A Thesis for the Degree of Ph.D. in Engineering

A study on gestures at a hands occupied situation for manually controlling a helping hand robot

March 2016

Graduate School of Science and Technology
Keio University

Mahisorn Wongphati

For mom and dad.

Acknowledgments

First of all, I would like to express my earnest gratitude to my supervisor, Professor Michita Imai, who gave me the best opportunity in my life to pursue my passion in robotics at the Keio University. This work will never been completed without his advice, supervision and supporting during the first three years in the Yagami campus and the two and a half years leaving in Thailand.

I also would like to thank Assistant Professor Hirotaka Osawa who provided many helps on my research, especially the robot arm used in the experiment. Without his helps, the realization of my helping hand robot might not be possible.

One of the important marks in my study at the Keio University is a request from Professor Yuichiro Anzai during the laboratory party in the year 2013.

> “I want to read your thesis”

By accepting the request, I could still remember the change that happened in my research decision. The word thank you is not enough for this invaluable trigger.

I gratefully acknowledge everyone in Anzai-Imai Laboratory, especially Takahiro Matsumoto and Yusuke Kanai, who have helped me in many things during my early days in our laboratory.

Finally, I would like to thank my friends and my family who always support me in every situation. I am sincerely grateful to them for their understanding and encouragement during my time in Japan.

Abstract

A helping hand robot is one of the main focuses in human robot collaboration (HRC) research and has high potential in both modern industry and daily living usages. The helping hand robot could be considered as an extra hand that gives helps when both hands of a user are occupied by a task. A multimodal interaction between a human and a robot is an important requirement in HRC because robots are usually working and interacting closely with a user, it is essential to find natural and intuitive user interfaces between a user and a robot in various usage scenarios.

From many usage scenarios and multiple interaction possibilities, this thesis studied the use of nonverbal interaction such as hand and body gestures for controlling basic movements (forward/backward, up/down, and left/right) of an end effector of a helping hand robot. We focused mainly on a scenario when both hands are occupied and manually control is needed, therefore explicit gestures such as a waving hand(s) gestures cannot be easily performed by a user.

To get an idea about the suitable natural gestures for controlling the movements of an end effector, we conducted a pilot study with laboratory members and found that gestures those the members were asked to freely perform vary substantially. With this cue, we (1) set up a video based experiment to survey gestures for controlling movements of the end effector and (2) developed a real helping hand robot system for evaluating the discovered gestures.

To allow the participants to freely use any gestures which they feel suitable, we used a guessability study methodology for extracting gestures from the participants in the gesture surveying experiment. The experiment showed the "effects" of gestures (a set of pre-recorded videos of an end effector movements) to participants and asked them to think about the "causes" or gestures (e.g. tilting body) which they thought suitable and intuitive. Although results from this methodology depend on the background and experiences of each participant, the results led to a set of common gestures which were generic and

intuitive for most participants.

By conducting the video based experiment with 19 voluntary participants, we captured and categorized 152 gestures. Our findings showed that a hand was a part of body used most often for gesture articulating even when the participants were holding tools and objects with both hands, that gestures for a pair of opposite movements such as up/down were consistently performed by most participants, and that the participants rarely care or aware of using one- or two-handed gestures interchangeable. From 152 gestures, we also found that there were many alternative gestures such as pursing lips, tilting head, and so on which could be useful for other situations such as a use for the handicapped persons.

By using results from the gesture surveying experiment, we implemented a helping hand robot using a small industrial robot for validating the discovered gestures. We used Microsoft Kinect sensors for sensing user's hand, and body movements. We implemented a gesture recognition algorithm using a state machine that checks distances of hands, arms, and body from their initial positions. The gestures that were a combination of hand, arm, and body movements could effectively be used to control movements of the helping hand robot by the participants.

With the implemented system, we conducted the second experiment with eight participants that used a real robot as a helping hand for a soldering task. The results showed that the selected hand and body gestures were easily accepted by the participants. The outcomes aligned with our expectation on the two most performed gestures from the video based experiment. However, in the real robot experiment, the body gestures were preferred over the hand gestures. This finding was unexpected but helped us confirm our intuition from the pilot study and the need of the real robot system implementation in the human robot interaction study.

Our findings could be useful as a guideline for acquiring natural gestures for controlling robots as a complementary for the multimodal interaction with a robot in HRC.

Contents

List of Figures

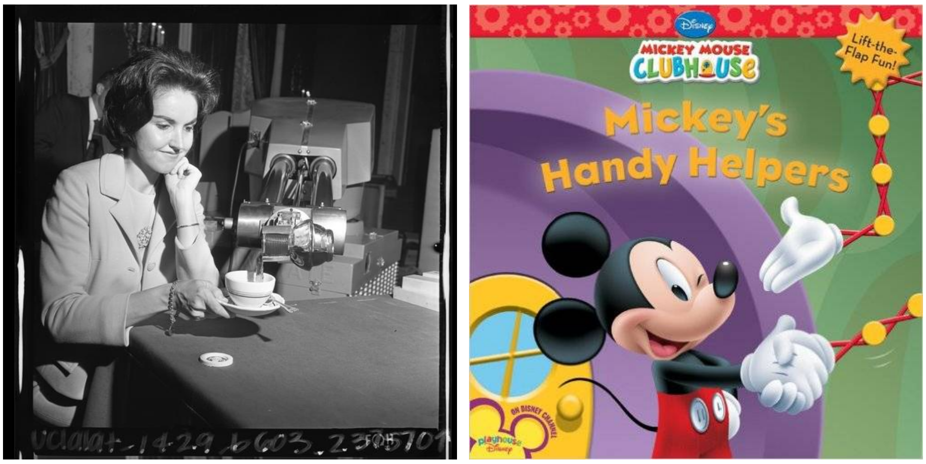
List of Tables

Chapter 1

Introduction

1.1 Motivation

A helping hand robot has been addressed since the early days of the introduction of an industrial robot in the 1960s [1]. Not only the real life concept has been portrayed (Figure 1.1a), the helping hand robot also appears in many sci-fi novels, movies, or even a cartoon for children (Figure 1.1b).



(a) (b)

Figure 1.1: (a) The Unimate robot helped pouring coffee for its mistress in 1967[a]. (b) Mickey also used the helping hand robots in his clubhouse[b].

---

[a] Source: http://lit250v.library.ucla.edu/islandora/object/edu.ucla.library.specialCollections.latimes%3A3491

[b] Source: http://www.amazon.com/Mickeys-Helpers-Disney-Mickey-Clubhouse/dp/142311017X

Not only the portrays and imaginations show the usefulness of the helping hand robot, the actual clinical trial of the helping hand robot for handicapped person (Figure 1.2) is also focused by many research groups since the 1980s.[2, 3]

Figure 1.2: The third generation of Desktop Vocational Assistant Robot (DeVAR-III) for helping a patient with disabilities in activities of daily living (ADL). Image courtesy of Hammel et al. [2].

The recent innovation on a helping hand robot for patient or disable person is a system from [4] that allows the helping hand robot to be directly controlled with brain waves for a tetraplegia patient (paralysis of both arms and both legs) as shown in Figure 1.3.

For a daily life of a normal person, the helping hand robot will no longer be just an imagination robot. There is an attempt to create a cooking robot system that can work side by side with its user as shown in Figure 1.4.

From the trend, it is not too risky to predict that the helping hand robots will become more and more important in our daily lives. The helping hand robot also has high potential in the modern industrial processes. Especially, the processes those are prepared for mass customization products. For such products, the customers will be allowed to customize almost every aspect of a product. This kind of requirement is not suitable for current robot systems that usually rely on a pre-programmed set up and not flexible.

To overcome the challenges, there are many attempts on a cognitive factory, a factory that can reconfigure itself to match with required outputs and available resources [5]. With currently state of the art, a fully autonomous factory is still infeasible for most daily use products those require complex assembly and testing processes. One of the key

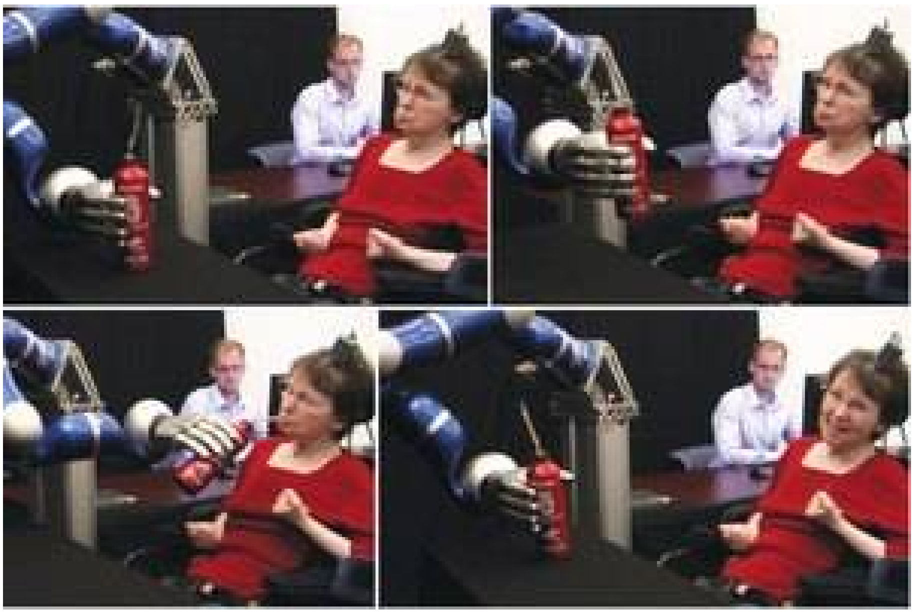


Figure 1.3: The patient with tetraplegia controlling a helping hand robot with her brain waves for the drink bottle. Image courtesy of Hochberg et al. [4].

Figure 1.4: The prototype of a helping hand robot system for cooking from Moley Robotics[a].

---

[a] Source `http://www.moley.com`

component to bridge the present need and the future technology is including human in the process with helps from a robot as shown in Figure 1.5.

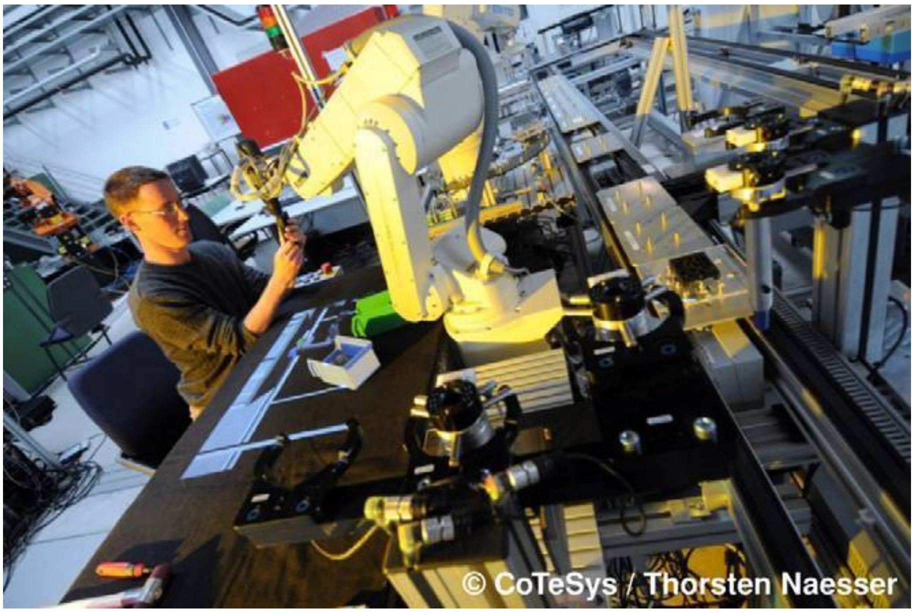


Figure 1.5: The worker working closely with an industrial robot in a re-configurable factory. Image courtesy of Wallhoff et al. [6].

For daily lives and industrial usages of a helping hand robot, there is one component that plays important role in the system, the interaction and communication between the human and the robot. The current industrial robot and sensor systems are already advanced enough to handle complex objects, tools, and human interactions. However, when the robot is required to handle natural language commands or interact directly with a human, the requirements become a very hard problem to tackle in robotics research. This created a new research field in robotics called human-robot interaction (HRI) [7]. Furthermore, when a collaboration between the robot and human, which is a core of the helping hand robot, is also needed, there is a sub-field that directly address issues called the human-robot collaboration (HRC) research [8].

From many usage scenarios of a helping hand robot in [9], we noticed one specific situation when both hands of a human are occupied and a helping hand is needed. There are many scenarios those a hands occupied situation can occur. For example, a handicraft task that a user is holding tools and objects with both hands and wants to adjust/control a helping hand robot (Figure 1.6a), a cooking task that a user wants to adjust angle of a helping hand robot that helps cooking while the user are working on the other plate with both hands (Figure

1.6b), a patient with reduced communication channel such as a person with broken arms or a person with some form of disabilities such as deaf or amputated arms (Figure 1.6c).


(a) (b) (c)

Figure 1.6: (a) Holding and gluing two objects together with both hands [a]. (b) Both hands are busy cooking [b]. (c) The amputated man was working on his new arms[c].

---

[a] Source: `http://www.model-space.com/gb/blog/2015/08/scale-model-basics-a-guide-to-glue/`

[b] Source: `http://www.theworldofchinese.com/2013/04/beijing-cuisine-beyond-peking-duck/`

[c] Source: `http://kotaku.com/5935251/chinese-farmer-is-both-luke-skywalker-and-iron-man`

In such scenarios, it is unlikely that a robot or even a human can perfectly determine what kind of manipulation that the user needs the robot to perform and apply to a particular object or environment without additional information after the help has been requested (e.g. “could you help me on this one?”). A dialog in Table 1.1 is one of the possible outcomes when a user is working with a helping hand robot for a soldering task.

Table 1.1: An example of a possible dialog between a user and a helping hand robot in a soldering task.

(1)	User:	“Add a soldering lead here”
(2)	Robot:	“Okay”
(3)	Robot:	(Trying to locate a soldering point)
(4)	Robot:	(Move an end-effector to a certain position)
(5)	User:	“left”
(6)	User:	“right”
(7)	User:	(repeated line 5 and 6 for several times)
(8)	User:	“override”
(9)	User:	(manually control the robot)

In the dialog, the user ask for a help from the helping hand robot in line 1 and the robot respond with line 2 before trying to locate a soldering point in line 3 and 4. The robot has to find the best target position based on available sensor information because the user is working on a task that the robot dose not had prior knowledge about it (e.g. Figure 1.7). In this case, the robot has failed to perform to meet the user expectation (line 5-7) and has been overridden by the user and being manually controlled in line 8.

Figure 1.7: A helping hand is needed! [a]

---

[a] Source: `http://portal.nifty.com/2010/11/20/b/`

The example dialog shows the need of the additional method for the user to manually control the end effector of the helping hand robot in certain situations. With the additional hands occupied constraint, it is clear that traditional methods such as a teaching pendant, a joystick, or a mouse are not suitable. The hand occupied situation also makes the tactile control methods such as a direct touch or a force torque sensor control inapplicable. Moreover, a speech command as shown in Table 1.1 could becomes tedious and very repetitive when it is used for conveying spatial information [10]. One of the reason is the fact that deictic terms such as “here”, “there”, and etc. are usually error-prone and hard to be interpreted by machines [11]. On the other hand, gestures are more suitable and less ambiguous when using conveying spatial information for controlling a robot or a computer device [12].

These motivations led us to the study on suitable gestures for manually controlling a helping hand robot in a hands occupied situation.

1.2 Summary

The summary of this thesis are as follows:

Possible usages of a helping hand robot We created two version of the online surveys to collect opinions and usage ideas of a helping hand in daily life from Thai and Japanese [9]. We asked two questions, “where do you want to use a robot arm?” and “what do you want from the robot?”. More than 240 responds has been collected and the results provided the motivations of this study. All collected usage ideas were described in Appendix A.

Finding of a set of user-defined gestures for manually controlling a helping hand robot We adopted a guessability study methodology [13] and used it to conduct a video based experiment to find gestures for manually controlling a virtual helping hand robot in a hands occupied situation that used a soldering task as an example scenario.

Implementation of a real helping hand robot system to confirm a usability of the user-defined gestures We implemented a helping hand robot system to confirm the usability of the collected user-defined gestures (hand, body, and some supplementary gestures). A generic algorithm has been developed for hand and body gestures recognition in the experiment and used as one of the controlled conditions for the participants. The experiment compared usability of the hand and body gesture methods with the traditional mouse/keyboard method for manually controlling the helping hand robot in a soldering task.

Evaluation and results analysis The gestures have been collected through common practices in HRI research. The experiments have been tested and prepared to collect and evaluate gestures for manually controlling a helping hand robot in a hands occupied situation. The results were analyzed in details from both qualitative and quantitative information collected during the experiment using interviewing, questionnaire, conversation, and recorded videos.

1.3 Contributions

We conducted the experiments on both visual and real robots to find suitable gestures for manually controlling a helping hand robot in the hands occupied situation that used a soldering task as an example scenario. The findings reveal a number of intuitive and counterintuitive characteristics of the user-defined gestures for manually controlling the helping hand robot.

The experimental and study processes also emphasize the need of a confirmation for a video-based experiment using a real robot experiment. In our study, the video based experiment unveiled two promising gesture sets, hand and body gestures, for manually controlling a helping hand robot. Even though hand gestures are counter intuitive for the hand occupied situation, it is conclusive from the video based experiment that hand gestures were the most used gestures. However, when the real robot system was implemented for testing the results, the study confirmed that body gestures (see Section 4.4.2) are more suitable gestures for manually controlling a helping hand robot in the hands occupied situations.

1.4 Organization

This thesis has seven chapters and two appendices. After the introduction in the first chapter, background of the study is provided in Chapter 2. Related work are extensive reviewed and summarized in Chapter 3. Chapters 4 provides detailed explanation of the gesture surveying experiment that utilized a video based experiment and a guessability study methodology. With the results from Chapter 4, Chapter 5 explains the usability testing of the collected gestures with a real helping hand robot. The limitations of the study and future work are provided in Chapters 6. Chapter 7 concludes the study.

Chapter 2

Background

2.1 Helping hand robots

A helping hand robots or an assistive manipulator has been a focus for many research group since the introduction of an industrial robot in the early 1960s. The implementation of the helping hand robot can roughly be separated into two groups, a desktop based robot and a mobile base robot.

The ideas of robotic assistants, especially, for a person with disability dates back to the early 1960s when Reswick et al. [14] proposed a power assisted exoskeleton for a disable person called Arm Aid robot (Figure 2.1). Although the robot is a power assistive tool for the disable person, it is one of the very early attempts in the modern days on helping people with robot system before the industrial robots and computers are widely available in the 1980s.

During the late 1970s and the early 1990s, there is a long term research project about a robotic assistant conducted at Rehabilitation Research and Development Center, California [3]. A number of rehabilitation robot systems have been developed and evaluated during the grant period of the research. One of the results of the project is a mobile manipulator called MoVAR which is one of the very first mobile manipulators that is designed as a mobile helping hand robot (Figure 2.2).

Hammel et at. [2] reported the clinical evaluation results of a desktop robotic assistant called DeVAR which is a part of the long term project. The system could be controlled by a quadriplegic (a person who is paralyzed in both arms and both legs) using various speech commands for daily lives tasks such as feeding, personal hygiene, vocational, recreational, and so on. A part of the implemented system is shown in Figure 1.2. The DeVAR robot is able to communicate

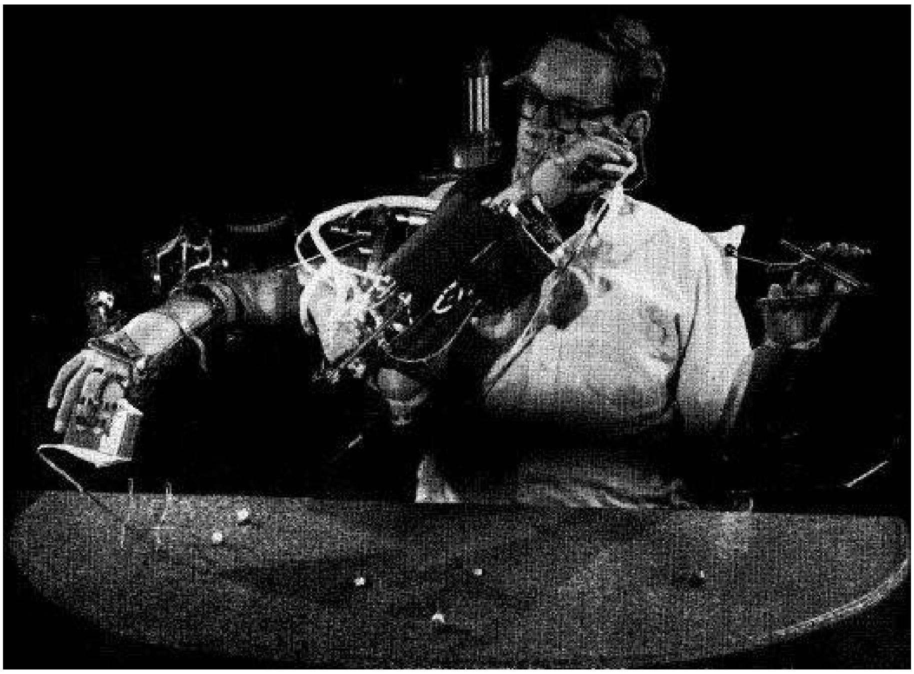


Figure 2.1: The Case Research Arm Aid, Mark I. Image courtesy of Reswick and Vodovnik [14].

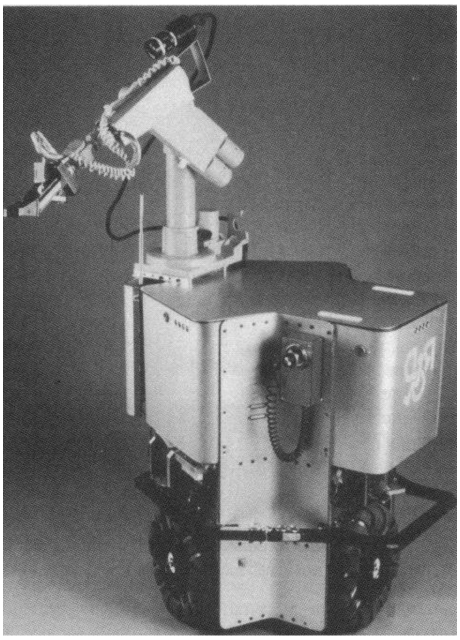


Figure 2.2: The MoVAR mobile manipulator robot. Image courtesy of der Loos [3].

and help its users through speech recognition and speech synthesis systems.

Before the 2000s, one of the main problems of a complex robot system was a lag of computation power with reasonable cost and size. This problem is more problematic with a helping hand robot with complex sensors for sensing and interacting with its environment and users. However, when CPU technology is getting better, sensors are more capable, and robots are becoming a mainstream in both manufacturing process and research, there are many active developments on helping hand robot systems after we entered the third millennium. More details about related robot systems are described in Chapter 3.

2.2 Human robot interaction

Human robot interaction or HRI is a field of study for understanding and evaluating robot systems those are working or living with humans. The interaction between humans and robots can be carried out in several forms and channels. One of the main factors of the interaction methods is the distance between the human and the robot, remotely operated or close proximity. Therefore, the interaction can separated into two groups, remote interaction and proximate interaction [7]. For more information about the remote interaction or teleoperation, which is also a very early robot development in HRI, comprehensive guideline and discussion can be found in [15].

In this study, we focused mainly on the close proximate interaction. In such interaction, the classic methods for communicating and controlling the robot are teaching pendants, joysticks, or mouse and keyboard. Such methods normally allow only manual control or perform a teaching and playback programming in an industrial domain [1].

When more complex interactions are needed and natural communication is preferred, a multimodal interaction based on five senses of the human (sight, hearing, taste, smell, and touch) have been incorporated into the development of a robot system. Normally, sight and hearing are the main channels for communicating and interacting with the robots through gestures, gazes, and speeches as shown in Figure 2.3 from [16]. Nevertheless, the touch or tactile is gaining more attraction when force-torque sensors are more advanced and affordable, especially for human-robot collaboration system implementation.

