Ethics Board (H10-00503).

The work presented in this thesis has been published/is to be published in the following.

Wesley P. Chan, Masayuki Inaba. *Status-Based Behavior Selection for Creating Culturally Aware Robots*. To appear in Proceedings of the 3rd International Workshop on Culture Aware Robotics (CARs 2015). 2015.

Wesley P. Chan, Kotaro Nagahama, Hiroaki Yaguchi, Yohei Kakiuchi, Kei Okada, Masayuki Inaba. *Implementation of a Framework for Learning Handover Grasp Configurations through Observation during Human-Robot Object Handovers*. In Proceedings of the 15th IEEE RAS International Conference on Humanoid Robots (Humanoids 2015). 2015.

Wesley P. Chan, Matthew K.X.J. Pan, Elizabeth A. Croft, and Masayuki Inaba. *Characterization of Handover Orientations used by Humans for Efficient Robot to Human Handovers*. In Proceedings of the 2015 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS 2015). pp. 1-6. 2015.

Wesley P. Chan, Yohei Kakiuchi, Kei Okada, Masayuki Inaba. *Determining Proper Grasp Configurations for Handovers through Observation of Object Movement Patterns and Inter-object Interactions during Usage*. To appear in Proceedings of the 2014 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS 2014). pp. 1355 – 1360. 2014.

Wesley P. Chan, Iori Kumagai, Shunichi Nozawa, Yohei Kakiuchi, Kei Okada, Masayuki Inaba. *Creating Socially Acceptable Robots: Learning Grasp Configurations for Object Handovers from Demonstrations*. In Proceedings of the Workshop on Advanced Robotics and its Social Impacts (ARSO 2013). pp. 94-99. 2013.

Table of Contents

Abstract ii

Preface iv

Table of Contents v

List of Tables ix

List of Figures x

Acknowledgements xiii

1 Introduction 1

1.1 Human-Robot Interaction 2

1.2 Object Handovers 3

1.3 Research Goals 4

1.4 Object Affordance 6

1.5 Contributions 7

1.6 Thesis Outline 8

2 Motivation and Related Works 10

2.1 Implications 10

2.1.1 Human Expectations 10

2.1.2 Efficiency 11

2.1.3 Safety 11

2.1.4 Social Conventions 12

2.2 Handover Components 12

2.2.1 Gaze 12

2.2.2 Body Approaching Motion 14

2.2.3 Arm Reaching Motion 14

2.2.4 Handover Cues 16

2.2.5 Grasp Configuration 17

2.2.6 Force Control 18

2.3 Challenges in Determining Proper Grasp Configurations 19

2.4 Existing Methods for Computing Handover Grasp Configurations 21

2.5 Summary 22

3 A Framework for Learning Object Handovers from Observations and Interactions 23
3.1 Learning from Handover Demonstrations 24
3.2 Object Grouping and Classification 24
3.3 Computing Proper Grasp Configurations from Observations 25
3.4 Discussion 25
3.5 Summary 27
4 Characterization of Handover Orientations used by Humans and Method for Computing Appropriate Handover Orientations from Observations 29
4.1 Objectives 29
4.2 Object Handover User Study 30
4.2.1 Experiment Design 30
4.2.2 Hypotheses 31
4.2.3 Motion Capture 32
4.2.4 Participant Population 32
4.2.5 Experimental Procedure 32
4.3 Data Analysis 34
4.3.1 Handover Orientation Extraction 34
4.3.2 Affordance Axes 36
4.3.3 Computation of Handover Orientation Means 37
4.3.4 Patterns in Handover Orientations 38
4.3.5 Comparison of Handover Orientations across Conditions 39
4.3.6 Variation in Measured Handover Orientations 39
4.4 Results 40
4.4.1 Handover Orientation and Affordance Axis 40
4.4.2 Patterns in Handover Orientations 41
4.4.3 Comparison of Handover Orientations across Conditions 42
4.4.4 Variations in Measured Handover Orientations 43
4.5 Discussion 43
4.5.1 Patterns in Handover Orientation 43
4.5.2 Comparison of Handover Orientations across Conditions 44
4.5.3 Variation in Measured Handover Orientations 46
4.5.4 Affordance Axis 46
4.5.5 Implications towards Building Intelligent Robots 46
4.6 Summary 47

5 An Object Grouping and Classification Method based on Observation of Object Movement Patterns and Inter-object Interactions during Usage for Generalization of Handover Grasp Configurations 48
5.1 A Usage-based Method for Grouping and Classifying Objects 48
5.1.1 Extracting Inter-object Interactions from Object Usage Demonstrations 49
5.1.2 Computing Action Features for Describing Objects base on Usage 51
5.1.3 Building a Knowledge Base of Handover Grasp Configurations 53
5.1.4 Classifying New Objects and Determining Handover Grasp Configurations 53
5.2 An Experiment on Generalizing Grasp Configurations to New Objects.................................. 54
5.2.1 Experimental Procedure 54
5.2.2 Training Set 54
5.2.3 Evaluation Set 56
5.3 Results 56
5.3.1 Knowledge Base Built from Training Set 56
5.3.2 Grasp Configurations Generated for New Objects 57
5.4 Summary 57
6 An Implementation of the Framework for Learning Object Handovers from Observations onto a Robot and its Validation 59
6.1 Observing Natural Handovers 59
6.2 Extracting Grasp Configurations from Natural Human-Robot Handovers 60
6.2.1 Object Detection 61
6.2.2 Object Tracking 62
6.2.3 Human Tracking and Grasp Detection 63
6.2.4 Handover Cue Detection 63
6.2.5 Handover Orientation Extraction 64
6.3 Experiment on Learning Handover Grasp Configurations 65
6.3.1 Hardware Platform 66
6.3.2 Procedure 67
6.3.3 Results 68
6.4 Experiment on Handover Execution 71
6.4.1 Determining Appropriate Handover Grasp Configuration from a set of Observations 72
6.4.2 Procedure 73
6.4.3 Handover Results of Known Objects 74
6.4.4 Handover Results of Unknown Objects 79
6.5 Consideration for Additional Non-Object-Related Factors as an Extension 81

6.5.1 Influence of User Identity on Handover Grasp Configuration 81
6.5.2 Grasp Configuration Selection Considering Additional Factors 82
6.5.3 Implementation 83
6.5.4 Experiment Procedure 83
6.5.5 Results 84
6.6 Integration into a Household Service Robot Application 86
6.7 Discussion 92
6.8 Summary 95
7 Conclusion 97
7.1 Framework for Learning Handover Grasp Configurations 97
7.2 Handover Grasp Configurations used by Humans 98
7.3 Affordance Based Object Grouping and Classification 98
7.4 Learning from Human-Robot Interactions 98
7.5 Contributions 99
7.6 Future Work 99
Bibliography 101
Appendix A Object Handover User Study Consent Form 113
Appendix B Computed Mean Handover Orientations 114
Appendix C Extracted Handover Grasp Configuration Data 115

List of Tables

Table 4-1 Representations of 3D rotational group SO(3). 37
Table 4-2 Comparison results of handover orientations among conditions. Significant results from t-test are highlighted. (Published in Chan et al. 2015 [107].) 42
Table 4-3 Computed averages, δ, and standard deviations, σ, in the spread of affordance axes in each condition. 43
Table 5-1 Description of grasp configurations specified in the experiment. (Published in Chan et al. 2014 [110].) 55
Table 5-2 Classification results and generated grasp configurations. Minimum distances are shown in bold font. (Published in Chan et al. 2014 [110].) 57
Table 6-1 Extracted grasp configurations from handover demonstrations, and the computed averages. (Published in Chan et al. 2015 [131].) 69
Table 6-2 Comparison of learned handover orientations, measured orientations, and computed errors in the quantitative evaluation of the first set of experiments. (Published in Chan et al. 2015 [131].) 75
Table 6-3 Computed spread of the affordance axes in the set of observations for each of the fifteen objects in the second set of experiments. 75
Table 6-4 Comparison of learned handover orientations, measured orientations, and computed errors in the quantitative evaluation of the second set of experiments. 79
Table 6-5 Quantitative evaluation results for handover execution of unknown objects, showing learned handover orientations, measured orientations, and computed errors. 80
Table 6-6 Handover configurations extracted from demonstrations and computed averages. (To be published in Chan et al. 2015 [46].) 85
Table 6-7 Comparison of proposed framework with existing works. 95

Table B-1 Computed mean handover orientations for all conditions from the handover user study. .114
Table C-1 Handover grasp configurations extracted from demonstrations and computed averages for the fifteen objects used in the second set of experiments. $C_{handover}$ gives the handover orientation, P_{grasp} gives the grasp point, and P_{grasp2} gives the placement of the giver's second hand at handover. 115

List of Figures

Figure 1-1 Comparison of handover grasp configurations for hammer and kitchen knife. (Published in Chan et al. 2013 [44].) 5
Figure 2-1 Comparison of handover grasp configurations for a can of coffee. (Published in Chan et al. 2013 [38].) 20
Figure 3-1 Proposed framework for learning handover grasp configurations from observations. (Modified from Chan et al. 2014 [110].) 24
Figure 3-2 Robot software design layers. 27
Figure 4-1 Everyday objects and common tools participants handed over in the user study, with infrared reflectors affixed. (Published in Chan et al. 2015 [113].) 30
Figure 4-2 Experimental setup of the user study. (Figure partially adopted from Chan et al. [113].) 33
Figure 4-3 A – Typical hand trajectories of giver and receiver. The instant of object transfer is found by locating the minimum distance between giver's and receiver's hands. B – Example object transfer. (Published in Chan et al. 2015 [113].) 34
Figure 4-4 Arbitrary coordinate frames assigned to the objects used in the user study. (Published in Chan et al. 2015 [113].) 35
Figure 4-5 Base frame defined by the giver's and receiver's locations. (Published in Chan et al. 2015 [113].) 35
Figure 4-6 Four repetitions of Monte Carlo simulation results showing histograms of θ_i. Twenty random handover orientations were generated for each simulation repetition. A spread of θ_i among all angles can be seen. (Published in Chan et al. 2015 [113].) 39
Figure 4-7 A – Teapot handover orientation, Condition C, one trial. Red, green, blue lines show x, y ,z axis respectively. B – Teapot handover orientations, Condition C, all trials. The thick red, green, and blue lines show computed mean $\bar{R}$, long thin line shows computed affordance axis ϕ_{Aff} in mean handover orientation frame $\bar{R}$. (Published in Chan et al. 2015 [113].) 40
Figure 4-8 Handover orientations for all objects, all trials. Bold coordinate frames show mean orientations. Long thin lines show computed affordance axes. (Published in Chan et al. 2015 [113].) 41
Figure 4-9 Histograms of the angles between $\bar{R}\hat{\phi}_{Aff}$ and $R_i\hat{\phi}_{Aff}$ (i.e., θ_i) for each object in Condition C. (Published in Chan et al. 2015 [113].) 42
Figure 5-1 Procedure for extracting object features from usage demonstrations. (Published in Chan et al. 2014 [110].) 50
Figure 5-2 Clustering results from k-means according to observed object usage. (Published in Chan et al. 2014 [110].) 55
Figure 5-3 Four discrete grasp configurations defined for handing over different types of objects. (Published in Chan et al. 2014 [110].) 55
Figure 5-4 Three new objects used as the test set for evaluating the proposed method's ability to classify new objects. (Published in Chan et al. 2014 [110].) 56
Figure 6-1 Flow diagram of system for extracting handover grasp configurations from handovers with users. 61
Figure 6-2 Three daily objects, a spray can, a mug, and a detergent bottle, used in first set of experiments testing the model based object detection method. (Published in Chan et al. 2015 [131].) 65

Figure 6-3 Fifteen daily objects used in the second set of experiments testing the model free object detection method. 66
Figure 6-4 Kawada Industries HRP2V Robot with ASUS Xtion Pro Live camera mounted on the head. (Published in Chan et al. 2015 [131].) 67
Figure 6-5 Stages of extracting grasp configurations from handover demonstrations. A – Object detection. B – Human detection and tracking. C – Grasp point detection. D – Direction of torso-to-hand vector for handover cue detection. E – Object orientation at handover. (Published in Chan et al. 2015 [131].) 68
Figure 6-6 A – Handover demonstrations of a spray can, mug, and detergent bottle. The spray can is handed over with the label facing the receiver, the mug with the handle towards the receiver, and the detergent bottle with the nozzle pointing away from the receiver. B – HRP2V's camera image at the moment when handover cues are detected, showing the extracted handover object orientations. (Published in Chan et al. 2015 [131].) 69
Figure 6-7 Handover orientations extract for the fifteen objects in the second set of experiments. Thin lines show extracted orientations from each trial, thick lines show computed averages. 70
Figure 6-8 Handover grasp points extract for the fifteen objects in the second set of experiments. Blue circles indicate grasp point, and red circles indicate position of the second hand where a two-handed handover is detected. Hollow circles show the data from each trial and solid circles show the computed average. 71
Figure 6-9 Handover execution in the first set of experiments using learned grasp configurations. The robot successfully handed over all three objects using grasp configurations that match the ones demonstrated to it by the person previously. (Published in Chan et al. 2015 [131].) 74
Figure 6-10 Handover execution using learned grasp configurations during qualitative evaluation in the second set of experiments. Figure shows the grasp configurations used for the first ten objects of the fifteen objects used in the second set of experiments. 77
Figure 6-11 Handover execution using learned grasp configurations during qualitative evaluation in the second set of experiments. Figure shows the grasp configurations used for the last five objects of the fifteen objects used in the second set of experiments. 78
Figure 6-12 Handover execution of new objects, showing grasp configuration used for the marker pen, screwdriver, and cutter from top to bottom. 80
Figure 6-13 Extended framework for considering additional non-object-related factors when selecting appropriate handover grasp configuration. 82
Figure 6-14 Grasp configurations demonstrated to the robot in each of the three cases. (To be published in Chan et al. 2015 [46].) 84
Figure 6-15 Handover grasp configurations extracted from all five demonstrations for each case. A shows object orientation, with thick long lines showing computed average. B shows Blue circles indicating grasp point, and red circles indicating position of the second hand where a two-handed handover is detected. Hollow circles show the data from each trial and solid circles show the computed averages. (To be published in Chan et al. 2015 [46].) 85
Figure 6-16 Handover execution. In each case, HRP2V handed over the mug using an appropirate grasp configuration matching the demonstrated one. 86
Figure 6-17 A household robot performs cleaning tasks. The users sends the command through a tablet device (A). The robot picks up an empty tray from the table (B), and brings the tray to the kitchen counter (C). Robot puts clothes into the washing machine (D), pick up the broom (E), sweeps underneath the table (F), and places broom back (G). 87
Figure 6-18 Improper handover of a mug. A – Robot inserts finger into mug. B – Robot hands over mug sideways. 88

Figure 6-19 Improper handover of a remote. A – robot hands over remote upside down with buttons facing away from user. B – After taking the remote, the user needs to first turn the remote upright. C –User then needs the turn the remote around to see the buttons. D – User can finally use the remote. 88
Figure 6-20 Knowledge transfer between robots. The handover grasp configurations learned by HRP2V are stored into a knowledge base. This knowledge is then transferred to PR2. PR2 then utilizes this transferred knowledge to execute handovers properly. 89
Figure 6-21 integration experiment. A – PR2 in room with person. B – Pick up mug. C – Hand over mug. D – Pick up tray. E – Place tray at counter. F – Pick up remote. G – Hand over remote. H – Place clothes in washing Machine. I – Move to kitchen. J – Hand over kitchen knife. 91
Figure 6-22 Examples demonstrating some failures of the object tracker. (Published in Chan et al. 2015 [131].) 94

Acknowledgements

I would like to thank my supervisor Dr. Masayuki Inaba, who has allowed me to conduct my research at the JSK Lab with much freedom. I would like to thank him for his advice on research and the support he has provided me over the past three years of my doctoral studies.

Dr. Kei Okada has also offered advice and provided support to different aspects of my studies, and I would also like to thank him.

Dr. Kotaro Nagahama, Dr. Shunichi Nozawa, Dr. Yohei kakiuchi, and Dr. Hiroaki Yaguchi has provided much support and assistance to me and from them I have also learned many things. To them I offer my thanks.

Yuki Asano has been a great tutor to me and from him I have received much help since the first year. Iori Kumagai has also taught me much about the lab's systems during the first year. To them I also offer my sincere thanks.

Many members of the JSK lab has also provided assistance to the work of this thesis and advice during my studies, and I would like to thank Youssef Ktiri, Yuki Furuta, Yuto Inagaki, Krishneel Chaudhary, and Xiangyu Chen for the support and encouragements they have given me.

I would like to thank the teachers and staffs at the Japanese Language Class of the School of Engineering, for the passion and enthusiasm they have shown to the student, especially Dr. Yumiko Furuichi, from whom I have learned more than just the language skills I needed.

I would like to thank Dr. Elizabeth Croft, who has allowed me to conduct the user study at the CARIS lab, Dr. Dinesh Pai, who has allowed us to use the motion capture system for data collection, my collaborator Matthew Pan, as well as the participants who have volunteered their time for the study.

I would also like to acknowledge the Ministry of Education, Culture, Sports, Science and Technology of Japan for providing financial support for a large part of my doctoral studies.

Finally, I would like to thank my parents and grandparents for giving me the freedom to pursue my interests, and for the continual, unconditional support they have provided all the years.

1 Introduction

Over the past two decades there has been increasingly more developments in service robotics. Researchers have been working towards many different applications including hospital delivery [1], intelligent factory assistants [2]–[4], surgical assistants, shopping mall robots [5]–[7], and companion robots for senior care [8]–[12]. Unlike traditional factory robots, which are only limited to working in confined, structured places and in the absence of humans, service robots will need to work in unstructured humans environments and in close proximity to users.

The development of service robots is driven by a few factors. One being the benefit of increasing work productivity through the use of cooperative robots. Humans and robots have different capabilities. Robots excel in executing repetitive, high-precision motions and are better at performing high power or heavy lifting tasks. Humans, on the other hand, are generally better in performing dexterous tasks and problem solving. Thus, we can leverage the different strengths of humans and robots to complement each other. For example, Cobots are a class of intelligent assistive devices designed to help factory workers [2]. Cobots provide power assist to allow human workers to lift heavy loads with less force exerted and provide motion guidance by creating virtual constraining surfaces. This allows the human workers to perform the decision making, such as how or where to move the workpiece, while providing positioning precision and reducing the risk of overexertion or repetitive motion injuries [13].

Another driving factor for the development of service robots is the expected labour shortage in many countries. With an increasing life expectancy and a lowered birth rate, many countries in the world are facing population ageing. A report published by the United Nations states that population ageing is a global issues and will intensify in the coming decades [14]. As the working age population decreases, the number of workers available in various industries will diminish. Especially in the area of healthcare, as existing care providers retire, the number of seniors requiring care also increases [15]. As a result, many countries have been investing in robotics to address the issue of labour shortage. By creating robots to automate certain tasks, we can help fill in some of the labour shortages. In the area of healthcare, there has been development of home care robots, therapy assistants, hospital deliverers, and companion robots [1], [11], [16], [17]. Such robots aim to lighten the work load of healthcare providers by helping them carry out non-critical tasks, thus, allowing the limited number of human healthcare providers to utilize their time more efficiently and attend to more critical tasks.

Service robots have a wide diversity in their target applications. Researchers have been working on home care robots for assisting seniors or patients with daily lives [9], museum or mall guide robots for engaging and interacting with visitors [18], intelligent assistants for helping human workers with factory assembly [2], [3], as well as astronaut robots for assisting space exploration [19]. Recently, many new service robot platforms have also emerged, including Toyota's HSR (Human Support Robot), Rethink Robotics' Baxter and Sawyer, Honda's Asimo, Willow Garage's PR2 (Personal Robot 2), General Motor's Robonaut, and Boston Dynamic's Atlas. Many academic research projects are being conducted in partnership with industrial companies [2], [3], [19], reflecting the industry's growing demand for service robots [20]–[22]. Market research reports also indicate that significant growth in the global service robot market is expected [23]. Thus, we can expect service robots to become more prevalent in the society in the near future.

1.1 Human-Robot Interaction

In many applications of service robots, the robots will work in close proximity with humans. These robots will have to collaborate with humans and will have many interactions with them. Human-robot interactions can be divided into two categories: non-physical interaction, and physical interaction. Non-physical interactions often take form of gesture, speech, or conversations. For example, museum tour guides explaining art pieces to visitors and engaging them in questions [18], [24], shopping mall service robots providing information to shoppers [25], or home companion robots for reminding elders about their schedules [10]. Physical interactions can include cooperative lifting and positioning with factory robots [2], [26], [27], passing of tools between assistants and mechanics [28], [29], dancing [30], or rehabilitation exercises administered to patients [16]. In non-physical interactions, if the robot does not function properly or behave appropriately, it may confuse the users or cause frustration, thus resulting in ineffective cooperation and poor user experience [31]. In physical interactions, if the robot malfunctions or acts in inappropriate ways, the consequences can be much more severe. The robot could cause damage to the environment, damage to itself, or worse, injure the user and other people around it. Thus it is important that we design proper robot behaviour for human-robot interaction in order to ensure efficient cooperation, and even more so for physical human-robot interaction to ensure safe cooperation.

Mistakes or injuries often occur when there is misunderstanding or when the robot acts in unexpected ways. Therefore, it is important to design robots that move in predictable ways and make interactions intuitive for users. By designing robots with intuitive behaviour, we can reduce the chances of causing misunderstanding and injury. Furthermore, intuitive robots can also minimize the time required for training users on how to use the robot, and reduce the mental load of the user during collaboration.

Studies have shown that people naturally anthropomorphize robots [31], [32]. When working with robots, people tend to expect robots to behave like humans and reason about the robots' actions [31], [33]. Thus, towards making robots more intuitive to users, many researchers have been using the approach of programing humanlike motions onto robots and making them behave more humanlike [34]–[37]. Indeed, studies have shown that when robots behave more humanlike, users tend to have a shorter reaction time, feel safer working with the robots, and indicate preference for the robots [34], [35]. Thus, programming robots to behave like humans or enabling robots to learn how to behave more humanlike can increase efficiency in human-robot cooperation and safety in physical interaction.

1.2 Object Handovers

As mentioned previously, there are many types of physical human-robot interactions. One type of such interactions that is commonly found in many service robot applications is object handover. For example, a shopping mall robot handing out pamphlets, a waiter robot serving drinks and snacks, a factory assistant or astronaut robot handing over tools, or a homecare robot bringing over the TV remote controller [17], [19], [25], [38]–[41]. Humans regularly perform many handovers each day, and in general they complete each handover safely, efficiently, and smoothly, without the need of meticulously planning each handover. However, object handover tasks actually involve many components.

Object handover is a task involving two parties – the giver, and the receiver – in which an object is transferred from the giver to the receiver. In a typical human-to-human object handover, as the giver prepares to hand over the object to the receiver, the giver and the receiver exchange gaze to communicate each other's readiness. If the two are too far from each other, then either the giver or the receiver or both need to approach the other to reduce the distance between them. The giver then needs to reach out his/her hand to present the object to the receiver, and the receiver also needs to reach out his/her hand to take the object. In most cases, the receiver reaches his/her arm out before the giver finishes reaching out, or even nearly at the same time [42]. As the receiver's hand comes into contact with the object and begins to grasp the object, the giver needs to adjust the applied grip force appropriately and control the release timing of the object properly in order to ensure a safe and efficient transfer of the object [43]. Releasing too early will risk dropping the object, and releasing too late could result in a rough transfer in which the receiver needs to pull too hard on the object. Finally, after object transfer, both giver and receiver need to retract their arms to complete the handover. Because of the many components involved, object handover is still a challenging task for robots, and even now, it is still difficult to achieve smooth, efficient human-robot handovers.

1.3 Research Goals

Since object handover is a basic common task occurring in many service robot applications, it is crucial that robots perform handovers well in order to guarantee safe and effective human-robot cooperation. Thus, the research goal of this thesis is to enable robots to handover objects properly when interacting with humans. Object handover involves many components including gaze, arm reaching motion, body posture, haptic exchange, and verbal communication. Among the many aspects of object handover, one important component is handover grasp configuration – that is, how the giver grasps the object, and how he/she orients the object when handing it over to the receiver. Handover grasp configuration affects the efficiency, safety, and receiver's grasp comfort in the handover. Thus, this thesis places an emphasis on handover grasp configuration.

To illustrate the importance of grasp configuration in handovers, consider the examples of handing over a hammer and handing over a kitchen knife. Figure 1-1 shows a robot handing over the objects to a person using different grasp configurations. In Figure 1-1A, the robot grasps the hammer handle and presents the hammer head to the person, while in Figure 1-1B, the robot grasps the hammer head and presents the handle. Comparing these two grasp configurations, the one used in Figure 1-1B is the better choice since it allows the receiver to more comfortably grasp onto the hammer by the handle. Furthermore, it is more efficient since once the receiver takes the hammer by the handle, it is in an orientation ready to be used. In Figure 1-1C, the robot hands over a kitchen knife with the knife tip pointing towards the receiver, while in Figure 1-1D the knife tip is pointed downwards. Comparing these two grasp configurations, the latter one is definitely preferred since it reduces the risk of injuring the receiver. Form these examples, it can be seen that when handing over an object to a person, the robot needs to determine a proper grasp configuration to use, in order to ensure a safe, efficient interaction.

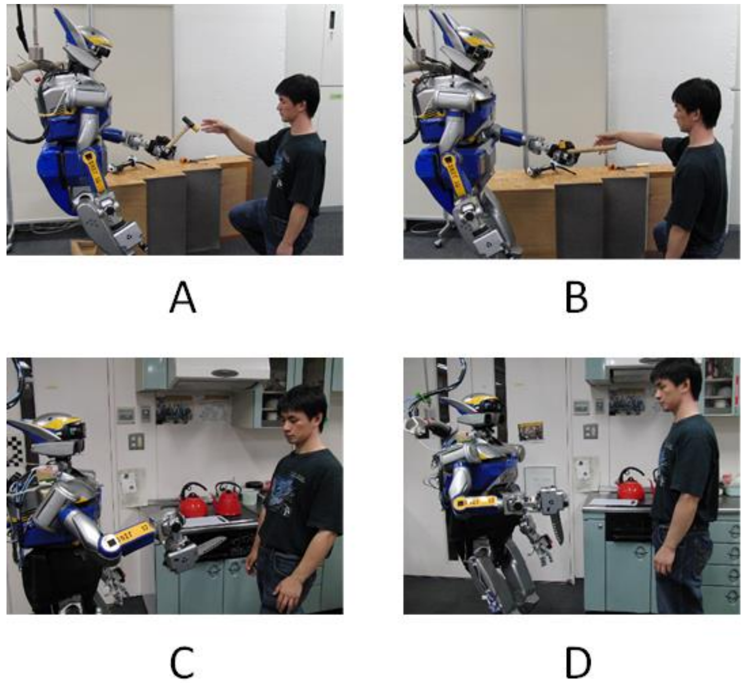

Figure 1-1 Comparison of handover grasp configurations for hammer and kitchen knife. (Published in Chan et al. 2013 [44].)

Determining the proper handover grasp configuration for different objects is not a simple task as it depends not only the physical properties of the object such as size, weight, geometry, and surface friction, but also more abstract factors such as object value and culture [45], [46]. To enable robots to hand over objects properly, existing approaches often rely on programming explicitly the grasp configurations to the robot, or specifying which part is to be presented or held for each object. However, for a large number of objects, these methods can take a lot of time and effort from the programmer and may not generalize well to unknown objects. Furthermore, given an object, there may not necessarily be only one appropriate grasp configuration. For objects that have more than one possible usages, depending on the receiver's intended use of the object, the appropriate grasp configuration may be different. As the applications of service robots expands, robot will have to work with larger number of objects, and they are bound to encounter new objects. Thus, to address the aforementioned issues, our goal is to construct a framework that will enable robots to automatically determine proper handover grasp configurations for various objects.

Generally, people know about the usages of the objects around them, and they know what the appropriate handover grasp configurations for the objects are. As robots enter society, they will have more time working around humans and interacting with them. Therefore, the approach of this thesis is to enable robots to learn proper handover grasp configurations by observing how humans use various objects, and how they handover various objects. This thesis proposes a framework for learning grasp configurations through observation of handover demonstrations, and generalization to new objects through observation of object usages.

In enabling robots to learn handover grasp configurations from observing how humans hand over objects, we first need to better understand the nature of the grasp configurations used by humans. To this purpose, this work investigates and compares the handover grasp configurations used by humans in different conditions. This thesis will introduce a novel method for identifying patterns in handover orientations and propose a distance minimization approach for computing appropriate handover orientations from a set of observations. This work will present the analysis results of the handover orientations used by humans for a set of common everyday objects.

Next, towards enabling robots to generalize handover grasp configurations to new objects, this thesis will present an object grouping and classification method based on object usage observations. The proposed method focuses on object movement and interactions between objects during usage, and uses a set of action features for describing each object. Experiment with a set of common objects validates the proposed method.

After presenting the methods for learning grasp configurations from handover observations and generalization based on object usages, this thesis will present the implementation onto a Kawada Industries HRP2V robot for observing and extracting handover grasp configurations from human-robot handovers, and the subsequent execution of handovers using proper grasp configurations for known and unknown objects. Finally to demonstrate its use in real world situations, this thesis will present an integration of this work into a household service robot application.

1.4 Object Affordance

The proper handover grasp configuration of an object depends on the object's affordances [47]. The term *affordance* was originally introduced in psychology to refer to the relation between an object or an environment with an agent, that permits the agent to carry out certain actions, where the agent can be a person or an animal [48]. The term has later been adopted in the field of human-machine interface, and the idea of perceivable affordance is further developed, based on the grounds that, from a design perspective, an affordance has to be perceivable by the agent to be effective [49]. In the field of robotics and human-robot interaction, the term has evolved to refer to the relation between an object, or a part of an object, and a human or robot agent, that allows the agent to perform specific actions, perhaps to achieve certain functions [47], [50]–[52]. This thesis will adopt the accepted definition in the field of robotics and human-robot interaction, where the term affordance is used to refer to the relation between an object or a part of the object, and a human or robot agent, that allows the agent to perform perceivable actions. For example, a hammer affords hammering of nails, thus, a hammer has the affordance of hammering. More specifically, the handle part has the affordance of grasping, while the hammer head has the affordance of hitting or hammering. The affordances of an object determines how an object can be grasped and how an object can used to achieve its intended function. Thus, an object's affordances influence its proper handover grasp configuration. For this reason, the approach of this thesis for determining proper handover grasp configuration for objects focuses on capturing the affordance information of objects.

1.5 Contributions

The contributions of this thesis includes the following. First, this thesis presents a framework for learning proper handover grasp configurations from observing handover demonstrations and object usages. Compared to existing methods that rely on users providing object specific information, the framework presented can significant reduce the time required from the programmer or user, and allows the robot to handle new objects as well.

Secondly, the user study conducted to investigate handover orientations used by humans in different condition has collected the orientations used in three different conditions for twenty common everyday objects. For identifying patterns in handover orientations and determining a proper handover orientation from a set of observations, this thesis also puts forward the novel idea of *affordance axes,* and an optimization-based approach for computing mean orientations. Computing mean orientation is known to be a difficult problem since orientations are rotations in 3D space belonging to the SO(3) group, and there is no standard method of computing averages for the SO(3) group. Results from the user study identified that grasp configurations used in natural handovers may not always be the most appropriate configurations, thus, indicating that addition consideration might be needed when a robot tries to learn from observing natural handovers. Through the process of investigating the grasp configurations used in different conditions, mean handover orientations for the twenty objects used in the experiment have been computed, and this thesis also presents these orientations, which can potentially be used by robots for achieving safer and more efficient robot-human handovers.

Thirdly, this thesis presents a method for grouping and classifying objects based on their observed usages. As proper handover grasp configuration depends on an object's affordances, the proposed method aims to capture objects' affordances by extracting actions features from object movements and interactions with surroundings during usage. The grouping and classification method allows a robot to determine the proper handover grasp configurations for unknown objects by generalizing those it has learned.

To validate the presented framework, this thesis presents an implementation on an HRP2V robot. Experimental results show that the robot is able to learn handover grasp configurations for fifteen objects from observations of human-robot handover demonstrations. To solve the challenging problem of identifying the handover grasp configurations used, the presented implementation uses a sequential Monte Carol technique to predict the object's pose, instead of using machine learning algorithms as most exiting object detection methods do. The experiment documented in this thesis also shows that subsequently when the robot is asked to handover the fifteen known objects and three new objects, it is able to handover the known objects using the appropriate grasp configurations computed from the observed demonstrations, and also the unknown objects using appropriate configurations generalized from the learned ones.

1.6 Thesis Outline

In the following of this thesis, Chapter 2 first provides the motivation to this work and a literature review. Chapter 3 presents a new framework for enabling robots to automatically determine appropriate handover grasp configurations for various objects. Chapter 4 documents a user study conducted to investigate and characterize the handover orientations used by human in different conditions, and at the same time presents novel approaches for identifying patterns in handover orientations and computing mean handover orientations. Chapter 5 then provides an object grouping and classification method to be used for generalizing learned handover grasp configurations to new objects. To verify the proposed framework, Chapter 6 realizes the presented framework by providing an implementation on an HRP2V robot, and a set of experiments conducted to test the implementation. Finally, Chapter 7 concludes this thesis.

2 Motivation and Related Works

This chapter first provides the motivation for the development of robots capable of performing handover properly with humans by discussing some of its implications in Section 2.1 from the perspectives of human expectations, efficiency, safety, and social conventions. Section 2.2 then provides a review of the literature, discussing existing studies on the various aspects of handovers. The review in Section 2.2 encompasses the handover components of gaze, body approaching motion, handover cues, grasp configurations, and force control. Section 2.3 then identifies grasp configuration as an aspect that can benefit from improvements, and discusses the challenges in determining proper handover grasp configurations for different objects. Section 2.4 then summarizes existing methods for computing handover grasp configurations, and discusses their advantages and limitations.

2.1 Implications

As object handover is a task arising frequently in many service robot applications, it is important that robots perform handovers well. This section discusses its importance from a few different perspectives including human expectations, efficiency, safety, and social conventions.

2.1.1 Human Expectations

Studies such as those conducted by Short et al. and Kahn et al. show that when interacting with robots, people have certain existing expectations of how the robots should behave, and that when the robots' actions do not meet their expectations, people try to reason about the robot's behaviours [31], [33]. Indeed, Fincannon et al. observed that during interaction, people tend to anthropomorphize robots [32]. This means that users naturally expect robots to have a certain degree of likeness to humans in their behaviours, and they expect robots to possess a certain degree of skills comparable to humans. Kiesler et al. and Huang et al. have also provided evidence showing that during collaboration when humans, robots implementing more humanlike behaviours are more acceptable to users and have greater work efficiency [53], [54]. In Chan et al.'s study comparing different robot behaviours in handovers, they also found that during object transfer, when a robot's behaviour does not match the person's expectations, it can cause frustration in the user, compromise their sense of safety of the task, and even cause the object to be dropped [35]. Thus, robots need to be able to perform handovers properly in a manner matching user expectations, from user acceptance and cooperation efficiency perspectives.