The close proximate interaction can be separated into two types, interacting and collaborating. The interacting is more general and it

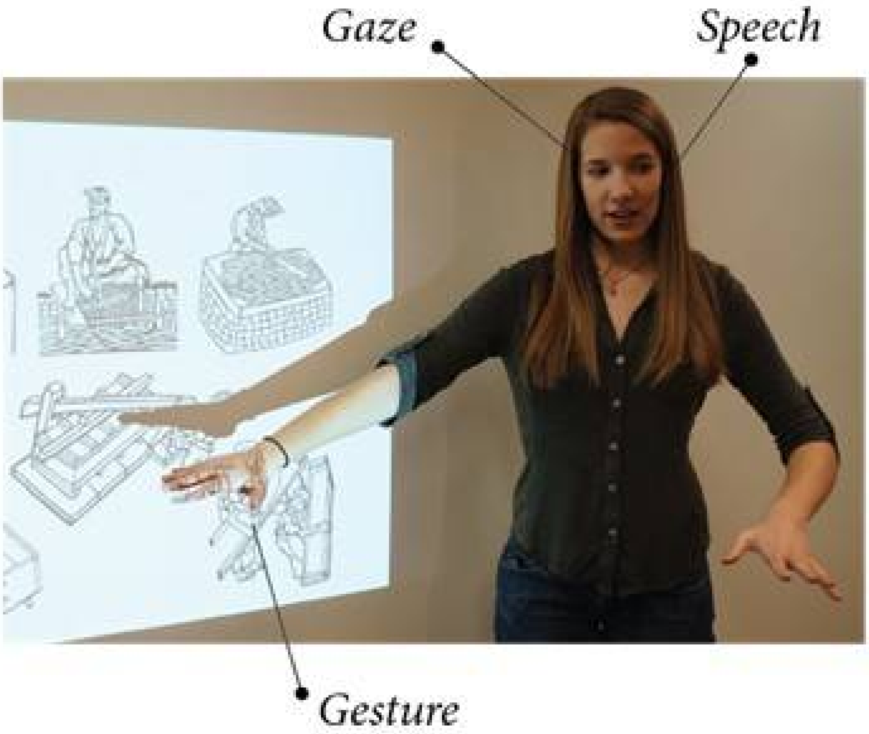


Figure 2.3: Multimodal interaction used by the presenter during the presentation for communicating intention and acquiring attention. Image courtesy of Huang and Mutlu [16]

is any kind of action that involves another human, system, or robot, which does not has to gain or aware about a benefit from the effort. The overview of the HRI can be found in [17, 7].

On the other hand, collaborating consists of members of a team with complementary skills who are committed to a common purpose, goal, and approach for which they share the benefit of working together. The same collaborative structure can be applied to human-robot teams that want to reach a common objective trough collaboration. The following quote from Bauer et al. [8] showed the important of having a robot system that is capable to create a joint intention that will lead to a joint action for reaching the objective of the human and robot collaboration.

> Efficient collaboration requires a common plan for all involved partners. To gain a joint intention they need to know the intentions of the other team members and what they are doing. Based on that knowledge a robot can plan its own actions that will eventually lead to fulfill the joint intention and reach a common goal. Therefore, they need the abilities of perceiving and comprehending their envi-

> ronment, decision-making, planning, learning, and reflection.

It can be infer from the quote that the multimodal interaction is a key for achieving the required joint action which will be described in the next section.

2.3 Multimodal interaction

When a team is working together, the important factor to guarantee the success of the collaboration is information sharing between members. The information can be the background knowledge of each member, the knowledge explained by the team leader, or the knowledge that each member has analyzed during the process. When objectives and information can be confirmed and all members understand their roles, synchronization between collaborators will happens through the multimodal interaction that includes both verbal and nonverbal communications. The same principal can be applied to a human and robot case.

However, with limited technologies, realizing the multimodal interaction on a robot system becomes a major challenge in HRI research that still need a number of efforts and contributions from the HRI community.

Chapter 3

Related work

3.1 Robot systems

The followings are some of the robot system those were designed as a companion or assistant on various tasks.

Care-O-bot [18], is a mobile robot platform for a home care system from Fraunhofer IPA, Germany. It is a long term research and development of a series of assistive robots for daily lives of elderly and handicapped people. After more than a decade of research efforts, the project expanded its domain to include industrial service robots that are working closely with human called rob@work. Both robot systems emphasis the need of the suitable multimodal communication and interaction between a human and a robot through numerous research. The efforts are based on both robot hardware construction and robot interaction research.

ALBERT [19], is a mobile manipulator platform developed in the early 2000s. The robot was equipped with a laser range finder, a stereo vision system, and a speech recognition system. The study proposed an event management that combines object recognition and gestures with speech input to allow the robot to communicate with a normal user. The system was capable to receive to grasp and place objects based on speech and gesture command from the user.

CoRA [20], is a robotic assistant for collaborating with a human on object manipulation tasks. The robot supported multimodal communication such as visual, audio, and haptic (touch). The robot was capable to recognize various objects, speeches, gestures, and gazes of its user. With a combination of the recognized information, the robot was able to pick up an object based on gaze and hand pointing gestures and bring it to the user hand when hand opening gesture can be detected. The user was also allowed to touch to change or interrupt robot motions directly.

SAC [21], is a two arm robotic system proposed by Zang and Knoll that focused on enabling the robot system to assemble objects by two arms and to understand unconstrained natural language of the human instructor. The system consisted of two six DoF robot arms installed overhead in a workstation cell. The robot system was capable to view objects and activities from the above of the working space and from the gripper point of view using cameras. The system could assembly many complex objects using multimodal dialogues with its instructor using speeches and hand gestures recognition functions.

Kawarazaki et al. [22] developed a cooperative welfare robot system that can be controlled by hand gestures. The gestures were detected and recognized using trinocular stereo vision that utilized three cameras. The robot was able to pick up objects according to a set of predefined hand gestures. The gestures allowed the user to point at an arbitrary object to instruct a seven DoF industrial robot arm to grab and bring object to the user hand.

ARMAR [23], is a humanoid robot with two arms and a mobile base. Stiefelhagen et at. presented an ongoing work for building a natural multimodal HRI for the robot. In the 2004, the robot was able to recognize speech and has dialogue with human. It also was able to recognize some gestures such as pointing gestures and head orientation. The system combined information from detected gestures and head orientation to elevate quality of speech recognition and correct ambiguity of sentences those contained deictic terms. The system was later improved and become one of the well-known robot system that are capable to perform various daily living tasks.

Recently, with advanced sensors and powerful small-size computers, a number of new robot systems were proposed with more capabilities for interacting with humans and environment. For example, HERB by Srinivasa et al. [24] is a mobile manipulator that could navigate, explore, and manipulate various objects in a common household, EL-E [25] is also a mobile manipulator that could be commanded by a user to pick up various objects from the floor. Cody [26] is a humanoid robot with mobile based that was used for many research objectives including bed baths for patient hygiene.

Some of the robots are continuously upgraded to incorporate new functions and more computation power such as ARMAR III [27], Care-O-bot 3 [28] or PR2 [29] and some of them became a standard platforms for HRI research and benchmark. Furthermore, the new generation robotic arms are safer or even designed for using as an assistive helping hand. A comprehensive survey for such robot can be found in [30].

The mentioned robot systems are some of related work those utilized multimodal communication, including gestures, in various aspects such as commanding, controlling, and cooperating with robots. Surprisingly, an option for manually controlling helping hand robots while working closely with them was not formally discussed in the mentioned studies, despite the fact that it is the last resort for users to overcome a glitch with their problem-solving skills when the robot's performance does not meet their expectation, as shown in the example dialog in Table 1.1.

It is also interesting to found that gestures in the mentioned works were designed by system developers who were familiar with the system's capability and normally the designers tend to select gestures based on ease of detection and distinguishability to increase recognition reliability (see Table 3.1 for more details). Therefore, the designer-designed gestures might not represent the real expectations of a user and might feel unnatural [13].

Table 3.1: Some of the designer designed gestures from the mentioned related studies. Most of the gestures were designed for giving a specific command to a robot such as a stop command or a command for requesting an object.

Research from	Gestures
Rogalla et al.	Hand gestures that can be recognized using the pre-defined contours of reference gestures [19].
Iossifidis et al.	Open hand and pointing gestures [20].
Zhang and Knoll	A small set of hand gestures [21].
Kawarazaki et al.	Four pre-defined hand gestures [22].

3.2 Gesture based interaction

The human robot interaction research (HRI) started gaining momentum during the late 1980s and the early 1990s [17]. A decade later, the interacting with gesture becomes one of the main focus for human-computer interaction (HCI) and HRI as a part of multimodal interaction. Many evaluation methodologies for gesture interaction between

HCI [31] and HRI [32, 33] have been proposed. Furthermore, with advanced sensors and more powerful computers, the multimodal interaction is also gaining more focus from the robotic research communities for both social robots [34] and industrial robots [8, 6] and gestures are a part of the effort.

Various methods were proposed for gesture recognition. In the early days, vision based gesture recognition system was a primary method for gesture interaction such the methods used in [19], [20], [22], and [35]. More information about vision based gesture recognition can found in [12]. After the arrival of consumer 3D sensors such as Kinect sensors, the more advanced gesture recognition method has been proposed for complex hand gestures [36]. A more intrusive method that requires sensing device to be attached was also investigated by a number of research, especially, for an outdoor interaction for field robots [37, 38].

Gestures are suitable for handling spatial referencing [10], especially, when deictic terms such as “here”, “over there”, and so on are involving in the interaction [39, 11]. A number of research showed the suitability of the gestures for a collaboration between a human and a robot as mentioned in the previous section. Interestingly, there was also an inverted investigation of the usability of *robot gestures* as an interaction modality for human-robot collaborative assembly [40].

At the time of the study, Gleeson et al. [41] proposed a closely related study that followed the same methodology from [13]. However, the study focused on developing a lexicon for communicating with an industrial robot instead of the controlling of an end effector as presented in this study.

Chapter 4

Gestures for a helping hand robot

4.1 Introduction

In this study, we assumed that a helping hand robot is working closely in front of a human as illustrated in Figure 4.1. This configuration is a common configuration for both human-human and human-robot interaction evaluations [42, 43].

Figure 4.1: The drawing of the human working closely with the helping hand robot on the architectural design process. Image courtesy of P. Sanoamueng.

Although the idea of controlling robots using gestures has been addressed by many research, directly adopting gestures those already have meanings or purposes, or designing a new set of gestures based only on the ease of detection and distinguishability, might not be a viable method, because different gestures could be chosen by different users in different situations [13]. Furthermore, gestures, like human languages, depend heavily on the user's background, and therefore, systems that can recognize gestures should be designed according to the gestures that are most likely to be used [12].

To investigate the characteristics of gestures that are needed for designing an efficient control method for a helping hand robot, we conducted an experimental study that adopts parts of the methodology from [13] to collect a set of user-defined gestures for manually controlling movements of an end effector. We used the guessability study methodology [13] to focus our experiment on how users are articulating their gestures after the causes of the gestures were shown. The "causes of the gesture" means a result (e.g. a robot moves its end effector to the right of a user) that elicits the user to perform a gesture that he/she feels suitable (e.g. sweeping right hand to the right).

In this study, we selected a soldering task as an example scenario for a survey that needs a hands occupied situation and the soldering task provides the following conditions for the situation.

1. The participant can hold objects independently with both hands (e.g. a soldering iron and a cable in the left and right hands before start soldering).

2. The participant left and right hands have to stay close together during the task (e.g. holding tips of the soldering iron and a cable together (see Figures 1.7 and 4.2b).

3. Both translation and orientation controls are needed for adjusting the helping hand robot.

In a scenario for the survey, we assumed that at a certain point of time during the soldering task, the participant wants to manually control the helping hand robot according to the following sequence.

1. The participant is trying to solder a cable to a particular point on a circuit board.

2. But both hands of the participant are occupied by a soldering iron and a cable.

3. The participant ask for a help from a helping hand robot.

4. The helping hand robot has some glitches and cannot perform according to the participant's expectation as show in Table 1.1.
5. The user wants to manually control the helping hand robot when it is unable to help feeding a soldering wire as expected.

4.2 Gesture surveying

4.2.1 Objectives

For this survey, we set up a video based experiment that continues from the point where the user are going to manually control the helping hand robot to observe, record, and collect information about gestures those the participants though suitable for the helping hand robot movements.

4.2.2 Setup and environment

The survey was conducted at the Keio Techno-Mall exhibition in 2011. The survey used a video based experiment to ensure reproducibility and safety of the participants [44]. The experiment was set up on a table in the exhibition booth (Figure 4.2). A PowerPoint presentation (Figure 4.4) was used for the experimental process and displayed on a vertical screen in Figure 4.2a. During the exhibition, the laboratory members helped inviting the visitors to participant in the experiment. The survey was conducted from 10:00 to 18:00 with one hour lunch break. There were two trained conductors helped conducting the survey (only one conductor at a time) during the seven hours experiment.

Six end effector movements (forward, backward, left, right, up, and down) and two gripper commands (open and close) were selected for the experiment. The selection of commands is based on a pilot study (see Section 4.5.3) that was conducted with the laboratory members to test a number of movements needed to give a good sense of an end effector control. The selection of the commands is also related to the limited exhibition time (8 hours) and the effort to provide a reasonable surveying session time for each participant (about 15 minutes). All end effector videos that were used during the experiment are shown in Figure 4.3. The videos were prepared using a software called V-REP from Coppelia Robotics[1].

---

[1] http://www.coppeliarobotics.com (accessed December 2015)

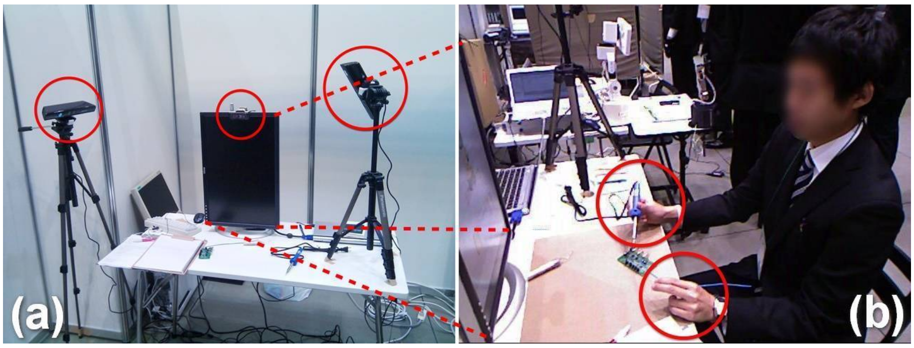
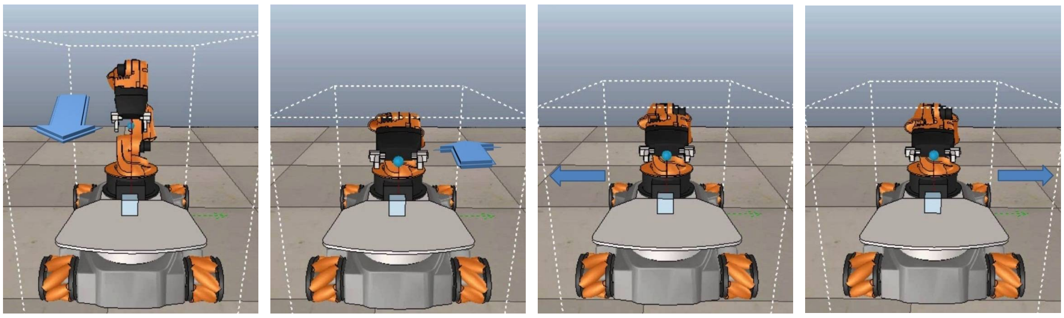
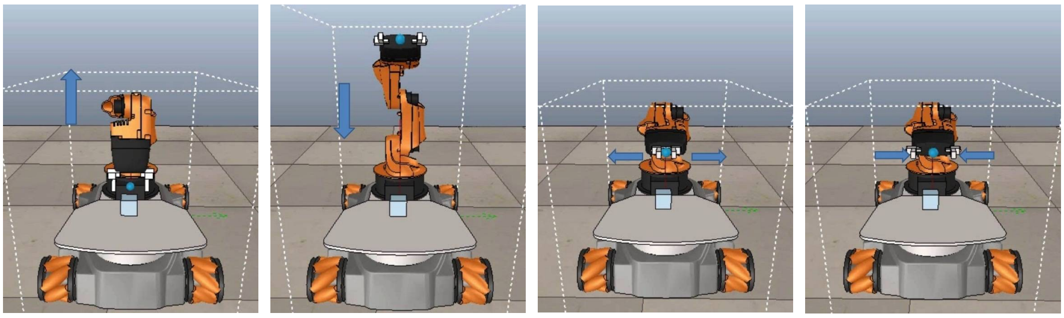


Figure 4.2: (a) The table equips with the 24-inch vertical screen for displaying pre-recorded videos of a visual helping hand robot and other instructions. Three cameras are used for recording videos from front, left, and right views (circled). (b) The participant was asked to think that he is doing the soldering task and had to hold the soldering iron and the cable with both hands during the experiment (circled).



(a) Forward (b) Backward (c) Left (d) Right



(e) Up (f) Down (g) Open (h) Close

Figure 4.3: Six movements and two gripper command for the virtual helping hand robot.

4.2.3 Participants

Nineteen visitors at the exhibition voluntarily participated in the survey: 16 males and three females. The ages of the participants are shown in Table 4.1. Two participants did not have any soldering experience. Seven participants had some experience with an industrial robot. The participants included university students, company employees, housewives, researchers, engineers, and retirees.

Table 4.1: The ages of the participants (years)

< 18	18–30	31–45	46–55	> 55
0	7	6	4	2

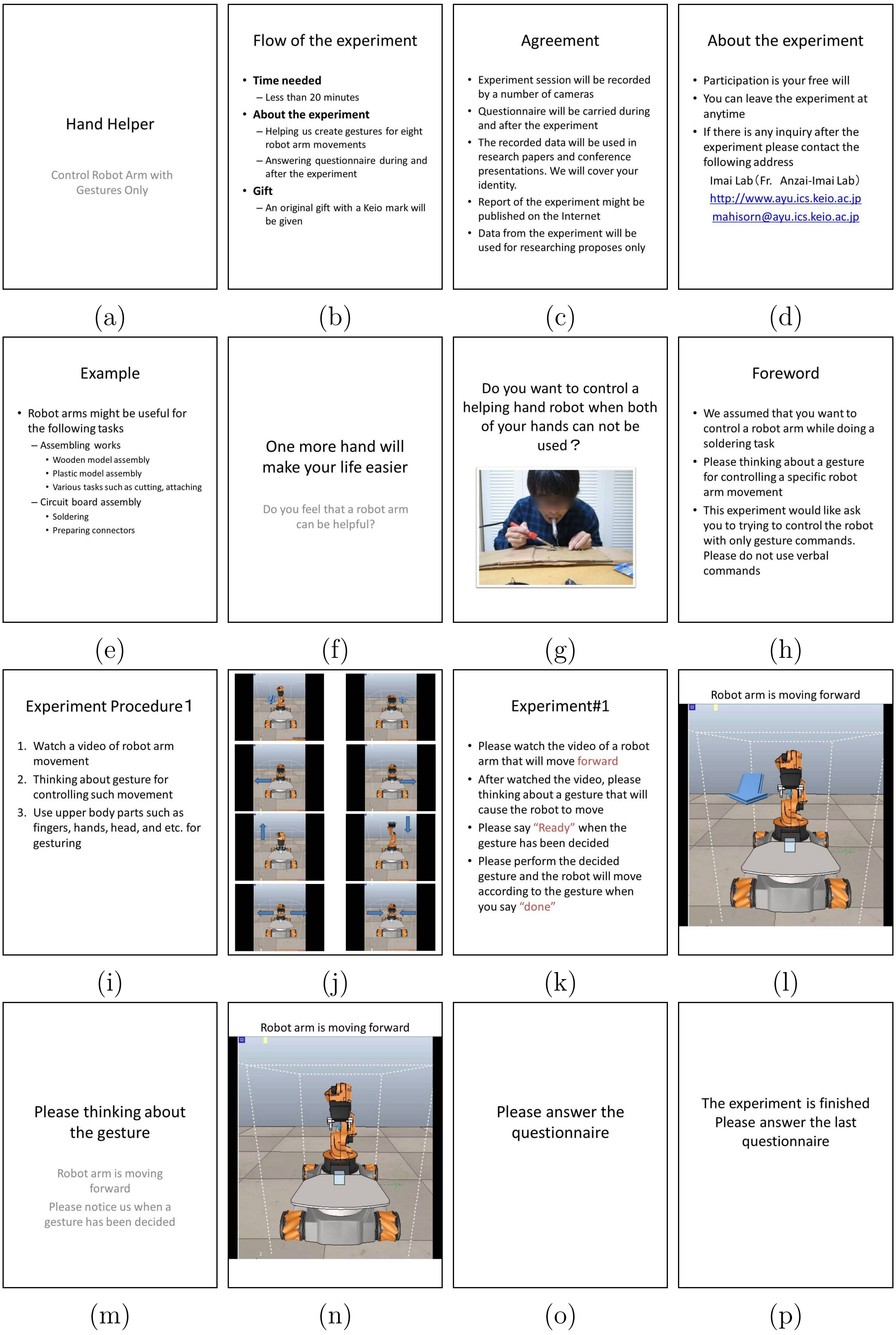
The non-professional participants helped finding out how the participants, without awareness about limitations of the robot and the sensors, articulate gestures in the hands occupied situation.

4.2.4 Procedure

After each participant agreed to volunteer in the experiment, one of the conductors explained the purpose and the introduction of the experiment using slides in Figures 4.4a to 4.4j. The participant was asked to hold a unplugged soldering iron and a cable during the experiment and pretend that the participant was going to solder the cable to a specific point on a circuit board as shown in Figure 4.2b.

During the experiment, the pre-recorded videos of the virtual helping hand robot movements were randomly selected and shown to the participants (Figure 4.4l). During each video, the conductor asked the participant to think about a gestures (causes) that will result the watched movement (effects) (Figure 4.4m). After that, the participant was given some time to think about how to plan the decided gesture. The participant then signaled the conductor when the participant was ready to perform the gesture. The conductor observed the gesture and manually trigger the corresponding video to simulate gesture recognition capability and the pre-recorded video will play accordingly (Figure 4.4n).

When finished each movement, the participant was asked to confirm the body part used for gesturing and rate two Likert scales (1-disagree to 7-agree) those could be read as “the gesture I have performed is easy to plan” and “the gesture I have performed can easily



Hand Helper

Control Robot Arm with Gestures Only

(a)

Flow of the experiment

- **Time needed**
  - Less than 20 minutes
- **About the experiment**
  - Helping us create gestures for eight robot arm movements
  - Answering questionnaire during and after the experiment
- **Gift**
  - An original gift with a Keio mark will be given

(b)

Agreement

- Experiment session will be recorded by a number of cameras
- Questionnaire will be carried during and after the experiment
- The recorded data will be used in research papers and conference presentations. We will cover your identity.
- Report of the experiment might be published on the Internet
- Data from the experiment will be used for researching proposes only

(c)

About the experiment

- Participation is your free will
- You can leave the experiment at anytime
- If there is any inquiry after the experiment please contact the following address

Imai Lab（Fr.　Anzai-Imai Lab）
http://www.ayu.ics.keio.ac.jp
mahisorn@ayu.ics.keio.ac.jp

(d)

Example

- Robot arms might be useful for the following tasks
  - Assembling works
    - Wooden model assembly
    - Plastic model assembly
    - Various tasks such as cutting, attaching
  - Circuit board assembly
    - Soldering
    - Preparing connectors

(e)

One more hand will make your life easier

Do you feel that a robot arm can be helpful?

(f)

Do you want to control a helping hand robot when both of your hands can not be used？

(g)

Foreword

- We assumed that you want to control a robot arm while doing a soldering task
- Please thinking about a gesture for controlling a specific robot arm movement
- This experiment would like ask you to trying to control the robot with only gesture commands. Please do not use verbal commands

(h)

Experiment Procedure１

1. Watch a video of robot arm movement
2. Thinking about gesture for controlling such movement
3. Use upper body parts such as fingers, hands, head, and etc. for gesturing

(i)

(j)

Experiment#1

- Please watch the video of a robot arm that will move forward
- After watched the video, please thinking about a gesture that will cause the robot to move
- Please say “Ready” when the gesture has been decided
- Please perform the decided gesture and the robot will move according to the gesture when you say “done”

(k)

Robot arm is moving forward

(l)

Please thinking about the gesture

Robot arm is moving forward

Please notice us when a gesture has been decided

(m)

Robot arm is moving forward

(n)

Please answer the questionnaire

(o)

The experiment is finished
Please answer the last questionnaire

(p)

Figure 4.4: The part slides from the PowerPoint presentation used during the experiment (translated from Japanese). The introduction of the experiment is from slides (a) to (j). The experiment for each helping hand movement is repeated randomly using slides (k) to (o) as a template.

be recognized by the robot". After watched and performed gestures for eight randomized videos, the participants were asked to answer demographic questions (Figure 4.4p) and an unexpected gift was awarded to the participant.

The author observed and took notes of all sessions on interactions, gestures, and conversations between the participants and the conductors for the further analysis.

ELAN software [45] as shown in Figure 4.5 was used to synchronize and annotate all the recorded videos. The gesture planning time for each gesture (the time that each participant took to think about the gesture) was manually marked using the synchronized videos.

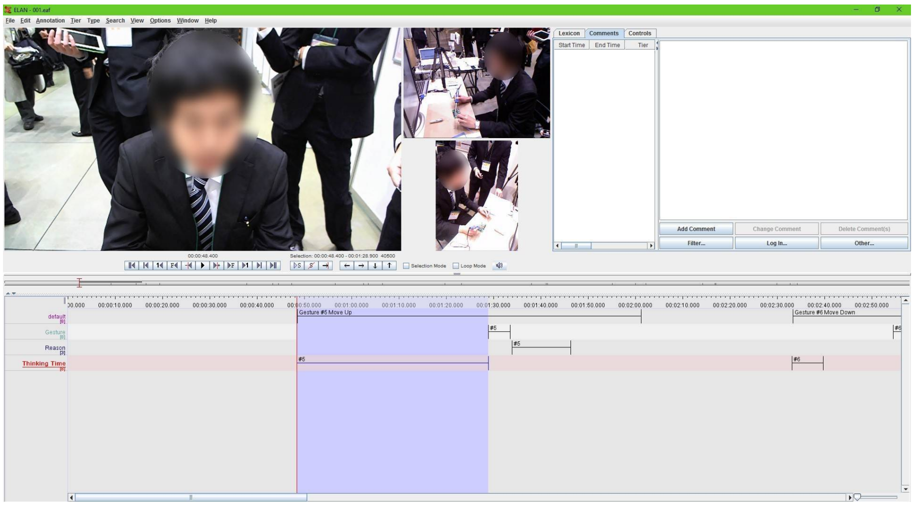


Figure 4.5: The ELAN software used for videos analysis.