2.1.2 Efficiency

The ability of a robot to perform handovers properly affects the efficiency of the handover both in terms of time and effort. In comparing different robot reaching arm trajectories, Huber et al. showed that by using an appropriate type of motion, the robot can shorten the amount of receiver reaction time [55]. Similarly, Cakmak et al.'s study comparing the use of different poses in handovers also showed that by using appropriate poses for carrying and handing over objects, the receiver's reaction time in taking the object is also shortened [56]. Specific to grasp configuration in handovers, Aleotti et al. showed that when a robot hands over an object with the handle part oriented towards the human receiver, the receiver also is able to more quickly take the object [57]. In Chan et al.'s study, they showed that by using an inappropriate handover behaviour, it increases the amount of time required for object transfer to more than twice the natural amount [35]. Furthermore, the study also revealed that the receiver may need to exert over 70% more effort to take the object when an inappropriate handover controller is used. Since object handover often occurs as a subtask in larger contexts, the efficiency of a handover also affects the efficiency of the larger task at hand. Thus efficient robot-human handover is required for efficient task completion and efficient human-robot cooperation.

2.1.3 Safety

During a handover, depending on how the robot behaves, the safety of the person, the safety of the object, and the safety of the robot itself is also affected. Edsinger and Kemp showed that when a robot does not time the release of the object properly, it can greatly increase the likelihood of the object being dropped, thus risking damage to the object [58]. The study by Chan et al. also shows that risk of dropping the object can result from improper force control by the robot during object transfer [35]. In addition to object safety, user safety may also become an issue depending on the object that is being handed over. For example, when handing over objects such as kitchen knives that have certain parts that are dangerous, the robot should take care not to point those parts at the human. This has been illustrated in Chapter 1 and is discussed by Kim et al. in their paper proposing a handover grasp planner [59]. Bohren et al.'s study on a bottled beverage delivery robot also reveals how user safety can be affected [38], [60]. Their handover controller focuses on object safety, trading off ease of handover. As a result, the person taking the bottle needs to pull very hard to take the object. When the robot finally releases the object, it causes a high-jerk motion, which in turn may cause carbonated drinks to spill when the user opens the bottle. If the same controller were used to hand over hot beverages in a cup, it would likely cause the liquid to spill on the robot and the person, thus causing injury to the person and damage to the robot. The above examples demonstrate that proper human-robot handovers are required to ensure safety in human-robot interactions.

2.1.4 Social Conventions

People's actions and behaviours are often influenced by various social conventions which may be determined by culture [61], [62]. Such conventions impose rules on how one should behave and affects how one expects others to behave and act. Since it is has been shown that users anthropomorphize robots and reason about robots' actions [31]–[33], users would expect robots to follow these social conventions as well. Handover is a task in which the participants' behaviours are govern by culture. Culture varies across country and region. In some cultures, younger people are required to show politeness and respect to older people and store clerks to customers, and when handing over or receiving an object, using two hands instead of one is expected to show politeness and respect [59]. Wang et al. showed that people are more receptive to recommendations given by a culturally conformant robot [63]. Trovato et al. and Eresha et al. have also found that people are more comfortable with and prefer a robot whose behaviours matches their culture [64], [65]. Therefore, robots need to observe such social conventions in order to be accepted by users and society.

2.2 Handover Components

Object handover involves many components, and the literature pertaining to object handover consists of studies focusing on different aspects. Among the different aspects are gaze, approaching motion, handover cues, grasp configuration, and force control. The following subsections provide a review of existing studies related to each of these aspects.

2.2.1 Gaze

Gaze is an important channel of communication among humans. As people naturally look at what they are focusing on, others usually try to determent where a person's attention lies by looking at their gaze direction [66]. Human eyes use saccade movements which are quick shifts in gaze fixation from one point to another [67], and saccade is one of the fastest motions the human body is capable of. Therefore, it often leads other body moments. For example, studies in psychology have shown that when driving, people direct their gaze at points they will be steering through [68], [69]. Investigations of gaze patterns of expert cricket batters also reveal that they anticipate the location of a pitched ball flying at high speed using saccade movement of the eyes [70]. Of course, people also direct their gaze ahead of them so that they can see ahead and plan their next movements. Thus, gaze often leads a person's actions and can be used to predict the person's movements and intentions.

Studies have shown that gaze can be used in human-robot interaction to detect attention, achieve engagement, and direct attention [24], [71]–[77]. These properties of gaze also applies to the task of handover. Gaze can be used for signaling or detecting handover intent [29], [78]–[80]. Strabala et al. constructed a predictor of handover events based on recorded video footages of human handovers [78]. They built a classifier capable of predicting when a handover is impending based on the events occurring in a prior time window. They reported that out of the four key features in their decision tree, two involved gaze. According to these key features, prior to a handover, the receiver's gaze must be directed towards the giver, and the giver's gaze is either directed at the object or the receiver.

To investigate how a robot's gaze affects human receivers' experiences in handover tasks, Moon et al. conducted an experiment to test different gaze patterns of a robot giver [79]. They compared three different gaze patterns: 1) No gaze, where the robot fixes its gaze at the initial position of the object in its hand for the entire duration of the handover. 2) Shared attention, where the robot starts with its gaze on the object's initial position, and as it reaches it arm over to hand over the object, it looks at the final position of its reach to signal the projected handover location to the receiver. 3) Turn taking, where the robot behaves as in the shared attention case, but towards the end of its reach, it quickly directs its gaze to the receiver's face. Results of their study show that participants reached for the object significantly earlier in the shared attention case, and that they show a trend of preferring handovers in the turn taking case. According to these findings, Moon et al. suggested that gaze can be used to improve user experience in human-robot handovers, and improve handover efficiency.

Gharbi et al. examined how different gaze patterns of the giver affects the perceived naturalness of the handover [81]. They showed videos of a person or a robot handing over an object by placing it on a table in front, while exhibiting different gaze behaviours, alternating between looking at the object and the receiver. They found that regardless of whether the giver was a person or a robot, people found most natural when the giver exhibits a gaze pattern of object-receiver, or receiver-object-receiver.

2.2.2 Body Approaching Motion

Basili et al. compared a giver's motion between handing over an object and placing an object on a table [55]. They tracked the giver's head and the object's positions, and results indicate that there is little difference between the two cases. A few researchers have also studied a robot's approaching motion for handing over objects [82]–[84]. In Koay et al.'s study, their robot handed over an object to a seated receiver. The robot approached the person from front, front left, front right, left, or right. Results indicate that participants preferred the robot to approach from the front, since it allows the robot's motion to be most visible. However, Walters et al.'s study had contradictory findings [83]. In their study, they tested a robot handing over when the person is seated and when the person is standing. The robot approached from the front, the left, and the right, and according to their results, seated participants preferred the robot approaching from the front the least since it makes the robot appear aggressive. When the participants are standing, then a frontal approach is more acceptable.

2.2.3 Arm Reaching Motion

Flash and Hogan showed that the human arm reaching motions can be model by a minimum jerk trajectory [85]. Based on this, researchers have modeled the minimum jerk trajectory for comparing humanlike trajectories with other typical robot trajectories [34], [86]. Shibata et al. used a one degree of freedom robot on a table to study different handover trajectories [86]. In their study, their robot handed over a glass to a person sitting across a table. Comparing a bell-shaped velocity pattern modeling a minimum jerk trajectory with triangular and trapezoidal velocity profiles, which are typical robotic trajectories, they found that with a bell-shaped velocity pattern, the robot is able to elicit better evaluations from users such as being more skilled, more careful, and more pleasant.

Extending Shibata et al.'s work to three dimensional space, Huber et al. tested humanlike handover trajectories using an anthropomorphic robot [34]. They programmed their robot to use a Cartesian minimum jerk trajectory and a trapezoidal joint velocity trajectory. Comparing the two types of trajectories, the researchers found that the minimum jerk trajectory allowed the human receivers to have a shorter response time and feel safer. Shibata et al.'s and Huber et al.'s studies show that employing humanlike motions to robots in handovers can make the task more time efficient and allow users to feel safer.

Kajikawa et al. also studied human arm reaching motions in handovers, but instead of using a quantitative approach, they identified a set of common characteristics qualitatively [87]. Form their study, they found that receives generally begin their reach after the giver has reached maximum approach velocity. Receivers tend to start with a quick trajectory with undetermined hand direction, and eventually they adjust their hand direction and slowdown their trajectory to match the giver's. Based on these findings they created a trajectory planner for generating robot handover motions and tested it in simulation.

Yamane et al. noted that in natural human-to-human handovers, the receiver often begins reaching for the object before the giver has completed their reach or the final handover location is known precisely [42]. In order to allow a robot receiver to respond in time and achieve natural human-robot handovers, the robot needs to recognize the giver's handover action and generate its own motion quick enough, otherwise, the person will have to be waiting holding the object. To address the timing and uncertainty issues, Yamane et al. used an approach where the robot synthesizes its motion using a database of handover motions. In their work, they first collected a set of handover motions of giver-receiver pairs and constructed a database of reaching motions. In the database, the reaching trajectories are segmented into small time steps, and the transition probabilities from segment to segment are computed. When executing a handover with a human giver, the robot observes the human's reaching motion and determines their most probable trajectory in the next time step. The robot then synthesizes its own motion using the corresponding receiver trajectories in the database. Using this approach, Yamane et al. were able to achieve human-robot handovers that more closely mimic the timing of natural handovers.

Towards determining handover location and time, Hart et al. attempted to create a predictor based on the human giver's reaching trajectory [39]. By using a skeleton tracker, they observed the giver's arm motion and tried to extrapolate the end point of the trajectory for predicting the handover location and time. Using their method, although they were able to achieve predictions with good precision, they pointed out that the algorithm required observation of most of the reaching trajectory, and that if the predictor were to be applied to human-robot handovers, it would need to be refined with a richer human motion model.

2.2.4 Handover Cues

Towards identifying cues that indicate a person's intent to handover an object, such that a robot receiver can recognize these cues and respond appropriately when being handed over an object, or that a robot giver can use these cues to signal a person that it is handing over an object, Strabala et al. built a classifier to predict when a handover is impending [29], [78]. They collected video data of pairs packing a picnic basket together, and based on features extracted over a time window prior to occurrences of handovers, they constructed a decision tree. From their decision tree, they identified four important features for detecting handovers: 1) A handover cannot be already occurring. 2) The giver must be holding an object and facing the receiver. 3) The giver must turn to face the receiver, and the receiver looking at the giver. 4) The giver is looking at his/her hand or the object. Using their constructed decision tree, they were able to predict when a handover will take place with good accuracy.

Micelli et al. developed a framework for enabling a robot receiver to take an object handed over by a human giver [88]. They used a skeleton tracker and point cloud data to determine the hand position of the giver, and they used a support vector machine (SVM) to determine whether the person was holding an object. To determine whether the human is handing over an object, they looked for three handover cues: 1) The person is near the robot. 2) The vector from the person's shoulder to his/her hand points towards the robot's upper body or hand. 3) The person's elbow joint angle is bent at less than 150°. By detecting when these cues persist for five consecutive frames (at a frame rate of five frames per second) their robot was able to detect when a person is handing over an object and take the object from the person.

To determine how a robot should hold onto an object while handing it over, Cakmak et al. conducted an online survey where they presented images of a robot holding a cylindrical object with different poses, and asked participants to label the robot as one of the four categories of handing over the object, looking at the object, showing the object to someone, or none of the above [56]. Different poses of the robot were created by varying the robot's arm extension, hand position, or the tilt of the object. Results show that an extended arm is the most prominent feature for conveying the intention of handing over the object.

2.2.5 Grasp Configuration

Studies found in the literature have shown that the use of a proper grasp configuration in handovers is important. Aleotti et al. conducted a user study to test the effects of proper grasp configurations in robot-human handovers [57]. In their experiment, the robot handed over a hammer and a detergent bottle to the participants. They compared a proposed configuration where the handle of the object is presented to the receiver, with an alternative configuration. They reported that with the proposed configurations, the participants felt safer interacting with the robot and required a shorter amount of time to take the object.

Cakmak et al. also showed that by using different grasp configurations to create spatial and temporal contrasts, the robot's handover intent can be made clearer [56]. Through an online survey, they identified grasp poses that are more commonly interpreted by users as carrying and those as handing over, and by programming the robot to use "carrying" grasp poses when carrying the object and "handing over" grasp poses when handing over, they created contrast. Results from a user study show that by using appropriate grasp poses to create contrasts, participants were able to more easily detect when the robot intends to handover the object, and they were able to complete the handovers in a shorter period of time.

The importance of proper grasp configuration in handovers is also recognized by Kim et al., and they have proposed a handover grasp planner for a dual arm robot [45]. They first labeled the affordances of each object, such as which part is to be grasped and which part is to be presented to the receiver, and provided these information to their grasp planner. The planner then computes the handover grasp configuration by searching for parallel flat surfaces on the object to grasp. The planner is capable of preforming mid-air re-grasps when needed to orient the object properly. Although the planner is able to plan three different grasps for different situations, including single and dual hand grasps, it does not include a method for determining when to use which one.

In another study by Cakmak et al., they have also proposed and compared two methods for computing handover configurations for objects [89]. The first method plans configurations using a human kinematic model. As for the second method, they first collected a set of good and bad examples from users for each of the five tested objects, by showing different configurations to the users or asking the users to adjust a set of robot parameters on a computer interface. The robot then learns what the good configurations are for each object based on the user provided examples. According to their user study results, they reported that while the planned configurations provide better reachability to the receiver, people may find them unnatural since consideration of object function is not included. Learned configurations, on the other hand, although produce preferred handovers, requires a tedious process of collecting the samples from users and may not scale well.

2.2.6 Force Control

Force control is another important aspect of handover, and other object manipulation tasks in general. In such tasks, the grip force needs to be regulated carefully to prevent the object from slipping and dropping, or from being crushed. Humans rely heavily on their tactile sensing capabilities for regulating the applied grip force. The inside of the human hand is densely covered with four different types of tactile afferents with densities that can reach over 100 afferents per cm^2, capable of sensing micro slips and deformations at frequencies up to 400 Hz [90]–[92]. It is known that individuals with compromised tactile perception or proprioception lose the ability to regulate their grip force properly and experience great difficulties even when simply picking up objects [91], [93], [94].

Investigating grip force control in handovers, Mason and MacKenzie measured how human givers and receivers control their grip forces with respect to the experienced load forces due to object weight and dynamic effects. They found that prior to and after object transfer, both giver and receiver maintain a steady grip-force-to-load-force ratio. Furthermore, the applied grip force is efficiently kept at the minimum that is required to prevent slip, with a small safety margin included. These results are in agreement with findings from studies in haptics measuring how people control their grip force with respect to load force when holding and moving an object [91], [95]–[99].

Extending Mason and MacKenzie's work, Chan et al. further investigated grip forces and load forces in handovers during object transfer [43]. Examining the haptic interaction between giver and receiver, they discovered that during object transfer, both the giver and the receiver actively adjust their grip forces according to their experienced load forces, such that an approximately linear relationship between the grip and load forces is observed. Chan et al. subsequently designed a human-inspired handover controller for robot givers based on the identified grip force control strategy, and through a user study, they showed that compared to existing strategies [17], [38], [58], [60], [100]–[102], their controller is able to achieve more robust, safer, smoother, and human preferred handovers [35], [103].

2.3 Challenges in Determining Proper Grasp Configurations

The literature review provided above shows that there has been many studies on the different aspects of handover, and indeed, the advances in the various components of handover have been building towards better interaction in human-robot handovers. However, one aspect still needing improvement is the determination of proper handover grasp configuration. Proper grasp configuration differs from object to object. Thus, a robot cannot simply use the same configuration for all objects. Although the importance of proper grasp configuration has been proven in various studies [56], [57], [89], a good method for computing them for different objects is still lacking. The reason for the lack of a good method for determining handover grasp configurations is due to the inherent challenges associated with the problem. This section discusses these challenges below.

One of the factors a robot needs to consider when computing handover grasp configuration is the physical properties of the object including size, weight, shape, surface friction, and structural strength. For example, given a candidate grasp point on the object, the robot needs to consider whether the size and weight distribution of the object permits a stable grasp, if the surface friction is sufficient to prevent the object from slipping, and whether the object has enough structural strength such that it will not break on its own weight when picked up by a certain part. Some of these properties such as surface friction and structural strength are not directly measureable by simply looking at an object. Humans rely on prior knowledge and experience when manipulating objects [91], but when an object is new or unknown, it becomes an even bigger challenge for both humans and robots.

When computing grasp configurations for handovers, the robot should also give consideration to the receiver. The goal of a handover is to give the object to the receiver, and in doing so, the robot should take into account the receiver's ease of grasping and the receiver's grasp comfort. Consider the comparison of two grasp configurations for handing over a can of coffee shown in Figure 2-1. In the left figure, the robots grasps the can by the middle, leaving no place for the receiver to grasp and take the object easily. In the right figure, the robot grasps the can by the bottom half, leaving the top half available for the receiver to grasp easily. Thus, the grasp configuration shown in the right figure is preferred. Grasping for handover is different from grasping for picking up or for using the object. Robotic grasping is in itself another challenging problem, and although there is a large body of research in grasping, and there exists many grasp planners [104]–[108], these planners are not necessarily suitable for the purpose of handover. Object stability is often the priority in grasping, and most grasp planners are based on object geometry [106], [107]. But the proper grasp configuration of an object is often dependent on its affordances. Thus, the computed grasps by these planners may not be the most considerate of the receiver or the most appropriate for handover.

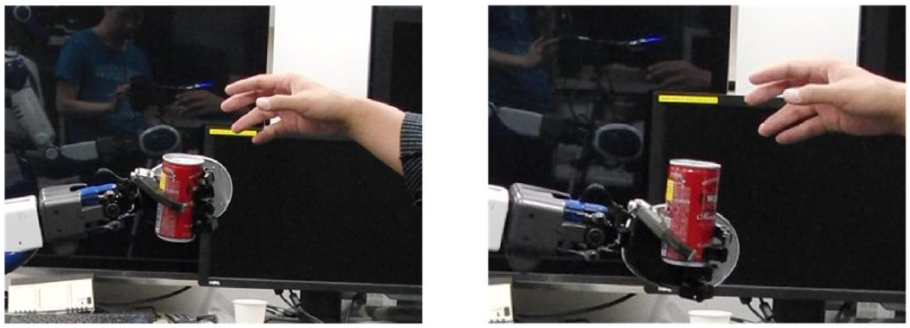

Figure 2-1 Comparison of handover grasp configurations for a can of coffee. (Published in Chan et al. 2013 [38].)

Object affordance is an important factor that should be considered when computing handover grasp configurations. Different objects have different usages and depending on how an object is intended to be used, it affects what people consider as appropriate configurations. For example, a plate is used to carry food with the concave surface upwards. Thus, people find it unnatural when a robot hands over a plate with the concave surface facing sideways [89]. Furthermore, the purpose of handing over an object is often to allow the receiver to use the object to complete some task, such as in the setting of a factory assistant handing over a tool [28]. Therefore, to ensure an efficient handover, the robot should use a configuration that allows the receiver to readily use the object after taking it without needing to re-grasp (such as presenting the handle of a hammer as in Figure 1-1B). Depending on the affordances of an object, the safety of the handover can also be affected by the grasp configuration used, such as with the kitchen knife handover example depicted in Figure 1-1. Object affordances cannot be accurately determine by simply observing the object's shape [109]. Objects having similar geometries might have very different usages (for example, consider a plate versus a flying disc), and thus, different proper handover grasp configurations. Therefore, it is difficult for a robot to determine the proper handover grasp configuration of an object only by looking at the object.

Proper handover grasp configuration can also be affected by culture and social conventions. Culture often influences how one should behave and how one is expected by others to behave [61], [62], [65]. This is true for handover behaviour as well. In certain cultures, when handing over an object, the giver is required to use both hands as a sign of respect and politeness [45]. Thus when handing over to elder people or to customers, the giver is required to use a two handed grasp configuration. In such cases, the proper handover grasp configuration for an object not only depends on the object itself, but also on the situation, the role of the robot, and identity of the receiver.

2.4 Existing Methods for Computing Handover Grasp Configurations

The literature provides a few different methods of determining proper handover grasp configurations of objects. The most straight forward method is to explicitly teach the proper grasp configuration of each object to the robot, such as in Aleotti et al.'s study on investigating the importance of proper grasp configurations in handovers [57]. This method has the advantage of allowing the user to precisely specify the grasp configuration for each object. However, for a larger number of objects, this would take a lot of time. Furthermore, when the robot encounters a new object, the robot would not know the proper grasp configuration for it.

Another approach found in the literature is to provide object affordance labels to the robot, and let the robot plan grasp configurations accordingly. In Kim et al.'s work, they labeled the grasp sites of each object and provided this information to their planner [45]. The planner is then capable of computing three different grasp configurations to present the labeled grasp site to the receiver. Kim et al.'s method has the advantage of enabling multiple grasp configurations for each object (e.g., one handed or two handed). However, this also has the drawback of being time consuming for a large number of objects and lacks the ability to handle new objects.

Similar to Kim et al., in another study, Aleotti et al. also proposed a method that plans handover grasp configurations using provided object labels [47]. Aleotti et al. provided the receiver grasp site labels of each object to their robot. The robot then segments the object into parts geometrically. Force-closure grasps are then sampled on all parts not containing the labelled grasp site, and the object is handed over in an orientation that directs the grasp site towards the receiver. To address the issue of generalizability, although not implemented, Aleotti et al. suggested that a part-based object classifier that they have developed separately [107] can be used for computing grasp configurations for new objects. While this approach may allow a robot to determine grasp configurations for new instances of know objects, it may still have trouble dealing with new types of object since their classifier is purely geometry based. Proper handover grasp configuration depends heavily on object affordances, and Bicici et al. has made the argument that object function cannot be determined from geometry alone [109]. Indeed, Aleotti et al. did find that their classifier can misrecognize a pair of pliers as a table.

Focusing on the receiver's ease of grasping, Cakmak et al. proposed a planner that computes handover configurations using a human kinematic model [89]. Their planner generates a set of configurations and simulates how the object can be taken by the receiver. Each configuration is evaluated base on how many ways the receiver can take the object. The configuration with the highest score is then selected. In the same paper, Cakmak et al. also proposed another method for learning handover configurations from human provided examples. Good and bad handover configuration examples for each object were collected from users. The users were either shown various handover configurations and asked to label them as good or bad, or the users were ask to provide good and bad examples by adjusting parameters of the robot through a computer interface. Using the collected positive and negative examples, the robot computes appropriate handover configurations for each object. Comparing the planned configurations with the learned configurations in a user study, results indicated that the planned configurations provide better reachability. The planning method also has to potential of computing handover configurations for new objects. However, the planned configurations were found to be lacking in appropriateness, usability, and naturalness, since object function is unaccounted for. Learned configurations, on the other hand, captures these aspects implicitly, and participants found them to be more natural and more appropriate, but lacks generalizability.

2.5 Summary

Proper handover grasp configurations depend on object affordances, and handover grasp configurations that fail to account for this may appear unnatural and inappropriate [89]. It is not easy for a robot to determine the affordances of an object by only looking at it [109]. To capture the affordance information, existing methods for computing handover grasp configurations rely on users to provide information specific to each object. This type of approach, however, lacks scalability and generalizability. To date, a demonstrated approach for determining proper handover grasp configurations that does not require users to teach the robot explicitly about each object is still lacking. In response to this, the following chapter presents a framework that aims to allow robots to automatically learn proper handover grasp configurations of different objects.

3 A Framework for Learning Object Handovers from Observations and Interactions

The objective of this work is to create a framework that allows robots to automatically determine proper handover grasp configurations of various objects to enable effective human-robot cooperation. Existing studies have shown that when a robot employs humanlike behaviour or uses behaviour learned from humans in handovers, human receivers feel safer interacting with the robot [34], have shorter task completion times [56], perceive the robot as more skillful [86], and indicate preference for the robot [35]. Furthermore, Cakmak et al. pointed out that configurations learned from users have usability, naturalness, and appropriateness implicitly encoded into them [89]. Therefore, the approach of this thesis for enabling robots to determine proper handover grasp configurations is to learn by observing humans.

Figure 3-1 shows the proposed framework for learning handover grasp configurations from observations. The framework can be divided into two phases – the learning phase and the application phase. During the learning phase, the robot first observes handover and usage demonstrations of objects. The handover observations can be obtained by either observing human-human handovers, or interacting with humans in handovers. From the observed handover demonstrations, the robot extracts and learns the handover grasp configurations used for each object. From the observed usage demonstrations, the robot tries to learn about the affordances of the objects by extracting a set of descriptive features for each object. The robot stores the learned handover grasp configurations and the extracted object features into a knowledge base, making the connections between objects usage and proper handover grasp configurations. Building up the knowledge base, the robot organizes its contents by grouping together objects with similar affordances.

Once the robot has built up a knowledge base, in the application phase when the robot is requested to handover an object, it first classifies the object into one of the groups in its knowledge base, according to features extracted from the observed usage demonstration of the object. The robot then recalls the observed handover grasp configurations of the instances of the classified group, and calculates an appropriate grasp configuration for the object based on the recalled data in its knowledge base. Once the appropriate grasp configuration for the object has been computed, it can then be passed to a hardware specific planner for determining an appropriate grasp plan and motion trajectory, which in turn can be used by the robot to hand the object over using the computed grasp configuration. The following sections discuss the requirements and challenges of the various components of the framework.

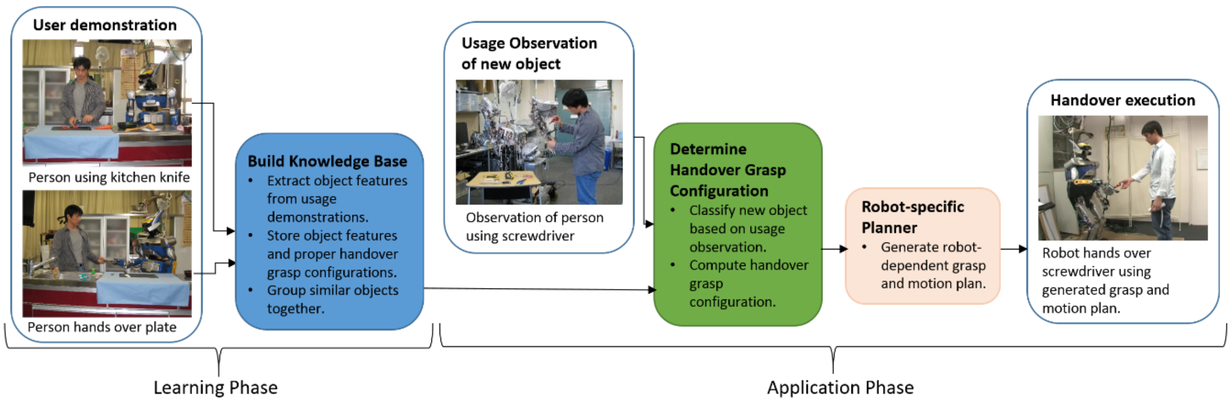

Figure 3-1 Proposed framework for learning handover grasp configurations from observations. (Modified from Chan et al. 2014 [110].)

3.1 Learning from Handover Demonstrations

When trying to learn handover grasp configurations by observing, the robot first needs to be able to recognize when a handover takes place. Human daily motion involves a lot of movements, thus, to be able to determine when a handover is occurring, the robot needs to be able to identify cues that signal a handover. Once the robot has identified that a handover is taking place, it then needs to be able to recognize the grasp configuration used. One challenge in extracting the grasp configuration used is in identifying the object orientation when it is being handed over. While there exists many object detection algorithms, most of them can only detect object location, but not orientation. Furthermore, most algorithms do not cope well with occlusions, but during handover, object occlusion by the giver’s hand is almost guaranteed. To overcome these challenges, the implementation Chapter 6 presents uses a sequential Monte Carlo method to predict the object’s orientation at the time of handover.

3.2 Object Grouping and Classification

After the robot has observed handover and usage demonstrations for a set of objects and accumulated them into the knowledge base, the knowledge base should be organized in a way such that objects with similar handover grasp configurations are grouped together. The classification scheme used for determining grasp configurations of new objects should also allow new objects to be classified into groups with similar proper grasp configurations. Methods that rely on geometrical information may not be the most suitable since an object's function is not necessarily reflected on its appearances [109]. The proposed framework groups and classifies objects by observing how the objects are use. Since function is revealed during usage, this allows the affordance information to be captured, and the grouping and classification results to be more suitable towards handover.

3.3 Computing Proper Grasp Configurations from Observations

During the application phase, when determining the appropriate grasp configuration of an object, after identifying the set of relevant handover grasp configurations, the robot needs to compute a sort of mean of the grasp point and handover orientation. However, computing the average of a set of orientations or rotations is known to be a difficult problem with no standard solution. Rotations in 3D space belong to the SO(3) group, which has many different representations including in $\mathbb{R}^3$ using Euler angles, in $\{ \mathbb{R}, \mathbb{R}^3 \}$ using axis angle representation, or in $\mathbb{R}^{3x3}$ using matrix representation. Calculating mean using any of these representations is not straight forward due to discontinuities (e.g., between 0° and 360°) or that fact the not all members in the set are rotations (e.g., not all $\mathbb{R}^{3x3}$ matrices represent pure rotations). To address this challenge, Chapter 4 will introduce the novel notion of object affordance axes and a distance minimization based method to be used for computing means of handover orientations and determining appropriate handover grasp configurations for objects.

3.4 Discussion

The presented framework for determining proper grasp configurations of objects aims to provide two main advantages over existing methods, which can be summarized as scalability and generalizability. As mentioned in Chapter 2, most existing methods require users to manually provide information for each object in order to determine the proper handover grasp configurations [57], [89], [107]. These methods can be tedious and time consuming when the robot is not limited to handling only a small number of predefined objects. By using a learn from observation approach, users can teach the robot through more natural interaction with the robot. Users can demonstrate the proper grasp configuration to the robot by simply handing over the object to the robot, or to another person, while the robot observes. Furthermore, this approach also has the potential of eliminating the need for explicit teaching. As robots enter the society and become more widespread, they will have more time working around people, and they will have more chances to observe how people handover different objects. By paying attention to these naturally occurring object handovers, the robot can learn handover grasp configurations automatically.

Another advantage of the proposed framework is that it allows the robot to handle new objects by generalizing learned grasp configurations. As the application of service robots widens, they are bound to encounter new objects. However, exiting methods lack the ability to generalize grasp configurations of known objects to new objects. The proposed framework allows generalization by identifying similarities between known objects and new objects. Since proper grasp configuration depends on object affordances, and object affordances are often revealed during usage, the proposed approach identifies similarities between objects based on their observed usages. Compared to geometry based methods, this allows a more accurate grouping and classification of objects for the purpose of handovers.

The presented framework also has the advantage of being hardware independent. In certain existing works, handover configurations are specified using robot specific parameters. In these cases, it may not be as straight forward when applying the learned grasp configurations to another robot. In the proposed framework, the representation chosen for the grasp configuration is free of robot hardware parameters. This makes the results robot independent and thus allows sharing of knowledge among different robots. From an hierarchical design perspective, robot software architecture can be organized into different layers as Figure 3-2 illustrates (c.f. Ueda 2010 [111], Inoue 1984 [112]). The proposed framework for determining proper handover grasp configurations would sit in the lower part of the action layer. When a handover action is required to achieve the goal of the task layer, the proposed framework generates a robot independent grasp configuration to be used. The grasp configuration can then be passed to the robot motion layer to plan hardware specific motions to achieve the grasp configuration. If necessary, the generated grasp configuration can be modified by the robot motion layer and the layers below to account for constraints such as collisions and joint limits.

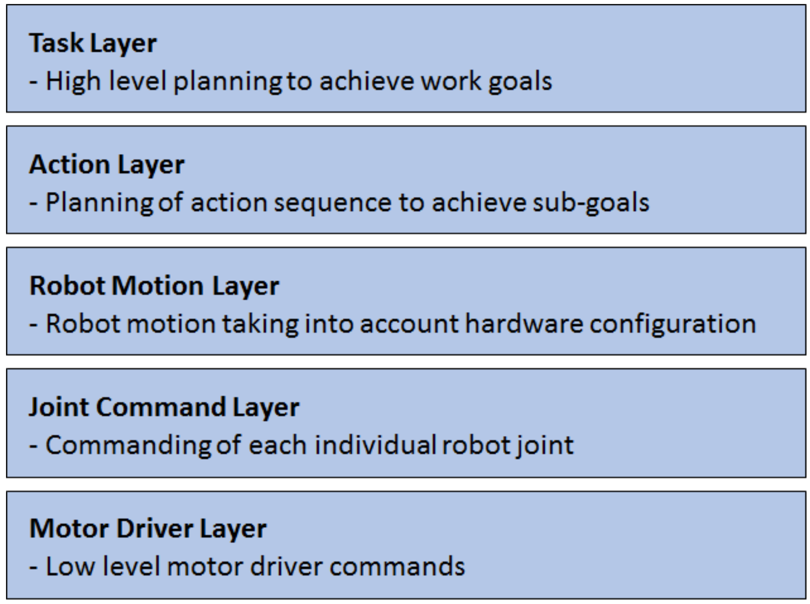

Figure 3-2 Robot software design layers.

While Figure 3-1 shows the learning phase and the application phases as two separate phases, in reality the separation may not be necessarily clear cut. While the robot is in operation and in the application phase, the robot may continue to observe handovers and object usages occurring in the surrounding. It can then remember the observed grasp configurations and usages and add them to its knowledge base, thus continuing to expand it. This allows the robot to continually learn and adapt.

3.5 Summary

This chapter has presented a framework that enables robots to determine proper grasp configurations for handing over objects. The framework aims to address the scalability and generalizability issues found in existing approaches. By enabling robots to learn handover grasp configurations through observations and natural interactions with humans, the proposed approach eliminates the need for users to explicitly specify the proper grasp configuration for each object. By allowing robots to learn about object affordances through observation of their usages, the proposed framework also enables robots to generalize learned grasp configurations to new objects. The proposed framework has the advantage of being hardware independent and can be applied to any robot platforms. Furthermore, it also allows sharing of knowledge among robots. The grasp configurations learned by one robot can be transferred and used by another robot for handing over objects.

In the following, Chapter 4 first presents a user study conducted to investigate and characterize the handover orientations used by humans, and devises a method for identifying patterns and computing mean handover orientations given a set of observations. Chapter 5 presents an affordance focused object grouping and classification method which will allow the generalization of learned handover grasp configurations to new objects. Using these building blocks, Chapter 6 then presents an implementation of the framework onto a robot hardware and the validation experiments.

4 Characterization of Handover Orientations used by Humans and Method for Computing Appropriate Handover Orientations from Observations

Chapter 3 presented a framework that aims to enable robots to automatically determine proper handover grasp configurations of various objects. The framework is based on observation of human handovers and object usages. The robot learns handover grasp configurations by observing the grasp configurations used in human handovers. In doing so, the robot presumes that the grasp configurations used by humans are appropriate. However, currently, a good understanding of the grasp configurations used by humans in natural handovers is still lacking. To first gain a better understanding, this chapter presents a user study conducted to investigate and characterize the grasp configurations used by humans, focusing on the handover orientations.

For identifying patterns in the handover orientations. This chapter puts forth the novel notion of *affordance axes* for capturing the affordance information of objects. A distance minimization method for computing a mean orientation from a set of observed handover orientations is also introduced. Results from this section will then be used in Chapter 6 for the implementation of the framework presented in Chapter 3.

4.1 Objectives

The user study commences with the objectives of surveying the handover orientations humans use for a set of twenty common objects, and to determine whether observable patterns exists in the handover orientations used. The findings are expected to guide the design of human-robot handovers. The driving research questions are:

1. How do people orient the object when handing over various common objects?

2. Do observable patterns exist in the handover orientations people use?

3. When handing over objects, are people naturally considerate of the receiver and use appropriate handover orientations?