4.3 Results

Nineteen participants made 152 gestures for manually controlling the end effector of the virtual helping hand robot. The percentages of body parts used by all participants for gesturing during the experiment are shown in Figure 4.6.

Hands were the most used body parts and 40% of hand gestures were one-handed gestures (using only left or right hand). The head and torso (upper body) accounted for 13% and 19% of the gestures, respectively. Other 10% of gestures were fingers (5%), mouth (3%) , arm (1%), and shoulder (1%). Most of the gestures were a pair of reversible gestures those had symmetrical movements as shown in Figure 4.7.

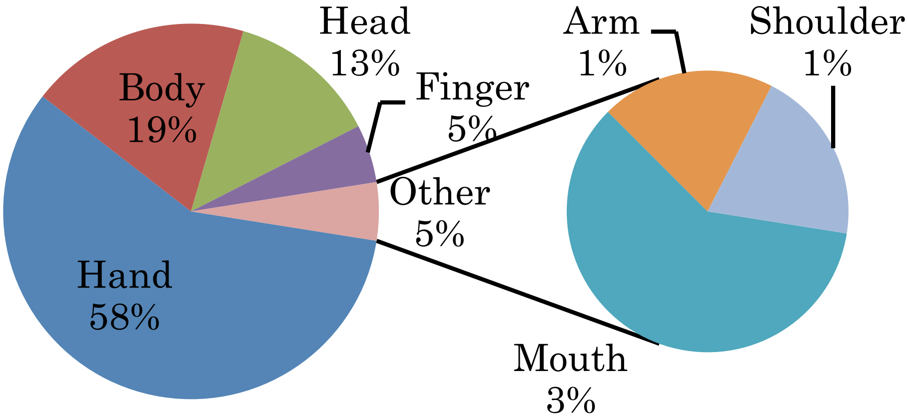


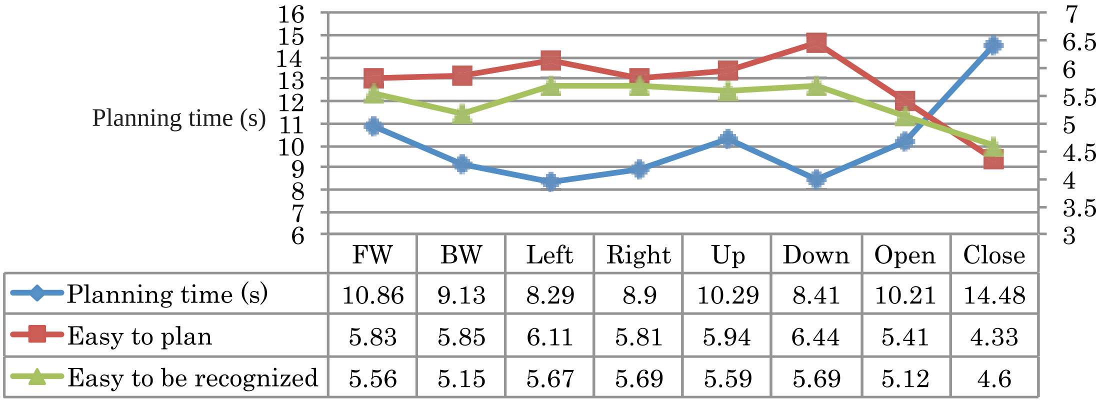
Figure 4.6: Percentages of body parts used for gesture articulation during the experiment

(a) (b)

Figure 4.7: The pair of reversible hand gestures for controlling left (a) and right (b) movements of the virtual helping hand robot

4.3.1 Gesture planning time and user expectation

Statistical data extracted from recorded videos in Figure 4.8 indicates that the average planning time to decide each gesture negatively correlates with the average score of “The gesture I have performed is easy to plan” from the questionnaire ($r = -0.9$). This correlation shows that the planning or thinking time of the gestures is also a good indicator of the complexity of the gesture for manually controlling a helping hand robot as suggested by [13].



	FW	BW	Left	Right	Up	Down	Open	Close
◆ Planning time (s)	10.86	9.13	8.29	8.9	10.29	8.41	10.21	14.48
■ Easy to plan	5.83	5.85	6.11	5.81	5.94	6.44	5.41	4.33
▲ Easy to be recognized	5.56	5.15	5.67	5.69	5.59	5.69	5.12	4.6

Figure 4.8: The average planning time (diamond marks), score of “The gestures I have performed is easy to plan” (square marks), and score of “The gesture I have performed can easily be recognized by the robot” (triangle marks) of all gestures (FW: forward, BW: backward)

The average score of the “The gesture I have performed is easy to plan” also correlates significantly with the score of “The gesture I have performed can easily be recognized by the robot” from the questionnaire ($r = 0.9$). This correlation implies that the participants expected a better gesture recognition system when they believed that a gesture is simple and easy to plan. This informs that the expectation on the system depends on how the users thinking about simplicity and intuitiveness of the interactions.

4.3.2 Hand and other body parts

Although the participants were informed at the beginning for the experiment that the scenario is the soldering task, and that the soldering iron and the cable have to be held during the experiment, most of the participants still used their hands as the main body part for gesture articulation (58% in Figure 4.6).

The outcomes are different from our expectation that most of the participants would not use their hands, but would try to use other body parts to articulate the gestures. The outcomes also different from the pilot study (see Section 4.5.3) with the laboratory members those were familiar with robots because they have tried hard to think about gestures that did not involving with hands.

For the hand gestures, the recorded data shows that the participants used two different approaches for gesture articulation. We call the first approach a fixing-tool approach (Figure 4.9a), where the participants tried to fix, or arrange, the tools and objects in their working position, while articulating the gesture. The second approach is called a moving-tool approach (Figure 4.9b), where the participants moved their hands freely with the soldering iron and the cable in their hands while articulating the gesture.

(a) The fixing-tool hand gesture for opening the gripper.

(b) The moving-tool hand gesture for opening the gripper.

Figure 4.9: (a) The fixing-tool gesture articulation. (b) The moving-tool gesture articulation.

These noticeable differences emphasize the importance of hand gestures and indicate that a gesture recognizer must be able to handle a situation when a user of the system is holding objects in their hands.

4.3.3 One- and two- handed gestures

The recorded videos and notes took during the experiment also show that many participants used one- and two-handed gestures interchangeably. Some of the participants also explicitly made both one- and two-handed gestures for the same movement during the experiment (Figure 4.10).

(a) The one-handed gesture for moving the end effector left.

(b) The two-handed gesture for moving the end effector left.

Figure 4.10: The one- and two-handed gestures for the same movement. The red and green arrows indicate moving direction of hands and the virtual robot in the screen, respectively

This interchangeable use of gestures implies that the gesture recognition systems should be able to handle both one-and two-handed gestures.

4.3.4 Gesturing pattern

The annotated data shows that the gestures articulations could be separated into two groups, namely “mirroring” and “directing” gestures. The mirroring gestures are a group of gestures that were articulated by the participants like a mirror reflection of the end effector movement. For example, sweeping hand(s) to the left to tell the robot to move to the left of the participant as happened in Figure 4.10. This group of gestures was used most often, and could be observed in all body parts used for gesture articulating. On the other hand, the directing gesture is a group of gestures representing movements those the participants wanted the end effector to follow in a non-mirroring manner. Gestures in this group can only be observed in hand gestures (Figure 4.11a) and body gestures (Figure 4.11b) for the forward and backward movements.

An important finding for these two groups of gestures is that the participants were not aware of the difference between the mirroring and directing gestures, as they performed the directing gestures (e.g. leaning *backward* to tell the robot to move *forward*) before and/or after the mirroring gesture (e.g. sweeping hand *left* to tell the robot to

(a) The hand gesture for moving the end effector forward.

(b) The body gesture for moving the end effector forward.

Figure 4.11: The directing gestures for the forward movement.

move to the *left*) in the randomized sequence of the end effector movements and the gripper commands. In other words, some participants performed mirroring gesture before directing gesture and vice versa.

This issue is important for the system designer to inform the system user about the system behavior to avoid any unexpected incident during the close interaction.

4.3.5 Gesturing consistency

The recorded data also reveals the gesture articulation consistency. The consistency of gesture articulation refers to the parts of the body and the methods used by the participants to articulate gestures for the same movements and commands, in particular, for the four pairs of reversible movements and commands: (forward/backward, left/right, up/down, and open/close). The consistency is one of the important information for selecting gestures and designing the gesture recognition system.

The analysis shows that nine of 19 participants used only their hands (including fingers) for the gesture articulation. The other ten participants used the combinations in Table 4.2 for the gesture articulation.

This observation shows that about 50% of the participants used only single body part for gesture articulation. However, when considering only the body parts used for gesture articulation between each pair of gestures, the data shows that 69 out of 76 gesture pairs, or less than 10%, were articulated using the different body parts as show in

Table 4.2: The counts for different body parts used by the participants who did not use only hand for gesture articulation.

Used body parts	Number of participants
Body and head	1
Body and hand	4
Body, head, and hand	2
Body, head, and arm	1
Body, head, and mouth	1
Head and mouth	1

Figure 4.12.

Another finding regarding the consistency of gesture articulation methods, particularly the direction and symmetry of gestures, is that the inconsistency between the participants can be observed only in the gestures for forward/backward movements (see Section 4.3.4). Nevertheless, the recorded data also shows that there was no inconsistency between pattern of each pair of gestures if a small number of outlier gestures (see Section 4.3.6) and confusing (Figures B.14c and B.14d are excluded.

(a) The body gestures for moving the end effector forward.

(b) The hand gestures for moving the end effector backward.

Figure 4.12: The inconsistency of body parts used for the pair of forward and backward movements.

Although many different body parts were used by the participants for some of the gestures, the articulation methods and used body parts is almost 100% consistent within each pair. This is useful information for the prediction the pattern of the gestures.

4.3.6 Outlier

Many gestures are completely different from the trends of their group or from the entire set of collected gestures (Figure 4.13). Gestures in this distinctive group might not be a good choice for continuously controlling of a robot end effector and some of them are obviously difficult for articulating, recognizing, and remembering.

(a) The head gesture for closing the gripper.

(b) The mouth pursing gesture for closing the gripper.

(c) The shoulder shrugging gesture for moving the end effector forward.

(d) The arm flapping gesture for opening the gripper.

Figure 4.13: Some of the outlier gestures collected during the experiment.

Nevertheless, when a handicapped person or a working condition that has very limited space or movement is considered, this distinctive group of gestures can be used as a guideline for special gestures development. For example, a gesture for the opening gripper command by flapping elbows in Figure 4.13d can be useful as a signal for opening or toggling function of the robot system when both hands are engaging or occupied by a task. The additional outlier gestures can be found in Appendix B.

4.4 Selected gestures

The gestures for manually controlling the end effector of the helping hand robot were selected by counting the number of matched gestures between each participant. Although the study adopted the methodology from Wobbrock et al. [13], who carried out more detailed analyses for the gesture selection, we were able to establish the following intuitive selection criteria.

The dominant gestures for all commands are hand gestures. They had a frequency at least 50% greater than that of any other type of gesture, as shown in Table 4.3.

Table 4.3: Counts of body parts used for each gesture articulation (FW: forward, BW: backward).

	FW	BW	Left	Right	Up	Down	Open	Close
Hand & finger	10	10	11	10	13	13	14	14
Head & body	8	9	8	9	6	6	2	2
Other	1	0	0	0	0	0	3	3

The consistency of gesture articulation described in Section 4.3.5 emphasizes that hand gestures were not only the gestures used most often, but also the gestures those were articulated with the most consistency by all participants. The following subsections describe the selected sets of primary and secondary gestures.

4.4.1 Primary gestures

The criteria provide a straightforward choice for selecting hand gestures as the primary gestures for manually controlling the end effector of the helping hand robot (Figures 4.14 and 4.15). One of the important characteristics of the selected primary gestures is that one- and two-handed gestures were used interchangeably by the participants for forward/backward, up/down, and left/right movements, but not for the gripper open and close commands.

As mentioned in Sections 4.3.3, the implementation of a gesture recognizer for the selected gestures should be able to handle both one- and two-handed gestures for a better usability. The directions of the forward and backward gestures in Figures 4.14a and 4.14b could confuse a user as described in Section 4.3.4 and the user should be informed to ensure safety.

4.4.2 Secondary gestures

The body gestures were selected as the secondary gestures as indicated by counts of the body parts used by the participants during the experiment (Table 4.3). For left/right gestures, some of the participants twisted their body toward the desired direction. For up/down gestures, some of the participants mentioned different body parts (body/head), but performed nearly indistinguishable gestures. Figure 4.16 shows the selected gestures and all of them have the fixing-tool property as described in Section 4.3.2.

Moreover, the number of upper-body gestures, which are articulated by the body and head, is significant when comparing with the number of hand gestures, especially for the six movements of the end effector. These gestures can be valuable, because they can be used as a backup for the selected hand gestures to achieve greater flexibility or as the main gestures in situations where both hands are occupied and cannot be used to articulate gestures.

As already mentioned in the primary gesture section, the directions of the forward and backward gestures in Figures 4.16a and 4.16b can confuse a user and the user should be informed to ensure safety.

4.5 Discussion

4.5.1 Results

The results show that the selected soldering task encouraged the participants to express wide range of gestures for manually controlling the virtual helping hand robot such as hand, body, and various body parts

(a) The hand gesture for moving the end effector forward.

(b) The hand gesture for moving the end effector backward.

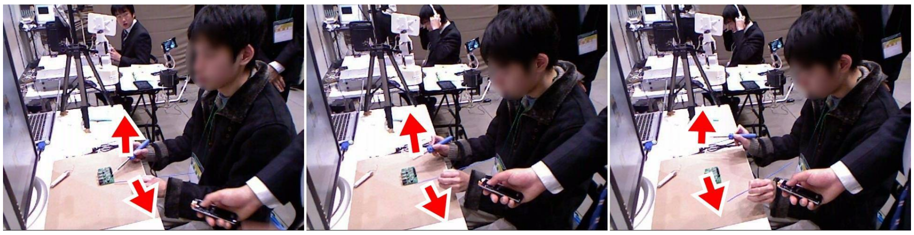
(c) The hand gesture for moving the end effector left.

(d) The hand gesture for moving the end effector right.

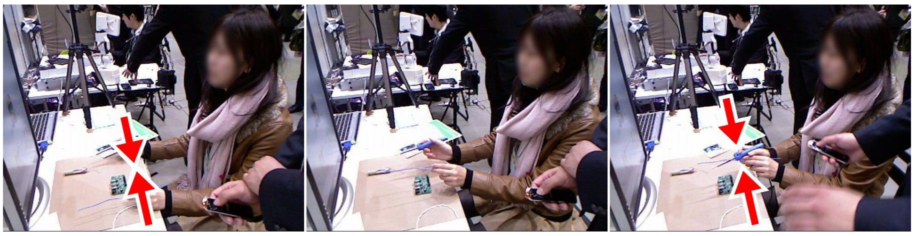
(e) The hand gesture for moving the end effector up.

(f) The hand gesture for moving the end effector down.

Figure 4.14: The selected primary gestures for the end effector movements



(a) The hand gesture for opening the gripper.



(b) The hand gesture for closing the gripper.

Figure 4.15: The selected primary gestures for the gripper commands.

gestures. In this study, the differences of precision needed for manually controlling/adjusting a hobby level soldering task should not be significant when comparing with other handicraft tasks that require a support from a helping hand robot such as gluing, sanding, drilling, and etc.

The results also show that hand gestures were the dominant gestures for controlling the basic movements (forward/backward, left/right, up/down) and the gripper open/close commands of the end effector of the virtual helping hand robot. In the example soldering task, the participants interchangeably used both one- and two-handed gestures to control the direction, but used only two-handed gestures for the open-close gripper commands. Some of the participants held objects (the soldering iron and the cable) between their thumb and index fingers before opening their hands to articulate gestures. A few participants pointed their fingers toward the desired directions during the gesture articulation.

One interesting finding from the group of participants is the dominant of the hands gestures. This is different from our expectations even though all participants were asked to pretend that they were soldering. We expected that the participants will try to use other body parts to perform the gesture, but it turned out that the participants did not care much about what they were holding in their hands. If an object is not fixed to a table and a tool is manipulable, the participants have a trend to use hand(s) for gesturing.

The alternative choices for controlling movements of the end effector in the soldering task are body and head gestures. Most of the

(a) The body gesture for moving the end effector forward.

(b) The body gesture for moving the end effector backward.

(c) The body gesture for moving the end effector left.

(d) The body gesture for moving the end effector right.

(e) The body gesture for moving the end effector up.

(f) The body gesture for moving the end effector down.

Figure 4.16: The selected secondary gestures for the end effector movements.

gestures made by this group of gestures can easily match with their corresponding end effector movements. However, for some participants, the body parts (torso and head) used for gesture articulation were not obvious and could not be easily distinguished without additional information provided by the participants, especially in the case of up/down gestures. The examples are shown Figure 4.17.

(a) The participant stated that head gesture was used for moving the end effector down.

(b) The participant stated that body gesture was used for moving the end effector down.

Figure 4.17: The down gestures those used of body parts cannot be easily distinguished between head and body.

The selected gestures provide a general idea and a guideline for preparing an initial set of gestures for an application that shares the characteristics of the soldering task. A hybrid assembly system in a manufacturing environment [6] is an example of the system that requires both automatic and manual operation of a robot. In such systems, the use of a simple helping hand to hold an object and move it according to a worker's preference could become easier if the direction and magnitude of the end effector motion could be directly controlled with natural gestures such as hand or body gestures. To be more specific, if only speech commands could be used, it would be tedious to instruct a robot to move in a 3D space without the use of gestures that are easier for conveying complex spatial information.

4.5.2 Video-based HRI

A video based experiment is a viable method for HRI study when the real system is in implementing stage and interactions between human and robot have to be evaluated during the process [46]. Studies from Dautenhahn's groups [44, 47] and other groups such as [33] and [48] addressed the use of video-based or simulation-based robots in their experiments.

Woods et al. [44] showed that results from video-based and real robot were equivalent in most aspects in the experiment about how should a robot approach human. The recent HRI experiment that uses the video-based method by Walters et al. [49] reported the following outcomes.

> The study also confirms that the video-based HRI (VHRI) methodology provides a valuable means to obtain early user feedback, even before fully working prototypes are available. This can usefully guide the future design work on robots, and associated verbal and nonverbal behaviors.

We are aware that the VHRI is not suitable for all HRI experiments. If the results depend on physical properties of the robots or the systems such as sound, smell, temperature, weight, and so on, it is inevitable to employ a real robot in the experiment. However, if the results depend on only visual feedback, such as our study, there should be no significant difference between the video and the actual robot experiments [33, 48, 49, 41].

Therefore, results from our video-based study should be equivalent with real robot system in most aspects, especially the gestures that the participants had performed. However, the real robot is still required for evaluating and confirming the results in the later development stage as we have done in this study [7, 46].

4.5.3 Pilot study

At the early stage of the gesture surveying study, we took a pilot survey with some of the laboratory members in the fourth quarter of 2011. The pilot study focused on gestures for manually controlling an industrial robot while holding tools/objects with both hands. The actual industrial robot (see Section 5.3.1) was used for demonstrating the movements and motions of an end effector of the robot.

The discussions about gestures and attempted gesturing efforts from both the author and the laboratory members showed that even simple the movements such as forward/backward, left/right, and up/down could generate a divert range of gestures from a small number of test subjects. Some of them were simple like the selected hand and body gestures (Section 4.4). Some of them were very difficult to perform or even remember like the outliers (Section 4.3.6).

These findings emphasize the need of the finding of generic gestures for manually controlling the movements of the helping hand robot because the system designer designed gestures might not suitable for the generic audiences.

4.6 Summary

The obtained experimental results reveal important characteristics of gestures in carrying out collaborative tasks between users and robotic arms in the desktop workspace. The details in Section 4.3 and the selected gestures in Section 4.4 could be used as guidelines for developing gesture recognition systems. The implementation of a real helping hand robot system for evaluating the selected gestures is described in the next chapter.

Chapter 5

Working with a helping hand robot

5.1 Introduction

With the results from Chapter 4, we continued the study by implementing a real helping hand robot system (Figure 5.1) to test the usability of the selected gestures from Section 4.4.

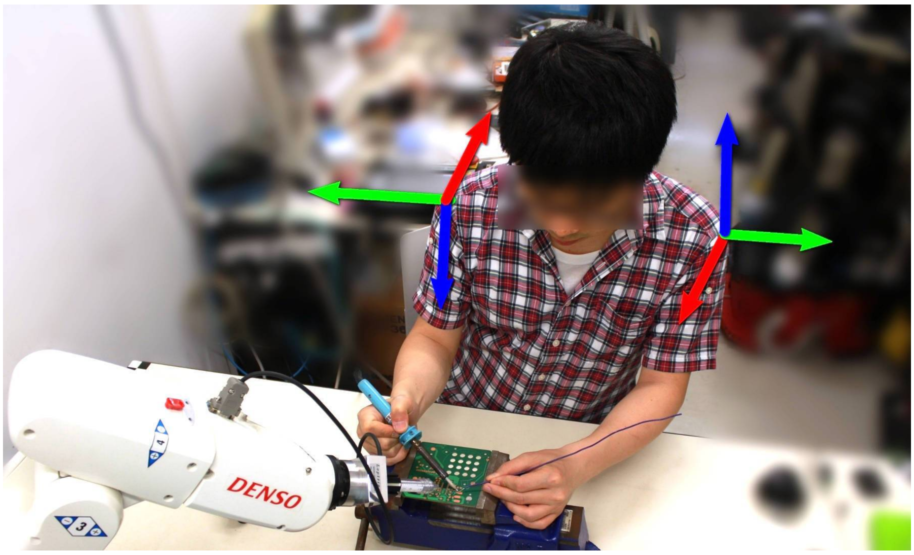


Figure 5.1: The user can tilt body forward/backward, left/right, and up/down to manually control and adjust position of the robot while both hands are occupied by the task.

In Section 4.4, the video based study collected gestures using the user centric study based on the methodology from [13]. The selected gestures were defined by the participants for manually controlling the basic movements (forward/backward, left/right, and up/down) and gripper commands (open and close) of the end effector of the virtual helping hand robot. The following summarizes important findings for gesture selection and the implementation of the recognition system when gestures are articulated while both hands are occupied during the soldering task.

- Hand gestures were the dominant gestures.
- Body gestures were the second most preformed gestures.
- Many participants who were holding objects in their hands from the beginning of the task would perform the gestures without releasing the objects.
- Many participants used left, right or both hands interchangeably for articulating hand gestures.
- The reversible gestures such as left and right gestures were consistently performed by the participants using the same body part. Only a few participants used different body parts for the pair of reversible gestures (see Figure 4.12).

In addition, experiences from the real industrial robot also show that trying to manually control an end effector with six degrees-of-freedom (DOFs) is not a trivial task. Therefore we have to observe the methods used by various systems to select and prepare a set of methods for manually controlling the helping hand robot for comparing with gestures in the experiment.

For the industrial robot, a teaching pendant allows its user to manually control an end effector by providing separated control buttons for controlling each joint (joint space control) or axis (Cartesian space control) as shown in Figure 5.2a. In 3D CAD/CAM software, it is common to expect an option for decoupling translation from orientation when manipulating objects inside the 3D scene and the software usually uses a virtual handle for a mouse and keyboard control as shown in Figure 5.2b. Interestingly, humans are also likely to transfer or move an object to a desired destination before/after aligning its orientation [50].

These observations and a trial-and-error testing with various additional control methods such as the 3D interactive marker in ROS [51] (Figure 5.2b), a 6 DOF 3D mouse from 3DConnexion[1], and a tactile

---

[1] http://www.3dconnexion.com/products/spacemouse.html (accessed December 2015)

(a) (b)

Figure 5.2: (a) The teaching pendant that allows the users to control each joint or axis separately by using the buttons in red rectangle. (b) The interactive marker for controlling the end effector of the robot in ROS (rviz). The arrows and rings are used for translation and orientations controls, respectively.

control that utilizes a force–torque sensor on the implemented helping hand robot system gave the following intuitive information about how to manually control an end effector.

- It is more intuitive when the translation and orientation controls cab be controlled separately.
- The translation and orientation controls should be able to switch between workspace and tool frames.
- Controlling the orientation of an end effector in the workspace frame is not intuitive.
- An individual axis (X, Y, Z) or a moving plan in 3D scene (XY, XZ, YZ) should be selectable by a user for the translation control.
- The orientation control is more intuitive and easier to handle when each rotation axis (roll, pitch, and yaw) is controlled separately.