In many cooperative scenarios, the purpose of handing over an object is to allow the receiver to utilize the object for completing the work at hand. Thus, for achieving efficient cooperation, one would favor the use of handover orientations that allow the receiver to easily take the object and readily use it once taken.

4.2 Object Handover User Study

4.2.1 Experiment Design

In the user study, participants working in pairs of "giver" and "receiver" handed over twenty everyday objects and common tools while a motion capture system tracked the object's orientation and the participants' motions. Figure 4-1 shows the items used in the study. The aim of this study is to examine the handover orientations people use in natural handovers and determine whether the handover orientations used vary depending on where the giver's focus is placed. Therefore, the study tests three conditions:

Figure 4-1 Everyday objects and common tools participants handed over in the user study, with infrared reflectors affixed. (Published in Chan et al. 2015 [113].)

Condition A: (Natural handovers) For the purpose of measuring natural handover orientations, in this condition, givers were not provided explicit instructions regarding how to hand over the objects.

Condition B: (Giver-centered handovers) For the purpose of measuring the handover orientations used when givers have their focus on themselves, in this condition, givers were asked to hand over the objects in a manner that is the easiest and most convenient to themselves.

Condition C: (Receiver-centered handovers) For the purpose of measuring handover orientations used when givers have their focus placed on the receiver, in this condition, givers were asked to hand over the objects in a manner that is most comfortable and convenient to the receiver, giving consideration to the object's usage and the function of its different parts.

4.2.2 Hypotheses

Since functional objects have specific affordances (such as for grasping, writing, hitting, cutting, etc.) associated to their different parts, this study hypothesizes that when givers have their focus placed on the receiver and consideration for the receiver is provided, givers would use an orientation that presents the handle or grasping part to the receiver, such that specific patterns can be observed in the handover orientations used in Condition C. If, however, the givers have their focus placed on themselves, they might simply use any arbitrary handover orientation convenient at the time, or hold on to the object by the part that is the easiest to grasp. Therefore, the handover orientations used in Condition B and Condition C are expected to differ. In addition, this study expects that in natural handovers, givers will have their focus placed either on themselves or on the receiver, thus the handover orientations observed in Condition A will either be similar to those in Condition B or Condition C. Formally stated, the hypotheses of the study are as follows:

H1: There are observable patterns in the handover orientations used by participants in receiver-centered handovers (Condition C).

H2: There is significant difference between the handover orientations used by participants in giver-centered handovers (Condition B) and the handover orientations used in receiver-centered handovers (Condition C).

H3: Handover orientations used by participants in natural handovers (Condition A) are either similar to handover orientations used in giver-centered handovers (Condition B) or handover orientations in receiver-centered handovers (Condition C).

4.2.3 Motion Capture

A Vicon motion capture system [114] captured the giver's and receiver's motions, as well as the object's position and orientation during the experiment. Reflective markers were attached to the objects and participants were fitted into a jacket and a cap with reflective markers for tracking. The Vicon system captured data at 300 Hz, and post-data collection, the Vicon Nexus software [115] was used for computing the object's orientation and the giver's and receiver's joint angles.

4.2.4 Participant Population

Participants in this study involved twenty healthy adults including nine females and eleven males. Volunteers were recruited through advertisement posted on the webpage of the Collaborative Advanced Robotics and Intelligent Systems Laboratory at the University of British Columbia, mailing lists, and by word of mouth. The age of the participants ranged from 19 to 61 and averaged at 28.2. Participants provided informed consent before beginning the experiment (refer to Appendix A for consent form), and each was given a candy bar as a token of appreciation upon completion of the experiment. Experiments were videotaped with participants' agreements.

4.2.5 Experimental Procedure

Participants cooperated in pairs with one arbitrarily appointed as giver first, and the other receiver. During the experiment, there were three tables placed around the giver, and the receiver stood opposite the giver in the beginning of each handover. Figure 4-2 shows the experimental setup. At the beginning of each handover trial, the object is placed randomly on one of the tables, at a random orientation. Each handover commenced with the experimenter saying "go". The giver picked up the object, handed over to the receiver, and after the receiver took the object, both giver and receiver returned to their starting positions to complete the trial. After handing over all twenty objects in a randomized order, the giver and the receiver switched roles and repeated the procedure with the twenty objects. For consistency, participants were instructed to perform the handovers using the right hand. To cover all three conditions, each pair of participants completed six sets of twenty handovers, totaling at one hundred and twenty handovers.

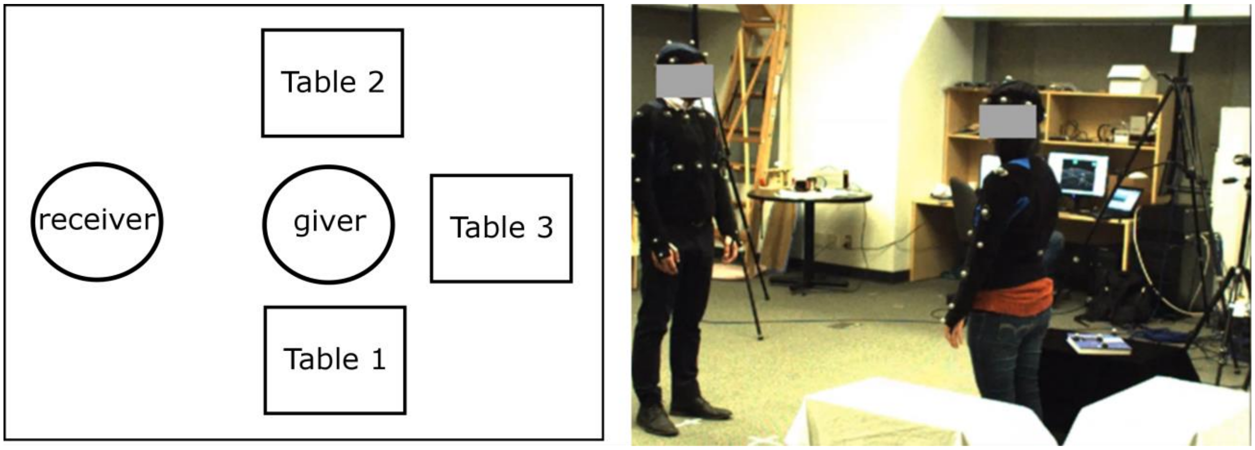

Figure 4-2 Experimental setup of the user study. (Figure partially adopted from Chan et al. [113].)

To be able to observe natural handovers with the participants unperturbed, Condition A was conducted as the first two sets of twenty handovers for each pair of participants. Subsequently, the remaining four sets varied between Condition B and Condition C in a counter-balanced order to account for carryover effects. Between each set of twenty handovers, there were short pauses as the experimenter explained the next condition, and participants were permitted to take additional breaks to prevent fatigue. However, no participants required additional breaks.

4.3 Data Analysis

4.3.1 Handover Orientation Extraction

For extracting the handover orientations from the motion capture data, the giver's and receiver's hand trajectories were first examined. Figure 4-3A demonstrates the typical hand trajectories of the giver and receiver in a handover trial. The bottom plot showing the distance between the giver's the receiver's hands reveal a characteristic trough. The instance of handover is determined by finding when the distance between the giver's and receiver's hands is minimized. The handover orientation is then extracted as the measured object orientation at this time.

Each of the twenty objects used in the study is assigned a coordinate frame arbitrarily as shown in Figure 4-4. To allow the results to be translated to any location in space, the extracted handover orientations are expressed in a base frame relative to the giver and the receiver. Figure 4-5 shows the base frame defined by the giver's and receiver's location. The z-axis of the base frame points in the same direction as the ground surface normal, the x-axis points from the receiver's torso to the giver's torso, and the y-axis is determined accordingly to form a right-handed coordinate frame.

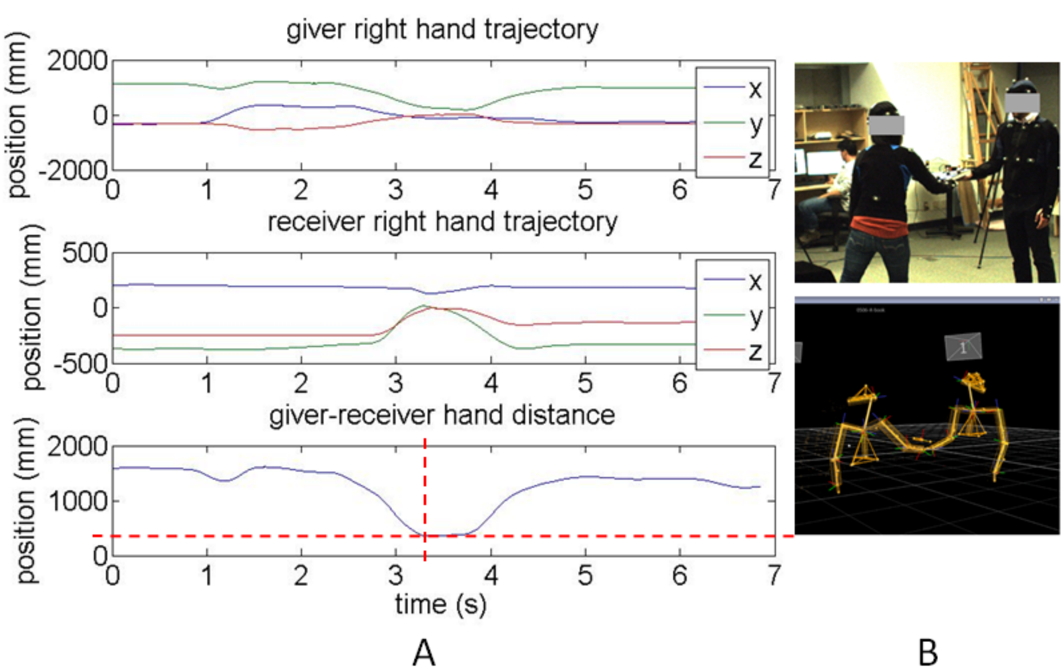

Figure 4-3 A – Typical hand trajectories of giver and receiver. The instant of object transfer is found by locating the minimum distance between giver's and receiver's hands. B – Example object transfer. (Published in Chan et al. 2015 [113].)

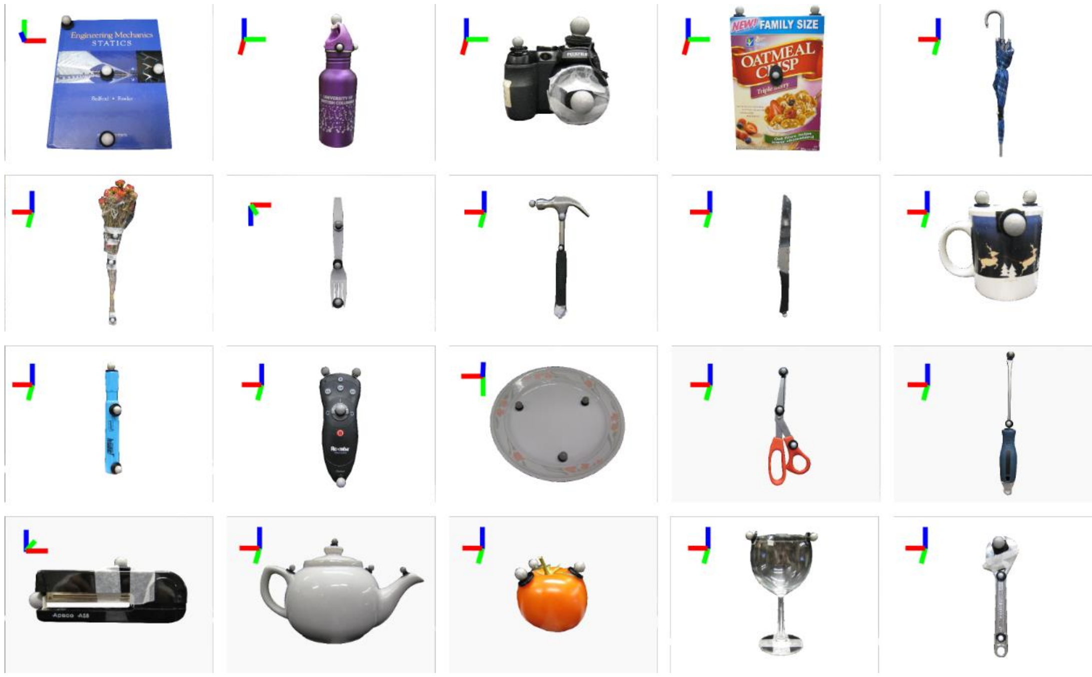

Figure 4-4 Arbitrary coordinate frames assigned to the objects used in the user study. (Published in Chan et al. 2015 [113].)

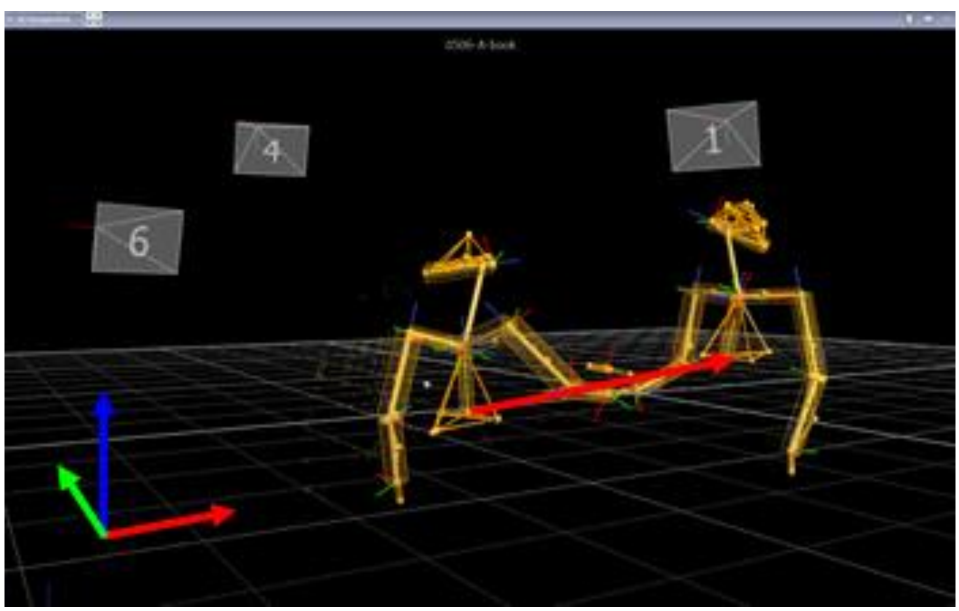

Figure 4-5 Base frame defined by the giver's and receiver's locations. (Published in Chan et al. 2015 [113].)

4.3.2 Affordance Axes

Preliminary Inspection of the data shows that observable patterns exist in the handover orientations used by participants for a majority of the objects. Taking the mug and the hammer as examples, most participants kept the mug in an upright orientation when handing it over, and most participants presented the handle to the receiver when handing over the hammer in Condition C. It appears that when handing over the objects, participants tend to align a certain imaginary axis of the object in a same general direction. While the pattern for some objects are more pronounced, it is less easily observed for some others. Thus, to facilitate the identification of patterns in the measured handover orientations, this thesis introduces the novel ideal of *affordance axes*:

> **Affordance Axis:** Given an object, the *affordance axis* ϕ_{Aff} of the object is an axis in the object's coordinate frame, which people tend to align in the same general direction when handing over the object, while giving consideration to the object's affordances.

Returning to the examples of the mug and the hammer, intuitively, the affordance axis of the mug would be an axis normal to the bottom surface, and the affordance axis of the hammer would align with the handle. To compute the affordance axes of the objects, this thesis defines ϕ_{Aff} mathematically as follows. Suppose that $\hat{\phi}_{Aff}$ is a vector of unit length pointing along ϕ_{Aff}. Then $\hat{\phi}_{Aff}$ can be computed as:

$$\hat{\phi}_{Aff} = \underset{\hat{\phi}}{argmin} \sum_{i} acos(\bar{R}\hat{\phi} \cdot R_{\boldsymbol{i}}\hat{\phi}) \tag{1}$$

where R_i is the handover orientation measured from the ith participant, and $\bar{R}$ is a "mean" orientation (explained in more details in the following section) computed for the object from handover orientations measured in all trials. In Equation (1), $\hat{\phi}$ is a unit vector in the object's frame, and $\bar{R}\hat{\phi}$ and $R_{\boldsymbol{i}}\hat{\phi}$ rotates $\hat{\phi}$ from the mean orientation frame and the ith measured handover orientation frame respectively into the world frame. If $\hat{\phi}$ truly aligns with the affordance axis, then the angle between $\bar{R}\hat{\phi}$ and each $R_{\boldsymbol{i}}\hat{\phi}$ should be small. Therefore, Equation (1) computes $\hat{\phi}_{Aff}$ by finding a $\hat{\phi}$ that minimizes the sum of angles between $\bar{R}\hat{\phi}$ and each $R_{\boldsymbol{i}}\hat{\phi}$. The affordance axes of the objects are computed using receiver-centered handover orientations obtained in Condition C.

4.3.3 Computation of Handover Orientation Means

Given a set of observed handover orientations, for the purpose of allowing a robot giver to determine an appropriate handover orientation to be used, a method for computing a "mean" of the given set of orientations is needed. However, computing the average or mean of a set of orientations is a challenging problem with no standard solution, and there exists few publications that focus specifically on this problem [116]–[118]. Orientations are rotations in space, and rotation in 3D space, often denoted as SO(3), can be expressed in many different representations. Table 4-1 provides a list of representations that can be used to express 3D rotations. All these representations are interconvertible, and there is no one best representation. However, computing the average of a set of rotations simply by summing all elements in the set and dividing by the size of the set using any of these representation will not guarantee good results. The reasons for this is because of the discontinuities in the representations, the non-unique mapping from representation to rotation, or because not all members of the representation are valid rotations. For example, in the Euler angles and axis angle pair representation, both $\theta=0°$ and $\theta=360°$ maps to the same rotation, but taking their average results in a completely different rotation. Even if θ is restricted to take its value from $[0°, 360°)$ to eliminate ambiguity, it still does not resolve the problem of taking the average of $\theta_1 = 1°$ and $\theta_2 = 359°$. With quaternion and matrix representations, because not all members in $\mathbb{R}^4$ are quaternions, and not all members in $\mathbb{R}^{3x3}$ are rotation matrices, by summing a set of rotations and divide by its size, the results is often no longer a valid rotation.

Table 4-1 Representations of 3D rotational group SO(3).

Euler Angles	Axis Angle Pair	Quaternion	Rotation Matrix
(θ,φ,ψ)	(v, θ)	(w, x, y, z)	$\begin{bmatrix} r_{11} & r_{12} & r_{13} \\ r_{21} & r_{22} & r_{23} \\ r_{31} & r_{32} & r_{33} \end{bmatrix}$
$\mathbb{R}^3$	$\{\mathbb{R}^3, \mathbb{R}\}$	$\mathbb{R}^4$	$\mathbb{R}^{3x3}$

In order to compute a mean of a set of handover orientations, a different approach is needed. Considering the alternative definition of an "average" in scalar space, an average $\bar{x}$ of a set of numbers $x_i \in \{x_1, \ldots, x_n\}$ is a number that minimizes the sum of distances from $\bar{x}$ to each x_i. In the scalar case, the distance measure is the L2-norm:

$$\bar{x} = \underset{x'}{argmin} \sum_i (x' - x_i)^2 \tag{2}$$

Using this alternative definition of average, this work calculates the mean handover orientation from a set of observations using a distance minimizing approach:

$$\bar{R} = \underset{R}{\operatorname{argmin}} \sum_i d(R, R_i) \tag{3}$$

where the measure of distance between two orientations R and R_i is chosen to be the angle of rotation between them:

$$d(R, R_i) = acos(\frac{tr(R^{-1}R_i)-1}{2}) \tag{4}$$

with $tr(R^{-1}R_i)$ giving the trace of the rotational matrix $R^{-1}R_i$. Equation (4) minimizes the rotation from the mean to each observed handover orientation. Thus it allows a mean orientation that most similarly resembles the measured handover orientations to be computed.

4.3.4 Patterns in Handover Orientations

To determine if there are observable patterns in the handover orientations used by participants in the receiver-centered handovers (Condition C) in response to hypothesis H1 stated in Section 4.2.2, the experimenter computed the angle between $\bar{R}\hat{\phi}_{Aff}$ and $R_i\hat{\phi}_{Aff}$ (designated as θ_i) for each object and examined the histograms of θ_i. If the histograms of θ_i show distinct modes in the distribution, it would suggest that $\bar{R}\hat{\phi}_{Aff}$ and $R_i\hat{\phi}_{Aff}$ for the objects are aligned in general in a certain orientation, and that there are patterns observable in the measured handover orientations R_i. This would then provide support to H1. However, if the histograms show distributions of θ_i spreading across a wide range of angles, it would suggest that $\boldsymbol{R_i}\hat{\phi}_{Aff}$ is not in general alignment with $\bar{R}\hat{\phi}_{Aff}$ and that $\boldsymbol{R_i}$ is distributed widely across rotational space. Indeed, Monte Carlo simulations confirm this expectation. Figure 4-6 shows the histograms of θ_i from four repetitions of the simulation of twenty random orientations. In Figure 4-6, the values of θ_i spread across a wide range of angles. Thus, if the data from the experiment exhibit a similar distribution, it would support that H1 is false. To examine the distributions of θ_i the experimenter carried out a Kuiper's test to determine if the measured distributions differ significantly from a uniform distribution.

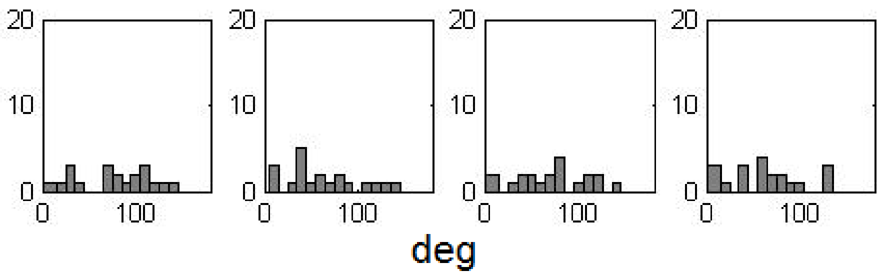

Figure 4-6 Four repetitions of Monte Carlo simulation results showing histograms of $\boldsymbol{\theta_i}$. Twenty random handover orientations were generated for each simulation repetition. A spread of $\boldsymbol{\theta_i}$ among all angles can be seen. (Published in Chan et al. 2015 [113].)

4.3.5 Comparison of Handover Orientations across Conditions

To test the hypotheses H2 and H3 and compare handover orientations used in the different conditions, the experimenter computed the differences between the handover orientations used. The difference in the handover orientations used for two conditions U and V is computed using the following equation:

$$\theta_{i\_UV} = acos(R_{i\_U}\hat{\phi}_{Aff} \cdot R_{i\_V}\hat{\phi}_{Aff}) \quad (5)$$

In Equation (5), R_{i_U} and R_{i_V} are the handover orienations used by the ith participant for Condition U and Condition V respectively. Equation (5), computes the angles between the affordance axes of the object in the handover orientations in Condition U and that in Condition V. Thus, if participants used similar handover orientations in Condition U and in Condition V, the mean $\bar{\theta}_{UV}$ should be smaller than some natural variance δ. As an estimation of δ, the experimenter used the average spread measured in the affordance axes in Condition C:

$$\delta_C = \frac{\sum_i acos(\bar{R}_C\hat{\phi}_{Aff} \cdot R_{i\_C}\hat{\phi}_{Aff})}{n} \quad (6)$$

with n being the number of total participants. To determine if handover orientations differed between conditions, the experimenter conducted t-tests to examine if $\bar{\theta}_{UV}$ is significantly larger than δ_C. An α level of 0.05 is used for determining statistical significance with the use of Bonferroni correction.

4.3.6 Variation in Measured Handover Orientations

One of the objectives of this user study is to determine whether handover orientations used in natural handovers are receiver-centered. If natural handover orientations are found to differ from receiver-centered orientations (which the results section will show is the case), then a robot will need to consider the quality of the observed handover demonstrations when trying to learn handover grasp configurations from observations. If a set of observed handover orientations contains random orientations and those that differ from receiver-centered orientations, then one can expect the variation of the observed handover orientations to be larger. Thus, a measure of variation of a set of handover orientations may provide an indication of the quality of the set, and could be used for guiding the robot when learning handover grasp configurations from observations. For this reason, the experimenter also computed and inspected the average spread of affordance axes, δ, in each condition, and compared the values among the three conditions using t-tests. An α level of 0.05 is used with the Bonferroni correction for determining statistical significance.

4.4 Results

4.4.1 Handover Orientation and Affordance Axis

To visualize the extracted handover orientations, the experimenter plotted the data in 3D plots. Figure 4-7A shows the extracted handover orientation for the teapot from one handover trial in Condition C. The plot shows that in this trial, the giver handed over the teapot upright, with the handle oriented towards the receiver, slightly to the right. Figure 4-7B plots the handover orientations extracted from all trials in Condition C, with the computed mean $\bar{R}$ shown in bold lines, and the computed affordance axis ϕ_{Aff} shown in a long thin line relative to the mean frame. For visualizing and comparing across different conditions, Figure 4-8 plots the extracted handover orientations from all trials, the

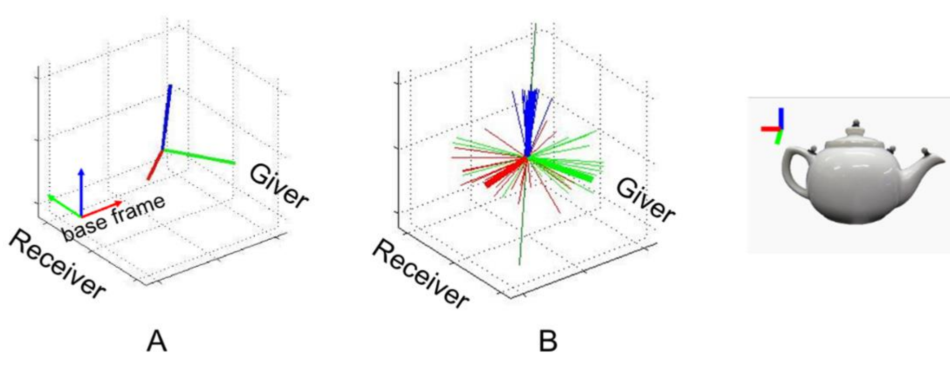

Figure 4-7 A – Teapot handover orientation, Condition C, one trial. Red, green, blue lines show x, y ,z axis respectively. B – Teapot handover orientations, Condition C, all trials. The thick red, green, and blue lines show computed mean $\bar{R}$, long thin line shows computed affordance axis ϕ_{Aff} in mean handover orientation frame $\bar{R}$. (Published in Chan et al. 2015 [113].)

computed mean orientations, and the computed affordances axes for all object in all three conditions. The computed mean orientations are provided in Appendix B.

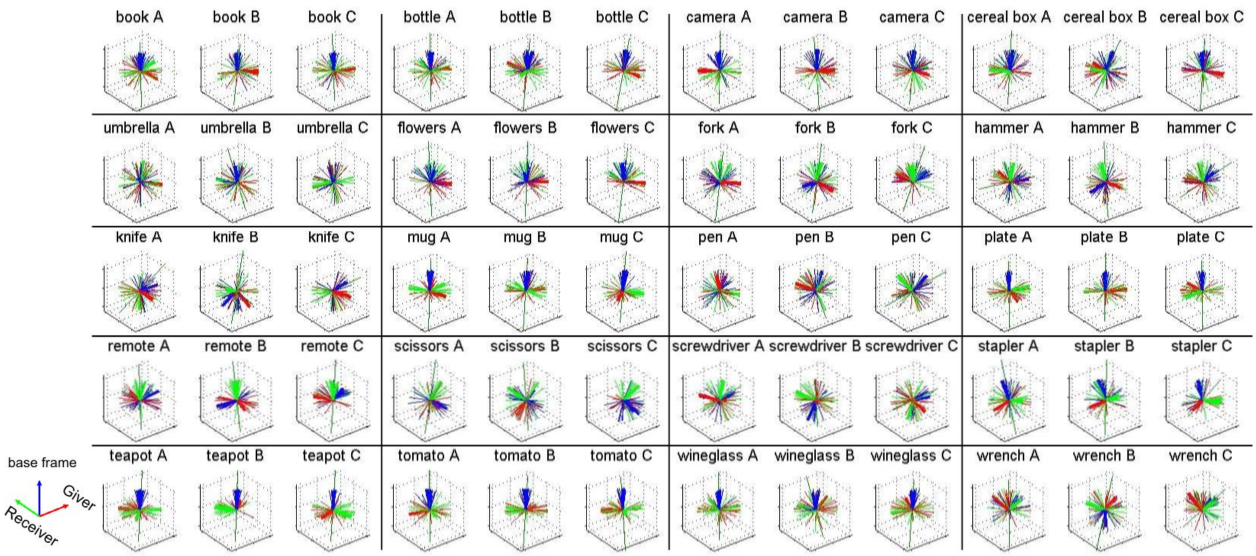

Figure 4-8 Handover orientations for all objects, all trials. Bold coordinate frames show mean orientations. Long thin lines show computed affordance axes. (Published in Chan et al. 2015 [113].)

4.4.2 Patterns in Handover Orientations

Visual inspection of Figure 4-8 reveals that for some objects, there are apparent patterns in the handover orientations. For example, most measured orientations for the mug and the teapot have the z-axis aligned roughly upright, indicating that participants handed over these object mostly upright. Indeed the affordance axis for these objects are found to roughly coincide with the z-axis. To help recognize if patterns exist in the handover orientations of the other objects as well, the experimenter plotted the histograms of θ_i in Condition C for all objects. Figure 4-9 shows the plotted histograms for all objects. Kuiper's test results reveal that the distributions of θ_i for all objects differ significantly from a uniform distribution. The p values for the knife and pen were 0.008 and 0.009 respectively, and the p values for all other objects were less than 0.005.

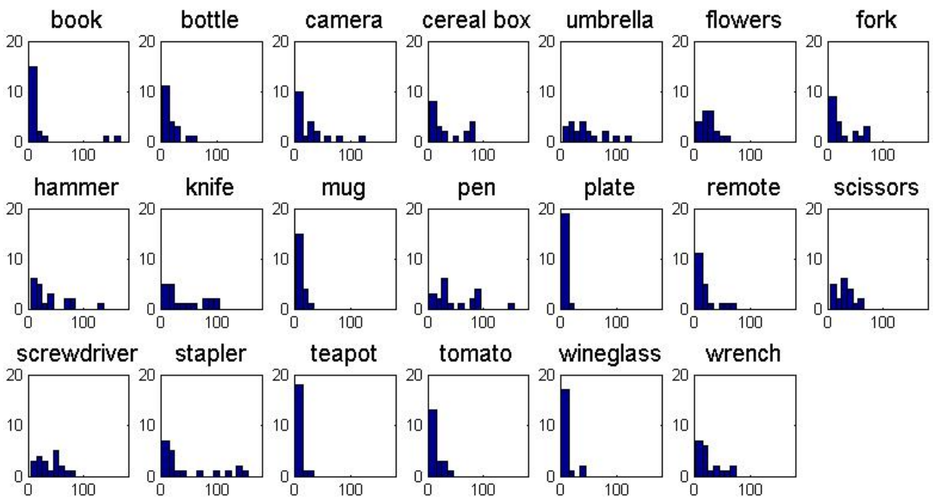

Figure 4-9 Histograms of the angles between $\overline{R}\widehat{\boldsymbol{\phi}}_{Aff}$ and $R_i\widehat{\boldsymbol{\phi}}_{Aff}$ (i.e., θ_i) for each object in Condition C. (Published in Chan et al. 2015 [113].)

4.4.3 Comparison of Handover Orientations across Conditions

Table 4-2 provides the means and standard deviations computed for $\bar{\theta}_{AB}$, $\bar{\theta}_{AC}$, and $\bar{\theta}_{BC}$, along with the t-test results, highlighting the statistically significant results. Table 4-2 shows that comparison of handover orientations among conditions yielded object dependent results. For the book, mug, and plate, comparison between any two conditions did not yield significant results. For the bottle, camera, fork, hammer, pen, remote, scissors, screwdriver, wineglass, and wrench, comparisons showed significant difference across all conditions. As for the rest of the objects, comparison results were mixed.

Table 4-2 Comparison results of handover orientations among conditions. Significant results from t-test are highlighted. (Published in Chan et al. 2015 [107].)

		book	bottle	camera	cereal box	umbrella	flowers	fork	hammer	knife	mug	pen	plate	remote	scissors	screw-driver	stapler	teapot	tomato	wine-glass	wrench
	δ_C (deg)	22.7	18.3	28.3	35.7	42.4	26.0	27.9	36.7	42.3	11.1	50.1	6.5	18.6	29.8	38.3	47.5	7.9	13.5	11.3	27.4
$\bar{\theta}_{AB}$ (deg)	Mean	34.4	32.6	53.7	65.8	81.9	39.5	52.7	99.1	72.2	19.9	82.0	13.8	31.6	74.4	78.0	68.8	11.3	31.2	33.8	88.4
	SD	41.8	24.5	39.5	42.5	53.4	28.3	39.2	58.5	45.1	28.3	51.2	23.0	18.2	42.3	60.9	53.0	8.7	22.7	34.8	44.9
	p val	0.112	0.008	0.005	0.003	0.002	0.023	0.005	0.000	0.004	0.090	0.006	0.084	0.002	0.000	0.004	0.045	0.047	0.001	0.005	0.000
	t(19)	1.25	2.62	2.87	3.17	3.31	2.13	2.83	4.76	2.96	1.39	2.79	1.43	3.21	4.72	2.91	1.79	1.76	3.48	2.88	6.08
$\bar{\theta}_{AC}$ (deg)	Mean	28.1	30.8	46.9	40.4	71.7	42.3	47.5	67.8	54.4	12.6	94.8	8.4	40.8	56.0	58.6	96.7	10.8	20.3	20.9	65.9
	SD	41.0	21.2	26.4	26.6	52.7	22.5	25.7	57.2	50.3	8.2	50.8	5.5	39.7	42.0	37.8	56.4	7.2	14.0	16.9	54.2
	p val	0.283	0.008	0.003	0.217	0.011	0.002	0.001	0.013	0.148	0.218	0.000	0.071	0.011	0.006	0.013	0.000	0.043	0.021	0.010	0.002
	t(19)	0.59	2.62	3.15	0.80	2.49	3.23	3.41	2.43	1.07	0.79	3.93	1.53	2.51	2.79	2.40	3.90	1.81	2.18	2.52	3.17
$\bar{\theta}_{BC}$ (deg)	Mean	46.4	35.8	46.3	62.2	65.2	53.0	50.8	95.2	88.3	19.2	90.8	12.5	35.1	74.8	79.0	71.8	10.7	26.3	32.6	97.8
	SD	53.1	28.6	33.8	39.5	46.1	31.8	39.8	58.1	61.8	23.8	55.9	21.0	32.1	44.9	58.2	59.6	3.9	20.6	31.6	59.3
	p val	0.030	0.007	0.014	0.004	0.020	0.001	0.009	0.000	0.002	0.072	0.002	0.107	0.016	0.000	0.003	0.042	0.002	0.006	0.004	0.000
	t(19)	2.00	2.73	2.38	3.00	2.22	3.79	2.57	4.50	3.33	1.52	3.25	1.29	2.31	4.48	3.13	1.82	3.20	2.78	3.02	5.32

4.4.4 Variations in Measured Handover Orientations

Table 4-3 gives the computed averages of the spread of affordance axes in each condition, δ. T-test results revealed that the average spread of affordance axes in all conditions differ significantly (δ_A vs δ_B, t-val = -2.313, p-val = 0.016, df = 19; δ_A vs δ_C, t-val = 3.853, p-val = 0.001, df = 19; δ_B vs δ_C, t-val = 5.694, p-val < 0.0005, df = 19). The spread of affordance axes in Condition C is found to be significantly smaller than those in Condition A and Condition B.

Table 4-3 Computed averages, δ, and standard deviations, σ, in the spread of affordance axes in each condition.

	book	bottle	camera	cereal box	umbrella	flowers	fork	hammer	knife	mug
δ_A (deg)	13.6	20.8	35.0	32.8	58.2	32.3	35.4	51.7	53.5	8.7
δ_B (deg)	29.5	31.3	34.1	48.4	52.3	39.2	39.1	51.2	56.4	17.7
δ_C (deg)	22.7	18.3	28.3	35.7	42.4	26.0	27.9	36.7	42.3	11.1

	Pen	plate	remote	scissors	screw-driver	stapler	teapot	tomato	wineglass	wrench	average	standard deviation
δ_A (deg)	68.2	5.1	30.5	40.1	50.1	52.6	7.1	17.2	14.6	45.9	33.7	18.8
δ_B (deg)	57.0	11.9	21.9	47.4	60.6	43.7	8.1	21.5	26.0	64.3	38.1	16.7
δ_C (deg)	50.1	6.5	18.6	29.8	38.3	47.5	7.9	13.5	11.3	27.4	27.1	13.3

4.5 Discussion

4.5.1 Patterns in Handover Orientation

Comparing Figure 4-6 with Figure 4-9, distributions computed from the handover orientations measured in Condition C clearly differs from the distributions computed from randomly generated handover orientations in Section 4.3.4. Particularly, the histograms for the book, bottle, camera, mug, plate, remote, teapot, tomato, and wineglass show a concentration around 0°, with 50% or more counts falling in the first bin of 0° < θ_i < 10°. For the remainder of the objects, other than the pen, although the degree of concentration towards 0° is less pronounced, the histograms still show the distributions being skewed towards 0°. These results suggest when handing over the objects, the givers do have a tendency of aligning the object affordance axis in the same direction, thus supporting the claim in H1. In identifying whether patterns exist in the handover orientations used by participants, this study used a Kuiper's test to compare the measured θ_i distributions with a uniform distribution. However, whether a uniform distribution is the best model for the distribution of random orientations is open for debate. The distribution of truly random orientations is not known exactly, and object symmetry may also affect the distribution. Nevertheless, the patterns in the distributions of objects such as the book, mug, plate, and teapot shown in Figure 4-9 is evident.

4.5.2 Comparison of Handover Orientations across Conditions

Analysis results for $\bar{\theta}_{BC}$ provided in Table 4-2 comparing the handover orientations used in Condition B and Condition C show significance for most of the objects, include the bottle, camera, cereal box, flowers, fork, hammer, knife, remote, scissors, screwdriver, teapot, tomato, wineglass, and wrench. This indicates that when handing over these objects, participants used different handover orientations depending on whether they had their focus on themselves or on the receiver. Upon inspection of the experiment videos, it can be confirmed that participants indeed do orient the objects differently in Condition B and in Condition C. For example when handing over the bottle in Condition B, participants orient it generally upright, but with a wider variance in the tilt direction. In Condition C, participants also orient the bottle generally upright, but there is less variance in the tilt direction, and most participants tilted the bottom part towards the receiver. Similarly, when handing over the hammer and the wrench in Condition B, participants tend to grasp onto the handle, but in Condition C, they tend to present the handle part to the receiver instead. Observing the computed mean orientations for these objects, it can be seen that the computed means do capture these characteristics of the handover orientations of the bottle, hammer, wrench, and for the other objects as well. Thus, this shows that the proposed method for computing handover orientation means is a suitable method, and that the means computed from handover orientations measured in Condition C can potentially be taught to robot givers for enabling receiver-centered handovers. For the objects mentioned above, results show that the handover orientations used by human givers vary depending on where their focus is placed, thus providing support to hypothesis H2.

Although t-test results did not show significance for the mug, the experimenter did observe that in more than 50% of the handovers in Condition B, the giver picked up the mug by its handle or rim and did not orient the handle towards the receiver. However, in Condition C, givers presented the handle to the receiver more than 75% of the time. These characteristics are indeed captured by the orientation means computed for Condition B and Condition C. The t-test result simply indicates that in both conditions, the affordance axis of the mug (computed to run roughly along the z-axis) is kept in the same general direction, and that the mug is kept upright. This is perhaps due to the fact that if the mug were not empty, orienting it otherwise would risk spilling its contents, thus revealing that safety is of higher priority than both the giver's and receiver's grasp comfort.

Examining $\bar{\theta}_{AB}$ and $\bar{\theta}_{AC}$ in Table 4-2 to compare the handover orientations used in Condition A with those used in Condition B and in Condition C, results show that the handover orientations used did not differ significantly between Condition A and Condition B, and between Condition A and Condition C for the book, mug, plate, and teapot. This provides support to H3 with respect to these objects, suggesting that a similar handover orientations is used for Condition A and Condition B or Condition C. Mug, plate, and teapot are all containers with the function of holding/carrying things, and there is an associated risk of spilling if they were holding contents. Therefore, in all conditions, participants handed over them upright in more than 85% of the time. As for the book, since people generally read beginning from the cover, givers handed it over with the cover facing up most of the time.

Examining the results for the cereal box, knife, and tomato, t-tests show that the handover orientations givers used in Condition A were similar to those in Condition C, but differ from those in Condition B. When handing over a knife, there is a potential risk of injuring the receiver if an improper handover orientation is used. Therefore, givers tend to point the knife tip away from the receiver in Condition C, and they naturally do similarly in Condition A. Only in Condition B when participants were given explicit instructions to focus on their own convenience and comfort did they use a different handover orientation.

Results indicate that for the bottle, camera, umbrella, fork, hammer, remote, scissors, screwdrivers, wineglass, and wrench, participants used a different handover orientation in Condition A in comparison to Condition B and Condition C. It seems that for some of these objects including the hammer and wrench, handover orientations observed in Condition A and Condition B are quite widely distributed. This perhaps indicates that these objects lack affordance characteristics strong enough to prompt the participants to use any specific handover orientations. On the other hand, for a few other objects such as the remote, it seems that the handover orientations used in Condition A consisted a mix of orientations from Condition B and Condition C. It appears that for such objects, givers naturally use either a giver-centered or a receiver-centered handover orientation, depending on the individual.

4.5.3 Variation in Measured Handover Orientations

Comparing handover orientations across conditions showed that handover orientations used in natural handovers may differ from those used in receiver-centered handovers. Thus, when learning handover orientations from observations, a robot needs to be able to distinguish between the two. Observation suggests that there is a larger variety in the handover orientations used in Condition A. indeed t-tests revealed that δ_C is significantly smaller that δ_A and δ_B. Thus, the computed δ of a set of observed handover orientations may help distinguish sets consisting of handover orientations that are receiver-centered, from sets that contain orientations that are not.

4.5.4 Affordance Axis

As previously mentioned, observations revealed that when handing over objects in Condition C, people tend to align an axis of the object in a same general direction. For example, the axis through the planar surface of the plate is aligned upright, and the axis along the handle of the hammer is aligned towards the receiver. Where this axis of the object is and how this axis is align during handover depends on the affordances of the object, for example, a plate carries food with the flat surface facing up, and a hammer is used by grasping the handle, and the affordance axis aims to capture such information. Compared to other axes computed based on object physical properties, such as the principal axes, while for some objects, the affordance axis may coincide with the major principal axis, for other objects, such as the hammer used in the experiment or an L-shaped drill, the principal axis is misaligned with the handle. Thus, the affordance axis more accurately captures the axis of the object that is important in handovers.

4.5.5 Implications towards Building Intelligent Robots

With regards to the greater goal of enabling robots to learn proper handover grasp configurations from observations of natural handovers, the user study presented in this chapter has offered a better understanding of the nature of hand orientations used by humans in natural handovers. This chapter has also introduced the novel notion of object affordance axes for identifying patterns in handover orientations. Data analysis showed that there are patterns in the receiver-centered handover orientations used by humans, and that natural handover orientations may differ from receiver-centered handover orientations. Thus a robot may need to distinguish between the two when learning handover orientations by observing human handovers. Results suggest that natural handover orientations and giver-centered handover orientations have higher variations than receiver-centered handover orientations. Thus the measured spread in affordance axes of a set of observed handover orientations may be used to determine the quality of the set.

This chapter presented an optimization based method for computing mean handover orientations, which has been tested and shown to capture the characteristics of various handover orientations well. A robot giver can use this method to compute the proper handover orientation of objects based on multiple demonstrations it has observed. Furthermore, the receiver-centered handover orientations computed for the twenty objects in the user study can potentially be used by robots for handing over these objects to facilitate more efficient and socially acceptable interactions with people.

4.6 Summary

This chapter has introduced the novel notion of object affordance axes, and a distance minimization based method for computing the mean of a set of observed handover orientations. A user study surveyed the handover orientations used by human givers in three different conditions for a set of common objects, and computed the mean handover orientations for the objects. Mean orientations computed from the receiver-centered handovers can potentially be used by robot givers for handing over these objects to allow more effective cooperation. Comparison of the three different handover conditions revealed that natural handover orientations are not necessarily receiver-centered. Thus, a robot may need to consider the quality of the observed handover orientations when trying to learn from observing human handovers. Results from this chapter will be used in Chapter 6 to implement and realize the framework presented in Chapter 3 for enabling robots to determine proper handover grasp configuration for objects automatically.

5 An Object Grouping and Classification Method based on Observation of Object Movement Patterns and Inter-object Interactions during Usage for Generalization of Handover Grasp Configurations

Chapter 4 presented a user study that investigated and characterized the handover orientations used by humans and presented a method for computing a mean handover orientation from a set of observed orientations. These results will be used toward learning handover grasp configurations from observations in Chapter 6 when implementing the framework Chapter 3 presented. To be able to generalize the learned handover orientations for handing over new objects, this Chapter presents an object grouping and classification method that focuses on object affordance.

5.1 A Usage-based Method for Grouping and Classifying Objects

As discussed in Chapter 2, the proper handover grasp configuration for an object depends on many factors including the object's physical properties, its function and affordances, the receiver's state, and even social conventions. Object function or object affordances especially have a heavy influence on determining what is considered as an appropriate handover grasp configuration for an object, since object affordance determines how an object can be grasp, how it is meant to use used, and how it can be functional. Thus, the object grouping and classification method presented in this chapter focuses on objects' function and affordances. With respect to the framework for enabling robots to determine proper handover grasp configurations shown in Figure 3-1, the work this chapter presents relates to building the knowledge base in the learning phase, and determining handover grasp configurations in the application phase.

5.1.1 Extracting Inter-object Interactions from Object Usage Demonstrations

Proper handover grasp configuration is dependent on object affordances, and object affordance information if often reflected in the object's movements during usage, and its interactions with the surrounding. For example, the function of a hammer is for hamming nails, and using a hammer involves grasping the hammer by the handle, swinging it in a swift motion, and contacting the nail with the head of the hammer quickly. Likewise, the function of a kitchen knife is for cutting, and using a kitchen knife to cut tomatoes involves contacting the tomato with the sharp edge of the knife, moving the knife in a sliding motion, and moving the blade into the tomato. Thus, the object grouping and classification method this chapter presents extracts object description features from usage demonstrations by observing the movements of the object, and its interactions with the environment.

Figure 5-1 illustrates the process of extracting object description features from a demonstration video. Given an object O, the state of the object, x_O, is defined by the position p_O, and the rotation r_O of the object. Thus, given a demonstration video of duration T, the object O is described by a sequence of object states $x_O(t)$, with $t \in [0, T]$. For a demonstration with multiple objects $O1$, $O2$, $O3$,… appearing, the data contains the object state sequences all the objects, $x_{O1}(t)$, $x_{O2}(t)$, $x_{O3}(t)$,…. Objects of interest that are tracked include the demonstrated object itself, the objects in the environment, and the demonstrator's hands.

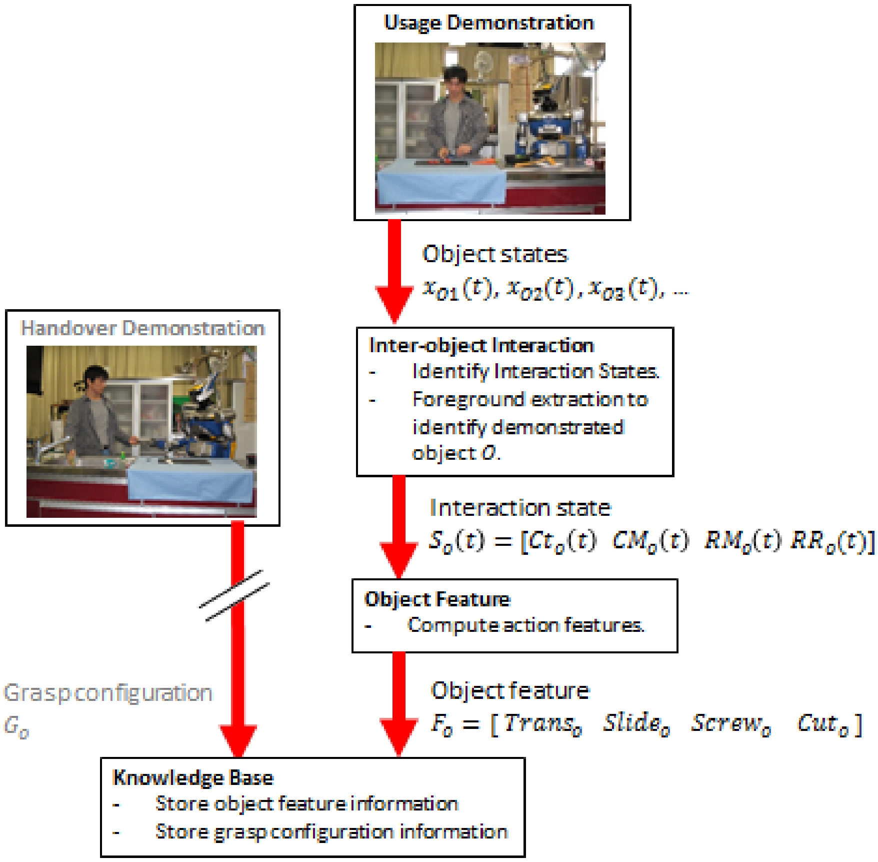

Figure 5-1 Procedure for extracting object features from usage demonstrations. (Published in Chan et al. 2014 [110].)

The algorithm for extracting object features first identifies the inter-object interactions from the object state data. This work defines four inter-object-interaction types:

1) **Contact**: Most objects function by bringing about actions to other objects. And to affect other objects, physical interaction, thus contact, is required.

2) **Co-movement**: Many objects such as plates, trays, and cups are used to hold and transport other objects. For such objects, there will be co-movement observed during transportation.

3) **Relative movement**: This type of interaction captures function of objects such as pen, knife, computer mouse, where usage of the object involves sliding it on another surface or object.

4) **Relative rotation**: Similar to relative movement, this captures another common class of object functions such as that of screwdriver, wrench, and cork opener

Let the Boolean variables Ct, CM, RM, RR, denote respectively the four inter-object interactions contact, co-movement, relative movement, and relative rotation. For a pair of objects On and Om, the four variables are computed as follows:

$$Ct = (||p_{On} - p_{Om}|| < \alpha_1) \tag{7}$$

$$CM = (||\dot{p_{On}} - \dot{p_{Om}}|| < \alpha_2) \wedge (||\dot{p_{On}}|| > \alpha_3) \tag{8}$$

$$RM = Ct \wedge (||\dot{p_{On}} - \dot{p_{Om}}|| > \alpha_2) \tag{9}$$

$$RR = Ct \wedge \left(|\omega_{On}^{Om}| > \alpha_4\right) \tag{10}$$

where ω_{On}^{Om} denotes the relative angular velocity between the objects On and Om, and α_1, α_2, α_3, and α_4 are constant thresholds. At each frame t' of the demonstration video, the algorithm computes the interaction state of the object O, using the above four equations, as:

$$S_o(t') = [Ct_o(t')\ \ CM_o(t')\ \ RM_o(t')\ RR_o(t')] \tag{11}$$

To allow the robot to focus its attention on relevant objects and filter out background objects, the algorithm extracts foreground event by performing the following. The robot first tracks the tool manipulation hand, which this study designates as the right hand, of the demonstrator. When the robot sees the demonstrator's hand come in contact with an object O in the surrounding, the robot identifies O as the object being demonstrated. Once the robot identifies the demonstrated object O, it tracks the object O and the objects in the environment O contacts. The interaction states sequence $S_o(t)$ is then computed for the object O by considering O and the objects it comes in contact with during the demonstration.

5.1.2 Computing Action Features for Describing Objects base on Usage

After extracting the sequence of interaction states $S_o(t)$, the algorithm then extracts a feature vector from $S_o(t)$ to describe the object. This algorithm defines four action features: Transport, Slide, Screw, and Cut. Let the variables $Trans$, $Slide$, $Screw$, Cut represent each of the four action features respectively. The variables $Trans$, $Slide$, $Screw$ are calculated as:

$$Trans = \frac{\sum_{t=0}^{T} CM_o(t)}{T} \tag{12}$$

$$Slide = \frac{\sum_{t=0}^{T} RM_o(t)}{T} \tag{13}$$

$$Screw = \frac{\sum_{t=0}^{T} RR_o(t)}{T} \tag{14}$$

For evaluating the feature variable Cut, the algorithm first defines two Boolean functions, $prior\_contact(O1, O2, t')$, and $disappear(O, t')$. The function $prior\_contact(O1, O2, t')$ returns true only if the two objects $O1$ and $O2$ were in contact prior to t' for a time window of 1 sec duration. The function $disappear(O, t')$ returns true only if the object O becomes no longer observable for $t > t'$. In the algorithm, when and object O splits into multiple parts at time t', its state data $x_O(t)$ becomes no longer available after time t', thus indicating that the object O is no longer observable. Using the above two Boolean functions, the feature variable Cut is evaluated as:

$$Cut = \begin{cases} 1, & \exists\ O', t'\ s.t. \\ & \quad prior\_contact(O, O', t')\ \wedge \\ & \quad disappear(O', t') \\ 0, & otherwise \end{cases} \tag{15}$$

Existing works on identifying object affordances have utilized object features including object location, object displacement, and inter-object position [119]; inter-object relations of being disconnected, in contact, and surrounding [120]; as well as specific predefined actions such as hammering and pouring [121]. This work has chosen its proposed object features to be actions that are general and permit intuitive interpretation. The feature variable $Trans$ would be indicative of affordances such as a plate carrying and transporting objects. A high value of $Silde$ would be observed for a piece of cloth cleaning the table or a pen writing on a piece of paper. A screwdriver or a wrench tightening a screw or bolt would give rise to a high $Screw$ value. Cut would capture the function of a kitchen knife cutting vegetables or fruits.

The robot extracts a features vector F_o to describe the demonstrated object O from each usage demonstration:

$$F_o = [\, Trans_o \quad Slide_o \quad Screw_o \quad Cut_o \,] \tag{16}$$

It then stores the entry $\{O, F_o\}$ into its knowledge base. Note that the robot can also handle multiple demonstrations for the same object. It would simply store the different F_o extracted from each demonstration as separate entries.

5.1.3 Building a Knowledge Base of Handover Grasp Configurations

As Figure 5-1 illustrates, the knowledge base includes observed usage information captured by F_o, stored as $\{O, F_o\}$ pairs, and observed handover grasp information expressed as G_o, stored as $\{O, G_o\}$ pairs. In the full implementation of the framework G_o will be obtained from observed handover demonstrations. For the experiment documented in this chapter, the focus is on the object grouping and classification method, and the robot uses pre-specified grasp configurations.

Upon observing the demonstrations for a set of objects and learning the object features and handover grasp configurations, the robot's knowledge base would be populated with the entries $\{O1, F_{O1}\}$, $\{O2, F_{O2}\}$, $\{O3, F_{O3}\}$, …$\{O1, G_{O1}\}$, $\{O2, G_{O2}\}$, $\{O3, G_{O3}\}$, …. For organizing its knowledge base, the robot groups together entries with similar object features. This is accomplished using k-means with initial seeding to cluster the entries according to the object features F. The algorithm first performs normalization of the feature vectors. For evaluating the distance between the feature vectors F_{O1} and F_{O2}, this experiment tested two distance measurements. The first measurement is the Euclidean distance between the two vectors, and the second measurements is the angle between the two vectors given by:

$$\Delta = \cos^{-1} \frac{F_{O1} * F_{O2}}{||F_{O1}||\ ||F_{O2}||} \tag{17}$$

Results, however, will show that both measurements yield the same grouping of objects in the knowledge base and classification outcomes. After grouping objects together in the knowledge base, the robot obtains the centroids of the group clusters C_1, C_2, C_3, …, C_n.

5.1.4 Classifying New Objects and Determining Handover Grasp Configurations

When the robot sees a new object O', the robot first observes the usage of the object and extracts the feature vector, $F_{O'}$, of the object following the same algorithm presented in Section 5.1.1 and Section 5.1.2. When the robot is asked to hand over the new object O', it first uses the nearest neighbour algorithm to classify O' into one of the clusters C_1, C_2, C_3, …, C_n in its knowledge base according to the distances Δ_i computed between $F_{O'}$ and each C_i. The new object O' is classified into the cluster giving the smallest Δ_i. After classification of O', the grasp configurations $G_{O'}$ for the new object is then computed based on the handover grasp configurations found in the knowledge base belonging to that cluster.

5.2 An Experiment on Generalizing Grasp Configurations to New Objects

This section presents the experiment conducted to test the object grouping and classification method proposed above. In the experiment, the robot first observes the usages of a set of different objects, and groups the objects together based on their affordances. Subsequently, when given a set of new objects, the robot then classifies them based on similarities in observed usages.

5.2.1 Experimental Procedure

The experimenter first provided object usage demonstrations by showing how the different objects are used, while a NaturalPoint OptiTrack system [122] tracked the objects and the demonstrator's hands. The tracking system included six V100:R2 infrared cameras [123]. The objects and the demonstrator's hands were marked with infrared reflectors for tracking. The OptiTrack system recorded data at approximately 30 Hz, and the NaturalPoint Tracking Tools software was used for computing the object states $x_O(t)$. For computing contact between objects, the algorithm uses spherical collision geometry defined for each object.

5.2.2 Training Set

The experimenter provide usage demonstrations for the ten objects shown in Figure 5-2 as the training set to test the proposed grouping method. The training set contains common objects including kitchen knives, screwdrivers, plates, cutting boards, and parker pens. For each object, the experimenter provided one to four different usage demonstrations. The demonstrations contained different usage examples. Kitchen knife usage included slicing of different objects demonstrated at different angles. Screwdriver demonstrations included tightening of screws on various objects held in the demonstrator's hand or placed on the table at different orientations. Demonstrations of the marker pens included drawing and writing. Demonstrations of cutting board and plate showed carrying and transporting of various items with different initial and final locations. The experimenter performed a total of twenty usage demonstrations for building the knowledge base.

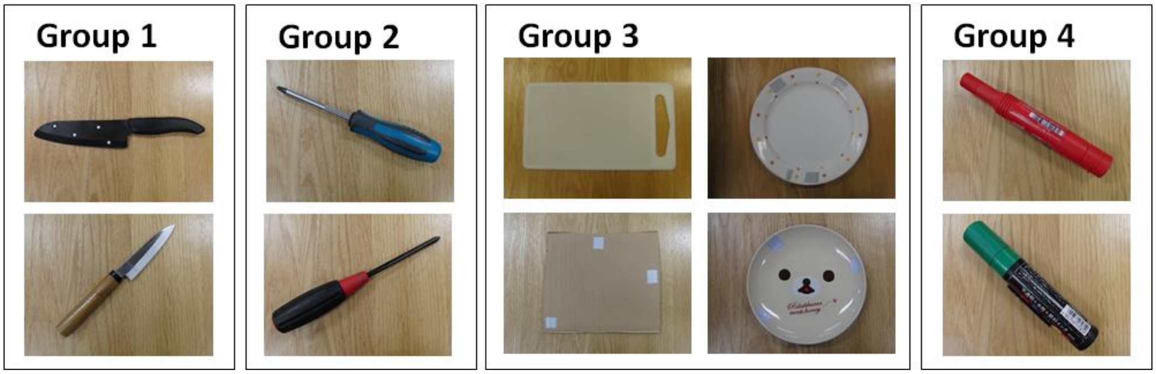

Figure 5-2 Clustering results from k-means according to observed object usage. (Published in Chan et al. 2014 [110].)

Figure 5-3 shows the grasp configurations defined in the experiment. Table 5-1 gives the description of each handover grasp configuration and lists the objects assigned with each grasp configuration. Grasp configurations for each object were chosen taking into consideration the affordances of the objects as well as the receiver's grasp comfort.

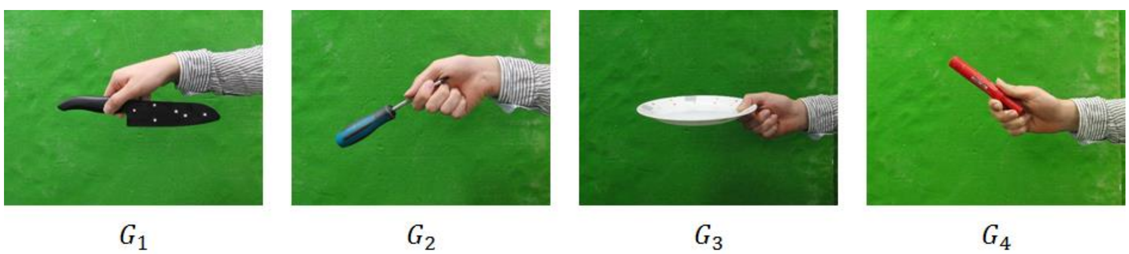

Figure 5-3 Four discrete grasp configurations defined for handing over different types of objects. (Published in Chan et al. 2014 [110].)

Table 5-1 Description of grasp configurations specified in the experiment. (Published in Chan et al. 2014 [110].)

Grasp Configuration	Associated Objects	Grasp Location	Orientation
G_1	Knifes	Knife handle, close to blade	Handle towards receiver, parallel to ground
G_2	Screwdrivers	Shaft of screwdriver (shank)	Handle pointed 30° downwards toward receiver
G_3	Plates, Cutting Boards	Edge of the object	Surface parallel to ground
G_4	Marker Pens	Close to writing tip	Non-writing end pointed 30° upwards towards receiver

5.2.3 Evaluation Set

To evaluate the proposed method's ability of classifying new objects, the experiment used three new objects that have not appeared in the training set. Figure 5-4 shows the three new objects, including a screwdriver, a marker pen, and a cutter. The experimenter demonstrated tightening of a screw with the screwdriver. For the marker pen, the experimenter demonstrated the usage of colouring. Colouring is a new usage that has not appeared in the training set. The screwdriver tests the proposed method's ability of classifying new instances of known types of objects, while the marker pen tests the proposed method's ability to classify objects, when a different usage is observed. As for the cutter, which is a new object that did not exist in the training set, the experimenter demonstrated using it to cut a piece of paper, which is also non-identical to any demonstrated usages provide in the training set. The cutter tests the proposed method's ability to classify new types of objects.

Feature vectors of these three objects are extracted from the usage demonstrations, and using the algorithm described in Section 5.1.4, each object is classified into one of the groups in the knowledge base constructed in the training stage. The handover grasp configuration for each object is then computed based on the classification results.

Figure 5-4 Three new objects used as the test set for evaluating the proposed method's ability to classify new objects. (Published in Chan et al. 2014 [110].)

5.3 Results

5.3.1 Knowledge Base Built from Training Set

Figure 5-3 shows the knowledge base built from the training set. The method proposed in this chapter was able to differentiate objects from their observed usages and group objects with similar affordances together. The proposed method successfully grouped screwdrivers into one group, different kitchen knives demonstrated to cut different objects into another group, marker pens used for the different actions of writing and drawing into one group, and plates and cutting boards having the common demonstrated function of carrying and transporting into another group.

5.3.2 Grasp Configurations Generated for New Objects

With the knowledge base built, the algorithm computed the appropriate handover grasp configurations for the three new objects. Table 5-2 presents the computed distances using the two different distance measurements between each object's feature vector and each cluster center, the classification results, and the computed grasp configurations for each of the new objects. Table 5-2 shows that classification using either the Euclidean distance or the angular distance yielded the same results for all three objects.

Table 5-2 Classification results and generated grasp configurations. Minimum distances are shown in bold font. (Published in Chan et al. 2014 [110].)

Object		Screwdriver	Marker Pen	Cutter
Distance (Euclidean)	Group 1	1.0815	0.9971	**0.0256**
	Group 2	**0.4512**	0.7999	1.3365
	Group 3	0.9385	0.9203	1.3771
	Group 4	0.6135	**0.7613**	1.3213
Distance (Angle)	Group 1	1.4609	1.4441	**0.0231**
	Group 2	**0.184**	0.689	1.4821
	Group 3	1.1877	1.1331	1.514
	Group 4	0.5582	**0.0528**	1.4351
Classification Result		Group 2	Group 4	Group 1
Computed Grasp Configuration		G_2	G_4	G_1

5.4 Summary

This chapter has presented an affordance based method for grouping and classifying objects for the purpose of determining proper handover grasp configurations of new objects. The presented method observes demonstrations of object usages and captures affordance information from the object's movements and interactions with other objects during usage. Handover grasp configurations of new objects are then computed based on the affordance features of the new objects. This chapter has also presented an experiment that tested the grouping of ten objects and classification of three new objects. Results showed that the presented method successfully created a knowledge base where objects with similar affordances are clustered into the same group, and new objects are classified appropriately into groups consisting of objects with similar affordances. Results demonstrated the proposed method's potential of classifying new objects and new usages. Thus, the proposed grouping and classification method can be used for computing appropriate handover grasp configurations for new objects. The next chapter will present an implementation of the framework Chapter 3 presented for enabling robots to determine handover grasp configurations automatically, using the components presented in this chapter and the previous chapter.

6 An Implementation of the Framework for Learning Object Handovers from Observations onto a Robot and its Validation

Chapter 3 presented a framework for enabling robots to determining proper handover grasp configurations for various objects automatically, and Chapter 4 and Chapter 5 provided the required building blocks for implementing this framework. This chapter will use the components of learning handover grasp configurations from observing human handovers and generalization to new objects Chapter 4 and Chapter 5 provided to implement the framework on a robot hardware.

6.1 Observing Natural Handovers

There are many challenges in robot perception when implementing a system for observing and learning handover grasp configurations from human handovers. These challenges deal with human recognition and tracking, object detection, and human intent recognition. The robot first needs to be able to recognize if a person is present. It needs to be able to track and follow the human's movements. Next, the robot also needs to be able to recognize the human's intention. When a person hands over an object, the robot needs to able to identify that a handover is taking place. Furthermore, when a person hands the object over, the robot needs to able to detect the object in the person's hand, and identify the grasp configuration used.

Face detection and silhouette tracking are a couple of existing methods for detecting and tracking humans. However, these method either do not provide rich enough information of the human's pose and motion, or can be computationally expensive. Recent advancements in skeleton tracking has provided a method for identifying and tracking humans that is computationally efficient and provides richer information. The literature provides examples of skeleton tracking being used for implementing human-robot handovers [124], and the implementation this chapter presents will also utilize skeleton tracking for obtaining human pose and motion information. Regarding handover intent, the literature includes studies related to construction of classifier for predicting occurrences of handovers, use of human or robot pose for signaling or recognizing handover intent, and prediction of handover location and time based on human motion [39], [56], [78], [124]. This chapter will use some of the results from these latest studies for determining when a handover is taking place. Recognizing the grasp configuration used in a handover is an especially difficult challenge, since the robot needs to recognize and identify the pose of the object that is held and occluded by the giver's hand. The performances of existing object detectors are often slow, and they cannot handle occlusions well. Furthermore, most existing object detectors are only capable of identifying object type and position, but not its orientation. In order to overcome these challenges, this thesis uses an alternative approach, where prior to the handover, the robot first identifies the object's pose before the giver picks it up at the pre-occluded state. A sequential Monte Carlo method is then used to predict and track the object's location and orientation at each subsequent instance. Using the approaches stated above, this chapter will present a successful implementation of the framework for learning handover grasp configurations from observations and interactions with humans.

6.2 Extracting Grasp Configurations from Natural Human-Robot Handovers

Figure 6-1 illustrates the system for learning handover grasp configurations from handovers with humans. At first, the robot looks at the environment, detects the objects in its surrounding, and starts tracking the objects' position and orientation. After the robot has started tracking the objects, it then begins to search for any person that enters the scene. When it detects a person, the robot starts tracking the person's skeleton. As the person reaches his/her hand over to an object and picks the object up, the robot determines where on the object the person grasps, and remembers the grasp point. Once the person picks up the object, the robot starts to pay attention to cues that would signal the person's intention of handing over the object. If the person at any time releases or places down the object, the robot goes back to the grasp detection phase. When the robot detects the cues signaling a handover from the person, it determines the orientation of the object in the person's hand, and stores it along with the previously extracted grasp point in its memory as the handover grasp configuration. After extracting the handover grasp configuration used by the person, the robot reaches its arm over and takes the object from the person. The following sections describe in more details the different components of the system.

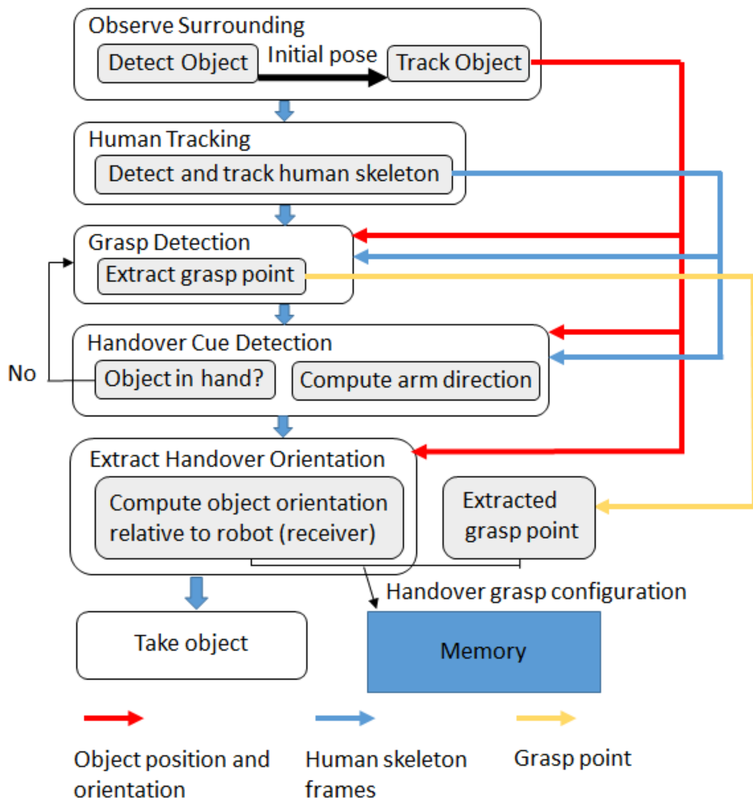

Figure 6-1 Flow diagram of system for extracting handover grasp configurations from handovers with users.