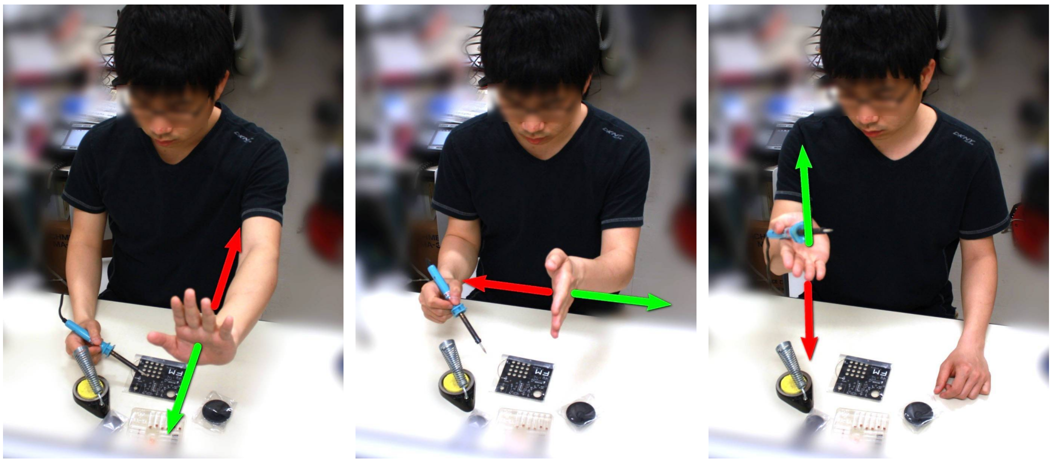
In this chapter, Section 5.2 explains selected gesture from Chapter 4. The implementation of the helping hand robot system is shown in Section 5.3. The experiment and the results are explained and showed in Sections 5.4 and 5.5). The last two sections discuss (5.6) and summarize (5.7) the results.

5.2 Gestures

In the real robot experiment, we focused mainly on how to allow a user of the system to manually control an end effector when one or both hands are occupied by the example soldering task. With results from Chapter 4, we selected hand and body gestures for manually controlling the translation of the end effector in the experiment.

The selected body gestures (Figure 5.1) allow the user to control translation movements of the end effector without interruption while dealing with tasks using both hands. However, the body gestures are limited by working postures for example, being seated. Although it is possible for the user to control both translation and orientation with only body gestures (e.g. twisting the body for controlling the yaw motion of the end effector), the preliminary testing indicates that it is difficult to maintain a good eye–hand coordination when comparing with the body gestures that only employ a simple tilting of the body for the translation control. Hence, the orientation control with body gestures was omitted.

To overcome the limitations of the body movement gestures, one- and two-handed gestures have been selected to allow the users to control translation (one-handed) and orientation (two-handed) of an end effector from any position in the workspace. The one-handed gestures for the translation control are shown in Figure 5.3.

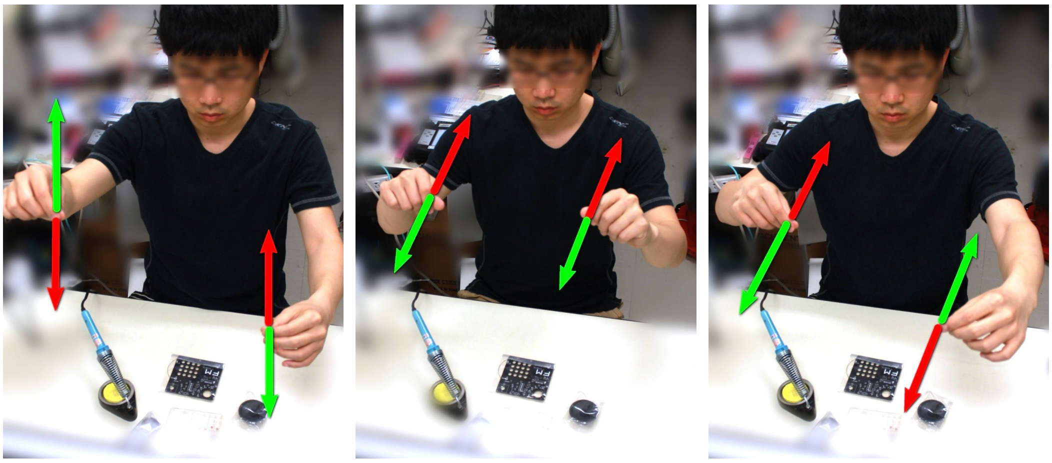


(a) (b) (c)

Figure 5.3: The one-handed gestures for translation control. The user can use either left or right hand to control an end effector while holding objects in one or both hands. (a) Forward/backward. (b) Left/right. (c) Up/down.

For the orientation control, a set of two-handed gestures (Figure

5.4 that resemble the action of holding a sheet of paper in both hands and flipping or rotating the paper around the X, Y, or Z axes were chosen. The implemention was based on hand gestures for CAD/CAM systems proposed by Wang et at. [52]. However, in this study, only translation control of the hand gestures was evaluated to be able to compare the results with the body gestures.



(a) Roll. (b) Pitch. (c) Yaw.

Figure 5.4: The two-handed gestures for orientation control. The user can use both hands to control the orientation of an end effector with a "sheet of paper" metaphor while holding objects in both hands. Note that pinching is not a necessary condition for gesture articulation.

In addition, as mentioned in the introduction of the chapter, there is a requirement for toggling between working space and tool frames while the user is manually controlling the end effector. This requirement usually arises when the user needs to move the tool along its main axes (e.g. feeding a solder wire which is aligned with the Z axis of the tool and $-Y$ axis the end effector) or the plane of the printed circuit board (e.g. changing a solder point) as shown in Figure 5.5.

With the need of the frame changing, we selected the flapping elbow gesture from the outlier in Section 4.3.6 as a toggle signal to allow the user to switch between the working space and tool frames while working with both hands (Figure 5.6).

Detailed discussion about the gestures recognition is described in Section 5.3.3.

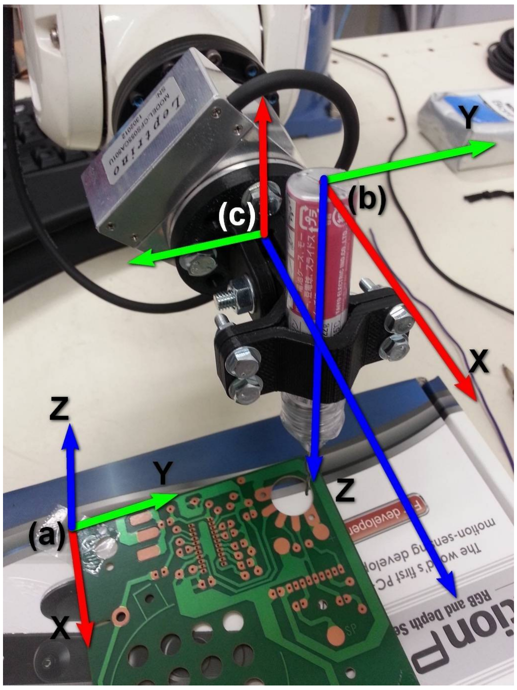
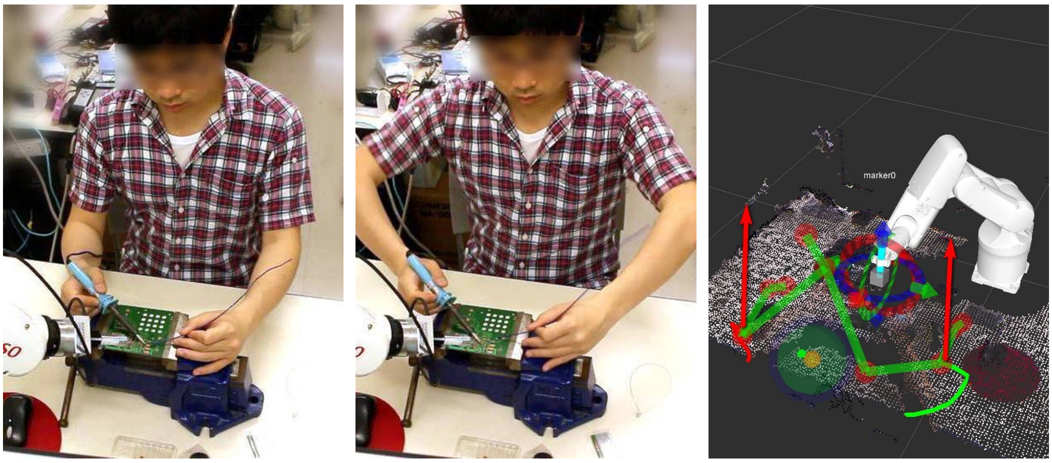


Figure 5.5: The example coordinate configuration of (a) the working space, (b) the tool, and (c) the end effector.



(a) (b) (c)

Figure 5.6: The flapping elbow gesture for switching between the workspace and tool frames.

5.3 Helping hand robot system

5.3.1 Hardware

The helping hand robot is a 6-DOF Denso VP-6424G industrial robot. Is is mounted on a table as shown in the lower right of Figure 5.7. Two Kinect sensors are used as the main sensors. The first Kinect is mounted over the workspace and connected to a main PC for data processing (Kinect 1 in Figure 5.7). Its raw point cloud and image data are used for objects detection, hand gestures recognition, and workspace calibration. The workspace calibation is based on a work from [53]. The point cloud library (PCL) [54] is used for processing the point cloud data from the first Kinect.

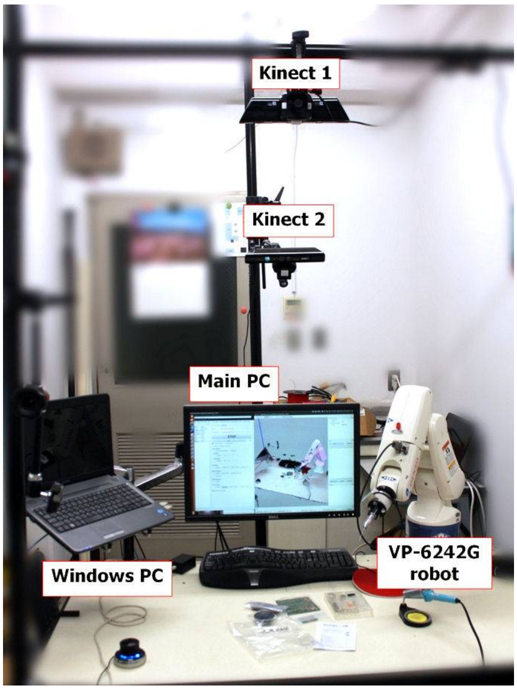


Figure 5.7: The helping hand robot setup.

The second Kinect is mounted in front of the workspace and pitched downward for detecting the upper body of a user (Kinect 2 in Figure 5.7). It is used for recognizing an upper body skeleton with Microsoft Kinect SDK[2] on a Windows PC because the SDK was available only on the Windows system at the time of the study. The recognized skeleton information (e.g. joint positions) is sent to the main PC for body gesture recognition. Detailed information about gesture recognition and algorithm will be discussed in Section 5.3.3.

---

[2] https://dev.windows.com/en-us/kinect (accessed December 2015)

The main PC is a Linux system with a robot operation system (ROS) [55] installed. The main PC handles all interactions between the user and the helping hand robot. After the target position of the end effector of the robot is computed from the interaction between the user and the system, the trajectory of the robot is generated and send to a real-time Linux PC to convert to joint commands to transmit to a robot controller at 1000 Hz. The need of the separated real-time Linux PC is caused by the computation load of the main PC that prevented it from sending joint commands with less than 2 ms jitter that is required by the robot controller.

A diagram of the system is shown in Fig. 5.8. All source codes of the implemented system are open-source and available online[3].

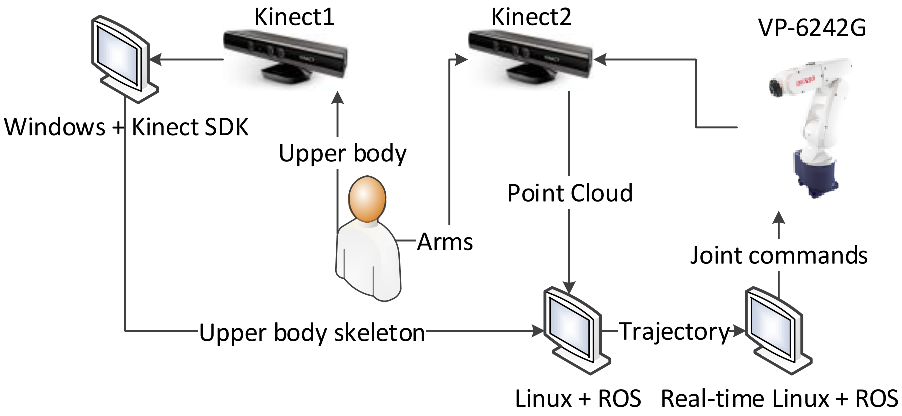


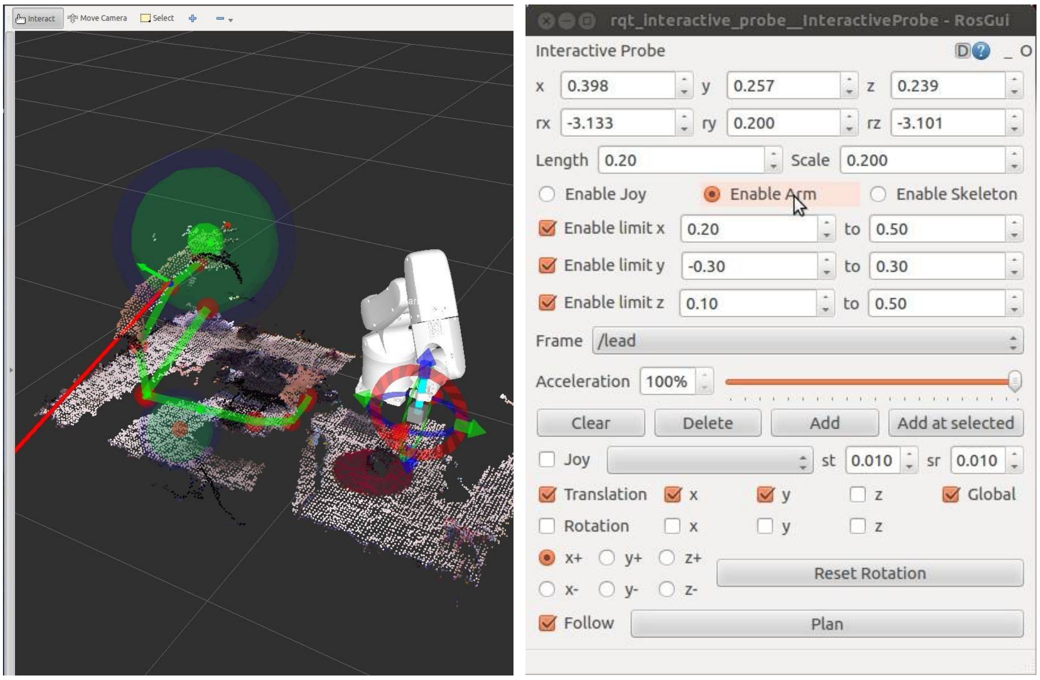
Figure 5.8: The diagram of the components and connections in the helping hand robot system.

5.3.2 Software

Software of the system is implemented with the C++ and Python languages. The user interface (UI) is based on a 3D visualization tool for ROS called rviz [56]. In rviz (Figure 5.9a), a user can perform common 3D CAD/CAM interface controls such as pan, tilt, zoom, and rotate the 3D scene to align and match with user's preferences and controlling methods. Robot states such as positions of joints are updated in real-time using data transmitted from the robot controller. The updated robot states are displayed with a 3D model of the robot in rviz (the white mesh in the right of Figure 5.9a). The real-time updated robot model is also used as a supplementary virtual feedback for the user to confirm states of the robot during the experiment.

---

[3] https://github.com/hiveground-ros-package (accessed December 2015)



x 0.398 y 0.257 z 0.239
rx -3.133 ry 0.200 rz -3.101
Length 0.20 Scale 0.200
○ Enable Joy ● Enable Arm ○ Enable Skeleton
☑ Enable limit x 0.20 to 0.50
☑ Enable limit y -0.30 to 0.30
☑ Enable limit z 0.10 to 0.50
Frame /lead
Acceleration 100%
☐ Joy st 0.010 sr 0.010
☑ Translation ☑ x ☑ y ☐ z ☑ Global
☐ Rotation ☐ x ☐ y ☐ z
● x+ ○ y+ ○ z+
○ x- ○ y- ○ z-
☑ Follow

(a) (b)

Figure 5.9: (a) The 3D rviz screen. (b) The control panel of the system.

An interactive marker at the end effector of the robot in Figure 5.9a (magnified in Figure 5.2b) is used to manually control or set up an end effector of the robot [51]. The user can drag the arrow or dial the ring to perform translation or orientation control. The interactive marker is used as one of the testing conditions in this experiment as a traditional mouse and keyboard control method. Obstacle avoidance, self-collision checking, inverse kinematic solving, and other related functions for controlling the industrial robot are based on ROS and MoveIt! libraries [57].

A control panel in Figure 5.9b is mainly used for setting up the robot system and selecting the interaction mode (the interactive marker, hand gestures, or body gestures). The user can also select a desired working frame (tool, workspace), axis (X, Y, Z), or plane (XY, XZ, YZ) on the control panel while interacting with the robot.

5.3.3 Gesture recognition

In this study, we proposed a rubber band model for implementing body and hand gestures recognition. The model allows the user to start and stop controlling an end effector at any point in the workspace. Gesture recognition states and information such as hand positions and a recognized user skeleton are also displayed in the rviz display as shown in Figure 5.10.

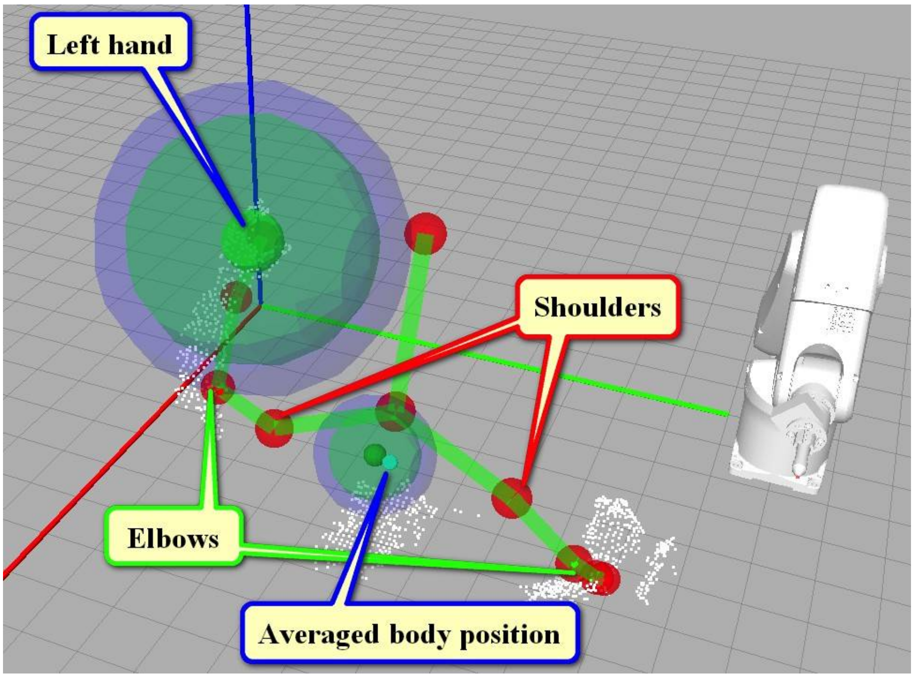


Figure 5.10: The upper body skeleton is detected with Microsoft Kinect SDK and displayed as the red dots connected with the green lines. The shoulder joints are averaged as the body position. The elbows joints are used to detect flapping elbow gestures. The detected left hand is initialized at the small green sphere.

The proposed model can be visualized using the metaphor of tying an object (e.g. hand) to a pivot point (initial position) with the rubber band. When the user moves the hand inside the workspace, the initial position (the smallest circle in Figure 5.11a) moves with the hand until it is held for a certain time for initialization (a green circle at the left hand in Figure 5.10). After the initialization, the position will be fixed as a pivot point for gesturing. At this state, the user can articulate gestures to control the end effector within the area between the middle and the large circles (Figure 5.11b). If the user wants to stop controlling, the user can either move the hand back to the pivot point for resting or move the hand outside the large circle to completely terminate the control (Figure 5.11c).

In other words, at the initial state (Figure 5.11a) the rubber band is not stretched enough to enable gesture control. This allows a gesture recognizer to deal with a noisy position measurement and unintended initialization. When the hand is moved further from the pivot point after initialized (Figure 5.11b), the direction and length of the rubber

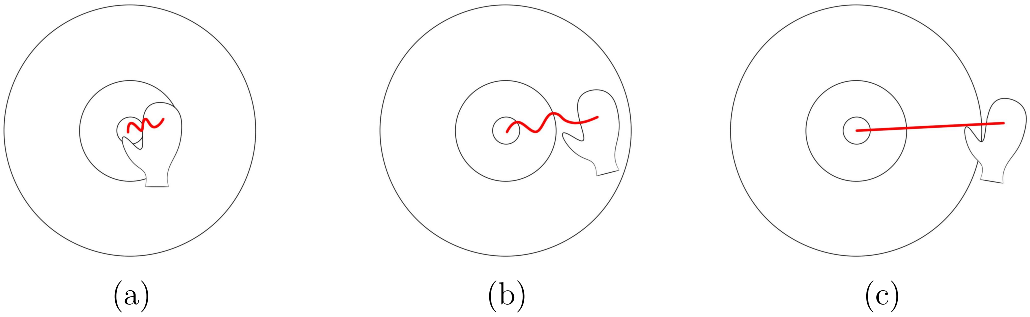

(a) (b) (c)

Figure 5.11: Three states of the rubber band model. The red string is for visualizing the rubber band. (a) Initial/idle (b) Control (c) Cancel

band can be used to control direction and velocity. The rubber band will rupture and a replacement (re-initialization) is needed if the hand is moved too far from the pivot point (Figure 5.11c). A certain initialization time is needed before starting to control the end effector to ensure that the manual control is intentionally activated.

Hand gestures

The selected hand gestures are recognized by functions in the point cloud library (PCL). The arm-like point cloud clusters are classified using the principal component analysis (PCA) function by searching for the elongated objects (e.g. a long point cloud cluster of an arm) that are floating above the working space (desk). All points those belong to the structure of the robot are filtered out with occupancy map monitoring functions in the MoveIt! library. This filtration helps the arm-like point cloud clusters become easier to detect. Hand positions are computed from the clusters of the point cloud near the end of the arm-like cluster as shown in Figure 5.12.

The detected hand positions are smoothed by the descrete Kalman Filter functions from the OpenCV library [58]. The filter smooths positions and velocities ($x_k = [x, y, z, v_x, v_y, v_z]^T$) of the hands by using a state transition matrix (F_x) in Equation 5.1. The non-constant velcity or accleration of the hand motion is modeled by a white noise assumption based on a report from [59]. The time step ($\mathrm{d}t$) is set to 30 Hz according to frame rate of the Kinect sensor. Inputs for the filter measurement update are the centroid (x, y, z) of the point cloud

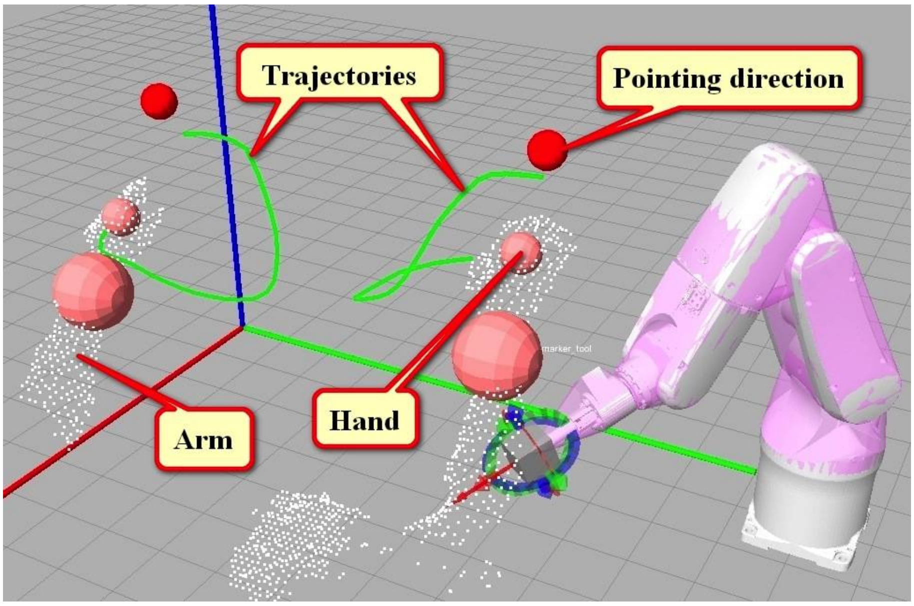


Figure 5.12: The detected hands inside the workspace. The white dots are the point cloud clusters of all the objects in the workspace. Hand positions are the small pink spheres. The smoothed hand trajectories are displayed with the green lines.

of each hand (the small pink spheres in Figure 5.12).

$$F_x = \begin{bmatrix} 1 & 0 & 0 & \mathrm{d}t & 0 & 0 \\ 0 & 1 & 0 & 0 & \mathrm{d}t & 0 \\ 0 & 0 & 1 & 0 & 0 & \mathrm{d}t \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix} \tag{5.1}$$

The filter is tuned to ensure a balance between the smoothness and responsiveness of detected hands. Additional information about the Kalman Filter can be found in [60]. Furthermore, we also implemented a state machine to track, update, correct, and reset state of the filters when the participant is moving hand(s) inside, into or from the working area. From the recorded videos, the participants can accurately control the tip of the solder lead down to the smallest distance between two IC pins (100 mils or 2.45 mm) in Figure 5.5.