6.2.1 Object Detection

This thesis tested two methods for detecting the object. The first one is a model based approach and the second one is a model free approach. For the model based approach, this work used the RoboEarth software package for model creation and object detection [125]–[127]. The robot is provided with pre-captured point cloud models of the objects. During object detection, the algorithm first extracts image feature points from the camera input. It then uses the RANSAC algorithm to compute a rigid transformation between the object model and the camera input for determining the object's pose. A minimum of five correspondence points is required by the RANSAC algorithm. While computing feature correspondence, the object detection algorithm uses depth information from the point cloud to ensure validity by filtering out those with large depth disparity.

In the model creation stage, the model of the object is captured by placing it on an augmented reality (AR) marker plate. However, testing with RoboEarth revealed that its performance in object detection is highly sensitive to background changes. As a result, during the experiment, the market plate had to be placed underneath the object in many trials to obtain detection. Due to the aforementioned, this thesis also tests an alternative method for object detection. In the second model free approach, a marker plate is used for initial object detection. At the beginning of each trial, the object is placed above the marker. The robot detects the location and the orientation of the marker plate and creates a bounding box above it. The robot then detects the object by segmenting the cloud points in the bounding box. After segmenting the point cloud of the target object, the robot determines the location of the object by computing the centroid of the point cloud and the orientation from the marker plate's orientation. This approach eliminates the need of creating object models and supplying them to the robot ahead of time.

6.2.2 Object Tracking

Upon detecting the object, the object's point cloud is given to a tracker to continually track the object's six degree of freedom (6 DoF) pose. For tracking the object, this thesis uses a particle filter [128]. The state of the object at each instance is defined by its positon $\boldsymbol{x}$ and rotation $\boldsymbol{r}$ in three dimensional space. At each time step t, the particle filter generates a set of particles from the object's state in the previous time step according to a uniform motion model. Each particle represents a potential state of the object at time t. For each particle i with state $\{\boldsymbol{x_i}, \boldsymbol{r_i}\}$, the algorithm generates a hypothesis point cloud by transforming the object's point cloud by $\{\boldsymbol{x_i}, \boldsymbol{r_i}\}$, and a weight is calculated for the particle according to:

$$w_i = \sum_j L_{dist}(\boldsymbol{p_j}, \boldsymbol{q_j}) L_{color}(\boldsymbol{p_j}, \boldsymbol{q_j}) \tag{18}$$

where

$$L_{dist} = \frac{1}{1 + \alpha |\boldsymbol{p_j} - \boldsymbol{q_j}|^2} \tag{19}$$

$$L_{color} = \frac{1}{1 + \beta |\boldsymbol{p_{jcolor}} - \boldsymbol{q_{jcolor}}|^2} \tag{20}$$

In Equation (19) and Equation (20), $\boldsymbol{p_j}$ gives the position, and $\boldsymbol{p_{jcolor}}$ the HSV colour values of the jth point in the observed point cloud. Similarly, $\boldsymbol{q_j}$ gives the position, and $\boldsymbol{q_{jcolor}}$ the HSV colour values of the jth point in the hypothesis point cloud. α and β are constants. The particle with the highest computed weight is the state with the highest probability of where the object actually is. The algorithm then returns the position and rotation of that particle as the predicted pose of the object in the current time step. The presented implementation uses this method for tracking the object's 6 DoF pose. Since this algorithm is robust to partial occlusions, it is capable of tracking the object even when the object is held in the giver's hand and moved by the giver, thus, enabling the robot to extract the handover orientation used when the human giver hands over the object. In comparison to some alternative methods [129], the algorithm described here also has the merit of not needing the user to first create virtual models of the objects in a CAD software or being limited to tracking objects with simple edges.

6.2.3 Human Tracking and Grasp Detection

Once the robot has begun tracking the object, it starts to search for people who enter the area. This implementation uses the OpenNI skeleton tracker for detecting and tracking humans [130]. The OpenNI tracker provides human joint location and orientation information. When the robot detects a person, it tracks the person, paying attention to the hands to determine when the person grasps the object and picks it up. The robot detects grasp by computing the distance between the person's hand and the object, and determines that the person has grasped the object when this distance reduces below an empirically determined threshold of $\varepsilon_{contact}$ = 13 cm. This work sets the value of $\varepsilon_{contact}$ base on the approximate width of a human hand. Experimentation with different values of $\varepsilon_{contact}$ shows that a large value causes the robot to detect grasp too early. On the other hand, a small value may cause the robot to fail to detect grasp all together, since the hand-to-object distance is computed using their centroids, and they cannot get closer than their physical sizes would permit.

When the robot detects that the person has grasped the object, it extracts the grasp point P_o^h. The grasp point P_o^h is expressed as the relative hand position of the person in the object's frame. To determine P_o^h, the robot uses the person's hand position in the camera frame, P_c^h, from the skeleton tracker, and the object's transformation in the camera frame, T_c^o, from the object tracker:

$$P_o^h = T_o^c P_c^h \tag{21}$$

where T_o^c = ${T_c^o}^{-1}$ gives the inverse transformation of T_c^o.

6.2.4 Handover Cue Detection

After the robot detects that the person has picked up the object, it starts to pay attention to cues that would signal the person's intention of handing over the object. The robot determines if the person is handing the object over to it by looking for the following cues:

1. The person is located near the robot.
2. The person is holding the object in his/her hand.
3. The vector pointing from the person's torso to the person's hand is directed towards the robot's torso.

These cues were designated considering existing studies found in the literature [29], [88]. To avoid the person's transient movements from triggering false detection of handovers, the robot requires a two second consecutive observation of these cues to consider it valid. If the person places down or release the object at this stage, the robot returns to the grasp detection phase. Once the robot detects the handover cues, it proceeds to extract the handover orientation used.

6.2.5 Handover Orientation Extraction

The robot extracts the handover orientation, T_r^o, as the transformation of the object in the receiver (robot) torso frame. To compute the handover orientation T_r^o, the robot uses T_c^o, its hardware parameters, and its joint states. Using its own hardware parameters and its joint states, the robot first computes T_t^c, the transformation between the frame of the camera mounted on its head and its torso. It then computes T_t^o as:

$$T_t^o = T_t^c T_c^o \tag{22}$$

Once the robot has extracted the handover orientation, it stores it along with the grasp point it previously extracted as the handover grasp configuration. After that, the robot takes the object from the person and completes the handover.

The next section documents two sets of experiments on learning handover grasp configurations. The procedure described above is used in the first set of experiments for extracting the handover orientation. Based on results from the first set of experiment, the experimenter made three modifications for the second set of experiments. First, when extracting the handover orientation, instead of computing the object's orientation in the robot's torso frame, the object's orientation is computed in a base frame defined by the giver's (human) and the receiver's (robot) torsos as described in Section 4.3.1. This takes into account different positions of the human giver. Second, due to the different arm segment proportions between humans and the robot, including the giver torso-to-object position, P_t^o (i.e., the translational component of T_t^o), as part of the handover grasp configuration was found to prevent the robot from solving the inverse kinematic problem in some cases when trying to position its hand for handover. Thus, the constraint of P_t^o is removed and only the rotational component of T_t^o is included in the grasp configuration. Third, the relative position of the giver's second hand to the object is also noted at handover, to enable learning of two-handed grasp configurations. The robot determines that a two-handed grasp configuration is used if the distance between the object and the giver's second hand is less than $\varepsilon_{contact}$, where $\varepsilon_{contact}$ is empirically set at 15 cm.

6.3 Experiment on Learning Handover Grasp Configurations

This section presents the implementation of the framework on a Kawada Industries HRP2V robot and the experiments conducted for testing learning of handover grasp configurations from observations of human-robot handovers. This section documents two sets of experiments testing the two object detection methods described in Section 6.2.1. The first set of experiments tests the model based object detection method using three different objects. Figure 6-2 shows the three objects used. The second set of experiments tests the model free object detection method with a larger set of objects containing fifteen items. Figure 6-3 shows the fifteen objects used in the second set of experiments. In the experiments, a human giver handed the objects over to HRP2V to provide the demonstrations. The human giver provided three demonstrations for each object in the first set of experiments and five demonstrations for each object in the second set of experiments.

Figure 6-2 Three daily objects, a spray can, a mug, and a detergent bottle, used in first set of experiments testing the model based object detection method. (Published in Chan et al. 2015 [131].)

Figure 6-3 Fifteen daily objects used in the second set of experiments testing the model free object detection method.

6.3.1 Hardware Platform

The Kawada Industries HRP2V robot, which Figure 6-4 shows, is an upper body humanoid robot with a movable base. The HRP2V used in this implementation has an ASUS Xtion Pro Live RGBD camera mounted in the head. This work uses the Robot Operating System (ROS) for implementing the software. HRP2V has two onboard computers, one used for controlling its motors, and one for running ROS and for vision processing. Although this implementation offloaded the vision processing and overall flow control to an external computer, it is possible to move all components onto the onboard computers and have the entire system standalone.

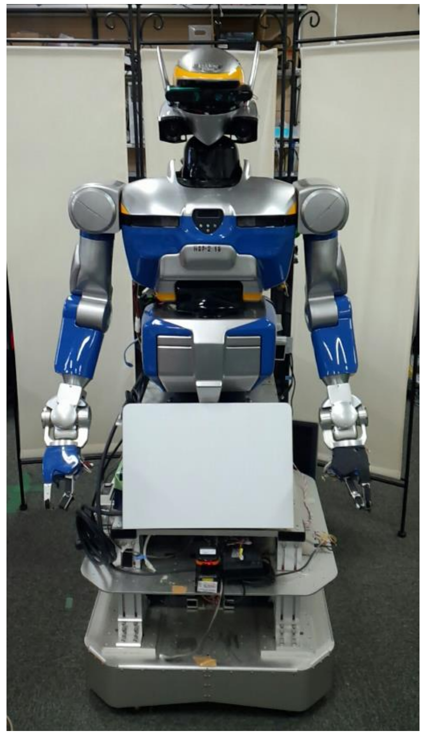

Figure 6-4 Kawada Industries HRP2V Robot with ASUS Xtion Pro Live camera mounted on the head. (Published in Chan et al. 2015 [131].)

6.3.2 Procedure

For the handover demonstrations, the experimenter first placed the object on a table located in front of HRP2V. HRP2V began by looking in the direction of the table. It detected and then began to track the object (Figure 6-5A). After it started to track the object, HRP2V looked up, keeping the object insight to avoid losing track, and tried to detect any humans that enter the scene. When HRP2V detected a person, it began to track the person's skeleton (Figure 6-5B). Once it saw the person grasping the object, it extracted the grasp point on the object (Figure 6-5C). HRP2V then looked for the handover cues that would indicate the person's intent to hand over the object. Once it detected the cues (Figure 6-5D), HRP2V identified the handover orientation used by the giver (Figure 6-5E). Finally, it reached over to take the object and completed the handover. The giver took care not to move too fast to avoid tracking problems. For the first set of experiments, the initial pose of the spray can and detergent bottle were kept consistent, and the initial pose of the mug was varied. In the second set of experiments the experimenter varied the initial pose of all objects for each trial.

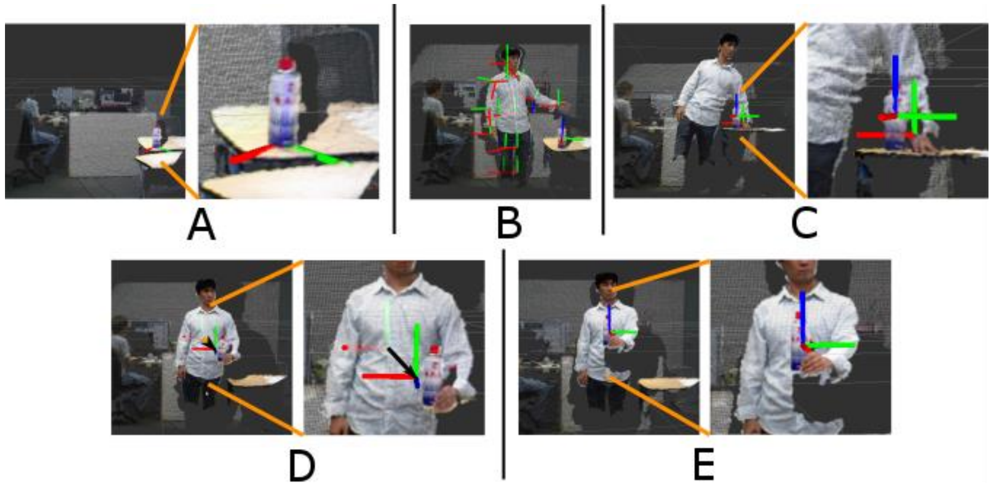

Figure 6-5 Stages of extracting grasp configurations from handover demonstrations. A – Object detection. B – Human detection and tracking. C – Grasp point detection. D – Direction of torso-to-hand vector for handover cue detection. E – Object orientation at handover. (Published in Chan et al. 2015 [131].)

6.3.3 Results

Figure 6-6A shows the giver demonstrating the handover grasp configurations to HRP2V by handing over the objects to HRP2V in the first set of experiments. Figure 6-6B shows the camera view of HRP2V at the time when it detected the handover cues and extracted the object orientations. The giver demonstrated the spray can handovers grasping the bottom of the can and facing the label towards HRP2V (receiver), the mug handovers presenting the handle to the receiver, grasping the opposite side, and the detergent bottle handovers with the nozzle pointing away from the receiver, grasping the body of the bottle. Table 6-1 lists the extracted grasp configurations from all demonstrations, along with the computed averages, with P_t^o and R_t^o giving the translational and rotational components of T_t^o respectively. The average of the orientations, R_t^o, are computed by expressing the rotations in the axis angle representation and averaging each of them.

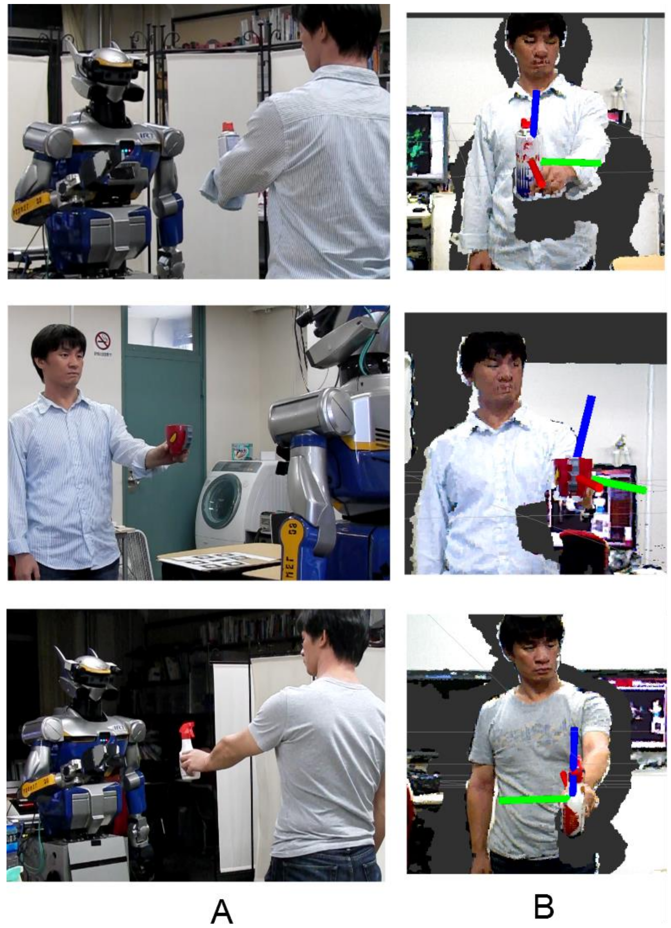

Figure 6-6 A – Handover demonstrations of a spray can, mug, and detergent bottle. The spray can is handed over with the label facing the receiver, the mug with the handle towards the receiver, and the detergent bottle with the nozzle pointing away from the receiver. B – HRP2V's camera image at the moment when handover cues are detected, showing the extracted handover object orientations. (Published in Chan et al. 2015 [131].)

Table 6-1 Extracted grasp configurations from handover demonstrations, and the computed averages. (Published in Chan et al. 2015 [131].)

		(mm)						*(deg)*	*(mm)*		
		$P_t^o(x)$	$P_t^o(y)$	$P_t^o(z)$	$R_t^o(x)$	$R_t^o(y)$	$R_t^o(z)$	$R_t^o(\theta)$	$P_o^h(x)$	$P_o^h(y)$	$P_o^h(z)$
Spray Can	Demo 1	887.40	-306.10	-38.67	-0.09	-0.01	1.00	161.0	-136.39	21.41	-57.05
	Demo 2	889.73	-194.33	28.11	-0.03	0.06	1.00	174.2	-142.09	36.18	-19.80
	Demo 3	868.11	-210.59	40.91	0.10	0.00	-1.00	163.9	-121.33	64.45	-54.19
	Avg	881.75	-237.01	10.12	-0.01	0.02	0.33	166.2	-133.27	40.68	-43.68
Mug	Demo 1	889.71	-55.20	126.49	0.01	0.06	-1.00	178.2	-132.90	63.97	-20.22
	Demo 2	949.13	-66.91	47.75	-0.08	-0.02	-1.00	159.9	-146.13	-22.36	-25.18
	Demo 3	878.99	-10.89	80.04	-0.09	-0.07	-0.99	186.8	-117.05	88.56	-19.09
	Avg	905.94	-44.33	84.76	-0.05	-0.01	-1.00	174.8	-132.02	43.39	-21.50
Detergent Bottle	Demo 1	838.10	-130.00	15.09	0.66	-0.27	0.70	50.4	142.32	46.06	-5.28
	Demo 2	818.81	-122.75	53.20	0.40	-0.21	0.89	31.5	131.78	50.48	-0.06
	Demo 3	814.20	-77.34	83.60	0.23	-0.03	0.97	32.1	113.23	89.43	-21.16
	Avg	823.70	-110.03	50.63	0.43	-0.17	0.86	38.4	129.11	61.99	-8.83

Figure 6-7 and Figure 6-8 shows the results for the second set of experiments. Figure 6-7 provides a visualization of the handover orientations in 3D plots, with thin lines showing the handover orientations extracted from each trail, and thick lines showing the averages computed using the distance minimization method developed in Section 4.3.3. Figure 6-8 shows the extracted grasp points in blue, and the positions of the second hand in red, for cases where a two-handed configuration was used by the giver. The hollow markers show the data from each trial, and the solid markers show the computed averages. Average for the position of the second hand is computed only if more than half of the observed demonstrations are identified as two-handed handovers. The raw data and computed averages for the handover grasp configurations are provided in Appendix A.

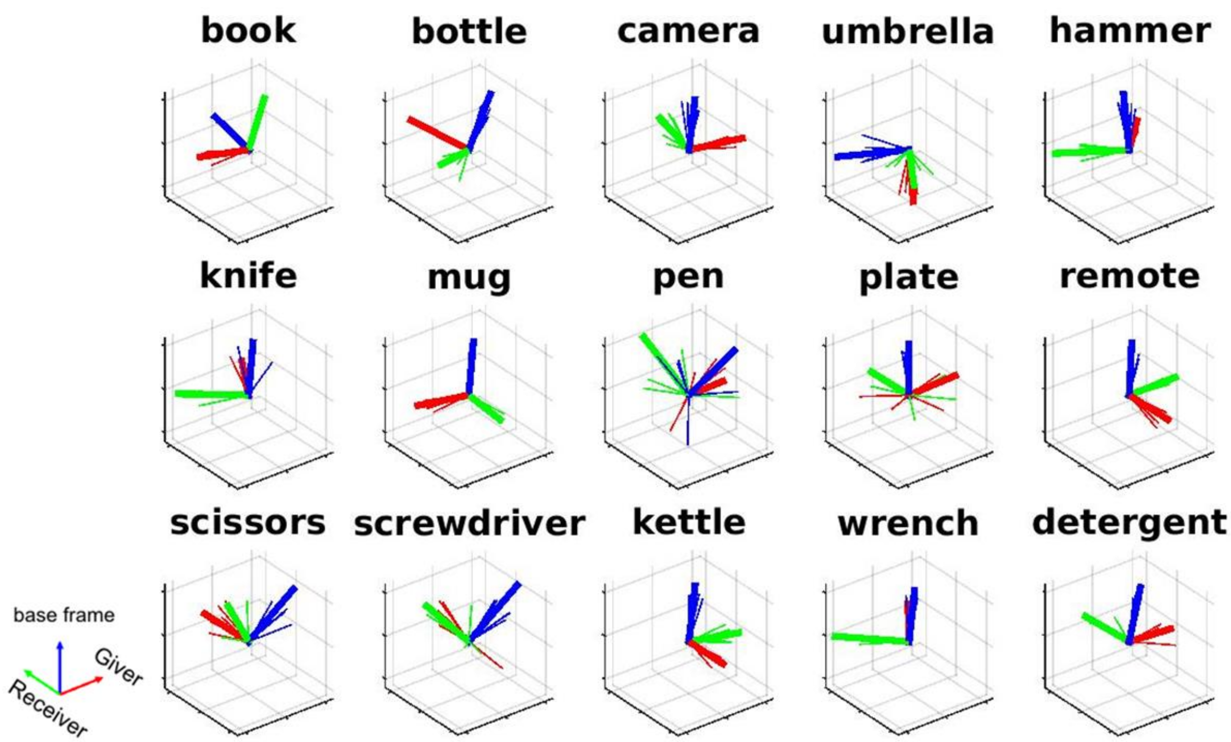

Figure 6-7 Handover orientations extract for the fifteen objects in the second set of experiments. Thin lines show extracted orientations from each trial, thick lines show computed averages.

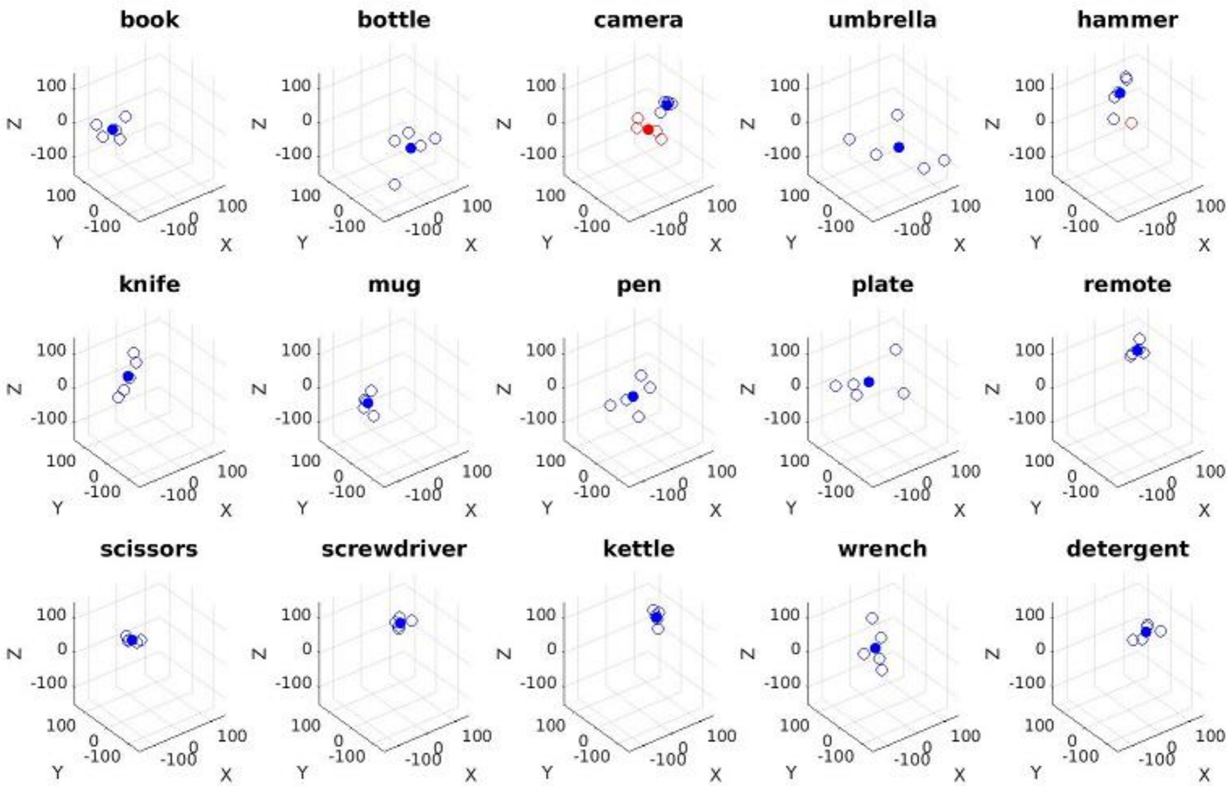

Figure 6-8 Handover grasp points extract for the fifteen objects in the second set of experiments. Blue circles indicate grasp point, and red circles indicate position of the second hand where a two-handed handover is detected. Hollow circles show the data from each trial and solid circles show the computed average.

6.4 Experiment on Handover Execution

This section presents the experiments conducted to test the robot's ability to execute handovers of various objects. First, to validate the robot's ability to hand over known objects, this section presents experiments where the robot executes handovers of the two sets of objects given in Section 6.3. Next, to validate the robot's ability to handover objects whose handover grasp configurations have not been observed, this section presents an experiment where the robot is asked to handover the three unknown objects given in Section 5.2.3. Using the object grouping and classification method and the observed usages of objects from Chapter 5, together with the observed handover demonstrations from this chapter, the robot determines the appropriate handover grasp configurations for the unknown objects and executes the handovers in the experiment.

6.4.1 Determining Appropriate Handover Grasp Configuration from a set of Observations

Given a set of observed handover grasp configurations for an object, the appropriate grasp configuration to be used can be determined by the robot by computing the mean of the handover orientations and the grasp points, assuming that the observed set contains receiver-centered grasp configurations. However, the user study Chapter 4 presented shows that humans do not necessarily naturally used receiver-centered handover orientations. Thus, the robot should give consideration to the quality of the observed demonstrations. If a set of observations is made up of rather arbitrary grasp configurations, or more than one distinct grasp configurations, the mean of the set may not necessarily be an appropriate grasp configuration, and in fact, the different parts of the computed mean grasp configuration may conflict with each other. For example, through a test where the experimenter provided a set of demonstration containing rather arbitrary grasp configurations to the robot, the experimenter encountered a case where the computed mean would require the robot to grasp and present the same side of the object to the receiver. This would be a very unnatural and inappropriate handover grasp configuration if the robot were to execute it, although, in this case, the grasp configuration actually presented an unsolvable inverse kinematic problem to the robot to begin with; after grasping one side of the object, the robot's arm did not have enough degrees of freedom/length to present the same side to the receiver.

Results of the user study in Chapter 4 revealed that a set of observations containing non-receiver-centered orientations tends to have larger variation, and that the average spread of affordance axes in the set, δ, is significantly larger. Thus, in this implementation, the robot uses the δ of a set of observed grasp configurations to evaluate the quality of the set. Given a set of observed handover grasp configurations, the robot computes the δ of the set, and compares it with a threshold, δ_{thres}. If δ is below δ_{thres}, then the set is deemed to be of good quality and likely to be consisted of receiver-centered grasp configurations, and the robot uses the computed mean of the set to handover the object. Otherwise, the set is deemed to be of poor quality and that the computed mean to be unsuitable. In such cases, the robot picks one of the observed grasp configurations from the set and uses it to handover the object, based on the assumption that, while the set may be of poor quality as a whole, possibly due to it containing multiple modes of grasp configurations, each grasp configuration in the set was observed from a human demonstration, and thus, should at least be sensible and adequate by itself.

To selected a value for the threshold δ_{thres}, this thesis uses the results from the user study presented in Chapter 4. Section 4.4.4 reported that in Condition C, δ_C had an average of 27.1 deg with a standard deviation of 13.3 deg. Taking the average of δ_C plus one standard deviation, this thesis uses the value of δ_{thres} = 40.4 deg and implements the check for the quality of observed grasp configuration set in the second set of experiments. Results of Chapter 4 shows that when handing over a pen, users tend to use a large variety of orientations, and that the δ_C for the pen has the largest value. According to this, the experimenter simulated a poor quality set of handover grasp configurations for the pen by providing a set of demonstrations with large variety.

6.4.2 Procedure

In the testing of the learned grasp configurations, the experimenter first placed the object on a table located close to HRP2V. The spray can and detergent bottle in the first set of experiments were placed diagonally in front of the robot, and the experimenter varied the initial pose of the mug in the first set of experiments and all objects in the second set of experiments. HRP2V began by looking in the direction of the table. It detected the object, and then calculated the grasp point in the world frame. HRP2V then positioned its hand end effector frame at the grasp point using inverse kinematics, moved its hand towards the object, and picked up the object. The hand end effector frame is located between the two opposing fingers of the hand. When HRP2V picked up the object, it remembered the transformation from its hand to the object. After picking up the object, HRP2V looked up towards where the person is. Using the transformation from the hand to the object, it then computed where it needed to place its hand to orient the object properly according to the desired handover grasp configuration. After determining where its hand needed to be, HRP2V used inverse kinematics to place its hand at the corresponding pose to hand the object over.

The experimenter performed two types of evaluations: a qualitative one and a quantitative one. In the qualitative evaluation, no markers were attached to the objects, and the resulting handover grasp configurations were compared qualitatively with the learned handover grasp configurations. In the quantitative evaluation, an AR marker was attached to the object to obtain a ground truth measurement of the achieved handover orientation R_m. These measurements are then compared with the learned handover orientation R_l by computing the error e as the rotation angle between R_m and R_l:

$$e(R_l, R_m) = acos(\frac{tr({R_l}^{-1}R_m)-1}{2}) \qquad (23)$$

6.4.3 Handover Results of Known Objects

Figure 6-9 shows HRP2V handing over the objects in the qualitative verification of the first set of experiments using the grasp configurations it has learned in Section 6.3. Examining the handover grasp configurations used by HRP2V, Figure 6-9 shows that HRP2V was able to hand over each object using a grasp configuration that matches the one demonstrated to it by the human. HRP2V grasped the spray can by the bottom part and handed it over with the label facing the receiver, it presented the mug handle to the receiver, grasping the opposite side, and it picked up the detergent bottle by the body of the bottle, handing it over with the nozzle directed away from the receiver.

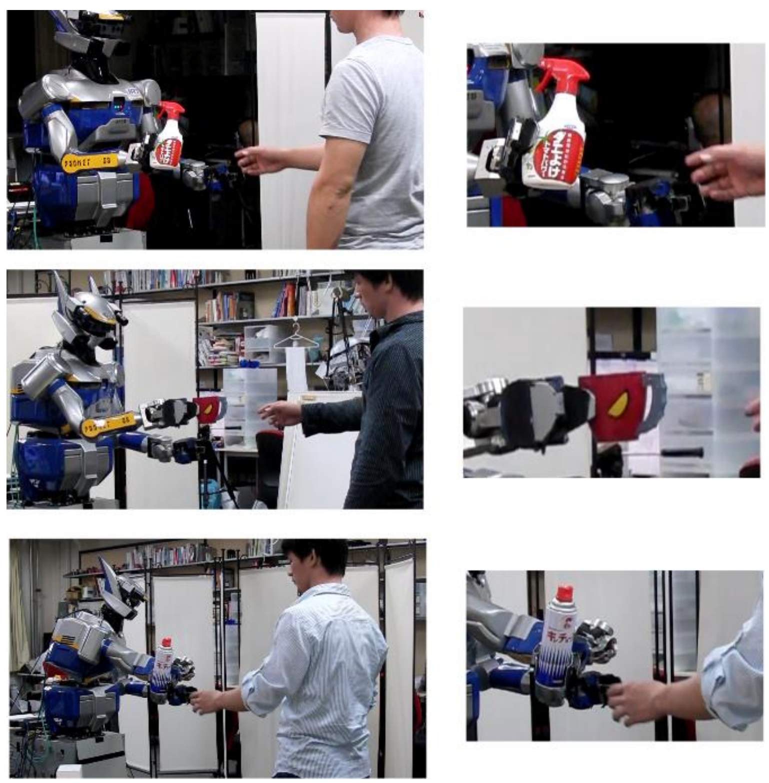

Figure 6-9 Handover execution in the first set of experiments using learned grasp configurations. The robot successfully handed over all three objects using grasp configurations that match the ones demonstrated to it by the person previously. (Published in Chan et al. 2015 [131].)

Table 6-2 presents the quantitative evaluation results of the first set of experiments, showing the learned handover orientation R_l, measured handover orientation R_m, and the computed error $e(R_l, R_m)$

between the two. The detergent bottle had the smallest error at 0.45 deg, the spray can had the largest error at 15.8 deg, and the mug had an error of 14.0 deg. The average error of the three objects is 10.1 deg.

Table 6-2 Comparison of learned handover orientations, measured orientations, and computed errors in the quantitative evaluation of the first set of experiments. (Published in Chan et al. 2015 [131].)

	$R_l(x)$	$R_l(y)$	$R_l(z)$	$R_l(\theta)$ (deg)	$R_m(x)$	$R_m(y)$	$R_m(z)$	$R_m(\theta)$ (deg)	$e(R_l, R_m)$ (deg)
Spray Can	-0.01	0.02	0.33	166.2	-0.09	-0.01	1.00	161.0	16.0
Mug	-0.05	-0.01	-1.00	174.8	-0.03	0.06	1.00	174.2	13.8
Detergent Bottle	0.43	-0.17	0.86	38.4	-0.19	-0.04	-0.98	179.3	0.45

In the second set of experiments, to determine whether the set of observed demonstrations for an object is of good or poor quality, the robot first computed the δ of the set and compared it with δ_{thres}. Table 6-3 shows the computed δ for each of the fifteen objects. For all objects except the pen, the δ is found to be smaller than δ_{thres}, indicating that the sets of demonstrations for those objects are of good quality. Thus, for all objects except the pen, the robot used the computed mean of the set of observed grasp configurations to handover the object. As for the pen, the δ is found to be greater than δ_{thres}, indicating that the set of observations contains a large variation and is of poor quality. Thus, instead of using the computed mean of the set, the robot used one of the observed grasp configurations from the set to handover the pen.

Table 6-3 Computed spread of the affordance axes in the set of observations for each of the fifteen objects in the second set of experiments.

	book	bottle	camera	umbrella	hammer	knife	mug	pen	plate	remote	scissors	screw-driver	kettle	wrench	detergent bottle
δ (deg)	2.8	5.5	6.2	6.5	6.8	3.6	2.8	48.6	5.1	3.3	10.6	17.7	5.2	2.9	3.4

Figure 6-10 and Figure 6-11 show HPR2V handing over the fifteen objects in the second set of experiments using the learned grasp configurations during qualitative evaluation. Results show that HRP2V was able to handover the objects using appropriate grasp configurations. For example, HRP2V was able to hand over the hammer, kitchen knife, and wrench with the handle towards the receiver and the head or tip pointed away from the receiver; it was able to handover the bottle, kettle, mug, and plate in an upright orientation, while pointing the detergent bottle nozzle away from the receiver, and presenting the mug handle to the receiver; and it was able to handover the camera, a delicate electronic device, using a two-handed configuration as demonstrated to it previously.

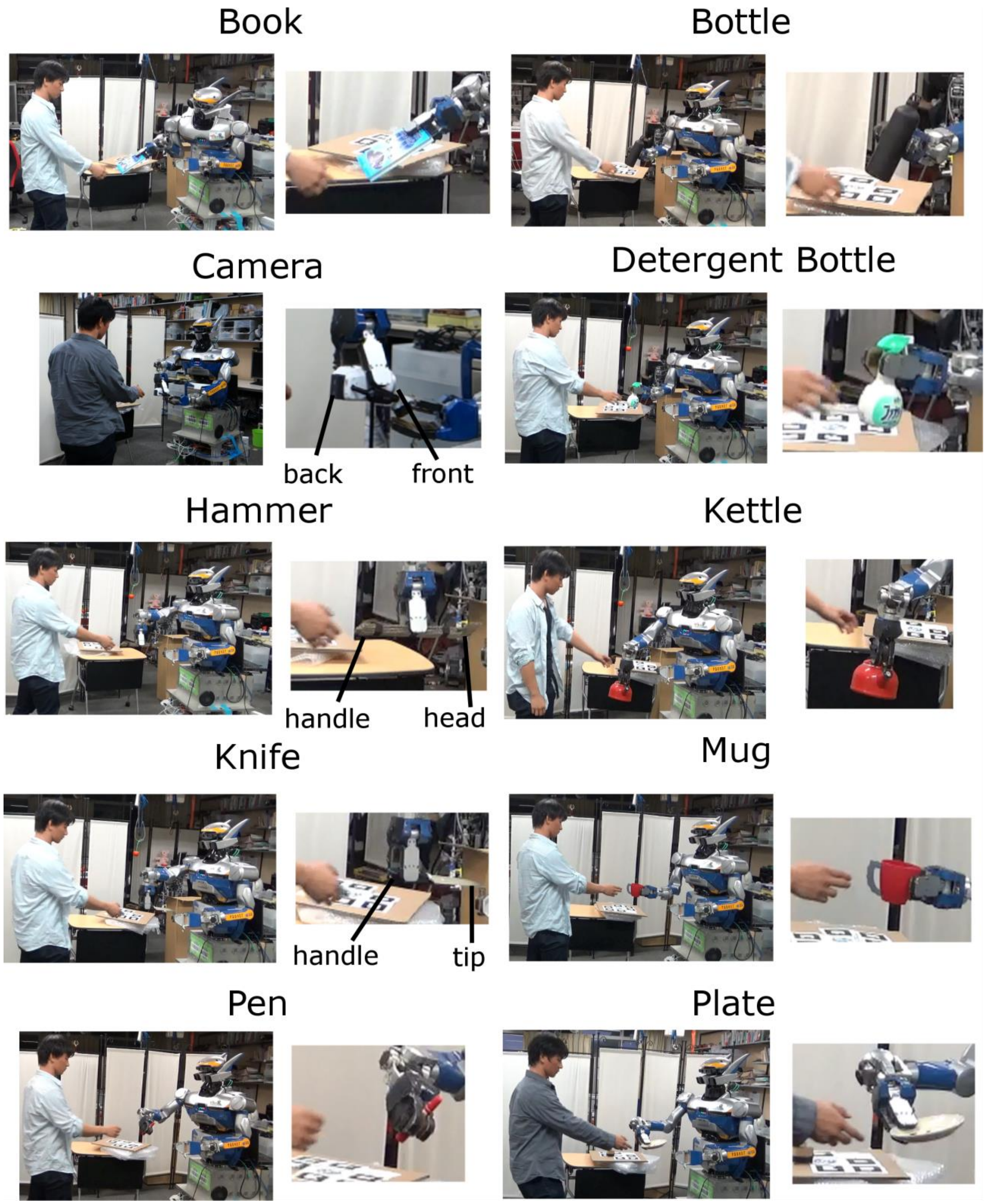

Figure 6-10 Handover execution using learned grasp configurations during qualitative evaluation in the second set of experiments. Figure shows the grasp configurations used for the first ten objects of the fifteen objects used in the second set of experiments.

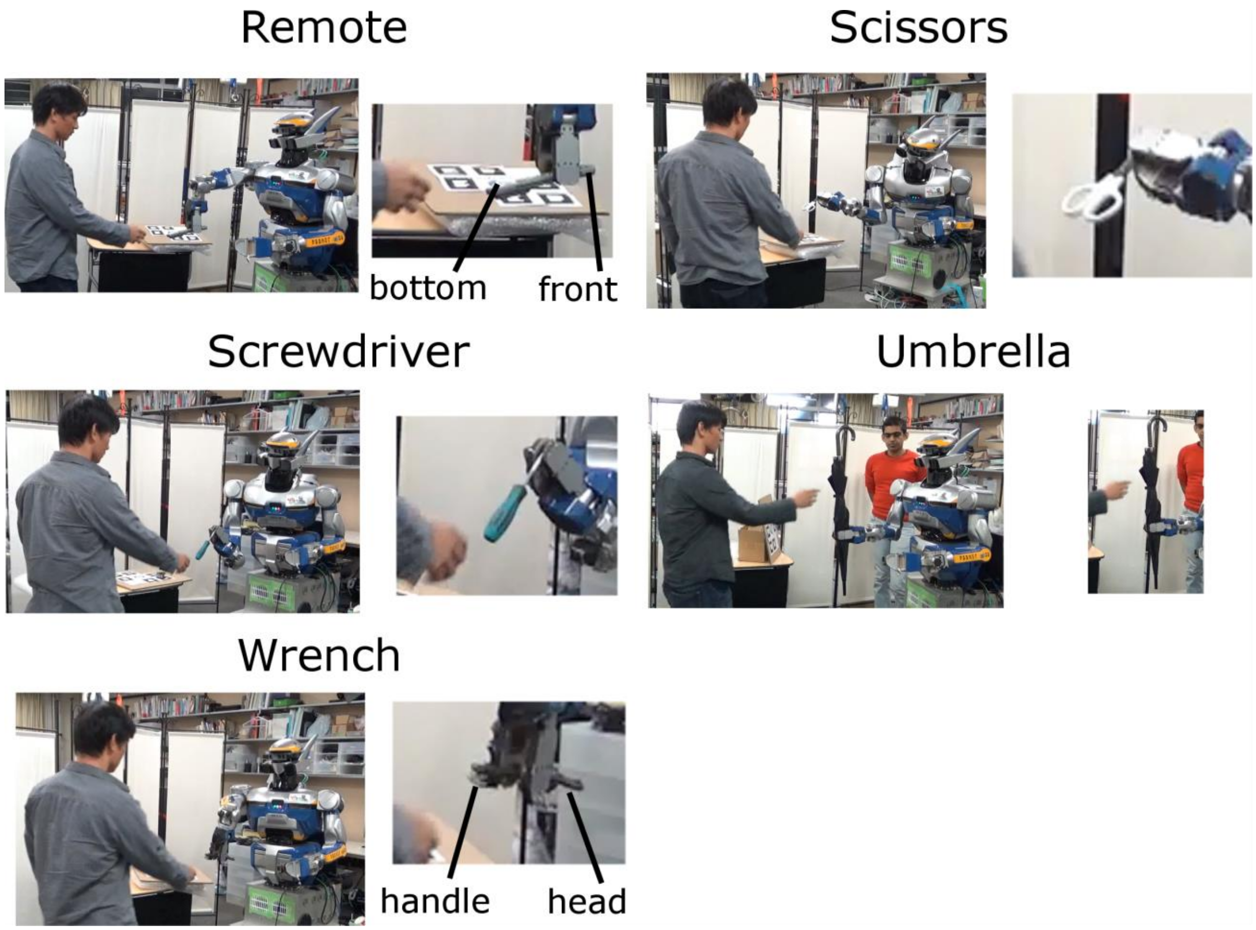

Figure 6-11 Handover execution using learned grasp configurations during qualitative evaluation in the second set of experiments. Figure shows the grasp configurations used for the last five objects of the fifteen objects used in the second set of experiments.

Table 6-4 shows the results of the quantitative evaluation of the fifteen objects in the second set of experiments. Comparing the learned handover orientations R_l with the measured handover orientations R_m, the table shows that the computed error $e(R_l, R_m)$ for the objects range from 4.0 deg for the hammer to 37.2 deg for the plate. The average error for all the objects is computed to be 16.2 deg.

Table 6-4 Comparison of learned handover orientations, measured orientations, and computed errors in the quantitative evaluation of the second set of experiments.

	$R_l(x)$	$R_l(y)$	$R_l(z)$	$R_l(\theta)$ (deg)	$R_m(x)$	$R_m(y)$	$R_m(z)$	$R_m(\theta)$ (deg)	$e(R_l, R_m)$ (deg)
bottle	0.21	0.06	0.98	118.0	0.19	0.03	0.98	112.3	6.3
book	-0.08	-0.64	-0.76	173.6	-0.18	-0.62	-0.76	171.9	11.5
camera	0.54	0.26	-0.80	10.9	-0.12	-0.19	-0.97	5.7	8.6
detergent bottle	0.09	0.71	0.70	10.9	0.16	-0.12	0.98	39.0	33.2
kettle	-0.10	0.05	-0.99	93.4	-0.09	-0.02	-1.00	99.7	9.2
knife	0.06	0.06	1.00	57.9	0.42	-0.28	0.86	67.0	32.1
hammer	-0.18	0.08	0.98	47.6	-0.22	0.07	0.97	51.0	4.0
wrench	0.09	0.04	0.99	53.3	0.37	-0.07	0.93	60.2	17.8
mug	0.05	0.01	1.00	174.2	0.06	-0.01	1.00	158.1	16.0
pen	0.15	-0.82	0.54	28.6	0.09	-0.97	0.23	22.9	10.9
plate	0.17	-0.28	0.94	2.3	-0.02	-0.99	0.11	37.8	37.2
remote	0.00	0.03	-1.00	87.7	0.04	-0.07	-1.00	90.0	9.2
scissors	0.43	0.39	0.82	114.0	0.45	0.30	0.84	108.9	10.9
screwdriver	0.43	0.36	0.83	94.0	0.37	0.28	0.88	88.8	10.9
umbrella	0.27	0.91	-0.30	91.1	0.40	0.75	-0.53	89.4	25.2

6.4.4 Handover Results of Unknown Objects

In the handover experiment of unknown objects, the experimenter asked the robot to hand over the three new objects presented in Section 5.2.3: a marker pen, a screwdriver, and a cutter. Using the classification results from Chapter 5 together with the observed handover grasp configurations from demonstrations in this chapter, the robot handed over the new objects using the grasp configurations learned in this chapter. From the results of Chapter 5, the marker pen, screwdriver, and cutter were classified into the groups of marker pens, screwdrivers, and kitchen knives respectively. Thus, the robot handed over the three objects using the grasp configurations learned in Section 6.3 for the pens, screwdrivers, and kitchen knives respectively.

Figure 6-12 shows the qualitative results for the handover executions of the three new objects. For the marker pen, the robot grasped the middle and tilted the top towards the receiver. For the screwdriver, the robot grasped the top half and oriented the handle towards the receiver. For the cutter, the robot picked it up by the middle part, and presented the handle end to the receiver, pointing the cutting tip away from the receiver.

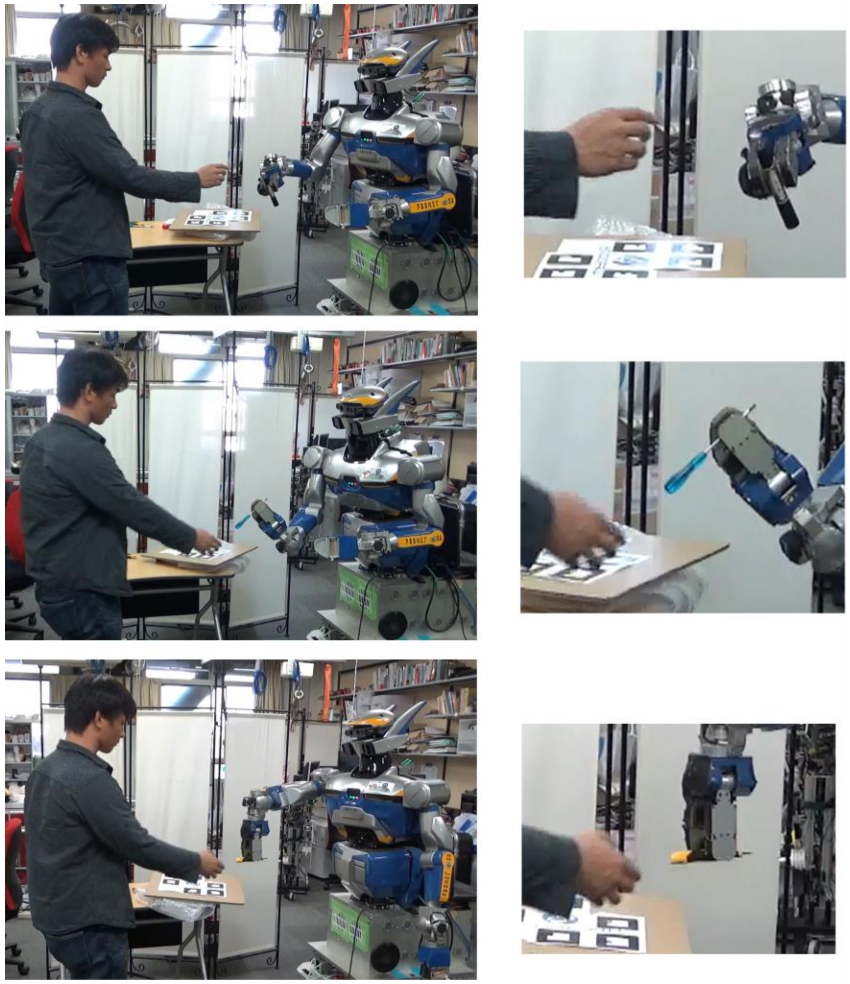

Figure 6-12 Handover execution of new objects, showing grasp configuration used for the marker pen, screwdriver, and cutter from top to bottom.

Table 6-5 provides the quantitative evaluation results for the handover execution of the three new objects. The error between the learned and the measured orientations are 7.3 deg for the pen, 15.7 deg for the screwdriver, and 7.4 deg for the cutter. The average error for the execution of unknown objects is 10.2 deg.

Table 6-5 Quantitative evaluation results for handover execution of unknown objects, showing learned handover orientations, measured orientations, and computed errors.

	$R_l(x)$	$R_l(y)$	$R_l(z)$	$R_l(\theta)$ (deg)	$R_m(x)$	$R_m(y)$	$R_m(z)$	$R_m(\theta)$ (deg)	$e(R_l, R_m)$ (deg)
pen	0.15	-0.82	0.54	28.8	-0.10	-0.85	0.52	27.4	7.3
screwdriver	0.43	0.36	0.83	94.0	0.40	0.24	0.88	104.3	15.7
cutter	0.06	0.06	1.00	58.1	-0.05	-0.02	1.00	56.0	7.4

6.5 Consideration for Additional Non-Object-Related Factors as an Extension

The proper handover grasp configuration of an object depends heavily on the properties of the object itself. However, as Chapter 2 has discussed, in addition to object properties, there are other factors that also affect whether a grasp configuration is considered to be appropriate or not. For example, receiver state, user identity, and intended use of the object are all such factors. This section demonstrates how the presented framework can be extended to account for such factors.

6.5.1 Influence of User Identity on Handover Grasp Configuration

Hierarchical structures can be found to exist in most societies, where each person's role and social status determine their level in the hierarchy [132]. In certain cultures, what is considered as a proper handover grasp configuration is dependent on who the receiver is, and it can be observed that the grasp configuration used by a giver changes depending on the relative statuses of the giver and receiver [46]. In some cultures, it is required that two hands be used when handing over an object in order to show politeness and respect. Thus, while it may be appropriate to hand over an object to a person of equal or lower status (such as handing over to a friend or a child) using one hand, it is considered to be inappropriate when handing over to a person of higher status (such as an elder or a customer), and that a two-handover configuration must be used. In such cases, the proper grasp configuration for handing over an object does not only depend on the properties of the object, but also the relative statuses of the receiver and the giver.

Given a giver status of S_{giver} and a receiver status of $S_{receiver}$, if $S_{receiver} > S_{giver}$, meaning that the receiver has a higher status than the giver (for example a store clerk handing over a merchandise to a customer), then it is more appropriate to hand over the object using a more formal two-handed grasp configuration. If $S_{receiver} = S_{giver}$ or $S_{receiver} < S_{giver}$, meaning that the receiver has an equal status, or lower status than the giver (for example, handing over to a friend or a child), then it is more appropriate to hand over the object using a more casual one-handed grasp configuration. Behaving formally towards someone of higher status can show politeness and casually towards someone of equal or lower status can show friendliness, whereas behaving formally towards someone of equal or lower status might appear unfriendly and behaving casually towards someone of higher status might appear impolite. Thus, it is important that an appropriate behaviour is selected for interaction based on the statuses of the interacting parties. The following sections show how the presented framework can take user identity into account when determining appropriate grasp configurations for handovers.

6.5.2 Grasp Configuration Selection Considering Additional Factors

Figure 6-13 shows the extended framework for determining appropriate handover grasp configurations, taking into account additional non-object-related factors. When handing over an object, the robot first identifies the object to be handed over. In the experiments Section 6.4 documented, the robot referenced its knowledge base, and determined the appropriate handover grasp configuration based only on what the object is. In the extended framework shown in Figure 6-13, the grasp configuration selector also takes into account additional factors when selecting the appropriate grasp configuration from the knowledge base. In the current case, the grasp configuration selector takes in the user's and the robot's statuses as additional inputs, and determines the proper grasp configuration based on the relative status of the robot (giver) and the user (receiver). The following sections describe an implementation of the extended framework and an experiment for testing the implementation.

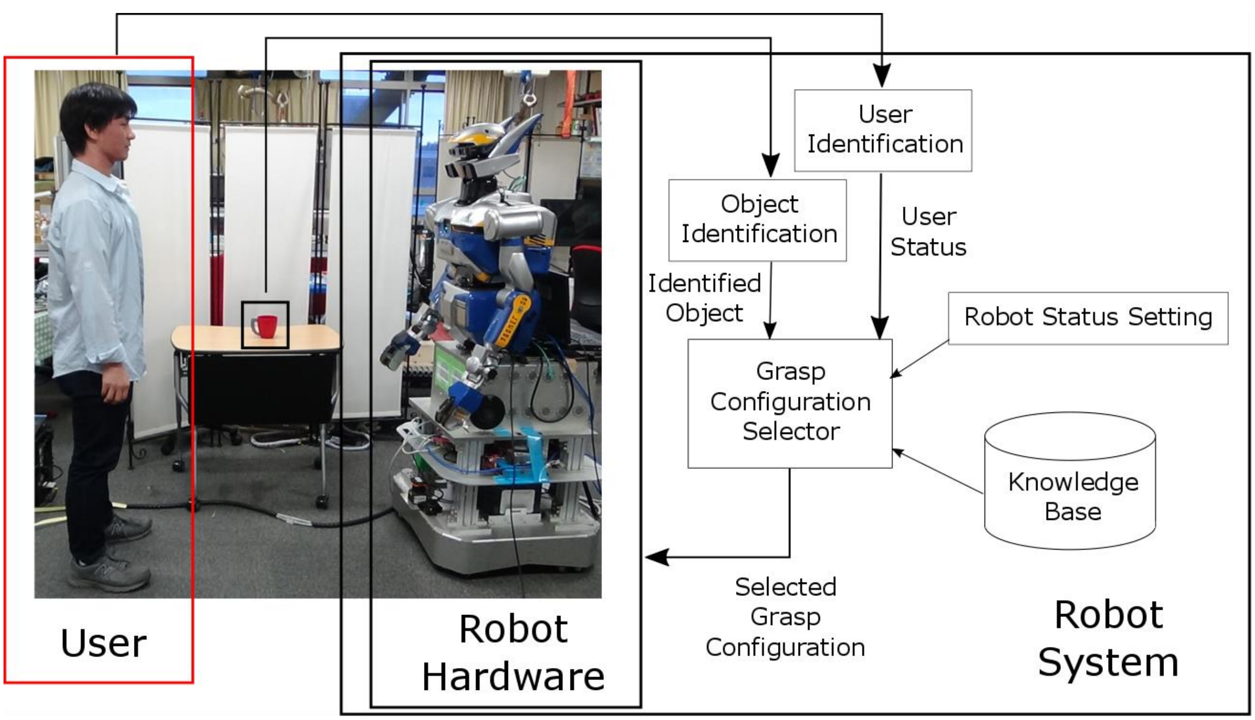

Figure 6-13 Extended framework for considering additional non-object-related factors when selecting appropriate handover grasp configuration.

6.5.3 Implementation

The implementation described in Sections 6.2 through 6.4 is extended through the following modifications. First, the robot is assigned a status setting of S_R. During handover demonstration when a person hands over an object to the robot, the robot identifies the status S_P of the person, and upon observing the demonstrated handover grasp configuration, it stores the relative giver (person)-receiver (robot) status along with the extracted grasp configuration into its knowledge base. The knowledge base is organized based on the relative giver-receiver status in addition to object type. Thus for each object type, the knowledge base holds grasp configuration information for each relative giver-receiver status. Subsequently, when executing a handover, the robot first identifies the user's status, and determines the relative giver (robot)-receiver (person) status. The robot then queries its knowledge base for the grasp configuration appropriate to the object and the relative giver-receiver status, and finally executes the handover using the returned grasp configuration.

6.5.4 Experiment Procedure

To test the extended implementation, this thesis uses the hierarchical social structure typically found in a university laboratory as an example. A laboratory commonly consists of doctor, master, and undergraduate students. In cultures where the sense of social hierarchy is more pronounced, doctor students have seniority and thus higher status than master students, who in turn have seniority and thus higher status than undergraduate students, and students of lower status tend to behave more formally towards student of higher status.

During handover demonstrations, an experimenter playing the role of master student demonstrated three different handover grasp configurations for a mug while varying the role of the robot among doctor, master, and undergraduate student to create the three cases of $S_{receiver} > S_{giver}$, $S_{receiver} = S_{giver}$, and $S_{receiver} < S_{giver}$. The experimenter demonstrated the handovers five times for each case. During handover execution, the experimenter then set the robot's role to master student, and asked the robot to hand over the mug to three different persons playing the roles of doctor, master, and undergraduate students. The achieved handover grasp configuration in each case are then evaluated by qualitatively comparing with the demonstrated grasp configurations.

6.5.5 Results

Figure 6-14 shows the handover grasp configurations demonstrated by the experimenter in each of the three cases. For the case $S_{giver} = S_{receiver}$, the experimenter handed over the mug presenting the handle to the receiver, grasping the opposite side of the mug, using a one-handed configuration. This grasp configuration is the same as the one demonstrated in Section 6.3. For the case $S_{giver} > S_{receiver}$, the experimenter handed over the mug grasping the handle, using a one-handed configuration. For the case $S_{giver} < S_{receiver}$, the experimenter handed over the mug presenting the handle to the receiver, grasping the opposite side of the mug, using a two-handed configuration. Figure 6-15 shows the handover orientations and grasp points extracted from each demonstration, as well as the computed means, while Table 6-6 gives the data.

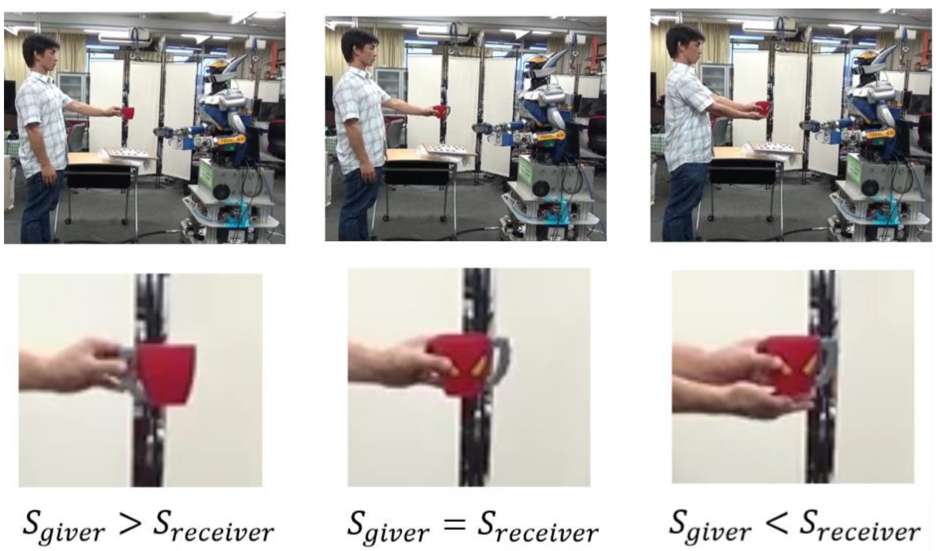

Figure 6-14 Grasp configurations demonstrated to the robot in each of the three cases. (To be published in Chan et al. 2015 [46].)

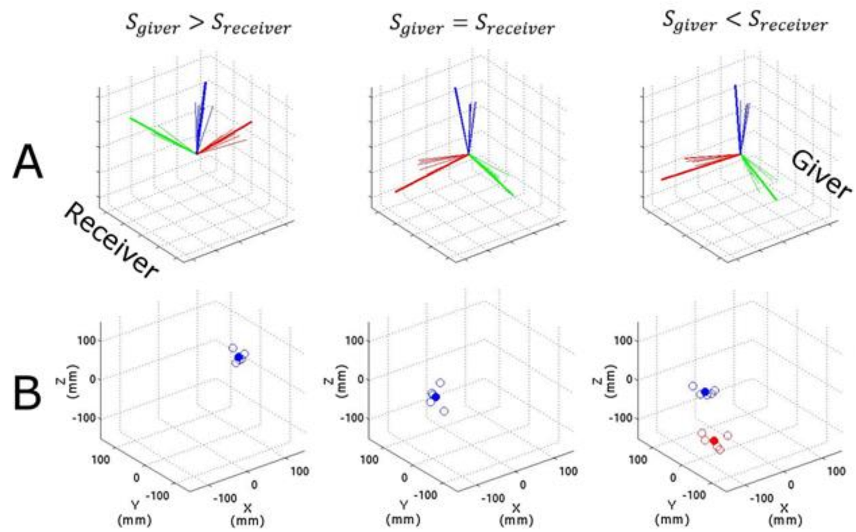

Figure 6-15 Handover grasp configurations extracted from all five demonstrations for each case. A shows object orientation, with thick long lines showing computed average. B shows Blue circles indicating grasp point, and red circles indicating position of the second hand where a two-handed handover is detected. Hollow circles show the data from each trial and solid circles show the computed average. (To be published in Chan et al. 2015 [46].)

Table 6-6 Handover configurations extracted from demonstrations and computed averages. (To be published in Chan et al. 2015 [46].)

	Trial	$C_{handover}$				P_{grasp} (mm)			P_{grasp2} (mm)		
		Angle (rad)	Axis								
$S_{giver} > S_{receiver}$	1	0.190	0.850	0.190	0.500	121.870	34.650	10.280	Null		
	2	0.260	0.240	-0.110	0.970	135.040	35.320	20.530			
	3	0.200	0.210	0.180	0.960	126.210	32.350	12.980			
	4	0.520	0.420	0.370	0.830	93.770	9.830	23.690			
	5	0.130	0.430	-0.100	0.900	114.660	46.850	37.270			
	Average	0.260	0.430	0.110	0.830	118.310	31.800	20.950			
$S_{giver} = S_{receiver}$	1	3.010	0.010	-0.010	1.000	-79.900	1.060	38.220	Null		
	2	3.040	-0.070	-0.010	-1.000	-95.500	4.370	10.450			
	3	2.990	0.030	0.020	1.000	-112.550	-12.720	29.100			
	4	3.050	0.000	-0.010	1.000	-88.430	-23.220	-21.420			
	5	3.000	-0.080	0.000	-1.000	-107.260	-2.290	-1.400			
	Average	3.020	-0.020	0.000	0.200	-96.730	-6.560	10.990			
$S_{giver} < S_{receiver}$	1	3.126	-0.063	0.024	-0.998	-61.137	3.947	10.055	-82.252	9.746	-113.249
	2	2.809	0.046	-0.025	0.999	-98.742	30.439	22.391	-39.961	-10.802	-106.967
	3	2.712	0.029	0.058	0.998	-72.831	14.476	-2.017	-91.163	-43.014	-102.186
	4	3.076	-0.078	0.029	-0.997	-95.441	8.769	8.928	-76.572	-32.801	-120.443
	5	2.840	0.009	-0.126	0.992	-62.298	14.232	-2.570	-78.251	25.230	-103.828
	Average	2.913	-0.011	-0.008	0.199	-78.090	14.373	7.358	-73.640	-10.328	-109.335

Figure 6-16Figure 6-14 shows the handover executions. From left to right, the figure shows HRP2V handing over the mug to a student playing the role of an undergraduate, master, and doctor student. During handover execution, the robot first identifies the person's status. Based on the relative status of the giver (robot) and receiver (person), HRP2V then selects the appropriate handover grasp configuration and executes the handover. Comparing with the demonstrated grasp configurations illustrated in Figure

6-14 shows that HRP2V was able to hand over the mug using a grasp configuration appropriate to each case depending on who the user is.

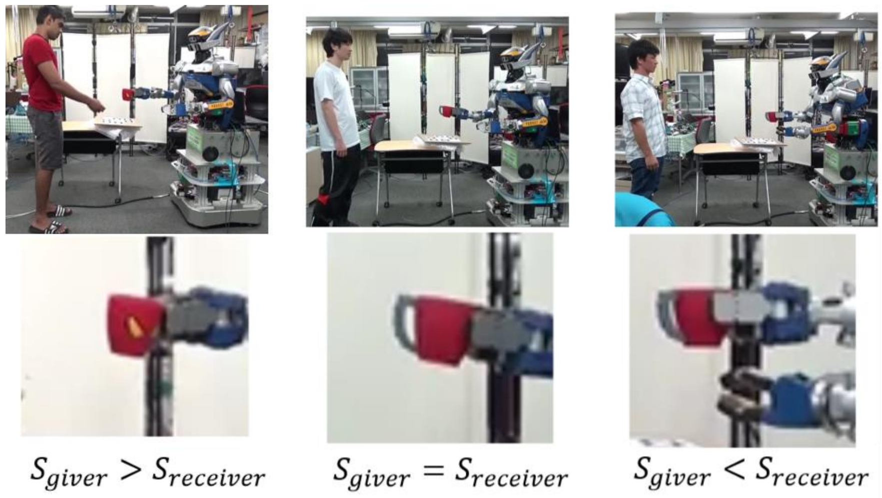

Figure 6-16 Handover execution. In each case, HRP2V handed over the mug using an appropirate grasp configuration matching the demonstrated one.

6.6 Integration into a Household Service Robot Application

As a demonstration of how the presented framework of this thesis can be utilized in a real world situation, this section presents an integration experiment where the work of this thesis is used to enhance a household service robot's capabilities in serving users. Daily assistive robotics is one of the research themes of the Jouhou System Kougaku (JSK) Laboratory. In one of its projects, the lab has developed a household robot application where a user can send commands from a tablet to a robot, and the robot autonomously performs several house cleaning tasks [133]. In the developed application, a PR2 robot performs the tasks of picking up an empty tray, bringing it to the kitchen counter, putting clothes into a washing machine, and sweeping the floor with a broom. Figure 6-17 depicts the tasks the robot performs. While the robot performs the tasks in an empty room without any humans, in a more realistic situation, there will be people who are living in the house present in the room. Household service robots will not only be expected to perform cleaning tasks in the absence of humans, but they will also be expected to performs other tasks such as serving drinks to people [38], and fetching the newspaper or TV remote controller [17], which involve physical human-robot interaction in the form of object handover.

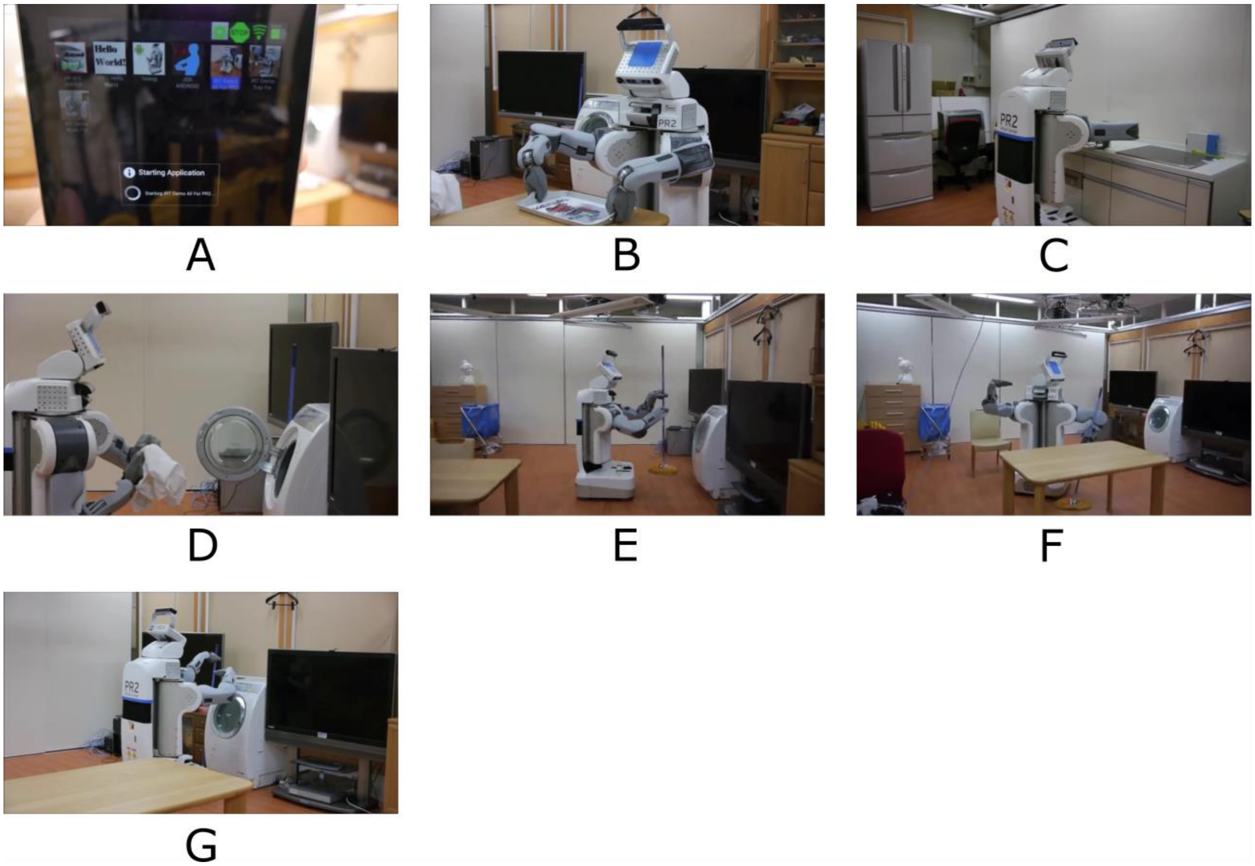

Figure 6-17 A household robot performs cleaning tasks. The users sends the command through a tablet device (A). The robot picks up an empty tray from the table (B), and brings the tray to the kitchen counter (C). Robot puts clothes into the washing machine (D), picks up the broom (E), sweeps underneath the table (F), and places broom back (G).