Using the rubber band model, spheres around the left hand in Figure 5.12 are the visualization of the model and used as feedback information for the user while interacting with the robot. The gesture recognizer interprets one- and two-handed gestures for translation (Figure 5.3) and orientation controls (Figure 5.4), respectively.

Body gestures

The body gestures are recognized using skeleton detecting functions in the Kinect SDK. The detected skeletons are filtered and smoothed using the built-in functions of the SDK before being sent to process and display on the main PC. In the current implementation, only wrists, elbows, shoulders, neck, and head joint information as shown with the small red spheres in Figure 5.10 are used.

The left and right shoulders of the skeleton are averaged as a reference point for body gestures recognition as shown in Figure 5.10. The shoulder joints are used because both joints are the most stable for detecting upper body motions with the SDK version 1.5. The spheres between the left and right shoulders in Figure 5.10 are the visualization of the rubber band model and are used as feedback information for users.

When the participants tilt, lean, or twist their bodies, the averaged shoulder position will move from its initial position. This can be visualized by replacing hand in Figure 5.11 with the averaged position of the shoulder at the neutral seating position. The displacement and direction of the averaged position are used for computing the moving direction and speed of the end effector.

Toggle gestures

The gesture for toggling between workspace and tool frames is recognized by detecting a flapping movement of the elbows (Figure 5.6). The detection is based on a one-shot state machine that uses elbows joints displacement and direction as it inputs (arrows in Figure 5.6c). A completed up and down cycle of the elbow joint is needed for triggering the state machine to output the recognized toggle gesture.

5.4 Experiment

5.4.1 Objectives

The experiment focused on the usability testing to validate the selected gestures with the real helping hand robot.

5.4.2 Setup and environment

We set up a soldering task to compare the selected hand gestures and body gestures by using the interactive marker as a reference. The experiment was conducted in the laboratory room as shown in Figure 5.13.

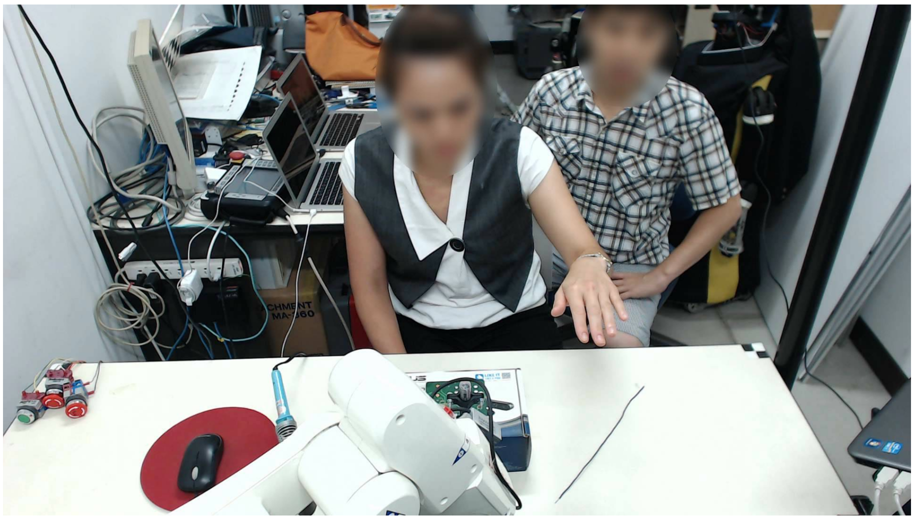


Figure 5.13: The scene captured from the experiment while the participant was controlling the robot with hand gestures. The author was monitoring the experiment behind the participant.

The author stayed close to the participant during the experiment to monitor and ensure safety of the participant. The experiment was a one by one session for each participant. The participants were asked to come to the experimental room based on their convenient schedule during the experimental week. System setup and software were not changed between each participant. The author also took notes during the experiment and all sessions were video recorded with two cameras from the front and left views of the participant for the further analysis.

5.4.3 Participants

Eight participants, all students of the Keio University, volunteered for the experiment. Three of them were women. The average age of the participants was $26.3, SD = 2.2$. All participants were familiar with computer systems but had no experience with an industrial robot. They had experience with the soldering task before the experiment. All participants had experience with 3D games or 3D software. They had gesture control experience with a modern game console such as Wii, Xbox, or PlayStation.

5.4.4 Procedure

The experiment began with an explanation of the purpose of the study before a demonstration of the selected gestures and system usage by the author. After the introduction, the author demonstrated how to manually control an end effector with body gestures, hand gestures, flapping elbows gestures, and the interactive marker. After the demonstration, the participants practiced the use of all the gestures and the interactive marker to ensure that they knew how to control the helping hand robot manually using all methods (Figure 5.14).

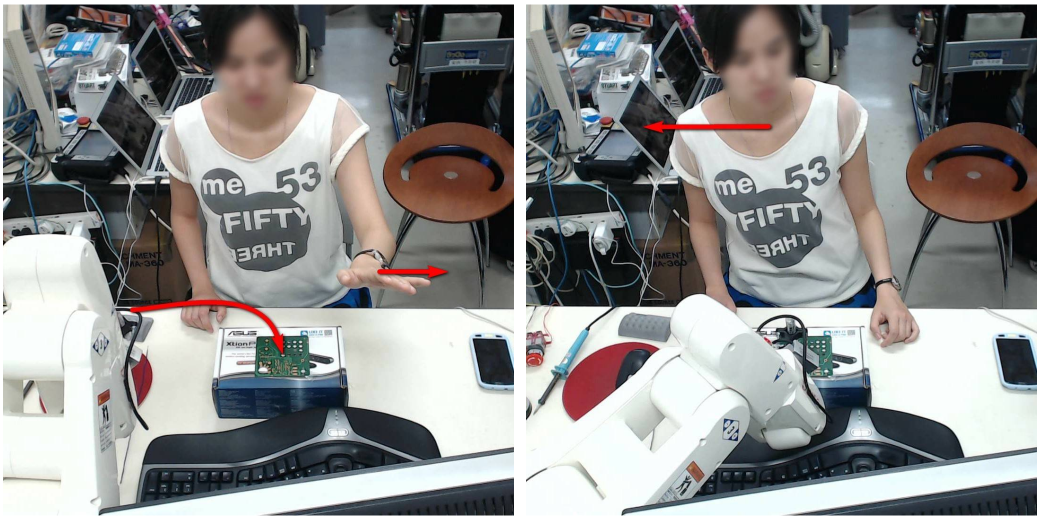

(a) (b)

Figure 5.14: The participant practiced before the beginning of the experiment. (a) Controlling the robot from its start position with hand gestures. (b) Controlling the robot above the circuit board area with body gestures.

When finished the practice session, the participants were asked to perform a simulated soldering task using an unplugged soldering iron. To be more specific with the manual control of the helping hand robot, we divided the manual control in the experiment into two steps for evaluating hand gestures, body gestures, and the interactive markers. The separated control steps are for testing the suitability of each control method in different conditions and done in sequence.

The setting up step is a situation that the user of the system is trying to manually moving the end effector from its initial position to a target working area.

The controlling step is a situation that the user of the system is trying is moving the end effector around in the working area based on the task's requirements.

In other words, in the setting up step the participants were asked to set up the end effector by manually controlling the robot from the start position to the area above the circuit board as shown in Figure 5.14a. In the controlling step, on the other hand, the participants were asked to control the end effector above the working area and solder a cable to three specified points on the circuit board with the help from the manually controlled robot (Figure 5.15).

The participants were not explicitly asked to hold the soldering iron and the cable in their hands before performing the experiment; this was to observe how the participants grasped and released objects in three soldering task trials using the interactive marker, hand gestures, and body gestures. The order of the experiment was not randomized because all participants already practiced all controlling methods (marker, hand and body) under supervision of the author before participating in the experiment.

After the experiments, the participants answered a questionnaire and discussed their opinions and suggestions for the system with the author.

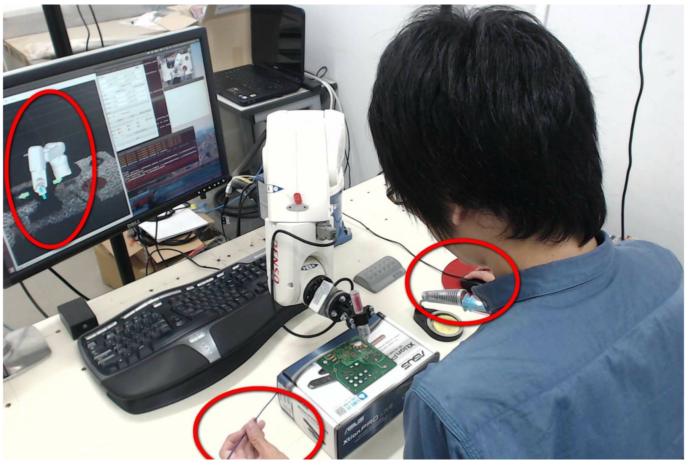
5.4.5 Metric

The participants were asked to rate their experiences using a set of seven-point Likert scales (1–disagree to 7–agree) and answer demographic questions after the experiment. The Likert scales begin with three pairs of scales for measuring opinions about the proposed gestures. The scales can be read as

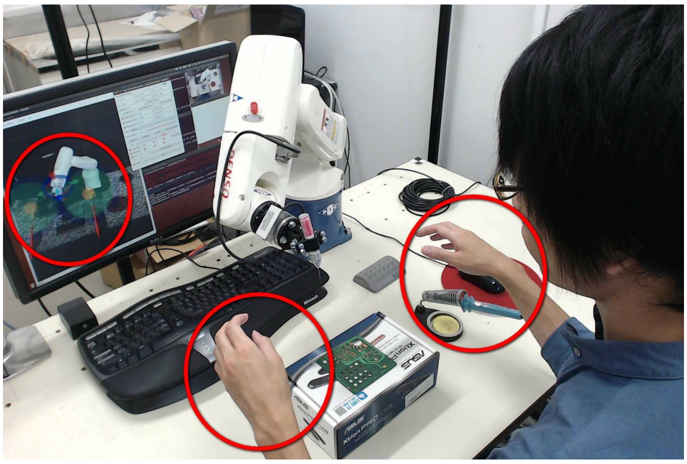
- "x gestures are suitable for the purpose"
- "x gestures are easy to remember and use"

where "x" are "hand", "body, and "elbow". The purposes of hand, body, and elbow gestures are translation and orientation control, translation control, and working frame toggling, respectively.

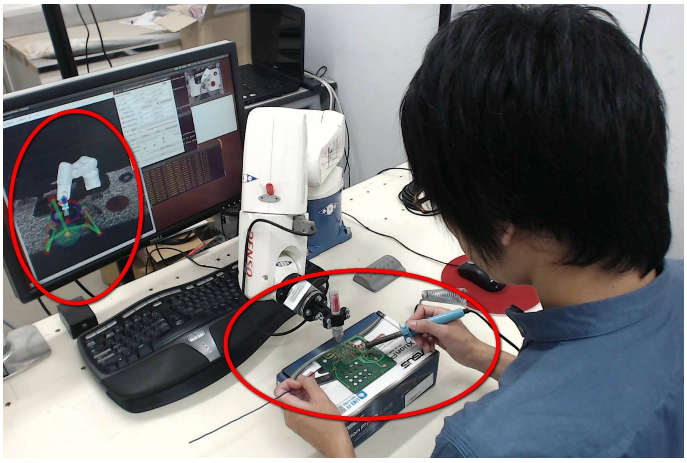
The questionnaire continues with six additional Likert scales for comparing the use of the hand gestures and body gestures with the interactive marker in both steps of the manual control. The scales are divided into two groups those can be read as



(a) The interactive marker.



(b) Hand gestures



(c) Body gestures

Figure 5.15: The participant tried all methods for manually controlling the helping hand robot during the hands occupied situation.

- “x is suitable for manually controlling the robot from the start position”
- “x is suitable for manually controlling the robot during the soldering task”

where “x” are “hand, “body, and “interactive marker”.

The questionnaire also asks if “it is acceptable to change the method for controlling the robot during the task”, for example, switching between hand and body movement gestures and the interactive marker as the participant see fit.

5.5 Results

5.5.1 Statistical results

The average score from the Likert scales indicates that hand gesture $[(M = 5.6, SE = 0.32), (M = 5.8, SE = 0.31)]$, body gestures $[(M = 5.9, SE = 0.40), (M = 5.8, SE = 0.49)]$, and the flapping elbow gesture $[(M = 6.1, SE = 0.35), (M = 6.1, SE = 0.40)]$ are suitable for their purposes. They can be remembered and used without difficulty. The average scores and standard error bars are shown in Figure 5.16.

Because each participant performed all manual control methods, we conducted the one-way within-subjects ANOVA to compare the preferences of the participants regarding manual control in the setting up and controlling steps. The post hoc analysis adjustments are based on the Bonferroni method.

The average scores of the setting up and controlling steps of the manual control are shown in Figure 5.17 and the differences between two steps were found to be statistically significant at the $p < 0.05$.

For the setting up step, there is a significant difference between the control methods, $F(2, 14) = 9.00$, $p < 0.05$. The post hoc analyses (Table 5.1) indicates that the interactive marker $(M = 6.4, SE = 0.18)$ is preferred over the hand gestures $(M = 4.5, SE = 0.50)$ and the body gestures $(M = 4.1,\ SE = 0.61)$ with statistical power $(\beta - 1)$ greater than 0.8. Hand gestures are slightly more preferred over body gestures, but the differences are not statistically significant in the setting up step with power less than 0.2. The statistical power is used to determine the type II error rejection of the test. Normally, $(\beta - 1 < 0.2)$ is too weak and $(\beta - 1 > 0.8)$ is strong enough for validating the study.

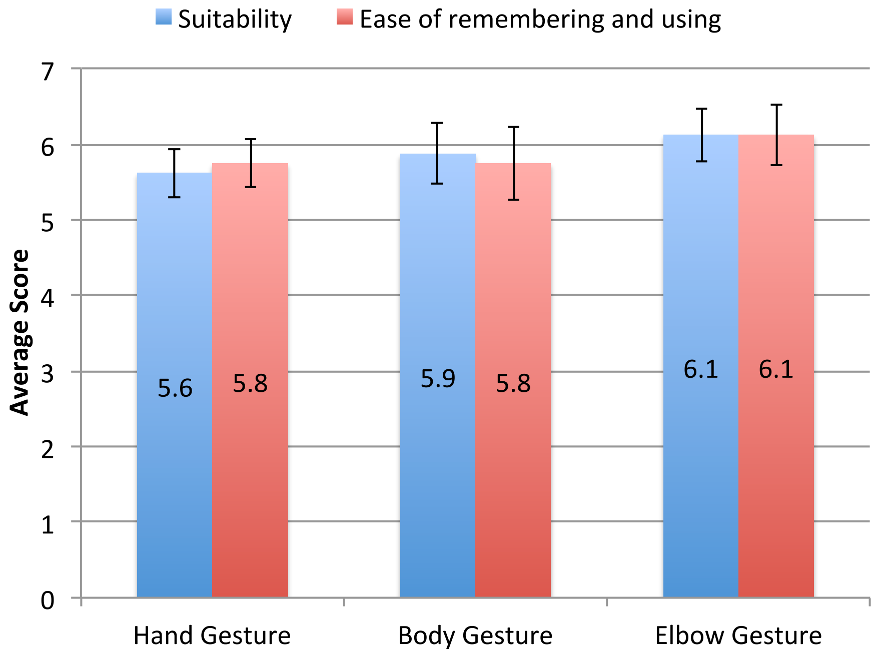


Figure 5.16: The average scores and the standard errors of the suitability and ease of remembering and using of the hand gestures, the body gestures, and the flapping elbow gesture.

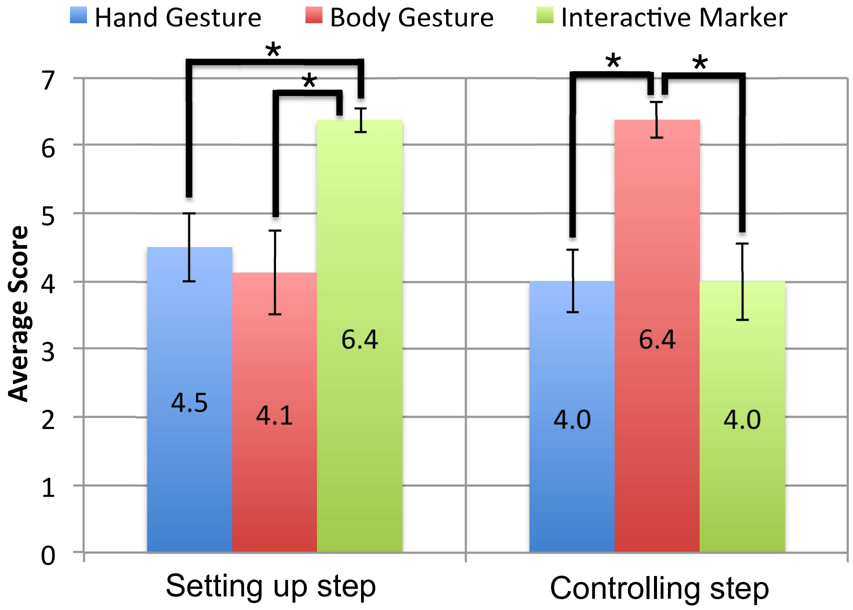


Figure 5.17: Average scores of the suitability of the selected gestures and the interactive marker in the setting up and controlling steps of the manual control experiment.

Table 5.1: T-test results of the setting up step. The (*) indicates that the difference is significant.

Pair	Result
Hand & Body	$t(7) = 0.75,\ p = .24, (\beta - 1) = 0.15$
Hand & Marker*	$t(7) = -3.91,\ p < .01, (\beta - 1) = 0.99$
Body & Marker*	$t(7) = -3.21,\ p < .01, (\beta - 1) = 0.98$

Table 5.2: T-test results of the controlling step. The (*) indicates that the difference is significant.

Pair	Result
Hand & Body*	$t(7) = -5.16,\ p < .01, (\beta - 1) = 1.00$
Hand & Marker	$t(7) = 0.00,\ p = .50, (\beta - 1) = 0.05$
Body & Marker*	$t(7) = 3.64,\ p < .01, (\beta - 1) = 1.00$

For the controlling step, there is a significant difference ($F(2, 14) = 10.93,\ p < 0.05$) between the body gestures ($M = 6.4,\ SE = 0.26$), the hand gestures ($M = 4.0,\ SE = 0.46$), and the interactive marker ($M = 4.0,\ SE = 0.56$). The post hoc analysis (Table 5.2) shows that the body gestures are preferred over the hand gestures and the interactive marker with significant differences and have statistical power greater than 0.8. There was no significant difference between the hand gestures and the interactive marker in the controlling step with statistical power less than 0.2.

Furthermore, the participants also showed that they were willing to switch between control methods if it helped complete the task and made their work easier ($M = 5.8, SE = 0.59$).

The used statistical power computation is based on the post hoc power analysis that computes the archived power using mean and standard deviation of each pair of the experiment [61].

5.5.2 Video analysis results

The detailed videos analysis reveals many interesting information that can be used as a guideline for the future development.

Practicing

The average teaching time (mm:ss) used by the author for each participant was 06:30 with $SD = 01{:}32$. The male participants needed slight less time ($M = 05{:}17, SD = 00{:}58$) than the female participants ($M = 08{:}07, SD = 00{:}09$). It is quite clear from the observation that the male participants asked slightly less questions about the system and the gestures and wanted to try the system right away. However, with eight participants, this trend might be useful as a noticeable difference only.

It can be seen that most of the participants can start controlling the helping hand robot with hand gestures without notice able difficulty. But it usually took longer for the participants to get familiar with the body gestures. The interesting finding is the participants were comfortable to use to flapping arms gestures to toggle between tool and workspace frames when using the body gestures.

Interactive marker

All participants did not have any noticeable difficulty with the interactive marker control. However, there is two issues those should be addressed to improve the use ability of the 3D interface and safety of the system.

The first issue is that the robot will move very fast from its start position to its target position when the participant drags the interactive marker without concerning about moving distance and speed of the robot as shown in Figure 5.18. This issue should be addressed by limiting drag-able distance and speed of the interactive marker based on the actual robot limitations and safety regulations.

The second issue is that the robot will move toward the participant very fast if the 3D scene is almost parallel with the actual robot setup as shown in Figure 5.19. When the participant drag the interactive marker toward screen a slightly change in x or y displacement of the mouse will cause a large displacement on the marker. This will cause the robot to move very fast on a particular direction (e.g. moving forward). It is like the first situation that the limited drag-able distance and speed of the interactive marker based on the actual robot limitations and safety regulations should be imposed. Furthermore, an automatic view angle changing based on a selected moving axis might help enhance the user experience. An example is changing the 3D scene to the left or right view when the participant is trying to move the robot forward or backward.

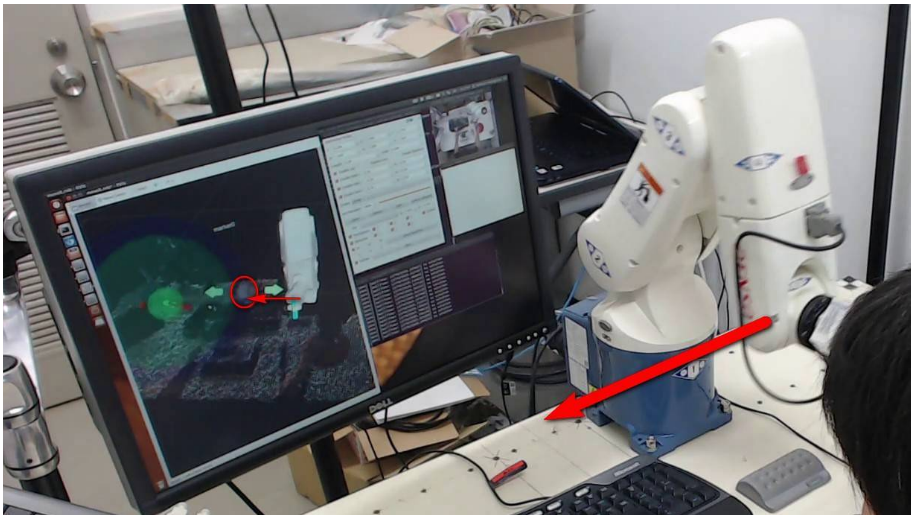


Figure 5.18: The "drag too fast" is a situation when the participant drags the interactive marker too fast and too far from its start point (the small circle and the arrow in the rviz display).

Figure 5.19: The "drag toward screen" is a situation when the participant drags the interactive marker toward the normal of the screen to move the helping hand forward.

Hand gestures

From the observation, the selected hand gestures can be used to control robot without difficulty as mentioned in Section 5.5.2. However, many participants mentioned about fatigue caused by the long hand gestures control session. Some of the participants performed hand gestures while resting their elbow on the table as show in Figure 5.20.

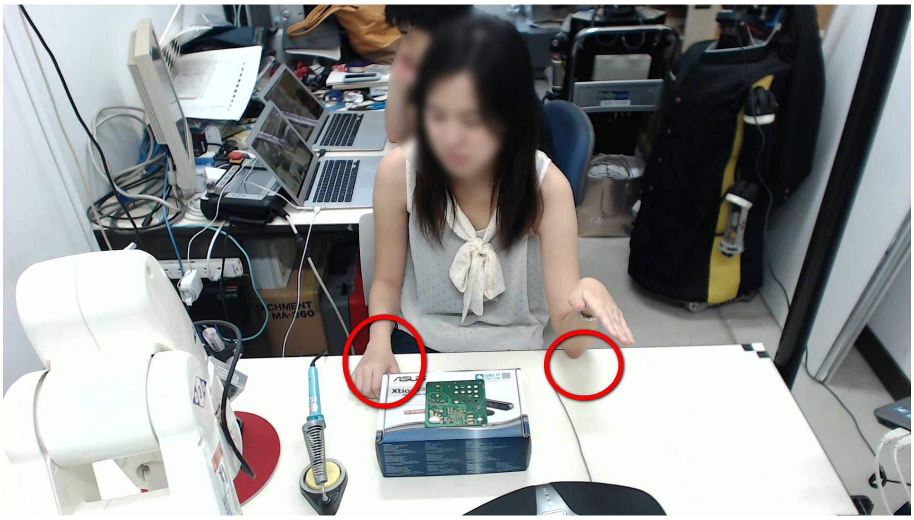


Figure 5.20: The participant was resting her elbow on the table during hand gesture controlling session.

The video data also confirms that interchangeable between left and right hand is important for one-handed gestures. In the experiment, when the robot was working in the center area of the workspace, controlling with the left or right hand was depended on the available space and the convenient of the participant (Figure 5.21). Without left or right hand limitation, the participants can freely changing hand based on the need of manual control and the possible fatigue.