In order to perform handover tasks properly, a household robot must consider the affordances of the objects. Figure 6-18 and Figure 6-19 demonstrates two instances of improper handovers, where object affordances are neglected. In Figure 6-18, the robot uses a grasp configuration that emphasizes on grasp stability. The robot inserts its finger into the mug to grasp it, and hands over the mug sideways instead of upright. If the mug actually contained water, it would be spilled, and the robot's gripper will likely incur water damage. In Figure 6-19, the robot only considers its own convenience and hands over the remote upside down and with the buttons faced away from the receiver. While this does allow the receiver to comfortably take the remote with a more neutral wrist angle, the receiver needs to first turn the remote upright, then turn the remote again to see the buttons before being able to use it. The following will describe the integration of three handover tasks into the household service robot application presented in [133] by using the framework presented in this thesis to enable the robot to carry out the handover tasks properly, while demonstrating an additional feature of the framework, which is knowledge transfer among robots.

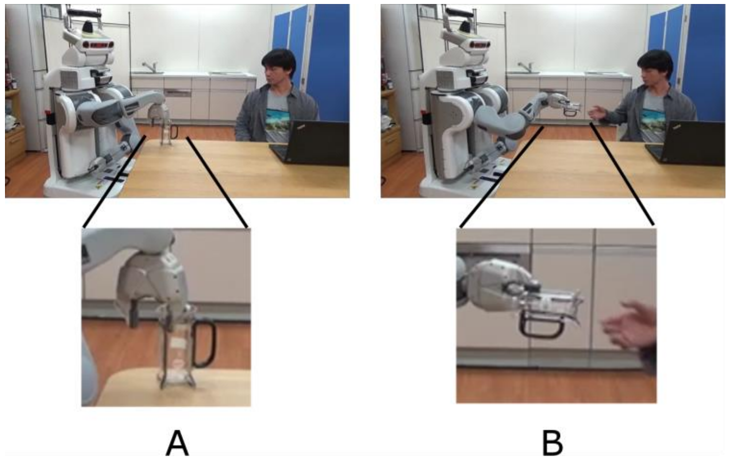

Figure 6-18 Improper handover of a mug. A – Robot inserts finger into mug. B – Robot hands over mug sideways.

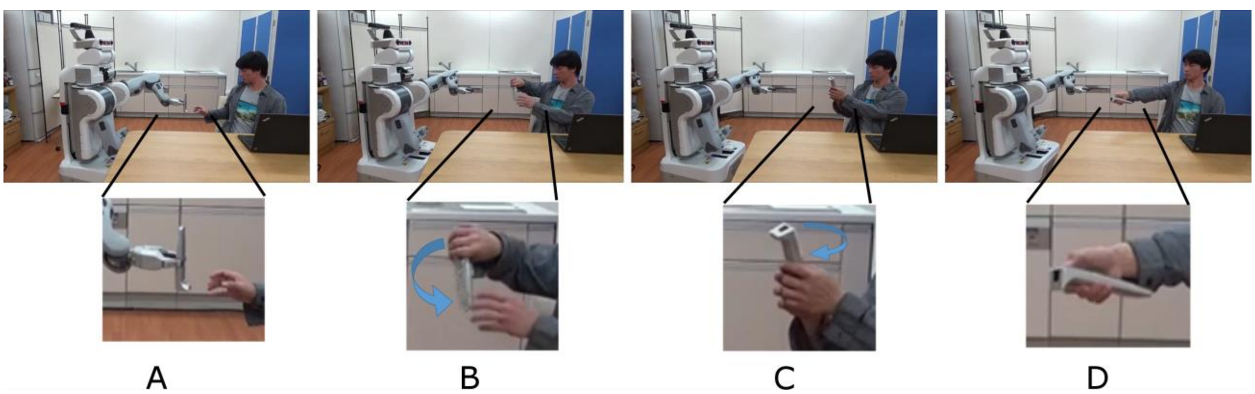

Figure 6-19 Improper handover of a remote. A – robot hands over remote upside down with buttons facing away from user. B – After taking the remote, the user needs to first turn the remote upright. C – User then needs the turn the remote around to see the buttons. D – Finally, user can use the remote.

One of the merits of the framework presented by this thesis is platform independence, and it allows knowledge learned by one robot to be transferred and used by other robots. For the current instance of integrating handover tasks into the household service robot in [133], the application is as such: first, the proper handover grasp configurations for three objects (a mug, a remote, and a kitchen knife) are demonstrated to an HRP2V robot. The HRP2V robot learns the proper grasp configurations for the objects from the demonstrations and builds a knowledge base. The knowledge is then transferred to the PR2 robot. Finally the PR2 robot utilizes the transferred knowledge to execute the handover tasks using proper grasp configurations (Figure 6-20).

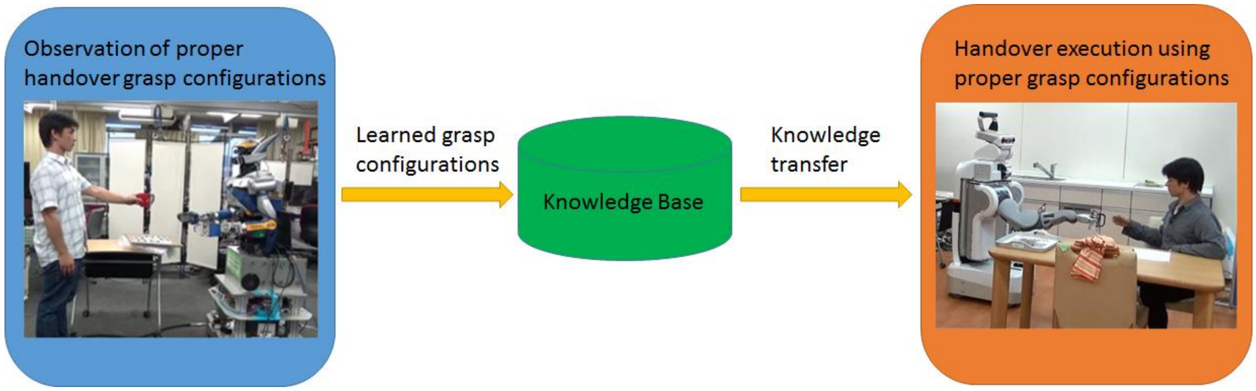

Figure 6-20 Knowledge transfer between robots. The handover grasp configurations learned by HRP2V are stored into a knowledge base. This knowledge is then transferred to PR2. PR2 then utilizes this transferred knowledge to execute handovers properly.

After PR2 obtained the transferred knowledge, the experimenter integrated three new handover tasks of a mug, remote, and kitchen knife into the household service robot application presented in [133] to expand PR2's capabilities. Figure 6-21 shows the execution results. With expanded capabilities, PR2 now works in a room inhabited by a person (Figure 6-21A). PR2 first proceeds to serve a drink to the person. Using the proper handover grasp configuration for a mug that was transferred to it from HRP2V, PR2 picks up the mug by grasping it from the side opposite to the handle (Figure 6-21B), and hands over the mug by presenting the handle to the person (Figure 6-21C). After serving the drink, it then cleans up the table by picking up the tray from the table (Figure 6-21D) and bringing it to the kitchen counter (Figure 6-21E). PR2 then picks up the TV remote controller from the table at the back corner of the room (Figure 6-21F) and brings it to the person. Using the transferred knowledge, PR2 hands over the remote by facing the buttons upwards, and directing the bottom of the remote towards the person (Figure 6-21G), thus allowing the person to conveniently use the remote immediately after receiving it. PR2 then proceeds to putting the clothes into the washing machine and sweeping the floor (Figure 6-21H). When PR2 finished sweeping the floor and returning the broom, the person moved to the kitchen and began to prepare food. PR2 then moves toward the kitchen as well to help with food preparation (Figure 6-21I). As the person brings out vegetables from a grocery bag, PR2 picks up a kitchen knife from the drawer, and once again, using the grasp configuration transferred from HPR2V, PR2 hands over the kitchen knife safely by presenting the handle and directing the knife tip away from the person (Figure 6-21J). As a result of the work of this thesis, PR2 is now able to extend the range of its capable tasks and hand over various objects to people properly. The results of this integration experiment demonstrates how the work of this thesis expands the capabilities of service robots.

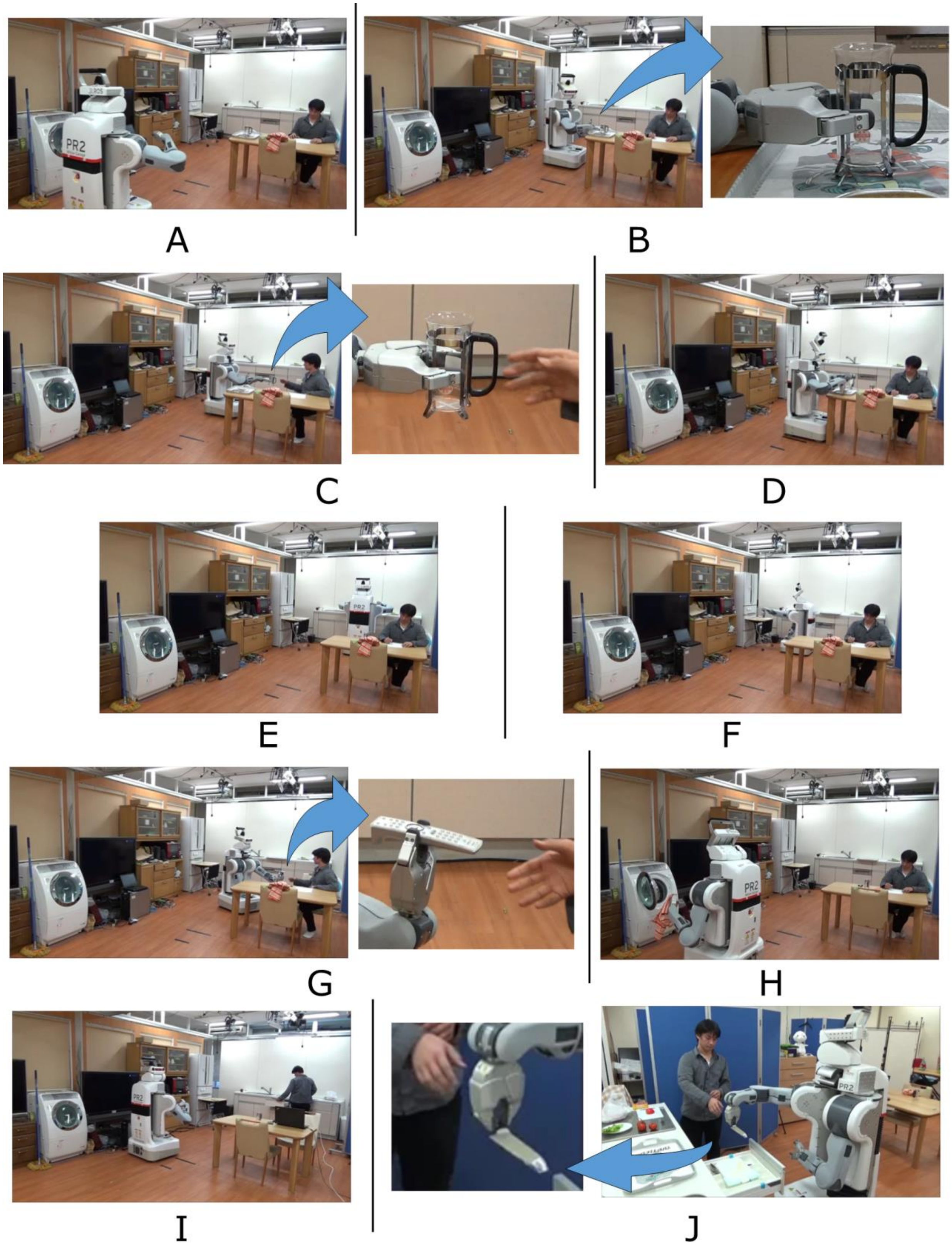

Figure 6-21 integration experiment. A – PR2 in room with person. B – Pick up mug. C – Hand over mug. D – Pick up tray. E – Place tray at counter. F – Pick up remote. G – Hand over remote. H – Place clothes in washing Machine. I – Move to kitchen. J – Hand over kitchen knife.

6.7 Discussion

This chapter has presented the implementation of the framework for learning proper handover grasp configurations through observations on an HRP2V robot. Validation experiments demonstrated that the implementation was successful and that HRP2V was capable of learning grasp configurations from interactions with humans and subsequently hand over objects using proper grasp configurations. Results from handover experiments with known objects showed that HRP2V was able to handover the objects using proper grasp configurations that matched the demonstrated ones. Furthermore, results from handover experiments with unknown objects showed that HRP2V was able to generalize the learned grasp configurations to unknown objects, and properly handover new objects using appropriate grasp configurations. Quantitative results showed that the error between the learned handover orientations and the measured handover orientations averaged at approximately 10 deg and 16 deg for the first and second set of objects respectively. Comparing these results with the natural variations found in handover orientations achieved by humans, δ_C, as shown in Table 4-3, the average δ_C for all the twenty objects is 27.1 deg, and the top three objects with the smallest variations were found to have a δ_C of around 10 deg. This shows that HRP2V was able to achieve a small enough error that is similar to the natural variations observed in human handovers.

Quantitative results on comparison of the achieved handover orientations with the learned handover orientations show that there is an error ranging from 0.45 deg to 37.2 deg. Observations during the experiments show that when the HRP2V reaches for the object on the table, its fingers sometimes comes in contact with the object and pushes the object slightly. This is due to inaccuracies in the detection of the object's position. Furthermore, as HRP2V closes its hand to grasp the object, due to the rigidity of its finger links, it forces the object's pose to change slightly. This change is not accounted for when the robot computes the required pose of its hand for handing over the object. The robot could improve the achieved handover grasp configurations by tracking the object's pose as it grasps the object or re-detecting the object's pose after it has grasped the object.

One of the major challenges in implementing the framework was in extracting the handover grasp configuration used by the giver from observations of handover demonstrations. Currently there does not exist a good method for identifying the orientation of an object held in a person's hand. The presented implementation overcame this challenge by using a different approach of first detecting the object at its pre-occluded state, and tracking the object using a particle filter. However, with this approach, if the giver comes into view with the object already in hand and hands over the object, the robot would not be able to identify the orientation of the object. The robot can potentially address this case by doing the following. When the giver hands over the object, the robot begins to record its camera data. After taking the object, the places it down and detects the pose of the object. Then, by working backwards in time and tracking the object's pose back to the point of handover, it can determine the grasp configuration used. This approach of sequential back-tracking can be found in existing study on detecting object positon [134].

Experiments with the skeleton tracker and object tracker showed that the skeleton tracker is quite computationally efficient and tracks quick human movements well. However, when the person is holding an object, the tracker tends to misidentify the end point of the object as the person's hand. The object tracker on the other hand, is robust to occlusions, but does not cope well with quick movements. Furthermore, when the object lacks distinct features, tracking results tend to drift. Figure 6-22 demonstrates two cases where the object tracker fails. In Figure 6-22A, the spray can has a cylindrical geometry similar to the person's arm, and after a while, the tracker drifts to the person's arm that is holding the object. In Figure 6-22B, the mug has similar dimensions in width, length, and height, and after some time, the track loses track of its orientation.

The implementation provided in this chapter has used separate trackers for tracking the human and the object. Using a combined approach and using information from each tracker to complement the other can potentially increase the accuracy of hand tracking and object tracking, and thus, improve the robot's performance on learning grasp configurations. A combined approach for object tracking and human tracking can be found in Micelli et al.'s work on human-robot handovers [88]. However, their system only identifies the object type and object position, but not its orientation. An improved accuracy in human tracking and object tracking would be important for cases where precise grasp point and object orientation is important, such as handing over a full cup of coffee or a kitchen knife.

Figure 6-22 Examples demonstrating some failures of the object tracker. (Published in Chan et al. 2015 [131].)

In determining the proper handover grasp configuration for an object from a set of observations, if the robot determines that the set of observations is of poor quality, it simply selects one of the grasp configurations from the set to use. A potential improvement might be to perform clustering on the observed handover grasp configurations to determine if multiple modes of handover grasp configurations exists, and then to compute the mean for each cluster.

Section 6.5 discussed how the implementation can be extended to take into consideration additional factors that influence handover grasp configuration and provided an example implementation. Experiment results showed that the robot was able to use different grasp configurations, appropriate to each case, to hand over the same object, depending on who the receiver is. While Section 6.5 have used user identity as an example, other factors can also be accounted for in a similar manner. For example, depending on the degree of formality of the occasion, individual user preference, and intended use of the object, the appropriate grasp configuration may also change. By modifying the grasp configuration selector to take in additional relevant inputs, the system can be further extended to account for these factors as well.

To evaluate the framework proposed in this thesis against other approaches, Table 6-7 provides a comparison of the proposed framework with existing works. Comparison shows that the framework presented in this thesis is the only one that captures object affordance information, does not rely on user provided object labels, can consider multiple grasp configurations for the same object by taking into account non-object related factors, and can generalize to new objects. While Kim et al.'s method can compute multiple grasp configurations, it lacks a method for determining which grasp configuration is to be used [45]. Cakmak et al.'s learning method does not rely on object labels, but instead relies on object specific grasp configurations examples labeled by users [89]. Their planning method, on the other hand, does not require object labels, and can theoretically handle new objects; however, it does not capture object affordance information well. Aleotti et al. briefly suggested a method for handling new objects, but the method has not been tested. The framework proposed in this thesis is the only method that addresses all aspects listed in Table 6-7. Comparing the performance of each method, the proposed framework has been able to handover a combined total of twenty known and unknown objects, which is four times as many as the next largest number. Thus the proposed framework has better performance over existing approaches considering the criteria shown in Table 6-7.

Table 6-7 Comparison of proposed framework with existing works.

	Kim et al. 2004 [45]	Cakmak et al. 2011 [89] (learning)	Cakmak et al. 2011 [89] (planning)	Aleotti et al. 2012 [57]	Aleotti et al. 2014 [47]	Proposed framework
Total number of objects handed over	0 (simulation only)	5	5	2	3	20
Captures affordance information	✓	✓	✗	✓	✓	✓
Rely on user proved object labels	✗	✓	✓	✗	✗	✓
Multiple grasp configurations for object	✓	✗	✗	✗	✗	✓
Generalization to new objects	✗	✗	✓	✗	✓	✓

6.8 Summary

This chapter has presented an implementation of the framework Chapter 3 presented for learning handover grasp configurations from observations on an HRP2V robot. In the implementation, when extracting grasp configurations from observed handover demonstrations, the robot determines the grasp point by tracking the giver's hand with a skeleton tracker and identifying where the giver grasps to pick up the object. For determining the object's orientation at handover, the robot first detects the object pose before it is held and occluded by the giver's hand, and tracks the object using a particle filter. Experiments validated that the robot was indeed able to learn handover grasp configurations for multiple objects, and subsequently handover the objects using the proper grasp configurations. Results also demonstrated the robot's ability to generalize learned grasp configurations to new objects. The presented system offers the advantages of not requiring any external cameras nor any markers attached to the person or object when learning grasp configurations from handover interactions with humans. The system presented in this chapter is the first demonstrated system that allows robots to automatically learn proper handover grasp configurations from interactions with humans, and determine proper grasp configurations for new objects.

7 Conclusion

This thesis has presented a framework for enabling robots to automatically learn proper object handovers. The approach of the framework is to learn from interactions with humans and from observations of handovers and object usages. As robots enter the society, they will have more chances of interacting with users and observing how objects are used around them. By enabling robots to learn from these surrounding events, the framework frees the user from having to program proper handover grasp configurations for each object to the robot.

Chapter 3 first presented the idea of the framework. The framework learns grasp configurations by observing how humans hand over the objects, and computes handover grasp configurations for new objects by generalizing learned configurations based on observed similarities in object usages. To provide the required components for implementing the framework, Chapter 4 presented a user study conducted to characterize the handover orientations used by humans and devised a method for computing proper handover orientations from observations, and Chapter 5 presented a method for grouping and classifying objects based on their observed usages. Using these components, Chapter 6 then presented an implementation of the framework onto a robot hardware and demonstrated the ability of the framework on learning proper handover grasp configurations. Finally, the chapter also showed how this work can be used in a real world scenario by integration into a household service robot application.

7.1 Framework for Learning Handover Grasp Configurations

To allow robots to hand over objects properly, existing methods require users to explicitly program or teach the proper grasp configurations of each object to the robot. These type of approach lack scalability and generalizability. The framework Chapter 3 presented aims to address these problems. As robots enter the society, they will have many opportunities to observe how humans use and handover various objects. This framework takes advantage of this by enabling robots to learn from observing these natural occurrences of handover and object usage. By observing the handover grasp configurations used by human givers, the robot learns grasp configurations of objects directly, and by observing how humans used various objects, it learns the affordance information of objects, which allows it to generalize learned grasp configurations to new objects.

7.2 Handover Grasp Configurations used by Humans

To be able to learn handover grasp configurations from observations, Chapter 4 presented a user study that surveyed and characterized the handover orientations used by human givers, and devised a method for computing mean handover orientations. The user study compared the handover orientations used by humans in three different conditions and found that the grasp configurations human givers use may change depending on whether they have their focus on themselves or on the receiver. The Chapter puts forth the novel concept of affordance axes for identifying patterns in handover orientations and offered a definition for computing them. Chapter 4 also presented a distance minimization method for computing mean handover orientations. The work presented in the chapter provided the first building block for implementing the framework presented in Chapter 3.

7.3 Affordance Based Object Grouping and Classification

Chapter 5 presented an object grouping and classification method that addresses the issue of generalizability of learned handover grasp configurations. Since proper handover grasp configurations of objects depend on the affordances of the objects, the proposed method focuses on object affordances. As object affordances are revealed during usage of the object, the propose method learns about objects by observing how humans use them. The algorithm extracts a set of action features from object movements and inter-object interactions during usages. It then uses the extracted action feature vector to group and classify objects. The experiment presented in the same chapter shows that the proposed method is capable of grouping and classifying objects properly according their affordances for the purpose of determining the objects' proper handover grasp configurations.

7.4 Learning from Human-Robot Interactions

Using the components presented in Chapter 4 and Chapter 5, Chapter 6 presented the implementation of the framework presented in Chapter 3. The implementation utilized recent advancements in skeleton tracking to track a human giver's body motions and identify handover cues. To overcome the major challenges of recognizing a held object's pose so that the robot can recognize the grasp configuration used by the giver in a handover, this thesis used a probabilistic predictive approach in which the object is first detected at the pre-occlusion state, and its pose is then tracked using a particle filter. Experiments demonstrated that using these methods, the implementation was able to learn handover grasp configurations by participating in human-robot handovers and observing the grasp configurations used by the human giver, and subsequently, able to hand over the objects properly using the grasp configurations learned. Furthermore, results also showed that the implementation was able to generalize the learned grasp configurations to new objects. An integration experiment finally demonstrated how grasp configurations learned by one robot can be transferred and used by other robots, and how the work of this thesis can be used in a real world household scenario.

7.5 Contributions

This thesis has presented the following contributions: First, this thesis has presented a framework for enabling robots to autonomously determine proper handover grasp configurations for various objects. The framework addresses both scalability and generalizability. Second, the user study on characterizing human handover grasp configurations has provided a better understanding of the grasp configurations used by human givers. At the same time, the study has also introduced the novel notion of object affordance axes, and a method for computing an appropriate handover orientation from a set of observations. Third, with regards to generalizability of learned grasp configurations, this thesis has presented an affordance focused method for grouping and classifying objects. Finally, this thesis has provided an implementation of the framework on a robot hardware, overcoming challenges of robot perception, and demonstrated the abilities of the framework.

7.6 Future Work

This thesis has documented a demonstrated approach for enabling robots to hand over objects to people using proper grasp configurations. However, there are still areas for improvement. The experiments presented in this thesis has dealt with common objects and tools which are small and light enough for a person to pick up using only their arms. To expand robots' capabilities of handing over objects, a possible extension would be to consider large, heavy objects. For example, when picking up and handing over a large heavy box, humans utilize not only their arms, but also their torso and their legs. Utilization of the whole body not only allows the giver to generate the required amount of force to handle the object, but the giver's body motion also communicates the properties of the object he/she is about the hand over to the receiver, thus allowing the receiver to better prepare himself/herself for the transfer of the object [135]. Enabling robots to observe and learn full body motion in handovers can allow robots to handle an even wider range of objects and situations.

Towards enabling better robot-human handovers, this thesis approached the problem from a grasp configuration aspect. However, there are many other aspects of handovers such as gaze [79], body posture and motion [56], [84], and arm movement as well [42], [55], [86]. Another direction for future work would involve devising a framework for learning these other aspects of handovers from observation of the surroundings and interactions with users in robots' daily operations.

Bibliography

[1] Y. Xu, W. Chan, T. W. Law, S. K. Tso, C. Y. Wang, W. Fung, Y. Leung, and M. Chow, "Development of a hospital service robot for transporting task," in *International Conference on Robotics, Intelligent Systems and Signal Processing*, 2003, pp. 628–633.

[2] P. Akella and M. Peshkin, "Cobots for the automobile assembly line," in *International Conference on Robotics & Automation*, 1999, pp. 728–733.

[3] R. Wilcox, S. Nikolaidis, and J. Shah, "Optimization of Temporal Dynamics for Adaptive Human-Robot Interaction in Assembly Manufacturing," in *Robotics: Science and Systems*, 2012.

[4] F. Wallhoff, J. Blume, A. Bannat, W. Rösel, C. Lenz, and A. Knoll, "A skill-based approach towards hybrid assembly," *Adv. Eng. Informatics*, vol. 24, no. 3, pp. 329–339, 2010.

[5] S. Satake, T. Kanda, D. F. Glas, M. Imai, H. Ishiguro, and N. Hagita, "How to Approach Humans?-Strategies for Social Robots to Initiate Interaction-," in *Human Robot Interaction*, 2010, pp. 109–116.

[6] C. Shi, M. Shiomi, C. Smith, T. Kanda, and H. Ishiguro, "A Model of Distributional Handing Interaction for a Mobile Robot," in *Robotics: Science and Systems*, 2013.

[7] D. Brščić, H. Kidokoro, Y. Suehiro, and T. Kanda, "Escaping from Children's Abuse of Social Robots," in *International Conference on Human-Robot Interaction*, 2015, pp. 59–66.

[8] K. Yamazaki, M. Kawashima, Y. Kuno, N. Akiya, M. Burdelski, A. Yamazaki, and H. Kuzuoka, "Prior-to-request and request behaviors within elderly day care : Implications for developing service robots for use in multiparty settings," in *European Conference on Computer Supported Cooperative Work*, 2007, pp. 24–28.

[9] Y. Iwamura, M. Shiomi, T. Kanda, H. Ishiguro, and N. Hagita, "Do Elderly People Prefer a Conversational Humanoid as a Shopping Assistant Partner in Supermarkets?," in *International Conference on Human-Robot Interaction*, 2011, pp. 449–456.

[10] Y.-H. Kao and W.-J. Wang, "Design and Implementation of a Family Robot," in *International Joint Conference on Computer Science and Software Engineering*, 2012, pp. 218–225.

[11] T. Shibata, Y. Kawaguchi, and K. Wada, “Investigation on people living with Paro at home effects of sex difference and owners’ animal preference,” in *International Workshop on Robot and Human Interactive Communication*, 2009, pp. 1131–1136.

[12] K. Wada, Y. Ikeda, K. Inoue, and R. Uehara, “Development and Preliminary Evaluation of a Caregiver’s Manual for Robot Therapy using the Therapeutic Seal Robot Paro,” in *International Workshop on Robot and Human Interactive Communication*, 2010, pp. 533–538.

[13] W. Wannasuphoprasit, P. Akella, M. Peshkin, and J. E. Colgate, “Cobots: A novel material handling technology,” in *International Mechanical Engineering Congress and Exposition*, 1998.

[14] United Nations, “World Population Ageing: 1950-2050,” Department of Economic and Social Affairs Population Division, New York, 2002.

[15] J. A. Mcmullin, M. Cooke, and R. Downie, “Labour Force Ageing and Skill Shortages in Canada and Ontario,” Canadian Policy Research Networks Inc., London, Ontario, 2004.

[16] Z. Yubo, W. Zixi, J. Linhong, and B. Sheng, “The Clinical Application of the Upper Extremity Compound Movements Rehabilitation Training Robot,” in *International Conference on Rehabilitation Robotics*, 2005, pp. 91–94.

[17] Y. S. Choi, T. Chen, A. Jain, C. Anderson, J. D. Glass, and C. C. Kemp, “Hand It Over or Set It Down: A User Study of Object Delivery with an Assistive Mobile Manipulator,” in *International Symposium on Robot and Human Interactive Communication*, 2009, pp. 736–743.

[18] K. Yamazaki, A. Yamazaki, M. Okada, and Y. Kuno, “Revealing Gauguin: Engaging Visitors in Robot Guide’s Explanation in an Art Museum,” in *Conference on Human Factors in Computing Systems*, 2009, pp. 1437–1446.

[19] M. Diftler, J. Mehling, and M. Abdallah, “Robonaut 2 - the first humanoid robot in space,” in *International Conference on Robotics and Automation*, 2011, pp. 2178–2183.

[20] Natural Sciences and Engineering Research Council of Canada, “NSERC Investments in Manufacturing (2010-11),” Canada, 2012.

[21] Robotics Trends’ News Sources, “Robotic Assistants May Adapt to Humans in the Factory,” *Robotics Trends*, 2012. [Online]. Available: http://www.roboticstrends.com/industry_manufacturing/article/robotic_assistants_may_adapt_to_humans_in_the_factory/?utm_source=RoboTrendsWeekly&utm_medium=email.

[22] D. Kara, "Thinking Twice On Rethink Robotics," *Robotics Business Review*, 2012. [Online]. Available: http://www.roboticsbusinessreview.com/article/thinking_twice_on_rethink_robotics.

[23] MarketsandMarkets, "Service Robotics Market (Personal & Professional) – Global Forecast & Assessment by Applications & Geography (2012 – 2017)," 2012.

[24] Y. Kobayashi, T. Shibata, Y. Hoshi, Y. Kuno, M. Okada, and K. Yamazaki, "'I will ask you' Choosing Answerers by Observing Gaze Responses using Integrated Sensors for Museum Guide Robots," in *International Symposium on Robot and Human Interactive Communication*, 2010, pp. 687–692.

[25] K. Takeuchi, T. Yoshimi, K. Hashimoto, and T. Tanaka, "Consideration and Evaluation of benefits of human shaped service robots [The report of experimental introduction to a shopping center]," in *Workshop on Advanced Robotics and its Social Impacts*, 2013, pp. 157–163.

[26] H. Arai, T. Takubo, Y. Hayashibara, and K. Tanie, "Human-Robot Cooperative Manipulation Using a Virtual Nonholonomic Constraint.pdf," in *International Conference on Robotics and Automation*, 2000, pp. 4063–4069.

[27] M. M. Rahman, R. Ikeura, and K. Mizutani, "Investigation of the impedance characteristic of human arm for development of robots to cooperate with humans," in *JSME International Journal Series C*, 2002, vol. 45, no. 2, pp. 510–518.

[28] A. Koene, S. Endo, A. Remazeilles, M. Prada, and A. M. Wing, "Experimental testing of the CogLaboration prototype system for fluent Human-Robot object handover interactions," in *International Symposium on Robot and Human Interactive Communication*, 2014, pp. 249–254.

[29] K. Strabala, M. K. Lee, A. Dragan, J. Forlizzi, S. S. Srinivasa, M. Cakmak, and V. Micelli, "Towards Seamless Human-Robot Handovers," *J. Human-Robot Interact.*, vol. 1, no. 1, pp. 1–23, 2013.

[30] T. Takeda, S. Member, Y. Hirata, and K. Kosuge, "Dance Step Estimation Method Based on HMM for Dance Partner Robot," in *Transactions on Industrial Electronics*, 2007, vol. 54, no. 2, pp. 699–706.

[31] E. Short, J. Hart, M. Vu, and B. Scassellati, "No fair!! An interaction with a cheating robot," in *International Conference on Human-Robot Interaction*, 2010, pp. 219–226.

[32] T. Fincannon, L. E. Barnes, R. R. Murphy, and D. L. Riddle, "Evidence of the Need for Social Intelligence in Rescue Robots," in *International Conference on Intelligent Robots and Systems*, 2004, pp. 1089–1095.

[33] P. H. Kahn, R. L. Severson, T. Kanda, H. Ishiguro, B. T. Gill, J. H. Ruckert, S. Shen, H. E. Gary, A. L. Reichert, and N. G. Freier, "Do people hold a humanoid robot morally accountable for the harm it causes?," in *International conference on Human-Robot Interaction - HRI '12*, 2012, p. 33.

[34] M. Huber, M. Rickert, A. Knoll, T. Brandt, and S. Glasauer, "Human-Robot Interaction in Handing-Over Tasks," in *International Symposium on Robot and Human Interactive Communication*, 2008, pp. 107–112.

[35] W. P. Chan, C. A. C. Parker, H. F. M. Van der Loos, and E. A. Croft, "A Human-Inspired Object Handover Controller," *Int. J. Rob. Res.*, vol. 32, no. 8, pp. 971–983, Jul. 2013.

[36] A. Moon, C. A. C. Parker, E. A. Croft, and H. F. M. Van der Loos, "Did you see it hesitate? - Empirically grounded design of hesitation trajectories for collaborative robots," in *International Conference on Intelligent Robots and Systems*, 2011, pp. 1994–1999.

[37] S. Sheikholeslami, A. Moon, and E. A. Croft, "Exploring the Effect of Robot Hand Configurations in Directional Gestures for Human-Robot Interaction," in *International Conference on Intelligent Robots and Systems*, 2015, pp. 3594–3599.

[38] J. Bohren, R. B. Rusu, E. G. Jones, E. Marder-Eppstein, C. Pantofaru, M. Wise, M. Lorenz, W. Meeussen, and S. Holzer, "Towards Autonomous Robotic Butlers: Lessons Learned with the PR2," in *International Conference on Robotics and Automation*, 2011, pp. 5568 – 5575.

[39] J. W. Hart, S. Sheikholeslami, M. K. X. J. Pan, W. P. Chan, and E. A. Croft, "Predictions of Human Task Performance and Handover Trajectories for Human-Robot Interaction," in *HRI 2015 Workshop on Human-Robot Teaming*, 2015.

[40] D. F. Franklin, R. E. Kahn, M. J. Swain, and R. J. Firby, "Happy patrons make better tippers: creating a robot waiter using Perseus and the Animate Agent architecture," in *International Conference on Automatic Face and Gesture Recognition*, 1996, pp. 253–258.

[41] M. K. Lee, J. Forlizzi, P. E. Rybski, F. Crabbe, W. Chung, J. Finkle, E. Glaser, and S. Kiesler, "The snackbot: documenting the design of a robot for long-term human-robot interaction," in *International Conference on Human Robot Interaction*, 2009, pp. 7–14.

[42] K. Yamane, M. Revfi, and T. Asfour, "Synthesizing object receiving motions of humanoid robots with human motion database," in *International Conference on Robotics and Automation*, 2013, pp. 1629–1636.