Body and toggle gestures

It is clear from the videos that the participants can start controlling the robot with the body and toggle gestures without noticeable difficulty as mentioned in Section 5.5.2.

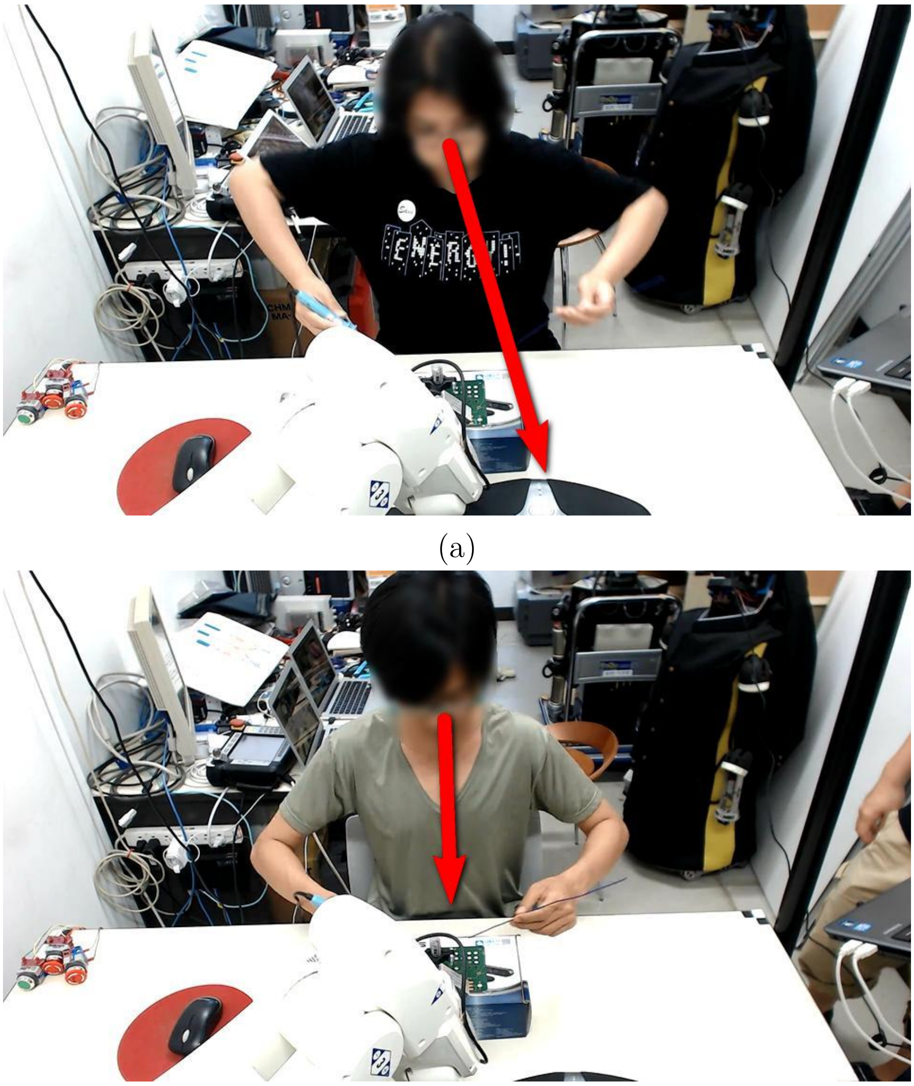
By looking at the improvement of eye-hand coordination that the body gestures might help, it is noticeable that some participants still needed to confirm their actions and system status by looking at the screen as shown in Figure 5.22. And the need of visual confirmation might not available for all situations. This informs the need of additional feedback channels such as sound, vibration, and so on for the

(a)

(b)

Figure 5.21: (a) The participant used the left hand to control the robot unit it was moved near his hand. (b) With limited controlling space in (a), the participant changed to the right hand for controlling the robot.

user of the system, particularly when the eye-hand coordination is really important such as working with dangerous chemical.


(a)

(b)

Figure 5.22: (a) The participant confirmed the toggle gesture by looking at the main screen. (b) The participant confidently performed the task without looking at the screen.

5.6 Discussion

5.6.1 Results

The participants were able to manually control the end effector with the proposed gestures without noticeable difficulty. The results show that the body gestures are preferred in the controlling step, whereas the interactive marker is preferred over gestures for setting up the end effector in the setting up step.

The behavioral observation shows that the participants have two distinguishable ways of holding and releasing objects during the experiment. For the *interactive marker* and hand gestures, all participants grasped and held the soldering iron and the cable only when they were performing the soldering task on each soldering point. The participants released the objects (put them on the table) immediately before starting to control the movement of the robot to the next soldering point. For the *body gestures*, after picking up the soldering iron and the cable for the first soldering point, only one participant released the soldering iron and the cable before starting to control the robot to the next soldering point. These findings emphasize that body movement gestures can be useful for manual tasks when there is a need to continuously hold tools and objects.

The participants also commented during the discussion that it would be more convenient if the setting up step can be performed automatically by the robot and it is acceptable if fine-tuning is needed. This qualitative data informs that the participants expected the robot to move automatically when the robot have to move in a long distance. However, gestures are acceptable when the robot is struggling in complicated situations. This comment supports the use of gestures in the example scenario in Table 1.1, which addressed a glitch in the interactive function of the robot system.

Although the selected hand gestures show no significant difference between the setting up and controlling steps, the average score of the gestures and the participants' comments still encourage the use of hand gestures as a supplementary or alternative control method when other methods are not appropriate.

5.6.2 User preferences

The helping hand robot utilizes the rubber band model (Figure 5.11) for gesture state recognition. The distance from the initial position for idle, control, and cancel states must be specified before using the system. From the observation and discussions with the participants, it is found that the participants exhibited noticeable preferences over the predefined distances. Some participants preferred small and precise displacement control, while others requested large and fast hand and body gestures control.

One of the issues from the discussions with the participants also shows that the main reason that the interactive marker was preferred over both types of gestures in the setting up step was the moving speed of the end effector that was set much slower when controlled with gestures for the safety reasons; this, in fact, can be altered to match each participant's preferences. However, additional effort for the safety and reliability of gesture recognition will be necessary.

5.6.3 Selected gestures

The gestures were derived from the results in Chapter 4 in which the experiment was conducted using the virtual helping hand robot. With the selected gestures, the participants can start and stop controlling the end effector whenever they need to. The body gestures allow the participants to control the end effector with high precision while both hands are occupied by tasks. Without restriction on hands and their initial positions, the participants can articulate one- or two-handed gestures while holding objects in their hands. The implementation also allows the participants to articulate dexterous body and hand gestures without additional devices such as gloves (e.g. [38]) or sensing devices (e.g. [37]).

5.6.4 System implementation

The current system was designed based on the ease of implementation and flexibility for the experiment. Its components are mostly off-the-shelf software and hardware. A more specific and efficient software implementation should be able to help reducing the number of hardware and system cost.

5.7 Summary

The obtained experimental results reveal the unexpected outcomes that the body gestures are preferred over the hand gestures even though the video based study in Chapter 4 suggests the hand gesture is the most performed gestures for manually controlling the helping hand robot. The details in Sections 5.5 and 5.6 can be used as guidelines for the future development of a helping hand robot that can interact with a human naturally.

Chapter 6

Limitations and future work

6.1 Limitations

The selected soldering task is an important constraint that limited the way participants can articulate gestures while sitting and holding objects in both hands. The gestures those were imposed by this constraint might not reflect all possible gestures that the participants might perform for different tasks and postures. For example, the participants might use only one hand for gesture articulation if the other hand is holding an object or is haptically connected to the robot (e.g. holding an object that is attached to an end effector of the robot).

This study describes the detailed implementation the helping hand robots that can be controlled manually with gestures. The experiment shows that the participants can freely control the end effector of the helping hand robot with the selected hand and body gestures. However, as suggested by some of the participants, an additional UI that will allow them to know the current state and status of the system, such as robot joint limits or the state of the gesture recognition, without looking at the computer screen would enhance the efficiency of the system. This suggestion implies that the implemented UI (Figure 5.9) might influence how the participants used gestures to control the robot. Although we expect that the proposed gestures for manual controlling the robot are natural and intuitive for the participants, additional studies of the system without a traditional UI are needed.

User preferences in Section 5.6.2 emphasize the need of customization functions for various aspects of the system. From the implemented system point of view, these kinds of adjustments are tedious and hence automatic calibration functions or setting methods will surely enhance the system efficiency and user experience. The rubber band model and robot moving speed are also the important topics those should be able to customize by users.

The current implementation is limited to a robot that is mounted on a table and facing its user (see Figure 5.1). Different robots and configurations, such as a robot that is mounted on a linear unit for extending the working envelope, a mobile manipulator robot, or a robot that is working side-by-side with the user, will require additional gestures and sensing effort to handle the additional DOFs and the variety of user positions with respect to the robot.

Switching between control methods is not a significant burden, as indicated by the results in Section 5.5, and hence multimodal manual control for assisting or setting up a helping hand robot with various methods, such as gestures and 3D user interfaces, could be more useful than using just one particular method. However, additional effort and further studies will be required to confirm this.

Results from the limited number of participants and the soldering task scenario can only be viewed as a guideline or hint for future study and development. Additional experiments using an extended capabilities robot system with a larger number of potential users, more end effector movements, and additional task scenarios will be required for validating and extending the results of the study.

Currently, the hand tracking function in Section 5.3.3 is relies on the raw position data of the point clouds from the Kinect sensor. Normally, the Kalman filter used for tracking and smoothing out the motion jitters is not support the non-liner nature of hand motions. However, it is used based on the assumption that the changes of translation (velocity and acceleration) are white noise [59]. The more accurate but relatively complex method such as the particle filter should be evaluated if more precise hand or object position is required.

6.2 Future work

Gestures or interaction methods that allow users to control the trajectory of the robot will open a new perspective for use of the helping hand robots. Industrial robots have already been used in various art-related domains such as cinematography, architecture, and installation arts. In such domains, expressing one's creativity through direct interaction with a robot using natural gestures might be more intuitive than tedious work with mouse clicks through a 3D virtual world in a traditional user interface.

The hand and body gestures can help specific handicapped persons such as deaf or semi-paralysis to interact with robots or machines easier. The rehabilitation such as a process for recovering motor skills (e.g. hand/arm movements) after injury could also benefit from robot motions if the robot could sense and move according to quality of patient motions (the analogy of user-defined gestures). This kind of applications will require additional experimental trials for validating and adjusting before clinical testing.

Chapter 7

Conclusion

The results obtained from the video based experiment show important characteristics of user-defined gestures in carrying out collaborative tasks between a user and a helping hand robot in a desktop workspace. The details in Section 4.3 and the selected gestures in Section 4.4 can be used as guidelines for developing a gesture recognition system for a robotic system that needs the natural gestures for manually controlling the robot, especially in a common desktop workspace where the user is working closely with robots while sitting and holding tools and objects in their hands.

We also presented an implementation of a helping hand robot system that can be manually controlled with a set of user-defined gestures that were derived from the video based experiment. The implemented system and selected gestures allow users to control an end effector while working closely with the robot using body movement and hand gestures. The gesture recognition module allows the user to start and stop controlling at any position within the workspace. In particular, the users were able to control the helping hand robot with body movement gestures even though both their hands were occupied with the task.

In addition, we conducted an experiment with a group of participants to confirm the benefit of our proposed system. The outcomes align with the expectation from two most used body parts for gesturing from the video based experiment. However, with the real robot experiment, the body gestures were preferred over the hand gestures. The outcome is different from the video based experiment. This finding is unexpected but helps confirming the intuition from the pilot study and the need of the real robot system implementation in human robot interaction study.

The discovered gestures and their characteristics can be useful as a complementary feature for the development of multimodal communication in HRI and HRC to make the robots interact more naturally with humans.

Appendices

Appendix A

What and where to use the helping hand robot

This appendix summarizes results from [9] which is an online survey research with the follow abstract.

> What users want from robotic arms if they are become part for their daily living and working life? We conducted an online survey to collect ideas from people about "place" that they want to use a robotic arm and "task" that they want the robot arm to perform. 96 participants anonymously volunteered in this study. More than 240 sets of place and task were extracted from two versions of web based questionnaires. Results suggest that household, workplace and working surface are, respectively, the three most mentioned places for robotic arm usage in daily life, that participants from different countries those familiar with information technology did not show significant difference in their answers, and that females had more interest in household and self-care tasks. Our findings can be useful as a guideline for future research and development that focuses on daily life tasks for robotic arm.

Total 96 volunteers participated in the online survey, 58 Thais and 38 Japanese. A complete demographic information of the participants is shown in Table A.1. From the table, 69% was male, 90% of the participants were 18–30 years old, only one participant had the education level under bachelor degree, and 80% of the participants did not have any experience with robotic arm.

Table A.1: Demographic of the participants.

Nationality	Gender		Age					Education			Experience	
			< 18	18-30	31-45	46-55	> 55	<	Bachelor	>	Yes	No
Thai	M	43	0	36	6	1	0	1	22	20	17	26
	F	15	0	15	0	0	0	0	3	12	2	13
Japanese	M	23	1	21	1	0	0	0	20	3	2	21
	F	15	0	14	1	0	0	0	12	3	0	15
Total		96	1	86	8	1	0	1	57	38	21	75
Total in percentage			1.04%	89.58%	8.33%	1.04%	0.00%	1.04%	59.38%	39.58%	21.88%	78.13%

The participants submitted 96 complete responses. We extracted more than 240 sets of *place* and *task* from all responses and grouped all places based on information in each participant's answer. Almost every answer clearly stated place name or usage purpose of the robotic arm, for example, “on a table", “laboratory", or “near my bed". The grouped places and purposes from Thai and Japanese questionnaire is shown in Fig. A.1. Fig. A.2 shows a combined answer from all participants.

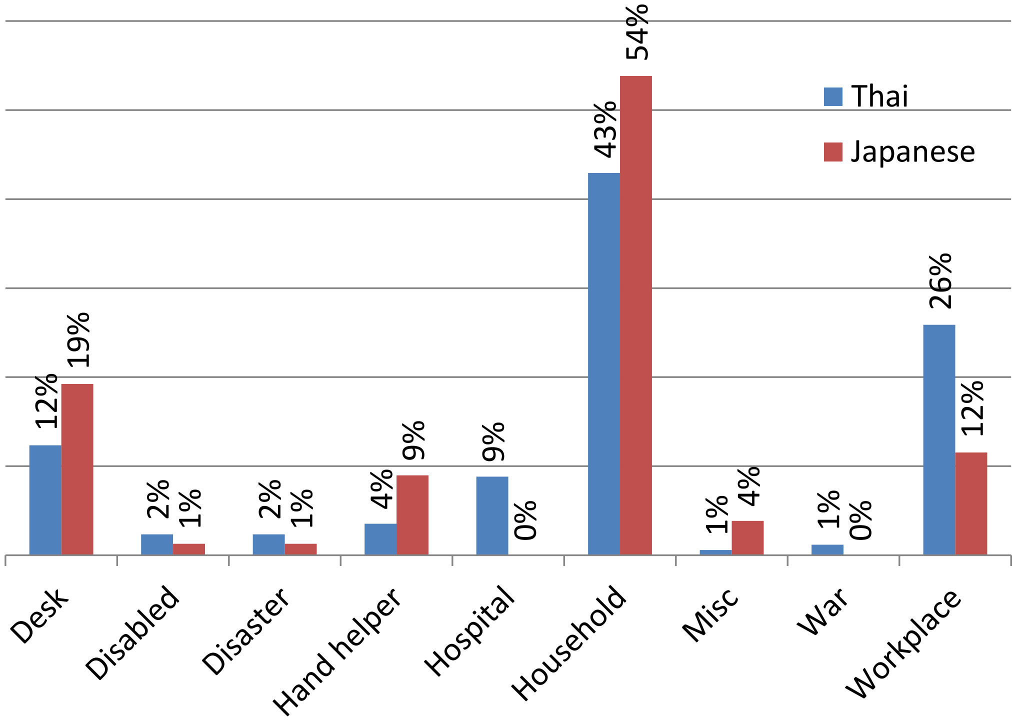


Figure A.1: Places and purposes of robotic arm usages those were answered by Thai and Japanese participants in percentage.

The following tables are functions of a helping hand robot grouped by places and purposes. Please see [9] for detailed description of each table.

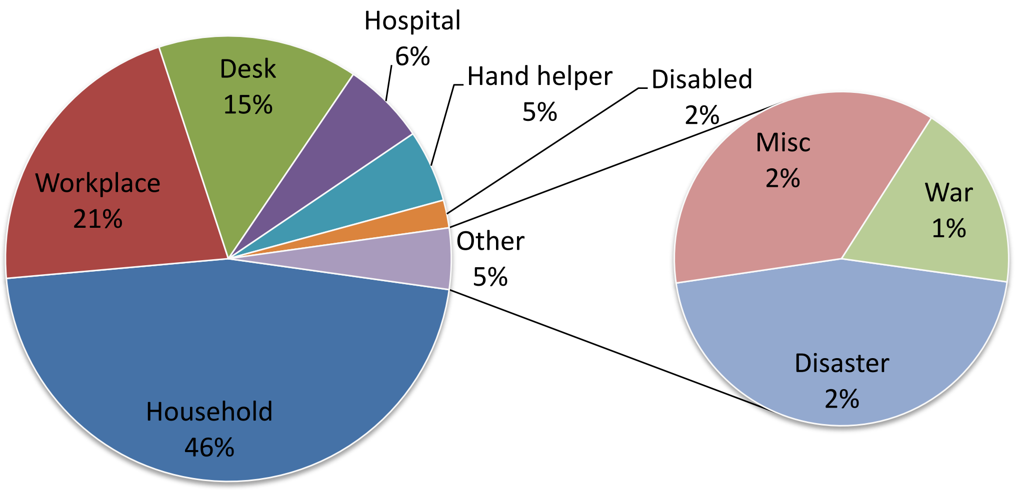


Figure A.2: Percentage of grouped places and purposes from all answers.

Table A.2: Disaster & War

Place	Usage
disaster	flood prevention tasks (e.g. set up sandbag) clear household and surrounding area rescue victim from collapsed buildings teleoperating in hazardous area (e.g. leaked nuclear power plant or chemical factory)
war	attached on UAV or mobile robot for helping solder to defusing bomb and removing trap

Table A.3: Hand helper & disability.

Place	Usage
prosthetic	replace disabled person limb(s) for daily life usage
additional arm	access or reach a place beyond arm length carry object while both hands are in use manipulate object that cannot be held normally (e.g. very high/low temperature)
portable arm	pick up potato chip while using a computer holding a camera or small object temporary driving a car

Table A.4: Household tasks.

Place	Usage
bathroom	showering, shampooing, scrubbing one's back, passing towel ensure safety of bathroom user (esp. elderly, disabled and patient)
bedroom	holding book/laptop/tablet while laying on bed picking up and passing remote, book alarm clock, massage
dressing room	makeup, haircut, hairdressing recommend/passing clothes, shoes, bags, accessories
garage	washing, cleaning, vacuuming load/unload heavy objects from car trunk
kitchen	preparing beverage (e.g. tea, coffee), ingredients (e.g. shrimp, squid, . . .) cooking: stir-fry, deep-fry, endless stirring, . . . bakery: dough preparation, cake decoration, . . . serving ready to eat meal that was prepared automatically by robot, table setup, washing, tidying hand helper when injured, pick up hot stuffs (e.g. pot, steamed foods, . . .)
laundry room	separating, washing, drying, collecting, ironing, folding and hanging
living room	opponent for playing card, game, . . . hand helper for load/unload CD, DVD, picking up remote, book
general tasks	cleaning kitchen, toilet, bathroom, bedroom, air conditioner and difficult to access area (e.g. ceiling, wall, . . .) tidying bedroom, bookshelf, shoes closet, storage room, shelf, and so on hand helper for searching, picking up, retrieving, and returning objects according to user's command
miscellaneous	protecting entrance(s) for home security take care of garden such such as tree shaping, mowing and watering feeding one's pet can be teleoperated when its owner is away from home

Table A.5: Workplace tasks.

Place	Usage
office	cleanup and tidying transporting, searching, picking up, retrieving, and returning objects (e.g. book, document, ...) massage when tired
sport	opponent for tennis, baseball, table tennis, ... trainer for body posture training/adjustment (e.g. golf, yoga)
art/music	enhance untrained person skill helping disable/injured artist paint, sculpt, play music instrument, ...
service	counter for money exchange, taller, ... automatic point of sale, kiosk, food stand logistic robot for load/unload, sort, search, object parking lot robot for assisting car parking mass ingredient preparation in hotel or restaurant, cooking assistant, serving food distribute handbill, pamphlet, brochure in public area, tourist spot
production	manufacturing process such as welding, drilling, cutting and milling heavy material manipulation such as iron ore, metal sheets, copper rods assembly line for car, airplane, building, train, ... utility installation such as pipe, cable, network, ... inventory management robot for small inventory that cannot be accessed by forklift or other vehicles
miscellaneous	working in extreme environment such as space, deep ocen, rescue, disaster zone ... teleoperation when human override is needed

Table A.6: Working surface tasks.

Place	Usage
object manipulation	holding, searching, picking up, passing, retrieving, returning
object assembly	small objects, complex objects, repetitive work, feeding parts
tool manipulation	measuring, drilling, soldering, lighting
tidying	put objects in a predefined place pre-tidying by arranging objects in a box that was prepared by user optimized working areas for space and ergonomic
writing	when printing is not allowed or handwriting is too bad for particular writing requirement

Table A.7: Hospital Tasks.

Place	Usage
emergency room	CPR, patient posture adjustment for X-ray and splint
operating room	passing surgery tools according to the surgeon's preference surgery assistant or surgery robot
laboratory	repetitive testing and analysis work, a test or process that uses or involves with hazardous material or chemical
pharmacy	storage management (e.g. store, retrieve, return), handle dangerous medicine (e.g. radio active)
rehabilitation	repetitive task, patient specific training session (e.g. force, path, angle, ...)
assisting patient	manipulate object based on patient command, ensure patient safety when he/she walking, getting up, using a toilet

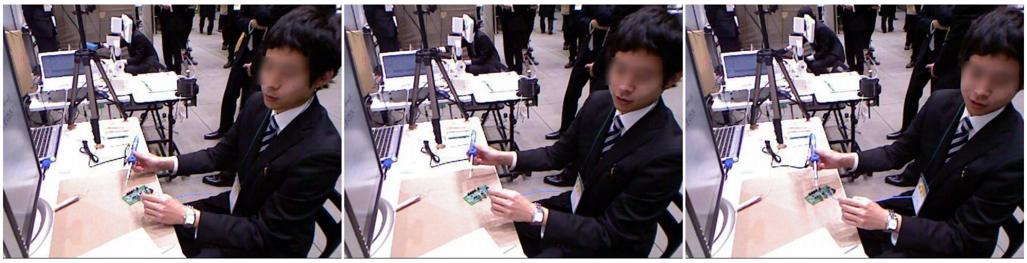
Appendix B

Observed gestures from the video based experiment

This appendix displays all gestures corrected from the video based experiment described in Chapter 4.

(a) The body gesture for moving the end effector forward.

(b) The body gesture for moving the end effector backward.



(c) The body gesture for moving the end effector left.

(d) The body gesture for moving the end effector right.

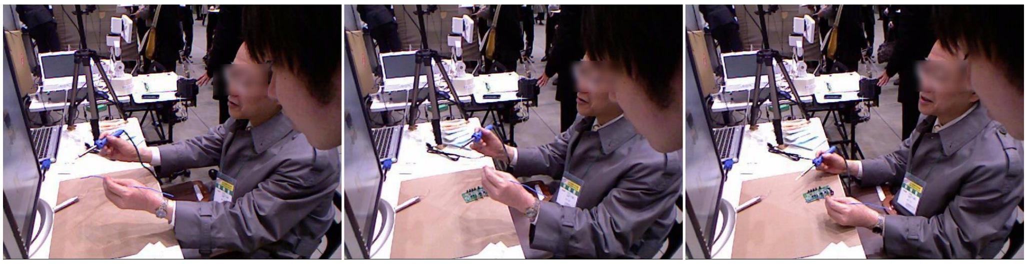
(e) The body gesture for moving the end effector up.

(f) The body gesture for moving the end effector down.

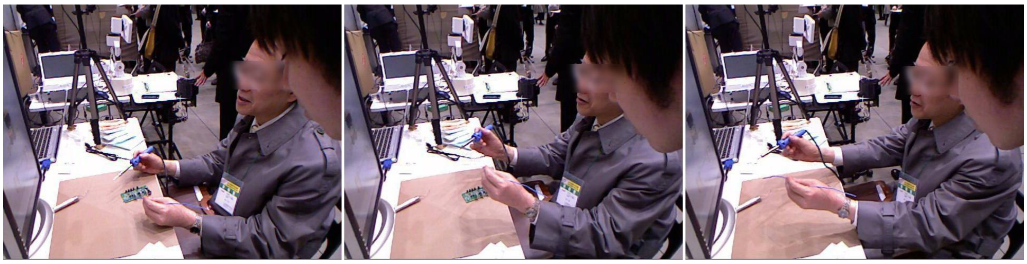
(g) The hand gesture for opening the gripper.

(h) The hand gesture for closing the gripper.

Figure B.1: Gestures from the 1st participants.



(a) The hand gesture for moving the end effector forward.



(b) The hand gesture for moving the end effector backward.

(c) The hand gesture for moving the end effector left.

(d) The hand gesture for moving the end effector right.

(e) The hand gesture for moving the end effector up.

(f) The hand gesture for moving the end effector down.

(g) The hand gesture for opening the gripper.

(h) The hand gesture for closing the gripper.

Figure B.2: Gestures from the 2nd participants.

(a) The body gesture for moving the end effector forward.

(b) The body gesture for moving the end effector backward.

(c) The body gesture for moving the end effector left.

(d) The body gesture for moving the end effector right.

(e) The head gesture for moving the end effector up.

(f) The body gesture for moving the end effector down.

(g) The head gesture for opening the gripper.

(h) The head gesture for closing the gripper.

Figure B.3: Gestures from the 3rd participants.


(a) The head gesture for moving the end effector forward.


(b) The body gesture for moving the end effector backward.


(c) The hand gesture for moving the end effector left.


(d) The hand gesture for moving the end effector right.

(e) The hand gesture for moving the end effector up.

(f) The hand gesture for moving the end effector down.

(g) The hand gesture for opening the gripper.

(h) The hand gesture for closing the gripper.

Figure B.4: Gestures from the 4th participants.

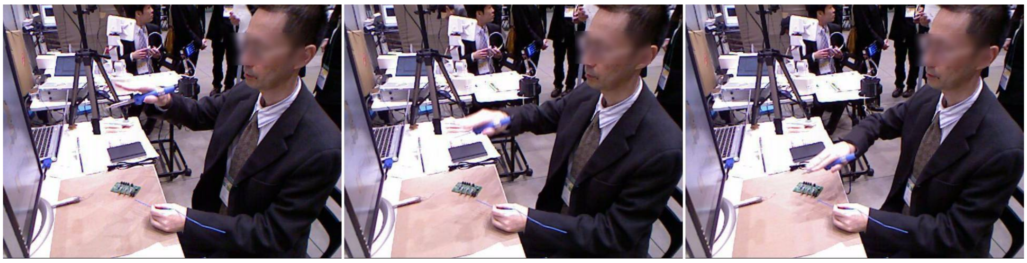
(a) The hand gesture for moving the end effector forward.

(b) The hand gesture for moving the end effector backward.

(c) The hand gesture for moving the end effector left.

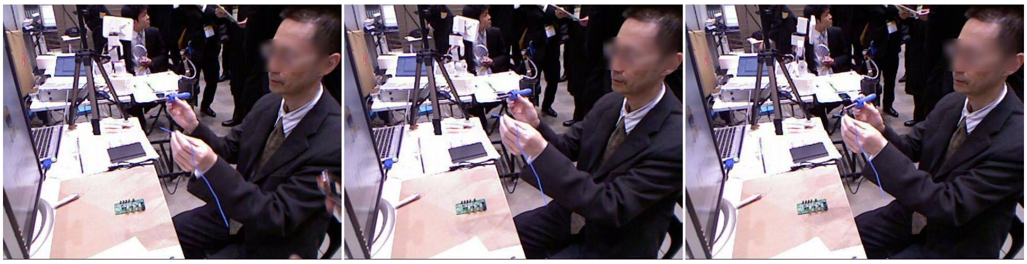
(d) The hand gesture for moving the end effector right.

(e) The hand gesture for moving the end effector up.



(f) The hand gesture for moving the end effector down.

(g) The hand gesture for opening the gripper.

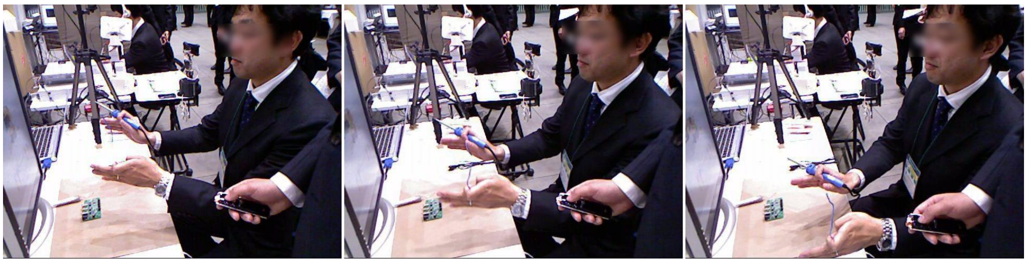


(h) The hand gesture for closing the gripper.

Figure B.5: Gestures from the 5th participants.

(a) The hand gesture for moving the end effector forward.

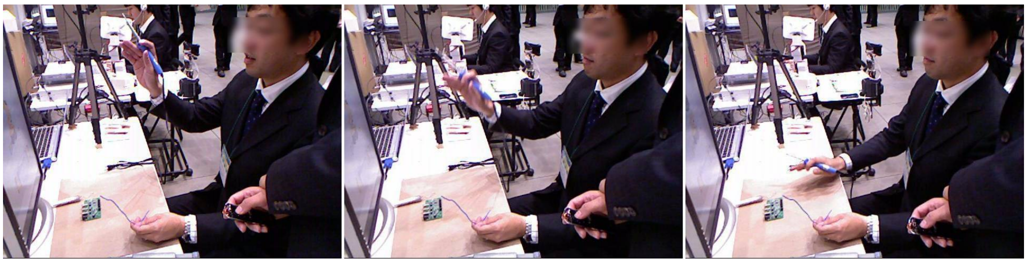
(b) The hand gesture for moving the end effector backward.



(c) The hand gesture for moving the end effector left.

(d) The hand gesture for moving the end effector right.

(e) The hand gesture for moving the end effector up.



(f) The hand gesture for moving the end effector down.

(g) The hand gesture for opening the gripper.

(h) The hand gesture for closing the gripper.

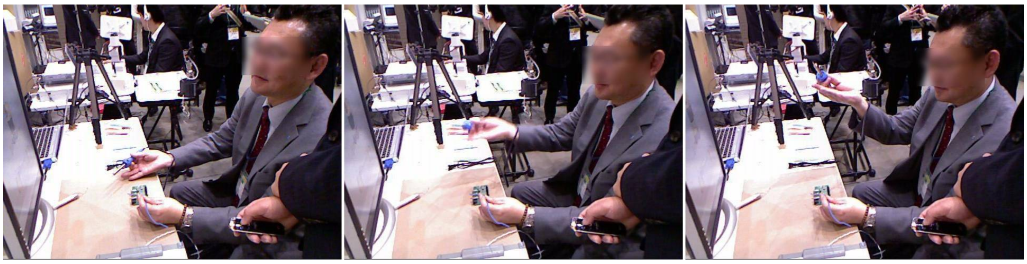
Figure B.6: Gestures from the 6th participants.

(a) The hand gesture for moving the end effector forward.

(b) The hand gesture for moving the end effector backward.

(c) The hand gesture for moving the end effector left.

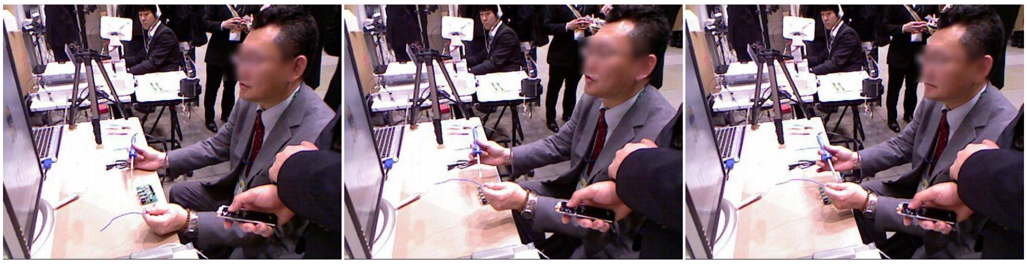
(d) The hand gesture for moving the end effector right.



(e) The hand gesture for moving the end effector up.

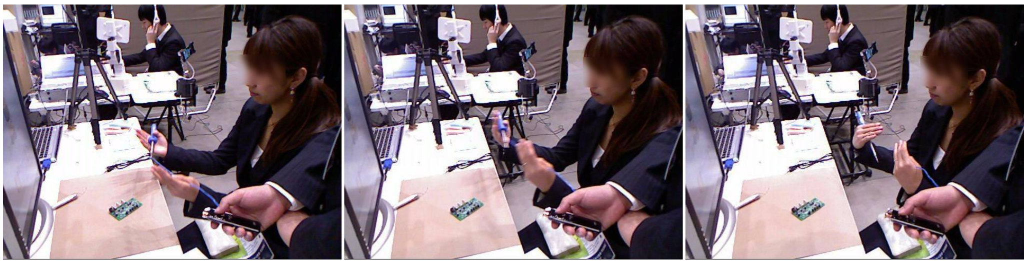
(f) The hand gesture for moving the end effector down.

(g) The hand gesture for opening the gripper.



(h) The hand gesture for closing the gripper.

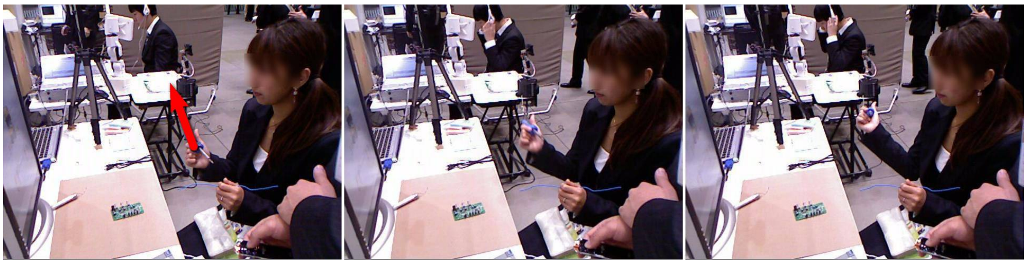
Figure B.7: Gestures from the 7th participants.



(a) The hand gesture for moving the end effector forward.

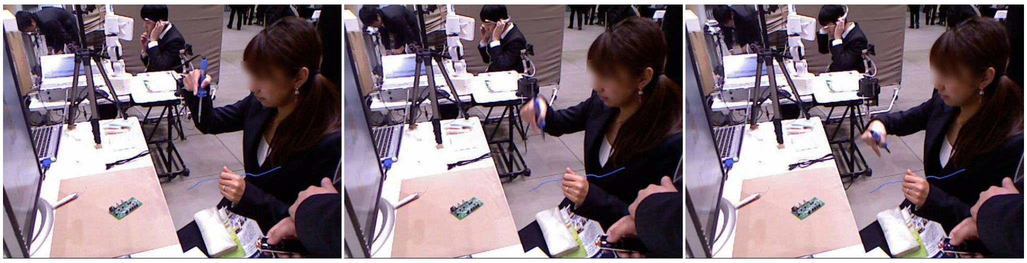
(b) The hand gesture for moving the end effector backward.

(c) The hand gesture for moving the end effector left.



(d) The hand gesture for moving the end effector right.

(e) The hand gesture for moving the end effector up.



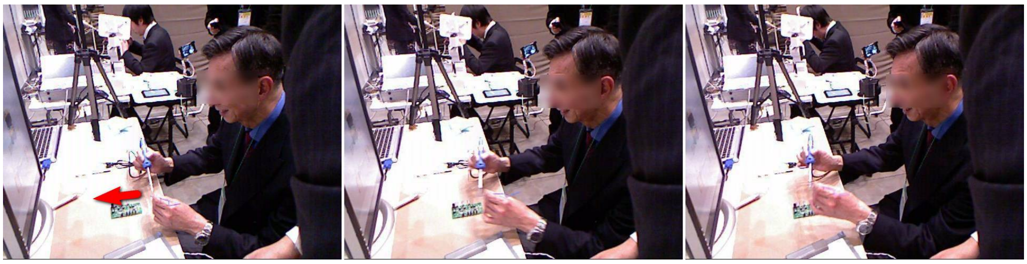
(f) The hand gesture for moving the end effector down.

(g) The hand gesture for opening the gripper.

(h) The hand gesture for closing the gripper.

Figure B.8: Gestures from the 8th participants.

(a) The hand gesture for moving the end effector forward.

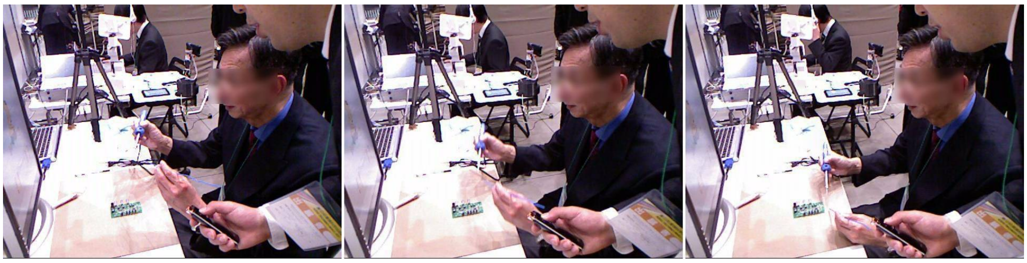


(b) The hand gesture for moving the end effector backward.

(c) The hand gesture for moving the end effector left.

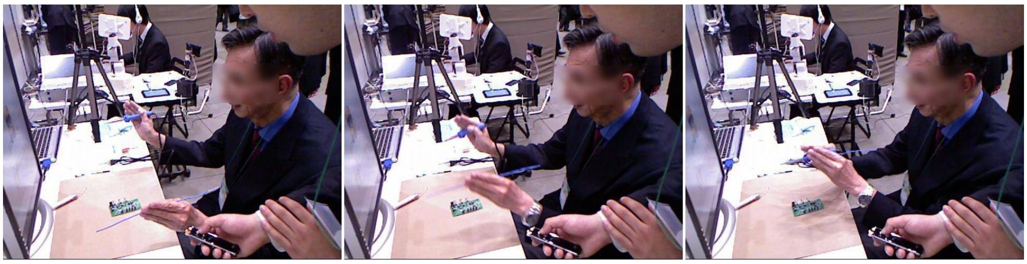
(d) The hand gesture for moving the end effector right.

(e) The hand gesture for moving the end effector up.



(f) The hand gesture for moving the end effector down.

(g) The hand gesture for opening the gripper.

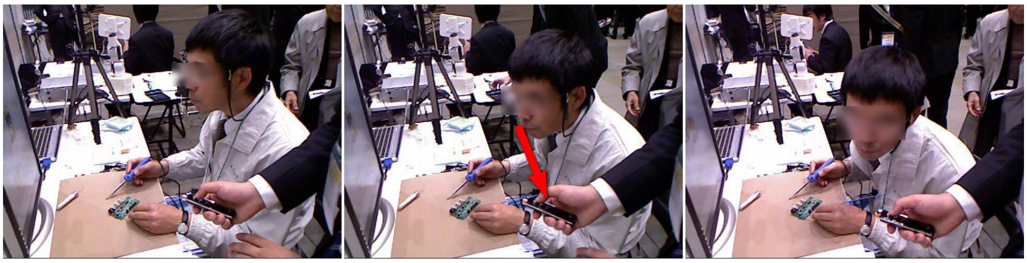


(h) The hand gesture for closing the gripper.

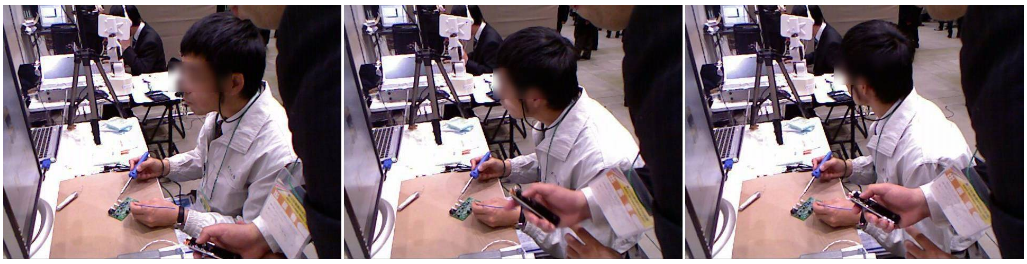
Figure B.9: Gestures from the 9th participants.

(a) The head gesture for moving the end effector forward.

(b) The head gesture for moving the end effector backward.

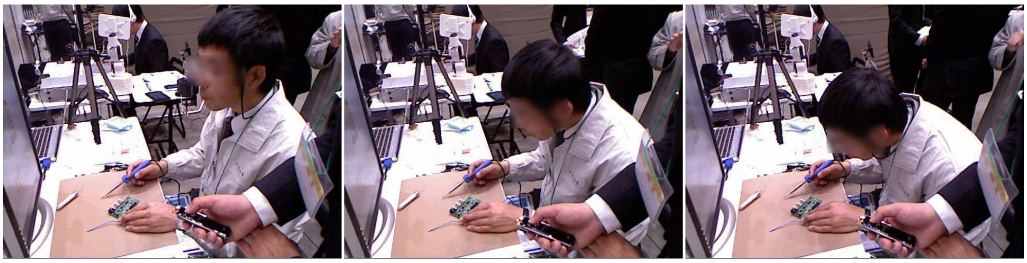


(c) The head gesture for moving the end effector left.



(d) The head gesture for moving the end effector right.

(e) The head gesture for moving the end effector up.



(f) The head gesture for moving the end effector down.

(g) The mouth gesture for opening the gripper.

(h) The mouth gesture for closing the gripper.

Figure B.10: Gestures from the 10th participants.

(a) The body gesture for moving the end effector forward.

(b) The body gesture for moving the end effector backward.

(c) The body gesture for moving the end effector left.

(d) The body gesture for moving the end effector right.

(e) The head gesture for moving the end effector up.

(f) The head gesture for moving the end effector down.

(g) The arm gesture for opening the gripper.

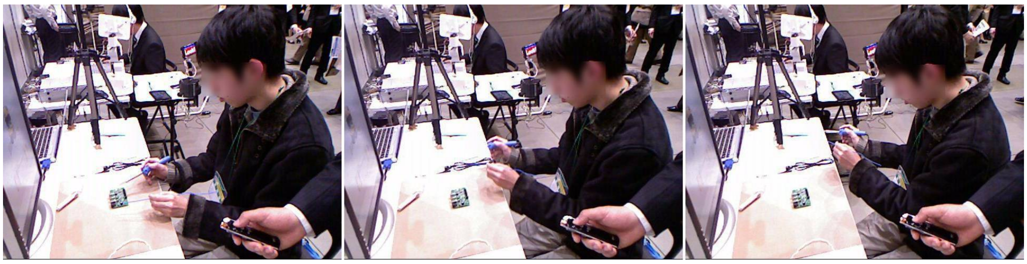
(h) The arm gesture for closing the gripper.

Figure B.11: Gestures from the 11th participants.

(a) The hand gesture for moving the end effector forward.

(b) The hand gesture for moving the end effector backward.

(c) The hand gesture for moving the end effector left.



(d) The hand gesture for moving the end effector right.

(e) The hand gesture for moving the end effector up.

(f) The hand gesture for moving the end effector down.

(g) The hand gesture for opening the gripper.

(h) The hand gesture for closing the gripper.

Figure B.12: Gestures from the 12th participants.

(a) The body gesture for moving the end effector forward.

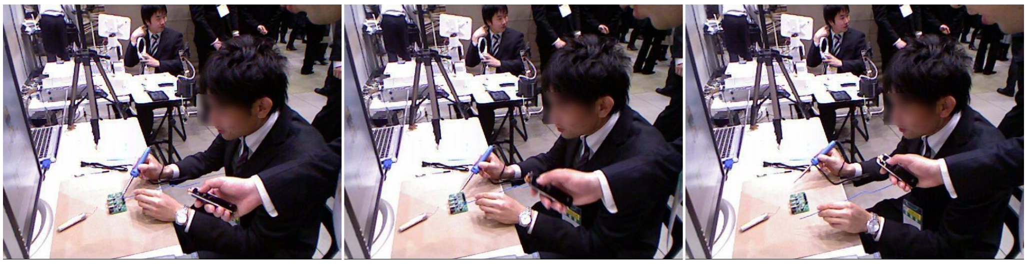
(b) The hand gesture for moving the end effector backward.

(c) The hand gesture for moving the end effector left.

(d) The body gesture for moving the end effector right.

(e) The hand gesture for moving the end effector up.

(f) The hand gesture for moving the end effector down.

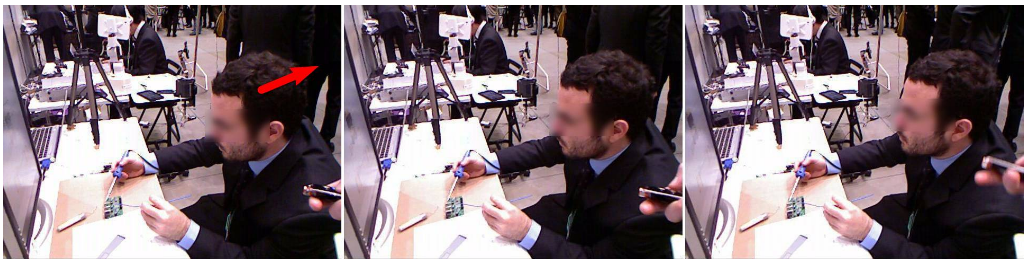


(g) The hand gesture for opening the gripper.

(h) The hand gesture for closing the gripper.

Figure B.13: Gestures from the 13th participants.

(a) The shoulder gesture for moving the end effector forward.



(b) The head gesture for moving the end effector backward.

(c) The head gesture for moving the end effector left.

(d) The head gesture for moving the end effector right.

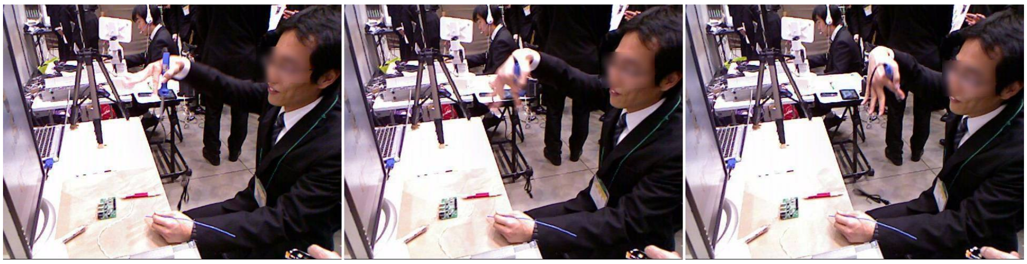
(e) The head gesture for moving the end effector up.

(f) The head gesture for moving the end effector down.

(g) The mouth gesture for opening the gripper.

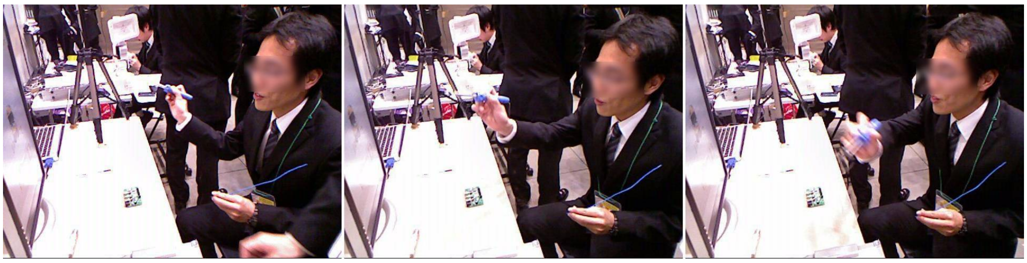
(h) The mouth gesture for closing the gripper.

Figure B.14: Gestures from the 14th participants.

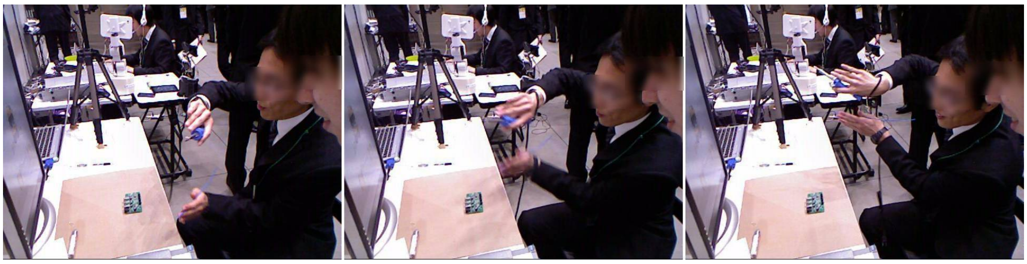


(a) The hand gesture for moving the end effector forward.

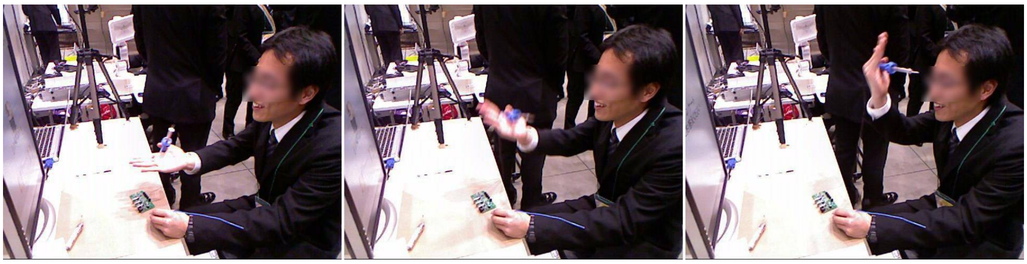
(b) The hand gesture for moving the end effector backward.