[43] W. P. Chan, C. A. C. Parker, H. F. M. Van der Loos, and E. A. Croft, "Grip Forces and Load Forces in Handovers: Implications for Designing Human-Robot Handover Controllers," in *International Conference on Human-Robot Interaction*, 2012, pp. 9–16.

[44] W. P. Chan, I. Kumagai, S. Nozawa, Y. Kakiuchi, K. Okada, and M. Inaba, “Creating socially acceptable robots: Leaning grasp configurations for object handovers from demonstrations,” in *Workshop on Advanced Robotics and its Social Impacts*, 2013, pp. 94–99.

[45] J. Kim, J. Park, Y. K. Hwang, and M. Lee, “Advanced Grasp Planning for Handover Operation Between Human and Robot: Three Handover Methods in Esteem Etiquettes Using Dual Arms and Hands of Home-Service Robot,” in *International Conference on Autonomous Robots and Agents*, 2004, pp. 34–39.

[46] W. P. Chan and M. Inaba, “Status-Based Behavior Selection for Creating Culturally Aware Robots,” in *International Workshop on Culture Aware Robotics*, 2015.

[47] J. Aleotti, V. Micelli, and S. Caselli, “An Affordance Sensitive System for Robot to Human Object Handover,” *Int. J. Soc. Robot.*, Jun. 2014.

[48] J. J. Gibson, “The Theory of Affordances,” in *The Ecological Approach to Visual Perception: Classic Edition*, Hillsdale, New Jersey: Lawrence Erlbaum Associates Inc., 1979.

[49] D. Norman, *Design of Everyday Things: Revised and Expanded*. United States of America: Basic Books, 2002.

[50] R. Jain, T. Graduate, A. Studies, and T. Inamura, “Robot Learning of Usage of Tools based on Bayesian Learning of Tool Affordances,” in *Conference of Robotics Society of Japan*, 2013, pp. 2S2–08.

[51] R. Jain and T. Inamura, “Bayesian learning of tool affordances based on generalization of functional feature to estimate effects of unseen tools,” *Artif. Life Robot.*, vol. 18, no. 1–2, pp. 95–103, Sep. 2013.

[52] R. Detry, D. Kraft, A. G. Buch, N. Kruger, and J. Piater, “Refining grasp affordance models by experience,” in *International Conference on Robotics and Automation*, 2010, pp. 2287–2293.

[53] C.-M. Huang and A. L. Thomaz, “Effects of Responding to, Initiating and Ensuring Joint Attention in Human-Robot Interaction,” *2011 Ro-Man*, pp. 65–71, Jul. 2011.

[54] S. Kiesler, A. Powers, and J. Goetz, “Matching robot appearance and behavior to tasks to improve human-robot cooperation,” in *International Workshop on Robot and Human Interactive Communication*, 2003, pp. 55–60.

[55] P. Basili, M. Huber, T. Brandt, S. Hirche, and S. Glasauer, “Investigating Human-Human Approach and Hand-Over,” in *Human Centered Robot Systems*, 2009, vol. 6, pp. 151–160.

[56] M. Cakmak, S. S. Srinivasa, M. K. Lee, S. Kiesler, and J. Forlizzi, “Using Spatial and Temporal Contrast for Fluent Robot-Human Hand-overs,” in *International Conference on Human Robot Interaction*, 2011, pp. 489–496.

[57] J. Aleotti, V. Micelli, and S. Caselli, “Comfortable robot to human object hand-over,” in *International Workshop on Robot and Human Interactive Communication*, 2012, pp. 771–776.

[58] A. Edsinger and C. C. Kemp, “Human-Robot Interaction for Cooperative Manipulation: Handing Objects to One Another,” in *International Symposium on Robot and Human Interactive Communication*, 2007, pp. 1167–1172.

[59] J. Kim, J. Park, Y. K. Hwang, and M. J. Lee, “Advanced grasp planning for handover operation between human and robot,” *2nd Int. Conf. Auton. Robot. Agents*, no. c, pp. 34–39, 2004.

[60] Willow Garage, “Beer Me, Robot,” 2010. [Online]. Available: http://www.willowgarage.com/blog/2010/07/06/beer-me-robot. [Accessed: 13-Jul-2012].

[61] Y. Ho, E. Sato-Shimokawara, T. Yamaguchi, and N. Tagawa, “Interaction Robot System Considering Culture Differences,” in *Workshop on Advanced Robotics and its Social Impacts (ARSO)*, 2013, pp. 170–174.

[62] J. Zhen, H. Aoki, E. Sato-Shimokawara, and T. Yamaguchi, “Sino-japanese culture differences in interaction robot system,” *Int. Symp. Syst. Integr. SII 2012*, pp. 295–300, 2012.

[63] L. Wang, P. P.-L. Rau, V. Evers, B. K. Robinson, and P. Hinds, “When in Rome: the role of culture & context in adherence to robot recommendations,” in *International Conference on Human-Robot Interaction*, 2010, pp. 359–366.

[64] G. Trovato, M. Zecca, S. Sessa, L. Jamone, J. Ham, K. Hashimoto, and A. Takanishi, “Towards culture-specific robot customisation: A study on greeting interaction with Egyptians,” in *International Workshop on Robot and Human Interactive Communication*, 2013, pp. 447–452.

[65] G. Eresha, M. Haring, B. Endrass, E. Andre, and M. Obaid, “Investigating the influence of culture on proxemic behaviors for humanoid robots,” in *International Symposium on Robot and Human Interactive Communication*, 2013, pp. 430–435.

[66] V. Navalpakkam, J. Rao, and M. Slaney, “Using Gaze Patterns to Study and Predict Reading Struggles due to Distraction,” in *Extended Abstracts on Human Factors in Computing Systems*, 2011, pp. 1705–1710.

[67] B. Fischer and B. Breitmeyer, “Mechanisms of visual attention revealed by saccadic eye movements,” *Neuropsychologia*, vol. 25, no. 1A, pp. 73–83, 1987.

[68] K. D. Robertshaw and R. M. Wilkie, “Does gaze influence steering around a bend?,” *J. Vis.*, vol. 8, no. 2008, pp. 1–13, 2008.

[69] R. M. Wilkie and J. P. Wann, “Driving as night falls: The contribution of retinal flow and visual direction to the control of steering,” *Curr. Biol.*, vol. 12, no. 23, pp. 2014–2017, 2002.

[70] D. L. Mann, W. Spratford, and B. Abernethy, “The Head Tracks and Gaze Predicts: How the World’s Best Batters Hit a Ball,” *PLoS One*, vol. 8, no. 3, 2013.

[71] C. L. Sidner, C. Lee, C. D. Kidd, N. Lesh, and C. Rich, “Explorations in Engagement for Humans and Robots,” *Artif. Intell.*, vol. 166, no. 1–2, pp. 140–164, 2005.

[72] M. Bennewitz, F. Faber, D. Joho, M. Schreiber, and S. Behnke, “Towards a Humanoid Museum Guide Robot that Interacts with Multiple Persons,” in *International Conference on Humanoid Robots*, 2005, pp. 418–423.

[73] M. Häring, J. Eichberg, and E. André, “Studies on Grounding with Gaze and Pointing Gestures in Human-Robot-Interaction,” in *International Conference on Social Robotics*, 2012, pp. 378–387.

[74] D. E. Karreman, G. U. Sepúlveda Bradford, E. M. a. G. van Dijk, M. Lohse, and V. Evers, “What Happens When a Robot Favors Someone?,” in *International Conference on Human-Robot Interaction*, 2013, pp. 157–158.

[75] M. Imai, T. Ono, and H. Ishiguro, “Physical relation and expression: Joint attention for human-robot interaction,” in *International Workshop on Robot and Human Interactive Communication*, 2001, vol. 50, no. 4, pp. 512–517.

[76] M. Zheng, A. Moon, B. Gleeson, D. M. Troniak, M. K. X. J. Pan, A. Benjamin, E. a Croft, and M. Q. H. Meng, “Human Behavioural Responses to Robot Head Gaze during Human-Robot Handovers,” in *International Conference on Robotics and Biomimetics*, 2014, p. 6.

[77] C. Rich, B. Ponsler, A. Holroyd, and C. L. Sidner, “Recognizing engagement in human-robot interaction,” in *International Conference on Human-Robot Interaction*, 2010, pp. 375–382.

[78] K. Strabala, M. K. Lee, A. Dragan, J. Forlizzi, and S. S. Srinivasa, “Learning the Communication of Intent Prior to Physical Collaboration,” in *International Symposium on Robot and Human Interactive Communication*, 2012, pp. 968–973.

[79] A. Moon, D. Troniak, and B. Gleeson, "Meet me where I'm gazing: how shared attention gaze affects human-robot handover timing," in *International Conference on Human-Robot Interaction*, 2014, pp. 334–341.

[80] M. K. Lee, J. Forlizzi, S. Kiesler, M. Cakmak, and S. Srinivasa, "Predictability or adaptivity?: designing robot handoffs modeled from trained dogs and people," in *Human Robot Interaction*, 2011, pp. 179–180.

[81] M. Gharbi, P.-V. Paubel, A. Clodic, O. Carreras, R. Alami, and J.-M. Cellier, "Toward a better understanding of the communication cues involved in a human-robot object transfer," in *Symposium on Robot and Human Interactive Communication*, 2015, pp. 319–324.

[82] K. L. Koay, E. A. Sisbot, D. S. Syrdal, M. L. Walters, K. Dautenhahn, and R. Alami, "Exploratory Studies of a Robot Approaching a Person in the Context of Handing Over an Object," in *AAAI Spring Symposium: Multidisciplinary Collaboration for Socially Assistive Robotics*, 2007, vol. 33, pp. 18–24.

[83] M. L. Walters, K. Dautenhahn, S. N. Woods, and K. L. Koay, "Robotic Etiquette: Results from User Studies Involving a Fetch and Carry Task," in *International Conference on Human-Robot Interaction*, 2007, pp. 317–324.

[84] K. Dautenhahn, M. Walters, S. Woods, K. L. Koay, C. L. Nehaniv, A. Sisbot, R. Alami, T. Siméon, C. Lane, and E. A. Sisbot, "How may I serve you? A robot companion approaching a seated person in a helping context," in *Human-Robot Interaction*, 2006, no. April 2005, p. 172.

[85] T. Flash and N. Hogan, "The coordination of arm movements: an experimentally confirmed mathematical model," *J. Neurosci.*, vol. 5, no. 7, pp. 1688–1703, 1985.

[86] S. Shibata, B. M. Sahbi, K. Tanaka, and A. Shimizu, "An Analysis of the Process of Handing Over an Object and its Application to Robot Motions," in *International Conference on Systems, Man, and Cybernetics*, 1997, pp. 64–69.

[87] S. Kajikawa, T. Okino, K. Ohba, and H. Inooka, "Motion planning for hand-over between human and robot," in *International Conference on Intelligent Robots and Systems. Human Robot Interaction and Cooperative Robots*, pp. 193–199.

[88] V. Micelli, K. Strabala, and S. Srinivasa, "Perception and Control Challenges for Effective Human-Robot," in *Robotics: Science and Systems*, 2011.

[89] M. Cakmak, S. S. Srinivasa, M. K. Lee, J. Forlizzi, and S. Kiesler, "Human Preferences for Robot-Human Hand-over Configurations," in *International Conference on Intelligent Robots and Systems*, 2011, pp. 1986–1993.

[90] R. S. Johansson and J. R. Flanagan, “Coding and use of tactile signals from the fingertips in object manipulation tasks.,” *Nat. Rev. Neurosci.*, vol. 10, no. 5, pp. 345–59, May 2009.

[91] R. S. Johansson and G. Westling, “Roles of Glabrous Skin Receptors and Sensorimotor Memory in Automatic Control of Precision Grip when Lifting Rougher or More Slippery Objects,” *Exp. Brain Res.*, vol. 56, no. 3, pp. 550–564, 1984.

[92] R. S. Johansson and G. Westling, “Signals in tactile afferents from the fingers eliciting adaptive motor responses during precision grip,” *Exp. Brain Res.*, vol. 66, pp. 141–154, 1987.

[93] M. Simoneau, J. Paillard, C. Bard, N. Teasdale, O. Martin, M. Fleury, and Y. Lamarre, “Role of the feedforward command and reafferent information in the coordination of a passing prehension task.,” *Exp. brain Res.*, vol. 128, no. 1–2, pp. 236–42, Sep. 1999.

[94] J. Monzée, Y. Lamarre, and A. M. A. Smith, “The effects of digital anesthesia on force control using a precision grip.,” *J. Neurophysiol.*, vol. 89, no. 2, pp. 672–83, Feb. 2003.

[95] J. Hermsdörfer, C. Marquardt, J. Philipp, A. Zierdt, D. Nowak, S. Glasauer, N. Mai, and Z. D. Nowak, “Moving weightless objects Grip force control during microgravity,” *Exp. Brain Res.*, vol. 132, pp. 52–64, 2000.

[96] R. S. Johansson and G. Westling, “Programmed and triggered actions to rapid load changes during precision grip,” *Exp. Brain Res.*, vol. 71, pp. 72–86, 1988.

[97] J. R. Flanagan, J. Tresilian, and A. M. Wing, “Coupling of grip force and load force during arm movements with grasped objects,” *Neurosci. Lett.*, vol. 152, no. 1–2, pp. 53–6, Apr. 1993.

[98] A. M. Wing and J. R. Flanagan, “Anticipating dynamic loads in handling objects,” in *Dynamic systems and Control Division*, 1998, pp. 139–143.

[99] A. H. Mason and C. L. MacKenzie, “Grip Forces when Passing an Object to a Partner,” *Exp. Brain Res.*, vol. 163, no. 2, pp. 173–87, May 2005.

[100] K. Nagata, Y. Oosaki, M. Kakikura, and H. Tsukune, “Delivery by Hand between Human and Robot Based on Fingertip Force-Torque Information,” in *International Conference on Intelligent Robots and Systems*, 1998, pp. 750–757.

[101] T. Deyle, H. Nguyen, M. S. Reynolds, and C. C. Kemp, “RFID-Guided Robots for Pervasive Automation,” *IEEE Pervasive Comput.*, vol. 9, no. 2, pp. 37–45, Apr. 2010.

[102] H. Nguyen, A. Jain, C. Anderson, and C. C. Kemp, “A clickable world: Behavior selection through pointing and context for mobile manipulation,” in *International Conference on Intelligent Robots and Systems*, 2008, pp. 787–793.

[103] W. P. Chan, “A Human-Inspired Controller for Robot-Human Object Handovers: A Study of Grip and Load Forces in Handovers and the Design and Implementation of a Novel Handover Controller,” University of British Columbia, 2012.

[104] K. Hsiao and T. Lozano-Perez, “Imitation Learning of Whole-Body Grasps,” in *2006 IEEE/RSJ International Conference on Intelligent Robots and Systems*, 2006, pp. 5657–5662.

[105] A. E. Leeper, K. Hsiao, M. Ciocarlie, L. Takayama, and D. Gossow, “Strategies for human-in-the-loop robotic grasping,” in *International Conference on Human-Robot Interaction*, 2012, p. 1.

[106] A. T. Miller, S. Knoop, H. I. Christensen, and P. K. Allen, “Automatic grasp planning using shape primitives,” in *International Conference on Robotics and Automation*, 2003, vol. 2, pp. 1824–1829.

[107] J. Aleotti and S. Caselli, “Part-based robot grasp planning from human demonstration,” in *International Conference on Robotics and Automation*, 2011, pp. 4554–4560.

[108] A. Saxena, J. Driemeyer, and A. Y. A. Ng, “Robotic grasping of novel objects using vision,” *Int. J. Rob. Res.*, vol. 27, no. 2, pp. 157–173, 2008.

[109] E. Bicici and R. S. Amant, “Reasoning About the Functionality of Tools and Physical Artifacts,” Department of Computer Science, North Carolina State University, 2003.

[110] W. P. Chan, Y. Kakiuchi, K. Okada, and M. Inaba, “Determining Proper Grasp Configurations for Handovers through Observation of Object Movement Patterns and Inter-object Interactions during Usage,” in *International Conference on Intelligent Robots and Systems*, 2014, pp. 1355–1360.

[111] R. Ueda, “Integrated System Architecture for Life-size Humanoid with Task and Motion Planner based on Predictive Fault Evaluation,” Master Thesis, Department of Information Science and Technology, University of Tokyo, Tokyo, Japan, 2010.

[112] H. Inoue, “Towards an Advanced Robot Programming,” *J. Robot. Soc. Japan*, vol. 2, no. 2, pp. 87–90, 1984.

[113] W. P. Chan, M. K. X. J. Pan, E. a Croft, and M. Inaba, “Characterization of Handover Orientations used by Humans for Efficient Robot to Human Handovers,” in *International Conference on Intelligent Robots and Systems*, 2015, pp. 1–6.

[114] “Vicon.” [Online]. Available: http://www.vicon.com/. [Accessed: 11-Oct-2015].

[115] Vicon, “Nexus.” [Online]. Available: http://www.vicon.com/products/software/nexus. [Accessed: 11-Oct-2015].

[116] I. Sharf, A. Wolf, and M. B. Rubin, "Arithmetic and geometric solutions for average rigid-body rotation," *Mech. Mach. Theory*, vol. 45, no. 9, pp. 1239–1251, 2010.

[117] W. D. Curtis, A. L. Janin, and K. Zikan, "A note on averaging rotations," in *Virtual Reality Annual International Symposium*, 1993, no. 2, pp. 377–385.

[118] C. Gramkow, "On averaging rotations," *J. Math. Imaging Vis.*, vol. 15, no. 1–2, pp. 7–16, 2001.

[119] H. S. Koppula, R. Gupta, and A. Saxena, "Learning Human Activities and Object Affordances from RGB-D Videos," *Int. J. Rob. Res.*, vol. 32, no. 8, pp. 951–970, 2013.

[120] M. Sridhar, A. Cohn, and D. Hogg, "Learning functional object categories from a relational spatio-temporal representation," in *European Conference on Artificial Intelligence*, 2008, pp. 606–610.

[121] H. Kjellström, J. Romero, and D. Kragić, "Visual object-action recognition: Inferring object affordances from human demonstration," *Comput. Vis. Image Underst.*, vol. 115, no. 1, pp. 81–90, 2011.

[122] Natural Point Inc., "OptiTrack." [Online]. Available: http://www.naturalpoint.com/optitrack/.

[123] Natural Point Inc., "Optitrack V100:R2." [Online]. Available: http://www.naturalpoint.com/optitrack/products/v100-r2/specs.html.

[124] V. Micelli, K. Strabala, and S. Srinivasa, "Perception and control challenges for effective human–robot handoffs," *RSS 2011 RGB-D Work.*, 2011.

[125] J. Civera and D. Gálvez-López, "Towards semantic SLAM using a monocular camera," in *International Conference on Intelligent Robots and System*, 2011, pp. 1277–1284.

[126] D. Di Marco and A. Koch, "Creating and using roboearth object models," in *International Conference on Robotics and Automation*, 2012, pp. 3549–3550.

[127] D. Di Marco and R. Janssen, "RoboEarth," 2014. [Online]. Available: http://wiki.ros.org/roboearth_stack. [Accessed: 24-Sep-2014].

[128] R. Ueda, "Tracking 3D objects with Point Cloud Library," 2012. [Online]. Available: http://www.willowgarage.com/blog/2012/01/17/tracking-3d-objects-point-cloud-library. [Accessed: 24-Sep-2014].

[129] T. Moulard, "Visp Tracker." [Online]. Available: http://wiki.ros.org/visp_tracker?distro=jade. [Accessed: 27-Sep-2015].

[130] Structure Sensor, "OpenNI." [Online]. Available: http://structure.io/openni. [Accessed: 18-Dec-2015].

[131] W. P. Chan, K. Nagahama, H. Yaguchi, Y. Kakiuchi, K. Okada, and M. Inaba, "Implementation of a Framework for Learning Handover Grasp Configurations through Observation during Human-Robot Object Handovers," in *Humanoids*, 2015.

[132] K. M. Slomczynski, *Social Structure and Mobility: Poland, Japan, and the United States*. Institute of Philosophy and Sociology, Polish Academy of Sciences, 1989.

[133] Y. Furuta, Y. Inagaki, Y. Kakiuchi, K. Okada, and M. Inaba, "Implementation of the IRT Assistant Robot Clean Up Task Sequence on PR2," in *Conference of Robotics Society of Japan*, 2013.

[134] K. Chaudhary, Y. Mae, M. Kojima, and T. Arai, "Autonomous acquisition of generic handheld objects in unstructured environments via sequential back-tracking for object recognition," in *International Conference on Robotics and Automation*, 2014, pp. 4953–4958.

[135] T. Matsumaru, "Handover movement informing receiver of weight load as informative motion study for human-friendly robot," in *International Symposium on Robot and Human Interactive Communication*, 2009, pp. 299–305.

Appendix A Object Handover User Study Consent Form

The University of British Columbia
Collaborative Advanced Robotics and Intelligent Systems (CARIS) Laboratory
Department of Mechanical Engineering, UBC
6250 Applied Science Lane, Vancouver, BC V6T 1Z4
Tel: [redacted] **Fax**: [redacted]
Web site: http://caris.mech.ubc.ca

HRI-Cues: Human-Human Handover Study II

Project Title: HRI-Cues: Human-Human Handover Study II
Principal Investigator: Elizabeth Croft, ([redacted]), [redacted],
Co-Investigators and Contact Persons: Wesley Chan ([redacted]) and Matthew Pan ([redacted]), [redacted]

Funding: This research is funded by the Natural Sciences and Engineering Research Council of Canada (NSERC).

Purpose: The purpose of this study is to observe and kinematic behaviours in human-human handovers, and to explore how a human giver and receiver negotiate handovers of several objects. Results from this study will be used in subsequent research to improve the ability of robotic assistants to interact with non-expert human users.

Procedures: Before the actual handover experiment, you will be asked to fill out a preliminary questionnaire which asks you for some demographic information (e.g., age, gender, etc). For the experiment, you will be paired with another participant to perform a series of handovers using several everyday objects while wearing a jacket and a cap with motion capture markers. One participant will play the role of the giver, while the other will be the receiver. During the experiment, both the giver and the receiver will be standing. Both the giver and receiver will start with their hands placed at their side. On a verbal "Go" signal, the giver will retrieve an object, and hand it over to the giver. After object transfer, both the giver and the receiver will return to their start positions.

The experiment will last approximately an hour. You may refuse to participate in this experiment and you may withdraw at any time. We will be recording motion capture and video data, although the latter is not required for your participation.

Potential Risks: Slight temporary fatigue from passing various objects back and forth.

Confidentiality: No identifying information will be collected or stored with your data. Data collected from the survey will be stored on a password protected computer or a locked cabinet in the CARIS Lab, which has restricted secure access and is locked at all times. If you have any questions or concerns about what we are asking of you, please contact the study leader or one of the study staff. The names and telephone numbers are listed at the top of the first page of this form. If you have any concerns about your rights as a research subject and/or your experiences while participating in this study, you may contact the Research Subject Information Line in the UBC Office of Research Services at [redacted] or if long distance e-mail [redacted] or call toll free [redacted].

Consent: By signing this form, you consent to participate in this study, and acknowledge you have received a copy of this consent form.

I agree to allow myself videotaped during this experiment (please circle one): **YES** **NO**

Name (print):_________________________________ Date:_______________

Signature:____

Appendix B Computed Mean Handover Orientations

Table B-1 Computed mean handover orientations for all conditions from the handover user study.

	Condition A			Condition B			Condition C		
	R	P	Y	R	P	Y	R	P	Y
book	2.5	-3.6	-77.6	5.0	-8.4	-47.2	1.3	-3.1	-27.9
bottle	2.3	2.8	-10.0	1.9	22.0	122.5	-9.3	12.4	-81.2
camera	-1.0	-1.7	137.4	-29.1	22.3	-66.5	-10.8	11.0	-79.7
Cereal box	-8.9	17.3	67.4	0.6	24.0	88.4	-3.0	-0.9	-56.7
umbrella	39.9	-31.8	-35.1	7.7	-11.2	37.6	11.9	-22.7	60.2
flowers	28.1	-7.8	-32.0	24.2	-23.9	-36.0	-3.0	-0.9	-45.3
fork	62.0	-44.8	-19.3	43.9	-73.5	-50.3	-53.5	74.2	149.6
hammer	144.0	78.2	-49.1	149.2	-79.5	65.0	272.2	66.8	-196.3
knife	-81.7	93.5	-7.5	24.9	-144.3	-94.1	-96.2	73.5	6.6
mug	-5.0	2.2	-80.2	-2.4	8.7	-9.5	-1.1	4.1	-142.1
pen	-63.4	54.8	-110.7	151.9	-220.3	-19.4	-3.9	61.3	63.2
plate	0.0	-1.6	-104.1	1.0	-0.4	-32.5	-0.6	0.9	94.9
remote	32.9	76.2	53.0	70.7	-107.8	-18.2	100.3	116.6	-9.9
scissors	92.2	8.7	-47.4	-92.4	60.2	120.0	113.8	11.9	-65.6
screwdriver	245.5	56.1	-153.7	-2.6	-148.6	-45.6	-5.8	100.1	-77.1
stapler	-20.4	-8.3	-177.4	-33.8	-1.2	-172.3	-11.9	-6.6	-154.4
teapot	-4.1	3.2	131.9	-3.1	5.3	44.4	0.7	3.9	-160.0
tomato	-10.5	2.2	161.0	-10.9	9.6	25.7	-13.3	10.6	150.8
wineglass	2.7	1.3	-9.4	3.6	-13.9	20.1	6.2	0.3	69.3

Appendix C Extracted Handover Grasp Configuration Data

Table C-1 Handover grasp configurations extracted from demonstrations and computed averages for the fifteen objects used in the second set of experiments. $C_{handover}$ gives the handover orientation, P_{grasp} gives the grasp point, and P_{grasp2} gives the placement of the giver's second hand at handover.

	Trial	$C_{handover}$ (deg)			P_{grasp} (mm)			P_{grasp2} (mm)		
		Roll	Pitch	Yaw	X	Y	Z	X	Y	Z
book	1	178.8	-11.5	80.8	-94.15	90.94	-31.52	-674.36	-317.82	88.37
	2	179.3	-16.0	80.8	-61.12	73.57	-16.82	-687.96	-311.92	182.99
	3	-175.3	-1.7	76.8	-107.34	101.68	4.17	-684.44	-394.47	61.51
	4	170.2	-12.0	86.5	-43.30	54.66	28.11	-651.99	-303.00	284.68
	5	175.9	-12.0	77.9	-60.51	55.01	-34.59	-654.21	-329.38	268.37
	Avg	178.2	-11.5	80.8	-73.29	75.17	-10.13		Null	
bottle	1	123.8	-22.9	13.8	-88.26	-119.49	-77.28	-421.38	-567.42	-156.16
	2	135.8	-25.2	-9.7	-6.16	-11.69	-26.25	-624.09	-559.03	-298.24
	3	115.7	-14.3	21.2	29.09	-86.25	-19.92	-130.53	-707.83	-190.30
	4	107.7	-8.6	12.6	50.90	-125.31	11.00	140.77	-750.24	-231.51
	5	107.1	-8.0	17.8	52.17	1.28	-29.17	-21.30	-825.06	-210.85
	Avg	114.6	-14.3	16.0	7.55	-68.29	-28.32		Null	
camera	1	-12.6	-1.7	20.1	112.62	11.76	30.01	13.91	26.03	-13.78
	2	-17.2	8.0	11.5	100.10	-21.71	48.91	104.20	55.07	-67.53
	3	-6.3	3.4	-8.6	115.24	13.92	32.91	7.29	11.59	23.79
	4	1.7	3.4	12.0	110.48	22.80	29.37	89.86	11.11	-67.47
	5	-9.7	0.6	-20.1	81.97	6.21	17.54	71.30	100.29	-164.42
	Avg	-8.6	3.4	5.7	104.08	6.60	31.75	53.81	25.95	-31.25
umbrella	1	116.3	85.4	-55.0	49.69	-2.03	24.98	9.42	-272.28	-485.07
	2	169.6	76.8	44.7	-99.66	23.40	-5.14	0.65	-555.72	-123.97
	3	-56.1	79.6	147.8	121.90	-29.94	-144.55	253.09	-288.73	-557.42
	4	-94.0	76.8	150.1	119.66	-126.38	-75.12	217.91	-498.18	-84.01
	5	-40.1	85.9	173.6	-101.70	-102.74	7.95	173.36	-502.17	-327.14
	Avg	-40.1	85.9	173.6	17.98	-47.54	-38.37		Null	
hammer	1	43.5	9.2	-8.0	6.03	17.71	143.56	367.41	-139.75	-332.67
	2	42.4	8.0	1.1	-3.66	2.31	146.16	550.20	-72.65	-221.45
	3	55.6	-5.2	6.9	-5.69	36.00	93.03	395.28	-389.48	-288.87
	4	52.7	17.2	-13.2	-66.13	-22.88	62.51	4.79	-11.45	22.41
	5	48.1	1.7	-9.7	-22.35	30.93	87.36	463.34	-133.74	-231.27
	Avg	46.4	6.9	-6.3	-18.36	12.81	106.53		Null	
knife	1	53.9	18.9	-12.0	-28.15	23.55	90.86	475.17	-327.43	-294.70
	2	51.6	3.4	4.6	-110.31	0.57	29.22	342.67	-415.83	-282.73
	3	57.9	1.1	4.0	-44.74	15.62	129.68	357.27	-360.31	-348.01
	4	58.4	1.7	4.6	-61.51	11.36	64.18	367.38	-292.24	-350.97
	5	75.6	-22.9	20.6	-46.28	57.63	0.69	96.22	-600.40	-270.63
	Avg	57.9	1.1	4.6	-58.20	21.75	62.93		Null	
mug	1	-171.9	-1.1	-1.7	-79.90	1.06	38.22	-608.17	-128.88	-449.98
	2	174.2	-8.0	1.7	-95.50	4.37	10.45	-717.45	-257.18	-382.28
	3	171.3	-3.4	2.3	-112.55	-12.72	29.10	-556.90	-459.96	-399.60
	4	174.8	0.0	-1.7	-88.43	-23.22	-21.42	-648.67	-310.80	-454.75
	5	171.9	-9.2	1.1	-107.26	-2.29	-1.40	-638.30	-353.27	-331.74
	Avg	174.2	-5.7	1.1	-96.73	-6.56	10.99		Null	

(continued next page)

	Trial	$C_{handover}$ (deg)			P_{grasp} (mm)			P_{grasp2} (mm)		
		Roll	Pitch	Yaw	X	Y	Z	X	Y	Z
pen	1	15.5	-24.1	1.1	77.16	88.29	-12.32	483.80	109.33	-615.67
	2	119.2	65.9	52.1	-9.92	-12.39	-56.36	340.43	-632.66	244.41
	3	-46.4	-43.5	-12.0	-62.92	47.83	-32.38	-111.75	550.24	-122.02
	4	60.7	39.5	-173.0	0.84	57.19	-42.75	810.04	2025.71	-1732.17
	5	45.3	2.9	81.4	68.12	36.90	-21.25	547.81	-410.48	321.90
	Avg	20.6	13.8	36.7	14.66	43.56	-33.01	Null		
plate	1	1.7	-0.6	0.6	-2.54	-97.15	48.04	573.72	226.34	-351.68
	2	149.0	-1.7	-0.6	-95.36	12.32	56.05	-429.84	-422.60	-353.62
	3	14.3	-2.3	-8.6	117.58	96.92	45.23	621.16	176.00	-437.52
	4	-84.2	-4.0	4.0	-110.78	72.88	29.92	-77.24	622.15	-298.19
	5	-148.4	-12.0	5.7	-126.44	-45.62	62.48	-655.70	128.44	-237.83
	Avg	1.7	-0.6	0.6	-43.51	7.87	48.34	Null		
remote	1	-80.2	0.0	-2.3	46.84	39.30	84.86	-77.40	578.07	-274.44
	2	-93.4	17.2	-1.7	54.93	16.51	134.46	-70.11	546.58	-282.40
	3	-87.7	-0.6	0.0	37.04	-0.54	118.45	-128.29	526.12	-312.69
	4	-91.7	10.3	-6.3	37.89	-22.51	118.91	-151.63	424.42	-200.81
	5	-88.8	-2.9	-3.4	12.78	2.18	105.60	-126.02	445.93	-336.14
	Avg	-87.7	1.7	-1.7	37.89	6.99	112.46			
scissors	1	84.8	-4.0	85.4	-74.07	0.88	78.73	163.26	-297.92	540.85
	2	104.9	8.6	60.2	-78.48	5.26	92.71	75.68	-575.91	268.21
	3	95.1	-26.4	51.6	-47.75	-22.11	80.58	-40.41	-530.27	345.14
	4	102.6	-5.2	57.9	-63.98	-24.11	80.98	-46.85	-529.42	244.97
	5	85.4	-30.9	35.0	-70.30	4.47	76.58	18.01	-594.40	8.78
	Avg	99.1	-7.4	57.9	-66.92	-7.12	81.92	Null		
screwdriver	1	88.8	5.2	48.1	-18.53	-45.61	115.01	137.44	-506.06	356.83
	2	131.8	-58.4	29.2	33.62	-36.97	117.68	-478.77	-490.98	345.24
	3	110.6	-46.4	35.5	-1.91	-27.10	133.82	-357.15	-585.98	232.43
	4	44.1	19.5	28.6	-7.77	-32.84	111.90	650.95	-301.70	-40.80
	5	-22.9	45.8	24.1	5.32	-0.32	103.80	571.19	485.86	-28.48
	Avg	82.5	-1.1	48.7	2.15	-28.57	116.44	Null		
kettle	1	-100.8	4.6	-16.0	66.37	6.43	98.38	-280.09	492.38	-194.31
	2	-87.1	-4.0	-18.3	63.26	-9.38	116.37	-239.27	698.58	-134.67
	3	-83.7	1.1	-0.6	65.55	18.18	110.95	-209.65	647.53	-284.49
	4	-94.5	-3.4	-9.2	66.55	2.23	92.25	-100.05	636.67	-256.04
	5	-91.1	-5.2	-4.0	57.01	-14.50	74.68	-163.90	582.22	-271.40
	Avg	-92.8	-2.9	-9.2	63.75	0.59	98.52	Null		
wrench	1	56.1	-2.3	6.9	-25.40	-22.03	80.87	268.59	-312.85	-35.42
	2	51.6	8.0	-3.4	-5.73	42.93	99.79	424.81	-312.13	-256.35
	3	51.0	-4.6	4.6	-47.31	-46.22	38.32	366.07	-278.61	-244.07
	4	55.6	-0.6	9.7	-67.63	-2.13	38.18	367.85	-350.16	-242.25
	5	49.3	4.0	1.7	-66.95	-81.24	27.50	394.68	-284.55	-253.98
	Avg	53.3	-0.6	5.2	-42.61	-21.74	56.93	Null		
Detergent bottle	1	10.9	12.0	5.7	112.52	57.69	27.48	669.05	-31.81	-393.15
	2	2.9	9.7	0.0	108.40	51.26	39.04	819.25	127.08	-410.01
	3	16.6	9.2	8.6	146.92	41.12	10.75	691.63	163.81	-296.04
	4	-20.1	14.3	0.6	50.06	41.97	17.21	658.10	455.12	-325.57
	5	13.8	2.9	4.0	89.88	50.88	2.66	587.07	264.49	-368.84
	Avg	9.7	10.3	4.6	101.56	48.58	19.43	Null		