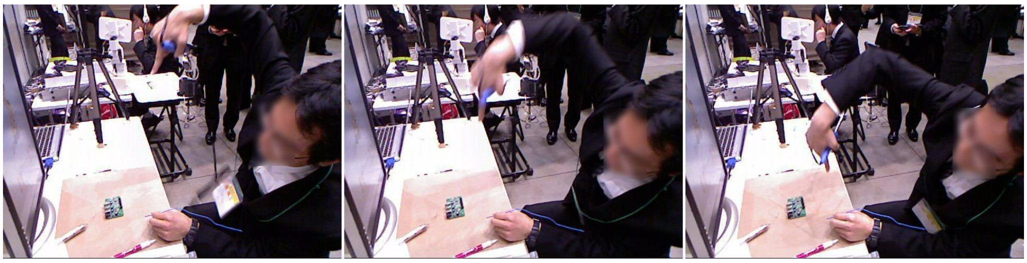
(c) The hand gesture for moving the end effector left.



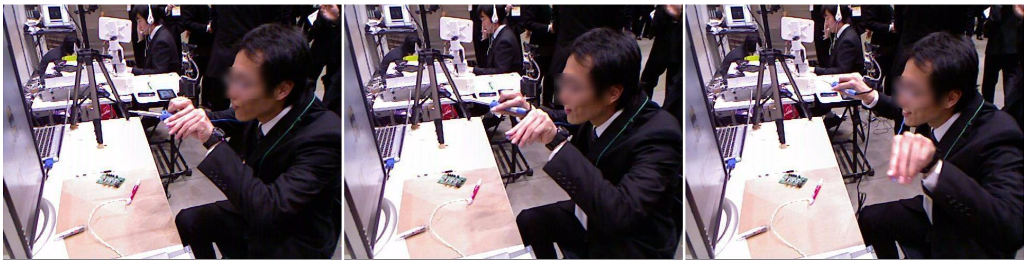
(d) The hand gesture for moving the end effector right.



(e) The hand gesture for moving the end effector up.



(f) The hand gesture for moving the end effector down.



(g) The hand gesture for opening the gripper.

(h) The hand gesture for closing the gripper.

Figure B.15: Gestures from the 15th participants.

(a) The body gesture for moving the end effector forward.

(b) The body gesture for moving the end effector backward.

(c) The head gesture for moving the end effector left.

(d) The head gesture for moving the end effector right.

(e) The body gesture for moving the end effector up.

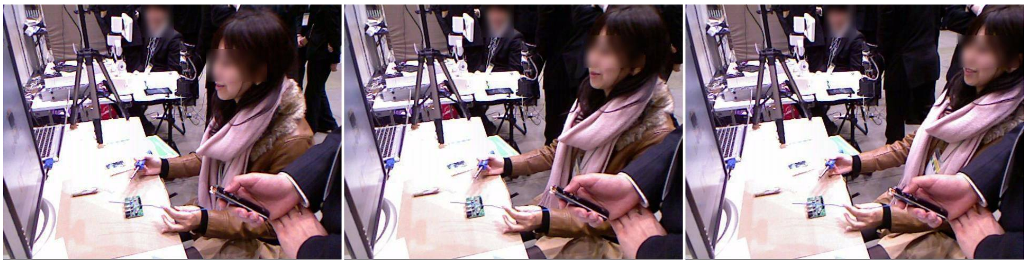
(f) The hand gesture for moving the end effector down.

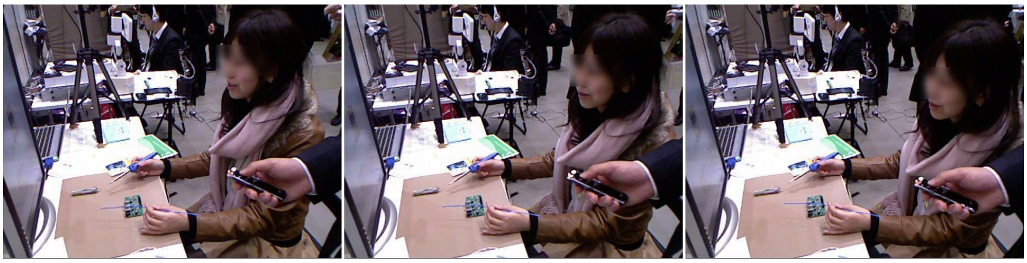
(g) The head gesture for opening the gripper.

(h) The head gesture for closing the gripper.

Figure B.16: Gestures from the 16th participants.

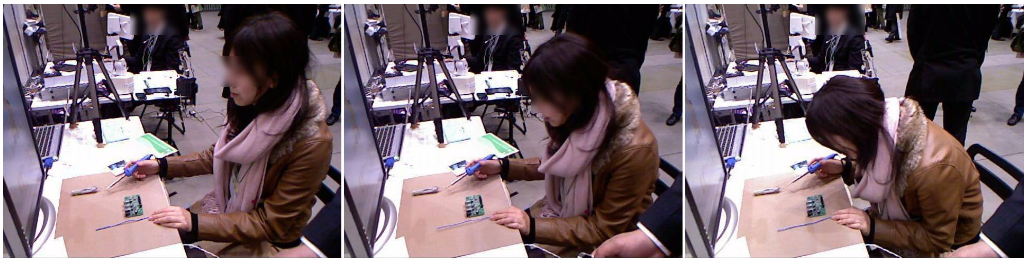

(a) The hand gesture for moving the end effector forward.


(b) The body gesture for moving the end effector backward.

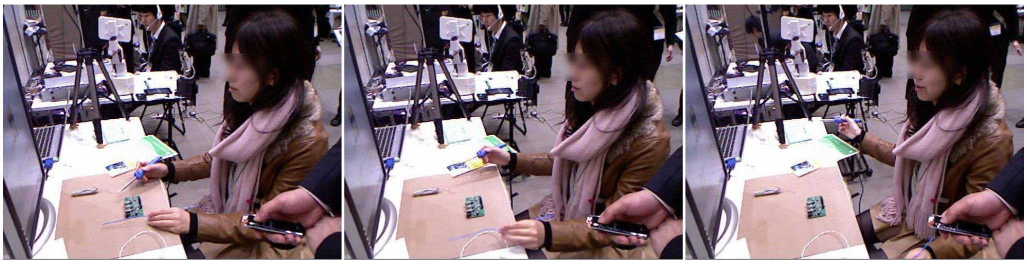

(c) The body gesture for moving the end effector left.


(d) The body gesture for moving the end effector right.

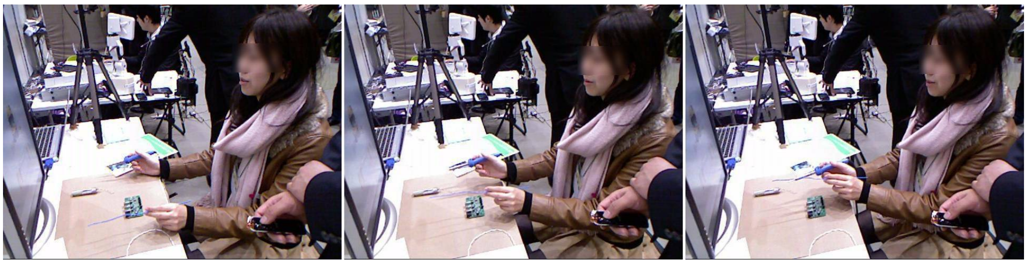
(e) The hand gesture for moving the end effector up.



(f) The body gesture for moving the end effector down.



(g) The hand gesture for opening the gripper.



(h) The hand gesture for closing the gripper.

Figure B.17: Gestures from the 17th participants.

(a) The hand gesture for moving the end effector forward.

(b) The hand gesture for moving the end effector backward.

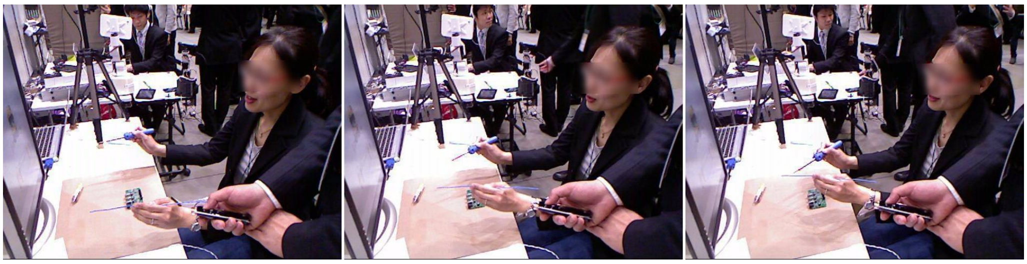
(c) The hand gesture for moving the end effector left.

(d) The hand gesture for moving the end effector right.

(e) The hand gesture for moving the end effector up.

(f) The hand gesture for moving the end effector down.

(g) The hand gesture for opening the gripper.



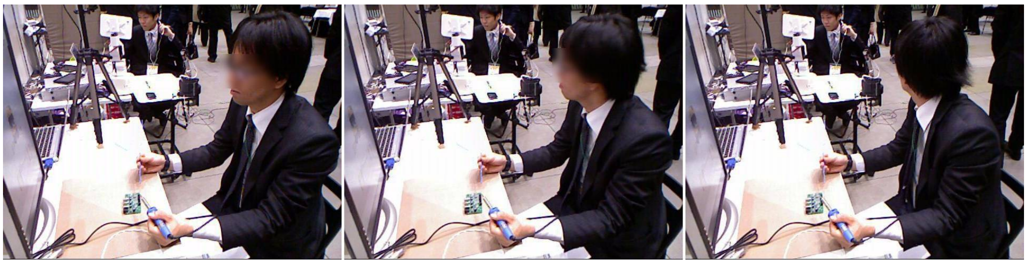
(h) The hand gesture for closing the gripper.

Figure B.18: Gestures from the 18th participants.

(a) The body gesture for moving the end effector forward.

(b) The body gesture for moving the end effector backward.

(c) The body gesture for moving the end effector left.

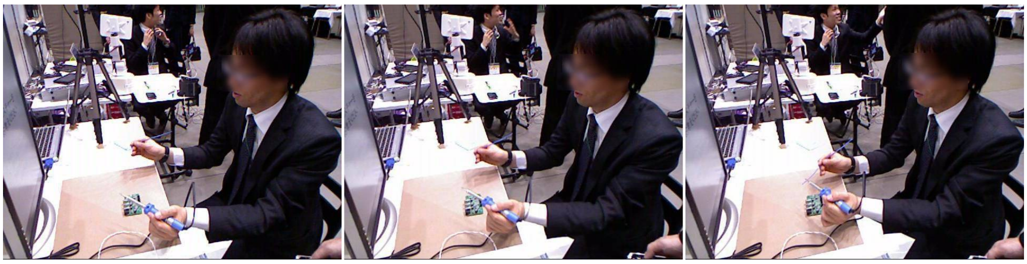


(d) The body gesture for moving the end effector right.

(e) The hand gesture for moving the end effector up.

(f) The hand gesture for moving the end effector down.

(g) The hand gesture for opening the gripper.



(h) The hand gesture for closing the gripper.

Figure B.19: Gestures from the 19th participants.

Publications

List of the published international conference papers and peer reviewed journals of the author.

[1] M. Wongphati, H. Osawa, and M. Imai. “Gestures for Manually Controlling a Helping Hand Robot”. English. In: *International Journal of Social Robotics* 7 (May 2015), pp. 731–742.

[2] M. Wongphati, H. Osawa, and M. Imai. “User-defined gestures for controlling primitive motions of an end effector”. In: *Advanced Robotics* 29.4 (2015), pp. 225–238.

[3] M. Wongphati et al. “Where do you want to use a robotic arm? And what do you want from the robot?” In: *Proceedings of the IEEE International Symposium on Robots and Human Interactive Communication (RO-MAN).* Sept. 2012, pp. 322–327.

[4] M. Wongphati et al. “Give Me a Hand – How Users Ask a Robotic Arm for Help with Gestures”. In: *Proceedings of the IEEE International Conference on Cyber Technology in Automation, Control, and Intelligent Systems (CYBER).* 2012, pp. 64–68.

[5] M. Wongphati, H. Osawa, and M. Imai. “3D low-profile evaluation system (LES) an unobtrusive measurement tool for HRI”. In: *Proceedings of the IEEE International Symposium on Robots and Human Interactive Communication (RO-MAN).* July 2011, pp. 162–167.

References

[1] S. Y. Nof. *Handbook of Industrial Robotics*. Second. John Wiley & Sons, 1999.

[2] J. Hammel et al. "Clinical evaluation of a desktop robotic assistant". In: *Rehabilitation Research and Development* 26 (1989), pp. 1–16.

[3] H. V. der Loos. "VA/Stanford Rehabilitation Robotics Research and Development Program: Lessons Learned in the Application of Robotics Technology to the Field of Rehabilitation". In: *IEEE Transactions on Rehabilitation Engineering* 3.1 (Mar. 1995), pp. 46–55.

[4] L. R. Hochberg et al. "Reach and grasp by people with tetraplegia using a neurally controlled robotic arm". In: *Nature* 485.7398 (May 2012), pp. 372–375.

[5] H. A. ElMaraghy, ed. *Changeable and Reconfigurable Manufacturing Systems*. Springer, 2009.

[6] F. Wallhoff et al. "A skill-based approach towards hybrid assembly". In: *Advanced Engineering Informatics* 24.3 (2010), pp. 329–339.

[7] M. A. Goodrich and A. C. Schultz. "Human-robot interaction: A survey". In: *Foundation and Trends in Human-Computer Interaction* 1 (3 Jan. 2007), pp. 203–275.

[8] A. Bauer, D. Wollherr, and M. Buss. "Human-robot collaboration: a survey". In: *Humanoid Robotics* 05.01 (2008), pp. 47–66.

[9] M. Wongphati et al. "Where do you want to use a robotic arm? And what do you want from the robot?" In: *Proceedings of the IEEE International Symposium on Robots and Human Interactive Communication (RO-MAN)*. Sept. 2012, pp. 322–327.

[10] S. Oviatt. "Ten Myths of Multimodal Interaction". In: *Communication of the ACM* 42.11 (Nov. 1999), pp. 74–81.

[11] Y. Hato et al. “Pointing to Space: Modeling of Deictic Interaction Referring to Regions”. In: *Proceedings of the ACM/IEEE International Conference on Human-Robot Interaction (HRI).* Osaka, Japan, 2010, pp. 301–308.

[12] J. P. Wachs et al. “Vision-Based Hand-Gesture Applications”. In: *Communications of the ACM* 54 (2 Feb. 2011), pp. 60–71.

[13] J. O. Wobbrock, M. R. Morris, and A. D. Wilson. “User-Defined Gestures for Surface Computing”. In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI).* Boston, MA, USA: ACM, 2009, pp. 1083–1092.

[14] J. B. Reswick and L. Vodovnik. “External Power in Prosthetics and Orthotics, an Overview”. In: *Artificial Limbs* 11.2 (1967), pp. 5–21.

[15] T. B. Sheridan. *Telerobotics, Automation, and Human Supervisory Control.* MIT Press, 1992.

[16] C. M. Huang and B. Mutlu. “Learning-based Modeling of Multimodal Behaviors for Humanlike Robots”. In: *Proceedings of the ACM/IEEE International Conference on Human-Robot Interaction (HRI).* Bielefeld, Germany: ACM, 2014, pp. 57–64.

[17] T. Fong, I. Nourbakhsh, and K. Dautenhahn. “A survey of socially interactive robots”. In: *Robotics and Autonomous Systems* 42.3–4 (2003), pp. 143–166.

[18] M. Hans, B. Graf, and R. Schraft. “Robotic home assistant Care-O-bot: past-present-future”. In: *Proceedings of the IEEE International Workshop on Robots and Human Interactive Communication (RO-MAN).* 2002, pp. 380–385.

[19] O. Rogalla et al. “Using Gesture and Speech Control for Commanding a Robot Assistant”. In: *Proceedings of the IEEE International Workshop on Robots and Human Interactive Communication (RO-MAN).* Berlin, Germany, 2002, pp. 454–459.

[20] I. Iossifidis et al. “Anthropomorphism as a pervasive design concept for a robotic assistant”. In: *Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS).* Vol. 3. 2003, pp. 3465–3472.

[21] J. Zhang and A. Knoll. “A two-arm situated artificial communicator for human-robot cooperative assembly”. In: *IEEE Transactions on Industrial Electronics* 50.4 (2003), pp. 651–658.

[22] N. Kawarazaki et al. “Cooperative Welfare Robot System Using Hand Gesture Instructions”. In: *Advances in Rehabilitation Robotics*. Ed. by Z. Bien and D. Stefanov. Vol. 306. Lecture Notes in Control and Information Sciences. Springer Berlin / Heidelberg, 2004, pp. 143–153.

[23] R. Stiefelhagen et al. “Natural Human-Robot Interaction Using Speech, Head Pose and Gestures”. In: *Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. Vol. 3. Sendai, Japan, 2004, pp. 2422–2427.

[24] S. Srinivasa et al. “HERB: a home exploring robotic butler”. In: *Autonomous Robots* 28 (Jan. 2010). doi: pp. 5–20.

[25] A. Jain and C. Kemp. “EL-E: an assistive mobile manipulator that autonomously fetches objects from flat surfaces”. In: *Autonomous Robots* 28 (1 2010), pp. 45–64.

[26] C.-H. King et al. “Towards an assistive robot that autonomously performs bed baths for patient hygiene”. In: *Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. 2010, pp. 319–324.

[27] T. Asfour et al. “Toward humanoid manipulation in human-centred environments”. In: *Robotics and Autonomous Systems* 56.1 (June 2008), pp. 54–65.

[28] U. Reiser et al. “Care-O-bot® 3 - creating a product vision for service robot applications by integrating design and technology”. In: *Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. Oct. 2009, pp. 1992–1998.

[29] J. Bohren et al. “Towards autonomous robotic butlers: Lessons learned with the PR2”. In: *Proceedings of the IEEE International Conference on Robotics and Automation (ICRA)*. May 2011, pp. 5568–5575.

[30] S. S. Groothuis, S. Stramigioli, and R. Carloni. “Lending a Helping Hand: Toward Novel Assistive Robotic Arms”. In: *IEEE Robotics Automation Magazine* 20.1 (2013), pp. 20–29.

[31] M. Nielsen et al. “A Procedure for Developing Intuitive and Ergonomic Gesture Interfaces for HCI”. In: *Gesture-Based Communication in Human-Computer Interaction*. Ed. by A. Camurri and G. Volpe. Vol. 2915. Lecture Notes in Computer Science. Springer Berlin / Heidelberg, 2004, pp. 105–106.

[32] C. L. Nehaniv et al. “A methodological approach relating the classification of gesture to identification of human intent in the context of human-robot interaction”. In: *Proceedings of the IEEE International Workshop on Robots and Human Interactive Communication (RO-MAN).* 2005, pp. 371–377.

[33] J. Wachs, H. Stern, and Y. Edan. “Cluster Labeling and Parameter Estimation for The Automated Setup of a Hand-Gesture Recognition System”. In: *IEEE Transactions on Systems, Man and Cybernetics, Part A: Systems and Humans* 35.6 (Nov. 2005), pp. 932–944.

[34] T. P. Spexard, M. . Hanheide, and G. . Sagerer. “Human-Oriented Interaction With an Anthropomorphic Robot”. In: *IEEE Transactions on Robotics* 23.5 (2007), pp. 852–862.

[35] B. Burger et al. “Two-handed gesture recognition and fusion with speech to command a robot”. English. In: *Autonomous Robots* 32.2 (2012), pp. 129–147.

[36] I. Oikonomidis, N. Kyriazis, and A. Argyros. “Tracking the articulated motion of two strongly interacting hands”. In: *Proceedings of the IEEE International Conference on Computer Vision and Pattern Recognition (CVPR).* June 2012, pp. 1862–1869.

[37] P. Neto, J. Pires, and A. Moreira. “Accelerometer-based control of an industrial robotic arm”. In: *Proceedings of the IEEE International Symposium on Robots and Human Interactive Communication (RO-MAN).* Oct. 2009, pp. 1192–1197.

[38] N. Tran et al. “Wireless Data Glove for Gesture-Based Robotic Control”. In: *Human-Computer Interaction. Novel Interaction Methods and Techniques.* Ed. by J. Jacko. Vol. 5611. Lecture Notes in Computer Science. Springer Berlin / Heidelberg, 2009, pp. 271–280.

[39] A. G. Brooks and C. Breazeal. “Working with robots and objects: revisiting deictic reference for achieving spatial common ground”. In: *Proceedings of the ACM/IEEE International Conference on Human-Robot Interaction (HRI).* Salt Lake City, Utah, USA: ACM, 2006, pp. 297–304.

[40] T. Ende et al. “A human-centered approach to robot gesture based communication within collaborative working processes”. In: *Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS).* 2011, pp. 3367–3374.

[41] B. Gleeson et al. “Gestures for Industry: Intuitive Human-Robot Communication from Human Observation”. In: *Proceedings of the ACM/IEEE International Conference on Human-Robot Interaction (HRI).* Tokyo, Japan, 2013, pp. 349–356.

[42] M. Huber et al. “Human Preferences in Industrial Human-Robot Interactions”. In: *Proceedings of the International Workshop on Cognition for Technical Systems.* 2008.

[43] S. Glasauer et al. “Interacting in time and space: Investigating human-human and human-robot joint action”. In: *Proceedings of the IEEE International Symposium on Robots and Human Interactive Communication (RO-MAN).* 2010, pp. 252–257.

[44] S. Woods et al. “Methodological Issues in HRI: A Comparison of Live and Video-Based Methods in Robot to Human Approach Direction Trials”. In: *Proceedings of the IEEE International Symposium on Robots and Human Interactive Communication (RO-MAN).* Hatfield, UK, Sept. 2006, pp. 51–58.

[45] H. Sloetjes and P. Wittenburg. “Annotation by category - ELAN and ISO DCR”. In: *Proceedings of the International Conference on Language Resources and Evaluation.* Marrakech, Morocco, 2008.

[46] K. Dautenhahn. “Methodology and Themes of Human-Robot Interaction: A Growing Research Field”. In: *Advanced Robotic Systems* 4.1 (2007), pp. 103–108.

[47] M. L. Walters et al. “Robotic Etiquette: Results from User Studies Involving a Fetch and Carry Task”. In: *Proceedings of the ACM/IEEE International Conference on Human-Robot Interaction (HRI).* Arlington, Virginia, USA: ACM, 2007, pp. 317–324.

[48] J. Drury, B. Keyes, and H. Yanco. “LASSOing HRI: Analyzing situation awareness in map-centric and video-centric interfaces”. In: *Proceedings of the ACM/IEEE International Conference on Human-Robot Interaction (HRI).* Washington DC, USA, Mar. 2007, pp. 279–286.

[49] M. L. Walters et al. “Evaluating the Robot Personality and Verbal Behavior of Domestic Robots Using Video-Based Studies”. In: *Advanced Robotics* 25.18 (2011), pp. 2233–2254.

[50] Y. Kuniyoshi, M. Inaba, and H. Inoue. “Learning by watching: extracting reusable task knowledge from visual observation of human performance”. In: *IEEE Transactions on Robotics and Automation* 10.6 (1994), pp. 799–822.

[51] D. Gossow et al. “Interactive Markers: 3-D User Interfaces for ROS Applications [ROS Topics]”. In: *IEEE Robotics Automation Magazine* 18.4 (Dec. 2011), pp. 14–15.

[52] R. Wang, S. Paris, and J. Popović. “6D hands: markerless hand-tracking for computer aided design”. In: *Proceedings of the ACM User Interface Software and Technology Symposium (UIST).* UIST ’11. Santa Barbara, California, USA: ACM, 2011, pp. 549–558.

[53] M. Wongphati, H. Osawa, and M. Imai. “3D low-profile evaluation system (LES) an unobtrusive measurement tool for HRI”. In: *Proceedings of the IEEE International Symposium on Robots and Human Interactive Communication (RO-MAN).* July 2011, pp. 162–167.

[54] R. B. Rusu and S. Cousins. “3D is here: Point Cloud Library (PCL)”. In: *Proceedings of the IEEE International Conference on Robotics and Automation (ICRA).* 2011, pp. 1–4.

[55] M. Quigley et al. “ROS: an open-source Robot Operating System”. In: *ICRA Workshop on Open Source Software.* 2009.

[56] E. Fernández et al. *Learning ROS for Robotics Programming - second edition.* Ed. by P. Publishing. Packt Publishing, Aug. 2015.

[57] Y. Pyo. *ROS Robot Programming.* Ruby Paper, Mar. 2015.

[58] G. Bradski and A. Kaehler. *Learning OpenCV.* O’Reilly Media, 2008.

[59] M. Kohler. *Using the Kalman Filter to track Human Interactive Motion – Modelling and Initialization of the Kalman Filter for Translational Motion.* Tech. rep. 1997.

[60] G. Welch and G. Bishop. *An introduction to the Kalman Filter.* Tech. rep. Department of Computer Science University of North Carolina at Chapel Hill Chapel Hill, 2006.

[61] F. Faul et al. “G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences”. English. In: *Behavior Research Methods* 39.2 (2007), pp. 175–191.